

# **Computational Statistical Physics**

## Part I: Statistical Physics and Phase Transitions

---

Marina Marinkovic

March 1, 2022

ETH Zürich

Institute for Theoretical Physics

HIT G 41.5

Wolfgang-Pauli-Strasse 27  
8093 Zürich



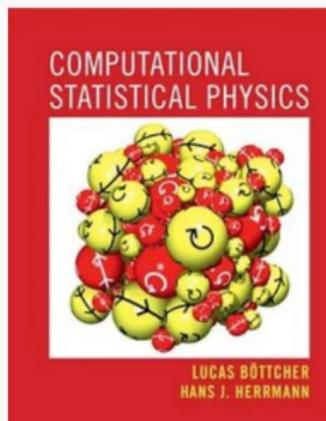
402-0812-00L  
FS 2022

## Dates and further information

- Lectures: Wednesday 9.45–11.45 in HCI J 7
- Exercises : Friday 9.45–11.45 in HPT C 103
- Tutors: Doruk Efe Gökmen and Pascal Engler
- Oral exams take place in Summer Exam Session 2022
- Moodle Course Page: <https://moodle-app2.let.ethz.ch/course/view.php?id=17219>

# Sources

- Hans J. Herrmann and Lucas Böttcher lecture notes
- Slides from FS 2019 by Lucas Böttcher
- Book by L.Böttcher and H. J. Herrmann "Computational Statistical Physics", Cambridge University Press, 2021



## **Monte Carlo methods**

---

# Monte Carlo Methods

Monte Carlo: a numerical method for estimating high-dimensional integrals by random sampling

## JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 247

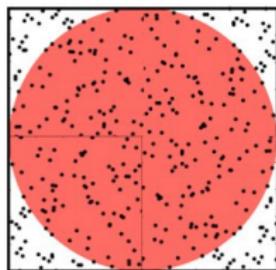
SEPTEMBER 1949

Volume 44

### THE MONTE CARLO METHOD

NICHOLAS METROPOLIS AND S. ULAM  
*Los Alamos Laboratory*

We shall present here the motivation and a general description of a method dealing with a class of problems in mathematical physics. The method is, essentially, a statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of the natural sciences.

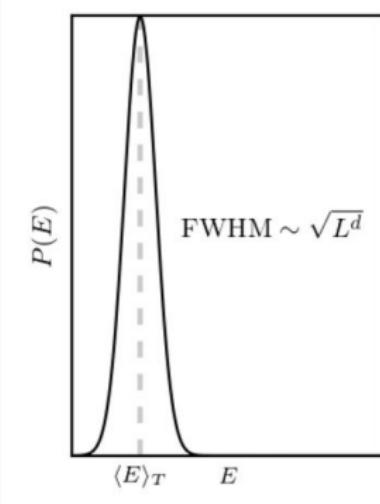


## Monte Carlo Methods

The main steps of the Monte Carlo sampling are

1. Choose randomly a new configuration in phase space based on a Markov chain.
2. Accept or reject the new configuration, depending on the strategy used (e.g., Glauber dynamics).
3. Compute the physical quantity and add it to the averaging procedure.
4. Repeat the previous steps.

# Markov chains



**Figure 1:** Example of an energy distribution with a system size  $L$  dependence of the distribution width which scales as  $\sim \sqrt{L^d}$  where  $d$  is the system dimension.

## Markov chains

In terms of a Markov chain, the transition probability from one state to another is given by the probability of a new state to be proposed ( $T$ ) and the probability of this state to be accepted ( $A$ ). Namely,  $T(X \rightarrow Y)$  is the probability that a new configuration  $Y$  is proposed, starting from configuration  $X$ . The transition probability fulfills three conditions:

1. *Ergodicity*: any configuration in the phase space must be reachable within a finite number of steps,
2. *Normalization*:  $\sum_Y T(X \rightarrow Y) = 1$ ,
3. *Reversibility*:  $T(X \rightarrow Y) = T(Y \rightarrow X)$ .

