STA2700 - Graphical Models / Assignment 4

Luis Alvaro Correia - No. 1006508566

November 25th 2020

1 Question 1

Consider a 5×5 homogeneous Ising model with periodic boundary conditions, in which $N = 25$ particles/variables $X_1, X_2, \ldots, X_{25}$ are arranged on the sites of a $2D$ lattice. As usual, $\mathcal{X} = \{+1, -1\}$ and only *adjacent* particles interact with each other.

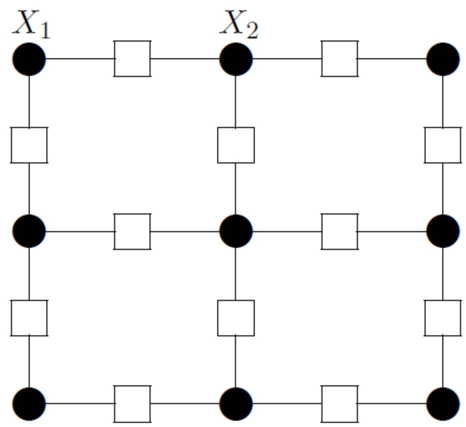
Let x_i represent a possible realization of X_i, and let $\boldsymbol{x}$ stand for $(x_1, x_2, \ldots, x_N$. The Hamiltonian (energy) of a configuration $\boldsymbol{x}$ is given by:

$$\mathcal{H}(\boldsymbol{x}) = -J \sum_{(i,j) adjacent} x_i y_j, \tag{1}$$

and the probability of a configuration $\boldsymbol{x}$ is given by:

$$p_B(\boldsymbol{x}) = \frac{e^{-\beta \mathcal{H}(\boldsymbol{x})}}{Z} = \frac{f(\boldsymbol{x})}{Z} \tag{2}$$

where Z is the partition function. For $N = 9$, the factor graph is illustrated below (excluding the boundary conditions).



(a) Compute the exact value of Z for different values of βJ via direct enumeration to fill out the following table.

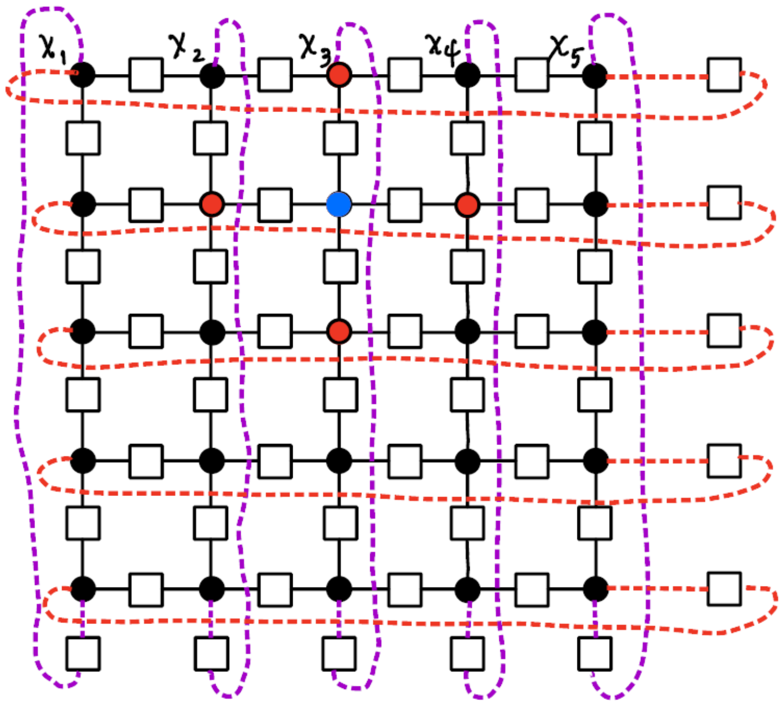
βJ	0.01	0.06	0.15	0.25	0.75
$\frac{\ln(Z)}{N}$	...	...	...	...	...

Table 1: Free Energy - Ising Model $2D$

Then plot $\frac{\ln(Z)}{N}$ as a function of βJ.

Solution.

The representation of our graph for $N = 25$ with boundary conditions can be understood as a lattice where the borders are connected in *horizontal* (here represented by red-dotted lines) and *vertical* directions (represented by purple-dotted lines).



The adjacent vertices are considered the ones connected by a factor. In this graph, the adjacent vertices of X_8 (blue vertice) are X_3, X_7, X_9 and X_{13} (red vertices).

