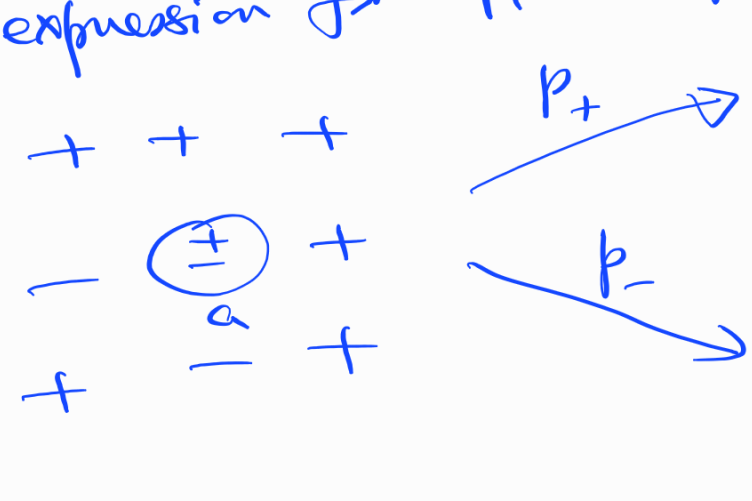
Computational Methods for Quantum Many-Body Physics

Lecture 5

June 14, 2023

Heatbath algorithm (on the 2D Ising model)

Again, we select a spin at random and then assign it to be + with prob. p_+ and − with prob p_-, independent of its previous orientation. Let us see how to set the expression for p_+ and p_- (clearly $p_+ + p_- = 1$)



$H = -\sum_{\langle ij \rangle} S_i S_j$ → $j \to a$ means j is nearest neighbor of a

$= -S_a \left(\sum_{j \to a} S_j \right) +$ other terms that do not involve S_a

$\sum_{j \to a} S_j$ then defines an "effective" h_a

$$h_a = \sum_{j \to a} S_j \quad \text{(local field on } S_a\text{)}$$

Heatbath algorithm aims to locally equilibrate S_a to the local field h_a

Then,

$$\left.\begin{aligned} p_+ &= \frac{1}{1+e^{-2\beta h a}} \\ \text{and } p_- &= \frac{1}{1+e^{2\beta h a}} \end{aligned}\right\}$$

This can be written a bit more ——— (1)
formally as

$$P(\mu \to \nu) = g(\mu \to \nu)\, A(\mu \to \nu) \quad \text{——— (2)}$$

where if $\mu \to \nu$ is rejected after having being considered, then the move $\mu \to \mu$ is chosen instead and the system remains in its original microstate.

For the heatbath algorithm,

$$g(\mu \to \nu) = \frac{1}{N} \quad \text{(where } N = \#\text{ of sites)} \quad \text{——— (3)}$$

for two microstates μ and ν connected by a single spin flip where the site is selected completely randomly. All other $g(\mu \to \nu) = 0$ [This is the same as the Metropolis algorithm.]

$$A(\mu \to \nu) = \frac{e^{-\beta E_\nu}}{e^{-\beta E_\mu} + e^{-\beta E_\nu}} \quad \text{——— (4)}$$

[This is different from the Metropolis's algorithm.]

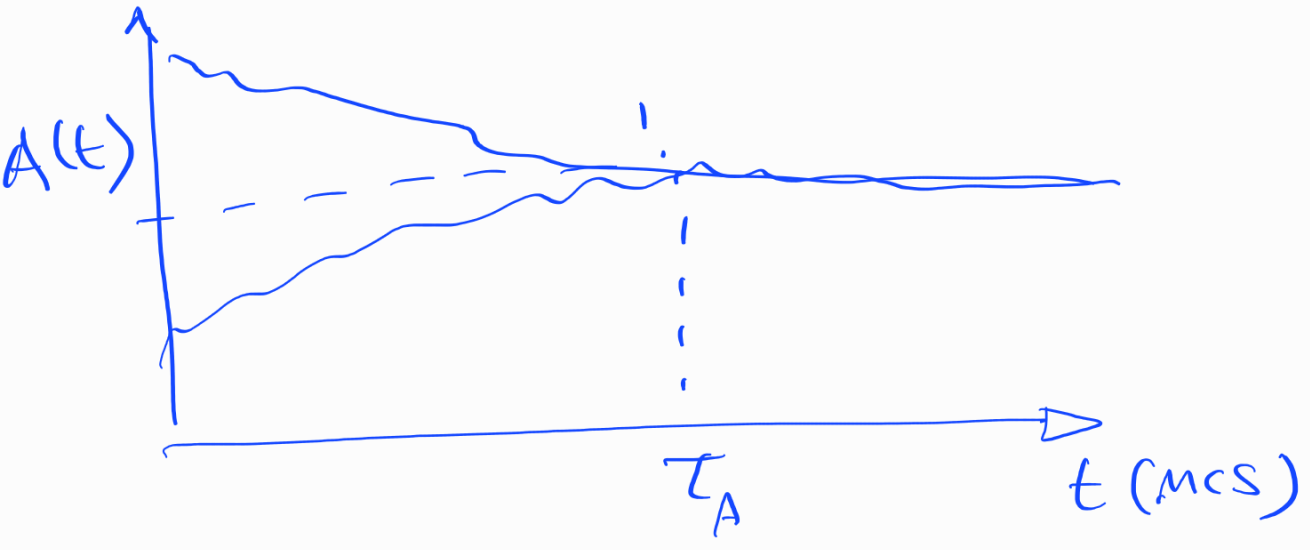
It is simple to show detailed balance and ergodicity for this algorithm which we leave as an exercise.

