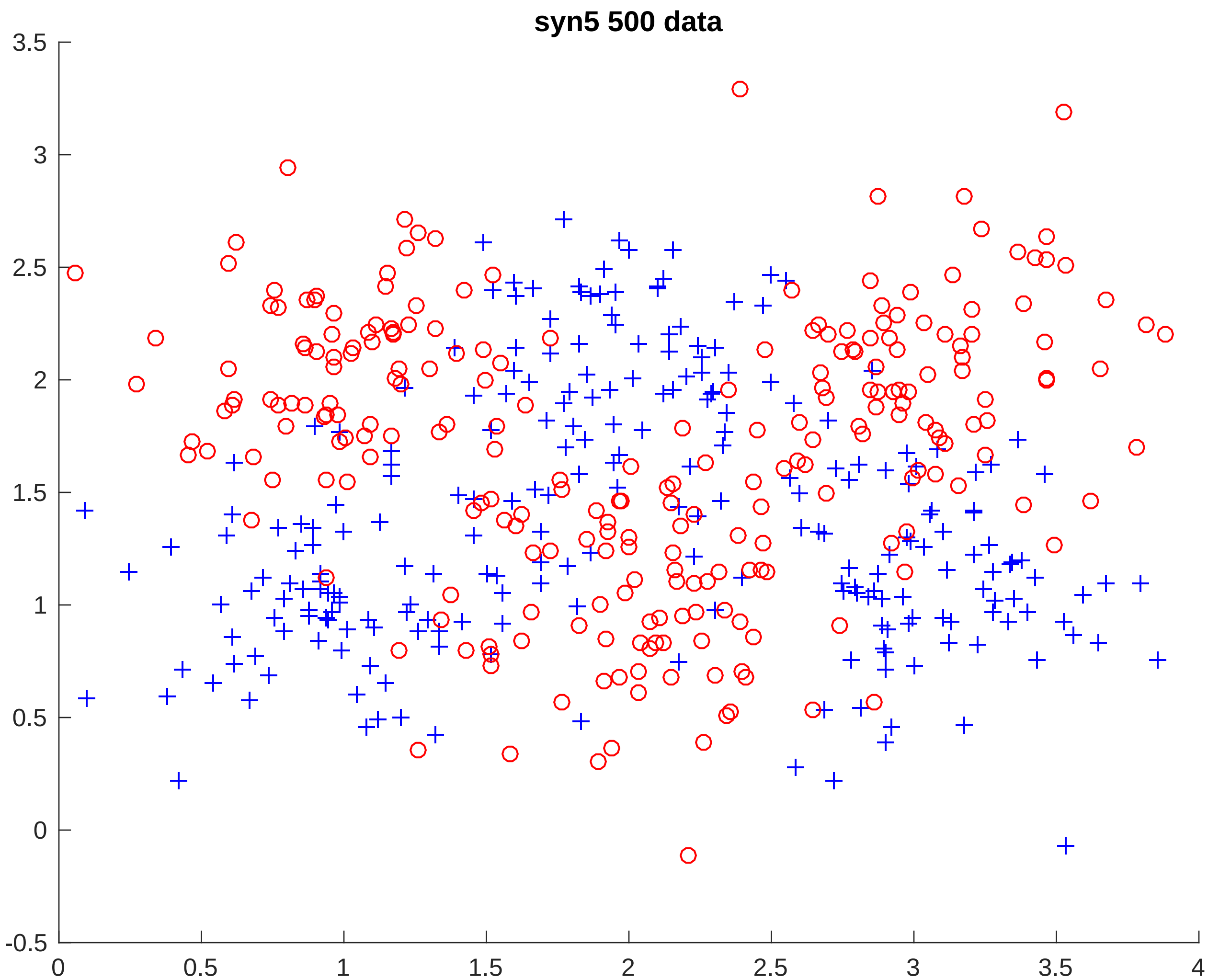


Figure 1 - This is an illustration of the binary classification problem used in the experiment. Six Gaussian “blobs”, three for each class. The problem is non-linear.

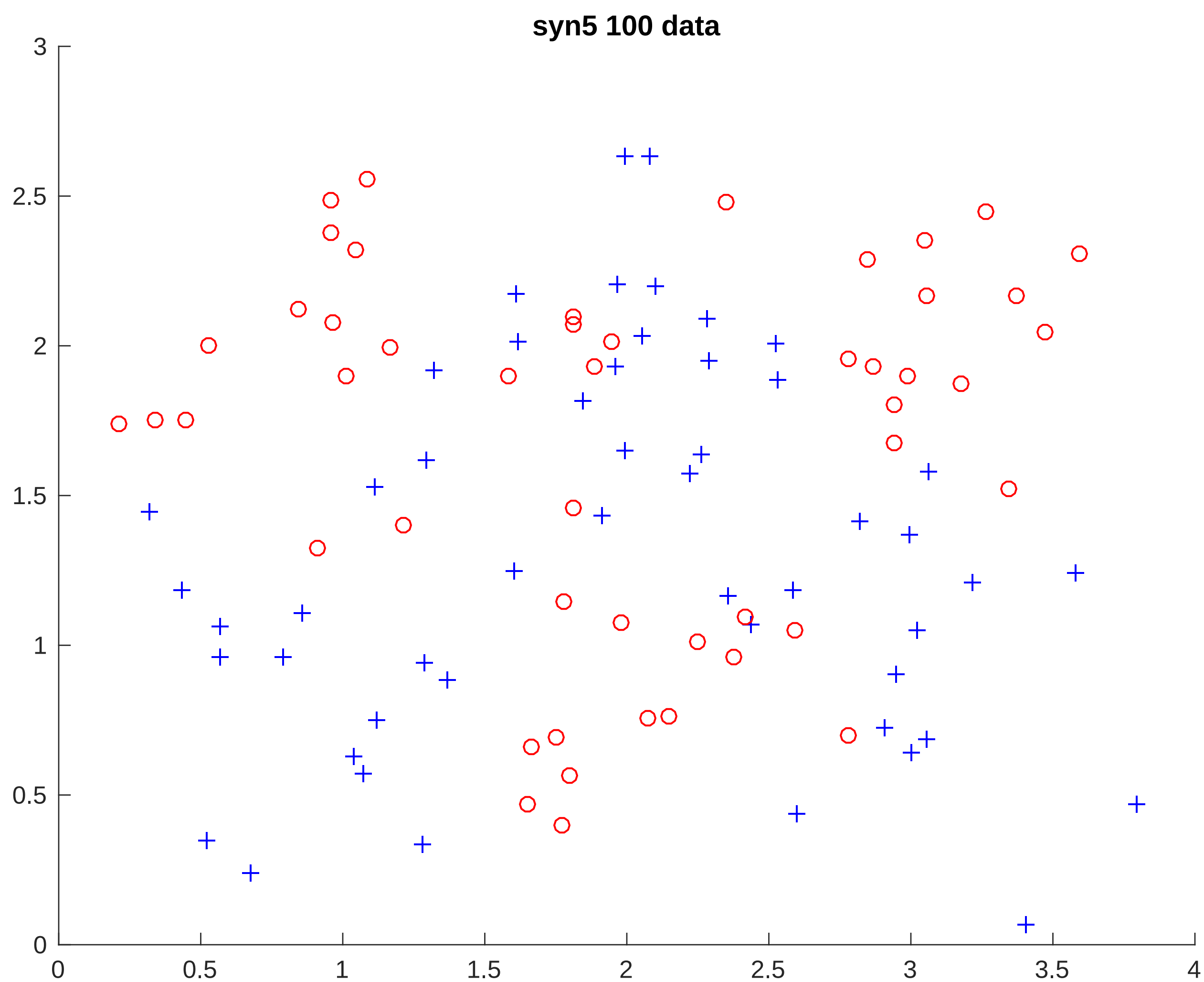


Figure 2 - The training data used for the first set of experiments.

