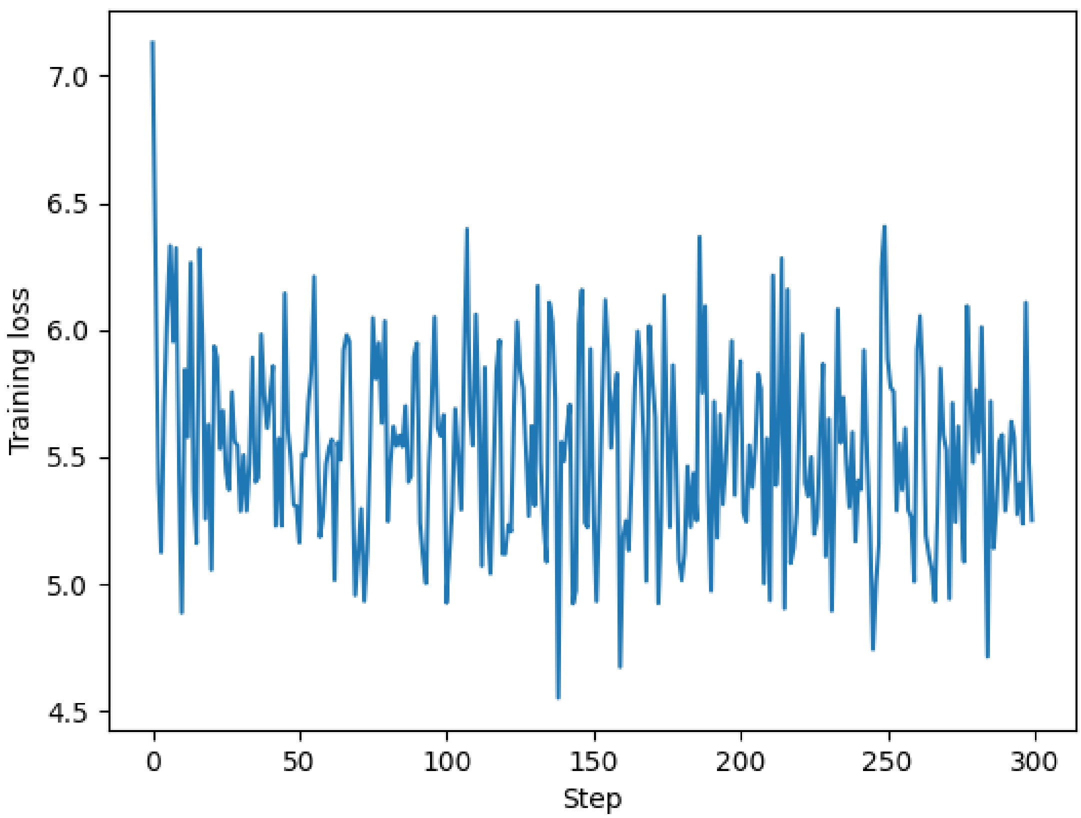
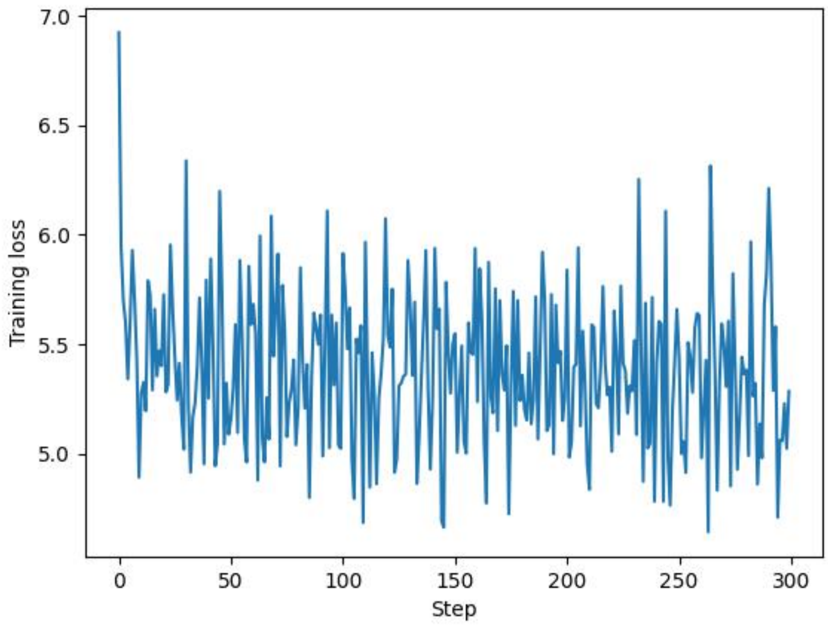
**2.3**