The Metropolis as well as the heat bath algorithms are both examples of local algorithms:

$g(\mu \to \nu)$ involve updating a finite # of spins (in these cases, one) to move from one microstate to another, no matter what the system size.

* Typically, in such cases, 1 MCS (Monte Carlo step) is defined as N spin flip attempts.

Some general comments about equilibration



In the Monte Carlo, since the next state μ_N depends on the state μ_{N-1}, there is a "memory effect". That is why, different observables (denoted by A) need different times ($\tau_A \to$ autocorrelation times) to become independent of the initial state.

Some useful "rules" to keep in mind:

* Check MC on small system sizes with exact methods whenever possible

* Check for independence of results of the MC for different initial conditions (hot start, cold start etc...)

* Check for independence of results for different random number seeds from the same initial condition.

* Always go from smaller systems to bigger systems in a MC study so that it may be possible to estimate those some internal relaxation timescales of the Markov process may increase with system size.

* Extremely important to use high quality random # generators like RANLUX or Mersenne Twister to get correct statistical results in a MC simulation.

E.g. see "Monte Carlo simulations: Hidden errors from "good" random number generators" by Ferrenberg, Landau & Wang PRL 69, 3382 (1992)

where results of a cluster algorithm for the 2D Ising model were compared

with exact results on a 16×16 lattice.
Some results were OK, but others were off by more than 100 standard deviations signalling that the random #s being used were not "sufficiently random".

This example highlights the importance of proper statistical error analysis of the MC data as well.

Useful ref: "Everything you wanted to know about data analysis and fitting but were afraid to ask" — Peter Young.
arXiv: 1210.3781

We will just summarize some of the useful formulas below.

Suppose there is a random variable A whose mean is $\langle A \rangle$ and whose

$$\mathrm{Var}(A) = \langle A^2 \rangle - \langle A \rangle^2 = \sigma^2 \quad \text{(5)}$$

when σ = width / standard deviation

* Now consider "n" "statistically independent" observations $\{A_i\}$ of this quantity A

* Unbiased estimator of the mean $\langle A \rangle$ is

$$\bar{A} = \frac{1}{n} \sum_{i=1}^{n} A_i \qquad \text{---} (6)$$

According to central limit theorem, $\bar{A}$ is normally distributed around $\langle A \rangle$ with the standard deviation Δ_A which goes as

$$\Delta_A = \sqrt{\frac{1}{n} \text{Var}(A)} \qquad \text{---} (7)$$

for "independent" A_i, the unbiased estimator for $\text{Var}(A)$ can be written easily to give the following for Eq. (7)

$$\Delta_A = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^{n} (A_i - \bar{A})^2} \qquad \text{---} (8)$$

However, this assumes that $\{A_i\}$ are statistically independent, i.e., $\langle A_i A_j \rangle = \langle A_i \rangle \langle A_j \rangle$

But, in reality, this is not true

$$\Delta_A^2 = \frac{\text{Var}(A)}{n} + \frac{1}{n^2} \sum_{i \neq j}^{n} \left(\langle A_i A_j \rangle - \langle A \rangle^2 \right)$$

$(i \neq j = 1)$

This can be simplified to give

$$\Delta_A = \sqrt{\frac{\mathrm{Var}(A)}{n}(1+2\tau_A)} \quad \text{— (9)}$$

Eq. (9) is the generalization of Eq. (7) when $\{A_i\}$ are correlated with each other.

Here, $\tau_A = \sum_{i=1}^{\infty} \frac{\langle A_1 A_{1+t}\rangle - \langle A\rangle^2}{\langle A^2\rangle - \langle A\rangle^2}$

= integrated autocorrelation time — (10)

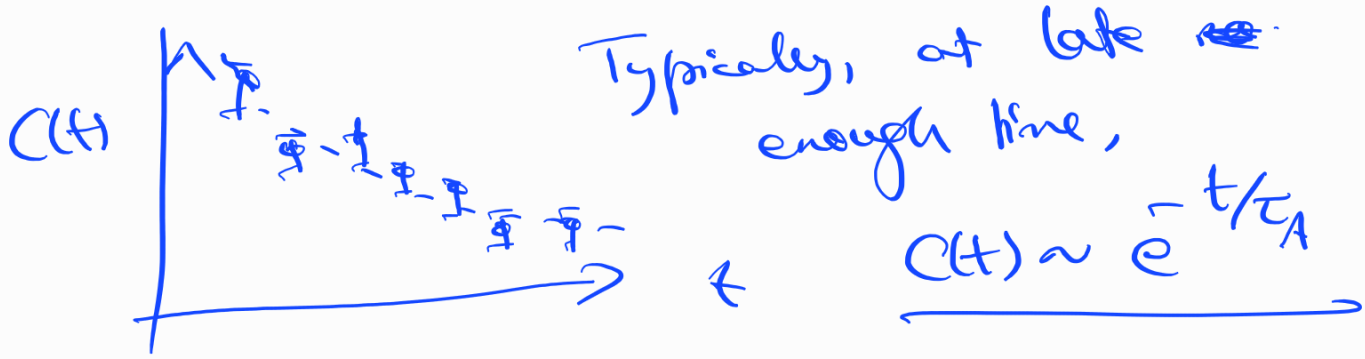
Thus, error bar increased by a factor of $\sqrt{1+2\tau_A}$ from the naive estimator.

