Cerebellar Learning and Applications

Matthew Hausknecht, Wenke Li,
Mike Mauk, Peter Stone

January 23, 2014

Center for Learning and Memory
The University of Texas at Austin
Mauk lab home
The University of Texas at Austin What Starts Here Changes The World

LARG
Learning Agents Research Group
The University of Texas at Austin

Motivation

- Introduces a novel learning agent: the cerebellum simulator.
- Study the successes and failures the cerebellum on machine learning tasks.
- Characterize the cerebellum’s capabilities and limitations.
- Develop a set of guidelines to help understand what tasks are amenable to cerebellar learning.

Outline

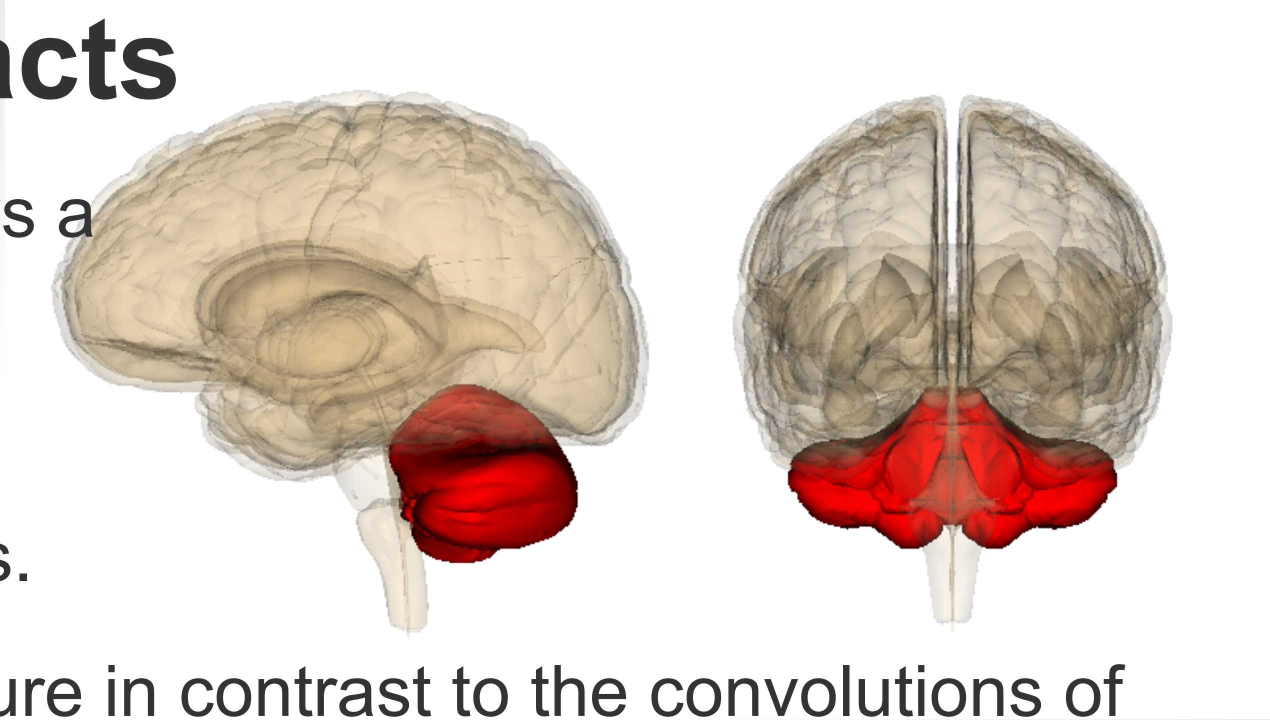
- Introduction: Biology of the cerebellum

- Cerebellum Simulator

- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition

- Conclusions

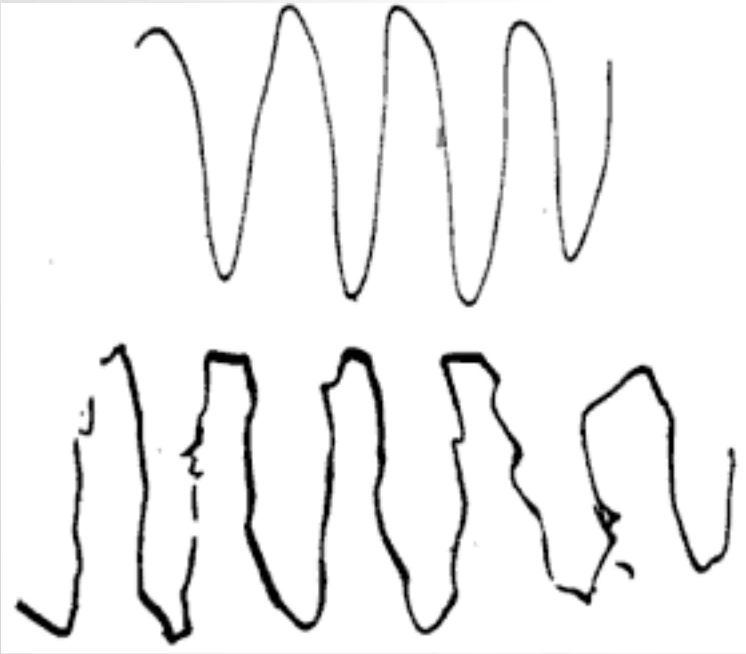
Cerebellum Facts



- Brain region that plays a role in motor control.
- Located beneath the cerebral hemispheres.
- Highly regular structure in contrast to the convolutions of the cerebral cortex.
- 10% of total brain volume but contains more neurons than rest of brain put together. (Half of the total neurons in brain are cerebellar granule cells)
- Does not initiate movement, but instead is responsible for fine tuning, timing, and coordinating fine motor skills.

Ataxia

Damage to the cerebellum results not in paralysis, but instead produces disorders fine movement, equilibrium, posture and motor learning.

