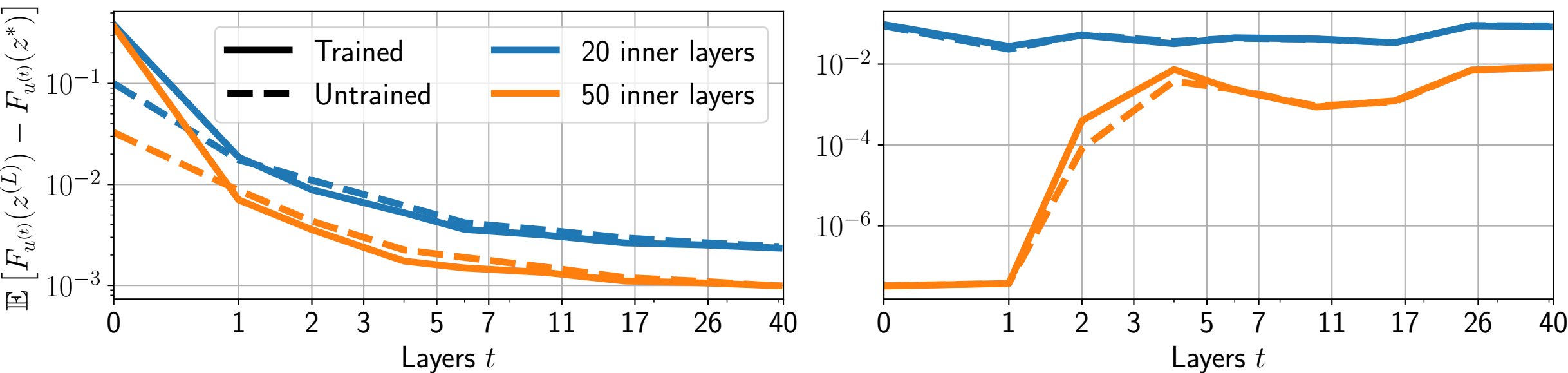
Trained
Untrained
20 inner layers
50 inner layers
$\mathbb{E}\left[F_{u^{(t)}}(z^{(L)}) - F_{u^{(t)}}(z^*)\right]$
10^{-1}
10^{-2}
10^{-3}
0
1
2
3
5
7
11
17
26
40
Layers t
10^{-2}
10^{-4}
10^{-6}
0
1
2
3
5
7
11
17
26
40
Layers t