Sampling the initial weights

Let me visually explain about how to select the initial weights in the gray-box approach. The objective function J defined in the gray-box approach is non-convex and very complex, and thus, we cannot guarantee to obtain the global minimum by Levenberg-Marquardt algorithm (Figure 1). The approach that is described in the Gray-Box paper is to sample 5,000 sets of initial parameter values of Rs and Cs uniformly within the predefined range, select the best set that achieves the minimum J, and apply Levenberg-Marquardt algorithm to further optimize. This approach might work if you sample very densely. However, sampling only 5,000 as described in the paper is very sparse if you consider the high dimensionality of the state space (i.e., 8 dimensions in our model). For each dimension, this number corresponds to $\sqrt[8]{5000} \approx 2.9$, which means that for each parameter of Rs and Cs, we select only 3 samples of initial values within the range.

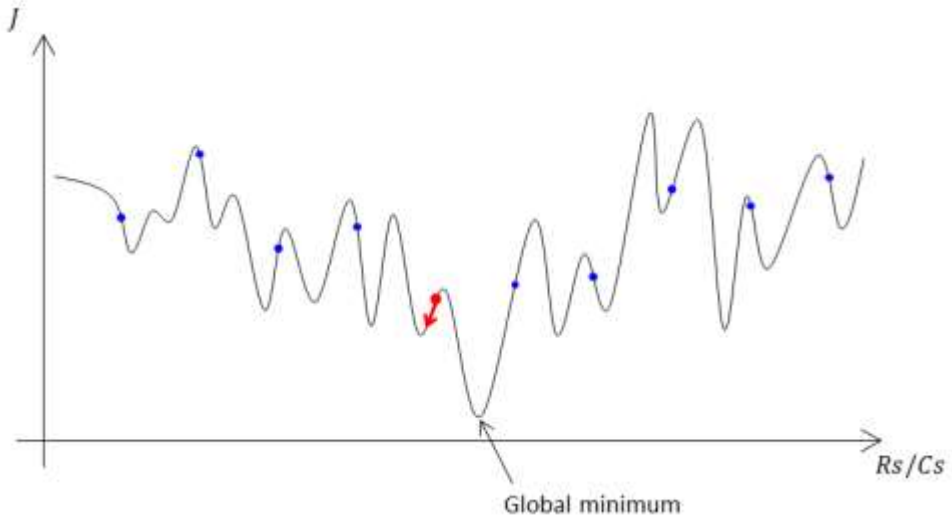

Figure 1: A simplified model of the objective function in the state space. *Once initial parameter values (blue dots) are sampled uniformly in the state space of Rs and Cs, the initial parameter values that achieve the minimum cost (red dot) are selected and Levenberg-Marquardt algorithm is applied to optimize the parameter values (red arrow). Notice that this approach does not guarantee to achieve the global minimum unless the initial parameter values are densely sampled.*

Thoughts

We got only poor results with a large residual error so far by using almost exactly same approach as the gray-box. Unfortunately I have no idea what is the cause of this problem. It might be the issue with our model, the training data, my implementation, or the gray-box algorithm. As I described above, only sparsely sampled initial estimate of parameter values do not guarantee to achieve a good result. Thus, I would increase the samples for the initial estimate of Rs and Cs in order to eliminate this concern. I have already started this idea running the optimization, which has been running already a half day and is still running. I will let you know once I get some results.