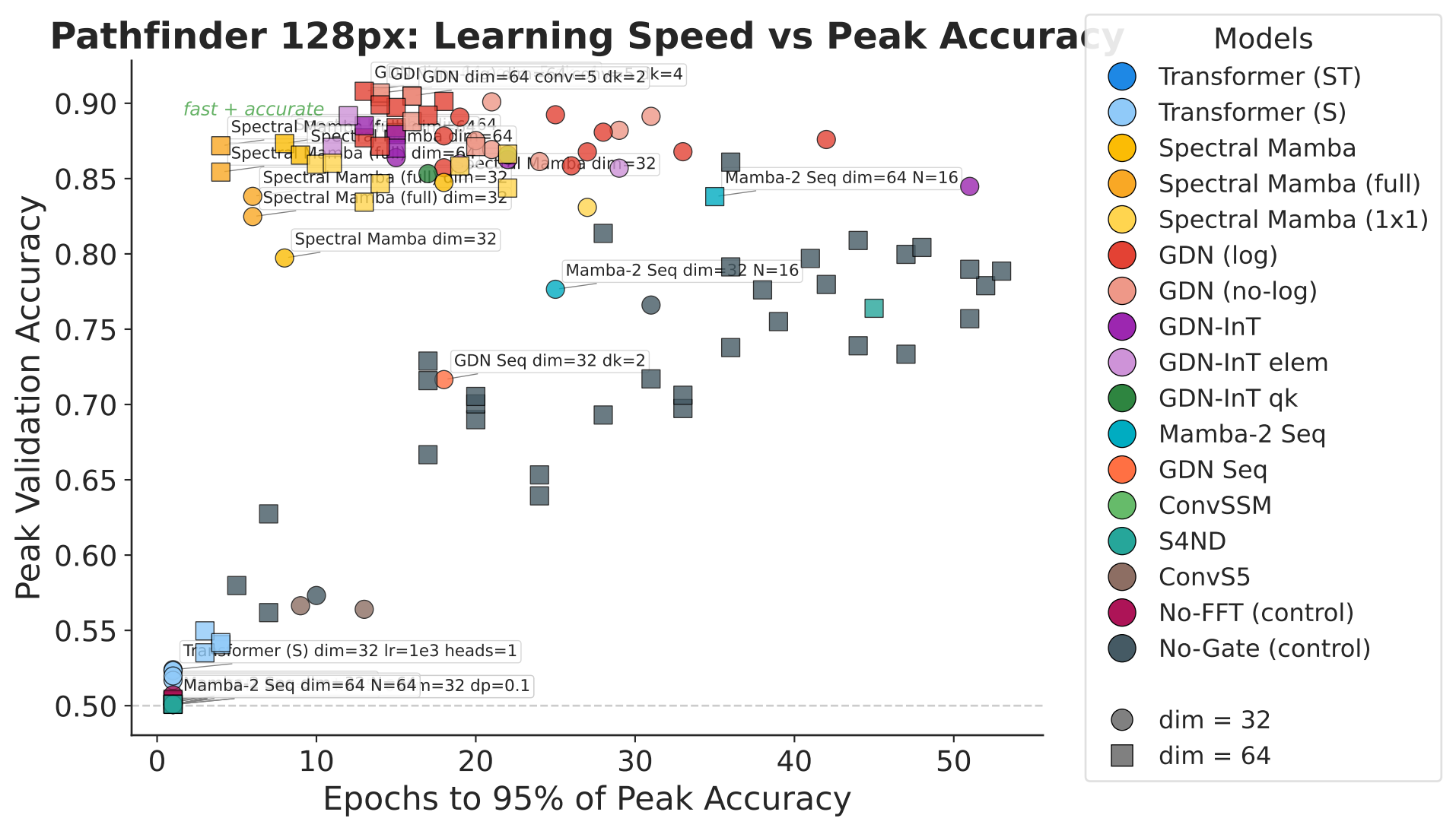
Pathfinder 128px: Learning Speed vs Peak Accuracy
Peak Validation Accuracy
Epochs to 95% of Peak Accuracy
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0
10
20
30
40
50
fast + accurate
GDN dim=64 conv=5 dk=2
Spectral Mamba dim=64
Spectral Mamba dim=32
Spectral Mamba (full) dim=32
Mamba-2 Seq dim=64 N=16
Spectral Mamba dim=32
Mamba-2 Seq dim=32 N=16
GDN Seq dim=32 dk=2
Transformer (S) dim=32 lr=1e3 heads=1
Mamba-2 Seq dim=64 N=64
Models
Transformer (ST)
Transformer (S)
Spectral Mamba
Spectral Mamba (full)
Spectral Mamba (1x1)
GDN (log)
GDN (no-log)
GDN-InT
GDN-InT elem
GDN-InT qk
Mamba-2 Seq
GDN Seq
ConvSSM
S4ND
ConvS5
No-FFT (control)
No-Gate (control)
dim = 32
dim = 64