

Finite-size scaling at fixed renormalization-group invariant

Francesco Parisen Toldin*

Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

(Received 7 January 2022; accepted 7 March 2022; published 24 March 2022)

Finite-size scaling at fixed renormalization-group invariant is a powerful and flexible technique to analyze Monte Carlo data at a critical point. It consists in fixing a given renormalization-group invariant quantity to a given value, thereby trading its statistical fluctuations with those of a parameter driving the transition. One remarkable feature is the observed significant improvement of statistical accuracy of various quantities, as compared to a standard analysis. We review the method, discussing in detail its implementation, the error analysis, and a previously introduced covariance-based optimization. Comprehensive benchmarks on the Ising model in two and three dimensions show large gains in the statistical accuracy, which are due to cross-correlations between observables. As an application, we compute an accurate estimate of the inverse critical temperature of the improved $O(2)$ ϕ^4 model on a three-dimensional cubic lattice.

DOI: [10.1103/PhysRevE.105.034137](https://doi.org/10.1103/PhysRevE.105.034137)**I. INTRODUCTION**

One of the most powerful techniques in numerical studies of systems with a large number of degrees of freedom is the (dynamic) Monte Carlo (MC) method [1]. It essentially consists in the realization of a Markov chain, which at equilibrium converges to a target distribution, typically the Gibbs measure. Due to its flexibility, MC methods are commonly used in numerical studies of critical phenomena [2]. Since truly singular behavior is found in the thermodynamic limit only, one needs extrapolation techniques in order to extract the critical behavior from numerical simulations which necessarily deal with a finite number of degrees of freedom.

For this purpose, one of the simplest approaches consists in extrapolating the infinite-volume limit from MC data at a finite size L . In the presence of a finite exponential correlation length ξ , and in the grand-canonical ensemble, observables generically exhibit an approach to their thermodynamic limit which is exponential in the ratio L/ξ ; i.e., their finite-size correction is proportional to $\exp(-L/\xi)$, for $\xi \ll L$ [3–5]. Nevertheless, finite-size corrections polynomial in $1/L$ are found for some specific observables, such as the second-moment correlation length [6,7], or in the absence of translational invariance; for instance, with open boundary conditions (BCs), subleading terms in the free energy give rise to corrections proportional to $1/L$. Interestingly, the statement of exponential approach to the infinite-volume limit does not hold in the canonical ensemble, where the constraint on the particle number induces finite-size corrections proportional to the inverse volume [8].

In the vicinity of a critical point, due to the divergence of the correlation length ξ , infinite-volume methods have a limited application. In this context, finite-size scaling (FSS) techniques play an important role [2,9]. FSS is formulated

by applying renormalization-group (RG) theory to a system of finite size, and close to a critical point. This leads to the prediction of a scaling form for the singular part of the free energy, and for derived quantities, such that they acquire a dependence on the ratio ξ/L . This scaling behavior can be then exploited to formulate specific FSS methods; we refer to Refs. [2,9] for a discussion of the various techniques. The great merit of FSS lies in the fact that, unlike infinite-volume methods mentioned above, it works in a parameter region where ξ is of the order of the linear size of the system L . For this reason, FSS is very often the method of choice for numerical studies of critical phenomena. In FSS a central role is played by the so-called RG-invariant observables. These are quantities which are scale-invariant at a critical point, and are commonly used to locate its position [2,9].

In the method of FSS at fixed RG-invariant, also known as FSS at fixed phenomenological coupling [10], one fixes the value of a chosen RG-invariant quantity R , thereby trading the statistical fluctuations of R with those of a parameter driving the transition. This method, introduced in Ref. [10], does not require additional computational time and has been used in several high-precision MC studies [11–20]. It shares some similarities with the method of phenomenological renormalization, where two system sizes are enforced to share a common value of $R = \xi/L$ [21,22]. One interesting feature of FSS at fixed RG-invariant is that, compared to a standard analysis, in some cases a significant reduction of error bars has been reported [11,13,16,23]. Considering that, for an equilibrated Markov chain, the accuracy of sampled observables is generically proportional to the inverse square root of the number of MC samples, even a (relatively) small improvement in the error bars may represent a significant gain in terms of computational time. This aspect is particularly important in the context of critical phenomena because of the occurrence of critical slowing down [1].

In Ref. [23] we have introduced an optimization of FSS at fixed RG-invariant which aims at minimizing the error bars of

*francesco.parisentoldin@physik.uni-wuerzburg.de

MC observables by determining the optimal RG-invariant to be fixed, as a linear combination of available RG-invariants. This is accomplished by an analysis of the statistical covariance between observables, and, as benchmarked in Ref. [23], leads to large gains in CPU time.

In this work we give a comprehensive overview of the method of FSS at fixed RG-invariant, benchmarking its performance in improving the statistical error bars. As a concrete application, we determine the location of the critical point in the improved $O(2)$ ϕ^4 lattice model, whose critical behavior belongs to the three-dimensional XY universality class.

This paper is organized as follows. In Sec. II we discuss the formulation of the method, paying particular attention to its practical implementation, and to the error analysis; we also describe in more detail the covariance optimization outlined in Ref. [23]. In Sec. III we report extensive benchmarks of the method on the Ising model in two and three dimensions. In Sec. IV we apply FSS at fixed RG-invariant to determine the critical inverse temperature of the ϕ^4 lattice model with two components in three dimensions. A summary of the present work is given in Sec. V. In the Appendix we give the proof of two formulas for the coefficients of an optimized linear combination of RG-invariants.

II. FSS AT FIXED RG-INVARIANT

A. Formulation of the method

A common approach in FSS methods [9] is the notion of a pseudocritical coupling. Generically, this is the (apparent) value of the coupling constant at the onset of criticality, as determined at a *finite* size L . Since no true critical point can appear in a finite size, the actual definition is not unique. For example, a pseudocritical coupling can be defined as the maximum of the susceptibility of the order parameter for a finite system. Another popular definition is provided by the crossing method: one defines a pseudocritical coupling by the intersection of the curves of an RG-invariant observable at size L and size αL , with fixed $\alpha > 1$ (see, e.g., Ref. [7]). By construction, pseudocritical couplings converge to the true critical-point constant for $L \rightarrow \infty$.

For concreteness, here and in the following we employ the language of classical critical phenomena at finite temperature, where the tuning parameter is the inverse temperature β ; the method discussed in the following can be readily applied also to quantum critical phenomena, where FSS holds and the tuning parameter is a coupling constant present in the Hamiltonian [24]. We define a pseudocritical inverse temperature $\beta_f(L)$ by fixing a given RG-invariant quantity R to a fixed value R_f :

$$R(\beta_f(L), L) = R_f. \quad (1)$$

According to RG theory and FSS, in the vicinity of a critical point located at $\beta = \beta_c$,

$$R(\beta, L) = f(tL^{1/\nu}) + L^{-\omega} g(tL^{1/\nu}), \quad (2)$$

where $1/\nu$ is the RG dimension of the relevant, non-symmetry-breaking, scaling field $t = a_0(\beta - \beta_c)/\beta_c$, and ω is the leading correction-to-scaling exponent. The scaling function $f(z)$ is universal, once the normalization a_0 on the argument z , the shape, and BCs are fixed [25–28]. We indicate

with $R^* \equiv f(0)$ the critical-point value of R . Inserting Eq. (2) in Eq. (1) and solving it for z one finds for $L \rightarrow \infty$

$$z = a_0 \left(\frac{\beta_f(L) - \beta_c}{\beta_c} \right) L^{1/\nu} \simeq f^{-1}(R_f) - L^{-\omega} \frac{g(f^{-1}(R_f))}{f'(f^{-1}(R_f))}. \quad (3)$$

