

A bio-inspired model towards vocal gesture learning in songbird

Silvia Pagliarini^{1,2,3}, Xavier Hinaut^{1,2,3}, Arthur Leblois³



1) Mnemosyne, Inria Bordeaux Sud-Ouest
2) LaBRI, UMR 5800, CNRS
3) IMN, UMR 5293, CNRS
Université de Bordeaux, France.

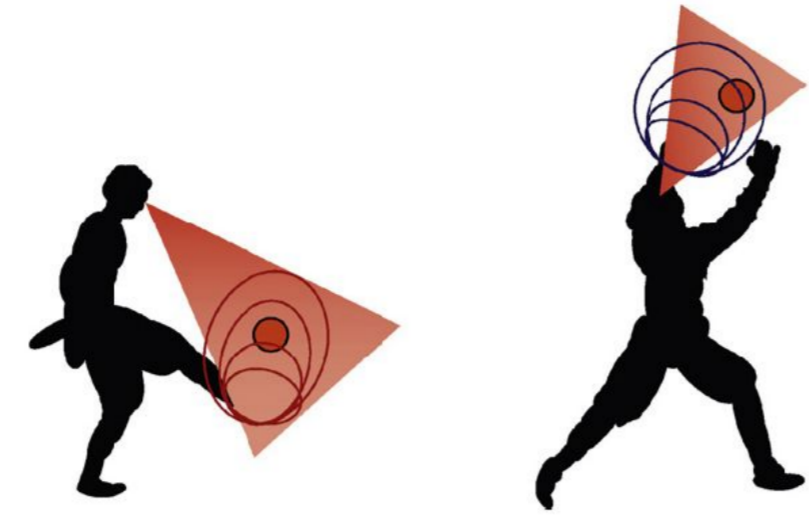


INTRODUCTION

Sensorimotor learning: control problem which maps a sensory input into a motor output.

Imitation: learning from a tutor using a feedback guided error.

Inverse model: the aim is to transform a sensory stimulus into the corresponding motor command.



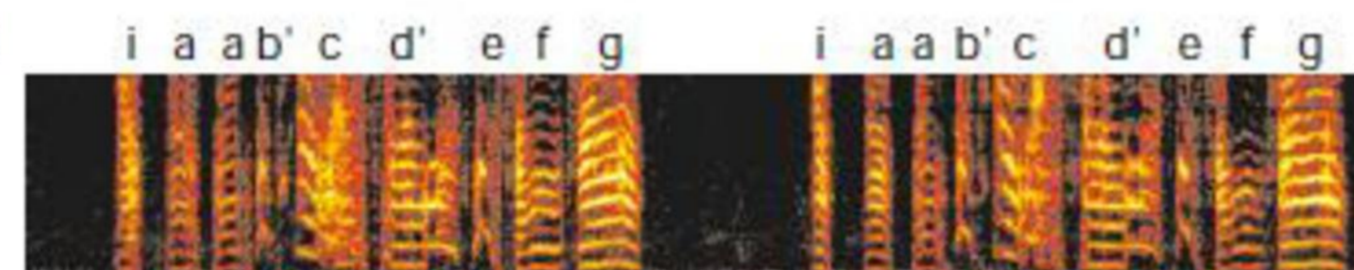
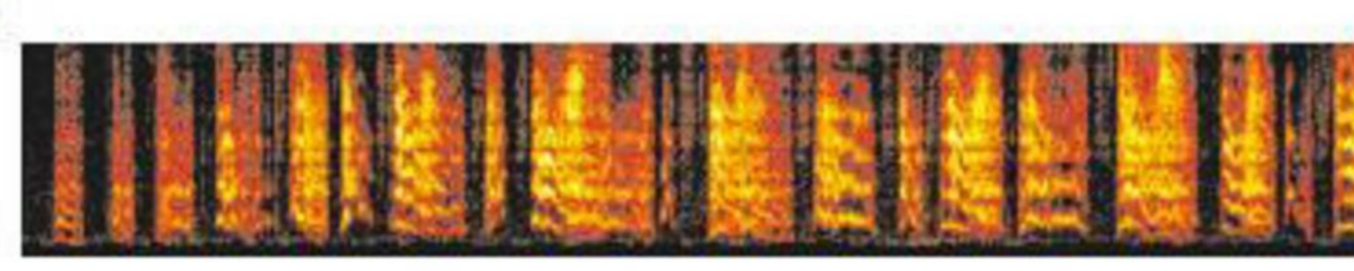
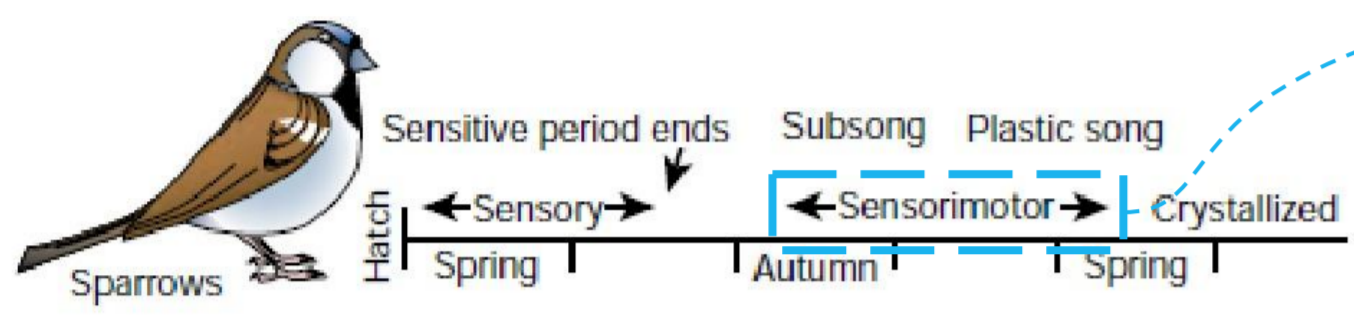
Da Cunha et al., 2010

AIM: implementation of an inverse model which describes the sensorimotor phase of learning in birdsong.

