

Component = Swiglu W_2 (First Six) – Scale = tiny



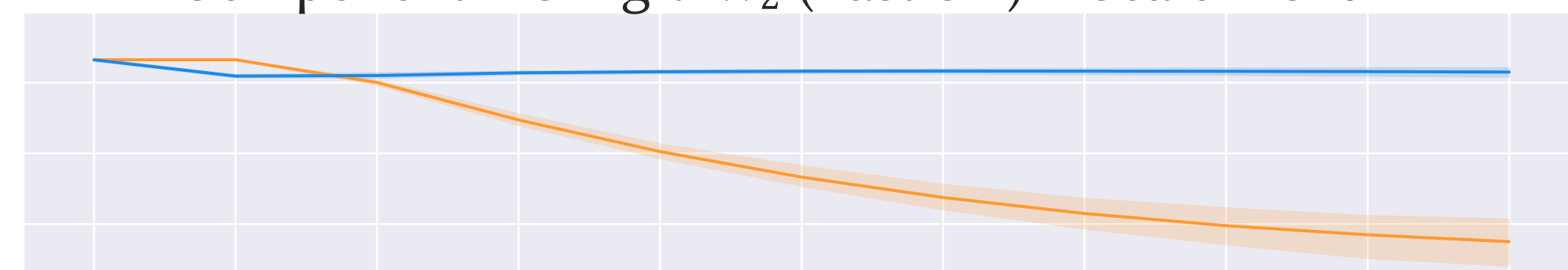
Component = Swiglu W_2 (First Six) – Scale = small



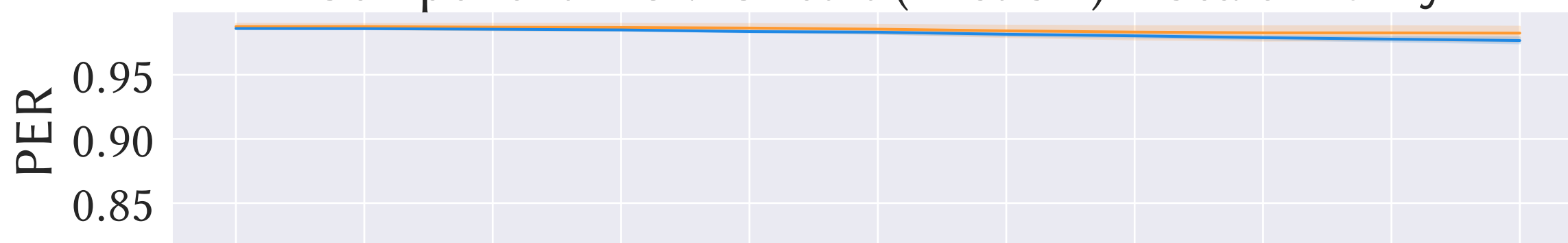
Component = Swiglu W_2 (Last Six) – Scale = tiny



Component = Swiglu W_2 (Last Six) – Scale = small



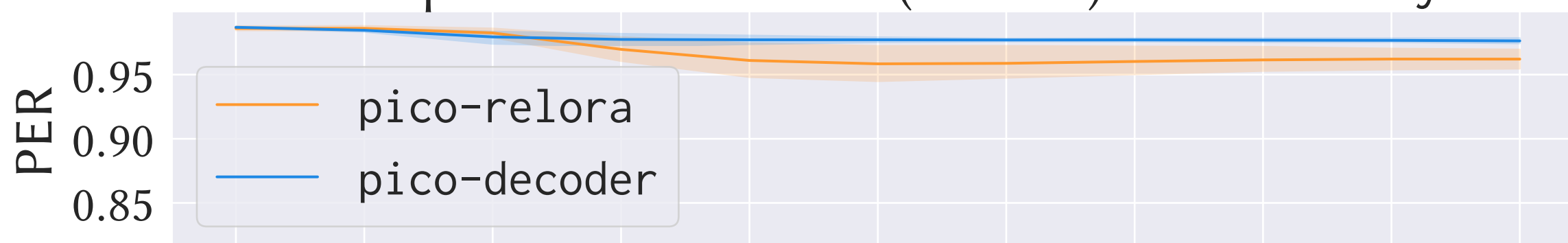
Component = OV Circuit (First Six) – Scale = tiny



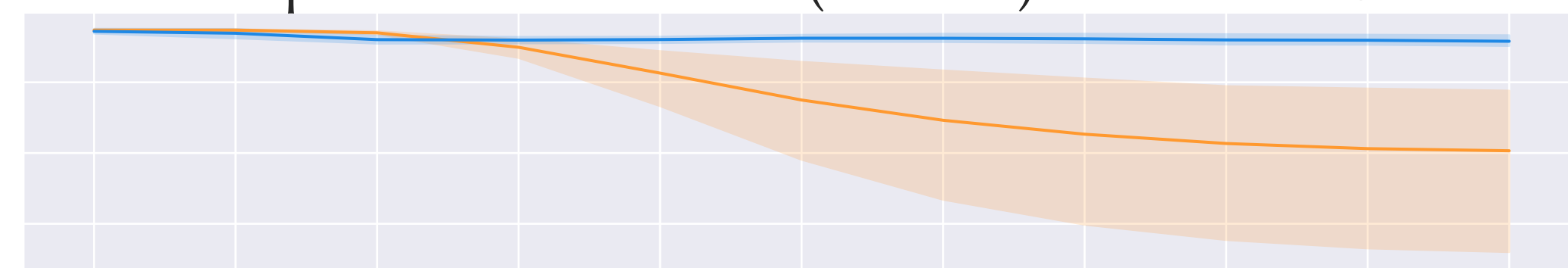
Component = OV Circuit (First Six) – Scale = small



Component = OV Circuit (Last Six) – Scale = tiny



Component = OV Circuit (Last Six) – Scale = small



— pico-relora
— pico-decoder

0 2 4 6 8 10 12 14 16 18 20

Checkpoint step / 1000

0 2 4 6 8 10 12 14 16 18 20

Checkpoint step / 1000