From Eq. (3) it follows that the pseudocritical inverse temperature $\beta_f(L)$ converges to β_c as

$$\beta_f(L) - \beta_c \propto L^{-1/\nu}, \quad \text{generic } R_f, \quad (4)$$

$$\beta_f(L) - \beta_c \propto L^{-1/\nu - \omega}, \quad R_f = R^*; \quad (5)$$

i.e., the convergence of $\beta_f(L)$ is improved when we fix R to its critical-point value R^* . Once $\beta_f(L)$ is determined as a solution of Eq. (1), the various observables are analyzed at $\beta = \beta_f(L)$, so that they depend on L only. In particular, let us consider a divergent observable O that in the FSS limit satisfies a scaling ansatz of the form

$$O(\beta, L) = L^e [f_O(tL^{1/\nu}) + L^{-\omega} g_O(tL^{1/\nu})], \quad (6)$$

where e is a critical exponent [29]. Using Eq. (3) in Eq. (6), we obtain the behavior of O at fixed $R = R_f$,

$$O_f(L) \equiv O(\beta_f(L), L) = AL^e (1 + BL^{-\omega}), \quad (7)$$

where A and B are observable-dependent constants. We notice that Eq. (7) holds irrespectively of the choice of R_f . In other words, unlike Eq. (5), the choice of $R_f = R^*$ does not suppress corrections in O at fixed $R = R_f$. Equation (7) is the starting point to determine the critical exponent e by fitting the size dependence of O_f . In Sec. III we consider as observables the susceptibility χ and the derivatives of RG-invariant quantities $dR/d\beta$, as estimators for the critical exponents η and ν .

B. Implementation of the method

Typically, MC simulations are performed at a fixed value of $\beta = \beta_{\text{run}}$. The implementation of the method requires the extrapolation of the observables at a different value of β , so to be able to solve Eq. (1) and to compute Eq. (7). For this purpose, one can sample the derivatives of observables with respect to β , and compute a Taylor expansion

$$O(\beta) \simeq O(\beta_{\text{run}}) + \frac{\partial O}{\partial \beta} (\beta - \beta_{\text{run}}) + \frac{1}{2} \frac{\partial^2 O}{\partial \beta^2} (\beta - \beta_{\text{run}})^2, \quad (8)$$

where $O(\beta_{\text{run}})$ indicates an observable O as sampled via MC simulation at $\beta = \beta_{\text{run}}$. Generally, a second-order expansion as in Eq. (8) should be sufficient to reliably extrapolate O to values of β close to the simulation parameter β_{run} . Nevertheless, it is advisable to repeat the calculations truncating the expansion to the first and second term in Eq. (8), so as to check the stability of the extrapolation. Alternatively, one can use reweighting techniques to compute $O(\beta)$.

An important aspect of the method is the correct determination of the uncertainty of observables O_f at fixed RG-invariant R . For this purpose, one can use the jackknife method [30]. This consists in dividing the MC measures into N_{bin} bins of equal size, after having discarded the initial measures to ensure the equilibration of the Markov chain. The size of each

bin needs to be larger than the autocorrelation time of the various observables, such that the bins can be regarded as statistically independent from each other. Then, one defines a jackknife bin j as the set of MC measures of all bins, excluded the bin j . For each jackknife bin j , one can compute the various observables, in particular the chosen RG-invariant quantity R . This results in N_{bin} jackknife estimates, which we denote with $R^{(j)}$. Employing the Taylor expansion of Eq. (8) (or a reweighting method) in Eq. (1), one determines the j th jackknife estimate of $\beta_f^{(j)}$. For a generic observable O , this is used in Eq. (8) to compute O_f , the value of O at fixed RG-invariant. Repeating this procedure for all N_{bin} jackknife bins results in N_{bin} estimates of $O_f^{(j)}$. Furthermore, we also consider the value of O_f , as obtained using the entire set of (equilibrated) MC data, and denote it with $O_f^{(\text{all})}$. The jackknife estimators for the central value of O_f and its variance are [30]

$$O_f^{\text{est}} = N_{\text{bin}} O_f^{(\text{all})} - (N_{\text{bin}} - 1) \overline{O_f^{(j)}}, \quad (9)$$

$$(\sigma_{O_f}^2)^{\text{est}} = \frac{N_{\text{bin}} - 1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} \left(O_f^{(j)} - \overline{O_f^{(j)}} \right)^2, \quad (10)$$

where $\overline{O_f^{(j)}}$ is the average over the jackknife bins

$$\overline{O_f^{(j)}} \equiv \frac{1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} O_f^{(j)}. \quad (11)$$

Similarly, the jackknife method offers an estimator for the covariance between two observables A and B ,

$$\begin{aligned} \text{COV}(A, B)^{\text{est}} \\ = \frac{N_{\text{bin}} - 1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} (A^{(j)} - \overline{A^{(j)}})(B^{(j)} - \overline{B^{(j)}}), \end{aligned} \quad (12)$$

where, analogous to the discussion above, $A^{(j)}$, $B^{(j)}$ are the estimates of A , B on the jackknife bin j , and $\overline{A^{(j)}}$, $\overline{B^{(j)}}$ their average [see Eq. (11)].

Alternative methods for estimating error bars are the bootstrap [30] and the Γ -method [31].

We end this section with a few technical remarks. As observed above, the determination of the pseudocritical inverse temperature β_f and of the observables at fixed RG-invariant O_f requires an extrapolation of MC data sampled at a given value of β . Irrespective of the method employed [Taylor expansion as in Eq. (8) or a reweighting technique], such an extrapolation can only be reliably performed for values of β close enough to β_{run} . If in performing this analysis the value of β_f is significantly distant from β_{run} , one may need to redo the simulation at value of β closer to the determined β_f . As a practical recipe, one can start from a small lattice size L_0 , where repeating a simulation does not incur in a significant overhead, determining $\beta_f(L_0)$. This value can be then employed as β_{run} for the simulation of the next lattice size $L_1 > L_0$. Like before, if the value of $\beta_f(L_1)$ found differs significantly from β_{run} , the simulation may need to be repeated at a value of β_{run} closer to the estimated $\beta_f(L_1)$. Following this procedure, one can systematically increase the lattice size,

determining $\beta_f(L)$ and $O_f(L)$. Equations (4) and (5) ensure a quick convergence of $\beta_f(L)$, so that the overhead of repeated simulations is limited to the smaller lattice sizes only.

As discussed in Sec. II A, the fixed value R_f can, in principle, be arbitrary. In practice, however, it is preferable to choose R_f as close as possible to the critical-point value R^* . There are at least three reasons for this. First of all, as shown in Eq. (5), when $R = R^*$, there is an improved convergence of $\beta_f(L)$ to β_c . Second, even if we choose a value $R_f \neq R^*$, the prefactor in front of the right-hand side of Eq. (4) is proportional to $R_f - R^*$, such that a value of R_f significantly different from the critical value R^* gives rise to a (numerically) slower approach of $\beta_f(L)$ to β_c (although the asymptotic power-law exponent $1/\nu$ remains independent of R_f when $R_f \neq R^*$). Third, the analysis of Sec. II A considers only the leading correction-to-scaling term. For a value of R_f sufficiently far from R^* , further subleading terms in Eq. (2) may be relevant and slow the approach to the FSS limit.

In a given model, when an estimate of R^* is not available from previous results, a common approach is to use standard FSS technique to determine the value of R^* , for example by using crossing methods, or by fitting a Taylor expansion of the right-hand side of Eq. (2) [17–20]. For this purpose one does not generally need large lattice sizes. Once an estimate of R^* is determined, one can proceed to analyze the MC data collected so far at $R = R^*$, possibly supplementing them with additional data obtained from simulations at larger lattice sizes. In this way, one optimizes the use of CPU time, reserving most of it for simulations at a single value of β_{run} and large lattice sizes.

