Legend: constant, β=1.00, warmup=5.0 — cosine, β=1.00, warmup=5.0 — linear, β=1.00, warmup=5.0 — sigmoid, β=1.00, warmup=5.0

Loss

MSE

KL Divergence

β