0.1 Experiments

I check if the assigned to the node value depends on degree of the node. In order to check this I counted average values of the node with fixed degree. On the figure is described how average value of the nodes with the same degree depends on this degree. According to the simulations we can observe that value of the node doesn't depend on its degree.

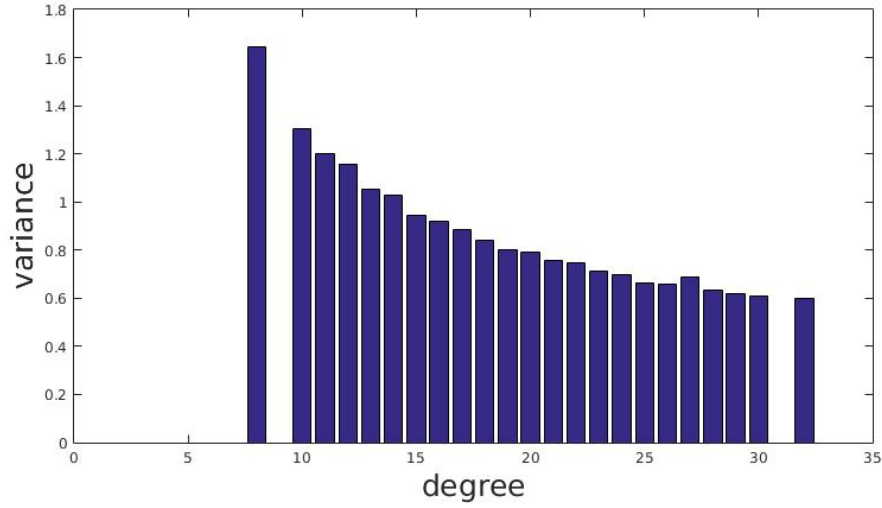However what depends on degree is variance of values. The smaller is degree the bigger is variance.

Figure 1: ER graph with 200 nodes, radius 0.1, average degree 20, 10 values

In all approximations I consider that there is no correlation between values of neighbors.

Let's look at the nodes i and j. It is reasonably to assume that expected values on all the nodes is the same.

$E[(x_j - x_i)^2] = E[x_j^2] + E[x_i^2] - 2E[x_i x_j] = Var[x_j] + (E[x_j])^2 + Var[x_i] + (E[x_i])^2 - 2(cov(x_i, x_j) + E[x_i]E[x_j]) = Var[x_j] + Var[x_i] - 2cov(x_i, x_j)$

In the last approximation

$$E_{apr}[(x_j - x_i)^2] = Var_{apr}[x_j] + Var_{apr}[x_i]$$

But according to experimentation $E_{apr}[(x_j - x_i)^2] = E_{real}[(x_j - x_i)^2]$. That means
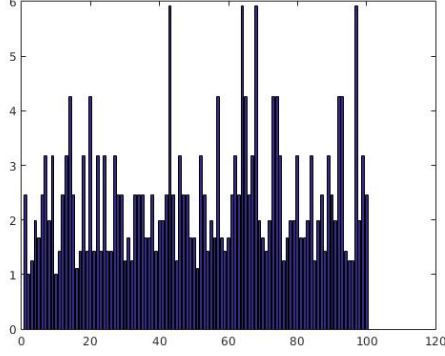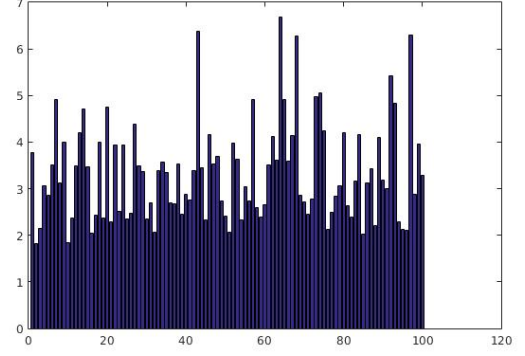
1

Figure 2: Approximated variance of the nodes

Figure 3: Real variance of the nodes

$$Var_{apr}[x_j] + Var_{apr}[x_i] = Var_{real}[x_j] + Var_{real}[x_i] - 2cov_{real}(x_i, x_j)$$

From the pictures I see that variances have similar shape. In reality variance is bigger than in our approximation.

$$Var_{apr}[x_j] <= Var_{real}[x_j]$$

But owing to the covariance in reality left and right part become surprisingly close.

$$Var_{apr}[x_j] + Var_{apr}[x_i] = Var_{real}[x_j] + Var_{real}[x_i] - 2cov_{real}(x_i, x_j)$$

Especially good this approximation works for ER-graph. So maybe I can count covariance at least in this case.