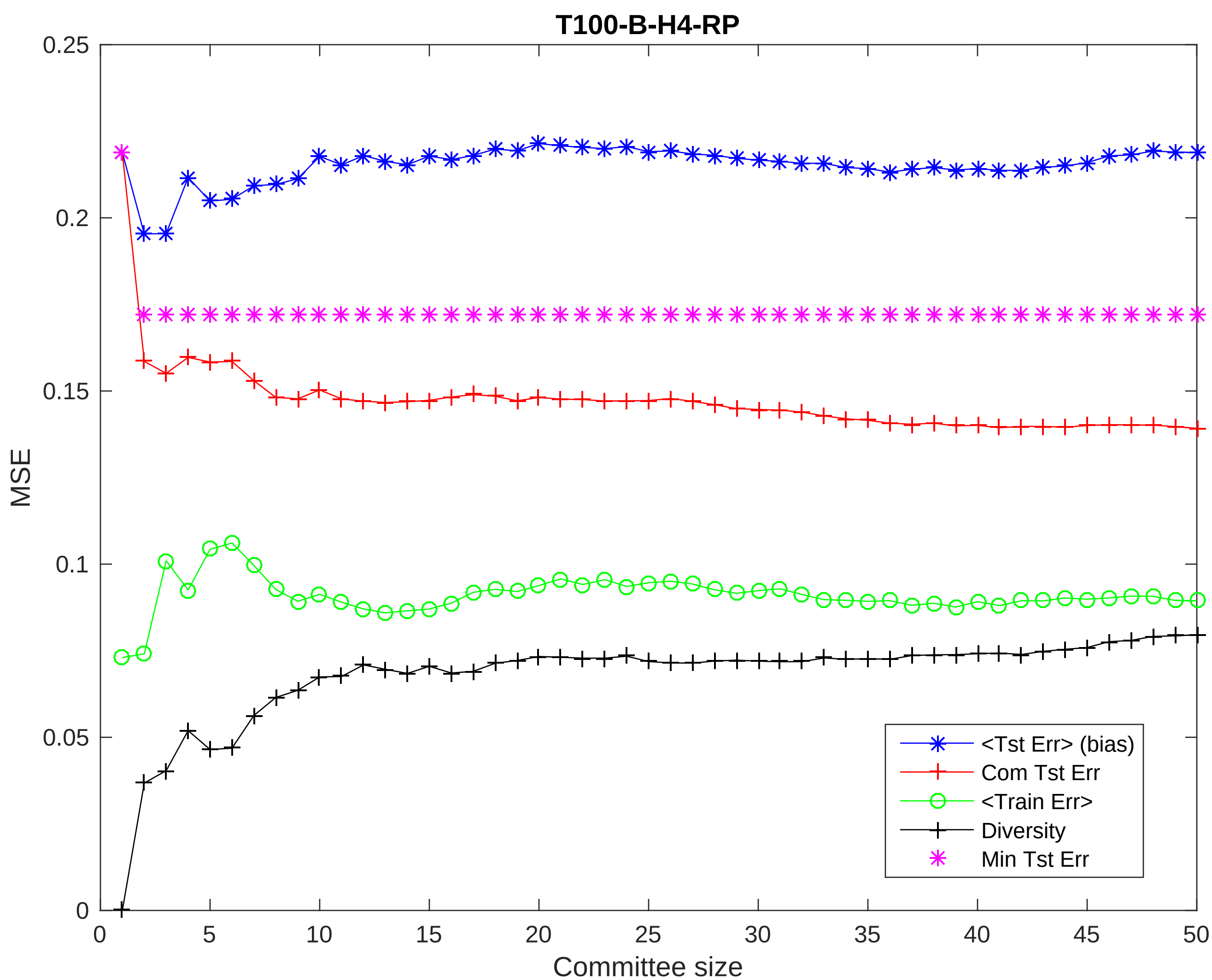


Figure 3 - The experiment (T100-B-H4-RP: Five hidden nodes) shows the effect of the ensemble: The general setup of the experiment is to start training MLPs for the classification problem. The more MLPs we train the larger the ensemble grows. Up to 50 members was trained for each experiment.
All graphs in this experiment follows the layout in this figure. There are five curves. On the x-axis we have the number of members of the ensemble, committee size. Y-axis is MSE (even though this is a classification problem).:
Red curve shows the ensemble test error. This error can be decomposed into the bias (blue curve) and the diversity (black curve). Red = blue – black!
The green curve shows the training error and can be used to indicate the level of overfitting (how?).
The magenta curve shows the current best individual MLP test error.
In this figure the 100-cases training dataset was used, bagging as the ensemble generator and 4 hidden nodes.

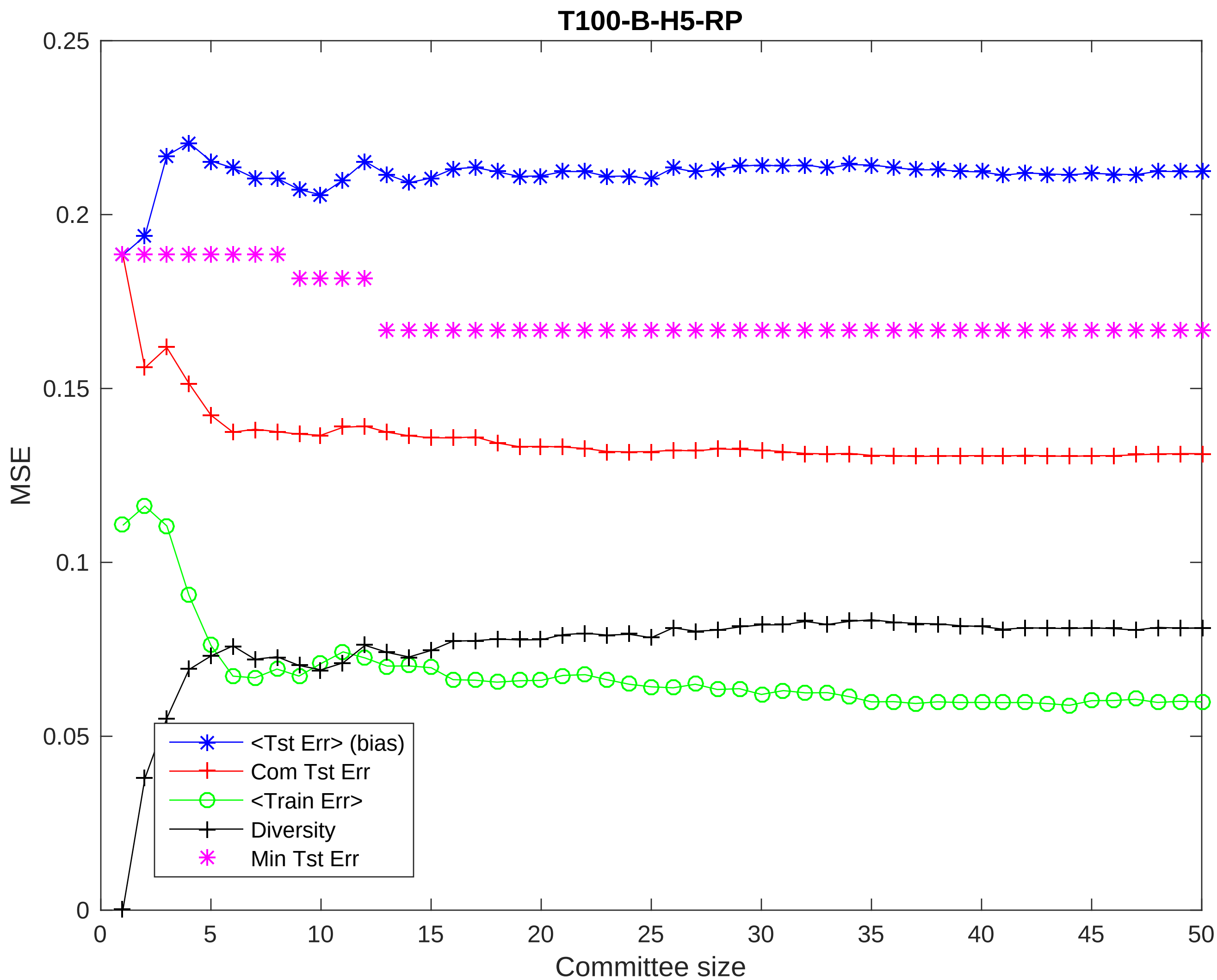


Figure 4 - Up to 20 hidden nodes: We can see that the training error is really small, indicating overfitting of the individual MLPs. Still the red curve shows almost the same test error.