SONG LEARNING IN BIRDS

• Sensorimotor phase is characterized by babbling.

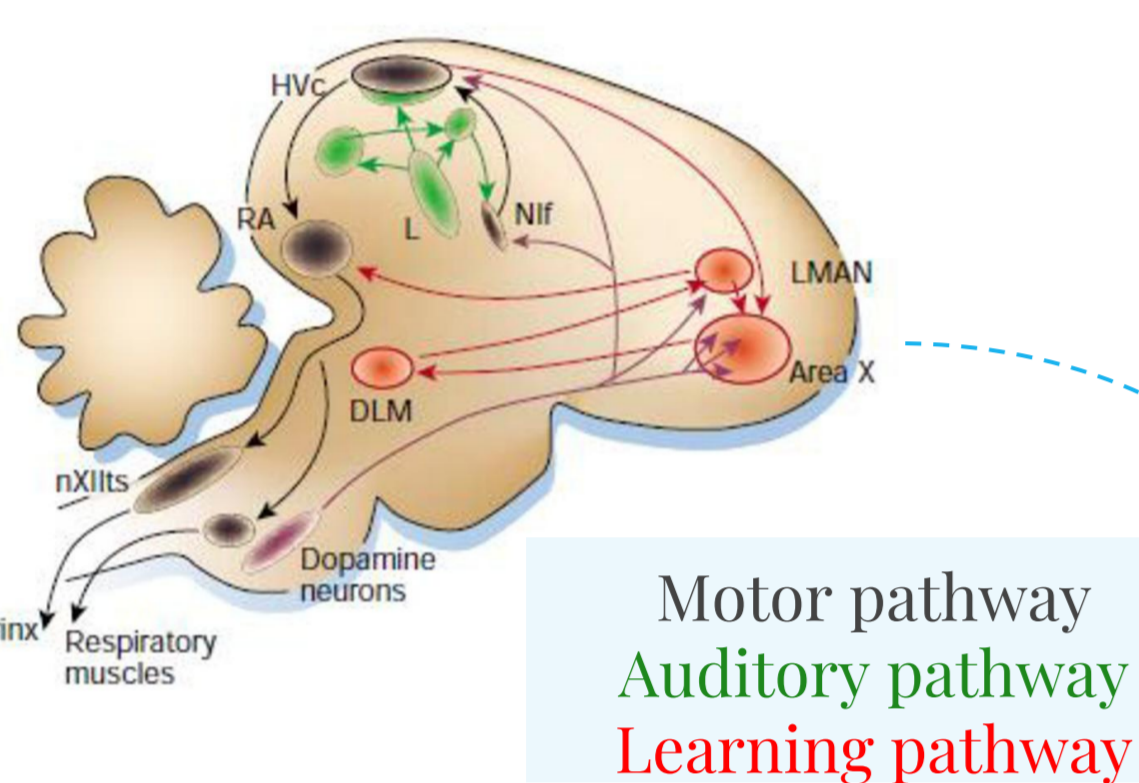
Learning by imitation



a. Adult zebra finch song (TUTOR).
b. Juvenile zebra finch song at an early stage of learning.
c. Song close to crystallization.

