Comparison Training Loss Over Iterations
300 samples, 6D points, 3 clusters, 3 sparse dimensions