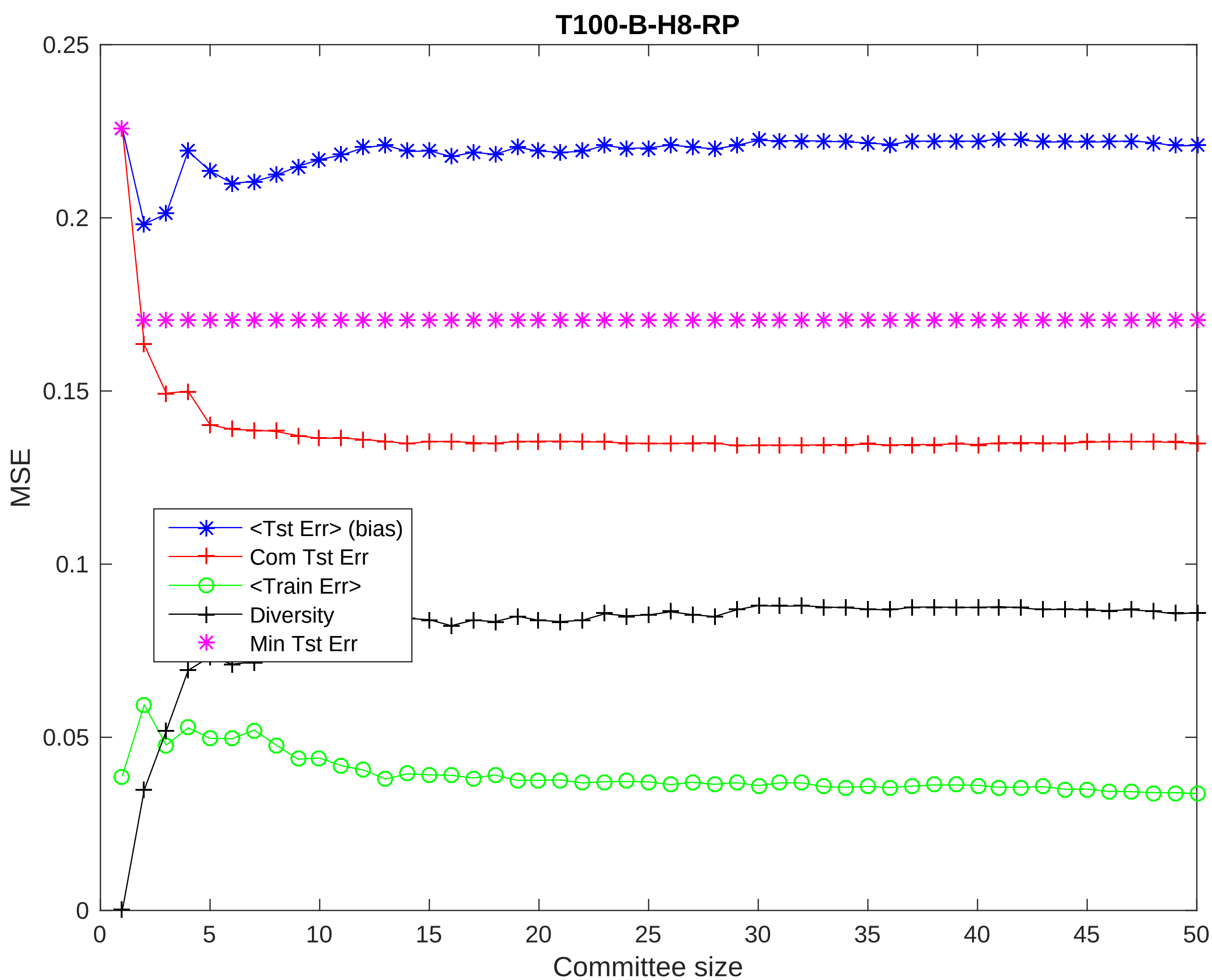


Figure 5 - Up to 20 hidden nodes: We can see that the training error is really small, indicating overfitting of the individual MLPs. Still the red curve shows almost the same test error.

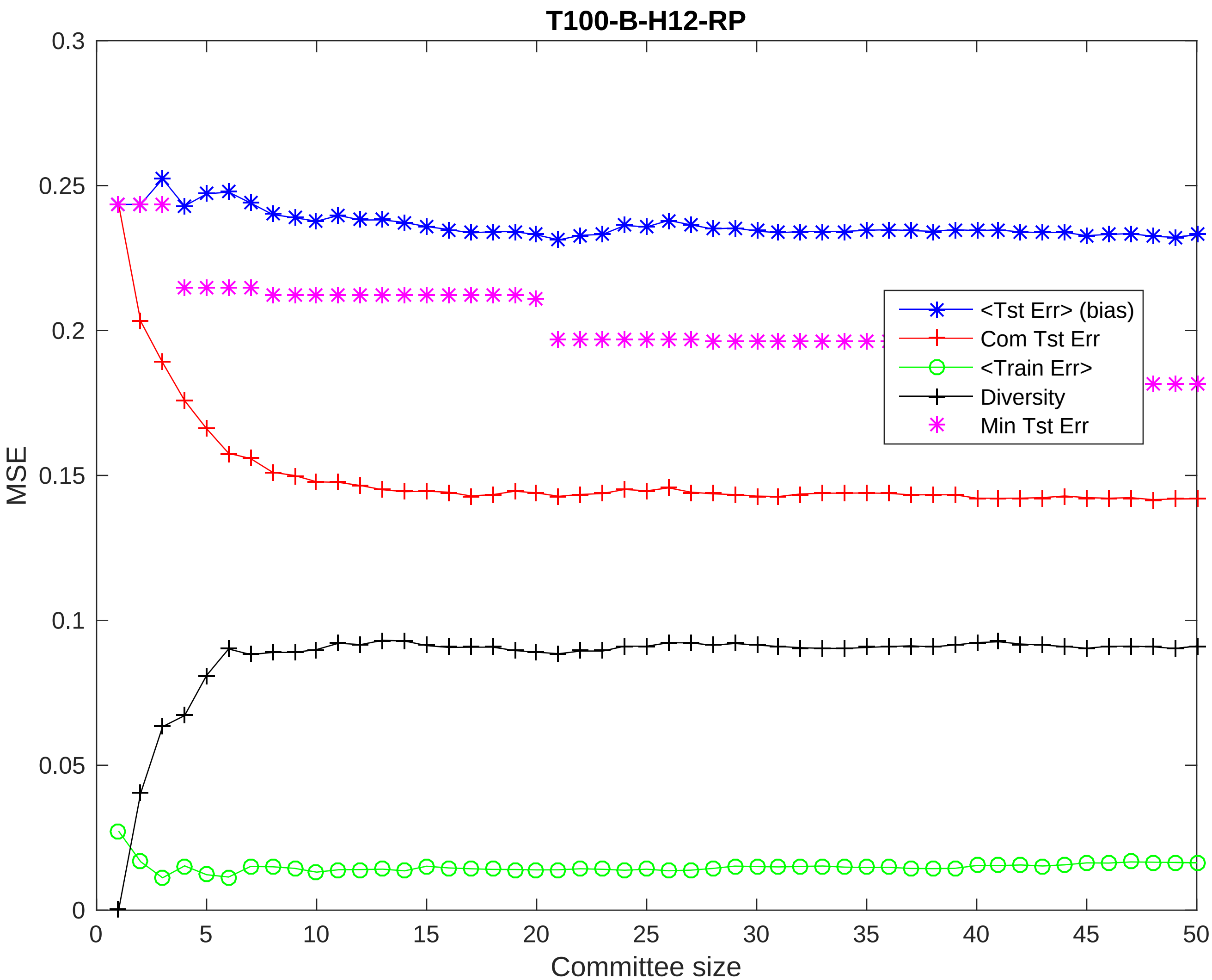


Figure 6 - Up to 20 hidden nodes: We can see that the training error is really small, indicating overfitting of the individual MLPs. Still the red curve shows almost the same test error.