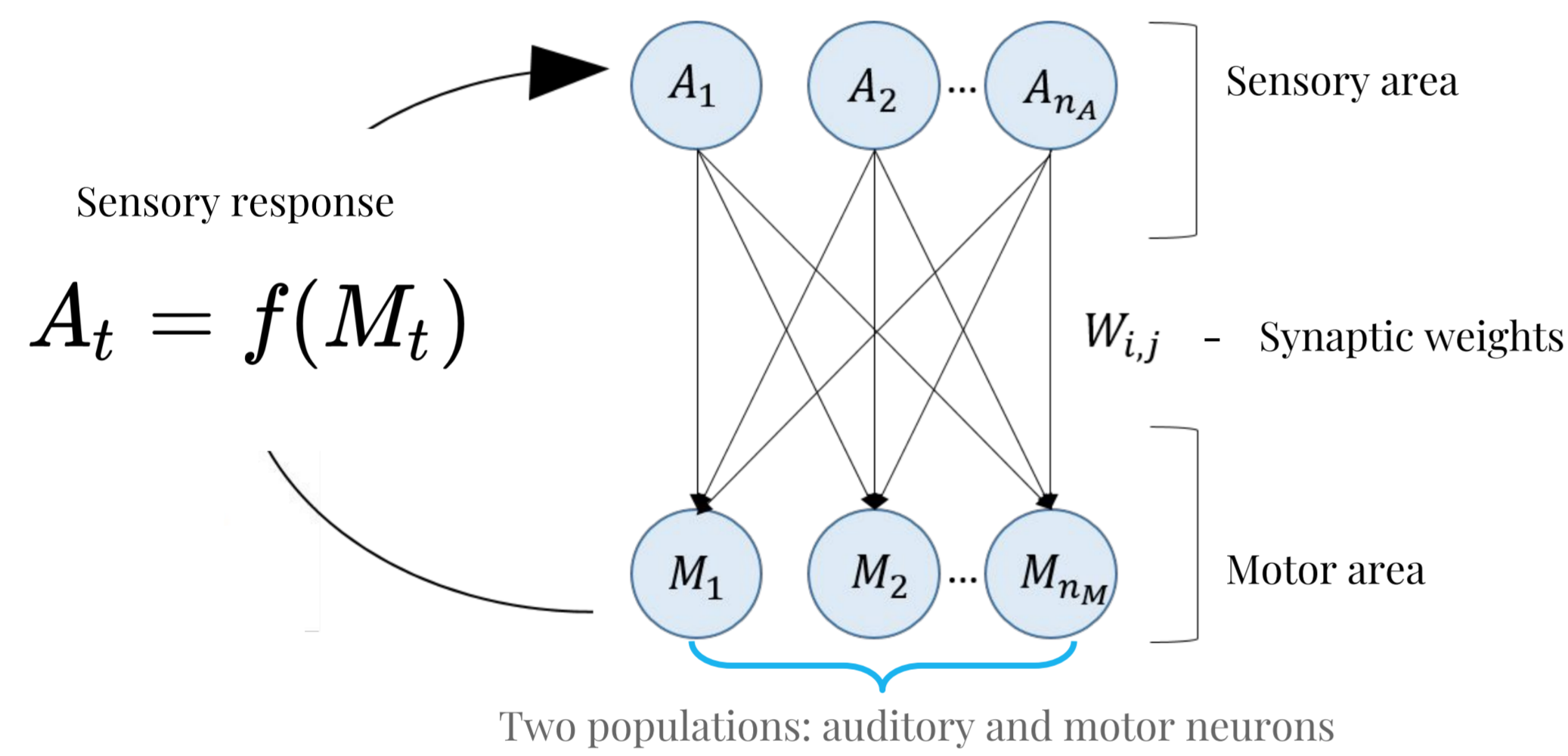
Brainard and Doupe, 2002

• Dedicated vocal circuit:



Motor pathway
Auditory pathway
Learning pathway

LEARNING AN INVERSE MODEL



$$A_t = f(M_t)$$

At each time step : t

Motor random exploration

- Auditory feedback: $A_t = f(M_t)$
- $\Delta W_t \propto \eta M_t A_t$
- Update weights: $W_t = W_{t-1} + \Delta W_t$

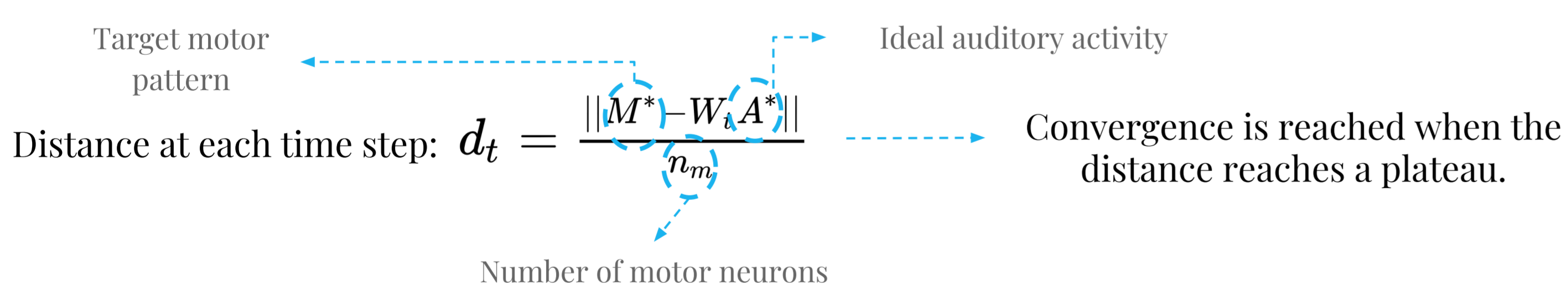
Hebbian learning rule

η : learning rate

SYNOPSIS WEIGHTS LIMITATION

Synaptic weights have a maximal value, related to the number of synaptic receptors one neuron is able to produce.

DISTANCE FROM THE TARGET AND CONVERGENCE



Distance at each time step: $d_t = \frac{\|M_t^* - W_t A_t\|}{n_m}$

Convergence is reached when the distance reaches a plateau.

Number of motor neurons

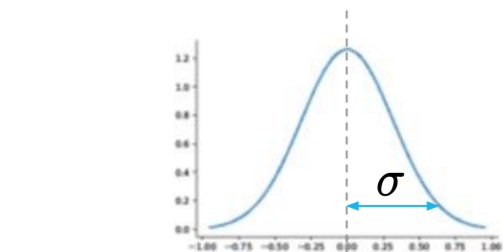
LINEAR (R. H. Hahnloser & S. Ganguli [4]) vs NONLINEAR MODEL

• Linear auditory response:

$$A_t = Q M_t$$

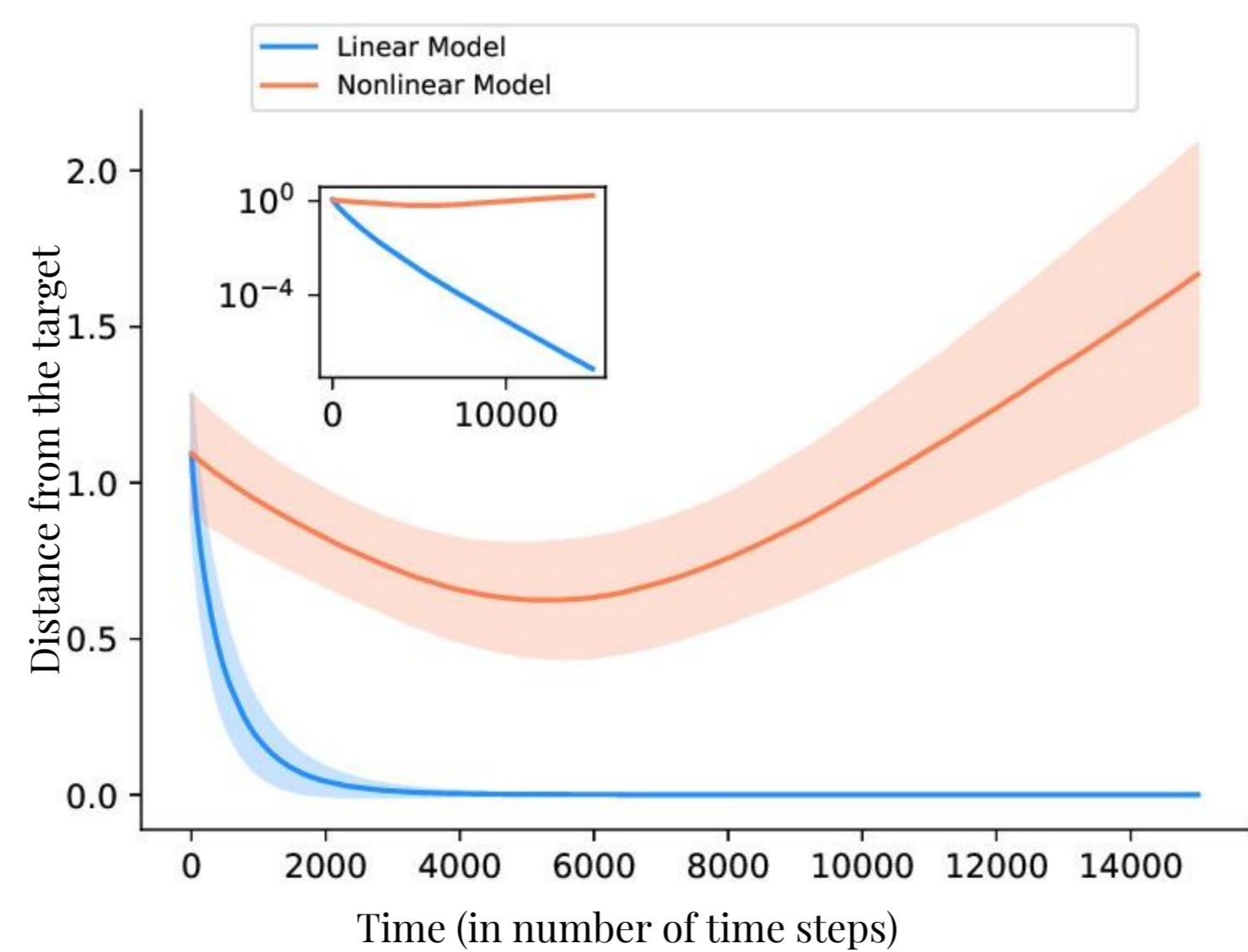
• Nonlinear auditory response:

$$A_{j,t} = \exp\left(\frac{-\|M_j^* - M_t\|^2}{2\sigma^2 n_m}\right)$$



σ selectivity tuning width which drives auditory selectivity.

