Total execution time as a function of validation loss 10¹ 6×10^{0} Execution time (hours) 10¹ 4×10^{0} 3×10^{0} 1.9 1.8 2.0 1.7 2.1 2.2 10⁰ B = 8B = 16B = 32B = 64 10^{-1} 5 6 Validation loss