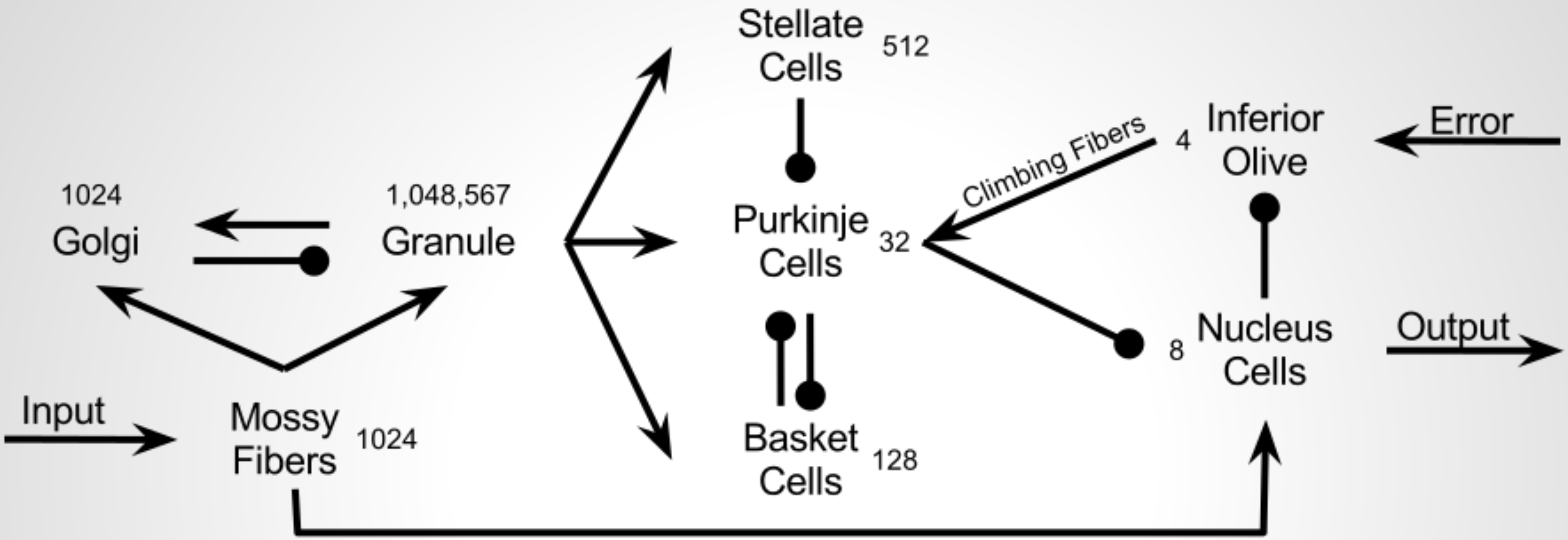


Top: Altered gate of woman with cerebellar disease.

Left: Attempt by cerebellar diseased patient to reproduce trace on top

Images: https://en.wikipedia.org/wiki/Ataxia

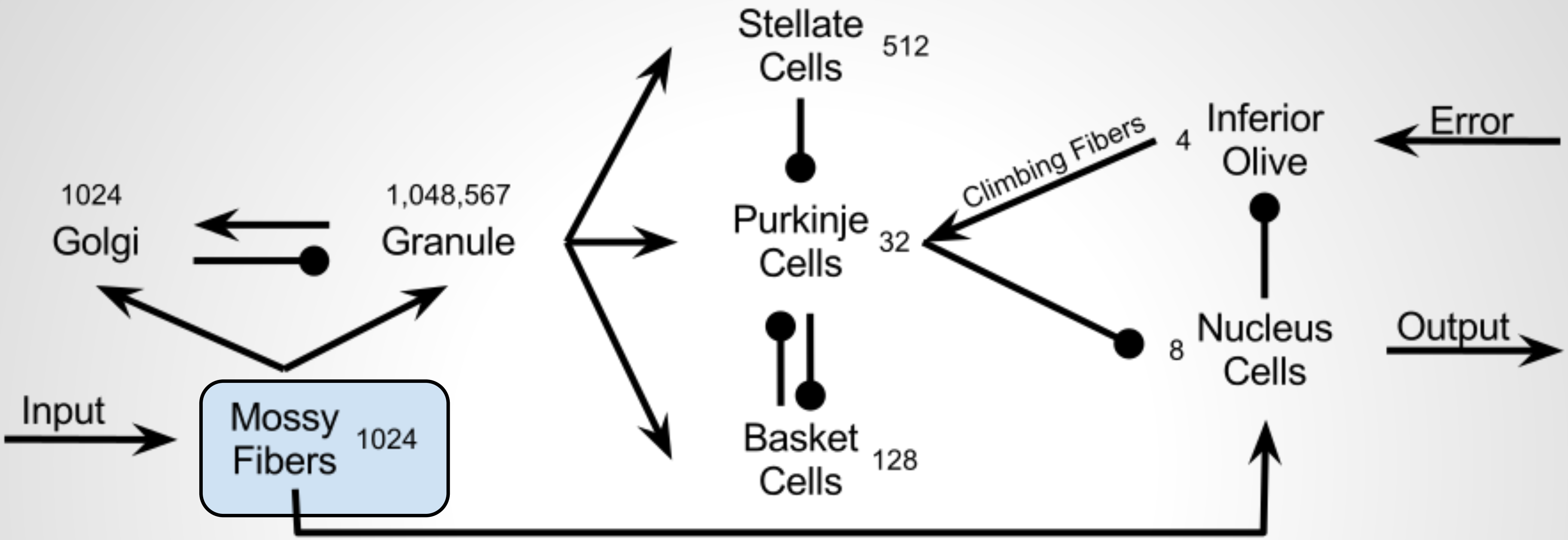
Synaptic Connectivity



- Cerebellar connectivity is highly regular with an enormous number of neurons but a limited number of neuron types.

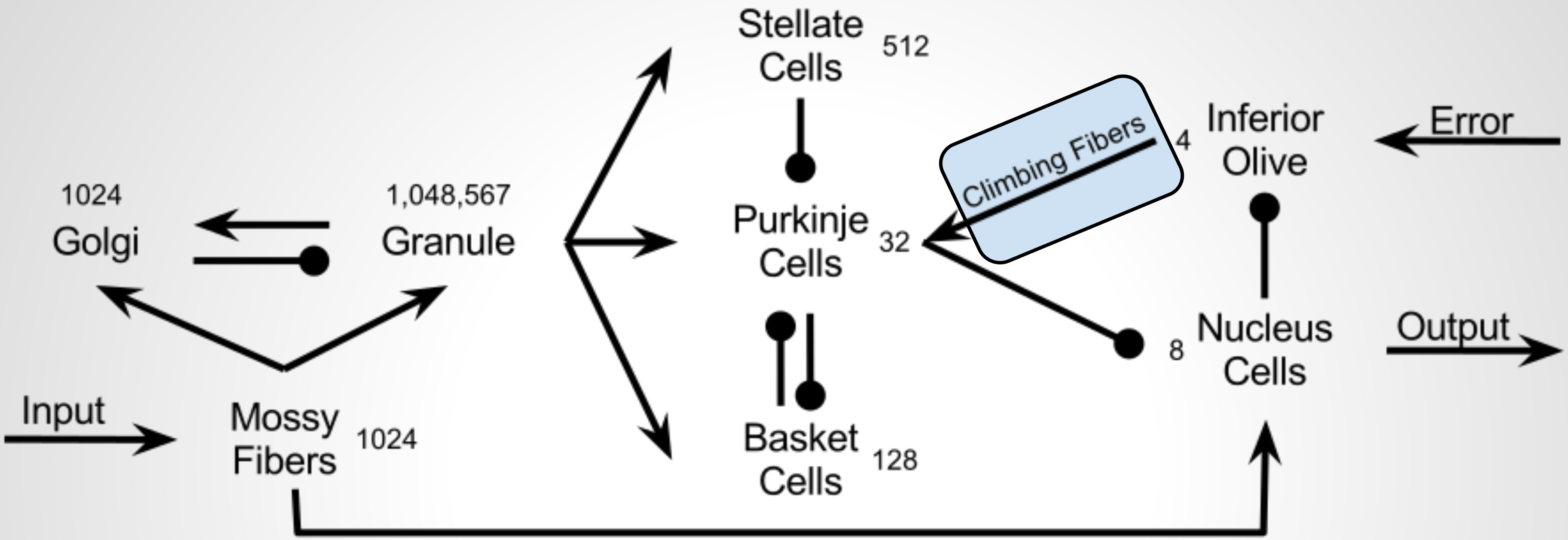
- Arrows denote excitatory connections while circles denote inhibitory connections. Numbers indicate number of simulated cells.

Mossy Fibers



- Carry external information about the state of the world to the rest of the cerebellum.