$$\Delta W_t = \eta (M_t - W_{t-1} A_t) A_t^T$$

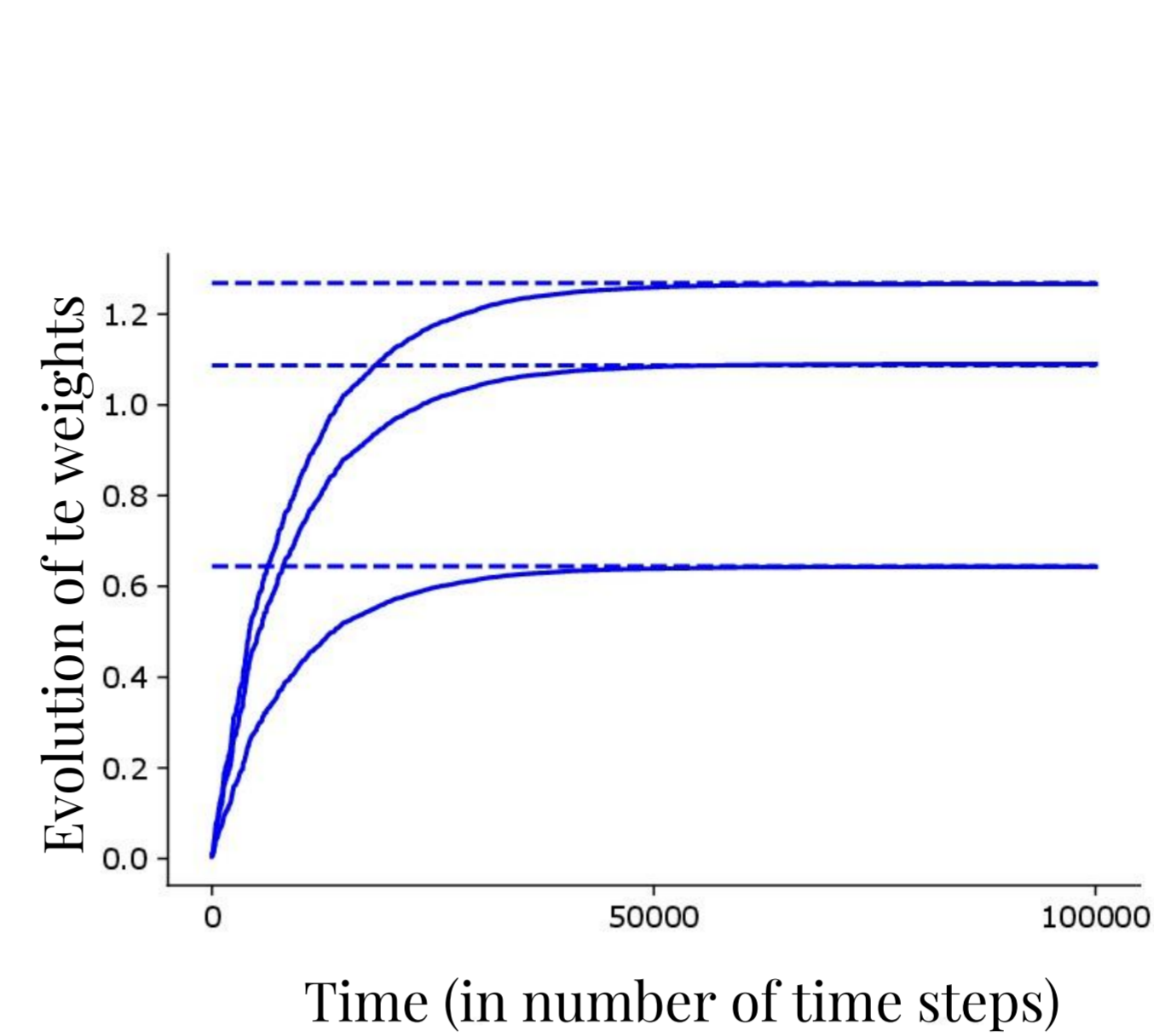


Nonlinear model does not converge applying this learning rule.

