A much better method uses the results of the previous Section. Indeed we know that the variance of the sample mean $\bar{f}$ is $C_{ff}(0)(2\tau_{int,f})$. $C_{ff}(0)$ is simply the static variance and is estimated by

$$C_{ff}(0) = \frac{1}{n} \sum_{i=1}^{n} \left(f_i - \bar{f}\right)^2 \tag{5.3}$$

Notice that this quantity depends uniquely on the model and not on the algorithm used to simulate it. What instead depends on the algorithm is $\tau_{int,f}$. To evaluate it we must first estimate $C_{ff}(t)$. The natural estimator is

$$\hat{C}_{ff}(t) = \frac{1}{n - |t|} \sum_{i=1}^{n-|t|} (f_i - \bar{f})(f_{i+|t|} - \bar{f}) \tag{5.4}$$

This is a biased estimator of $C(t)$ the bias being of order $1/n$. The natural estimator of $\rho_{ff}(t) = C_{ff}(t)/C_{ff}(0)$ is thus

$$\hat{\rho}_{ff}(t) = \frac{\hat{C}_{ff}(t)}{\hat{C}_{ff}(0)} \tag{5.5}$$

Then the natural estimator of τ_{int} would seem to be

$$\hat{\tau}_{int,f} = \frac{1}{2} \sum_{t=(1-n)}^{n-1} \hat{\rho}_{ff}(t) \tag{5.6}$$

but this is wrong. Indeed this estimator has a variance that does not go to zero as $n \to \infty$. Roughly speaking this is because the sample autocorrelation $\hat{\rho}_{ff}(t)$ for $|t| >> \tau_{int,f}$ contains much noise but little signal (see the following exercise).

Exercise 1: Let f_i be independent Gaussian variables of mean μ/n and variance σ^2/n. Compute the mean and the variance of $\sum_{i=1}^{n} f_i$.

The solution is to cut off the sum defining

$$\hat{\tau}_{int,f} = \frac{1}{2} \sum_{t=-M}^{M} \hat{\rho}_{ff}(t) \tag{5.7}$$

Of course this cut off introduces a bias and in general this formula will underestimate the correct value. However we can choose M in such a way that the bias is small. This can be achieved using the authomatic windowing procedure by Madras and Sokal: choose M self-consistently as the smallest integer such that $M \geq c\hat{\tau}_{int,f}(M)$. If $\rho_{ff}(t)$ were a pure exponential one could take for instance $c = 4$, in this case making an error of the $e^{-4} = 2\%$. However in many cases $\rho_{ff}(t)$ has a slower preasymptotic decay and thus in these cases c must be larger, for instance $c = 4 - 6$ in order to have an error of the same magnitude. In general the determination of c requires the study of the behaviour of the autocorrelation function $C_{ff}(t)$. Indeed c must not be too small, otherwise the bias would be too large, but neither too large otherwise we

would include some noise. What small and large mean, depends on the number of iterations and in general, increasing the number of iterations one must at the same time increase the value of c in order to keep the systematic error smaller than the statistical one. Let us finally quote the variance of $\hat{\tau}_{int,f}$:

$$\text{Var}(\hat{\tau}_{int,f}) = \frac{2(2M+1)}{n}\tau_{int,f}^2 \tag{5.8}$$

valid in the regime $\tau_{int,f} << n$.

As a final remark, notice that this procedure works well only if $n >> \tau_{int,f}$. Empirically one usually finds that the results are reliable if $n \gtrsim 1000\tau_{int,f}$.

From what we have said it is clear that the performance of the algorithm is completely characterized by $\tau_{int,f}$ measured in CPU-time units. The higher $\tau_{int,f}$ the less efficient is the algorithm. Now what is the behaviour of $\tau_{int,f}$? Generically, away from phase transitions (i.e. in the regions where the correlation length is small) $\tau_{int,f}$ remains reasonably small. However near a critical point the autocorrelation time diverges as

$$\tau \sim \xi^z(\beta) \tag{5.9}$$

where $\xi(\beta)$ is the correlation length of the infinite volume system at temperature $1/\beta$ and z is a dynamical critical exponent. This phenomenon is called dynamic critical slowing down and it is the most severe limitation to Monte Carlo studies. We will see in the next Section that for local Monte Carlo one usually have $z \gtrsim 2$ (overrelaxation is an exception with $z \gtrsim 1$) while using non-local methods (cluster algorithms, Fourier acceleration, multigrid ...) one can achieve sensible improvements (in many cases $z \approx 0$).