C. Error analysis

MC simulations are a realization of a Markov chain, whose equilibrium distribution is the target Boltzmann-Gibbs measure [1]. The outcome of the simulations, and thus the various statistical estimators discussed in the previous section, are actually stochastic variables, whose expectation value is equal to the quantity of interest. We indicate with $E[X]$ the expectation value of a stochastic variable X , and its fluctuations around the mean value with $\delta X \equiv X - E[X]$. As discussed in Sec. II B, Eq. (1) can be reliably solved only if the simulation parameter β_{run} is sufficiently close to $\beta_f(L)$. In this case, to the leading order in the Taylor expansion of an RG-invariant quantity, we have

$$\beta_f(L) \simeq \beta_{\text{run}} + \frac{R_f - R(\beta_{\text{run}}, L)}{\frac{\partial R}{\partial \beta}(\beta_{\text{run}}, L)}. \quad (13)$$

By expanding Eq. (13) around the expectation value, and keeping the lowest order in the fluctuations, we obtain the fluctuations of β_f ,

$$\delta \beta_f \simeq - \frac{\delta R}{E\left[\frac{\partial R}{\partial \beta}\right]} + \frac{R_f - E[R]}{E\left[\frac{\partial R}{\partial \beta}\right]^2} \delta \frac{\partial R}{\partial \beta}, \quad (14)$$

where to keep a light notation, observables on the right-hand side are meant to be computed at the parameter of the simulation β_{run} , and at a given lattice size L . The second term on the right-hand side of Eq. (14) represents a correction and, according to the discussion above, we can neglect it as long as

$R \simeq R_f$. The variance of $\beta_f(L)$ is then

$$\sigma_{\beta_f}^2 = \frac{\sigma_R^2}{E\left[\frac{\partial R}{\partial \beta}\right]^2}. \quad (15)$$

Given an observable O , its value at fixed $R = R_f$ can be computed by inserting Eq. (13) in the Taylor expansion of Eq. (8). To the lowest order in $\beta_{\text{run}} - \beta_f(L)$, we have

$$O_f(L) \simeq O(\beta_{\text{run}}, L) + \frac{\partial O}{\partial \beta} \frac{R_f - R(\beta_{\text{run}}, L)}{\frac{\partial R}{\partial \beta}(\beta_{\text{run}}, L)}. \quad (16)$$

Expanding around the expectation value, we have, to the lowest order in the fluctuations,

$$\begin{aligned} \delta O_f \simeq \delta O - E\left[\frac{\partial O}{\partial \beta}\right] \frac{\delta R}{E\left[\frac{\partial R}{\partial \beta}\right]} + \delta \frac{\partial O}{\partial \beta} \frac{R_f - E[R]}{E\left[\frac{\partial R}{\partial \beta}\right]} \\ - E\left[\frac{\partial O}{\partial \beta}\right] \frac{R_f - E[R]}{E\left[\frac{\partial R}{\partial \beta}\right]^2} \delta \frac{\partial R}{\partial \beta}. \end{aligned} \quad (17)$$

Similar to Eq. (14), the second and the third term on the right-hand side of Eq. (17) can be neglected in first approximation, leading to the variance of O_f :

$$\sigma_{O_f}^2 = \sigma_O^2 + \frac{E\left[\frac{\partial O}{\partial \beta}\right]^2}{E\left[\frac{\partial R}{\partial \beta}\right]^2} \sigma_R^2 - 2 \frac{E\left[\frac{\partial O}{\partial \beta}\right]}{E\left[\frac{\partial R}{\partial \beta}\right]} \text{COV}(O, R). \quad (18)$$

The variance of an observable O at fixed $R = R_f$ is thus a sum of the variance of the observable O , i.e., the one arising from a standard analysis at fixed Hamiltonian parameters, and two additional terms. The first term is due to the fluctuations of the RG-invariant quantity R and it is always positive. The second term, proportional to the covariance between O and R , is the term that is, potentially, responsible for the reduction of error bars. It is worth noticing that if there is little covariance between O and R , the final uncertainty of O at fixed $R = R_f$ is actually larger than the standard one. In fact, the large covariance found between observables and RG-invariant quantity is the origin of the drastic reduction of error bars reported in some cases [13,23].

D. Covariance optimization

In Ref. [23] we have proposed a method to optimize the FSS at fixed RG-invariant, by minimizing the error bar of the resulting observable O_f . For this purpose, one considers a set of RG-invariant observables $\{R_i\}$ and a linear combination of them,

$$R(\{\lambda_i\}) \equiv \sum_i \lambda_i R_i. \quad (19)$$

We search in the space of linear combinations of $\{R_i\}$ the optimal coefficients λ_i that minimize the variance of an observable O at fixed $R(\{\lambda_i\})$. The solution to this problem is, up to an inessential normalization of $\{\lambda_i\}$ [23],

$$\lambda_i = - \frac{\mathbf{R}^T \mathbf{M}^{-1} \mathbf{N} - E[\partial O / \partial \beta]}{\mathbf{R}^T \mathbf{M}^{-1} \mathbf{R}'} (\mathbf{M}^{-1} \mathbf{R}')_i + (\mathbf{M}^{-1} \mathbf{N})_i, \quad (20)$$

where the matrix \mathbf{M} and the vectors \mathbf{N} and \mathbf{R}' are defined as

$$\begin{aligned} \mathbf{M}_{ij} &\equiv \text{COV}(R_i, R_j), \\ \mathbf{N}_i &\equiv \text{COV}(O, R_i), \\ \mathbf{R}'_i &\equiv E\left[\frac{\partial R_i}{\partial \beta}\right]. \end{aligned} \quad (21)$$

Analogously, one can consider the problem of finding the optimal linear combinations of $\{R_i\}$ that minimizes the variance of β_f . In this case, the solution is

$$\lambda_i = \frac{(\mathbf{M}^{-1} \mathbf{R}')_i}{\mathbf{R}^T \mathbf{M}^{-1} \mathbf{R}'}, \quad (22)$$

with \mathbf{M} and \mathbf{R}' as in Eq. (21). In the Appendix we prove Eq. (20), whose proof was only sketched in Ref. [23], as well as Eq. (22). The covariances in the definitions of Eq. (21) can be estimated within the jackknife method using Eq. (12). Covariances and estimated error bars are in general related to the transition matrix defining the Markov chain [1]. To illustrate this aspect, in Sec. III we test the method using different sampling algorithms.

In implementing the covariance-optimized FSS at fixed RG-invariant, it is important to note that the chosen RG-invariant observable R must be the same for all lattice sizes L . On the other hand, the optimal choice of $\{\lambda_i\}$ obtained from Eq. (20) does in general depend on L . Thus, a possible strategy is to choose the values of $\{\lambda_i\}$ that minimize the error bar on O for the largest available lattice size, where typically the most part of the CPU time is consumed, and then use the resulting RG-invariant in the FSS analysis of all lattice sizes. On the other hand, it is possible (and useful) to optimize $\{\lambda_i\}$ differently for each observable O considered, as long as one does not combine different observables in the scaling analysis. This is the strategy that we used in Ref. [23] to test the covariance optimization.

III. BENCHMARKS

A. Model and observables

To benchmark the method, we have simulated the standard Ising model, whose Hamiltonian on a d -dimensional lattice is

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \quad \sigma_i = \pm 1, \quad (23)$$

where the sum is over the nearest-neighbors sites. Without loss of generality, we set $J = 1$. We apply periodic BCs on a finite lattice with linear extension L . To test the influence of the sampling algorithm on the performance of the method, we have simulated the model using either the standard Metropolis, or the Wolff single-cluster updates [32]. Both are ergodic algorithms for the Ising model [1].