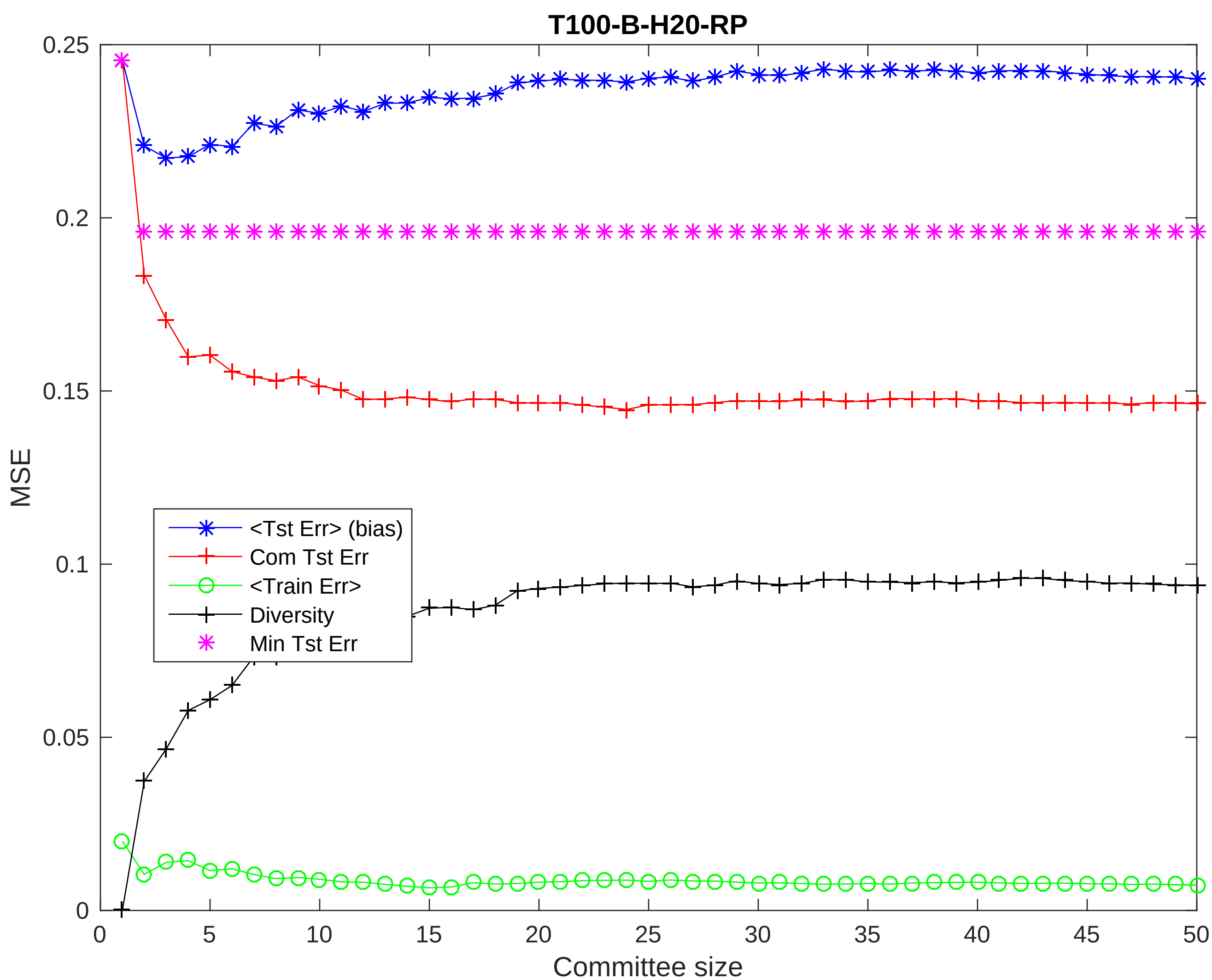


Figure 7 - Up to 20 hidden nodes: We can see that the training error is really small, indicating overfitting of the individual MLPs. Still the red curve shows almost the same test error.

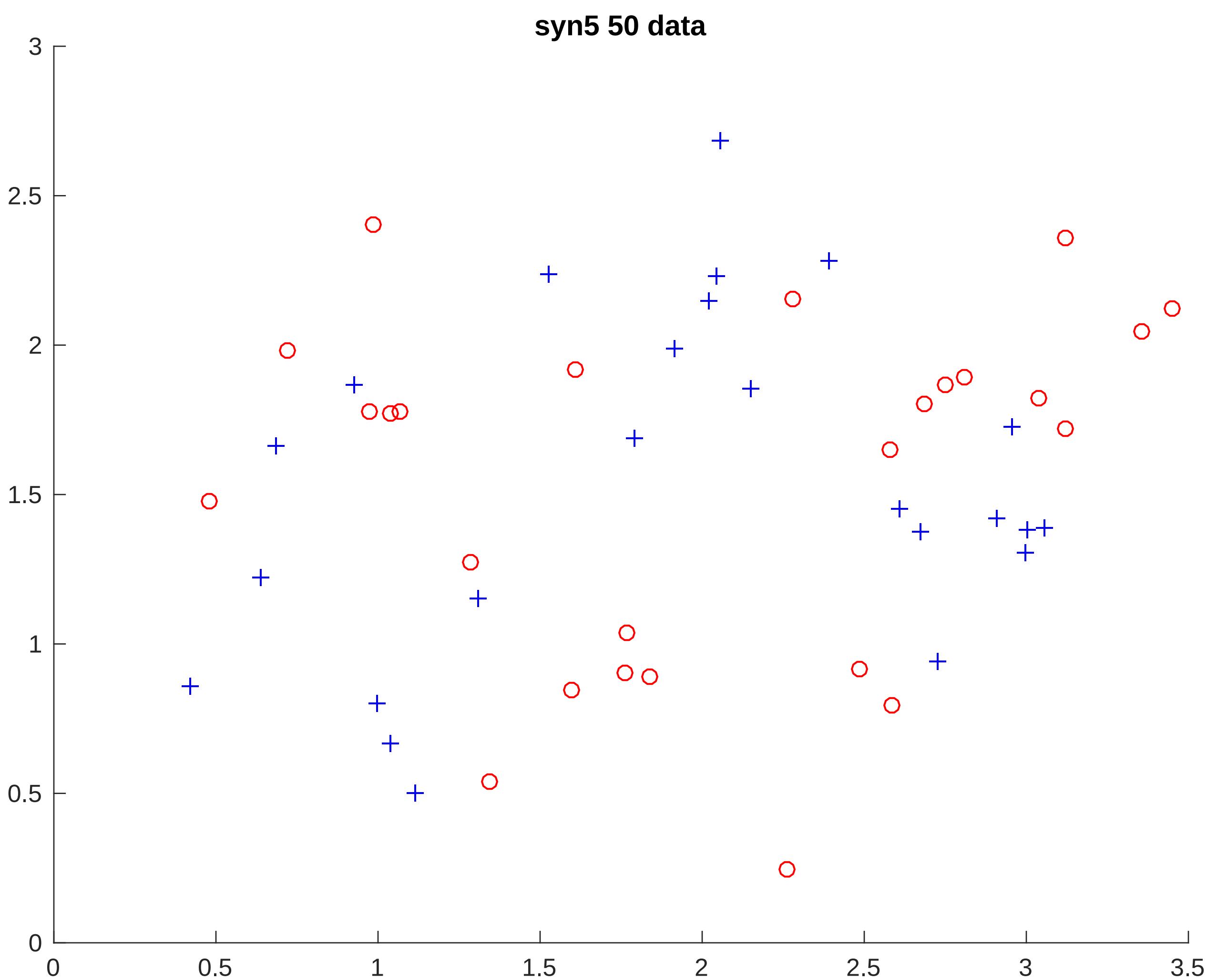


Figure 8 - Now an even smaller training dataset

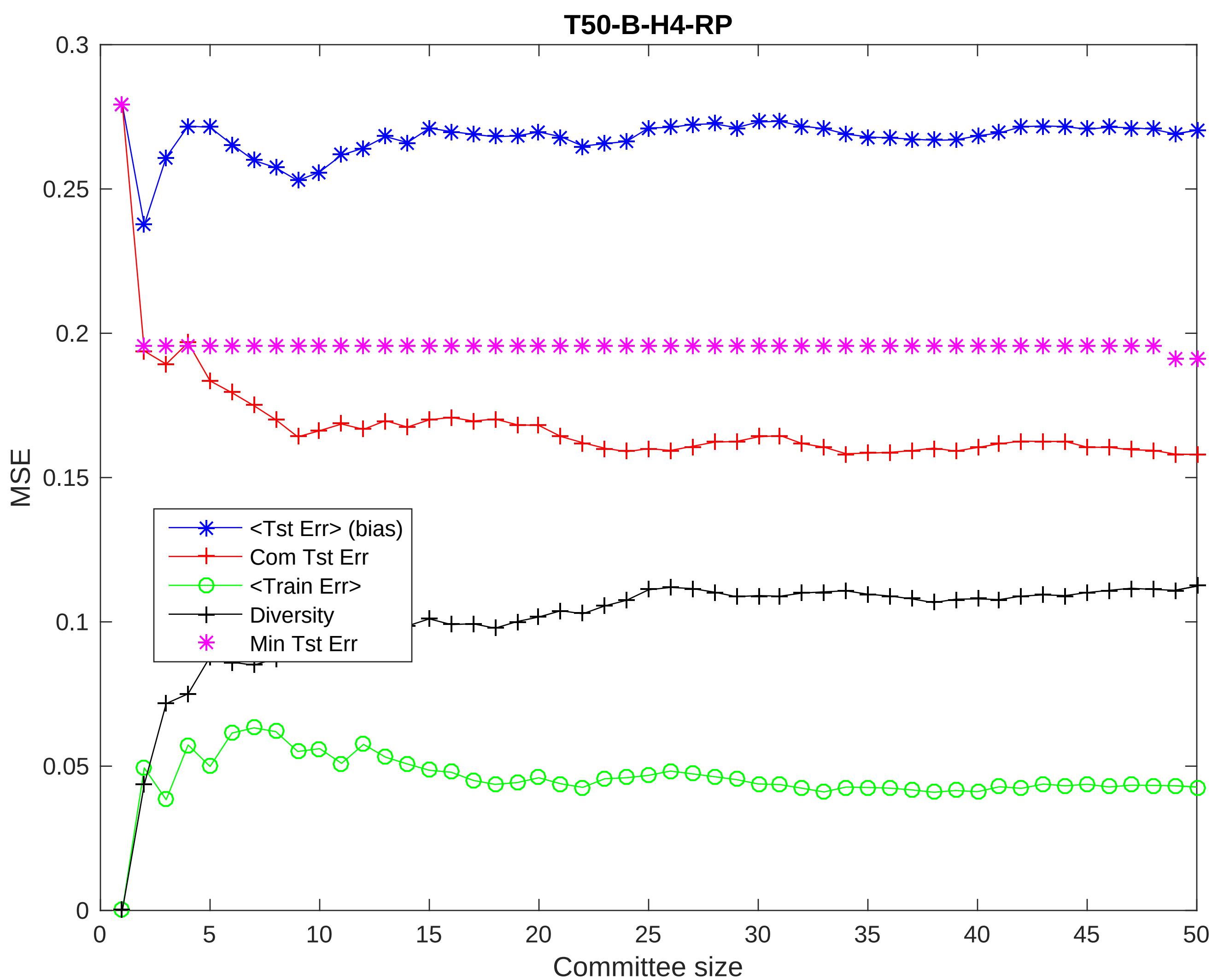


Figure 9 - Up to 20 hidden nodes. Again, for the 20 hidden MLPs we have very low training error, but the ensemble does not overfit.