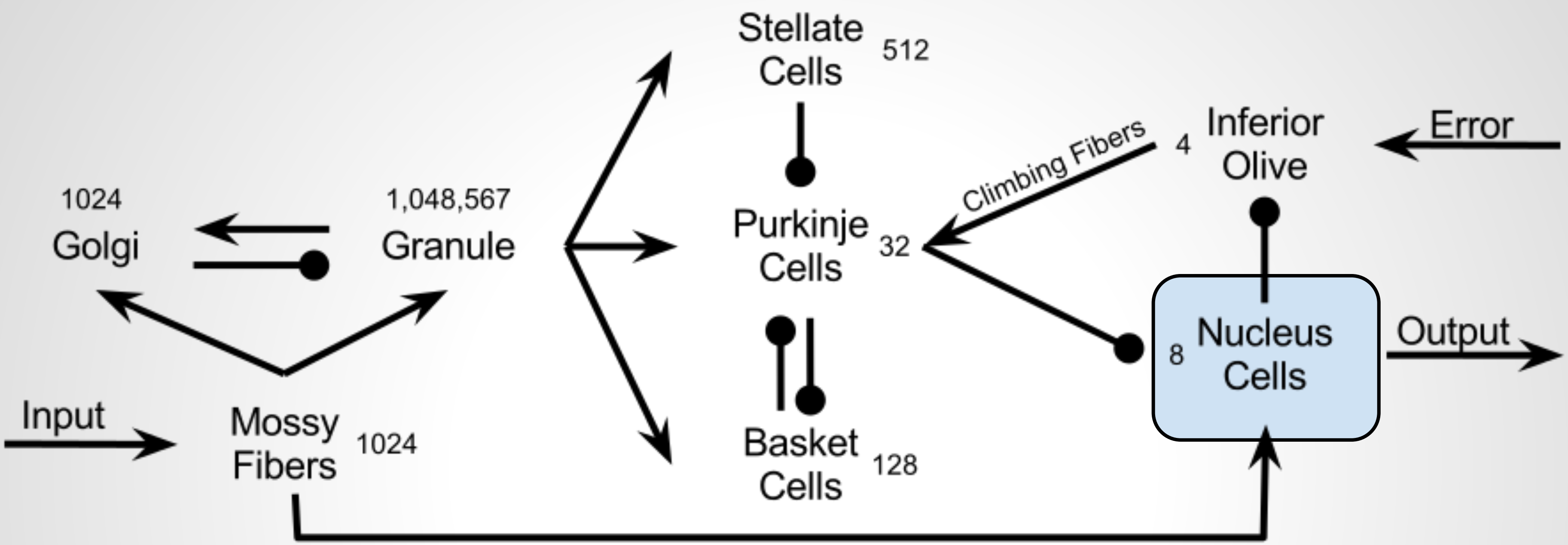
Climbing Fibers



- Teaching signals originate in the Inferior Olive and are transmitted via the Climbing Fibers.

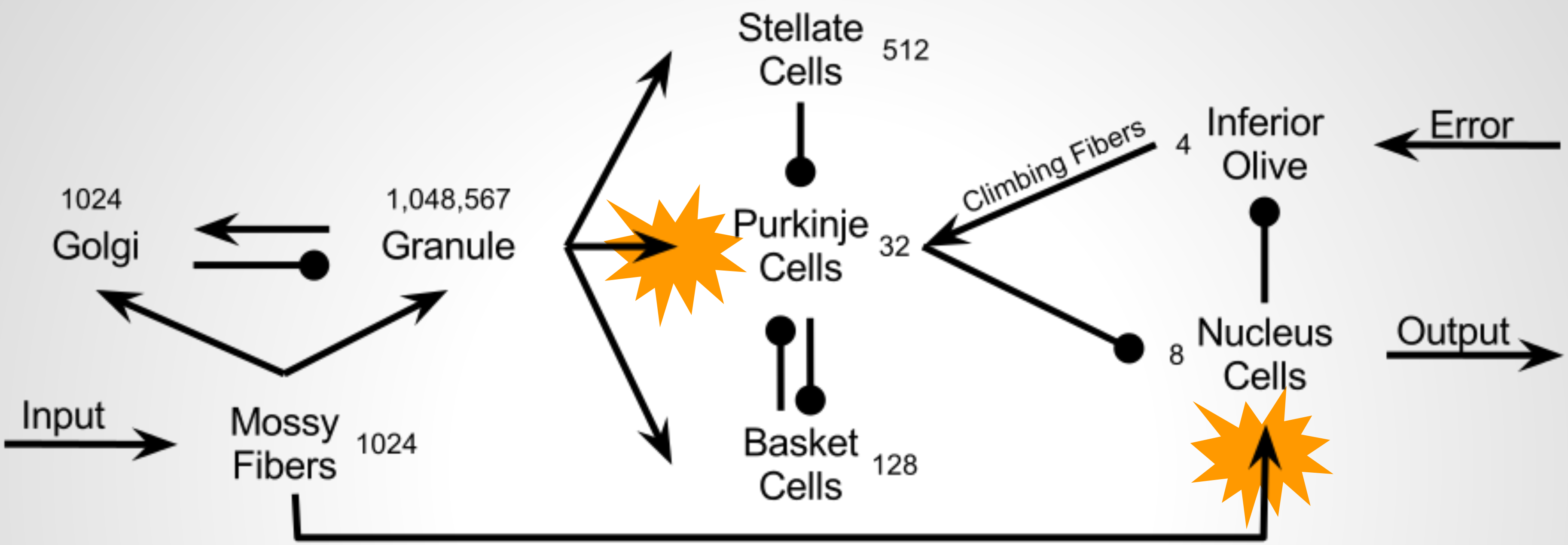
- Teaching signals indicate the need for changes in synaptic plasticity and ultimately behavior.

Nucleus Cells



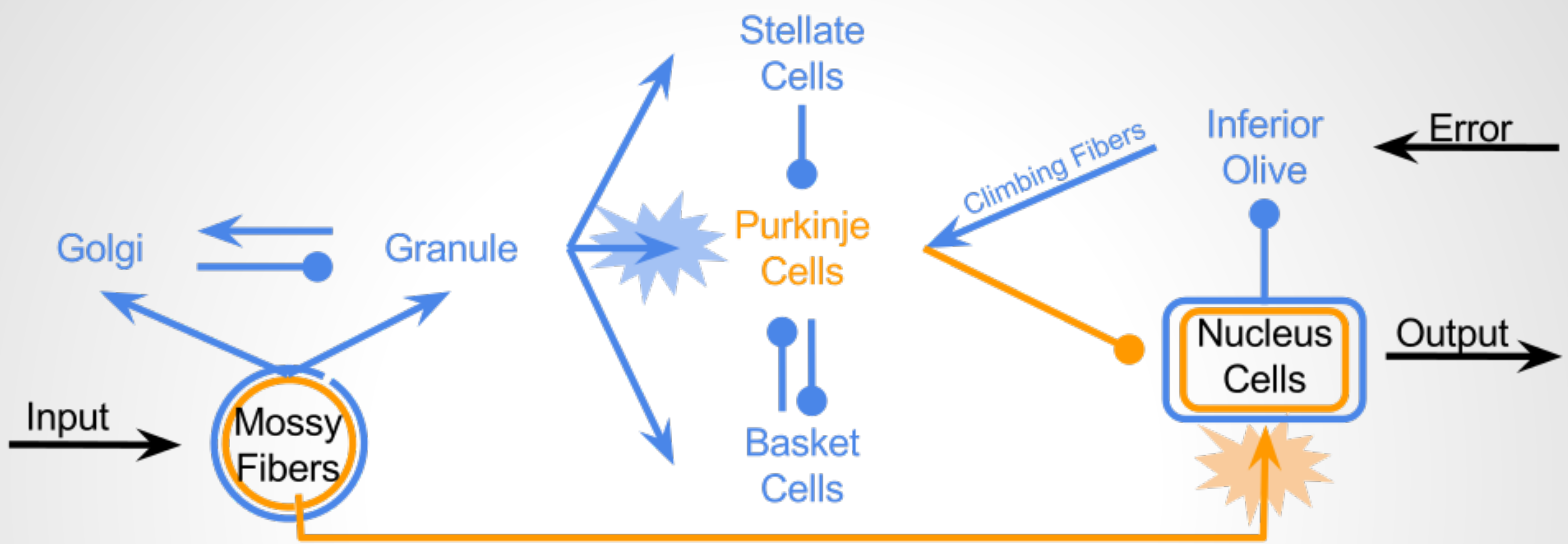
- Outputs from the nucleus cells form the basis of muscle control.