Another handle to τ_A is provided by monitoring the autocorrelation function

$$C_A(t) = \frac{\langle A(t)A(0)\rangle - \langle A\rangle^2}{\langle A^2\rangle - \langle A\rangle^2} \quad \text{— (11)}$$

It is clear from Eq (11) that

$C_A(t) \to 1$ as $t \to 0$
and $C_A(t) \to$ [illegible] as $t \to \infty$ — (12)



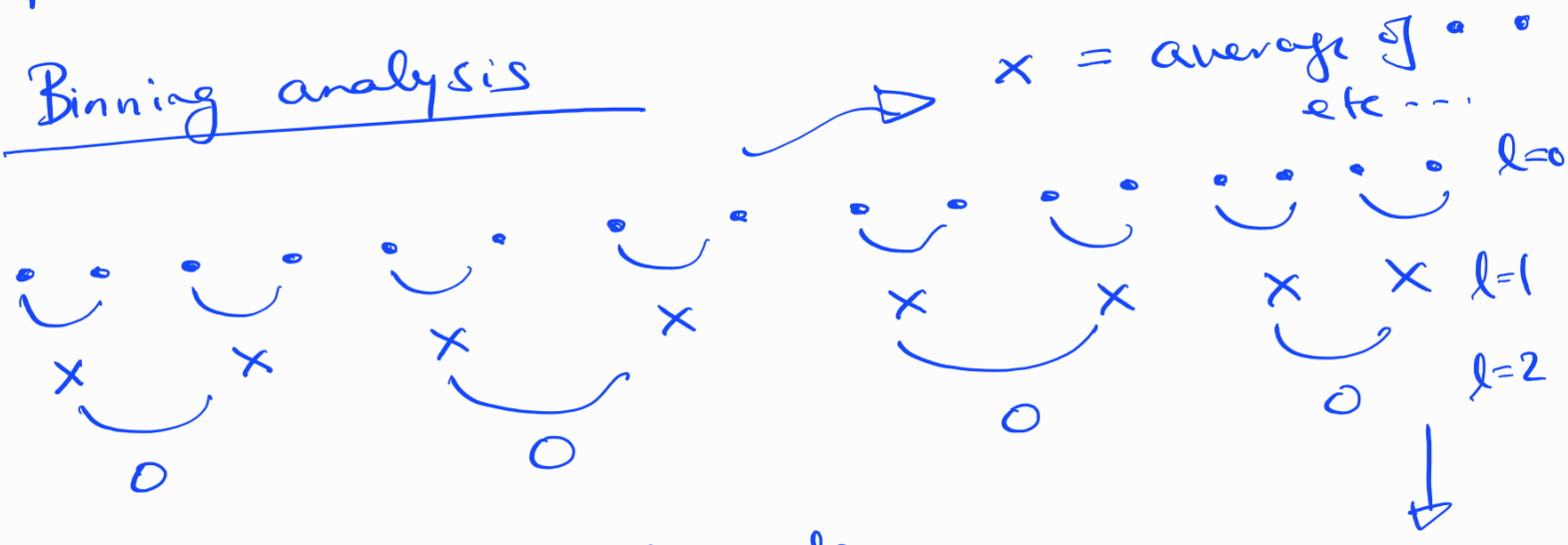
Typically, at late enough time,

$C(t) \sim e^{-t/\tau_A}$

However, reliably estimating τ_A is very difficult in any MC simulation.

We will now look at an alternate route to compute Δ_A for correlated data sets obtained from MC.