## Markov chains

Once a configuration is proposed, we can accept the new configuration with probability  $A(X \rightarrow Y)$  or reject it with probability  $1 - A(X \rightarrow Y)$ . The *probability of the Markov chain* is then given by

$$W(X \rightarrow Y) = T(X \rightarrow Y) \cdot A(X \rightarrow Y). \quad (1)$$

## Markov chains

We denote the probability to find the system in a certain configuration  $X$  at virtual time  $\tau$  by  $p(X, \tau)$ . The *master equation* describes the time evolution of  $p(X, \tau)$  and is given by

$$\frac{dp(X, \tau)}{d\tau} = \sum_Y p(Y)W(Y \rightarrow X) - \sum_Y p(X)W(X \rightarrow Y). \quad (2)$$

## Markov chains

A stationary state  $p_{\text{st}}$  is reached if  $\frac{dp(X,\tau)}{d\tau} = 0$ . The probability of the Markov chain fulfills the following properties:

1. *Ergodicity*: any configuration must be reachable:  $\forall X, Y : W(X \rightarrow Y) \geq 0$ ,
2. *Normalization*:  $\sum_Y W(X \rightarrow Y) = 1$ ,
3. *Homogeneity*:  $\sum_Y p_{\text{st}}(Y)W(Y \rightarrow X) = p_{\text{st}}(X)$ .

## Markov chains

To sample all the relevant regions of the phase space of our system, the Markov chain probability  $W(\cdot)$  has to depend on the system properties. To achieve that, we impose the distribution of the stationary states  $p_{\text{st}}$  as the equilibrium distribution of the physical system  $p_{\text{eq}}$  (a real and measurable distribution):

$$\frac{dp(X, \tau)}{d\tau} = 0 \Leftrightarrow p_{\text{st}} \stackrel{!}{=} p_{\text{eq}}. \quad (3)$$

## Markov chains

It then follows from the stationary state condition ( $\frac{dp(X,\tau)}{d\tau} = 0$ ) that

$$\sum_Y p_{\text{eq}}(Y)W(Y \rightarrow X) = \sum_Y p_{\text{eq}}(X)W(X \rightarrow Y).$$

A sufficient condition for this to be true is

$$p_{\text{eq}}(Y)W(Y \rightarrow X) = p_{\text{eq}}(X)W(X \rightarrow Y), \quad (4)$$

which is referred to as a *detailed balance condition*.

## Markov chains

As an example, in a canonical ensemble at fixed Temperature T, the equilibrium distribution is given by the Boltzmann factor

$$p_{\text{eq}}(X) = \frac{1}{Z_T} \exp \left[ -\frac{E(X)}{k_B T} \right] \quad (5)$$

with the partition function  $Z_T = \sum_X \exp \left[ -\frac{E(X)}{k_B T} \right]$ .

## M(RT)<sup>2</sup> algorithm

One possible choice of the acceptance probability fulfilling the detailed balance condition is given by

$$A(X \rightarrow Y) = \min \left[ 1, \frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} \right]. \quad (6)$$

which can be obtained by rewriting Eq. (4).

# M(RT)<sup>2</sup> algorithm (*Metropolis-Hastings algorithm*)

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

## Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,  
*Los Alamos Scientific Laboratory, Los Alamos, New Mexico*

AND

EDWARD TELLER,\* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.



## M(RT)<sup>2</sup> algorithm

In the case of the canonical ensemble with

$p_{\text{eq}}(X) = \frac{1}{Z_T} \exp\left[-\frac{E(X)}{k_B T}\right]$ , the acceptance probability becomes

$$A(X \rightarrow Y) = \min\left[1, \exp\left(-\frac{\Delta E}{k_B T}\right)\right], \quad (7)$$

where  $\Delta E = E(Y) - E(X)$ . The last equation implies that the step is always accepted if the energy decreases, and if the energy increases, it is accepted with probability  $\exp\left(-\frac{\Delta E}{k_B T}\right)$ .

