Peer Review of CA 4 – Assignment by Group 2, Review by Group 4

4a)

Problem Setup:

The problem in terms of workers and loss functions used are well explained.the loss and loss gradients formulas are correct.

Data Preprocessing:

The MNIST data preprocessing in terms of normalization and data selection is presented in detail. Various statistics of the data are shown, including class histograms for test and train sets.

Distribution over workers:

The workers data structures are presented in detail in terms of dimension and data used for each worker.

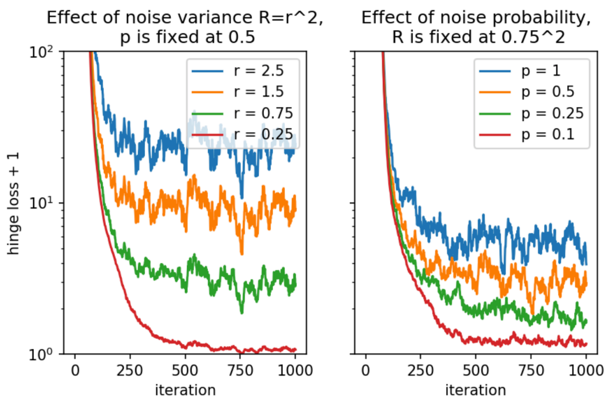
Computation of gradients:

It is well explained how the subgradients are computed at the workers and aggregated at the master.

Robustification:

Unfortunately robustification of your solution is missing. It would have been nice to see how your solution performed with this robustification step. However an high-level explanation of how to implement this robustification is provided.

Results:



I really like the way you present the results! It Clearly shown the dependence on both R and p and the results show consistency. Also the hyperparameter used and in terms of MonteCarlo iterations and lambda and step size are clearly stated! Very good job!

4b),c)

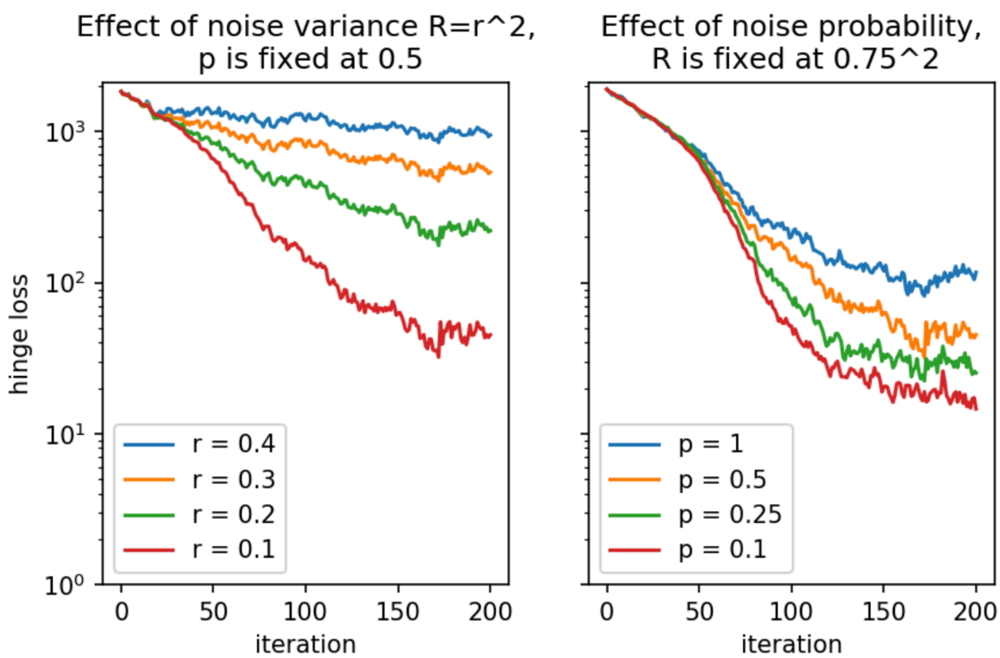
Problem Setup & Data preprocessing:
The problem setupo and data processing steps are the same as in 4a).

Distribution over workers:
The a boolean connectivity matrix A of shape (10,10). Defining the star topology is clearly presented.

Implementation:
The implementation of the algorithm is presented in the following pseudocode snipped.

- for each worker `i=1,...,10`:
  - get the degree `d_i` defined as the number of neighbors + 1
  - get all neighbors local parameters (corrupted by additive Gaussian noise (variance `R`) with probability `p`.
  - set `w_bar` to the average of all the received parameters together with the local parameter vector.
  - compute the gradient `g` w.r.t. `w_bar`.
  - set the local parameter vetor to `w_bar - alpha * g`, where `alpha` is the step size.

Results:



Similar and consistent behavior with respect to p and R of the hinge loss across iteration is shown. Slightly less level of details is provided in terms of hyperparameters.

The answer to 4c) is correct and well motivated.