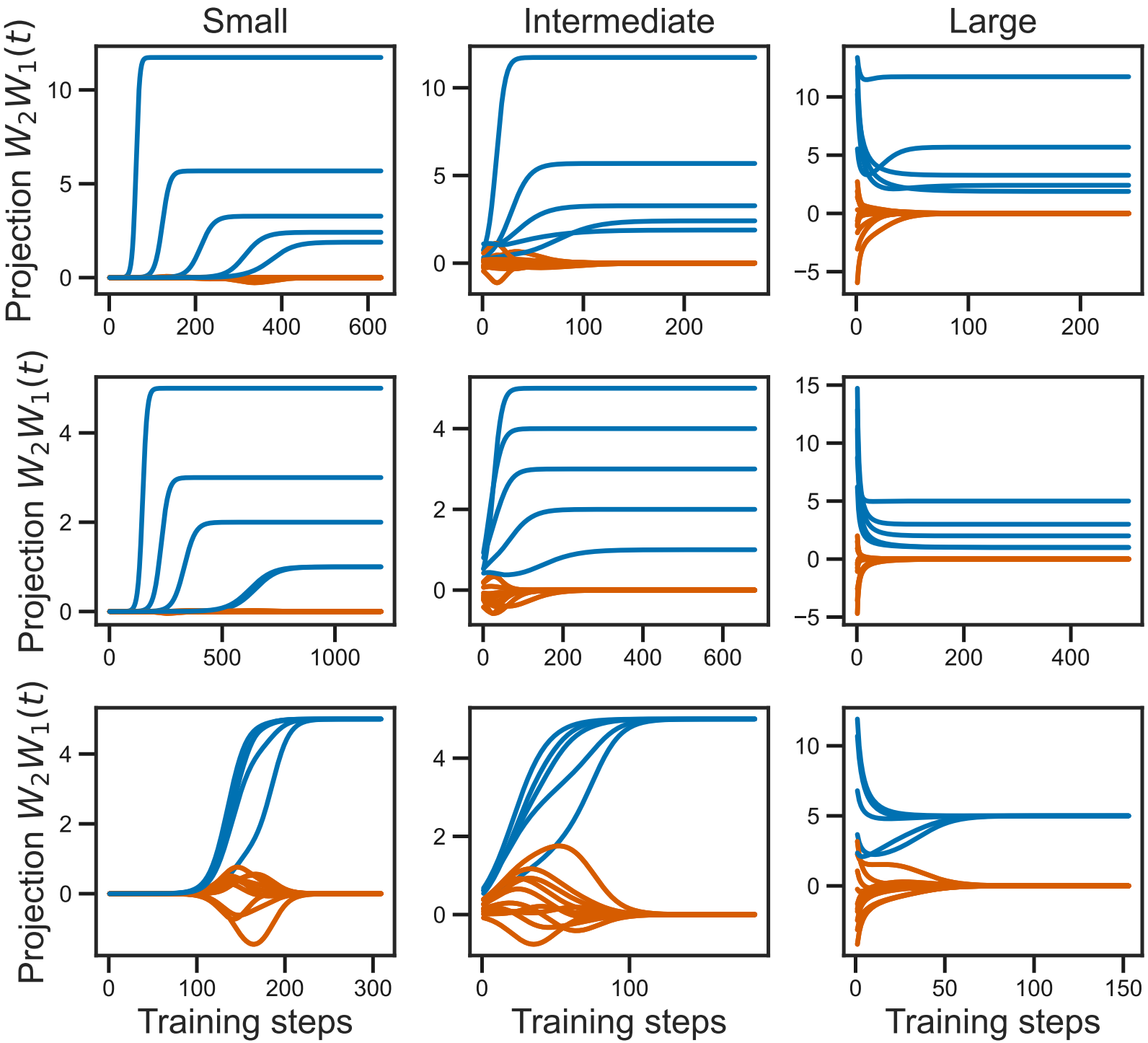
Small
Intermediate
Large
Projection $W_2W_1(t)$
Training steps