## M(RT)<sup>2</sup> algorithm

In summary, the steps of the M(RT)<sup>2</sup> algorithm applied to the Ising model are

### M(RT)<sup>2</sup> algorithm

- Randomly choose a lattice site  $i$ ,
- Compute  $\Delta E = E(Y) - E(X) = 2J\sigma_i h_i$ ,
- Flip the spin if  $\Delta E \leq 0$ , otherwise accept it with probability  $\exp\left(-\frac{\Delta E}{k_B T}\right)$ ,

with  $h_i = \sum_{\langle i,j \rangle} \sigma_j$  and  $E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$ .

## Glauber dynamics

The Metropolis algorithm is not the only possible choice to fulfill the detailed balance condition. Another acceptance probability given by

$$A_G(X \rightarrow Y) = \frac{\exp\left(-\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(-\frac{\Delta E}{k_B T}\right)} \quad (8)$$

has been suggested by Glauber in 1963.

## Glauber dynamics

In contrast to the  $M(RT)^2$  acceptance probability, updates with  $\Delta E = 0$  are not always accepted but with probability 1/2.

To prove that Eq. (8) satisfies the condition of detailed balance, we have to show that

$$p_{\text{eq}}(Y)A_G(Y \rightarrow X) = p_{\text{eq}}(X)A_G(X \rightarrow Y) \quad (9)$$

since  $T(Y \rightarrow X) = T(X \rightarrow Y)$ .

## Glauber dynamics

The previous equation is equivalent to

$$\frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} = \frac{A_G(X \rightarrow Y)}{A_G(Y \rightarrow X)} \quad (10)$$

which is fulfilled since

$$\frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} = \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (11)$$

and

$$\frac{A_G(X \rightarrow Y)}{A_G(Y \rightarrow X)} = \frac{\exp\left(-\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(-\frac{\Delta E}{k_B T}\right)} \left[ \frac{\exp\left(\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(\frac{\Delta E}{k_B T}\right)} \right]^{-1} = \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (12)$$

## Glauber dynamics

As in the M(RT)<sup>2</sup> algorithm, only the local configuration around the lattice site is relevant for the update procedure.

Furthermore, with  $J = 1$ , the probability to flip spin  $\sigma_i$  is

$$A_G(X \rightarrow Y) = \frac{\exp\left(\frac{-2\sigma_i h_i}{k_B T}\right)}{1 + \exp\left(\frac{-2\sigma_i h_i}{k_B T}\right)} \quad (13)$$

with  $h_i = \sum_{\langle i,j \rangle} \sigma_j$  being the local field and

$X = \{\dots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \dots\}$  and  $Y = \{\dots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \dots\}$

the initial and final configuration, respectively.

## Glauber dynamics

We abbreviate the probability defined by Eq. (13) as  $p_i$ . The spin flip and no flip probabilities can then be expressed as

$$p_{\text{flip}} = \begin{cases} p_i & \text{for } \sigma_i = -1 \\ 1 - p_i & \text{for } \sigma_i = +1 \end{cases} \quad \text{and} \quad p_{\text{no-flip}} = \begin{cases} 1 - p_i & \text{for } \sigma_i = -1 \\ p_i & \text{for } \sigma_i = +1 \end{cases} \quad (14)$$

