Worksheet 4: Error Analysis and Langevin Thermostat

April Cooper, Patrick Kreissl und Sebastian Weber

January 9, 2013
University of Stuttgart
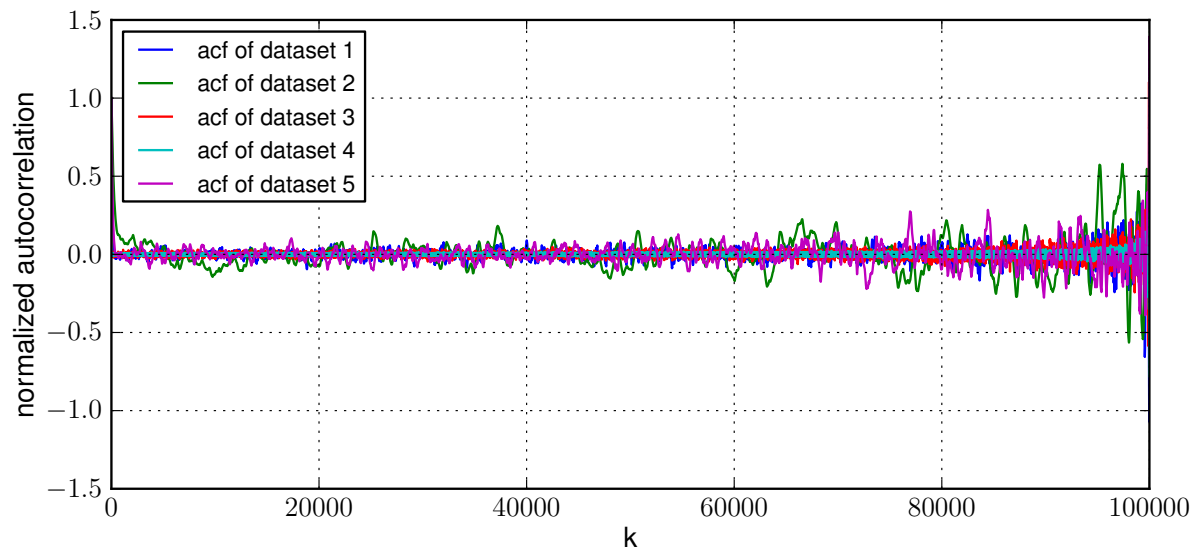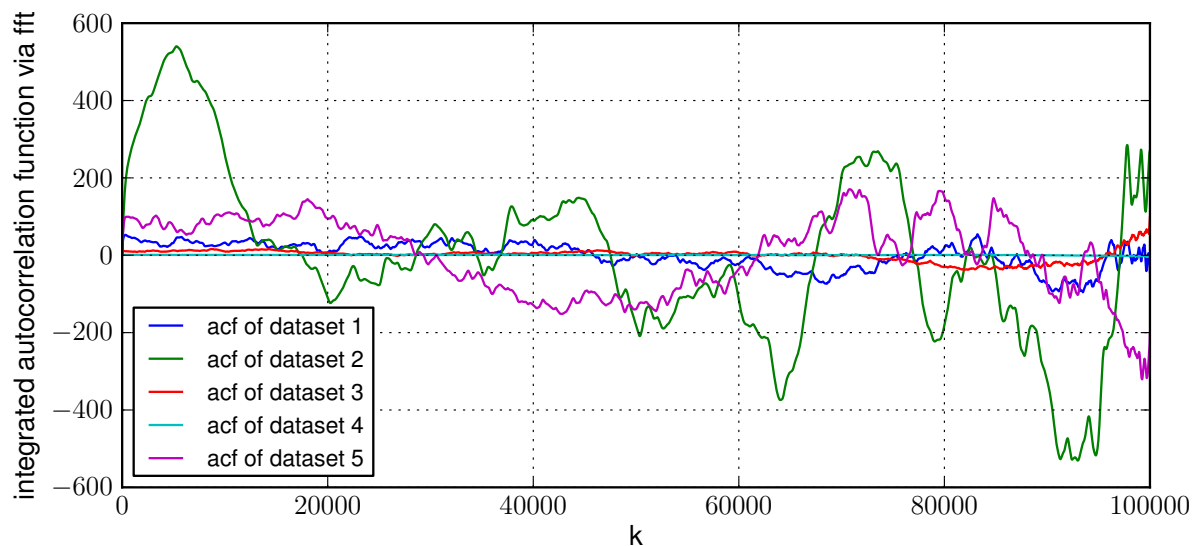
Contents