Cerebellar Learning Mechanisms



- Learning takes place by updating synaptic plasticity at two sites: GR:Purkinje and MF:Nucleus.
- Synaptic plasticity is the ability of the connection or synapses between two neurons to change in strength.

Learning Pathways



- Direct pathway:

$$\Delta w_i^{mf} = \delta_-^{mf} \cdot \mathrm{MF}_i \cdot \Theta_{LTD}^{PKJ}(50) + \delta_+^{mf} \cdot \mathrm{MF}_i \cdot \Theta_{LTP}^{PKJ}(50)$$

- Indirect pathway:

$$\Delta w_i^{gr} = \delta_-^{gr} \cdot \mathrm{GR}_i \cdot \mathrm{CF}(100) + \delta_+^{gr} \cdot \mathrm{GR}_i \cdot (1 - \mathrm{CF}(100))$$

Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition
- Conclusions

Cerebellum Simulator

- Cellular level simulation of the cerebellum.
- Based on a previous simulator built by Buonomano and Mauk[1].
- Primary difference from previous simulator is a nearly 100x increase in the number of cells: from 12,000 to 1,048,567.
- At this scale divergence/convergence ratios of granule cell connectivity more closely approximate those in the brain.
- Developed and parallelized by Wenke Li.

[1]Dean V. Buonomano and Michael D. Mauk. *Neural network model of the cerebellum: temporal discrimination and the timing of motor responses.* Neural Comput., 6:38–55, January 1994.

Parallel Implementation

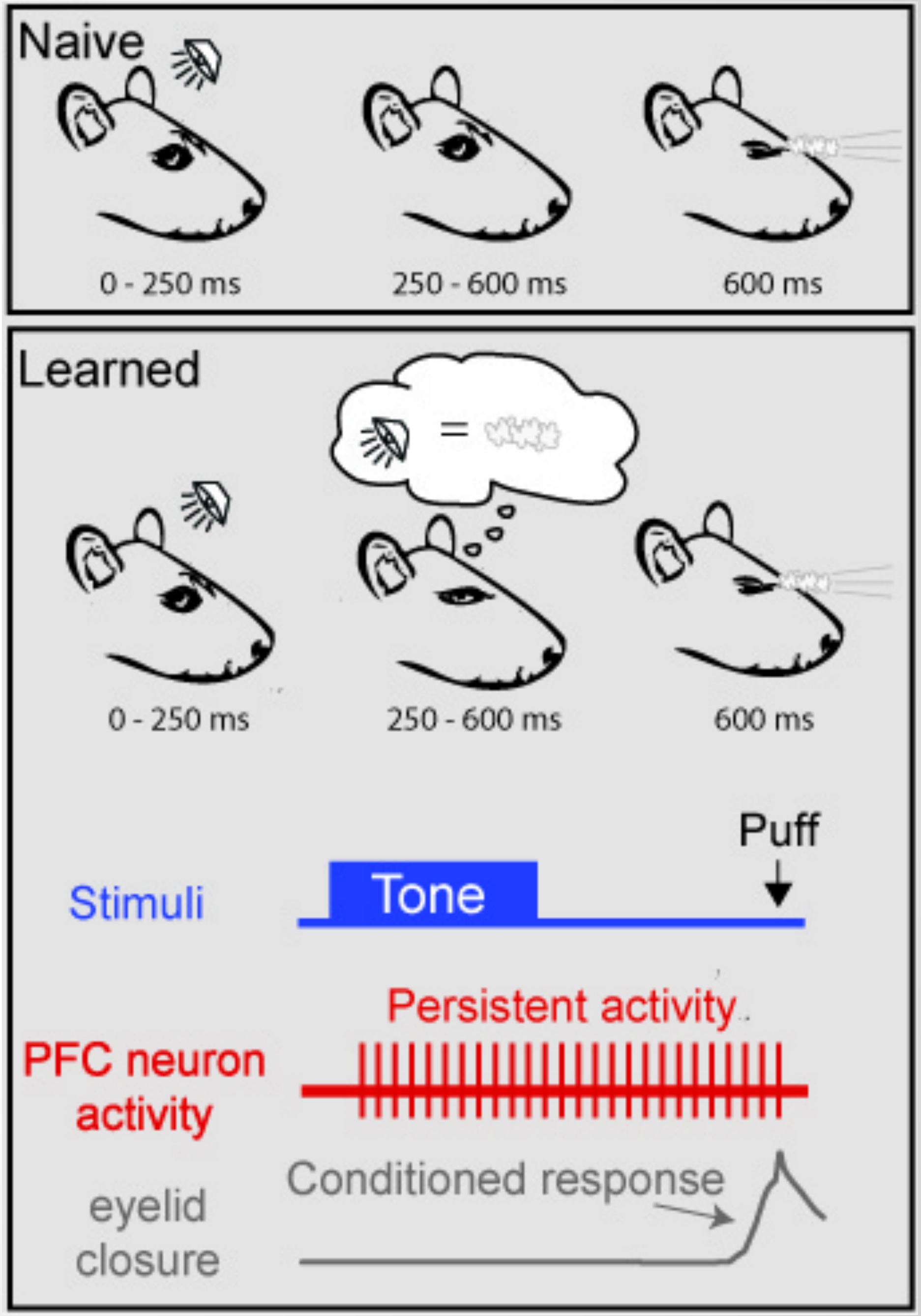
- Relies on Nvidia Cuda GPUs to compute granule cell firings in parallel.
- Traditional parallel programming approach (OpenMP etc) were inadequate due to high memory bandwidth required ~128 GB/s for real-time operation.
- GPU computation provides necessary memory bandwidth as well as several hundred cores.
- A single Nvidia Fermi GTX580 GPU brings the simulation to 50% real-time speed.

Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition
- Conclusions

Eyelid Conditioning

- Rabbits learn to close their eyes in response to a tone being played.
- Lesioning of cerebellum renders animals incapable of learning responses[1].
- Unpaired CS+US results in extinction.
- Simulator tuned from to match experimental data collected from rabbits.



[1] McCormick et al. (1981)

Outline

- Introduction: Biology of the cerebellum

- Cerebellum Simulator

- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition

- Conclusions

Inverted Pendulum Balancing

- Objective: keep an inverted pole balanced for as long as possible.
- Forces are applied to the cart along the axis of movement.
- Differs from Eyelid conditioning in that forces now need to applied in two directions.