## Glauber dynamics

A possible implementation is

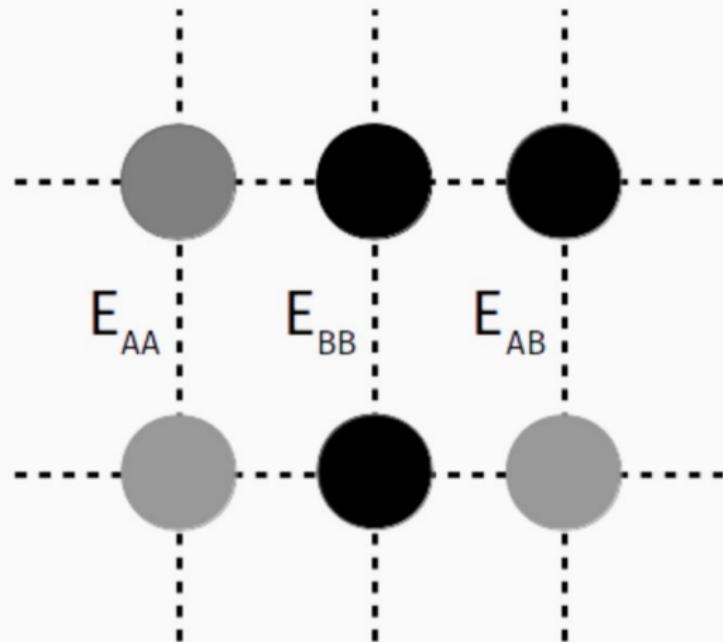
$$\sigma_i(\tau + 1) = -\sigma_i(\tau) \cdot \text{sign}(p_i - z), \quad (15)$$

with  $z \in (0, 1)$  being a uniformly distributed random number, or

$$\sigma_i(\tau+1) = \begin{cases} +1 & \text{with probability } p_i \\ -1 & \text{with probability } 1 - p_i \end{cases} \quad \text{and} \quad p_i = \frac{\exp(2\beta h_i)}{1 + \exp(2\beta h_i)}. \quad (16)$$

This method does not depend on the spin value at time  $t$  and is called *heat-bath Monte Carlo*.

## Binary mixtures



**Figure 2:** An example of a binary mixture consisting of two different atoms A and B.

## Binary mixtures

### Kawasaki dynamics

- Choose a  $A - B$  bond,
- Compute  $\Delta E$  for  $A - B \rightarrow B - A$ ,
- Metropolis: If  $\Delta E \leq 0$  flip, else flip with probability

$$p = \exp\left(\frac{-\Delta E}{k_B T}\right),$$

- Glauber: Flip with probability

$$p = \exp\left(-\frac{\Delta E}{k_B T}\right) / \left[1 + \exp\left(-\frac{\Delta E}{k_B T}\right)\right].$$

This procedure is very similar to the previously discussed update schemes. The only difference is that the magnetization is kept constant.

# Creutz Algorithm

VOLUME 50, NUMBER 19

PHYSICAL REVIEW LETTERS

9 MAY 1983

## Microcanonical Monte Carlo Simulation

Michael Creutz

*Department of Physics, Brookhaven National Laboratory, Upton, New York 11973*

(Received 24 February 1983)

A new algorithm for the simulation of statistical systems is presented. The procedure produces a random walk through configurations of a constant total energy. It is computationally simple and applicable to systems of both discrete and continuous variables.

**Figure 3:** An algorithm to perform microcanonical Monte Carlo simulations, i.e., system at constant energy.

## Creutz Algorithm

The movement in phase space is in fact not strictly constrained to a subspace of constant energy but there is a certain additional volume in which we can freely move. The condition often constant energy is softened by introducing a so-called *demon* which corresponds to a small reservoir of energy  $E_D$  that can store a certain maximum energy  $E_{\max}$ .