Binning analysis



$l=0$ has n elements

$l=1$ has $\frac{n}{2}$ elements

$l=2$ has $\frac{n}{2^2}$ elements

$\Rightarrow n_l = \frac{n}{2^l}$

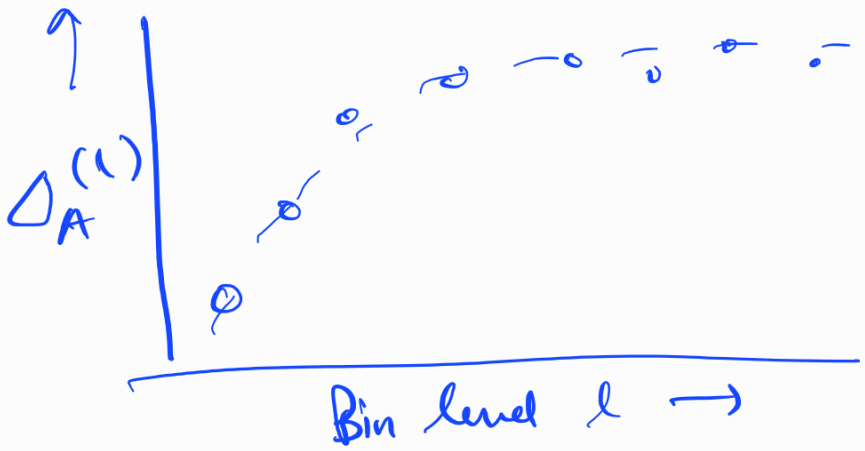
Calculate

$$\Delta_A^{(l)} = \sqrt{\frac{1}{n_l(n_l-1)} \sum_{i=1}^{n_l} (A_i^{(l)} - \bar{A})^2} \quad \text{---(13)}$$

Note that $\bar{A}$ does not change with binning.

Now, $\Delta_A = \lim_{l \to \infty} \Delta_A^{(l)}$ ---(14)

In practice, if bin size exceeds τ_A, then $\Delta_A^{(l)}$ saturates.



Also, its better to anyway store observables as bins by averaging over every M MCS where M can be chosen to be some number to (a) have a smaller data file and (b) make entries of that data file less correlated than if M=1

Several interesting quantities like

$$C_V = \frac{\partial \langle E \rangle}{\partial T} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{T^2}$$

$$\chi = \frac{\langle M^2 \rangle - \langle |M| \rangle^2}{T}$$

Binder cumulant

$$U_L = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$$

(15)

are examples of quantities which are not linear functions of quantities stored like m, E etc. In such cases, the error bars are correctly estimated using Bootstrap or

Jacknife method. Here we just discuss the Bootstrap Method

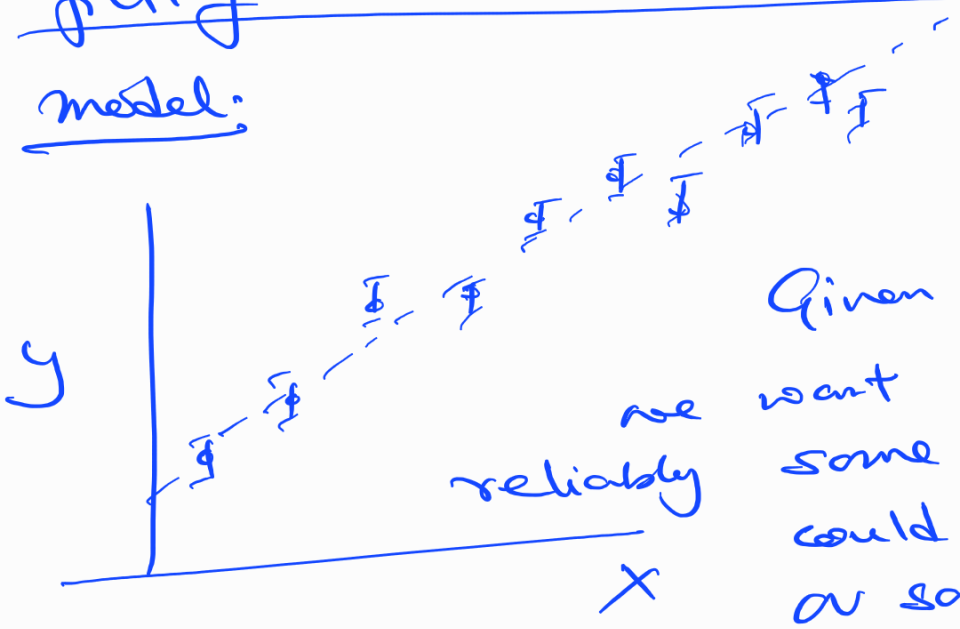
Suppose, we have "n" data points which may even represent binned data. Then, bootstrap method produces synthetic data sets, each with n data points again. For each synthetic data set, pick n data points at random where we allow picking a point more than once as well. Calculate observable like C_l, χ etc from this synthetic data set and repeat over many such data sets (typically 100 to 500 such bootstraped data sets). Then the error bar over any quantity, let us call that σ is calculated as

$$\sigma = \sqrt{\overline{C^2} - \bar{C}^2} \qquad \text{---} (16)$$

(Notice there is no extra $\frac{1}{n-1}$ factor here)

A good way to check whether this error bar is reliable is to repeat this by binning the data and checking for the stability of σ from Eq. 16.

We next move to the question of fitting the data to a particular model:



Given x_i, y_i, σ_i we want to see how reliably some $f(x)$ [which could be polynomial or something more complicated] captures the data.

Let us say that the fitting function $f(x)$ depends on M "fitting parameters"

Then χ^2 per dof is defined as

$$\chi^2 = \frac{1}{N-M} \sum_{i=1}^{N} \frac{(y_i - f(x_i))^2}{\sigma_i^2} \quad \text{— (17)}$$

Typically, chi-sq per dof $\sim O(1)$ is deemed good whereas $\chi^2 \gg 1$ is statistically unreliable. $\chi^2 \ll 1$ or

How to put error bars on the fit parameters of the function $f(x)$?