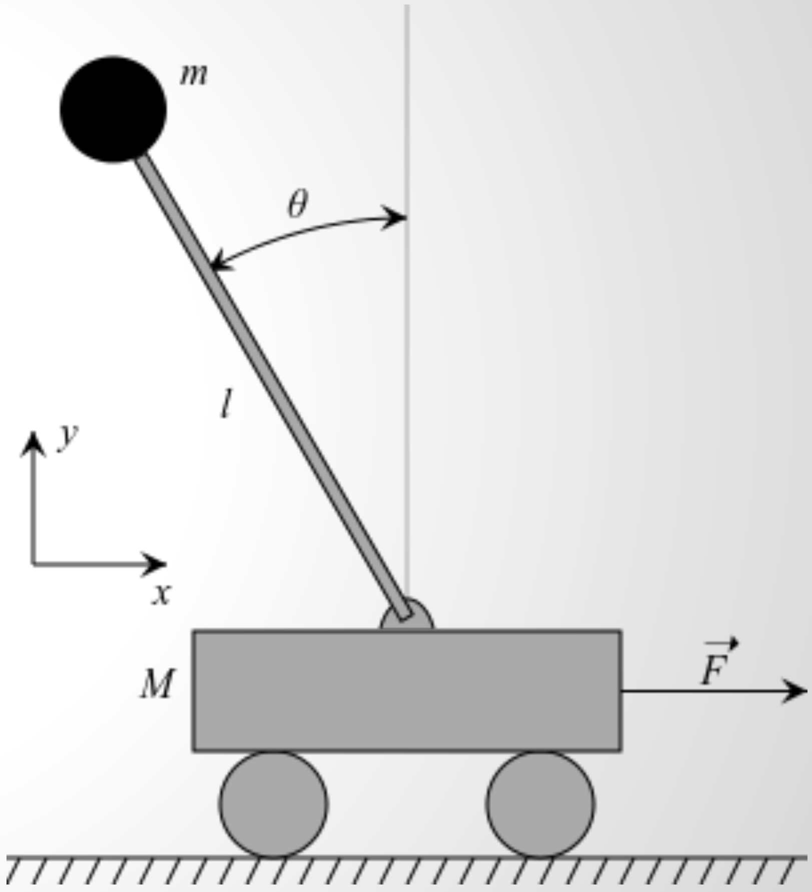


Image: https://en.wikipedia.org/wiki/Inverted_pendulum

Inverted Pendulum Balancing

- Main challenge: How best to interface the cerebellum simulator to the inverted pendulum domain?

- Three main questions:
  1. How to encode state of cart & pole?
  2. How and when to deliver error signals?
  3. How to interpret outputs as forces?

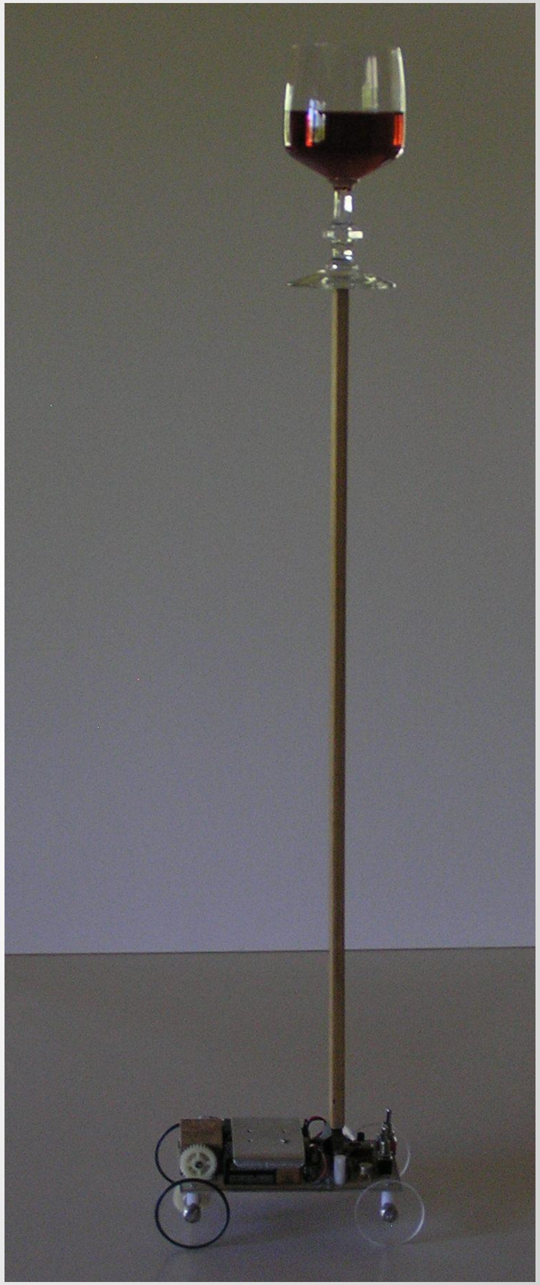
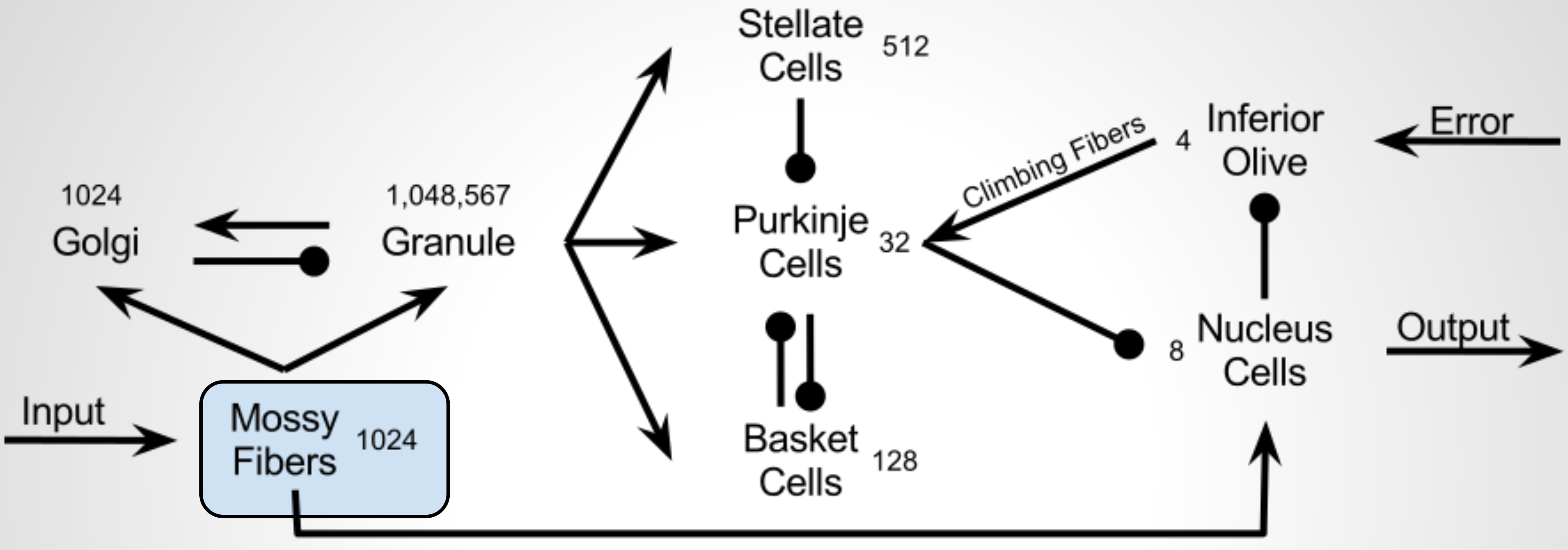


