Improvement of Monte Carlo estimates with covariance-optimized finite-size scaling at fixed phenomenological coupling

Francesco Parisen Toldin

Max Planck Institute for Physics of Complex Systems
Dresden

25 November 2011

Monte Carlo methods

- Powerful and flexible method to study critical phenomena
- General feature:

$$\text{accuracy} \propto \frac{1}{\sqrt{\text{computational time}}}$$

- In order to improve the accuracy: improved estimators
- Use of covariance analysis:
  - add control variates, whose expectation value vanish[1]
  - compute the optimal weighted average of different estimates of a critical exponent[2]
- Our method:
  - Optimization of a Finite-Size Scaling method
  - It allows for a significant reduction of the error bars
  - It does not require additional computational time

[1] L. A. Fernández, V. Martín-Mayor, Phys. Rev. E **79**, 051109 (2009)
[2] M. Weigel, W. Janke, Phys. Rev. Lett. **102**, 100601 (2009); Phys. Rev. E **81**, 066701 (2010)

Finite-Size Scaling

- A system in a finite volume of linear size L
- Finite-Size Scaling (FSS) works in region of parameters where $\xi \sim L$
- FSS behavior of long-ranged quantities

$$O = L^x f(tL^{1/\nu}), \quad t \equiv \frac{T - T_c}{T_c}$$

E.g. Susceptibility $\chi \propto L^{2-\eta} f(tL^{1/\nu})$

- Renormalization-Group invariant quantities R

$$R = F(tL^{1/\nu})$$

E.g. $R = \xi/L$, Binder ratios, etc...
Also called phenomenological couplings

Finite-Size Scaling at fixed phenomenological coupling

- We fix a phenomenological coupling R to a constant R_f

$$R(\beta, L) = R_f$$

$\Rightarrow \beta_f(L)$ such that $R(\beta_f(L), L) = R_f$

- From FSS relation $R = F(tL^{1/\nu})$, $t = (T - T_c)/T_c = \beta_c/\beta - 1$:

$$\beta_f(L) - \beta_c \propto L^{-1/\nu}, \qquad \text{generic } R_f$$

$$\beta_f(L) - \beta_c \propto L^{-1/\nu-\omega}, \qquad R_f = R^* = F(0)$$

- All the other observables $O(\beta, L)$ are calculated at $\beta = \beta_f(L)$

$$\chi(\beta_f(L), L) \propto L^{2-\eta}, \qquad \text{susceptibility}$$

$$\frac{\partial R}{\partial \beta}(\beta_f(L), L) \propto L^{1/\nu}$$

M. Hasenbusch, J. Phys. A **32**, 4851 (1999)

Finite-Size Scaling at fixed phenomenological coupling

- It does not require a precise knowledge of β_c
- Reduced error bars by fixing ξ/L:
  - Fully frustrated XY model[1]
  - Randomly Dilute Ising Universality class[2]
  - Ising model in $d = 3, 4, 5$[3]
- Similarities with the Phenomenological Renormalization method:
  - Two system sizes are enforced to share a common value of ξ/L
  - Reported reduction of error bars[4]

[1] M. Hasenbusch, A. Pelissetto, E. Vicari, JSTAT **P12002** (2005)
[2] M. Hasenbusch, F. Parisen Toldin, A. Pelissetto, E. Vicari, JSTAT **P02016** (2007)
[3] U. Wolff, Phys. Rev. D **79**, 105002 (2009)
[4] H. G. Ballesteros, L. A. Fernández, V. Martín-Mayor, A. Muñoz-Sudupe, Nucl. Phys. B **483**, 707 (1997)

Finite-Size Scaling at fixed phenomenological coupling

Why there is a reduction of the error bars?

Notation:

- Simulations at $\beta = \beta_{\rm run}$
- An observable O is calculated from a statistical estimator $\widehat{O}$

$$\widehat{O} = O + \widehat{\delta O}, \qquad O = E[\widehat{O}]$$

- We choose to fix a phenomenological coupling R sampled using the estimator $\widehat{R}$
- We fix $R(\beta = \beta_f(L), L) = R_f$ and calculate $O_f(L) \equiv O(\beta = \beta_f, L)$

Finite-Size Scaling at fixed phenomenological coupling

- For β close to $\beta_{\rm run}$

$$\widehat{R}(\beta) \simeq \widehat{R} + \widehat{R}'(\beta - \beta_{\rm run}), \qquad \widehat{R} = \widehat{R}(\beta_{\rm run}), \quad R' = \partial R/\partial\beta.$$

- Solving $\widehat{R} = R_f$

$$\widehat{\beta}_f = \beta_{\rm run} - \frac{\widehat{R} - R_f}{\widehat{R}'}$$

$\Rightarrow$ we trade the fluctuations of R for the fluctuations of β_f

- For a generic observable O calculated at β_f

$$\widehat{O}_f \simeq \widehat{O} + \widehat{O'}(\beta - \beta_{\rm run}) = \widehat{O} - \widehat{O'}\frac{\widehat{R} - R_f}{\widehat{R}'}.$$

- The variance of $\widehat{O}_f$ is, to the lowest order in the fluctuations

$$VAR[\widehat{O}_f] = VAR[\widehat{O}] + \left(\frac{O'}{R'}\right)^2 VAR[\widehat{R}] - 2\frac{O'}{R'}COV[\widehat{O}, \widehat{R}]$$

Optimization of Finite-Size Scaling

Can we optimize the Finite-Size Scaling?