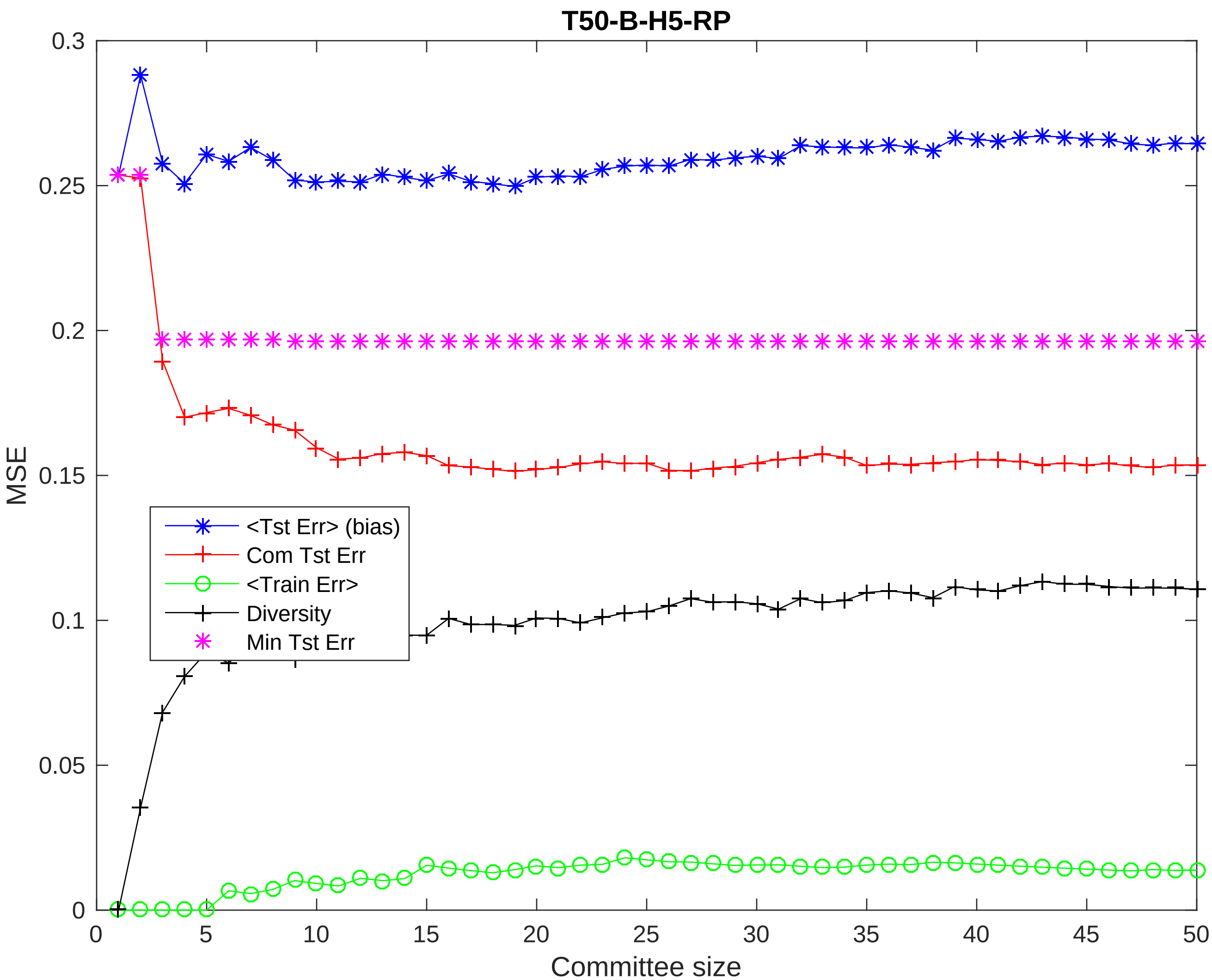


Figure 10 - Up to 20 hidden nodes. Again, for the 20 hidden MLPs we have very low training error, but the ensemble does not overfit.

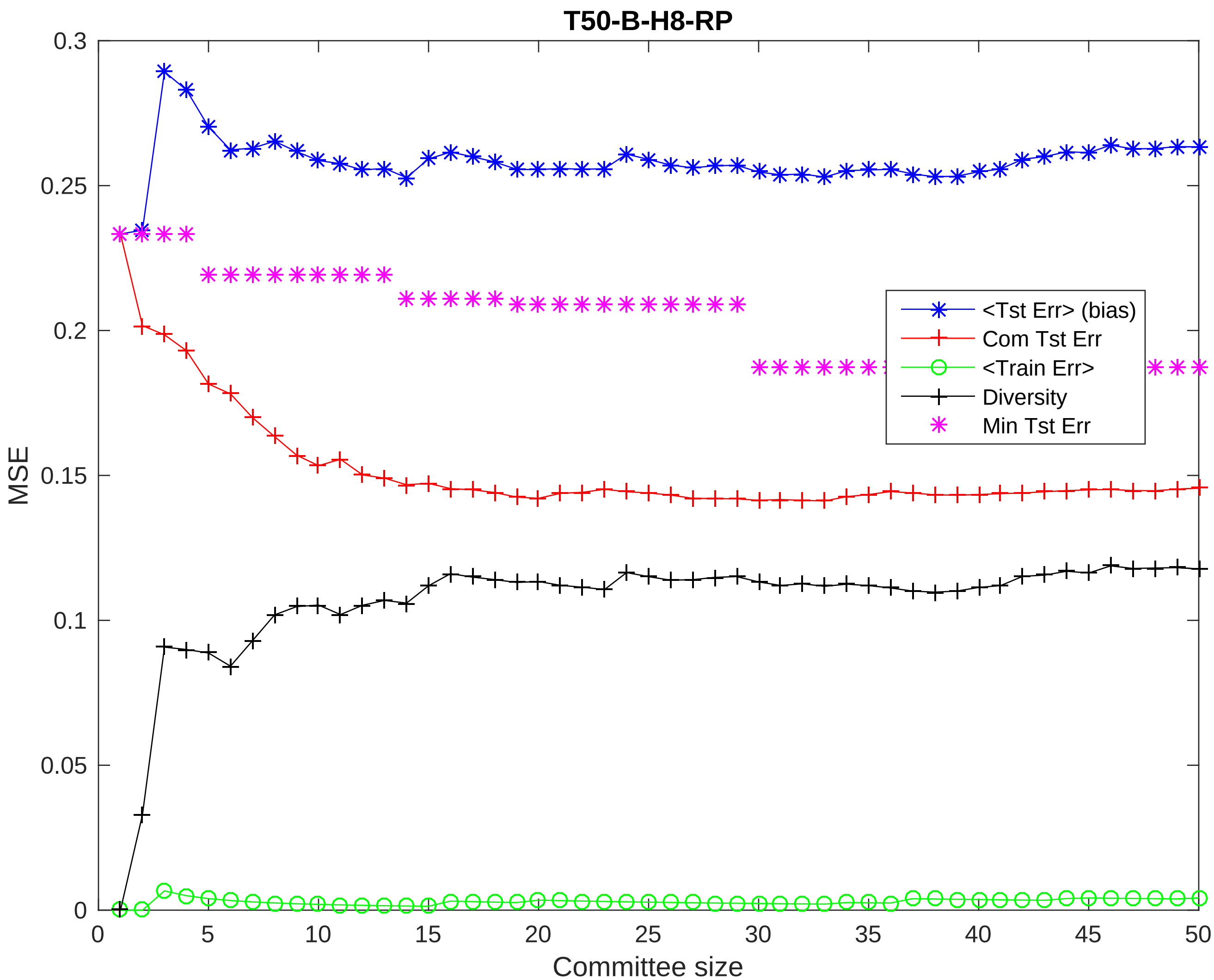


Figure 11 - Up to 20 hidden nodes. Again, for the 20 hidden MLPs we have very low training error, but the ensemble does not overfit.