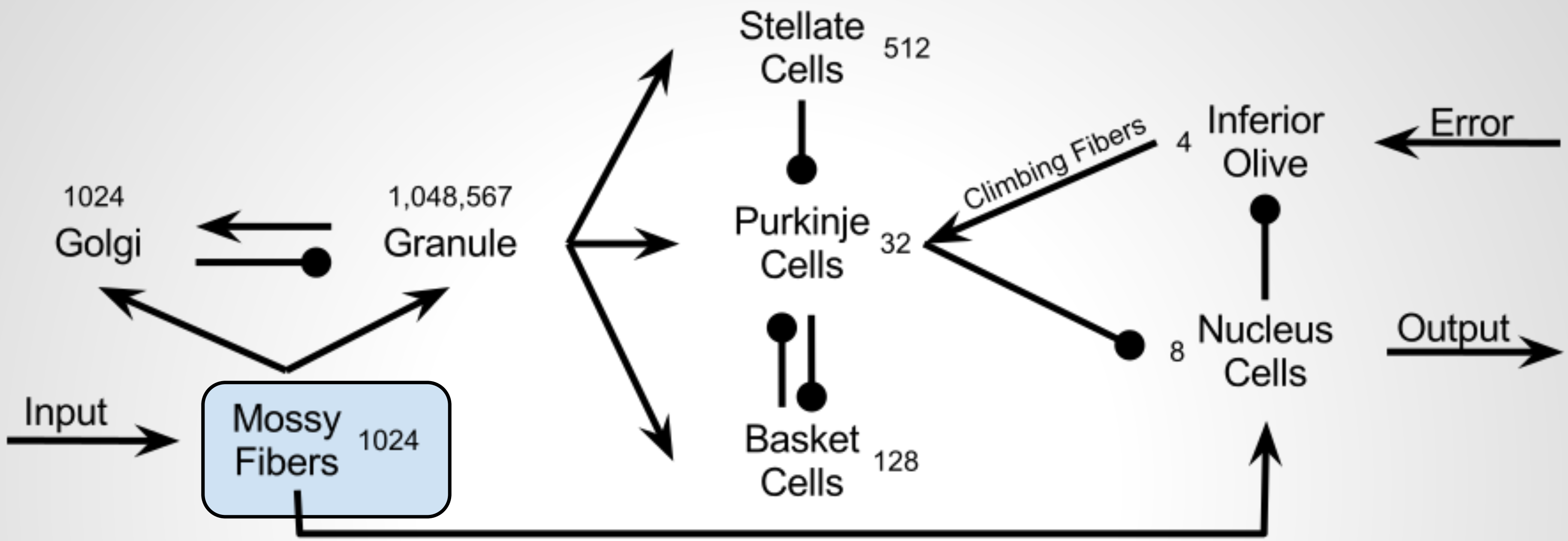
Image: https://en.wikipedia.org/wiki/Inverted_pendulum

Mossy Fibers



- Carry external information about the state of the world to the rest of the cerebellum.

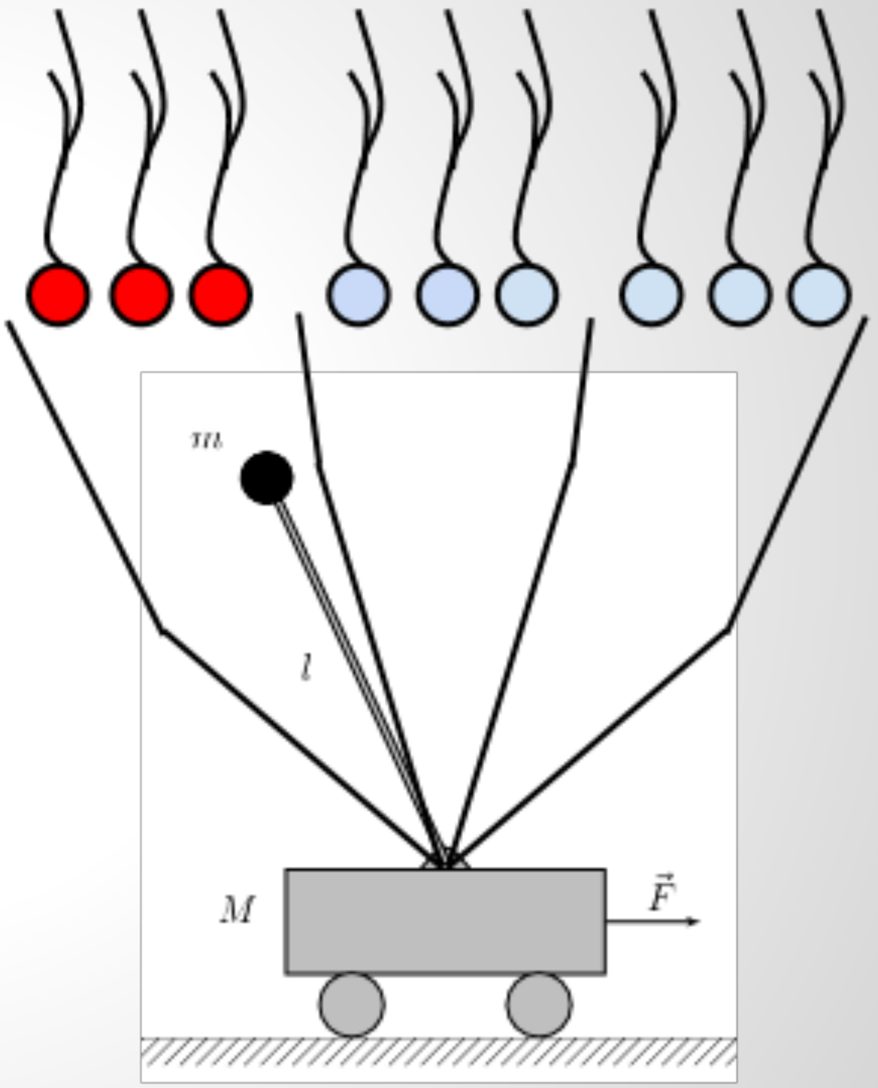
State Signal Interface



- Challenge: Convey Pole Angle, Pole Velocity, Cart Position, and Cart Velocity.
- 1024 Mossy Fibers (MFs) available.
- When at rest MFs fire with a low background frequency.
- When excited, MF firing rate increases.
- Need to selectively excite MFs.

