

Layer:model.layers.0.attention.wo.weight alpha=8.73, ks\_distance=0.08

