

MSE

$1.08 \times 10^{-2}$

$1.06 \times 10^{-2}$

$1.04 \times 10^{-2}$

$1.02 \times 10^{-2}$

$10^{-2}$

- Learning schedule
- With moment
- Adagrad
- RMSprop
- Adam

$\eta_0 = 0.1$

Batch size = 20

200

400

600

800

1000

Epochs

