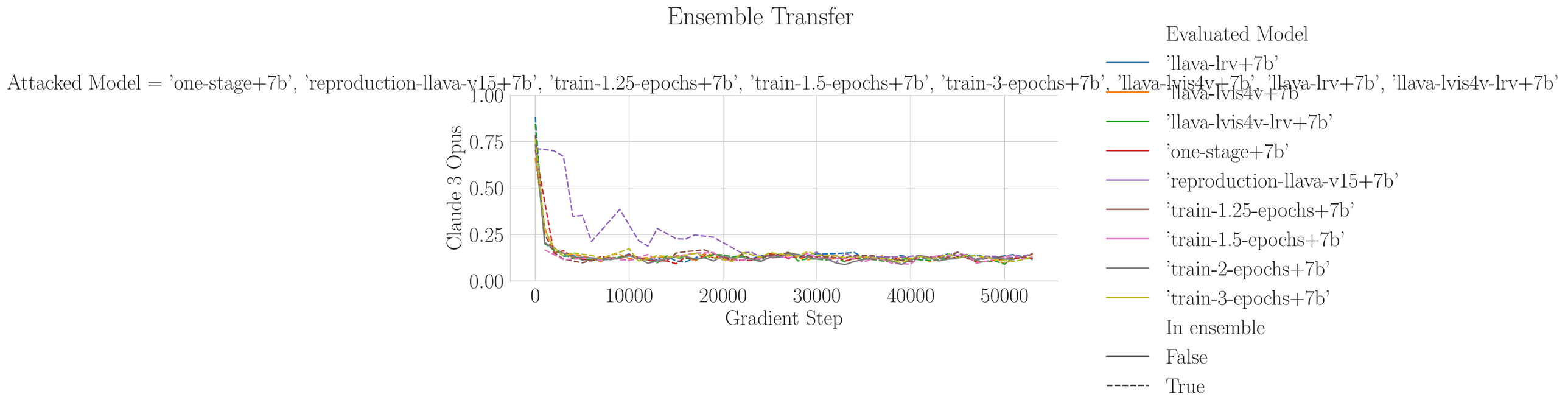
Ensemble Transfer
Attacked Model = 'one-stage+7b', 'reproduction-llava-v15+7b', 'train-1.25-epochs+7b', 'train-1.5-epochs+7b', 'train-3-epochs+7b', 'llava-lvis4v+7b', 'llava-lrv+7b', 'llava-lvis4v-lrv+7b'
Claude 3 Opus
1.00
0.75
0.50
0.25
0.00
0
10000
20000
30000
40000
50000
Gradient Step
Evaluated Model
'llava-lrv+7b'
'llava-lvis4v+7b'
'llava-lvis4v-lrv+7b'
'one-stage+7b'
'reproduction-llava-v15+7b'
'train-1.25-epochs+7b'
'train-1.5-epochs+7b'
'train-2-epochs+7b'
'train-3-epochs+7b'
In ensemble
False
True