- Consider N phenomenological couplings $R_1, \ldots, R_N$
- We define the RG-invariant quantity $R(\{\lambda_i\}) \equiv \sum_i \lambda_i R_i$
- We consider FSS at fixed phenomenological coupling $R(\{\lambda_i\})$

Problem: given an observable O, what are the coefficients $\{\lambda_i\}$ that minimize the value of O_f?

Minimize: $$VAR[\widehat{O}_f] = VAR[\widehat{O}] + \left(\frac{O'}{R'}\right)^2 VAR[\widehat{R}] - 2\frac{O'}{R'} COV[\widehat{O}, \widehat{R}]$$

Solution: $$\lambda_i = -\frac{\mathbf{R'}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{N} - O'}{\mathbf{R'}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{R'}} \left(\mathbf{M}^{-1}\mathbf{R'}\right)_i + \left(\mathbf{M}^{-1}\mathbf{N}\right)_i ,$$

with $\mathbf{M}_{ij} \equiv COV[\widehat{R}_i, \widehat{R}_j]$, $\mathbf{N}_i \equiv COV[\widehat{O}, \widehat{R}_i]$, $\mathbf{R'}_i \equiv R'_i$.

Optimization of Finite-Size Scaling

Can we optimize the Finite-Size Scaling?

- Consider N phenomenological couplings $R_1, \ldots, R_N$
- We define the RG-invariant quantity $R(\{\lambda_i\}) \equiv \sum_i \lambda_i R_i$
- We consider FSS at fixed phenomenological coupling $R(\{\lambda_i\})$

Problem: given an observable O, what are the coefficients $\{\lambda_i\}$ that minimize the value of O_f?

$$\text{Minimize:} \quad VAR[\widehat{O}_f] = VAR[\widehat{O}] + \left(\frac{O'}{R'}\right)^2 VAR[\widehat{R}] - 2\frac{O'}{R'} COV[\widehat{O}, \widehat{R}]$$

$$\text{Solution:} \quad \lambda_i = -\frac{\mathbf{R'}^\mathbf{T}\mathbf{M}^{-1}\mathbf{N} - O'}{\mathbf{R'}^\mathbf{T}\mathbf{M}^{-1}\mathbf{R'}} \left(\mathbf{M}^{-1}\mathbf{R'}\right)_i + \left(\mathbf{M}^{-1}\mathbf{N}\right)_i,$$

with $\mathbf{M}_{ij} \equiv COV[\widehat{R}_i, \widehat{R}_j]$, $\mathbf{N}_i \equiv COV[\widehat{O}, \widehat{R}_i]$, $\mathbf{R'}_i \equiv R'_i$.

Optimization of Finite-Size Scaling

Can we optimize the Finite-Size Scaling?

- Consider N phenomenological couplings $R_1, \ldots, R_N$
- We define the RG-invariant quantity $R(\{\lambda_i\}) \equiv \sum_i \lambda_i R_i$
- We consider FSS at fixed phenomenological coupling $R(\{\lambda_i\})$

Problem: given an observable O, what are the coefficients $\{\lambda_i\}$ that minimize the value of O_f?

Minimize: $$VAR[\widehat{O}_f] = VAR[\widehat{O}] + \left(\frac{O'}{R'}\right)^2 VAR[\widehat{R}] - 2\frac{O'}{R'} COV[\widehat{O}, \widehat{R}]$$

Solution: $$\lambda_i = -\frac{\mathbf{R'}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{N} - O'}{\mathbf{R'}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{R'}} \left(\mathbf{M}^{-1}\mathbf{R'}\right)_i + \left(\mathbf{M}^{-1}\mathbf{N}\right)_i,$$

with $\mathbf{M}_{ij} \equiv COV[\widehat{R}_i, \widehat{R}_j]$, $\mathbf{N}_i \equiv COV[\widehat{O}, \widehat{R}_i]$, $\mathbf{R'}_i \equiv R'_i$.