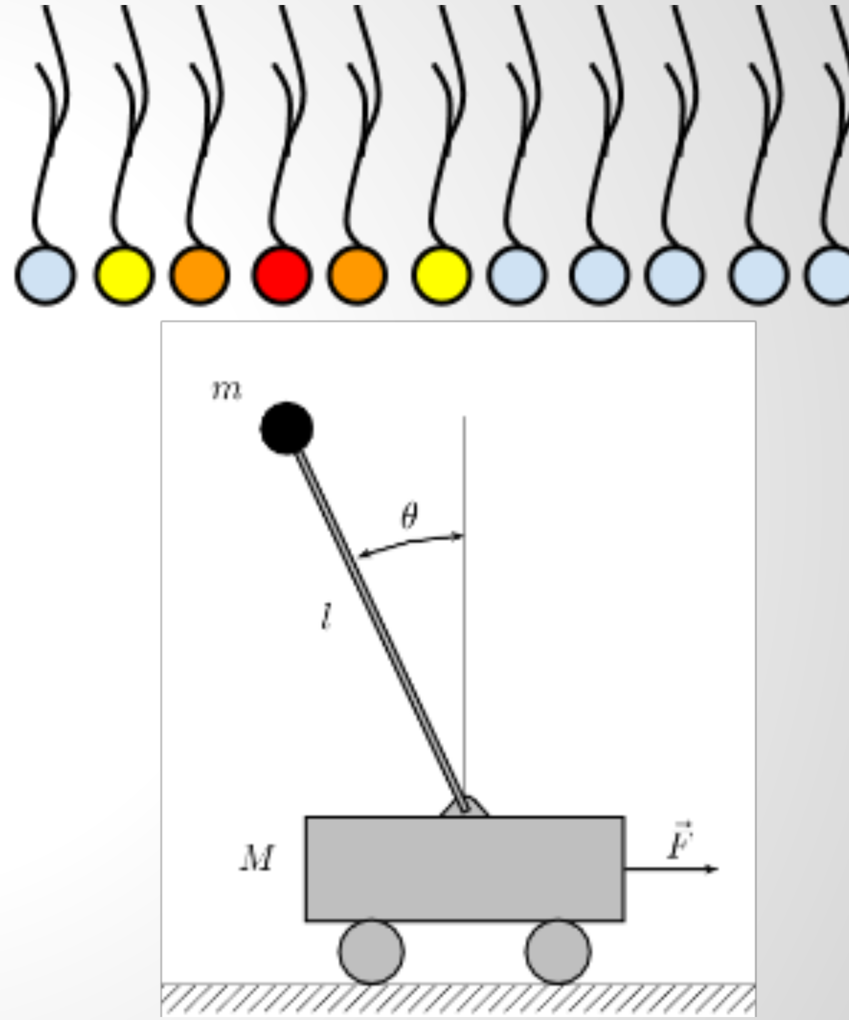
Boolean State Encoding

- Has 3 receptive zones (tiles).
- Increases firing rates of MFs in the active zone.
- Conveys rough information about the location of the pole.



Gaussian State Encoding

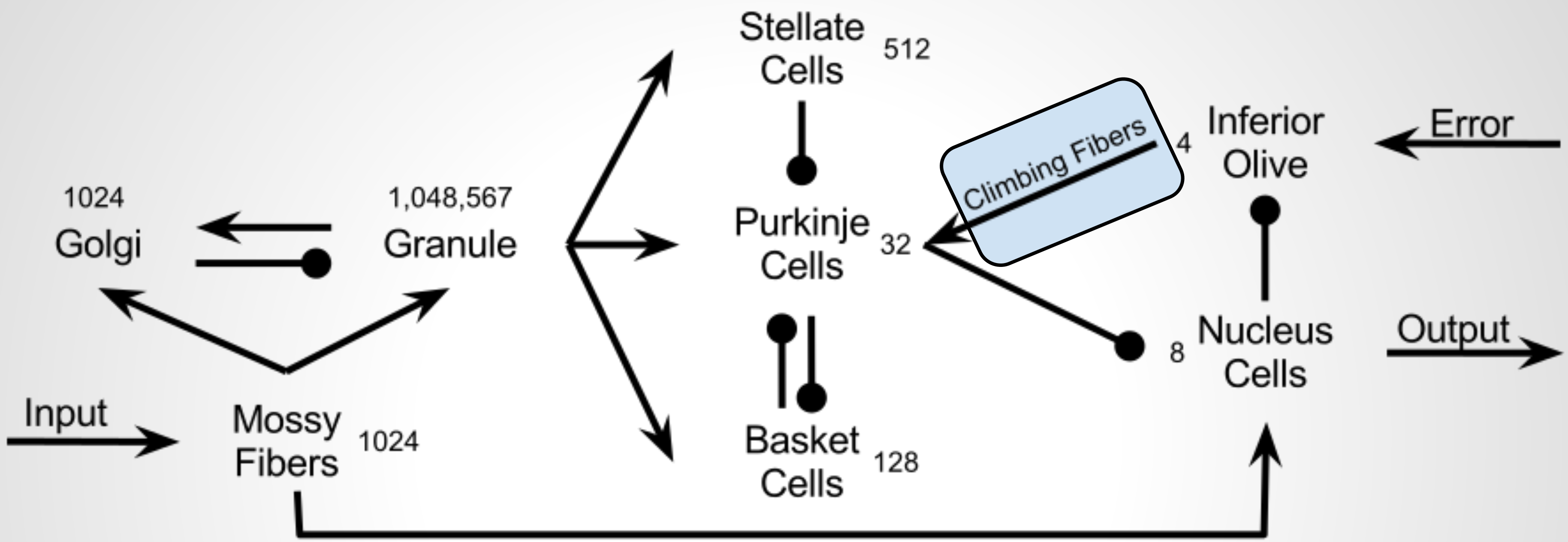
- Multiple receptive zones (tiles).
- Assign MFs values in 'input space.'
- Each MF fires proportional to how close the pole angle value is to its value in input space.
- Conveys fine-grained information about the location of the pole.



State Signal Interface

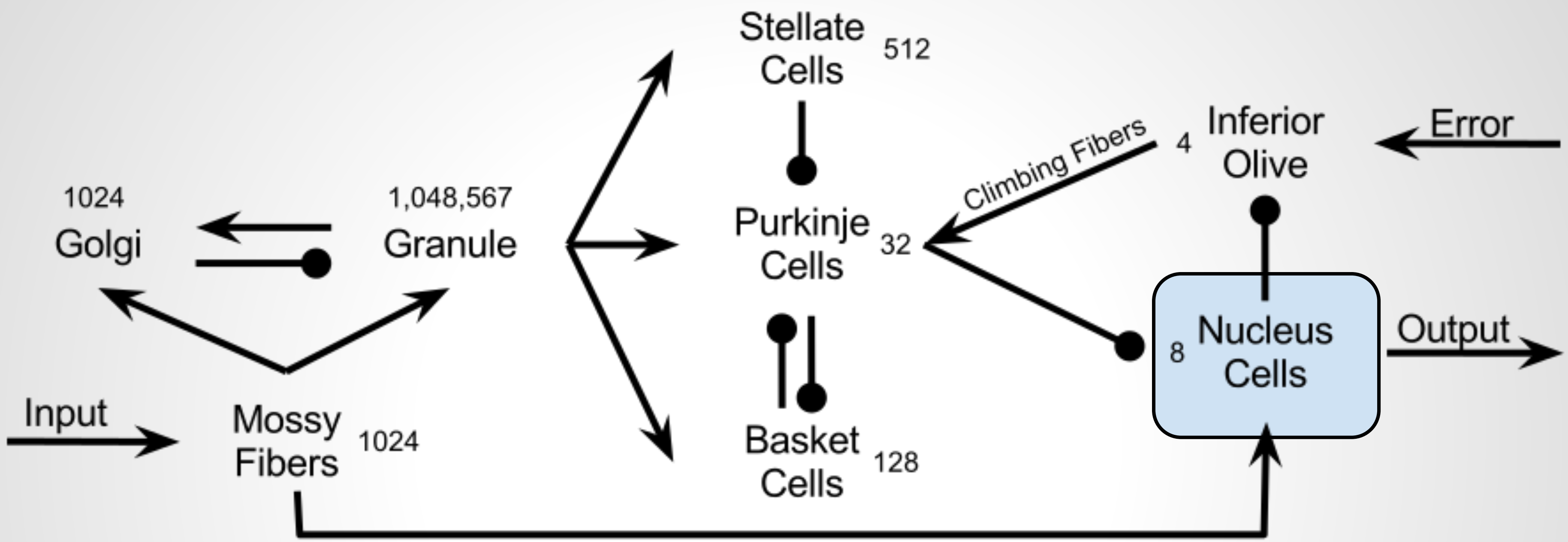
- 1024 total Mossy Fibers (MFs) process input.
- We assign 30 random MFs each to encode pole angle, pole velocity, cart position, and cart velocity.
- Lastly we have 30 MFs which fire with high frequency regardless of state.
- MFs for each state variable are randomly distributed throughout the 1024, so the cerebellum must decided which MFs carry signal and which do not.
- Both Boolean and Gaussian encodings have proved successful.