Use synthetic data with Gaussian noise (dictated by σ_i) to generate

synthetic data sets and minimize χ^2 to get fit parameters for each such data set.

Then $\sqrt{\text{variance}}$ (similar to Eq. 16) of the fit parameter values gives the error bar on that particular parameter.

IMP: Local detailed balance enough

E.g. sequential updating scheme for Ising spins does not satisfy detailed balance but nonetheless, this dynamic Markov process drives the system to equilibrium.

E.g. See "strict detailed balance is unnecessary in Monte Carlo simulations" by Manousiouthakis + Deem arXiv: cond-mat/9809240v2

Finite size scaling

↓

Useful to extract T_c and critical exponents from data at finite system sizes.

Let us review some basics first ($L \to \infty$ below)

The correlation length $\xi(T)$ scales as

$$\xi(T) \sim |t|^{-\nu} \quad \text{where} \quad t = \frac{T - T_c}{T_c} \qquad (18)$$

Also, two spin correlation

$$C(r) \sim \frac{1}{r^{d-2+\eta}} \quad \text{at } T = T_c \qquad (19)$$

$C(r) \leftarrow \langle S(r) S(0) \rangle$

Below T_c, $\lim_{r \to \infty} C(r) = \langle m^2 \rangle$

When $t \to 0^-$, $\langle m \rangle \sim |t|^{\beta}$ — (20)

Also, $\chi = \left. \frac{d\langle m \rangle}{dh} \right|_{h \to 0}$

$$= \frac{N}{T} \left(\langle m^2 \rangle - \langle |m| \rangle^2 \right)$$

$$\chi \sim |t|^{-\gamma} \qquad (21)$$

At critical point $T = T_c$, $\langle m \rangle \neq \chi h$

$$\langle m \rangle \sim h^{1/\delta} \qquad (22) \quad \text{at } T = T_c$$

Further, $C_v \sim |t|^{-\alpha}$ — (23)

(Not all exponents are independent of each other)

Finite size scaling (FSS) hypothesis assumes two important length scales ξ and L.

Consider a quantity Q which exhibits a

divergence at T_c.

$$Q \sim |t|^{-\kappa} \quad \text{— (24)}$$ in thermodynamic limit

Now, $|t| \sim \xi^{-1/\nu} \therefore \xi \sim |t|^{-\nu}$

$$\Rightarrow Q \sim \xi^{\kappa/\nu}$$

This form applies when $\xi \ll L$ but when $\xi \sim L$,

$$\boxed{Q_{max}(L) \sim L^{\kappa/\nu}} \quad \text{— (25)}$$

Also, $t_{max}(L)$ i.e. when $\xi \sim L$ scales as

$$\boxed{t_{max}(L) \sim A L^{-1/\nu}} \quad \text{— (26)}$$

* One can use both these relations (Eq. (25) & (26)) for χ and C_V peaks to extract both ν and $\frac{\alpha}{\nu}, \frac{\gamma}{\nu}$ respectively.

Suppose $Q(t, L \to \infty) \sim |t|^{-\kappa}$

where κ is negative for singular non-divergent quantities (E.g. $\kappa = -\beta$ for order parameter)

Then, $Q(t, L) = L^{\sigma} f(\xi/L)$ — (27)

Using $\xi \sim |t|^{-\nu}$

$$Q(t, L) = L^{\sigma} g(t L^{1/\nu}) \quad \text{— (28)}$$

Exactly at T_c, $Q(0, L) \sim L^{\sigma}$ from Eq (28)

Also, $Q(t, L \to \infty) \sim |t|^{-K}$ in thermodynamic limit

$\hookrightarrow \therefore$ of this, $g(x) \sim x^{-K}$ for $x \to \infty$

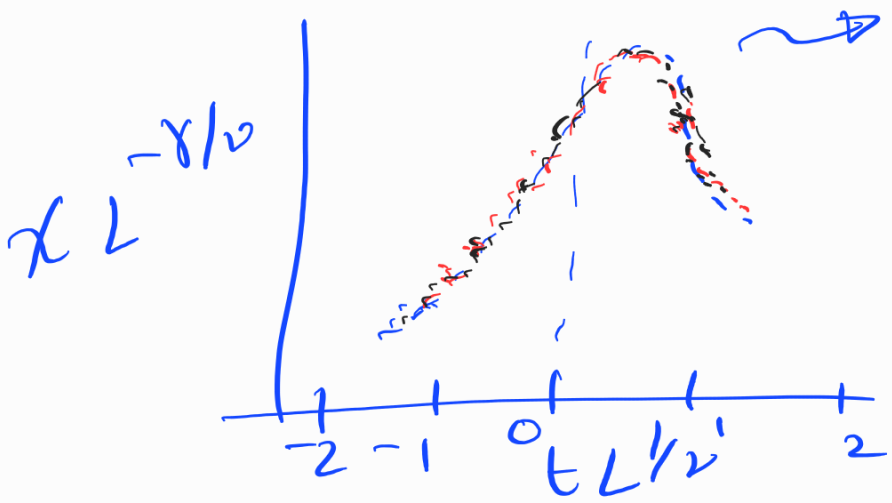
Not just that, we need to cancel the L factors in this same limit. This gives