Here we consider four RG-invariant observables: ξ/L , U_4 , U_6 , and Z_a/Z_p , defined as follows. ξ is the second-moment correlation length at a finite size, defined as

$$\xi = \frac{1}{2 \sin(\hat{p}/2)} \sqrt{\frac{C(0)}{C(\hat{p})}} - 1, \quad (24)$$

where $C(p)$ is the Fourier transform of the two-point function of the order parameter σ_i , and $\hat{p} = 2\pi/L$ is the minimum

nonzero momentum on a finite lattice with size L ; see Appendix A of Ref. [7] for a discussion on the definition of ξ in a finite lattice. The generalized Binder ratios U_4 and U_6 are defined by

$$U_{2j} \equiv \frac{\langle M^{2j} \rangle}{\langle M^2 \rangle^j}, \quad (25)$$

where $M \equiv \sum_i \sigma_i$ is the total magnetization of the model, and the brackets $\langle \ \rangle$ indicate the thermal average. Finally, Z_a/Z_p is the ratio of the partition function of the model with antiperiodic BCs on one direction and periodic BCs on the remaining ones, over the partition function with full periodic BCs. This quantity can be efficiently sampled with the boundary-flip algorithm of Ref. [33,34].

As for the observables O which we analyze at fixed RG-invariant, we consider here the susceptibility χ

$$\chi \equiv \frac{1}{L^d} \sum_{ij} \sigma_i \sigma_j, \quad (26)$$

which in the FSS limit is $\propto L^{2-\eta}$, thus serving as an estimator of the critical exponent η . We also compute the first derivative of the RG-invariant quantities defined above with respect to β : $d(\xi/L)/d\beta$, $dU_4/d\beta$, $dU_6/d\beta$, and $d(Z_a/Z_p)/d\beta$. These are $\propto L^{1/\nu}$ in the FSS limit and are commonly used as estimators of the critical exponent ν [9].

B. Results

We have simulated the Ising model in $d = 2$, at the critical point $\beta_{\text{run}} = 0.4406867935$ [35], and for lattice sizes $L = 8$ –128. The analysis has been carried out using the fixed-point values of the RG-invariants, which thanks to the available exact solutions are known to a high numerical accuracy. We use $(\xi/L)^* = 0.9050488292(4)$, $U_4^* = 1.167923(5)$, $U_6^* = 1.455649(7)$ [36], and $(Z_a/Z_p)^* \simeq 0.37288488$ [37]. For each observable, we compute its value fixing one of the RG-invariants. Since in the MC method the uncertainty of an observable is $\propto 1/\sqrt{N_{\text{steps}}}$, with N_{steps} the number of MC steps in the sampling, the effective gain in CPU time can be defined as

$$\text{CPU gain} = \left(\frac{\sigma_{\text{std}}}{\sigma_{\text{fixed-RG}}} \right)^2, \quad (27)$$

where σ_{std} , $\sigma_{\text{fixed-RG}}$ are the error bar estimates for a standard analysis and for the analysis at fixed RG-invariant. For comparison we also show the CPU gains obtained using the covariance-optimized linear combination of RG-invariants; for observables χ , $d(\xi/L)/d\beta$, $dU_4/d\beta$, $dU_6/d\beta$ these are taken from Ref. [23]. As discussed in Sec. IID, for each observable, the coefficients of the linear combination are fixed by the minimization of the resulting error bar for the largest lattice size.

In Table I we report our benchmarks obtained by sampling the model with the standard Metropolis algorithm. The CPU gains appear to depend strongly on the observable, and for each observable, there is considerable dependence on the chosen RG-invariant R to be fixed. In particular, for χ , $dU_4/d\beta$, and $dU_6/d\beta$ we observe that at least one of the four possible choices R gives a significant reduction of the error bars and,

TABLE I. Benchmark results for the 2D Ising model, sampled with a Metropolis dynamics. We report the effective gain in CPU time [Eq. (27)] in the analysis of various observables at fixed RG-invariant R , as compared to the standard analysis at fixed MC parameters. For each lattice size L , we compare the gains obtained by choosing $R = \xi/L, U_4, U_6, Z_a/Z_p$, and by considering the covariance-optimized linear combination of these four RG-invariants. The results for the covariance-optimized FSS of χ , $d(\xi/L)/d\beta$, $dU_4/d\beta$, and $dU_6/d\beta$ are taken from Ref. [23].

χ					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	5.7	2.4	2.8	3.8	11.5
16	3.5	3.3	1.9	2.6	18.4
32	3.2	4.7	1.7	2.3	27.2
64	2.4	2.7	1.5	2.0	13.3
128	2.7	4.8	1.5	2.0	22.1
$d(\xi/L)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.1	1.2	0.7	0.8	6.0
16	0.8	1.1	0.5	0.6	4.7
32	0.7	1.1	0.5	0.5	4.5
64	0.8	1.0	0.6	0.6	2.6
128	0.6	1.0	0.4	0.5	2.9
$dU_4/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	3.0	1.6	6.7	5.1	35.4
16	3.0	1.5	6.2	4.8	43.6
32	3.5	1.7	7.0	5.4	54.9
64	3.2	1.5	5.8	4.7	22.9
128	3.7	1.6	6.8	5.4	39.8
$dU_6/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	3.7	1.7	11.4	7.9	47.6
16	3.8	1.6	10.3	7.2	54.2
32	4.5	1.8	11.8	8.4	68.2
64	4.1	1.6	9.3	7.1	27.5
128	5.0	1.8	11.5	8.4	48.9
$d(Z_a/Z_p)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	0.7	1.5	0.7	0.7	1.2
16	0.8	1.4	0.8	0.9	1.4
32	1.2	1.6	1.4	1.4	2.1
64	1.1	1.4	1.2	1.3	1.7
128	1.5	1.5	1.8	1.8	2.2

consequently, an effective gain in CPU time. Moreover, the use of the covariance-optimized RG-invariant greatly improve further the CPU gains, in particular for $dU_4/d\beta$ and $dU_6/d\beta$ for which we observe a rather large CPU gain. For the observables $d(\xi/L)/d\beta$, $d(Z_a/Z_p)/d\beta$ the CPU gain obtained choosing one of the four RG-invariants is instead limited, and in some cases it is less than 1, meaning that there is actually a (small) increase of the error bars, compared to the standard analysis. Nevertheless, also for these observables the choice of the covariance-optimized RG-invariant provides a non-negligible gain in CPU time. Remarkably,

TABLE II. Same as Table I for a Wolff dynamics.

χ					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	3.1	0.7	1.8	2.3	5.9
16	3.1	0.8	1.7	2.1	5.4
32	2.9	0.9	1.5	1.8	5.9
64	3.4	1.2	1.9	2.3	6.5
128	4.3	1.3	1.8	2.2	7.7
$d(\xi/L)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.4	0.8	0.8	0.8	2.5
16	1.2	0.9	0.7	0.8	2.2
32	1.0	0.9	0.7	0.7	2.4
64	1.1	0.9	0.7	0.8	2.1
128	1.0	0.9	0.7	0.7	2.0
$dU_4/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.5	1.0	3.4	2.7	6.3
16	1.5	1.1	3.1	2.6	8.4
32	1.5	1.1	2.7	2.4	7.6
64	1.5	1.2	2.6	2.4	5.7
128	1.5	1.3	2.4	2.2	5.6
$dU_6/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.6	1.0	5.1	3.9	8.5
16	1.7	1.2	4.4	3.5	9.6
32	1.6	1.1	3.6	3.1	8.7
64	1.6	1.3	3.4	3.0	6.7
128	1.7	1.3	3.1	2.8	6.5
$d(Z_a/Z_p)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	0.8	1.5	0.8	0.8	1.2
16	0.9	1.3	1.0	1.0	1.4
32	1.0	1.1	1.1	1.1	1.3
64	1.0	1.2	1.2	1.2	1.4
128	1.1	1.2	1.3	1.3	1.6

TABLE III. Same as Table I for $d = 3$.

