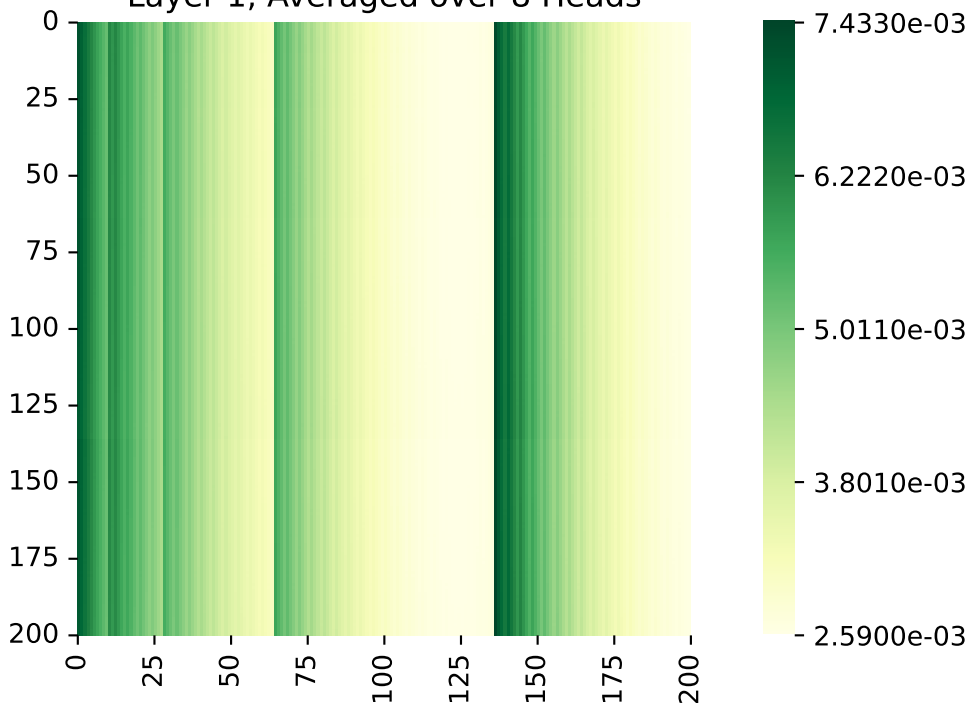


Layer 1, Averaged over 8 Heads



Layer 2, Averaged over 8 Heads