$$Q(t, L) = L^{K/\nu}\, g(tL^{1/\nu}) \quad \text{— (29)}$$

To extract scaling function $g(x)$ using numerical data, define

$$\left.\begin{aligned} y_L &= Q(t, L)\, L^{-K/\nu} \\ x_L &= tL^{1/\nu} \end{aligned}\right\} \quad \text{— (30)}$$

Plot y_L versus x_L



$\rightsquigarrow \gamma = \frac{7}{4}, \nu = 1$ for 2D Ising

Note $y_L = g(x_L)$

where g is a smooth function

$= a_0 + a_1 x_L + a_2 x_L^2 + \cdots$

Take a small symmetric range in x_L get χ^2 fit after that.

Correction to FSS

$$Q(t,L)\, L^{-K/\nu} = g\left(t L^{\frac{1}{\nu}}\right)\left(1 + a L^{-\omega}\right) \quad \text{---(31)}$$

One possibility

Its particularly useful to perform scaling collapse of dim-0 observables where $K=0$.

Two particularly useful ones are $\frac{\xi}{L}$ and U_2.

Let $S(q) = \langle \sigma_{-q}\, \sigma_q \rangle = \sum_r \cos(\bar{q}\cdot\bar{r})\, C(r)$

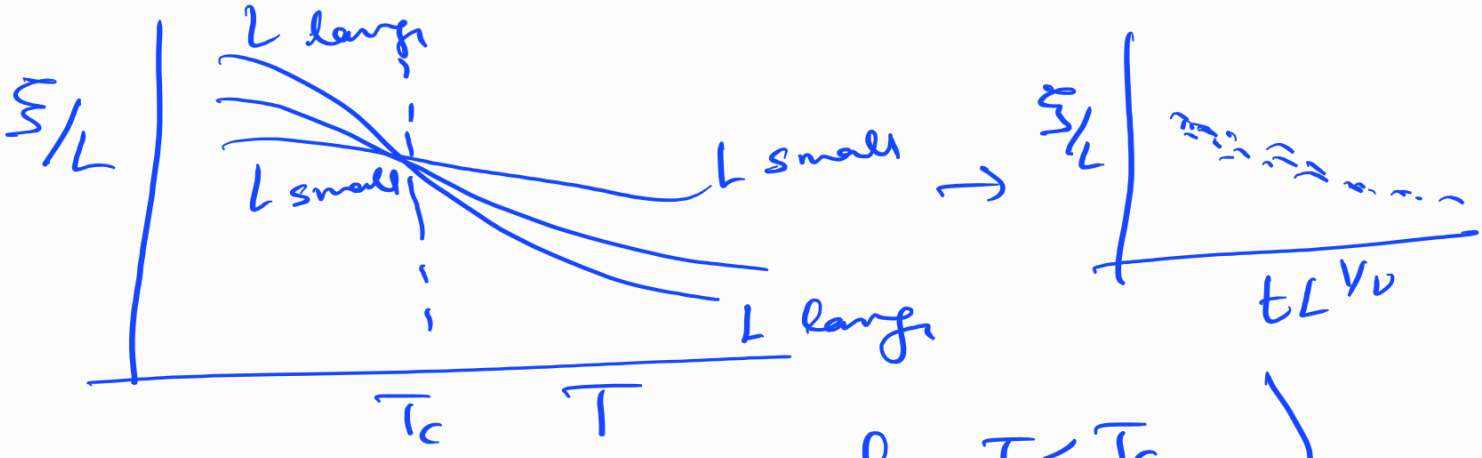
where $\sigma_q = \frac{1}{\sqrt{N}} \sum_j \sigma_j\, e^{-i\bar{q}\cdot\bar{r}_j}$ — (32)

Let $\bar{q}_1 = \left(\frac{2\pi}{L}, 0, 0 \cdots\right)$

Then, $\xi = \frac{1}{q_1}\sqrt{\frac{S(0,0,\cdots)}{S(q_1,0,0\cdots)} - 1}$ — (33)

Assume $S(q) \sim \frac{1}{q^2 + m^2}$ $\Rightarrow$ Then Eq (33)