χ					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	9.2	4.5	1.8	2.2	21.2
16	9.4	6.9	1.7	2.1	27.9
32	8.2	7.7	1.3	1.6	26.9
64	8.6	6.3	1.4	1.7	22.2
128	9.1	8.3	1.4	1.7	25.5
$d(\xi/L)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.2	1.1	0.6	0.6	4.2
16	1.1	1.1	0.5	0.6	4.9
32	1.0	1.0	0.4	0.5	6.2
64	0.8	0.9	0.4	0.4	4.2
128	1.0	1.1	0.4	0.5	6.1
$dU_4/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.3	1.2	1.5	1.4	24.2
16	1.3	1.2	1.5	1.4	9.1
32	1.2	1.2	1.5	1.4	23.0
64	1.3	1.2	1.7	1.6	16.1
128	1.2	1.2	1.5	1.5	19.7
$dU_6/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.8	1.5	3.1	2.5	27.6
16	1.7	1.5	3.0	2.5	24.2
32	1.7	1.5	3.0	2.6	25.8
64	1.9	1.6	3.5	2.9	25.2
128	1.6	1.4	3.1	2.7	23.0
$d(Z_a/Z_p)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.8	1.3	1.0	1.1	2.2
16	1.4	1.2	0.7	0.7	2.2
32	1.1	1.1	0.5	0.6	2.4
64	1.0	1.0	0.5	0.6	1.9
128	0.9	1.0	0.5	0.5	2.0

the CPU gains are stable on increasing the lattice size. This especially underscores the effectiveness of the FSS method at fixed RG-invariant, for a reduction of error bars at small lattice sizes only would severely limit the usefulness of (any) improvement method. In Table II we report our results for the Wolff dynamics. On a qualitative level, the observed pattern in the CPU gains is very similar to the case of a Metropolis dynamics. Overall, the reduction of error bars is however smaller than in the former case. Simulations in $d = 3$ have been carried out at $\beta_{\text{run}} = 0.2216544$, close to the accurate estimate of Ref. [38] of the critical inverse temperature $\beta_c = 0.221654626(5)$. For the FSS analysis at fixed RG-invariant, we have employed the critical-point values $(\xi/L)^* = 0.6431(1)$, $U_4^* = 1.6036(1)$, $U_6^* = 3.1053(5)$, $(Z_a/Z_p)^* = 0.5425(1)$ [17]. The results of the benchmarks are reported in Table III for a Metropolis sampling and in Table IV for a Wolff dynamics [39]. A comparison with the results for $d = 2$ (Tables I and II) reveals a striking similarity. For

each observable, there is a choice of RG-invariant to be fixed, among ξ/L , Z_a/Z_p , U_4 , U_6 , where the CPU gain is maximized. Such an optimal choice appears, for a given observable, to be the same across $d = 2, 3$, and for the two dynamics tested here. The covariance-optimized linear combination of RG-invariants provides a further important gain in error bars, in line with the results of $d = 2$. The fact that these observations hold irrespective of dimension and sampling algorithm hints at a possible underlying physical mechanism explaining the significant statistical covariance between observables; this, indeed, according to Eq. (18) is ultimately responsible for the reduction of error bars.

Another quantity of interest is the pseudocritical inverse temperature $\beta_f(L)$, which converges to β_c for $L \rightarrow \infty$ [see Eqs. (4) and (5)]. In Table V we report the statistical error bars on $\beta_f(L)$ as obtained by a FSS at fixed RG-invariant, for the same MC data used in the previous benchmarks. We also test the idea of a covariance-optimized linear combination of

TABLE IV. Same as Table II for $d = 3$.

χ					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	4.2	1.0	1.0	1.1	6.3
16	4.4	1.2	1.0	1.1	7.0
32	5.0	1.4	1.2	1.3	7.8
64	5.7	1.7	1.1	1.1	9.5
128	6.5	2.1	1.2	1.3	10.1
$d(\xi/L)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.2	0.9	0.6	0.6	2.7
16	1.1	0.8	0.5	0.5	3.1
32	1.2	0.8	0.6	0.6	3.1
64	1.1	0.8	0.5	0.5	3.2
128	1.1	0.9	0.5	0.6	3.1
$dU_4/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.1	1.1	1.2	1.1	7.0
16	1.1	1.1	1.2	1.2	5.4
32	1.1	1.1	1.2	1.1	4.2
64	1.1	1.1	1.3	1.2	4.2
128	1.0	1.1	1.2	1.2	4.6
$dU_6/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.3	1.2	1.9	1.6	7.7
16	1.3	1.2	2.0	1.8	6.7
32	1.2	1.2	1.8	1.6	5.5
64	1.2	1.2	1.9	1.7	6.3
128	1.2	1.2	1.9	1.7	5.8
$d(Z_a/Z_p)/d\beta$					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.2	1.0	0.8	0.8	1.4
16	1.1	1.0	0.8	0.8	1.4
32	1.1	0.9	0.7	0.8	1.4
64	1.0	0.9	0.7	0.7	1.4
128	1.0	0.9	0.7	0.7	1.4

$\{R_i\}$ that minimizes the error bar in $\beta_f(L)$ [Eq. (22)]; analogous to the benchmarks in Tables I–IV, the coefficients are fixed by optimizing $\beta_f(L)$ for the largest available lattice size $L = 128$. In analyzing the results in Table V, it is not particularly meaningful to compare numbers between different lattice size: indeed, the resulting precision in $\beta_f(L)$ depends on the size of the MC sample. Inspecting Table V, we observe that generically the precision in $\beta_f(L)$ fixing ξ/L or Z_a/Z_p is only slightly better than obtained by choosing $R = U_4$ or $R = U_6$. The covariance-optimized choice of R appears to give an appreciable reduction of error bar only for the case of the 2D Ising model sampled with a Metropolis dynamics; in the other cases the improvement is limited.

IV. IMPROVED THREE-DIMENSIONAL XY model

We simulate the $O(2)$ ϕ^4 model on a three-dimensional lattice of linear size L , applying periodic BCs along all direc-

tions. The reduced Hamiltonian \mathcal{H} , such that the Gibbs weight is $\exp(-\mathcal{H})$, is

$$\mathcal{H} = -\beta \sum_{(i,j)} \vec{\phi}_i \cdot \vec{\phi}_j + \sum_i [\vec{\phi}_i^2 + \lambda(\vec{\phi}_i^2 - 1)^2], \quad (28)$$