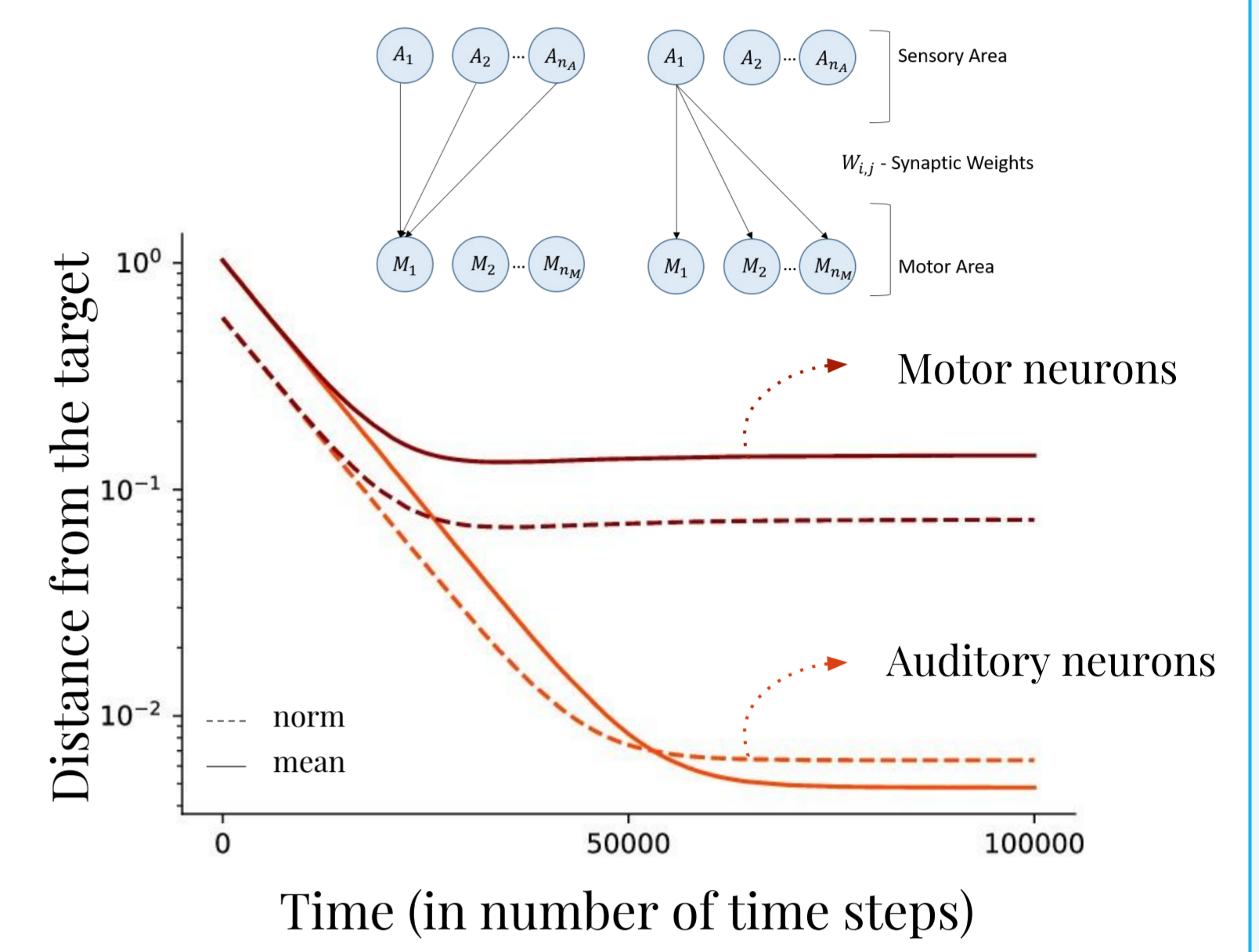
RESULTS

NORMALIZED INVERSE MODEL

Evolution of weights

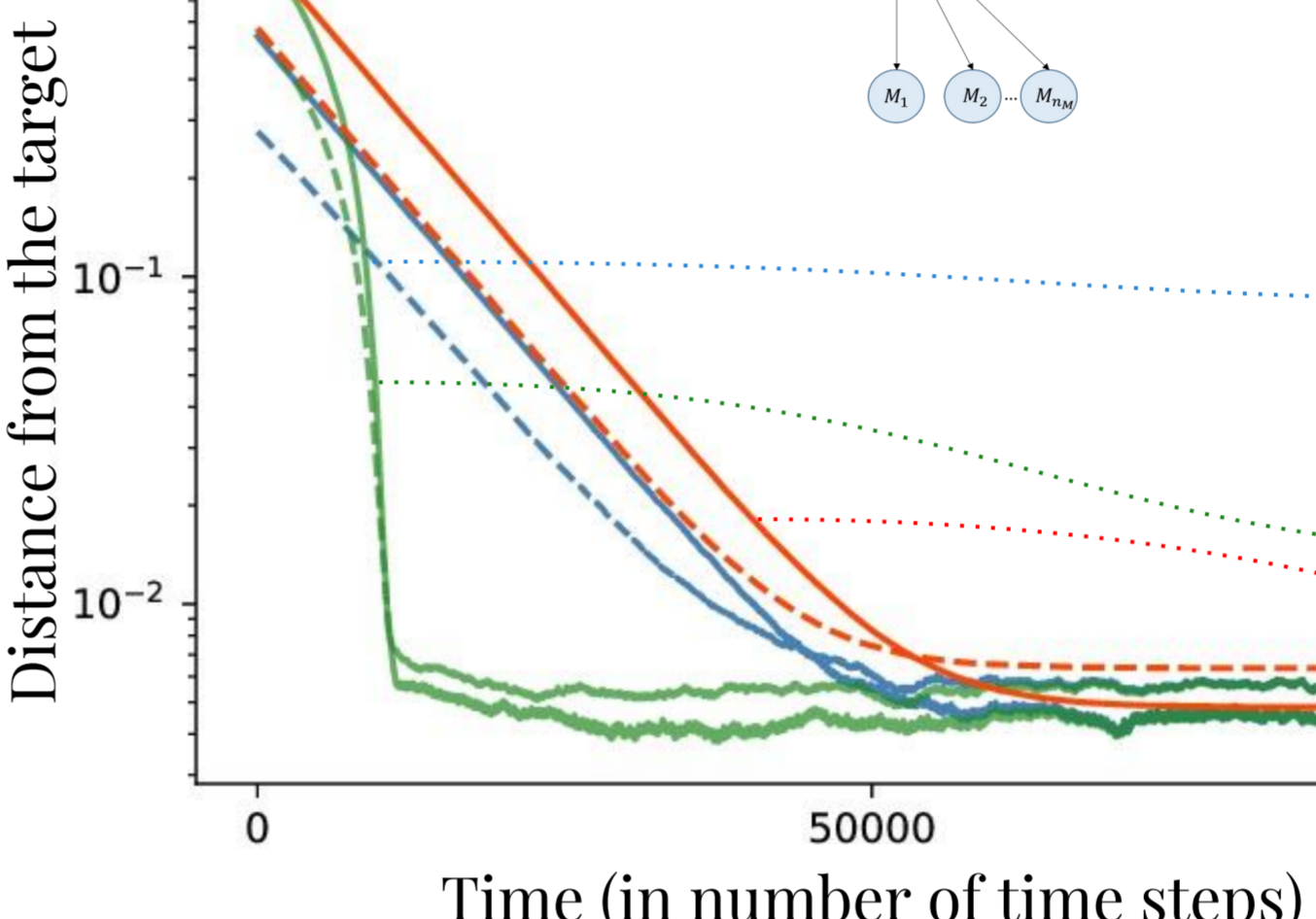


Motor VS Auditory



Normalization applied over auditory neurons gives better performance.

Distance from the target



Normalization applied over auditory neurons

Three normalizations

Mean of synaptic weights columns

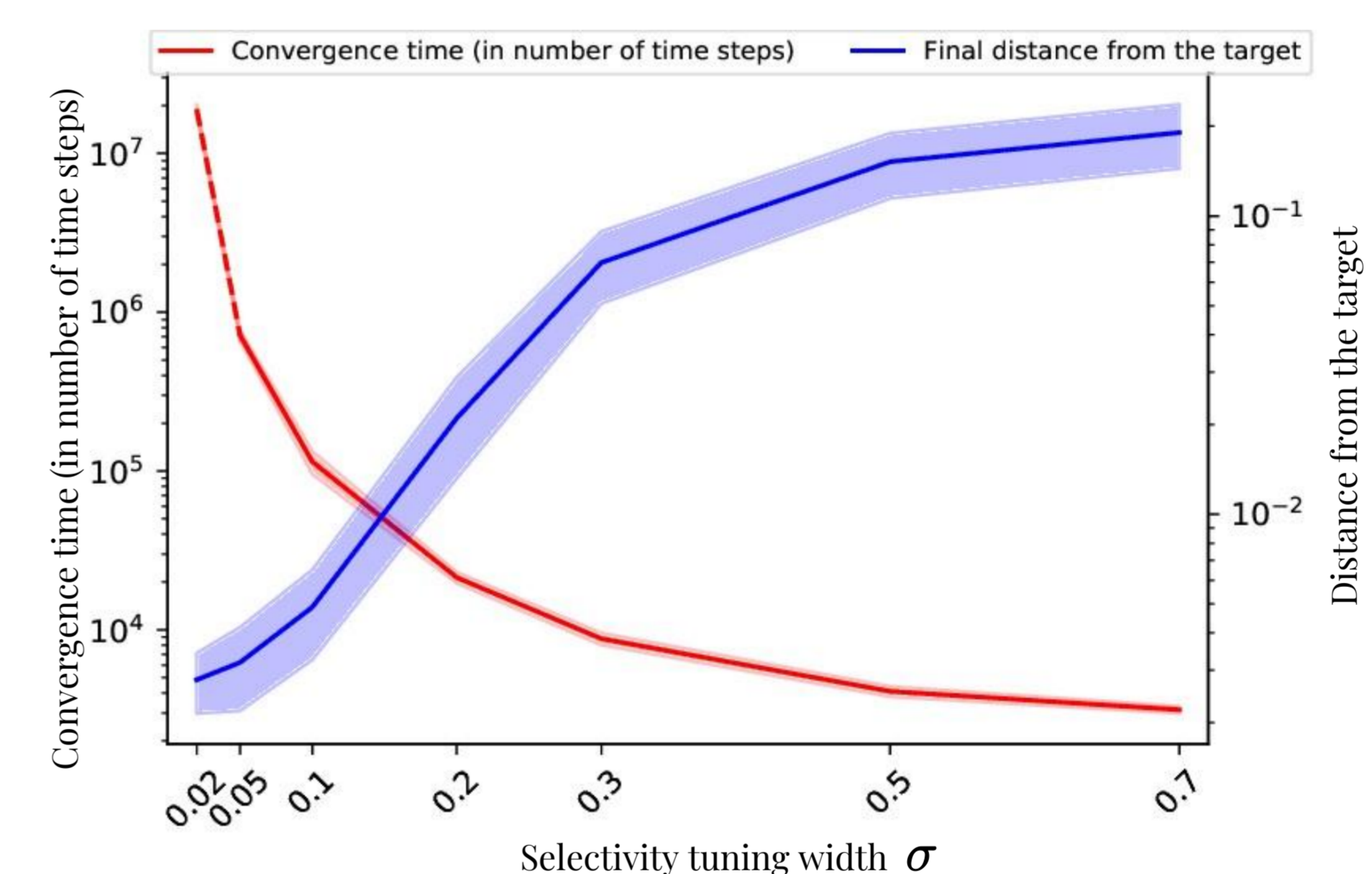
$$W_{i,j} = \frac{W_{i,j}}{\langle \mathbb{E}(W) \rangle}$$