What I see from the graph (check): When temperature is big enough true variance equals to approximated (look at each node).

Covariance is strange when temperature is near zero. probably cov can be also negative.

I want to check if the time series from values that I see during the random walk on the graph with field is stationary. In order to do this I need to see that mean does not change over time (seems like a true) and autocovariance function depends on difference.

Both structure of the graph and the fact that values of neighbors are correlated bring error to the estimations.

Suppose, that we know correlation between the nodes in the graph. For example we know correlation between nodes, depending on the distance between them.
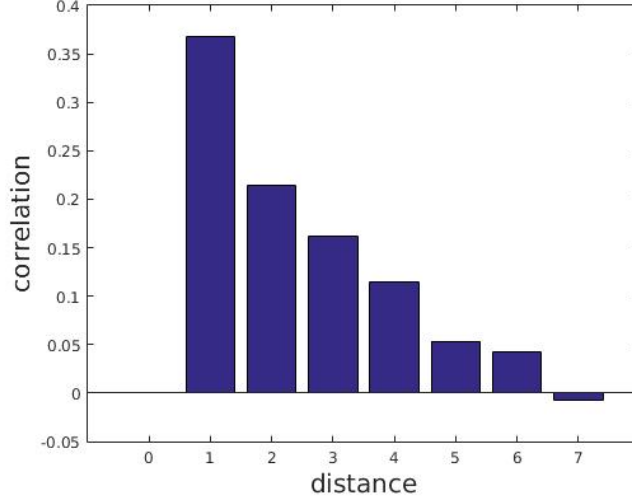
2

Figure 4: Dependence of correlation between nodes on the distance between them

With this information what can we say about error of estimation using Y_1, Y_2, Y_3, Y_n (values collected during the random walk)? It is clear that correlation between Y_i and Y_j corresponds to the correlation between node with distance 1. But, for example Y_1 can be connected to Y_5. So what we can say about correlation of nodes that we collected. Obviously, it depends on graph structure.

Is seems for me that bias depends on structure and variance depends on correlation between values (check with simulations).

[check that bias does not depend on values correlation on the chain]

One more noticed equality $Var[x_i - x_j] = E[(x_i - x_j)^2]$

$$Var\left[\frac{Y_1 + Y_2 + ... + Y_n}{n}\right] = \frac{\sum\limits_{i=1..n} Y_i}{n^2} + \sum_{i=1..n}\sum_{j=1..n} cov[Y_i, Y_j]$$
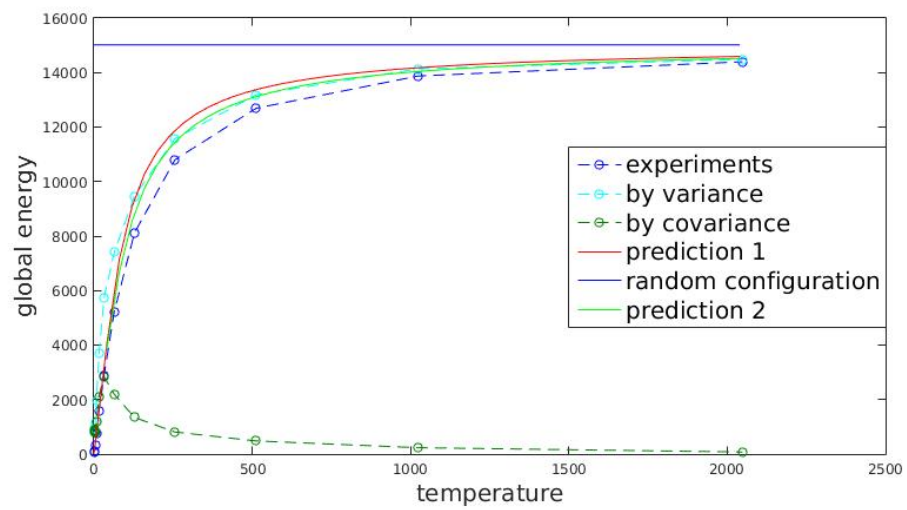
My attempts to calculate error knowing covariance.

3

Figure 5: RGG different energies