where $\vec{\phi}_x$ is a 2-component real field on the lattice site x , the first sum extends over the nearest-neighbor pairs, and the last term is a sum over all lattice sites. MC simulations have been carried out combining Metropolis, overrelaxation, and Wolff single-cluster updates [32]; technical details of the sampling algorithm can be found in Ref. [20]. In the limit $\lambda \rightarrow \infty$ the Hamiltonian (28) reduces to the standard XY model. Starting from $\lambda = \infty$, the (β, λ) phase diagram exhibits a line of continuous phase transitions in the classical XY universality class. For $\lambda = 2.15(5)$ [12] the model is *improved*; i.e., the leading irrelevant scaling field with dimension $y_i = -0.789(4)$ [18] is suppressed. Improved models are instrumental in high-precision numerical studies of critical phenomena [2]. For $\lambda = 2.1$ the model is critical at $\beta_c = 0.5091503(3)$ and for $\lambda = 2.2$, it is critical at $\beta_c = 0.5083355(3)$ [12]. Here, we have used FSS at fixed RG to compute the critical value of β_c for the estimated central value of $\lambda = 2.15$ at which the model is improved. To this end, we have simulated the model (28) for lattice sizes $L = 32$ – 384 with high statistics, and analyzed the data at fixed ξ/L and Z_a/Z_p , employing the fixed-point values $(\xi/L)^* = 0.59238(7)$ and $(Z_a/Z_p)^* = 0.32037(6)$ [18]. According to Eq. (5), we expect an improved approach of $\beta_f(L)$ to the inverse critical temperature. Furthermore, being the model improved, the leading irrelevant scaling field, associated with the cubic anisotropy of the lattice, has a negative scaling dimension $\omega = \omega_{nr} \approx 2$ [18]. Therefore, we expect a rather fast convergence of $\beta_f(L) \rightarrow \beta_c$ on increasing L . This is confirmed in Fig. 1, where we show $\beta_f(L)$ as obtained by fixing ξ/L or Z_a/Z_p . For $L \geq 96$, data points of $\beta_f(L)$ are mutually compatible within the error bars. Accordingly, we can safely average the data for the two largest lattice sizes $L = 256, 384$, so as to obtain a statistically improved value of β_c . Doing so for data at fixed ξ/L we obtain $\beta_c = 0.50874988(6)$, while using the analysis at fixed Z_a/Z_p we get $\beta_c = 0.50874988(5)$. From these two values, we take as a conservative estimate

$$\beta_c = 0.50874988(6). \quad (29)$$

Including also the data at $L = 192$ in the average does not change the final estimate: in fact, as shown in Fig. 1, the data points for $L = 192$ lie well within one error bar of the estimated β_c . This further strengthens the robustness of the final result given in Eq. (29).

V. SUMMARY

FSS at fixed RG-invariant is a convenient and powerful method to analyze MC data of a critical model. Its implementation, discussed in detail in Sec. II, essentially consists in a specific analysis of MC data, which does not require additional computational costs. The two main advantages of this method over more standard FSS techniques are in the observed statistical improvement of critical observables, and in its convenience for computing critical parameters.

TABLE V. Benchmark of the statistical precision of $\beta_f(L)$ as obtained by a FSS analysis at fixed RG-invariant R , with $R = \xi/L, U_4, U_6, Z_a/Z_p$, as well as choosing R as the covariance-optimized linear combination of these four RG-invariants.

Ising 2D Metropolis dynamics					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	4.2×10^{-4}	2.5×10^{-4}	4.6×10^{-4}	4.4×10^{-4}	2.5×10^{-4}
16	4.0×10^{-4}	2.1×10^{-4}	4.4×10^{-4}	4.1×10^{-4}	1.8×10^{-4}
32	2.4×10^{-4}	1.3×10^{-4}	2.7×10^{-4}	2.6×10^{-4}	1.0×10^{-4}
64	8.6×10^{-5}	4.5×10^{-5}	9.4×10^{-5}	8.8×10^{-5}	4.3×10^{-5}
128	5.8×10^{-5}	3.0×10^{-5}	6.4×10^{-5}	6.1×10^{-5}	2.3×10^{-5}
Ising 2D Wolff dynamics					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.9×10^{-4}	1.8×10^{-4}	2.0×10^{-4}	1.9×10^{-4}	1.5×10^{-4}
16	1.3×10^{-4}	1.3×10^{-4}	1.5×10^{-4}	1.4×10^{-4}	1.1×10^{-4}
32	7.1×10^{-5}	6.5×10^{-5}	8.1×10^{-5}	7.8×10^{-5}	5.7×10^{-5}
64	3.7×10^{-5}	3.7×10^{-5}	4.3×10^{-5}	4.1×10^{-5}	3.2×10^{-5}
128	2.0×10^{-5}	1.9×10^{-5}	2.4×10^{-5}	2.3×10^{-5}	1.7×10^{-5}
Ising 3D Metropolis dynamics					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	1.2×10^{-4}	8.1×10^{-5}	1.5×10^{-4}	1.4×10^{-4}	7.8×10^{-5}
16	4.6×10^{-5}	3.2×10^{-5}	5.8×10^{-5}	5.6×10^{-5}	2.7×10^{-5}
32	1.6×10^{-5}	1.2×10^{-5}	2.1×10^{-5}	2.0×10^{-5}	1.0×10^{-5}
64	4.5×10^{-6}	3.4×10^{-6}	5.9×10^{-6}	5.7×10^{-6}	3.0×10^{-6}
128	1.7×10^{-6}	1.3×10^{-6}	2.2×10^{-6}	2.1×10^{-6}	1.1×10^{-6}
Ising 3D Wolff dynamics					
L	Fixed ξ/L	Fixed Z_a/Z_p	Fixed U_4	Fixed U_6	Cov. opt.
8	4.6×10^{-5}	4.6×10^{-5}	6.5×10^{-5}	6.3×10^{-5}	4.1×10^{-5}
16	1.6×10^{-5}	1.7×10^{-5}	2.2×10^{-5}	2.2×10^{-5}	1.5×10^{-5}
32	5.6×10^{-6}	5.8×10^{-6}	7.3×10^{-6}	7.1×10^{-6}	5.2×10^{-6}
64	2.0×10^{-6}	2.1×10^{-6}	2.6×10^{-6}	2.6×10^{-6}	1.9×10^{-6}
128	7.0×10^{-7}	7.2×10^{-7}	9.3×10^{-7}	9.2×10^{-7}	6.6×10^{-7}

Extensive benchmarks carried out in this work show that the method offers in many cases a significant reduction of statistical error bars, as compared to a standard analysis at fixed Hamiltonian parameters. The improvement of statistical precision is due to the cross-correlations of MC observables. An optimization of the method put forward in Ref. [23], based on the covariance analysis, allows one to obtain rather

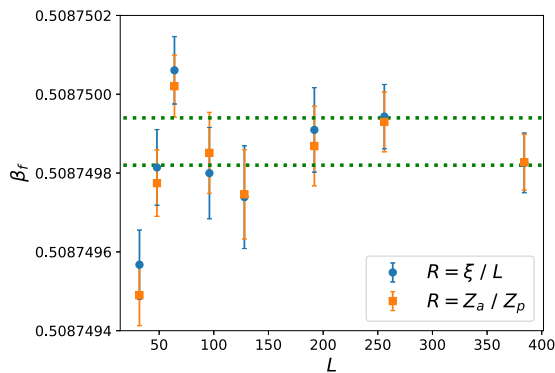


FIG. 1. Pseudocritical inverse temperature $\beta_f(L)$ for the improved XY model of Eq. (1), as obtained by FSS at fixed ξ/L and fixed Z_a/Z_p . Dotted lines indicate the final estimate of β_c [Eq. (29)].

large gains in the error bars or, equivalently, in the computational time. In this context, covariance analysis has been used to reduce the statistical error of an observable by adding a so-called control variate, i.e., a quantity whose mean value vanishes [40], and to determine the optimal weighted average of critical exponents estimators [41,42]. As a tool to improve statistical precision, FSS at fixed RG-invariant largely outperform the aforementioned methods, especially in the case of the covariance-optimized version [23]. It is worth mentioning that, in principle, one could combine the present method with those of Refs. [40–42].

FSS at fixed RG-invariant provides also a convenient way to compute the critical parameters of a model. As shown in Eqs. (4), (5), and (7), within this framework pseudocritical couplings and critical observables acquire a simple size dependence, which is the starting point for fitting critical exponents and critical couplings. As a concrete application, we have employed the method to accurately determine the critical inverse temperature [Eq. (29)] of an improved model in the three-dimensional XY universality class.

