Snapshot evolution: residual stream + BP grad over training
(top: with terminal LN — DFA explodes; bottom: no terminal LN — DFA still grows but BP grad does NOT collapse)