JSB Chorales, clip = 1.0, difference from no clipping 0.0025 explicit implicit 0.0020 -Training set loss 0.0015 -0.0010 -0.0005 -0.0000 -0.0005-0.0010100 200 300 400 500 Learning rate