In order to calculate the exact value of $\dfrac{\ln(Z)}{N}$ we will use a program written in Python due to performance issues and robustness our current problem demands.

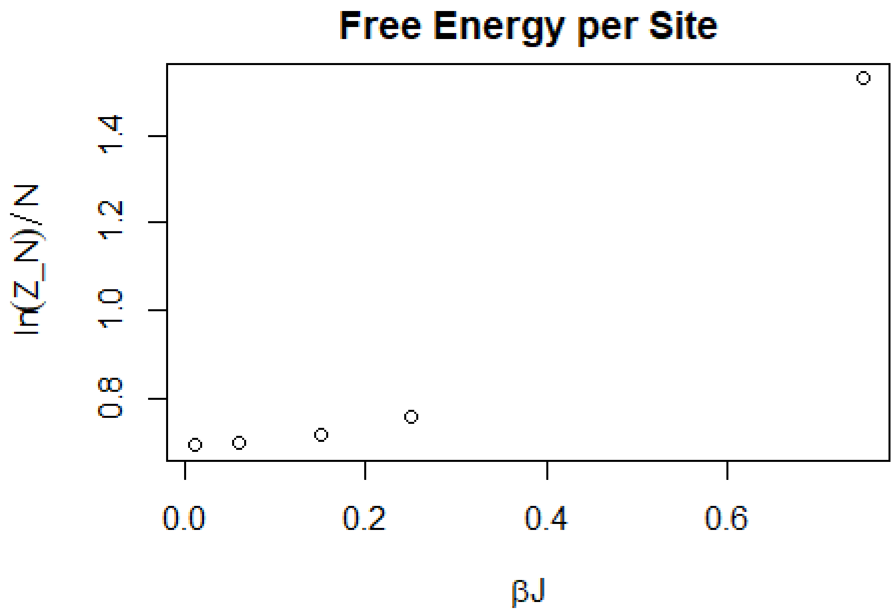
The values obtained for the free-energy are represented in the table below:

βJ	0.01	0.06	0.15	0.25	0.75
$\dfrac{\ln(Z)}{N}$	0.693247	0.696758	0.716122	0.760061	1.530484

Table 2: Free Energy - Ising Model $2D$

The plot $\dfrac{\ln(Z)}{N}$ as a function of βJ becomes as follows:

Figure 1: Free Energy as function of βJ



(b) Suppose $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(l)}$ are drawn according with $p_B(\boldsymbol{x})$. Prove that

$$\hat{\Gamma} = \frac{1}{L|\mathcal{X}|^N} \sum_{l=1}^{L} \frac{1}{f(\boldsymbol{x}^{(l)})}$$

is an unbiased estimator of $1/Z$. (In general, $\hat{\Gamma}$ is a very bad estimator, but we are going to use it in this assignment anyway.)

Solution.

In order to prove the Ogata-Tanemura estimator $\hat{\Gamma}$ is unbiased for $\frac{1}{Z}$, need to calculate it's expectation and check if it is equal to $\frac{1}{Z}$.

Then we have:

$$\begin{aligned} E[\hat{\Gamma}] &= E\left[\frac{1}{L|\mathcal{X}|^N} \sum_{l=1}^{L} \frac{1}{f(\boldsymbol{x}^{(l)})}\right] \\ &= \frac{1}{L} \sum_{l=1}^{L} E\left[\frac{1}{|\mathcal{X}|^N} \frac{1}{f(\boldsymbol{x}^{(l)})}\right] \end{aligned}$$

$$\implies E[\hat{\Gamma}] = \frac{1}{L} \sum_{l=1}^{L} E\left[\frac{\frac{1}{|\mathcal{X}|^N}}{f(\boldsymbol{x}^{(l)})}\right] \tag{3}$$

We know that

(i) if f and h have compatible supports (i.e., $h(\boldsymbol{x}) > 0$ when $f(\boldsymbol{x}) > 0$);

(ii) There is a constant M with $f(\boldsymbol{x})/h(\boldsymbol{x}) \leq M$ for all $\boldsymbol{x}$;

Then we can define a uniform distribution $\boldsymbol{h}$ with the same support as $f(\boldsymbol{x})$, such that:

$$h(\boldsymbol{x}) = \frac{1}{|\mathcal{X}|^N} \tag{4}$$

In our case, by doing $M = Z$ and substituting (4) in (3) we have:

$$
\begin{aligned}
E[\hat{\Gamma}] &= \frac{1}{L}\sum_{l=1}^{L} E\left[\frac{h(\boldsymbol{x}^{(l)})}{f(\boldsymbol{x}^{(l)})}\right] \\
&= \frac{1}{L}\sum_{l=1}^{L} E\left[\frac{1}{Z}\right] \\
&= \frac{1}{L} L\left(\frac{1}{Z}\right) \\
&= \frac{1}{Z}
\end{aligned}
$$

As result, we have that the Ogata-Tanemura estimator $\hat{\Gamma}$ is an unbiased.

