Training Loss Comparison: Standard vs Cold vs Tempered VI Standard (Standard) Cold (Cold T=0.8) 10<sup>6</sup> · Tempered (Tempered  $\lambda$ =0.8) 10<sup>5</sup> · Negative ELBO (log scale) 10<sup>3</sup> ·  $10^{2}$ 2500 10000 20000 0 5000 7500 12500 15000 17500 Epoch