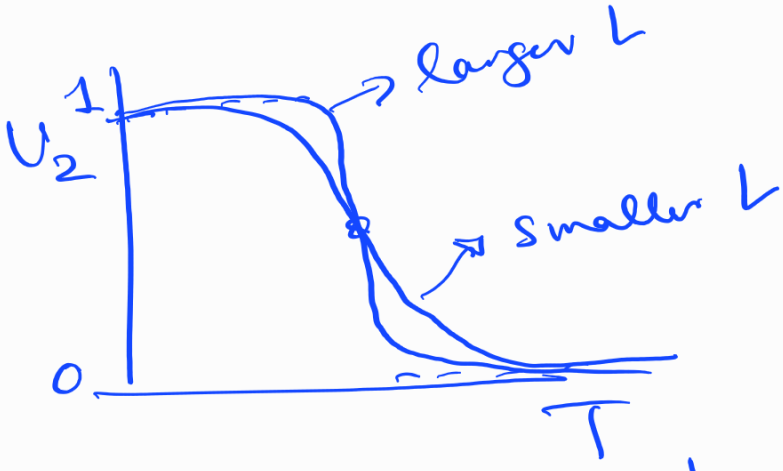
gives $\xi = \frac{1}{m}$



$$\left.\begin{aligned} \frac{\xi}{L} &\to \infty \quad \text{as } L \to \infty \text{ for } T < T_c \\ \frac{\xi}{L} &\to 0 \quad \text{as } L \to \infty \text{ for } T > T_c \\ \frac{\xi}{L} &\to G(1) \quad \text{as } L \to \infty \text{ for } T = T_c \end{aligned}\right\} \quad (34)$$

$$U_2 = \frac{3}{2}\left(1 - \frac{1}{3}\frac{\langle m^4\rangle}{\langle m^2\rangle^2}\right) \quad (35)$$

$$\left.\begin{aligned} U_2 &\to 1 \quad \text{as } L \to \infty \text{ in ordered phase} \\ U_2 &\to 0 \quad \text{as } L \to \infty \text{ in disordered phase} \\ U_2 &\to (U_2)^* \quad \text{as } L \to \infty \text{ as } T = T_c \end{aligned}\right\} \quad (36)$$

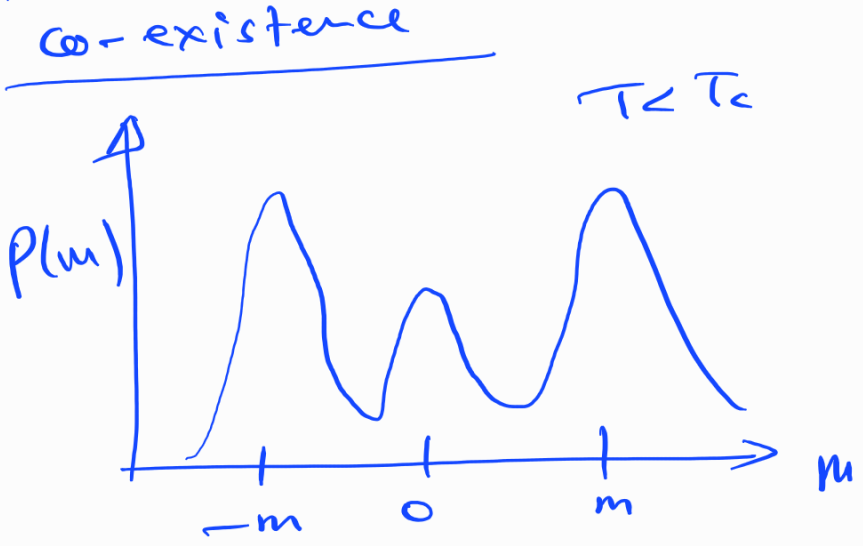


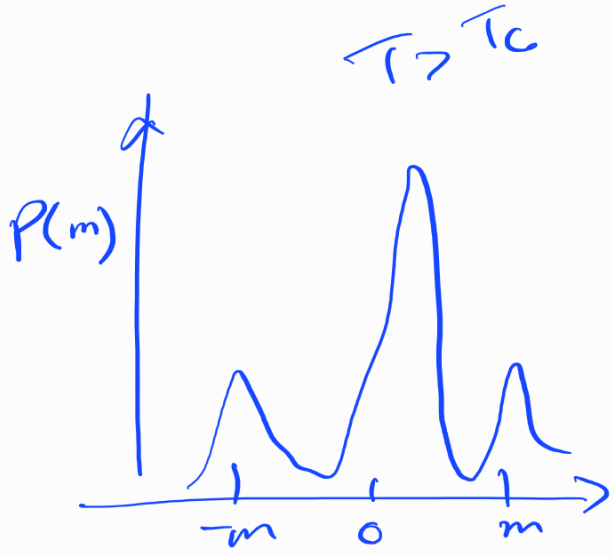
FSS - first order transitions

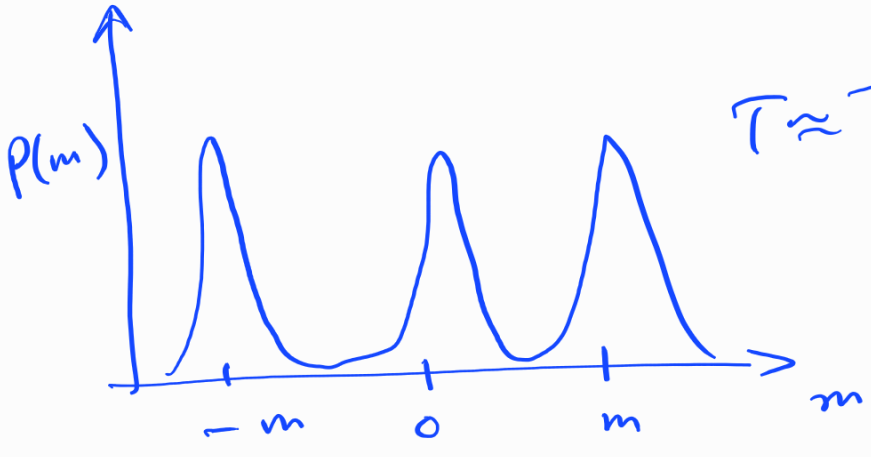
At first order (discontinuous) transition, the correlation length $\xi(T)$ remains finite and the order parameter shows a discontinuous jump.