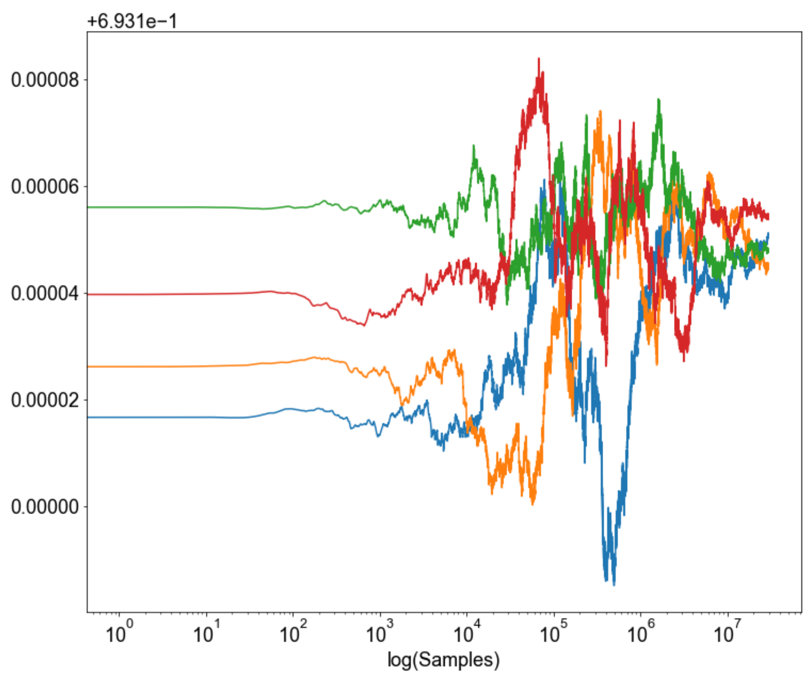
(c) For different values of βJ given in the above table, run the Gibbs sampling algorithm to draw samples according to $p_B(\boldsymbol{x})$. Then employ $\hat{\Gamma}$ to compute an estimate of $\ln(Z)/N$. Specifically, for each value of βJ, report four sample paths in a single plot that show the convergence of the estimator. Note that your plots should contain four sample paths, which show the estimated quantity $\ln(Z)/N$ in Y-axis vs. the number of samples (in log scale) in X-axis. Set the required number of samples $L = 3 \times 10^7$ in all simulations. At the end, report five plots (for each value of βJ), where each plots contains four different sample paths.

Solution.

In this question I will plot the graphs obtained with $L = 3 * 10^7$ samples and $J = 1.0$. In order to do some comparisons I also added additional simulations with $L = 10^5$ and different values of J to identify how the coupling energy affects convergence in multiple scenarios of temperature.

The Sample Paths for each beta requested are presented in the following graphs[1]:

Figure 2: Free Energy - $\beta = 0.01$



[1]Sometimes, the number presented on top of the graph will represent the magnitude of Y-scale, specially when the convergence is strong and the variability of sample paths is small.