Now, how can we determine z? First of all, as we have already remarked, z depends on the type of autocorrelation time we are considering and thus we must distinguish between z_{exp} and $z_{int,f}$ and in many algorithms different quantities f have different exponents (for instance in Wolff's algorithms for RP^n σ-models the exponent of the susceptibility is ~ 1 while the exponent of the energy is ~ 0). Secondly, notice that (5.9) is true only when $L \to \infty$. In practice experience with two-dimensional models shows that corrections to scaling are usually very strong. For instance, using Wolff's algorithm for the $O(4)$ σ-model one finds for $L = 32$, $\beta = 1.70$ $\xi = 3.54$ and $\tau_{int,\chi} = 4.53(16)$ while for $L = 32$, $\beta = 2.20$ $\xi = 11.59$ and $\tau_{int,\chi} = 10.61(70)$. Using (5.9) and these two values we would get $z_{int,\chi} \sim 0.8$ but this is wrong. Indeed for this model one has $z_{int,\chi} \lesssim 0.1$. The wrong result is due to the fact that we have neglected the finite-size corrections. The correct way of performing the analysis uses a finite-size scaling Ansatz of the form

$$\tau_{int,f}(\beta, L) \sim \xi(\beta, L)^{z_{int,f}} g_f(\xi(\beta, L)/L) \tag{5.10}$$

where g_f is an unknown scaling function with $g_f(0)$ supposed to be finite and non-zero. To determine $z_{int,f}$ one plots $\tau_{int,f}(\beta, L)/\xi(\beta, L)^{z_{int,f}}$ as a function of $\xi(\beta, L)/L$ fixing $z_{int,f}$ so that all the points lie on a unique curve within error bars. In order to have a reliable estimate many different values of L must be used and sufficiently accurate estimates of $\tau_{int,f}$ are required.

Let us now pass to a second subject: so far we have discussed how to compute from the Monte Carlo data the mean values of the observables we are interested in and their error bars. Now let us discuss how to use them in some practical problem.

In many cases we want to compute invariant ratios, i.e. quantities of the form

$$R = \frac{\langle A \rangle^p}{\langle B \rangle^q} \tag{5.11}$$

where A and B are two different observables and p and q some exponents. Of course we estimate this quantity from the sample means $\bar{A}$ and $\bar{B}$:

$$R_{est} = \frac{\bar{A}^p}{\bar{B}^q} \tag{5.12}$$

But then, what is the error bar on R_{est}? The main problem in its estimation is that $\bar{A}$ and $\bar{B}$ come from the same run; consequently they are correlated and this must be kept into account in setting the error bars. Let us compute the variance of R_{est}. By definition

$$\mathrm{Var}\,(R_{est}) = \langle \frac{\bar{A}^{2p}}{\bar{B}^{2q}} \rangle - \langle R_{est} \rangle^2 \tag{5.13}$$

Then introduce

$$\Delta A = \frac{\bar{A} - \langle A \rangle}{\langle A \rangle} \tag{5.14}$$

$$\Delta B = \frac{\bar{B} - \langle B \rangle}{\langle B \rangle} \tag{5.15}$$

Expanding in ΔA and ΔB (valid in the large sample limit) we get

$$\mathrm{Var}\,(R_{est}) = \frac{\langle A \rangle^{2p}}{\langle B \rangle^{2q}} \left(p^2 \langle \Delta A^2 \rangle + q^2 \langle \Delta B^2 \rangle - 2pq \langle \Delta A \Delta B \rangle \right) + o(1/n) \tag{5.16}$$

Notice that $\langle \Delta A^2 \rangle = \sigma_A^2 / \langle A \rangle^2$ and $\langle \Delta B^2 \rangle = \sigma_B^2 / \langle B \rangle^2$ while the last term keeps into account the correlation between the estimates of A and B. Using the fact that (Schwartz inequality)

$$|\langle \Delta A \Delta B \rangle| \leq \langle \Delta A^2 \rangle^{1/2} \langle \Delta B^2 \rangle^{1/2} \tag{5.17}$$

we obtain an upper bound in terms of the variances of A and B alone

$$\mathrm{Var}\,(R_{est}) \leq \frac{\langle A \rangle^{2p}}{\langle B \rangle^{2q}} \left(p \frac{\sigma_A}{\langle A \rangle} + q \frac{\sigma_B}{\langle B \rangle} \right)^2 \tag{5.18}$$

