The Quality of Finite-Size Data Collapse

Andreas Sorge

September 2, 2014

1 Finite-size scaling

The finite-size scaling ansatz

$$A_L(\varrho) = L^{\zeta/\nu} \tilde{f}\left(L^{1/\nu}(\varrho - \varrho_c)\right)$$

postulates how a physical quantity $A_L(\varrho)$ observed in a system of finite size scales with system size L and paramater ϱ according to a scaling function $\tilde{f}$, the critical parameter ϱ_c, the critical exponent ζ of the quantity itself, and the critical exponent ν of the correlation length ξ [1], [2].

Finite-size scaling analysis concerns experimental data a_{L_i,ϱ_j} at system sizes L_i and parameter values ϱ_j. Plotting $L_i^{-\zeta/\nu} a_{L_i,\varrho_j}$ against $L_i^{1/nu}(\varrho - \varrho_c)$ with the right choice of ϱ_c, ν, ζ should let the data collapse onto a single curve. The single curve of course is the scaling function $\tilde{f}$ from the finite-size scaling ansatz. In the following, we present a measure by Houdayer & Hartmann [3] for the quality of the data collapse. Melchert [4] refers to some alternative measures, for example [5], [6], and to some applications of these measures in the literature.

2 The quality function

Houdayer & Hartmann [3] refine a method proposed by Kawashima & Ito [7]. They define the quality as the reduced χ^2 statistic

$$S = \frac{1}{\mathcal{N}} \sum_{i,j} \frac{(y_{ij} - Y_{ij})^2}{dy_{ij}^2 + dY_{ij}^2}, \tag{1}$$

where the values y_{ij}, dy_{ij} are the scaled observations and its standard errors at x_{ij}, and the values Y_{ij}, dY_{ij} are the estimated value of the master curve and its standard error at x_{ij}.

The quality S is the mean square of the weighted deviations from the master curve. As we expect the individual deviations $y_{ij} - Y_{ij}$ to be of the order of the individual error $\sqrt{dy_{ij}^2 + dY_{ij}^2}$ for an optimal fit, the quality S should attain its minimum $S_{\min}$ at around 1 and be much larger otherwise [8].

Let i enumerate the system sizes L_i, $i = 1, \ldots, k$ and let j enumerate the parameters ϱ_j, $j = 1, \ldots, n$ with $\varrho_1 < \varrho_2 < \ldots < \varrho_n$. The scaled data are

$$y_{ij} := L_i^{-\zeta/\nu} a_{L_i,\varrho_j} \tag{2}$$

$$dy_{ij} := L_i^{-\zeta/\nu} da_{L_i,\varrho_j} \tag{3}$$

$$x_{ij} := L_i^{1/\nu}(\varrho_j - \varrho_c). \tag{4}$$

The sum in the quality function S only involves terms for which the estimated value Y_{ij} of the master curve at x_{ij} is defined. The number of such terms is $\mathcal{N}$.

The master curve itself depends on the scaled data. For a given i, L_i, we estimate the master curve at x_{ij} by the two respective data from all the other system sizes which respectively enclose x_{ij}: for each $i \neq i$, let j' be such that $x_{i'j'} \leq x_{ij} \leq x_{i'(j'+1)}$, and select the points $(x_{i'j'}, y_{i'j'}, dy_{i'j'}), (x_{i'(j'+1)}, y_{i'(j'+1)}, dy_{i'(j'+1)})$. Do not select points for some i', if there is no such j'. If there is no such j' for all i', the master curve remains undefined at x_{ij}.

Given the selected points (x_l, y_l, dy_l), the local approximation of the master curve is the linear fit

$$y = mx + b$$

with weighted least squares [9]. The weights w_l are the reciprocal variances, $w_l := 1/dy_{ij}^2$. The estimates and (co)variances of the slope m and intercept b are

$$\hat{b} = \frac{1}{\Delta}(K_{xx}K_y - K_x K_{xy})$$
$$\hat{m} = \frac{1}{\Delta}(K K_{xy} - K_x K_y)$$

$$\hat{\sigma}_b^2 = \frac{K_{xx}}{\Delta}, \hat{\sigma}_m^2 = \frac{K}{\Delta}, \hat{\sigma}_{bm} = -\frac{K_x}{\Delta}$$

with $K_{nm} := \sum w_l x_l^n y_l^m$, $K := K_{00}$, $K_x := K_{10}$, $K_y := K_{01}$, $K_{xx} := K_{20}$, $K_{xy} := K_{11}$, $\Delta := K K_{xx} - K_x^2$.

Hence, the estimated value of the master curve at x_{ij} is

$$Y_{ij} = \hat{m} x_{ij} + \hat{b}$$

with error propagation

$$dY_{ij}^2 = \hat{\sigma}^2 x_{ij}^2 + 2\hat{\sigma}_{bm} x_{ij} + \hat{\sigma}_b^2.$$

References

[1] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, 1999).

[2] K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics* (Springer, Berlin, Heidelberg, 2010).

[3] J. Houdayer and A. Hartmann, Physical Review B **70**, 014418+ (2004).

[4] O. Melchert, "autoScale.py - a program for automatic finite-size scaling analyses: A user's guide," (2009), arXiv:0910.5403 .

[5] S. M. Bhattacharjee and F. Seno, Journal of Physics A: Mathematical and General **34**, 6375 (2001).

[6] S. Wenzel, E. Bittner, W. Janke, and A. M. J. Schakel, Nuclear Physics B **793**, 344 (2008).

[7] N. Kawashima and N. Ito, Journal of the Physical Society of Japan **62**, 435 (1993).

[8] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed. (McGraw-Hill, Boston, 2003).

[9] T. Strutz, *Data fitting and uncertainty : a practical introduction to weighted least squares and beyond* (Vieweg + Teubner, Wiesbaden, 2011).