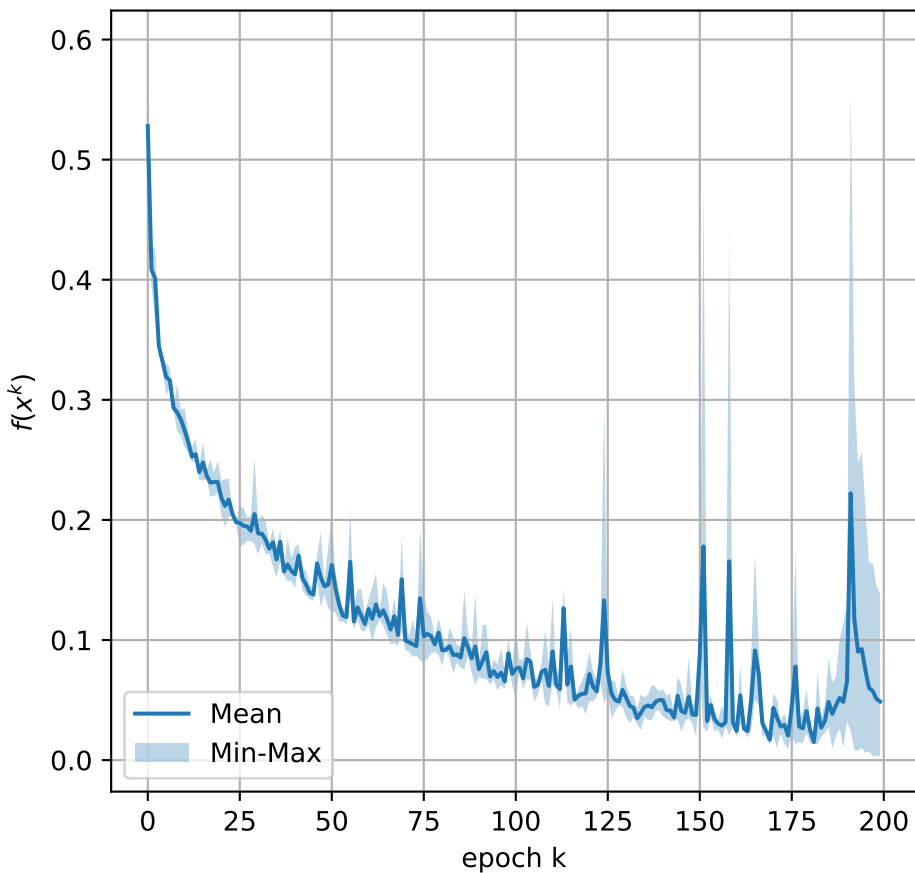
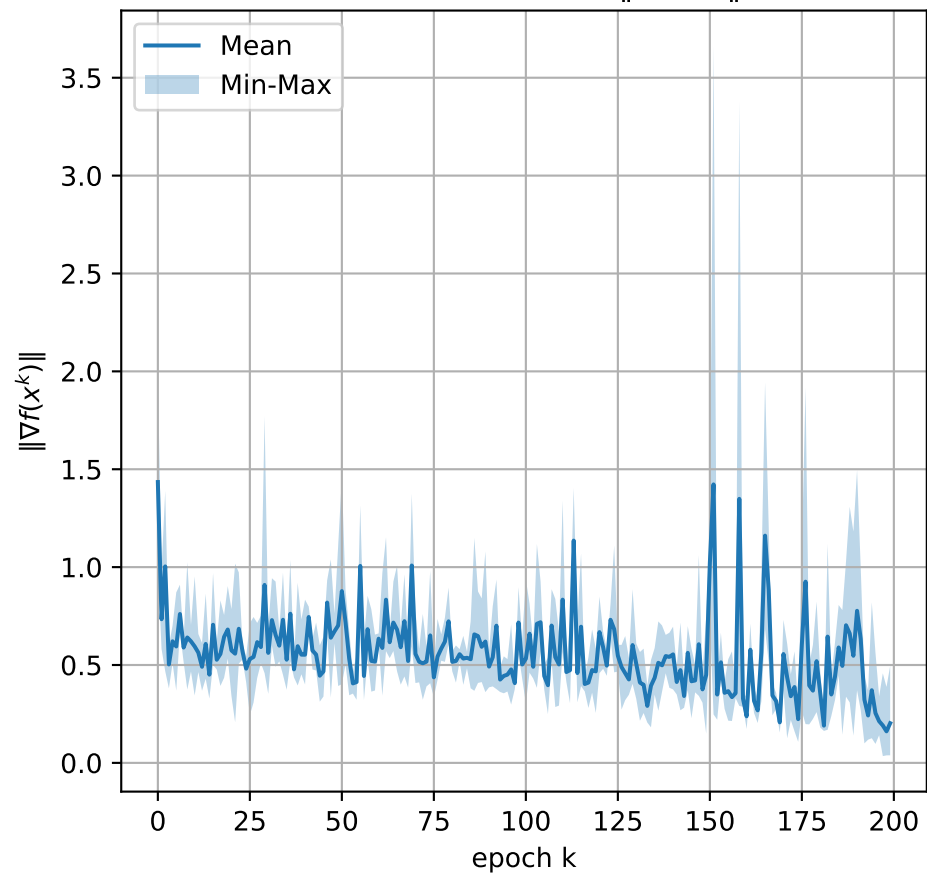
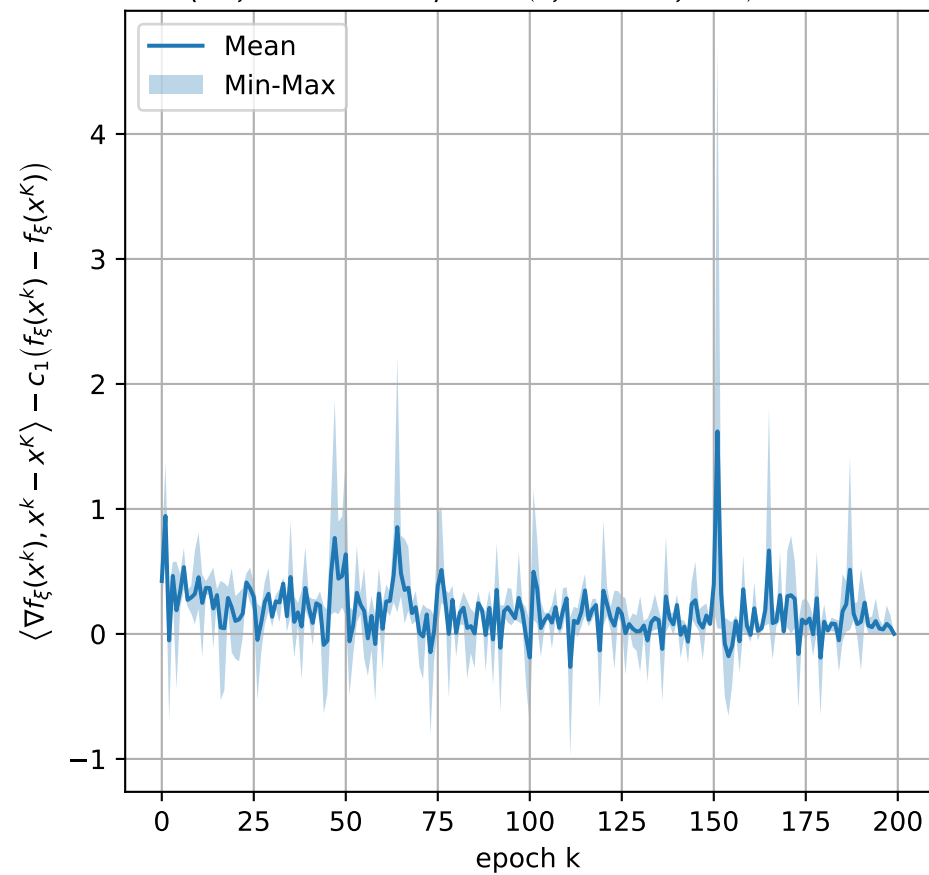


Full Loss $f(x^k)$ Full Gradient Norm $\|\nabla f(x^k)\|$  $\langle \nabla f_{\xi}(x^k), x^k - x^K \rangle - c_1(f_{\xi}(x^k) - f_{\xi}(x^K)), c_1 = 1$  $\langle \nabla f_{\xi}(x^k), x^k - x^K \rangle - c_1\|\nabla f_{\xi}(x^k)\|^2, c_1 = 0.01$ 