Convergence of the training as function of the execution time B = 8B = 1610 B = 32B = 648 Training loss 2 30 10 20 40 50 Execution time (hours)