



Figure 1: The relationship between initial input activation loss and post-optimization output loss. By randomly sampling initial activations for ViT-Small and Swin-Small's last block and optimizing with a 4/4 bit quantization setup, we observe a positive correlation between the magnitude of initial loss and final optimization loss.