Figure 3: Free Energy - $\beta = 0.06$

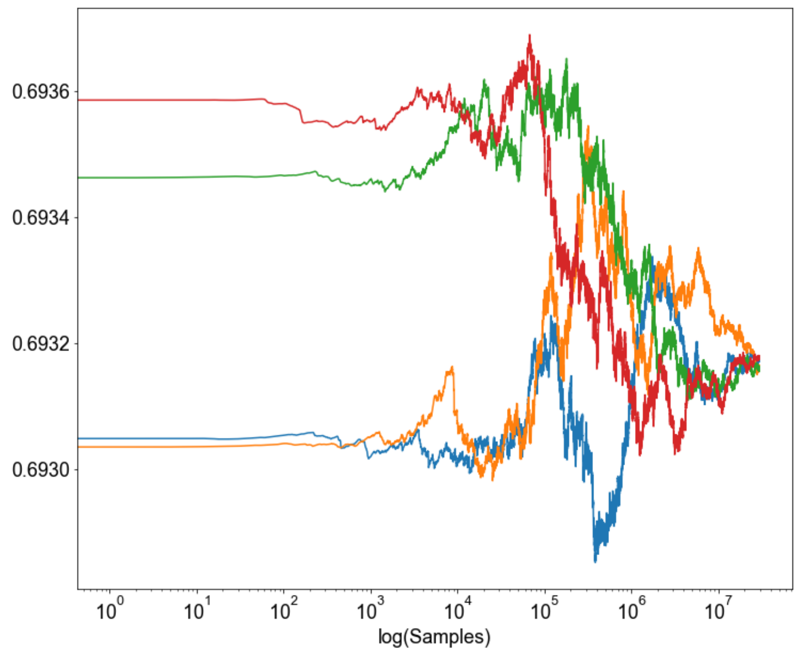


Figure 4: Free Energy - $\beta = 0.15$

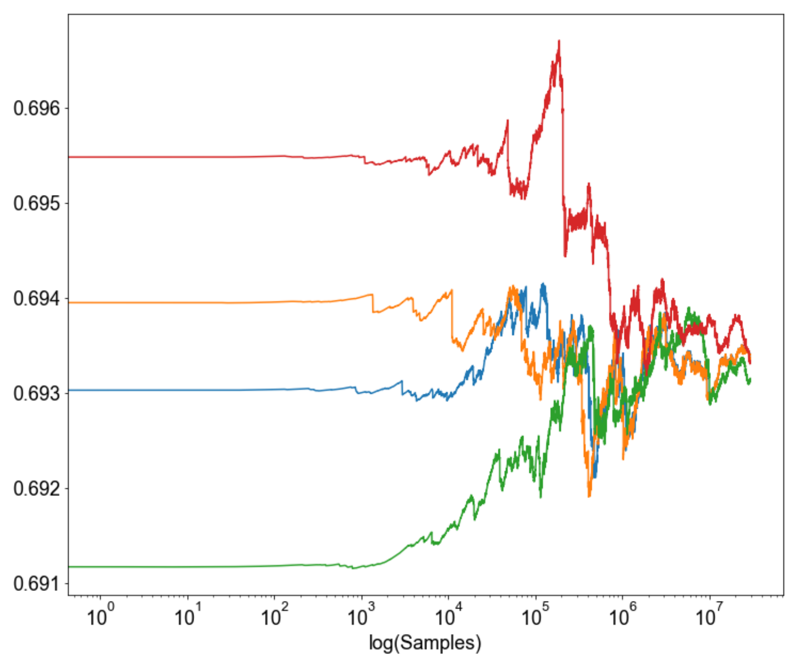


Figure 5: Free Energy - $\beta = 0.25$

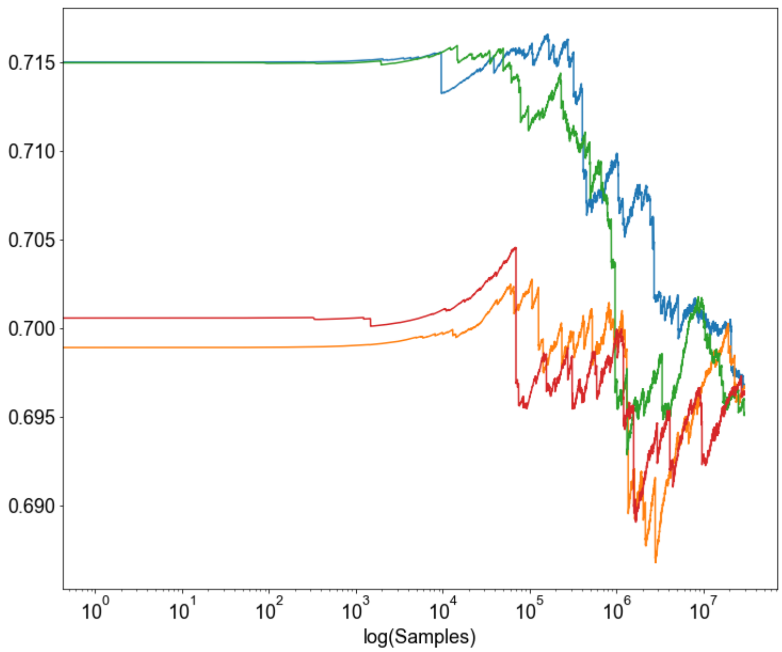
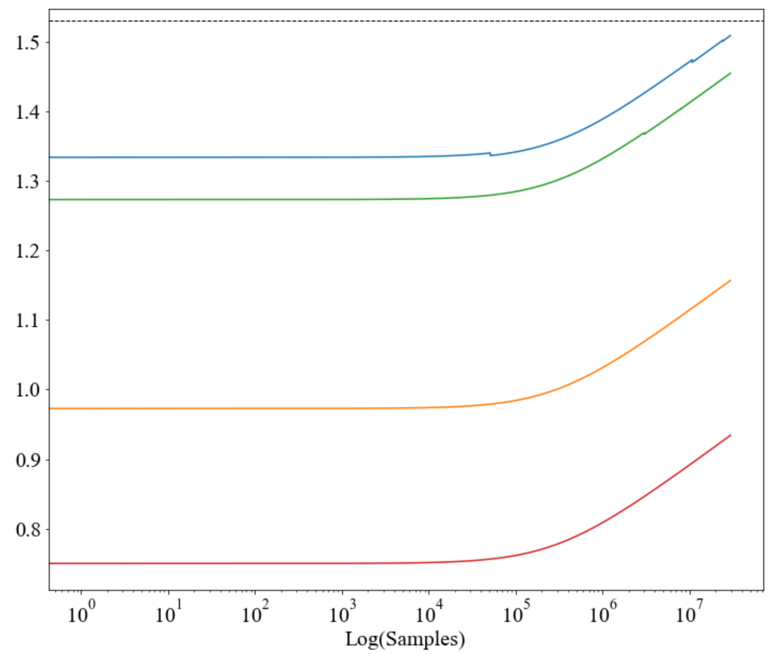


Figure 6: Free Energy - $\beta = 0.75$



(d) What are your observations regarding the convergence of the estimator for different values of βJ? (Discuss in terms of low-temperature and high-temperature regimes.)

Solution.

To answer this question it will be presented an additional set of simulations to illustrate how the convergence is affected by different values of temperature and coupling energy.

For the sake of simplicity, I concentrated on *lowest* and *highest* temperature, i.e. $\beta = 0.75$ and $\beta = 0.01$, respectively and also defined 02 arbitrary values for J ($J = 0.43$ and $J = 0.0084$) to verify how they affect the convergence of the estimator Ogata-Tanemura $\hat{\Gamma}$.

For each pairwise combination, it was also calculated the *Absolute Error*, *Relative Errors* and *Histogram of Sample Paths*.

We will see the results on the following sub-sessions.