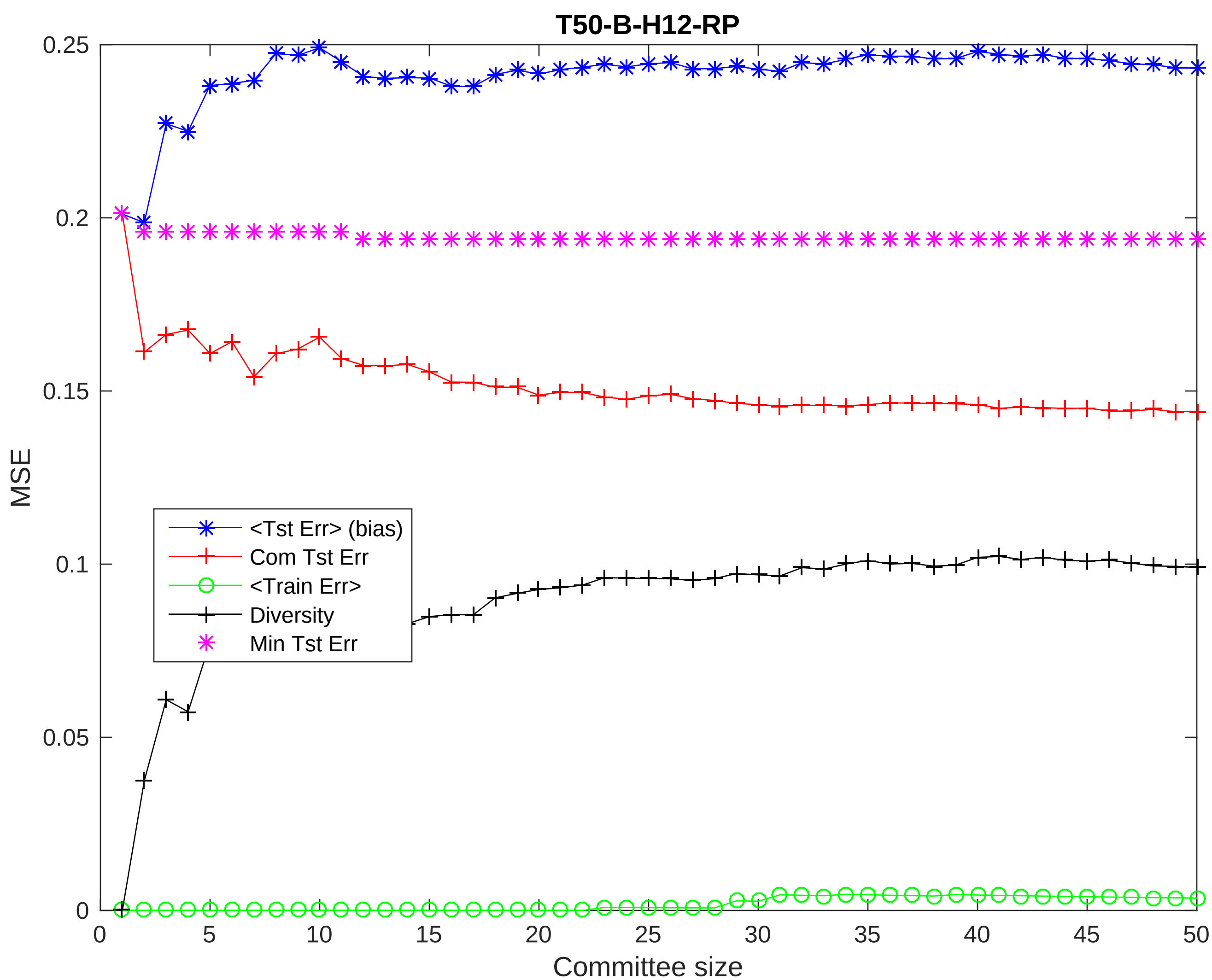


Figure 12 - Up to 20 hidden nodes. Again, for the 20 hidden MLPs we have very low training error, but the ensemble does not overfit.

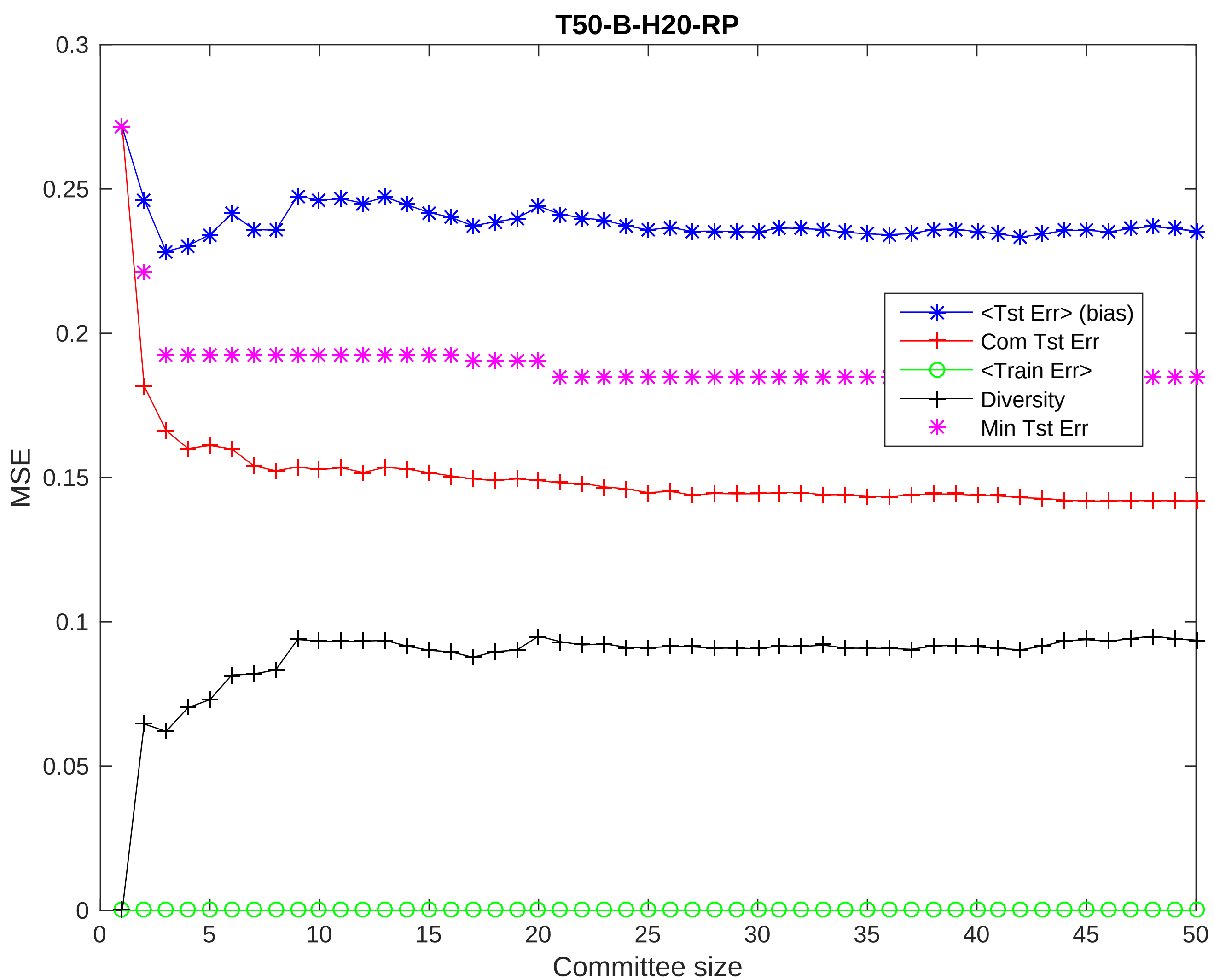


Figure 13 - Up to 20 hidden nodes. Again, for the 20 hidden MLPs we have very low training error, but the ensemble does not overfit.