Number of training steps: 300 ($E = 30, Q = 2000, Q_t = 200$)

Training loss plot:

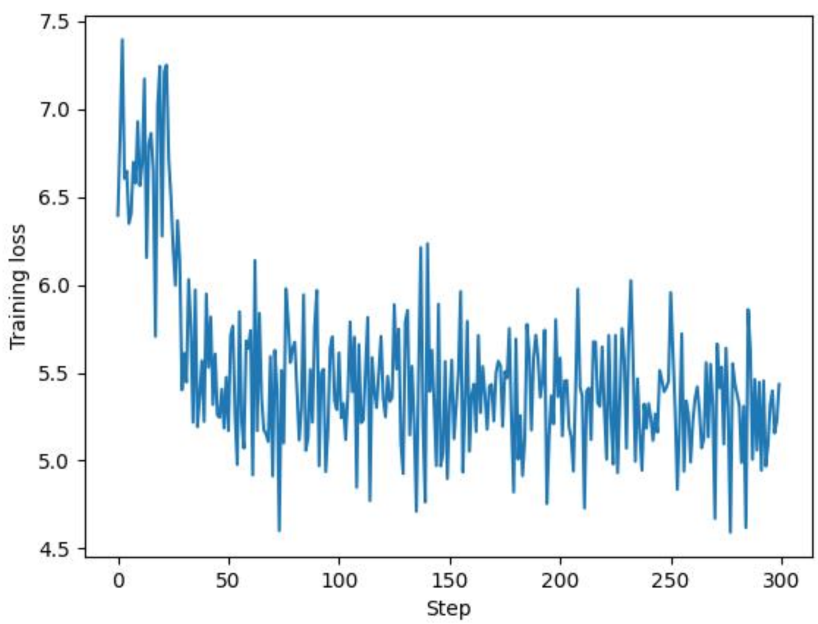


**3.1**



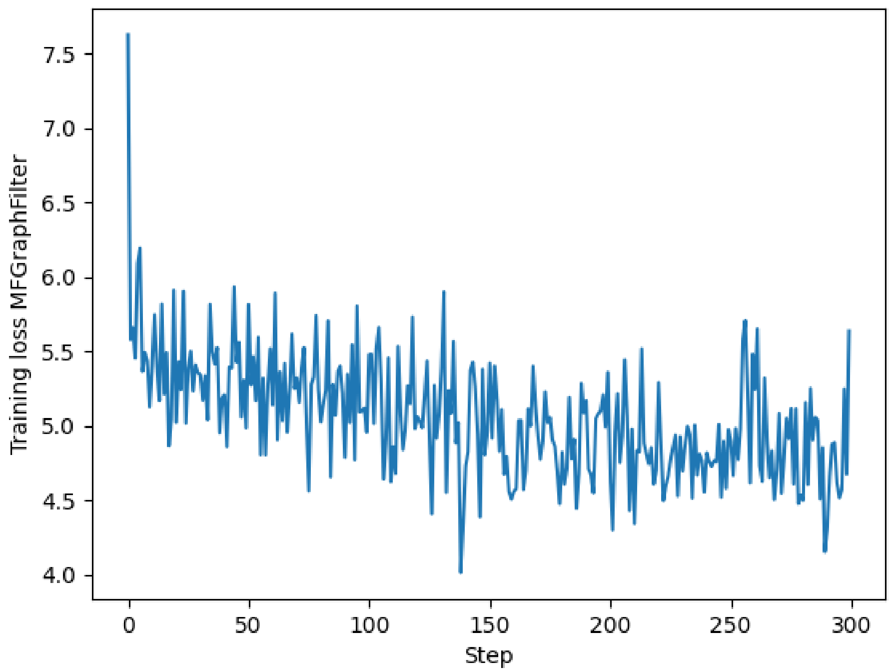
7.0
6.5
6.0
5.5
5.0
Training loss
0
50
100
150
200
250
300
Step