Simulations with $K = 10^5$, $J = 0.43$

Figure 7: Sample Paths

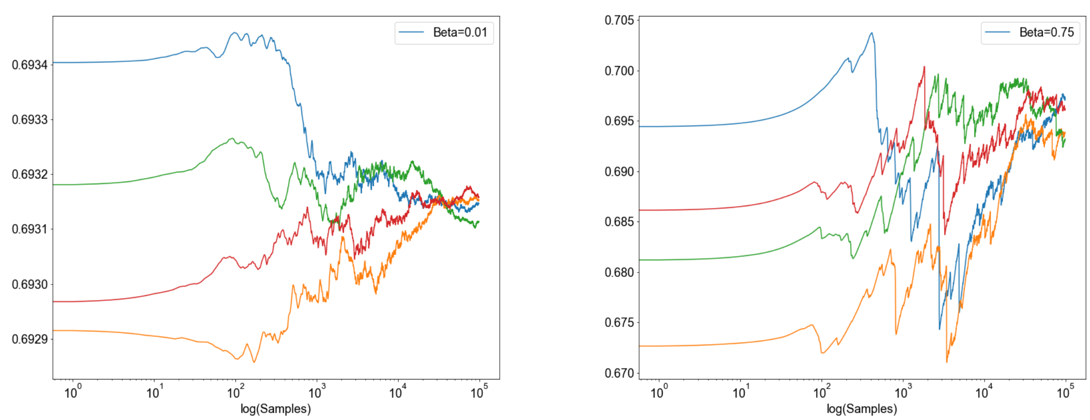


Figure 8: Absolute Error

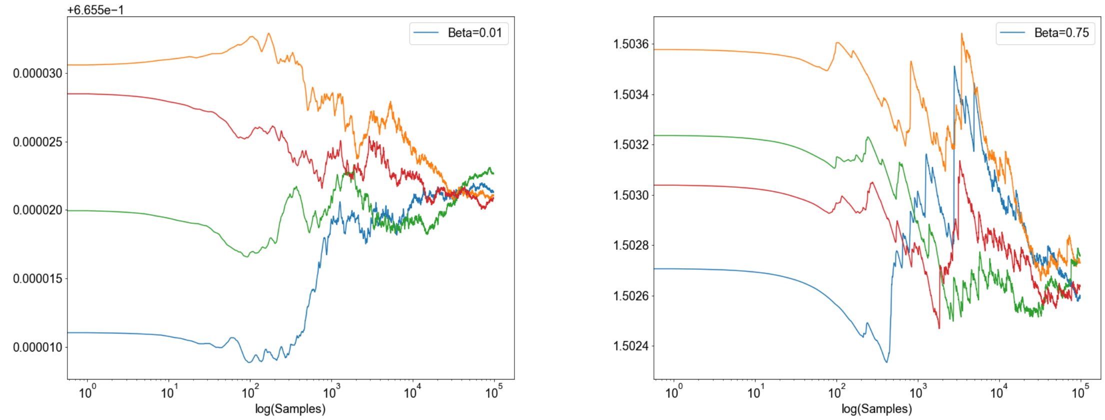


Figure 9: Relative Error

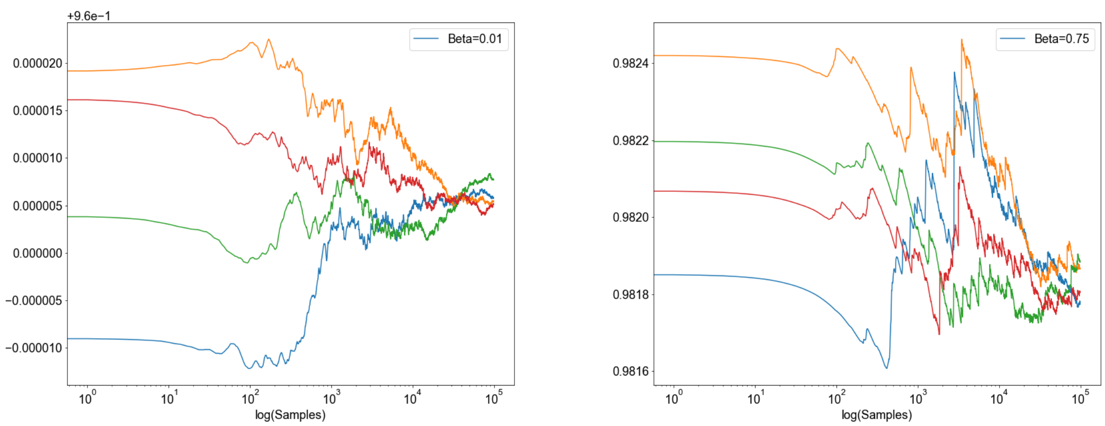
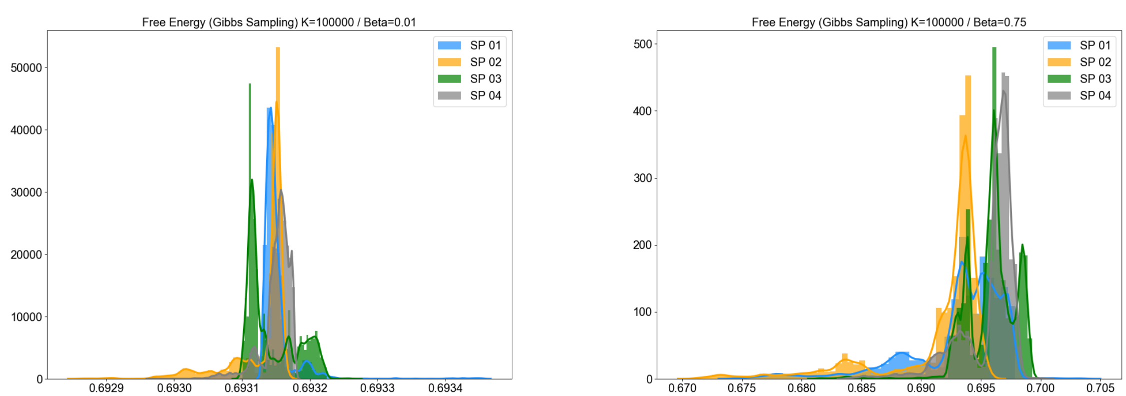