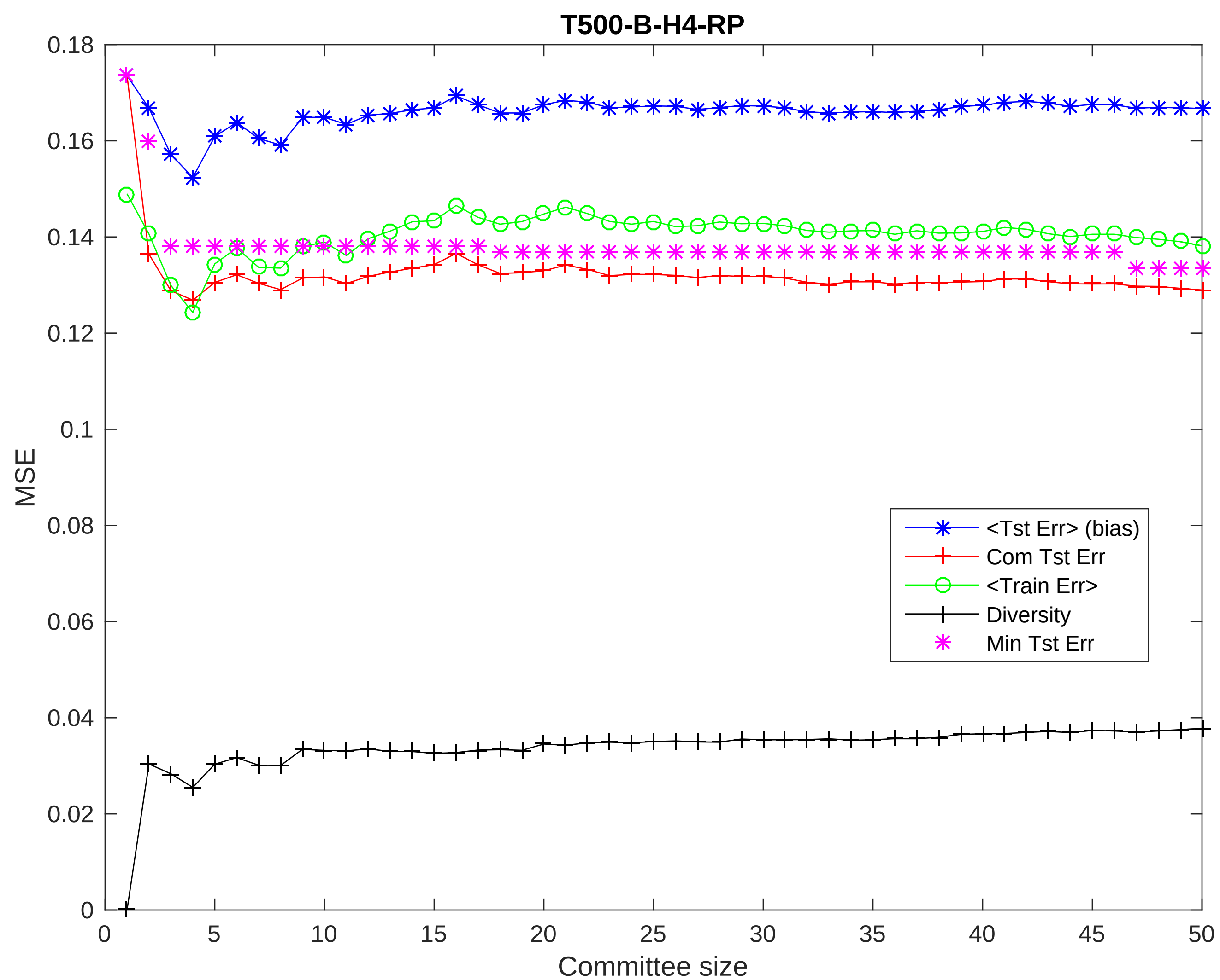


Figure 14 - Same as before, but with a larger training dataset. The ensemble effect is now less visible.

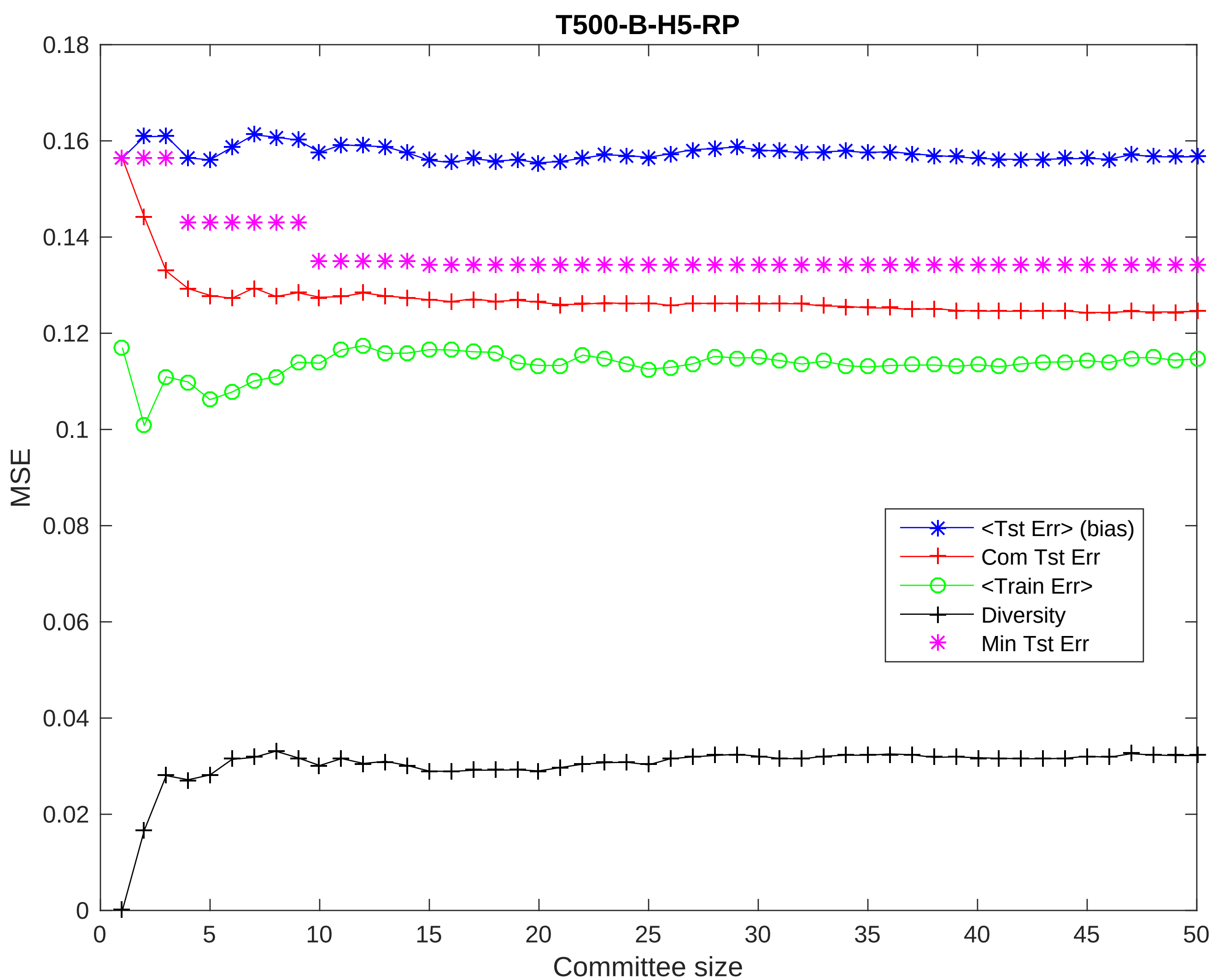


Figure 15 - Same as before, but with a larger training dataset. The ensemble effect is now less visible.

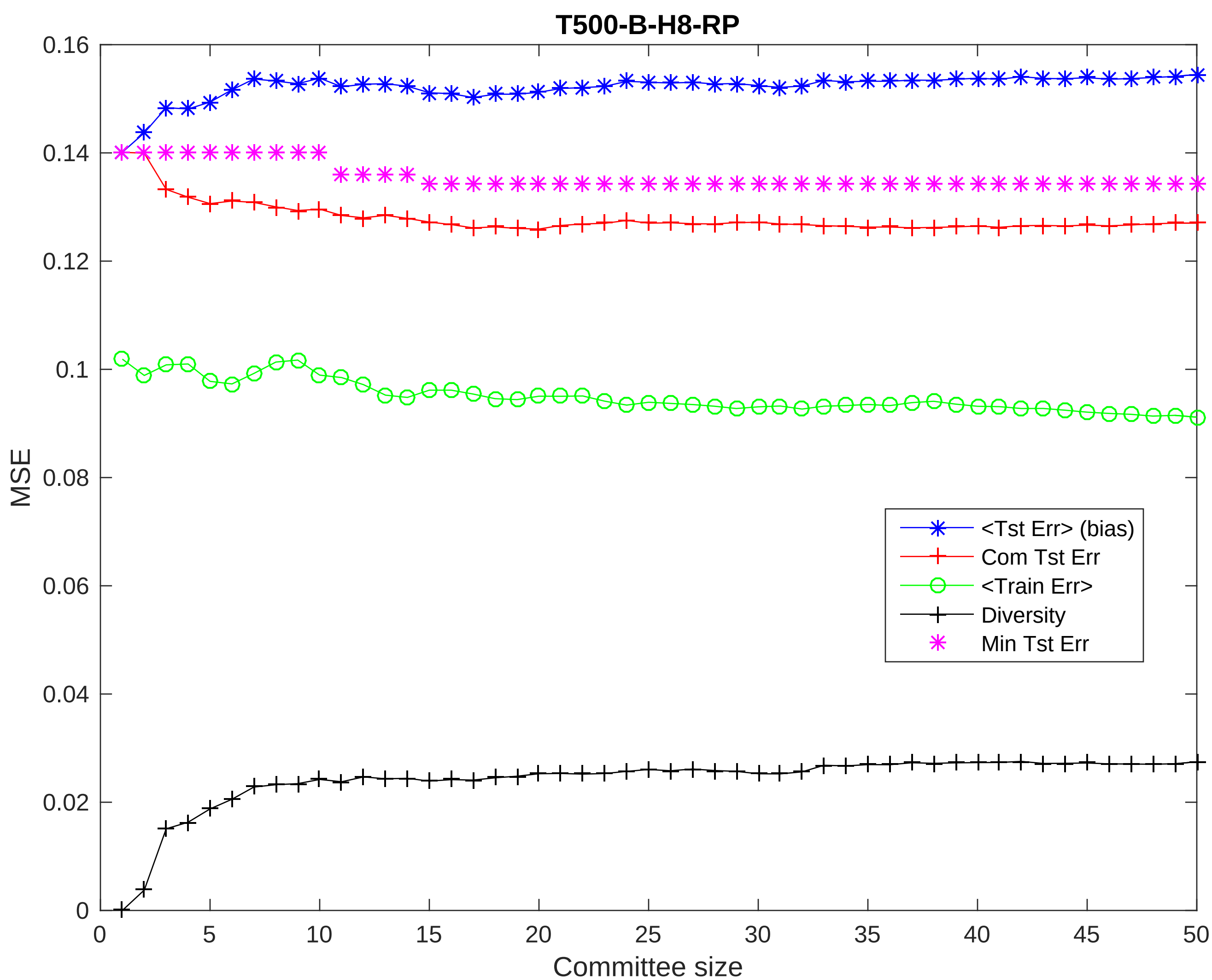


Figure 16 - Same as before, but with a larger training dataset. The ensemble effect is now less visible.