ACKNOWLEDGMENTS

F.P.T. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project No.

414456783. The author thanks U. Wolff for useful communications. The author gratefully acknowledges the Gauss Centre for Supercomputing e.V. [43] for funding this project by providing computing time on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre [44], where the simulations of the ϕ^4 model have been carried out.

APPENDIX: PROOF OF THE FORMULAS FOR THE OPTIMAL COEFFICIENTS

In this Appendix we prove Eq. (20) which was reported in Ref. [23], and Eq. (22). The space of real stochastic variables with null expectation value is a real vector space equipped with a scalar product $\langle \cdot, \cdot \rangle$, defined as the covariance

$$\langle \delta X, \delta Y \rangle \equiv \text{COV}(\delta X, \delta Y) = E[\delta X \delta Y]. \quad (\text{A1})$$

In particular, the variance of a stochastic variable is the length square computed with the scalar product of Eq. (A1). Thus, minimizing the variance is equivalent to the problem of minimizing the length of a vector.

Let us first consider the minimization of σ_{β_f} . Inserting Eq. (19) into Eq. (14), and neglecting the second term on the right-hand side (see the discussion in Sec. IIC), we obtain the fluctuations of β_f as a function of the parameters $\{\lambda_i\}$,

$$\delta\beta_f = -\frac{\sum_i \lambda_i \delta R_i}{\sum_i \lambda_i \mathbf{R}'_i}, \quad (\text{A2})$$

with the definition of Eq. (21). Clearly, the right-hand side of Eq. (A2) is invariant under a global rescaling $\lambda_i \rightarrow c\lambda_i$: this simply reflects the trivial equivalence of fixing $R = R_f$ or $cR = cR_f$. To minimize the variance, it is useful to fix the normalization of λ_i as

$$\sum_i \lambda_i \mathbf{R}'_i = 1, \quad (\text{A3})$$

such that we need to minimize the length of the vector

$$\delta\beta_f = -\sum_i \lambda_i \delta R_i \quad (\text{A4})$$

subject to the constraint of Eq. (A3). An immediate solution to this problem is obtained by exploiting a geometrical interpretation. In the vector space spanned by the N fluctuations δR_i , the coefficients $-\lambda_i$ of Eq. (A4) are the coordinates in the (nonorthogonal) base $\{\delta R_i\}$. Accordingly, the constraint of Eq. (A3) is an affine subspace of codimension 1, which contains the vectors $\delta v \equiv \lambda_i \delta R_i$ satisfying

$$\sum_{ij} \lambda_i \mathbf{M}_{ij} (\mathbf{M}^{-1} \mathbf{R}')_j = \langle \delta v, \delta w \rangle = 1, \quad (\text{A5})$$

where the vector δw ,

$$\delta w = \sum_i (\mathbf{M}^{-1} \mathbf{R}')_i \delta R_i, \quad (\text{A6})$$

is orthogonal to the affine subspace of the constraint. An illustration of the problem is shown in Fig. 2. The minimum length of $\delta\beta_f$ is readily found when $\delta\beta_f$ is orthogonal to the affine subspace, i.e., when

$$\delta\beta_f = \mu \delta w. \quad (\text{A7})$$

The coefficient μ is easily obtained by inserting Eq. (A7) into Eq. (A5) or Eq. (A3), obtaining the optimal coefficients $\{\lambda_i\}$ given in Eq. (22).

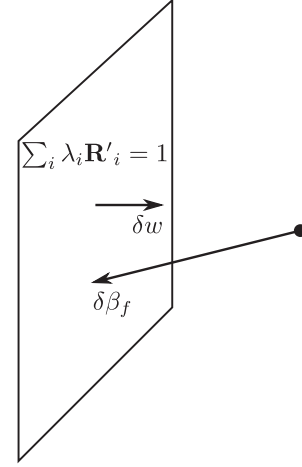


FIG. 2. Geometrical illustration of the problem of minimizing the fluctuations of β_f . The hyperplane defined by the constraint of Eq. (A3) is an affine subspace of codimension 1, orthogonal to the vector δw . The fluctuations $\delta\beta_f$ are a vector from the origin to the hyperplane. Its length is minimized when $\delta\beta_f$ is orthogonal to the hyperplane.

To prove Eq. (20) one proceeds analogously. Using Eq. (19) in Eq. (17), and keeping only the leading term, we obtain the fluctuations of O_f ,

$$\delta O_f = \delta O - E \left[\frac{\partial O}{\partial \beta} \right] \frac{\sum_i \lambda_i \delta R_i}{\sum_i \lambda_i \mathbf{R}'_i}. \quad (\text{A8})$$

This time we fix the normalization of $\{\lambda_i\}$ as

$$\sum_i \lambda_i \mathbf{R}'_i = E \left[\frac{\partial O}{\partial \beta} \right], \quad (\text{A9})$$

such that we need to minimize the length of the vector δO_f ,

$$\delta O_f = \delta O - \sum_i \lambda_i \delta R_i, \quad (\text{A10})$$

subject to the constraint of Eq. (A9), which represents an affine subspace, parallel to the one of Eq. (A3). Once again, the minimum length of δO_f is found when the right-hand side of Eq. (A10) is parallel to δw defined in Eq. (A6),

$$\delta O_f = \delta O - \sum_i \lambda_i \delta R_i = \mu \delta w. \quad (\text{A11})$$

By taking the scalar product of both sides of Eq. (A11) with δR_i we obtain an equation for $\{\lambda_i\}$,

$$\mathbf{N}_i - \sum_j \mathbf{M}_{ij} \lambda_j = \mu \mathbf{R}'_i, \quad (\text{A12})$$

where we have used Eq. (A6) and the definitions of Eq. (21). Equation (A12) can be inverted to obtain

$$\lambda_i = [\mathbf{M}^{-1} (\mathbf{N} - \mu \mathbf{R}')]_i. \quad (\text{A13})$$

Inserting Eq. (A13) into Eq. (A9) we find the value of the constant μ ,

$$\mu = \frac{\mathbf{R}^T \mathbf{M}^{-1} \mathbf{N} - E[\partial O / \partial \beta]}{\mathbf{R}^T \mathbf{M}^{-1} \mathbf{R}'}. \quad (\text{A14})$$

Finally, using this result in Eq. (A13) we obtain Eq. (20).

