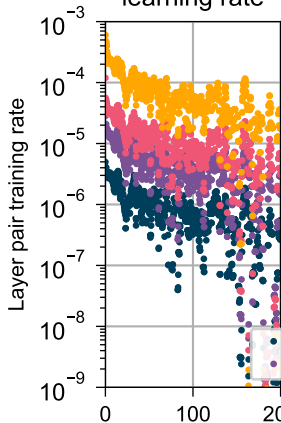
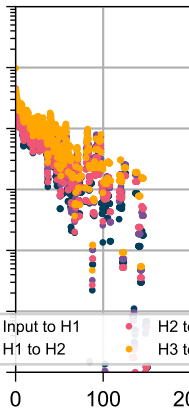


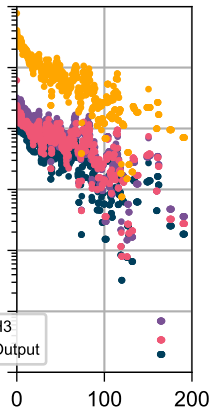
Layered, global
learning rate



Layered, per-layer
rates



Our topology,
 $p = 7.56\%$



Input to H1

H1 to H2

H2 to H3

H3 to Output

Epoch