Team 3 in Explore AI @ NYCU - Questions

1. Complete the code and try to find the best result by changing the optimizer and learning rate.

   Done.

2. Why CORONA is called unfolded network or model-driven network?

   CORONA is called an unfolded network because it constructs a network by stacking multiple layers that perform specific tasks. This method eliminates the need for explicit iterations and enables parallel processing, making it more efficient. CORONA can be called a model-driven network because its structure is designed to identify particular patterns in the data, guided by the model's assumptions or prior knowledge about the problem.

3. If we use an iterative algorithm, we may need to iterate at least 100 times to get the result, why do you think we only need 3 layers here?

   Each layer detects it's set of specific features, which together allow the model to capture the relevant patterns in data, and decompose it into low rank and sparse components. While more layers would generally be a good practice, it's important to not add too much of those, because it might result in overfitting, where the model just learns the specific features of the training data too much. This approach allows efficient parallel execution, and therefore, avoids the need for additional explicit iterations.

4. In the code, α (a hyperparameter in the loss function) is set to 0.5, what do you think will be the difference if it is changed to 0.2 or 0.8? why?

   The parameter α controls the balance between the reconstruction error of the low-rank component and the sparse component. When α is set to 0.5, both components are considered equally important. If α is decreased to 0.2, more emphasis is given to the sparse component, resulting in a more aggressive thresholding in the SFT step, making the sparse component even sparser. Increasing α to 0.8 gives greater importance to the low-rank component, prioritizing better reconstruction of the underlying structure.

5. In the code, there are three layers in the network, what difference will it make if it is changed to two layers? why? Please run a simulation to test your ideas.

   Each layer in CORONA makes the model more complicated. In our case we have 3 layers of size 5 kernels, which is represented by the kernel_list = [5]*3 variable. In our simulation, reducing the number of layers led to a decrease in the model's ability to capture complex patterns or result in a less effective decomposition of the input data.