1 Error Analysis

1.1 Autocorrelation Analysis

In this section we had to implement a Python function that computes the autocorrelaton function of a given time series and use this function on the given datasets. The resulting plot can be seen below:


Furthermore we were supposed to plot the integrated autocorrelation function which can be seen below:

As you can see in the plots, it is not useful to integrate the function over the whole interval because the lager the autocorellations shift parameter k becomes the lareger the fluctuation of the autocorrelation function becomes. This is because for larger ks due to the autocorrelations definition less data is used to compute the function. Therefore the statistical fluctuation of a single datapoint has more influence on the computed autocorrelation function. This larger fluctuation influences also manifest in the integrated autocorrelation function. The integrated function should converge against the autocorrelation time. So for a visual estimate of the autocorrelation time the only relevant part of the interval is the

---

one in which the autocorrelation funcion approximately acts like an exponetial function and where the influence of the intrisic error of the function due to the increased influence of statistical fluctuations is still relatively small. In the following plots we zoomed in to the relevant interval:
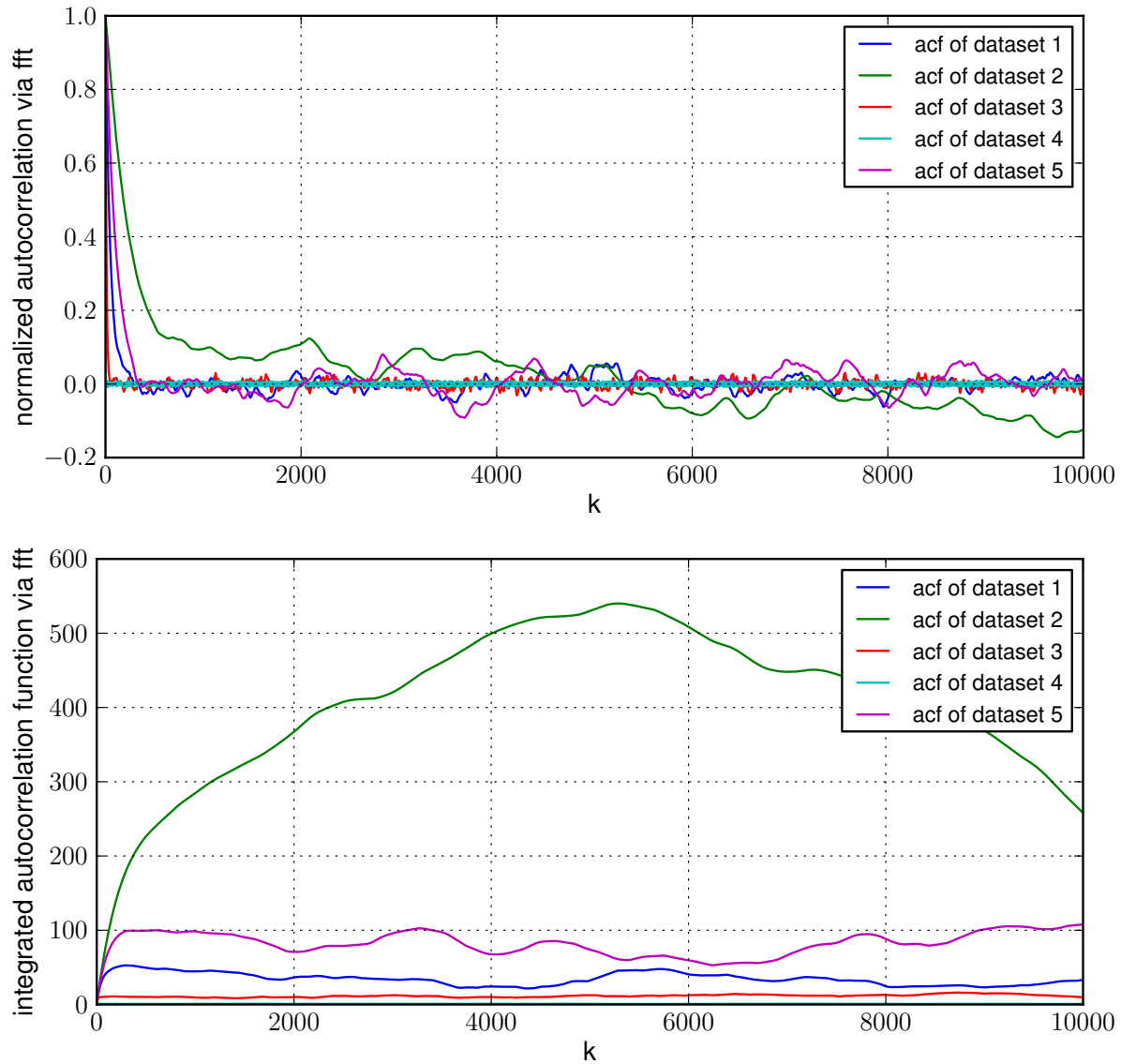
The autocorrelation times of the datasets can be visually guessed as follows:

