## Critical training compute $T_{crit}$ (FP16 FLOP) V100 SXM 10<sup>38</sup> A100 SXM H100 SXM $10^{36}$ $10^{34}$ $10^{32}$ 0.2 -0.2-0.10.0 0.1 Exponent difference ( $\alpha = \alpha_b - \alpha_L$ )