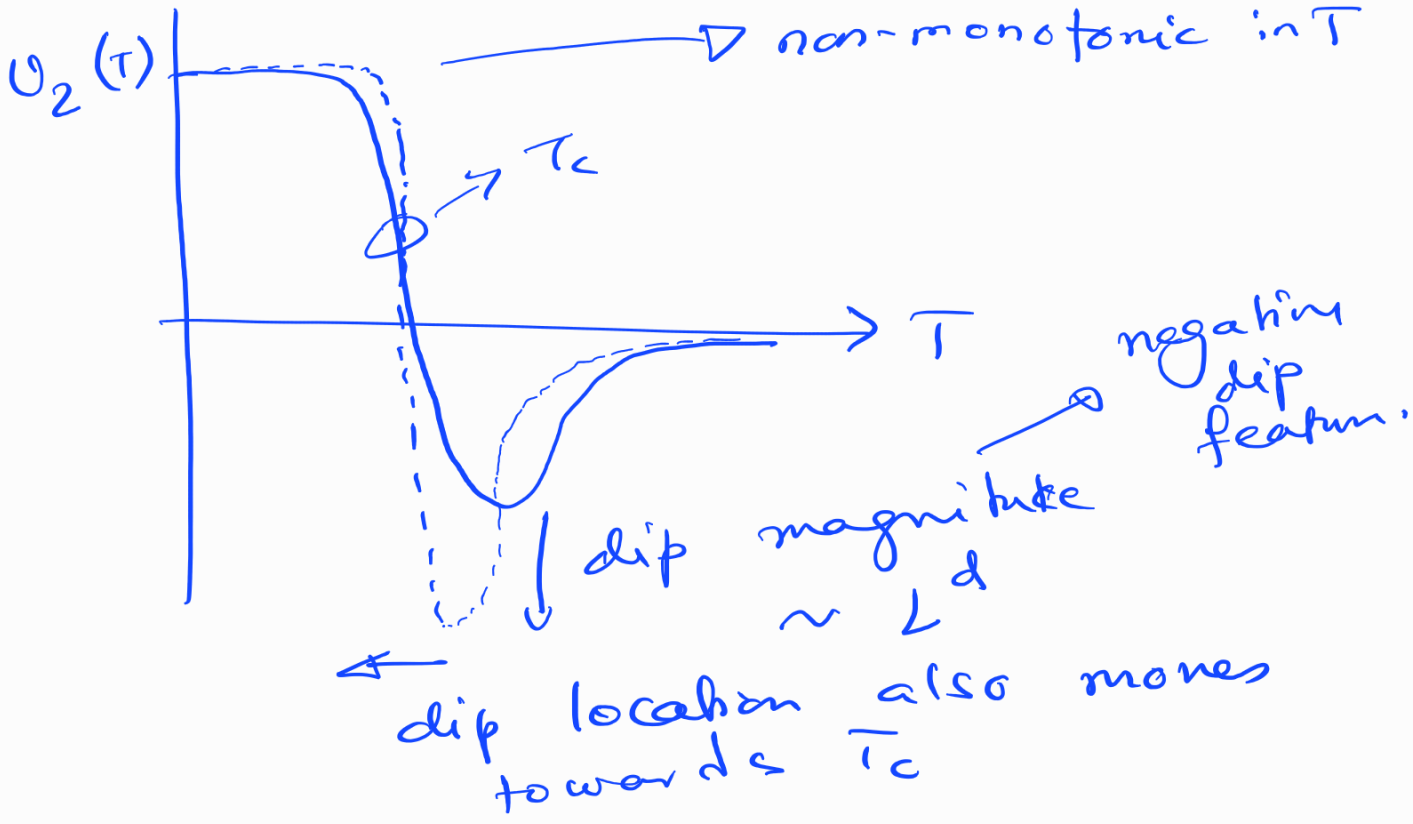
* $C_{max}(L)$ diverges as L^d instead of $L^{\alpha/\nu}$
* Shifts in critical point with system size $\sim L^{-d}$ instead of $L^{-1/\nu}$
* Weak first order difficult to distinguish from continuous transitions.
* Strong first order transitions suffers from metastability.

Binder cumulant signature of phase co-existence








$v_2(T)$
non-monotonic in T
T_c
T
negative dip feature
dip magnitude
$\sim L_d$
dip location also moves
towards T_c