

Figure 1a: Training Stability Comparison
Standard KD vs Decoupled KD