Figure 10: Histogram of Sample Paths



Simulations with $K = 10^5$, $J = 0.0084$

Figure 11: Sample Paths

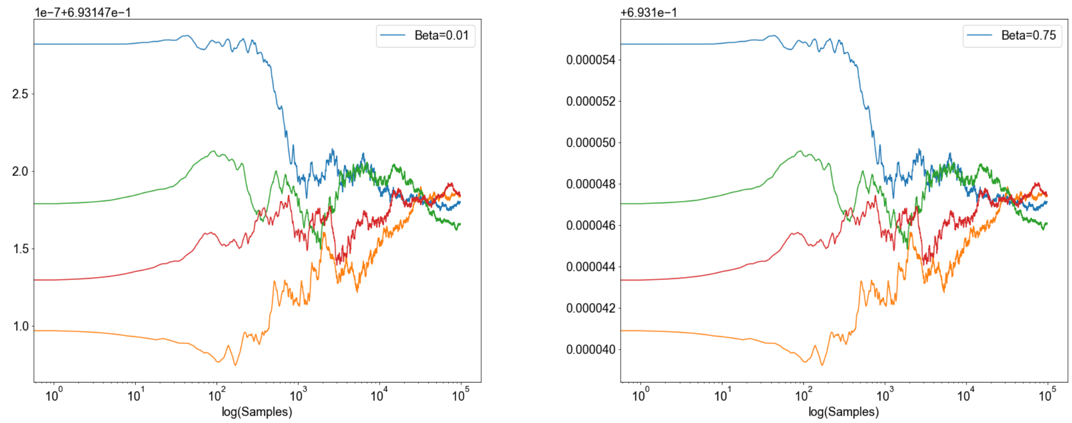


Figure 12: Absolute Error

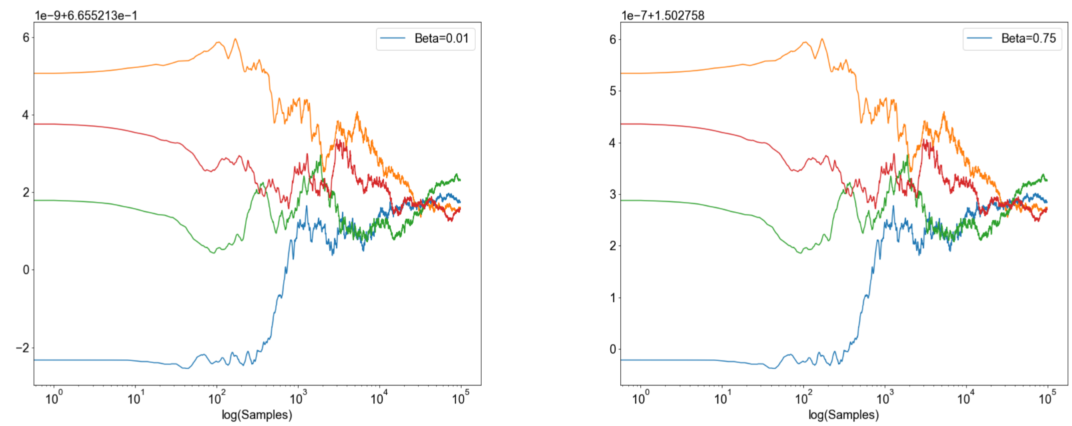


Figure 13: Relative Error

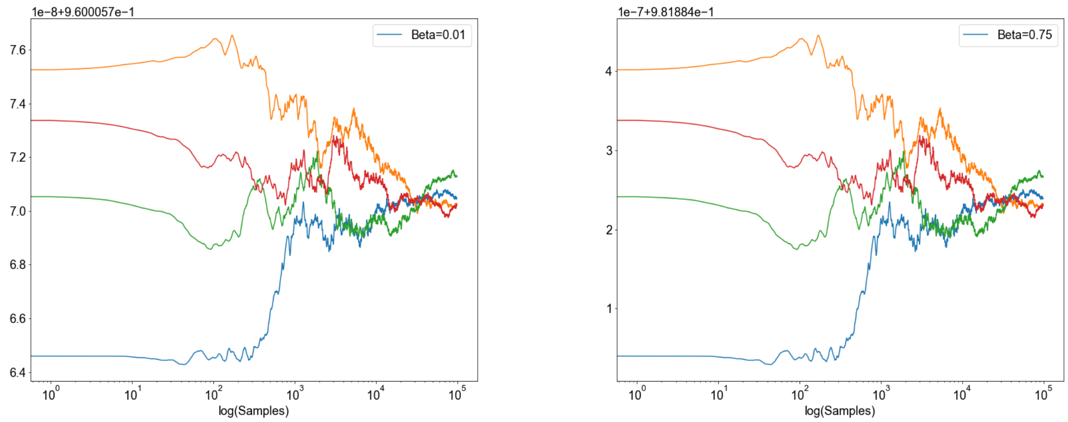
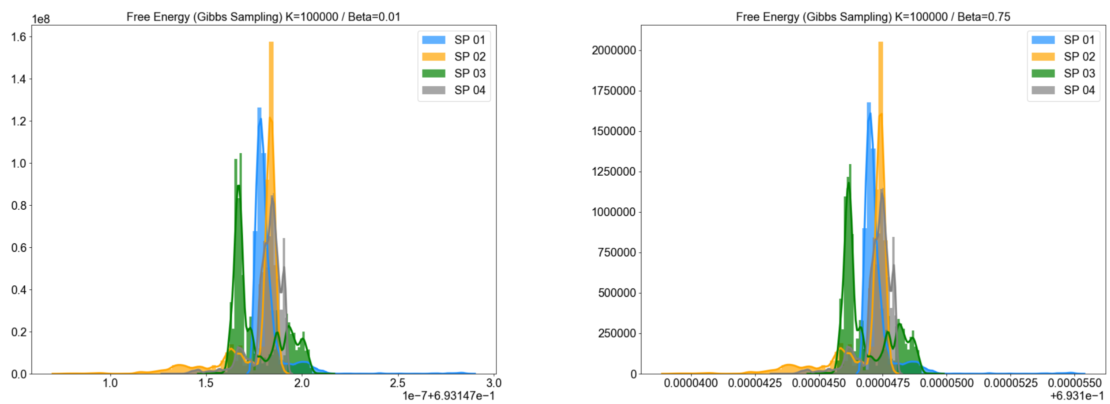


Figure 14: Histogram of Sample Paths



Comments

We can see some interesting patterns when observing the behaviour of the estimated free energy throughout the graphs with different β's and J's:

1. When comparing the graphs for $J = 0.43$ and $J = 0.0084$ e observed that sample paths converged faster in the second one.
2. Both *Relative* and *Absolute* errors are greater for smaller temperature (i.e., largest β)
3. The distribution of samples showed smaller variability for larger values of J (i.e., less disperse in related to the true value of the free energy) than when at higher temperature;

4. For the same level of coupling energy, we observed in item (c) that higher temperature leads to faster convergence; on the other hand, the opposite also applies: lower temperature turns difficult the convergence of Gibbs Sampling (in some cases, not converged at all affecting the stability of the model, see $\beta = 0.75$ for $L = 3 * 10^7$).

5. In general, an additional conclusion can be drafted for lower values of J (weak coupling energy) which enables a very fast convergence. In this case, the samples generated by the Monte Carlo algorithm are *more independent* which benefits the convergence of Markov Chain during simulations;