# Creutz Algorithm

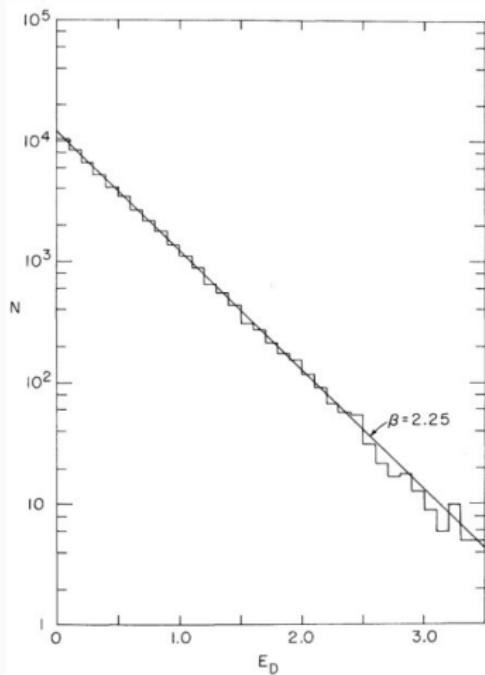
## Creutz algorithm

- Choose a site,
- Compute  $\Delta E$  for the spin flip,
- Accept the change if  $E_{\max} \geq E_D - \Delta E \geq 0$ .

**Pro:** Besides the fact that we can randomly choose a site, this method involves no random numbers and is thus said to be completely deterministic and therefore reversible.

**Con:** The temperature of the system is not known.

## Creutz Algorithm



**Figure 4:** The distribution of the demon energy  $E_D$  is exponentially distributed. Based on the Boltzmann factor, it is possible to extract the inverse temperature  $\beta = (k_B T)^{-1} = 2.25$ . The figure is taken from Ref. shown in Figure 3.

## Boundary conditions

For finite lattices, the following boundary conditions might be used:

- Open boundaries, i.e., no neighbors at the edges of the system,
- fixed boundary conditions,
- and periodic boundaries.

## Correlation length

The correlation function is defined by

$$G(r_1, r_2; T, H) = \langle \sigma_1 \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle, \quad (25)$$

where the vectors  $r_1$  and  $r_2$  pointing in the direction of lattice sites 1 and 2. If the system is translational and rotational invariant, the correlation function only depends on  $r = |r_1 - r_2|$ . At the critical point, the correlation function decays as

$$G(r; T_c, 0) \sim r^{-d+2-\eta}, \quad (26)$$

where  $\eta$  is another critical exponent and  $d$  the dimension of the system.

- 2D:  $\eta = 1/4$
- 3D:  $\eta \approx 0.036$

## Correlation length

For temperatures away from the critical temperature, the correlation function exhibits an exponential decay

$$G(r; T, 0) \sim r^{-\vartheta} e^{-r/\xi}, \quad (27)$$

where  $\xi$  defines the *correlation length*. The exponent  $\vartheta$  equals 2 above and 1/2 below the transition point. In the vicinity of  $T_c$ , the correlation length  $\xi$  diverges since

$$\xi(T) \sim |T - T_c|^{-\nu} \quad (28)$$

- 2D:  $\nu = 1$
- 3D:  $\nu \approx 0.63$

## Critical exponents and universality

The aforementioned six critical exponents are connected by four scaling laws

$$\alpha + 2\beta + \gamma = 2 \quad (\text{Rushbrooke}), \quad (29)$$

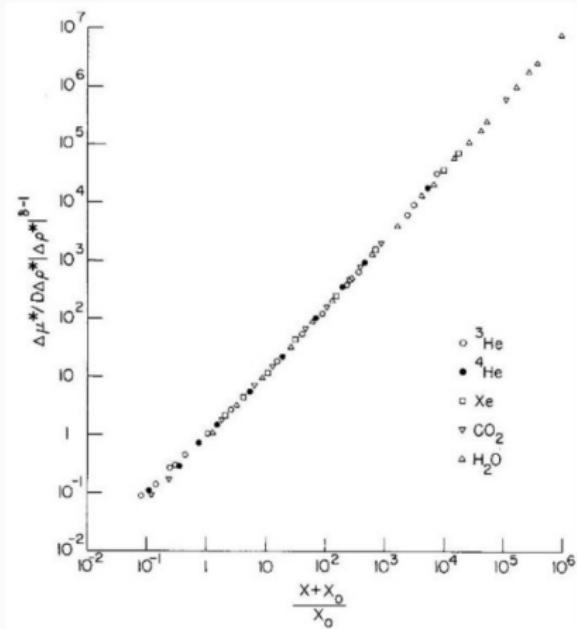
$$\gamma = \beta(\delta - 1) \quad (\text{Widom}), \quad (30)$$