$$W_{i,j} = \begin{cases} W_{i,j} & \text{if } \mathbb{E}(W) < 1 \\ \frac{W_{i,j}}{\mathbb{E}(W)} & \text{otherwise} \end{cases}$$

$$\Delta W_{i,j} = \eta M_t A_t (1 - \mathbb{E}(W))$$

Decreasing factor normalization gives better performance.

AUDITORY SELECTIVITY ANALYSIS

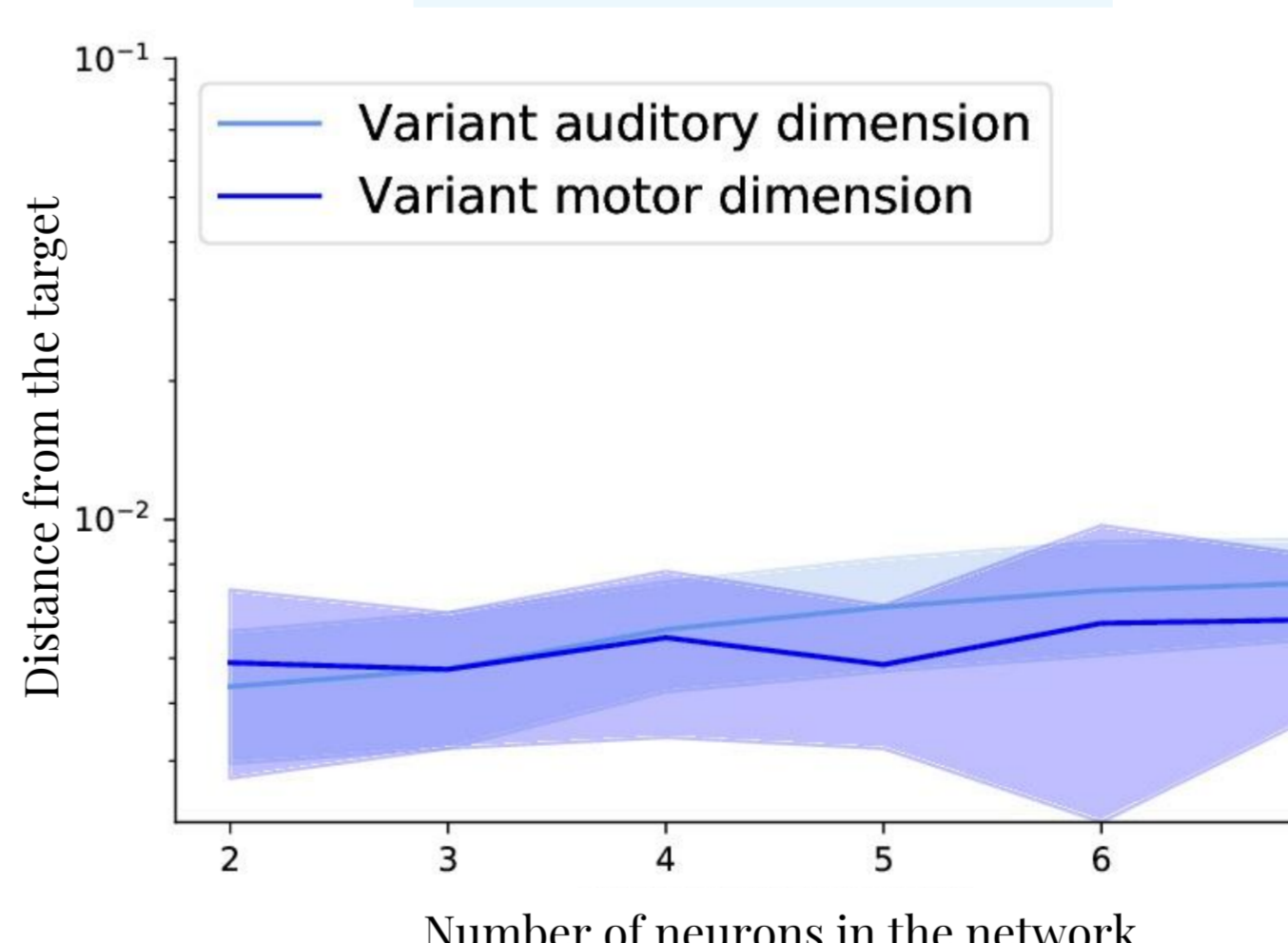


For simulations the decreasing factor normalization has been applied.