This upper bound is usually very far from being sharp since A and B are strongly correlated (in some cases (5.18) is ten times larger than the correct formula (5.16)). For this reason it is always convenient to use (5.16). However in this case we must compute the covariance $\langle \Delta A \Delta B \rangle$. This can be avoided using a little trick, i.e. rewriting (5.16) as

$$\mathrm{Var}\,(R_{est}) = \frac{\langle A \rangle^{2p}}{\langle B \rangle^{2q}} \,\mathrm{Var}\,(p \Delta A - q \Delta B) \tag{5.19}$$

and this last variance can be computed applying the time-series methods we have presented before to the time series

$$p\,\frac{A_i - \bar{A}}{\bar{A}} \;-\; q\,\frac{B_i - \bar{B}}{\bar{B}} \tag{5.20}$$

In this way one usually gets smaller (but correct!) error bars. We want to emphasize that these results are not connected with dynamic Monte Carlo, in the sense that the important thing here is the static correlation of A and B (usually $p\Delta A - q\Delta B$ has an autocorrelation time which is of the order of the largest of $\tau_{int,A}$ and $\tau_{int,B}$) and thus it is important to keep it into account also in the case of independent sampling.

These formulas can easily be extended to generic functions of $\langle A \rangle$ and $\langle B \rangle$ and to functions of more than two observables.

Exercise 2: In lattice gauge theories one usually defines the Creutz ratio

$$R(I,J) \;=\; \frac{W(I,J)W(I-1,J-1)}{W(I,J-1)W(I-1,J)} \tag{5.21}$$

where $W(I,J)$ denotes the expectation of a rectangular Wilson loop of lattice dimensions I and J. Compute the variance of $R(I,J)$ in the large sample limit expressing the result in terms of variances of suitably defined observables for which the usual time-series analysis can be applied.

Let us now present two applications of the previous results. The first one is concerned with the so-called hystogram method introduced by Falcioni et al. and Ferrenberg and Swendsen. Suppose you have made a simulation of the system at a given value of β. Can this run be used to get some information on the mean values of the various observables we have measured at a different temperature β_0? The answer is yes, and it is based on the following observation: if $\mathcal{O}$ is any observable we have

$$\langle \mathcal{O} \rangle_{\beta_0} \;=\; \frac{1}{Z(\beta_0)} \sum_{\{\sigma\}} \mathcal{O}(\sigma) \exp(-\beta_0 H(\sigma)) \tag{5.22}$$

$$=\; \frac{\sum_{\{\sigma\}} \mathcal{O}(\sigma) \exp((\beta - \beta_0)H(\sigma)) \exp(-\beta H(\sigma))}{\sum_{\{\sigma\}} \exp((\beta - \beta_0)H(\sigma)) \exp(-\beta H(\sigma))} \tag{5.23}$$

$$=\; \frac{\langle \exp((\beta - \beta_0)H)\mathcal{O} \rangle_\beta}{\langle \exp((\beta - \beta_0)H) \rangle_\beta} \tag{5.24}$$

where $\langle \ldots \rangle_\beta$ is the estimate at inverse temperature β. Thus an estimate of $\langle \mathcal{O} \rangle_{\beta_0}$ is obtained as a ratio of the sample means of $\exp((\beta - \beta_0)H)\mathcal{O}$ and $\exp((\beta - \beta_0)H)$ and the error bar is computed using (5.16).

One can also compute $\langle \mathcal{O} \rangle_{\beta_0}$ using runs at different, but nearby, values β_i. In this case the estimates are independent and thus it is very simple to write down the final estimate and its error bar. If $\mathcal{O}_i$ is the estimate of $\langle \mathcal{O} \rangle$ obtained using the runs at temperature $1/\beta_i$ and σ_i^2 its variance we get the final estimate

$$\mathcal{O}_{est} \;=\; \sum_i \frac{\mathcal{O}_i}{\sigma_i^2} \left(\sum_i \frac{1}{\sigma_i^2} \right)^{-1} \tag{5.25}$$