$$\gamma = (2 - \eta)\nu \quad (\text{Fisher}), \quad (31)$$

$$2 - \alpha = d\nu \quad (\text{Josephson}), \quad (32)$$

which have been derived in the context of the phenomenological scaling theory for ferromagnetic systems. Due to these relations, the number of independent exponents reduces to two.

## Critical exponents and universality



**Figure 6:** Universal scaling for five different gases. The scaling variable is defined as  $x = \Delta T |\Delta \rho|^{-1/\beta}$  and  $x_0$  depends on the amplitude  $B$  of the power-law for the coexistence curve  $\Delta \rho = B \Delta T^\beta$

[Sengers, Sengers, Croxton, Prog. in liquid physics, Wiley, 1978]

## Critical exponents and universality

**Table 1:** The critical exponents of the Ising model in two and three dimensions [Pelissetto, Vicari, Phys. Rep. 368, 549–727 (2002)]

Exponent	$d = 2$	$d = 3$
$\alpha$	0	0.110(1)
$\beta$	1/8	0.3265(3)
$\gamma$	7/4	1.2372(5)
$\delta$	15	4.789(2)
$\eta$	1/4	0.0364(5)
$\nu$	1	0.6301(4)

## Temporal Correlations

According to the definition of a Markov chain, the dependence of a quantity  $A$  on virtual time  $\tau$  is given by

$$\langle A(\tau) \rangle = \sum_X p(X, \tau) A(X) = \sum_X p(X, \tau_0) A(X(\tau)). \quad (17)$$

In the second step of the latter equation, we used the fact that the average is taken over an ensemble of initial configurations  $X(\tau_0)$  which evolve according to Eq. (2).

## Temporal Correlations

For some  $\tau_0 < \tau$ , the *non-linear correlation function*

$$\Phi_A^{\text{nl}}(\tau) = \frac{\langle A(\tau) \rangle - \langle A(\infty) \rangle}{\langle A(\tau_0) \rangle - \langle A(\infty) \rangle} \quad (18)$$

is a measure to quantify the deviation of  $A(\tau)$  from  $A(\infty)$  relative to the deviation of  $A(\tau_0)$  from  $A(\infty)$ .

## Temporal Correlations

The *non-linear* correlation time  $\tau_A^{\text{nl}}$  describes the relaxation towards equilibrium and is defined as<sup>1</sup>

$$\tau_A^{\text{nl}} = \int_0^\infty \Phi_A^{\text{nl}}(\tau) d\tau. \quad (20)$$

---

<sup>1</sup>If we consider an exponential decay of  $\Phi_A^{\text{nl}}(\tau)$ , we find that this definition is meaningful since

$$\int_0^\infty \exp\left(-\tau/\tau_A^{\text{nl}}\right) d\tau = \tau_A^{\text{nl}}. \quad (19)$$

## Temporal Correlations

In the vicinity of the critical temperature  $T_c$ , we observe the so-called *critical slowing down* of our dynamics, i.e., the non-linear correlation time is described by power law

$$\tau_A^{\text{nl}} \sim |T - T_c|^{-z_A^{\text{nl}}} \quad (21)$$

with  $z_A^{\text{nl}}$  being the non-linear dynamical critical exponent. This implies that the time needed to reach equilibrium diverges at  $T_c$ !