dataset	guessed autocorrelation time
1	50
2	450
3	10
4	1
5	90

The next task was to implement a Python function that performs automatic error analysis via autocorrelation analysis of a given time series of an observable. This function should compute an estimate of the outocorrelation time via the running autocorrelation time estimator, where the integration is cut of when $k_{max} \geq 6\hat{\tau}_{\mathcal{O},\text{int}}(k_{max})$. The function was supposed to return the

mean value of the observable $\bar{\mathcal{O}}$, the error of the mean value $\epsilon_{\bar{\mathcal{O}}}$, the estimated autocorrelation time $\hat{\tau}_{\mathcal{O},\text{int}}$, its error $\epsilon_{\hat{\tau}_{\mathcal{O},\text{int}}}$ and the effective statistics N_{eff}. The results shown in the folowing table:

dataset	$\bar{\mathcal{O}}$	$\epsilon_{\bar{\mathcal{O}}}$	$\hat{\tau}_{\mathcal{O},\text{int}}$	$\epsilon_{\hat{\tau}_{\mathcal{O},\text{int}}}$	N_{eff}
1	$8.2468 \cdot 10^{-15}$	0.0031	52.134	5.8381	956
2	$3.9756 \cdot 10^{-14}$	0.0032	403.139	125.4145	123
3	$4.0010 \cdot 10^{-15}$	0.0031	9.707	0.4736	5042
4	$-1.4214 \cdot 10^{-13}$	0.0032	0.5039	0.0068	66666
5	$-3.1724 \cdot 10^{-14}$	0.0032	99.4101	15.3684	502

As you can see the visually estimated autocorrelation times are very similar to the computed ones. The dataset for witch it was most difficult to guess the autocorrelation time was dataset 2. This manifests in the computed error of the autocorrelation time of that dataset, too. It is nearly by factor 10 larger than the error of dataset 5, and even by factor 100 for the datasets 3 and 4.