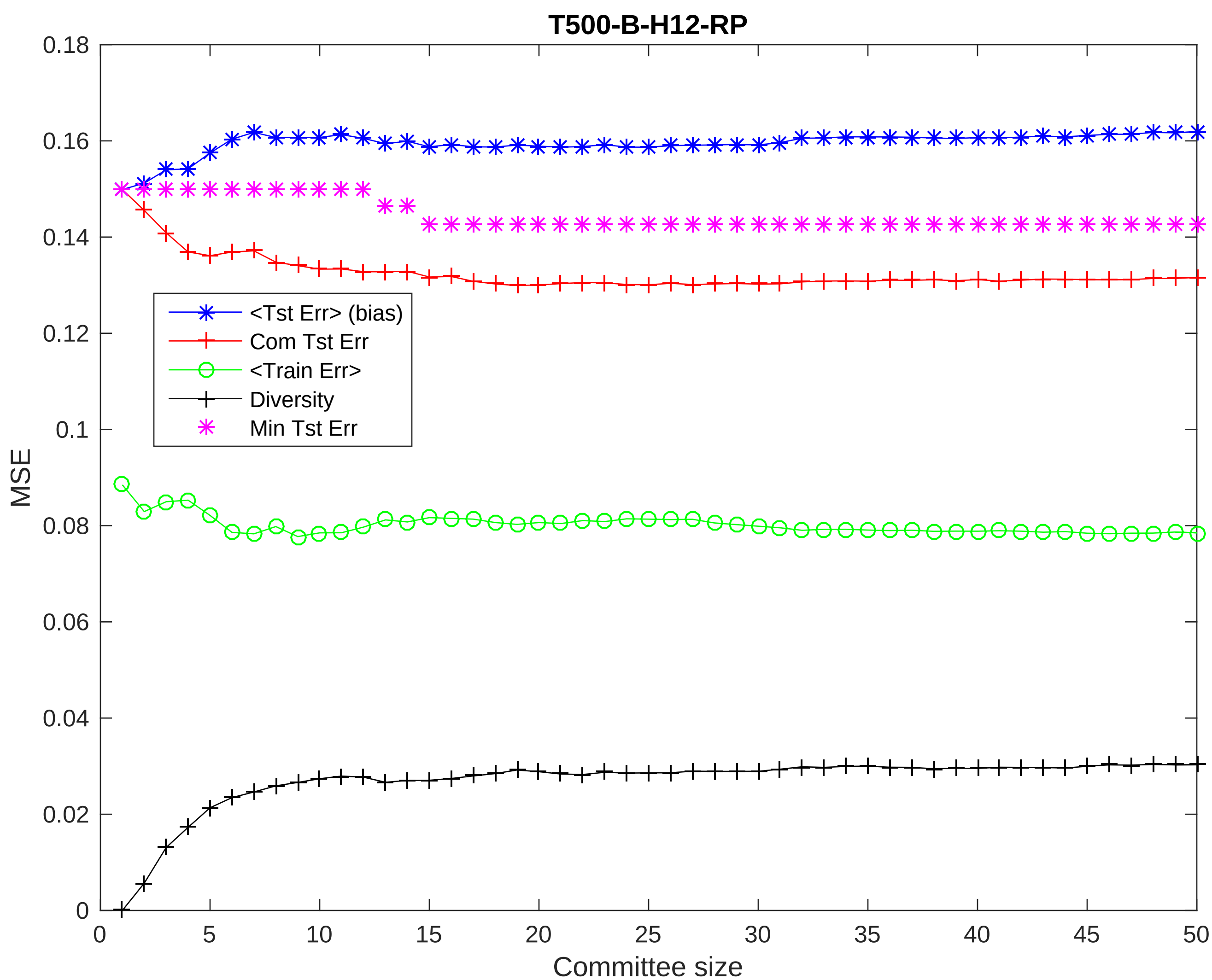


Figure 17 - Same as before, but with a larger training dataset. The ensemble effect is now less visible.

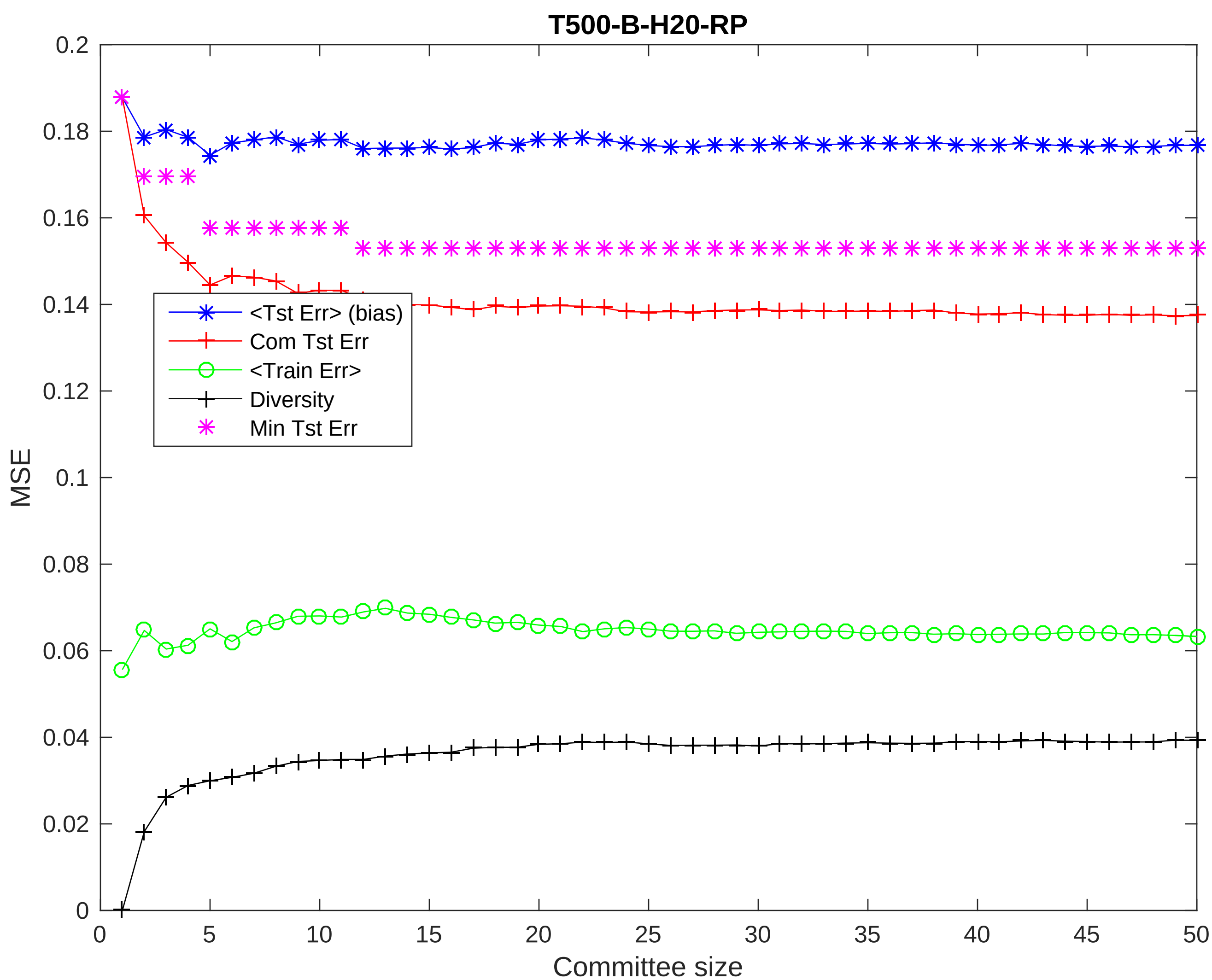


Figure 18 - Same as before, but with a larger training dataset. The ensemble effect is now less visible.