## Temporal Correlations

The linear correlation function of two values  $A, B$  is defined as

$$\Phi_{AB}(\tau) = \frac{\langle A(\tau_0)B(\tau) \rangle - \langle A \rangle \langle B \rangle}{\langle AB \rangle - \langle A \rangle \langle B \rangle} \quad (22)$$

with

$$\langle A(\tau_0)B(\tau) \rangle = \sum_X p(X, \tau_0) A(X(\tau_0)) B(X(\tau)).$$

As  $\tau$  goes to infinity,  $\Phi_{AB}(\tau)$  decreases from unity to zero.

## Temporal Correlations

If  $A = B$ , we call Eq. (22) the *autocorrelation function*. For the spin-spin correlation in the Ising model we obtain

$$\Phi_{\sigma}(\tau) = \frac{\langle \sigma(\tau_0)\sigma(\tau) \rangle - \langle \sigma(\tau_0) \rangle^2}{\langle \sigma^2(\tau_0) \rangle - \langle \sigma(\tau_0) \rangle^2}$$

## Temporal Correlations

The *linear* correlation time  $\tau_A^{\text{nl}}$  describes the relaxation towards equilibrium

$$\tau_{AB} = \int_0^\infty \Phi_{AB}(\tau) d\tau. \quad (23)$$

## Temporal Correlations

As in the case of the non-linear correlation time, in the vicinity of  $T_c$ , we observe a *critical slowing down*, i.e.,

$$\tau_{AB} \sim |T - T_c|^{-z_A}. \quad (24)$$

with  $z_A$  being the *linear* dynamical critical exponent.

## Temporal Correlations

The dynamical exponents for spin correlations turn out to be

$$z_\sigma = 2.16 \text{ (2D)},$$

$$z_\sigma = 2.09 \text{ (3D)}.$$

There is a conjectured relation between the Ising critical exponents and the critical dynamical exponents for spin  $\sigma$  and energy correlations  $E$ . The relations

$$z_\sigma - z_\sigma^{\text{nl}} = \beta, \tag{25}$$

$$z_E - z_E^{\text{nl}} = 1 - \alpha, \tag{26}$$

$$\tag{27}$$

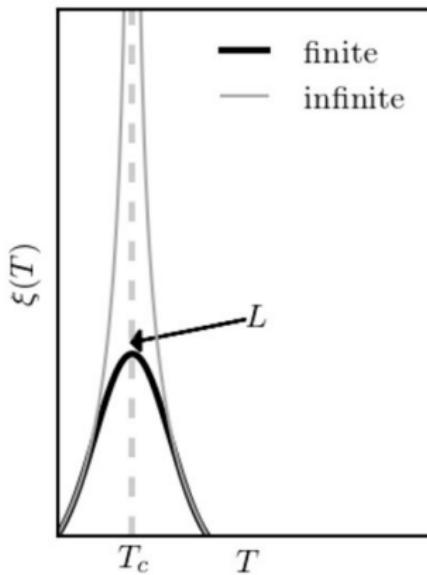
are numerically well-established, however, not yet analytically proven.

## Decorrelated configurations

Connecting this behavior with the one observed for the correlation time described by Eq. (23) yields

$$\tau_{AB} \sim |T - T_c|^{-z_{AB}} \sim L^{\frac{z_{AB}}{\nu}} \quad (28)$$

## Decorrelated configurations



**Figure 5:** The correlation length diverges in an infinite system at  $T_c$  according to the definition of the correlation length from last week's lecture. In a finite system, however, we observe a round off and the correlation length approaches the system size  $L$  at  $T_c$ .

## Decorrelated configurations

To ensure not to sample correlated configurations one should

- first reach equilibrium (discard  $n_0 = c\tau^{\text{nl}}(T)$  configurations),
- only sample every  $n_e^{\text{th}} = c\tau(T)$  configuration,
- and at  $T_c$  use  $n_0 = cL^{\frac{z^{\text{nl}}}{\nu}}$  and  $n_e = cL^{\frac{z}{\nu}}$

where  $c \approx 3$  is a "safety factor" to make sure to discard enough samples.