1.2 Binning Analysis

We were supposed to implement a Python function that computes the block variance σ_B^2 for a given block size k. From that block variance the estimated autocorrelation time $\tau_{\mathcal{O},\text{int}}$ and error of the mean value $\epsilon_{\bar{\mathcal{O}}}$ can be computed. Our task was to plot theese for $1 < k < 2000$. The resulting plots are shown below:
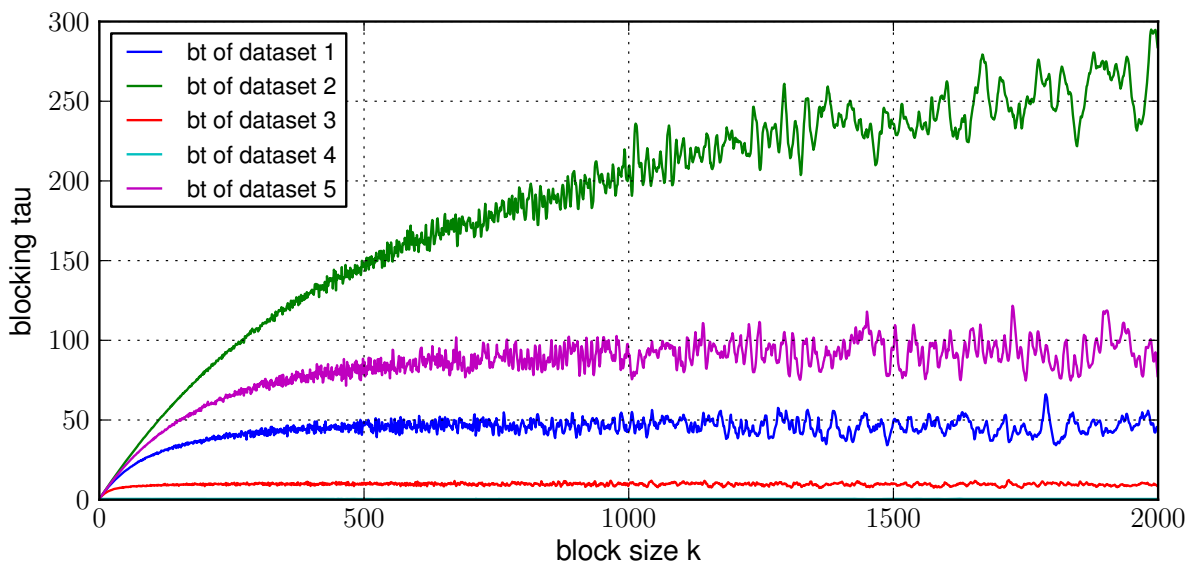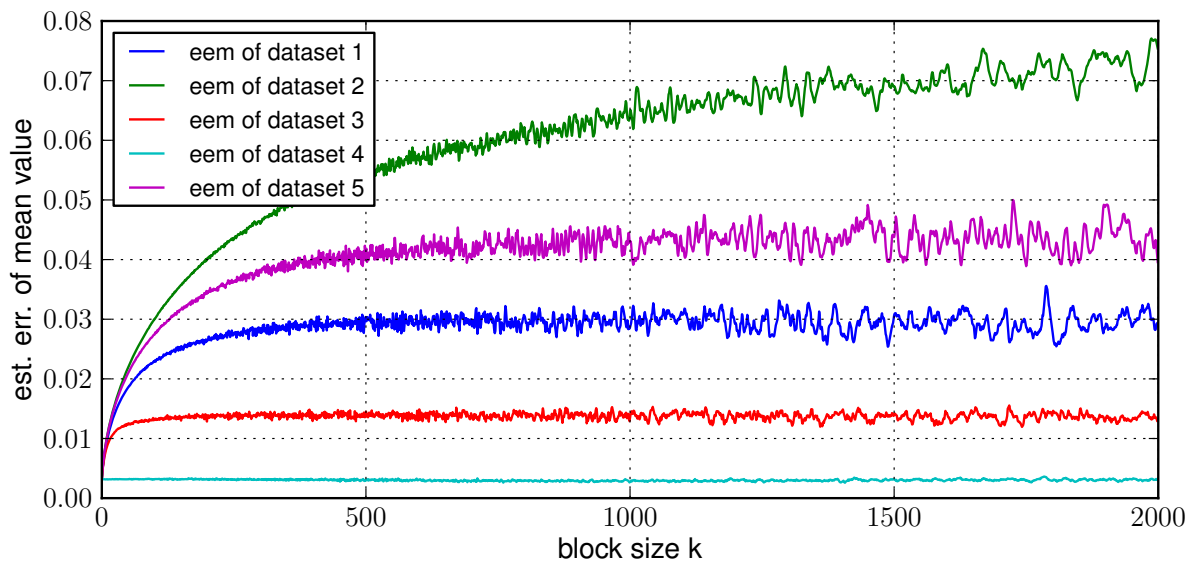
From the plots the autocorrelation time and the error of the mean value can be guessed as follows:

dataset	$\tau_{\mathcal{O},\text{int}}$	$\epsilon_{\bar{\mathcal{O}}}$
1	49	0,030
2	-	-
3	10	0,014
4	1	0,004
5	90	0.042

For both, autocorrelation time and the error of the mean value, for dataset 2 at the end of the interval ($k = 2000$) the functions do still not converge, while the fluctuation of the data becomes already relatively large. Therefore for dataset 2 no values could be estimated.

For the others the autocorrelation time matches the values obtained in the previous tasks. For the error of the mean value only the error of dataset 1 matches the previous estimated value. The error of the other is smaller than previously obtained.

1.3 Jackknife Analysis

In this section we had to implement a Python funcion that computes the Jackknife error $\epsilon_{\bar{\mathcal{O}}}$ for a given time series and block size k. The resulting plot is shown below:
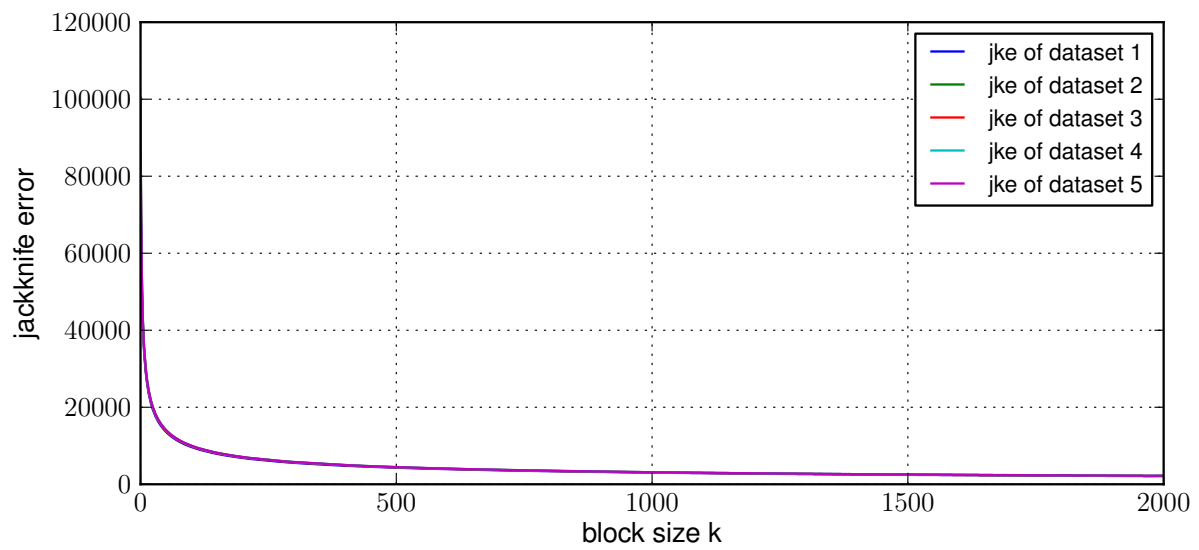
The computed errors would range from 2126 to 2203 at $k = 2000$. That seems hardly plausible and does not match the values obtained in the other tasks at all.