6. On the other side, higher values of J turns the samples generated by Gibbs-Sampling algorithm more dependent and, as a consequence, harder to converge;

7. Another interesting aspect regarding the temperature, when considering the same value of J, is that *higher temperatures (i.e., lower β) generate faster convergence*. The opposite is also true: *lower temperatures (i.e., higher β) generate slow or even no-convergence*;

8. As we already knew, the estimator $\hat{\Gamma}$ is not an *state-of-the-art* predictor. This could be verified by the *Error Graphs (absolute and relative)* which did not converged to *zero*, even with larger amounts of samples. In fact, in some cases the error grew with increasing number of samples (L), suggesting the estimator might be propagating its error as the number of samples increases.

References

[1] Molkaraie, M., Gomez, V. *Monte Carlo Methods for the Ferromagnetic Potts Model Using Factor Graph Duality*. arXiv:1506.07044v4, 2018.

[2] Ogata, Y., Tanemura, M. *Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure*. Inst. Statist. Math., vol. 33, pp. 315–338, 1981

[3] Molkaraie, M., Loeliger,H. A. *Monte Carlo Algorithms for the Partition Function and Information Rates of Two-Dimensional Channels*. arXiv:1105.5542v2, 2012.

[4] Robert, C. P., Casella, G.. *Introducing Monte Carlo Methods*. Springer, 2013.