Some observations

- Finite-Size Scaling at fixed $R(\{\lambda_i\}) \equiv \sum_i \lambda_i R_i$, with

$$\lambda_i = -\frac{\mathbf{R}'^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{N} - O'}{\mathbf{R}'^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{R}'}\left(\mathbf{M}^{-1}\mathbf{R}'\right)_i + \left(\mathbf{M}^{-1}\mathbf{N}\right)_i,$$
$$\mathbf{M}_{ij} \equiv COV[\widehat{R}_i, \widehat{R}_j], \qquad \mathbf{N}_i \equiv COV[\widehat{O}, \widehat{R}_i].$$

- M and N are related to the transition matrix of the Markov chain $\Rightarrow$ they depend on the model and on the dynamics
- $\{\lambda_i\}$ can be optimized separately for every observable O
- FSS limit is correctly defined when $\{\lambda_i\}$ are the same for all sizes.
  The optimal $\{\lambda_i\}$ depend on the lattice size.
  Possible strategy:
  choose the optimal $\{\lambda_i\}$ obtained from the largest available lattice.

Test: Ising model in $d = 2, 3$

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \qquad \sigma_i = \pm 1.$$

- Four RG-invariant quantities:

$$U_4 \equiv \frac{\langle M^4 \rangle}{\langle M^2 \rangle^2}, \qquad U_6 \equiv \frac{\langle M^6 \rangle}{\langle M^2 \rangle^3}$$

$$R_\xi \equiv \xi / L, \qquad R_Z \equiv Z_a / Z_p^1,$$

with $M \equiv \sum_i \sigma_i$ magnetization, ξ second-moment correlation length
Z_a partition function with antiperiodic b.c. on one direction,
Z_p partition function with fully periodic b.c.

- Observables

$$\chi \equiv \sum_i \sigma_i \sigma_j / V \propto L^{2-\eta}$$

$$\frac{\partial U_4}{\partial \beta}, \frac{\partial U_6}{\partial \beta}, \frac{\partial R_\xi}{\partial \beta} \propto L^{1/\nu}$$

[1] M.Hasenbusch, Physica A **197**, 423 (1993)

Results $d = 2$

- Gain in CPU time is given by $\left(\frac{\text{standard error bar}}{\text{error bar at fixed } R}\right)^2$
- CPU gain for Metropolis dynamics

$$\begin{aligned} \chi\text{: gain} &\sim 20; & \text{gain at fixed } R_\xi &\lesssim 4 \\ \partial U_4/\partial\beta, \partial U_6/\partial\beta\text{: gain} &\sim 30-50; & \text{gain at fixed } U_4, U_6 &\sim 6-10 \\ \partial R_\xi/\partial\beta\text{: gain} &\sim 2-3; & \text{gain at fixed } U_4, U_6, R_\xi, R_Z &\lesssim 1 \end{aligned}$$

- CPU gain for Wolff single-cluster dynamics

$$\begin{aligned} &\chi\text{: gain} \sim 6; & \text{gain at fixed } R_\xi &\sim 3 \\ &\partial U_4/\partial\beta, \partial U_6/\partial\beta\text{: gain} \sim 6-9; & \text{gain at fixed } U_4, U_6 &\sim 2-4 \\ &\partial R_\xi/\partial\beta\text{: gain} \sim 2-3; & \text{gain at fixed } U_4, U_6, R_\xi, R_Z &\lesssim 1 \end{aligned}$$

- Gain in computational time roughly independent on $L = 8 - 128$

Results $d = 3$

- Gain in CPU time is given by $\left(\frac{\text{standard error bar}}{\text{error bar at fixed } R}\right)^2$
- CPU gain for Metropolis dynamics

$$\begin{aligned} \chi\text{: gain} \sim 20-30; &\qquad \text{gain at fixed } R_\xi \lesssim 9 \\ \partial U_4/\partial\beta, \partial U_6/\partial\beta\text{: gain} \sim 20-30; &\qquad \text{gain at fixed } U_4, U_6 \lesssim 3 \\ \partial R_\xi/\partial\beta\text{: gain} \sim 4-6; &\qquad \text{gain at fixed } U_4, U_6, R_\xi, R_Z \lesssim 1 \end{aligned}$$

- CPU gain for Wolff single-cluster dynamics

$$\begin{aligned} \chi\text{: gain} \sim 6-10; &\qquad \text{gain at fixed } R_\xi \sim 5 \\ \partial U_4/\partial\beta, \partial U_6/\partial\beta\text{: gain} \sim 5-8; &\qquad \text{gain at fixed } U_4, U_6 \sim 1-2 \\ \partial R_\xi/\partial\beta\text{: gain} \sim 3; &\qquad \text{gain at fixed } U_4, U_6, R_\xi, R_Z \lesssim 1 \end{aligned}$$

- Gain in computational time roughly independent on $L = 8 - 128$

Outlook

- Substantial reduction of the error bars
- It does not require additional computational time
- Test on the Ising model
  - CPU gains are roughly independent on the lattice L
  - CPU gains are more pronounced for Metropolis update
- Other possible applications:
  - "Improved models"
  - Models with quenched disorder
  - ...

Ref: F. Parisen Toldin, Phys. Rev. E **84**, 025703(R)