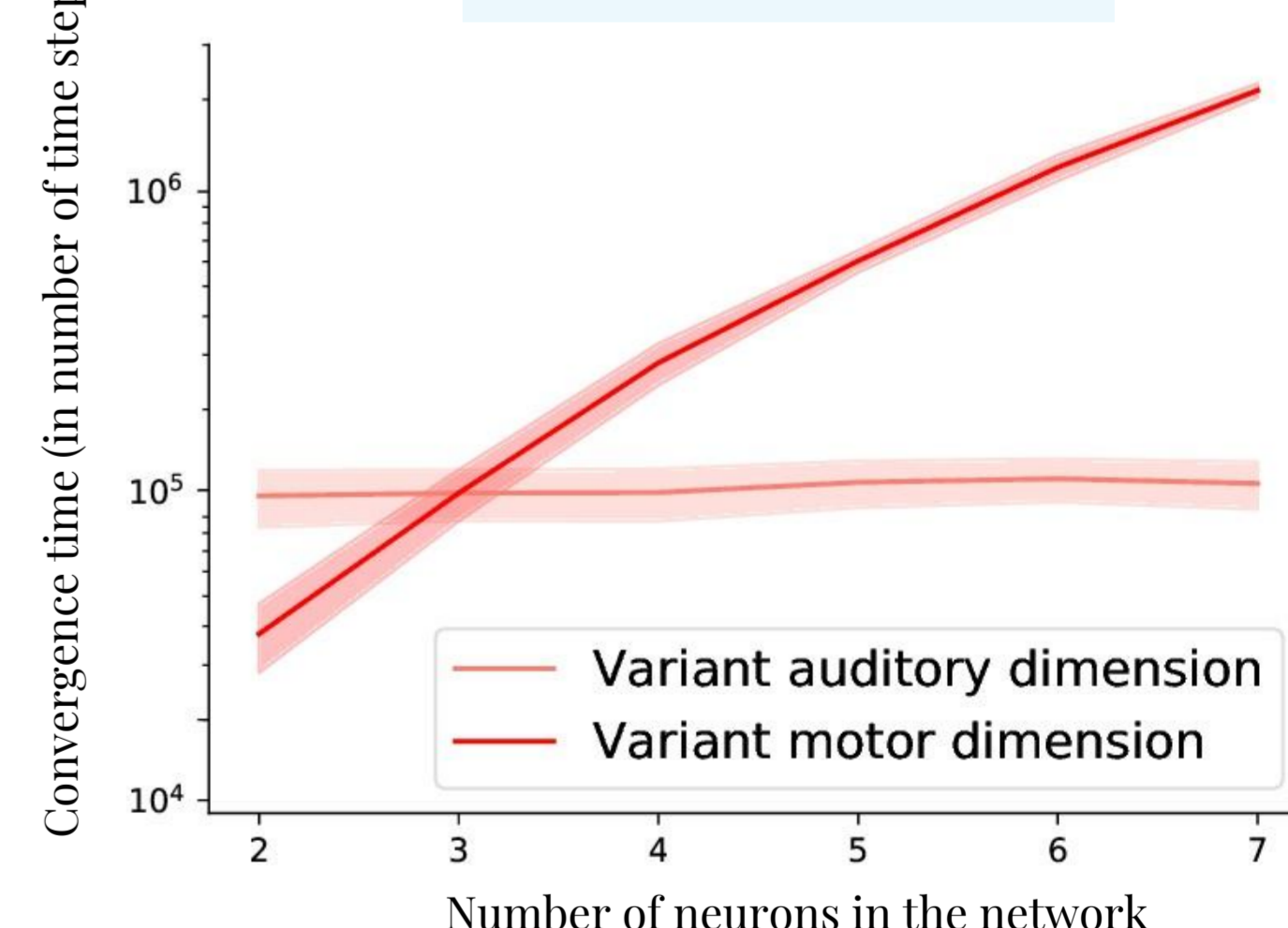
High selectivity - which means low tuning width - makes learning slower but more accurate.

VARYING INPUT/OUTPUT DIMENSION

Learning accuracy



Learning speed



Motor dimension influences learning making it slower and slower as it increases.

SUMMARY

- Normalization gives better performance when applied over auditory neurons.
- Low selectivity tuning width makes learning slower but more accurate.
- Motor dimension influences learning making it exponentially increasing.

WHAT'S NEXT?

- Exploration of motor dimension, dividing it in term of behavioral output and motor commands.
- Implement a realistic vocal production system (motor control model).

BIBLIOGRAPHY

- (1) Brainard, M. S. and Doupe, A. J., *What songbirds teach us about learning*, Nature, 2002
- (2) Mooney R., *Neural mechanisms for learned birdsong*, Cold Spring Harbor Laboratory Press, 2009
- (3) Wolpert D. M., Diedrichsen J. and Flanagan J. R., *Principles of sensorimotor learning*, Nature Reviews, 2011
- (4) Hahnloser R. H. R. and Ganguli S., *Vocal learning with inverse models*, *Principles of Neural Coding*, 2013, CRC Press Boca Raton