

Per-layer average log learning rate vs Iteration