- [1] A. Sokal, Monte Carlo methods in statistical mechanics: Foundations and new algorithms, in *Functional Integration: Basics and Applications*, edited by C. DeWitt-Morette, P. Cartier, and A. Folacci (Springer US, Boston, MA, 1997), pp. 131–192.
- [2] A. Pelissetto and E. Vicari, Critical phenomena and renormalization-group theory, *Phys. Rep.* **368**, 549 (2002).
- [3] M. Lüscher, Volume dependence of the energy spectrum in massive quantum field theories: I. Stable particle states, *Commun. Math. Phys.* **104**, 177 (1986).
- [4] H. Neuberger, Finite size effects in massive field theory, *Phys. Lett. B* **233**, 183 (1989).
- [5] G. Münster, The size of finite size effects in lattice gauge theories, *Nucl. Phys. B* **249**, 659 (1985).
- [6] S. Caracciolo and A. Pelissetto, Corrections to finite-size scaling in the lattice N -vector model for $N = \infty$, *Phys. Rev. D* **58**, 105007 (1998).
- [7] F. Parisen Toldin, M. Hohenadler, F. F. Assaad, and I. F. Herbut, Fermionic quantum criticality in honeycomb and π -flux Hubbard models: Finite-size scaling of renormalization-group-invariant observables from quantum Monte Carlo, *Phys. Rev. B* **91**, 165108 (2015).
- [8] Z. Wang, F. F. Assaad, and F. Parisen Toldin, Finite-size effects in canonical and grand-canonical quantum Monte Carlo simulations for fermions, *Phys. Rev. E* **96**, 042131 (2017).
- [9] V. Privman, Finite-size scaling theory, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, Singapore, 1990), p. 1.
- [10] M. Hasenbusch, A Monte Carlo study of leading order scaling corrections of ϕ^4 theory on a three-dimensional lattice, *J. Phys. A: Math. Gen.* **32**, 4851 (1999).
- [11] M. Hasenbusch, A. Pelissetto, and E. Vicari, Instability of $O(5)$ multicritical behavior in $SO(5)$ theory of high- T_c superconductors, *Phys. Rev. B* **72**, 014532 (2005).
- [12] M. Campostrini, M. Hasenbusch, A. Pelissetto, and E. Vicari, Theoretical estimates of the critical exponents of the superfluid transition in ^4He by lattice methods, *Phys. Rev. B* **74**, 144506 (2006).
- [13] M. Hasenbusch, F. Parisen Toldin, A. Pelissetto, and E. Vicari, The universality class of 3D site-diluted and bond-diluted Ising systems, *J. Stat. Mech.: Theory Exp.* (2007) P02016.
- [14] M. Hasenbusch, F. Parisen Toldin, A. Pelissetto, and E. Vicari, Critical behavior of the three-dimensional $\pm J$ Ising model at the paramagnetic-ferromagnetic transition line, *Phys. Rev. B* **76**, 094402 (2007).
- [15] M. Hasenbusch, F. Parisen Toldin, A. Pelissetto, and E. Vicari, Universal dependence on disorder of two-dimensional randomly diluted and random-bond $\pm J$ Ising models, *Phys. Rev. E* **78**, 011110 (2008).
- [16] U. Wolff, Precision check on the triviality of the ϕ^4 theory by a new simulation method, *Phys. Rev. D* **79**, 105002 (2009).
- [17] M. Hasenbusch, Finite size scaling study of lattice models in the three-dimensional Ising universality class, *Phys. Rev. B* **82**, 174433 (2010).
- [18] M. Hasenbusch, Monte Carlo study of an improved clock model in three dimensions, *Phys. Rev. B* **100**, 224517 (2019).
- [19] M. Hasenbusch, Monte Carlo study of a generalized icosahedral model on the simple cubic lattice, *Phys. Rev. B* **102**, 024406 (2020).
- [20] F. Parisen Toldin, Boundary Critical Behavior of the Three-Dimensional Heisenberg Universality Class, *Phys. Rev. Lett.* **126**, 135701 (2021).
- [21] M. P. Nightingale, Scaling theory and finite systems, *Phys. A (Amsterdam)* **83**, 561 (1976).
- [22] M. P. Nightingale, Non-universality for Ising-like spin systems, *Phys. Lett. A* **59**, 486 (1977).
- [23] F. Parisen Toldin, Improvement of Monte Carlo estimates with covariance-optimized finite-size scaling at fixed phenomenological coupling, *Phys. Rev. E* **84**, 025703(R) (2011).
- [24] M. Campostrini, A. Pelissetto, and E. Vicari, Finite-size scaling at quantum transitions, *Phys. Rev. B* **89**, 094516 (2014).
- [25] K. Kaneda, Y. Okabe, and M. Kikuchi, Shape effects of finite-size scaling functions for anisotropic three-dimensional Ising models, *J. Phys. A: Math. Gen.* **32**, 7263 (1999).
- [26] F. Parisen Toldin and S. Dietrich, Critical Casimir forces and adsorption profiles in the presence of a chemically structured substrate, *J. Stat. Mech.: Theory Exp.* (2010) P11003.
- [27] A. Hucht, D. Grüneberg, and F. M. Schmidt, Aspect-ratio dependence of thermodynamic Casimir forces, *Phys. Rev. E* **83**, 051101 (2011).
- [28] A. Malakis, N. G. Fytas, and G. Gülpinar, Critical Binder cumulant and universality: Fortuin-Kasteleyn clusters and order-parameter fluctuations, *Phys. Rev. E* **89**, 042103 (2014).
- [29] In Eq. (6) we have for simplicity assumed that the leading correction-to-scaling in O decays with the same exponent ω appearing in Eq. (2). This is indeed the case when the dominant corrections are given by the leading irrelevant scaling field, which generically enters in the FSS limit of all observables. Nevertheless, for some specific observables the leading scaling correction has another origin; for example it could be due to the nonsingular part of the free energy. This is the case of the susceptibility in the Ising model on the square lattice [2], or of the zero-momentum correlations of the order parameter in the ground state of the Hubbard model on the honeycomb lattice, at the onset of the quantum phase transition between a semimetallic and an antiferromagnetic phase [7]. In such cases, the leading correction-to-scaling exponent ω entering in Eq. (7) is the smallest between those appearing in Eq. (2) and Eq. (6).
- [30] A. P. Young, *Everything You Wanted to Know About Data Analysis and Fitting but Were Afraid to Ask*, Springer Briefs in Physics (Springer International Publishing, Heidelberg, 2015).
- [31] U. Wolff and ALPHA Collaboration, Monte Carlo errors with less errors, *Comput. Phys. Commun.* **156**, 143 (2004) Erratum to “Monte Carlo errors with less errors”, **176**, 383 (2007).
- [32] U. Wolff, Collective Monte Carlo Updating for Spin Systems, *Phys. Rev. Lett.* **62**, 361 (1989).
- [33] M. Hasenbusch, Monte Carlo simulation with fluctuating boundary conditions, *Phys. A (Amsterdam)* **197**, 423 (1993).
- [34] M. Hasenbusch, Monte Carlo studies of the three-dimensional Ising model in equilibrium, *Int. J. Mod. Phys. C* **12**, 911 (2001).
- [35] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Academic Press, 1989).
- [36] J. Salas and A. D. Sokal, Universal amplitude ratios in the critical two-dimensional Ising model on a torus, *J. Stat. Phys.* **98**, 551 (2000).
- [37] M. Caselle and M. Hasenbusch, Deconfinement transition and dimensional cross-over in the 3D gauge Ising model, *Nucl. Phys. B* **470**, 435 (1996).

- [38] A. M. Ferrenberg, J. Xu, and D. P. Landau, Pushing the limits of Monte Carlo simulations for the three-dimensional Ising model, *Phys. Rev. E* **97**, 043301 (2018).
- [39] Due to a trivial mistake, the CPU gains reported in Table II of Ref. [23] for the covariance-optimized FSS of the three-dimensional Ising model at $L = 8$, sampled with Metropolis dynamics, are underestimated by a factor 2. Also, due to a typo in the analysis, in the case of the three-dimensional Ising model and Wolff dynamics there is a small numerical error in the values of $dU_4/d\beta$ at fixed covariance-optimized RG-invariant, as reported in Table II of Ref. [23]. In both cases, Table IV here contains the corrected values.
- [40] L. A. Fernandez and V. Martin-Mayor, Mean-value identities as an opportunity for Monte Carlo error reduction, *Phys. Rev. E* **79**, 051109 (2009).
- [41] M. Weigel and W. Janke, Cross Correlations in Scaling Analyses of Phase Transitions, *Phys. Rev. Lett.* **102**, 100601 (2009).
- [42] M. Weigel and W. Janke, Error estimation and reduction with cross correlations, *Phys. Rev. E* **81**, 066701 (2010).
- [43] See <http://www.gauss-centre.eu>.
- [44] See <http://www.lrz.de>.