

BP-trained: protocol passes everywhere

ResMLP $d = 256$ (terminal LN)	✓	✓	✓	✓
ResMLP $d = 512$ (terminal LN)	✓	✓	✓	✓
ViT-Mini (cls + LN)	✓	✓	✓	✓
ResMLP $d = 256$ (no terminal LN)	✓	✓	✓	✓
CNN BatchNorm (no terminal LN)	✓	✓	✓	✓
	(a) scale	(b) $\ g\ $ floor	(c) drift	(d) frozen

DFA-trained: protocol verdict by architecture

ResMLP $d = 256$ (terminal LN)	WB	WB	✓	WB
ResMLP $d = 512$ (terminal LN)	WB	WB	✓	WB
ViT-Mini (cls + LN)	WB	WB	—	WB
ResMLP $d = 256$ (no terminal LN)	WB	✓	—	WB
CNN BatchNorm (no terminal LN)	WB	✓	—	—
	(a) scale	(b) $\ g\ $ floor	(c) drift	(d) frozen

Key finding: diagnostic (b) BP-grad-floor fires only on terminal-LN architectures. Of the 5 architectures audited, (b) is restricted to the with-terminal-LN family. Cells marked — are not applicable for that architecture (no matched frozen-blocks baseline / no stability run).