

Weight decay train 0e+00 1e-04

 10^{4}

1e-03 1e-02 1e-01

0.2

0.0

10°

10¹

10²

Optimization Steps

10³