**3.2**



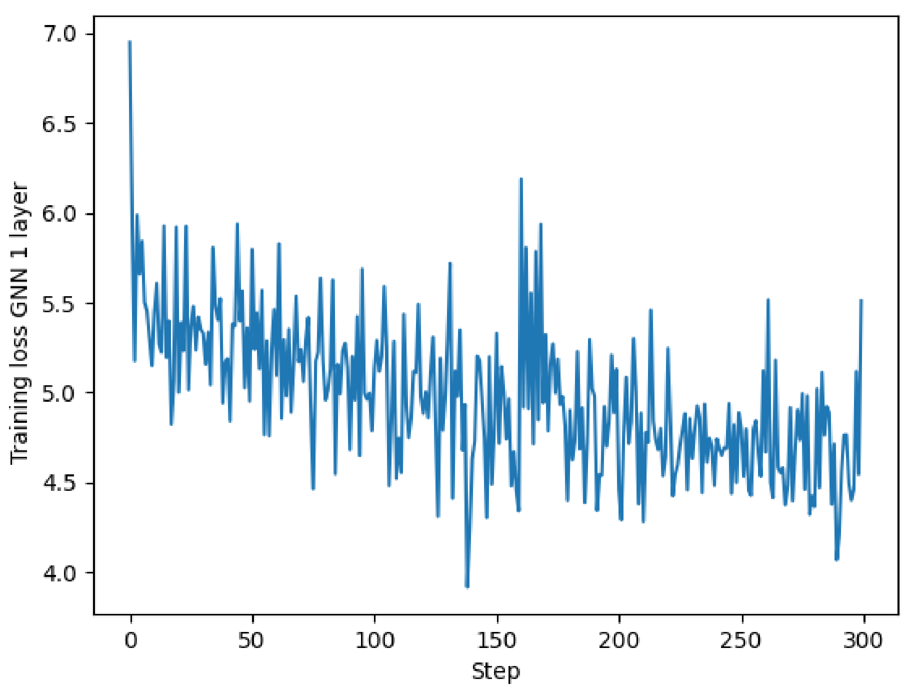
7.5
7.0
6.5
6.0
5.5
5.0
4.5
Training loss
0
50
100
150
200
250
300
Step

