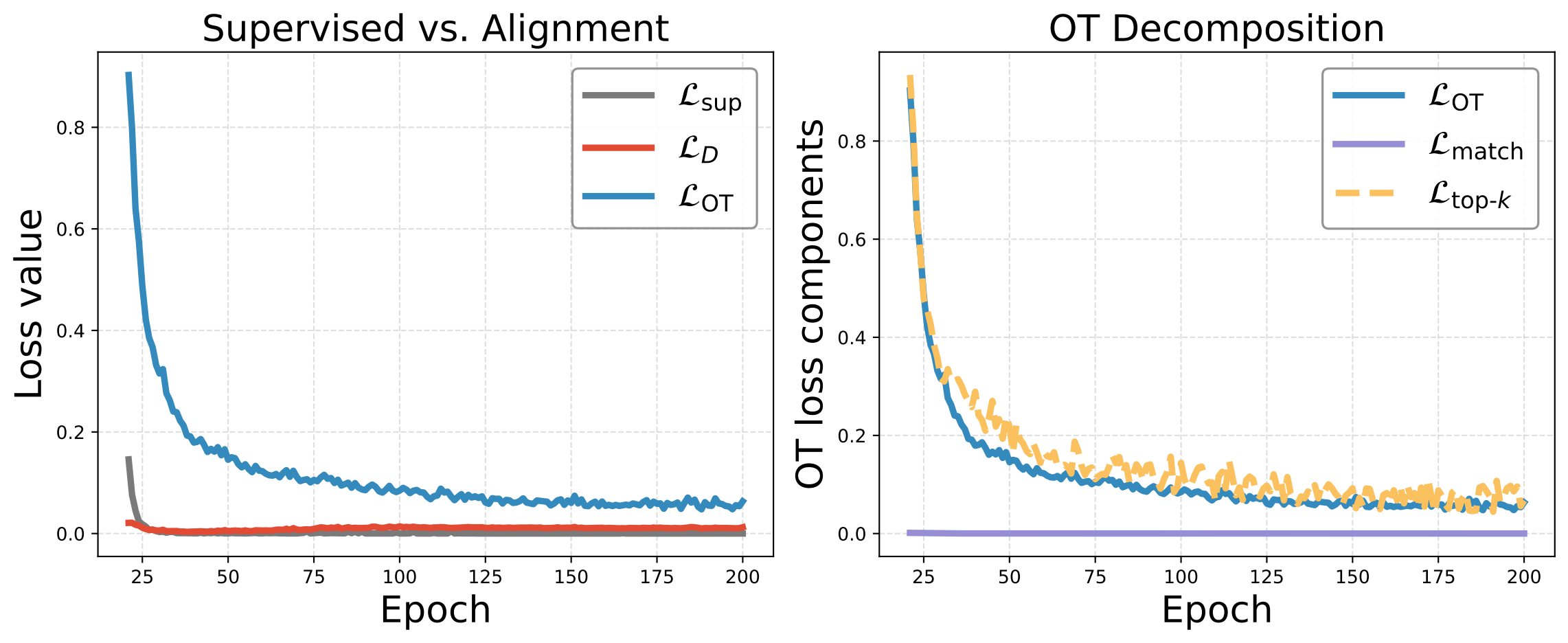
Supervised vs. Alignment
Loss value
Epoch
$\mathcal{L}_{\text{sup}}$
$\mathcal{L}_D$
$\mathcal{L}_{\text{OT}}$
0.0
0.2
0.4
0.6
0.8
25
50
75
100
125
150
175
200
OT Decomposition
OT loss components
Epoch
$\mathcal{L}_{\text{OT}}$
$\mathcal{L}_{\text{match}}$
$\mathcal{L}_{\text{top-}k}$