**4.1**



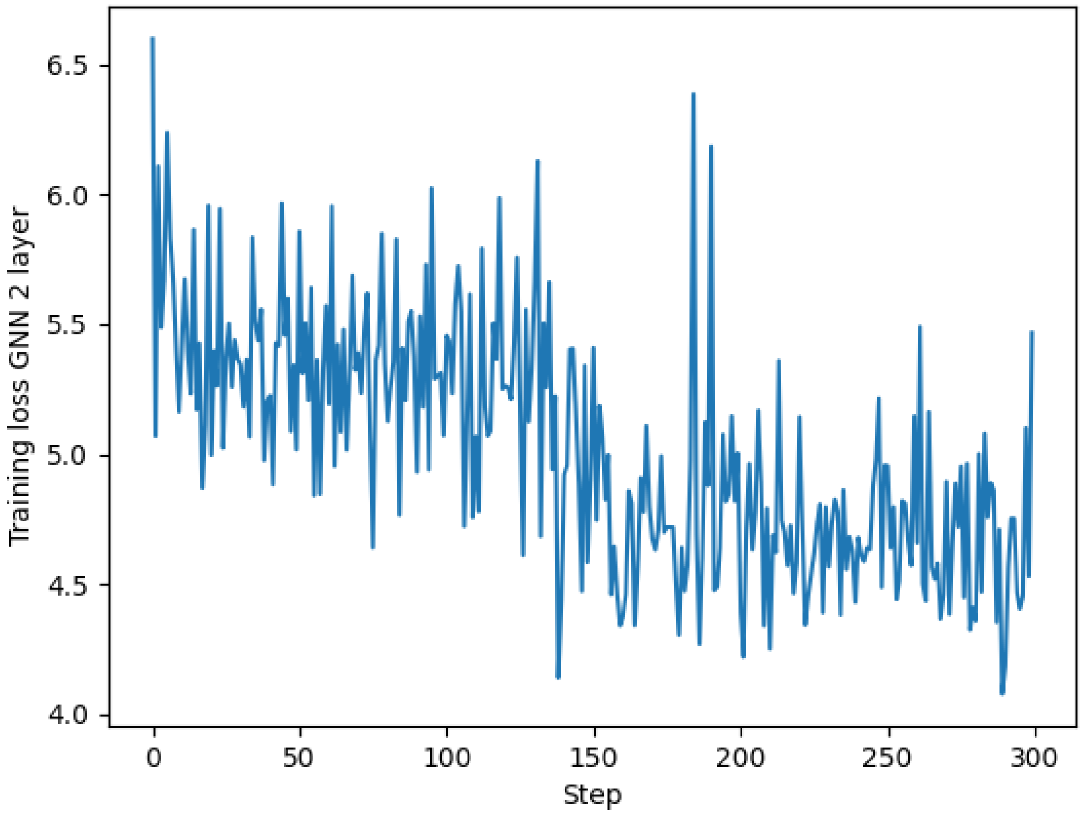
Training loss MFGraphFilter
Step
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
0
50
100
150
200
250
300


**4.2**



Training loss GNN 1 layer
Step
7.0
6.5
6.0
5.5
5.0
4.5
4.0
0
50
100
150
200
250
300

**4.3**



Training loss GNN 2 layer
Step
6.5
6.0
5.5
5.0
4.5
4.0
0
50
100
150
200
250
300

**5.1**

Number of parameters:

- LinearLayer: 2550
- Graph Filter: 321

Loss over test set:

- LinearLayer: 5.97
- Graph Filter: 5.49

Comment: The graph filter achieves lower error than a generic linear parametrization because it leverages the graph structure and permutation equivariance, leading to better generalization.

**5.2**

Number of parameters:

- FCNN: 2575
- GNN (2-layer): 425

Loss over test set:

- FCNN: 6.23
- GNN (2-layer): 5.17

Comment: The GNN outperforms the fully connected network by a significant margin. Incorporating the graph shift operator makes the model more expressive and better suited for the source localization task.

**5.3**

Number of parameters:

- Graph Filter: 321
- GNN (1-layer): 321
- GNN (2-layer): 425

Loss over test set:

- Graph Filter: 5.49
- GNN (1-layer): 5.38
- GNN (2-layer): 5.17

Comment: GNNs consistently outperform graph filters. The nonlinearities and layer stacking improve expressive power, resulting in lower test error.

**5.5**

Loss over test set:

- GNN (1-layer): 5.35
- GNN (2-layer): 5.15

**5.6**

Loss over test set:

- GNN (1-layer): 4.87
- GNN (2-layer): 4.96

Comment: Both GNNs transfer successfully to larger graphs, but the 1-layer GNN generalizes slightly better. The 2-layer model shows mild overfitting to the smaller training graph.

**5.7**

Loss over test set:

- GNN (1-layer): 5.36
- GNN (2-layer): 5.46

Comment: Transferability holds even for very large graphs. Again, the 1-layer GNN is more stable and yields lower error, suggesting that shallower GNNs may generalize more effectively across graph sizes.