Error Signal Interface



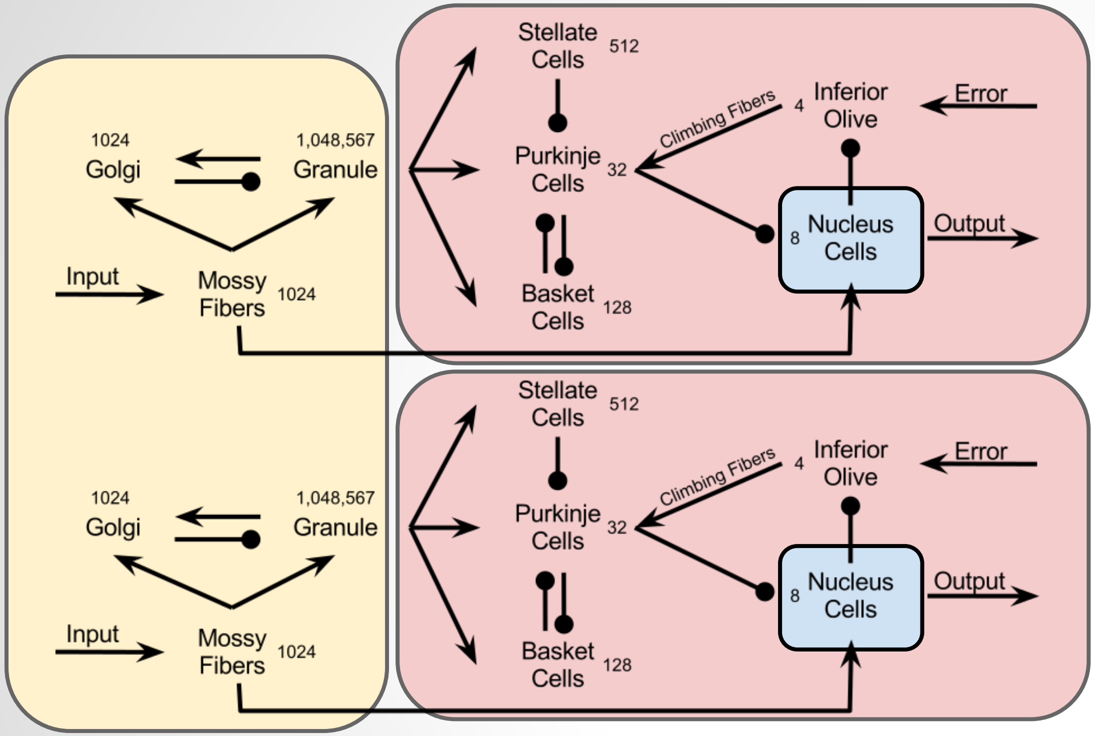
- Four Climbing Fibers transmit error input.
- Inverted pendulum domain receives error with probability proportional to how far the pole differs from upright.
- Errors are boolean in nature, so at each timestep if error is received either all 4 climbing fibers activate or none.

Output Signal Interface



- Output is produced by 8 Nucleus Cells.
- Combine NC firings into a single output force in range [0,1]: *NumberFiringNCs / 8*.
- This provides a single output force, but Inverted Pendulum requires two opposing forces.

Microzones

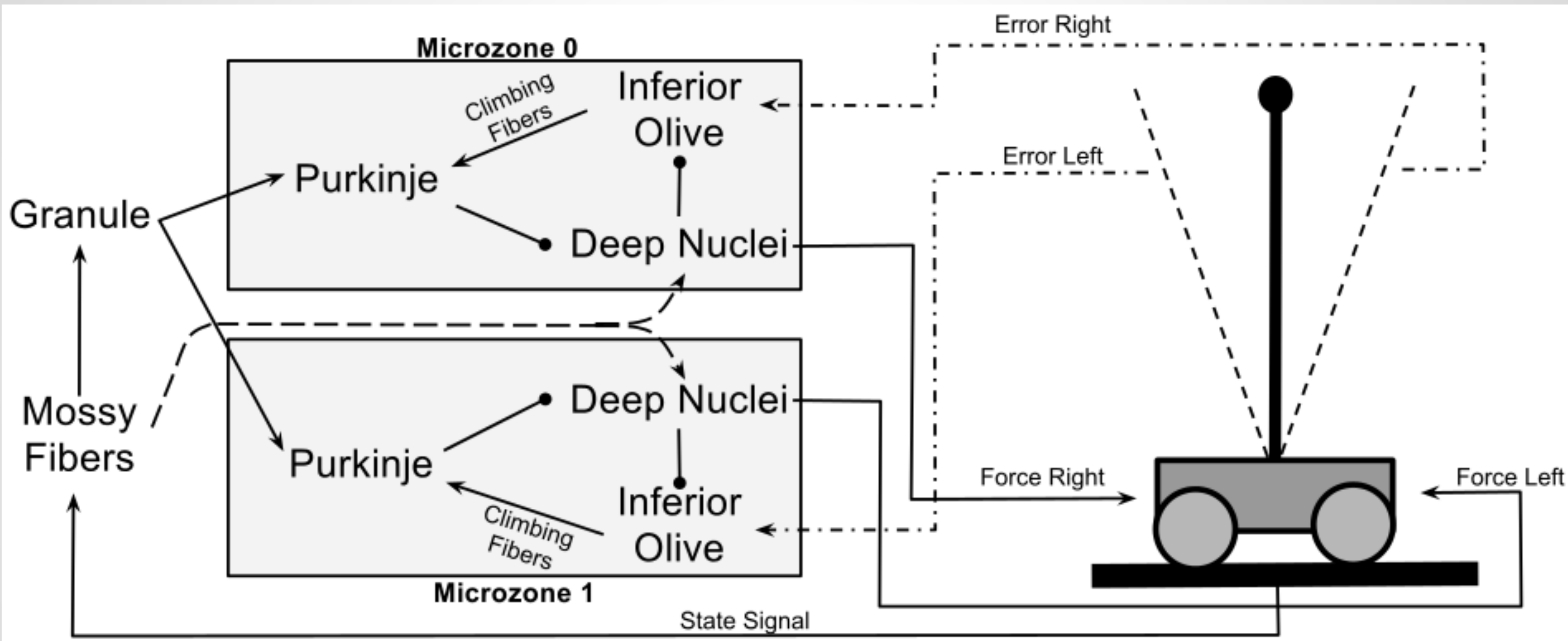


Output Network 1

Output Network 2

- Frequently need to control 2 or more effectors
- Group common input cells and duplicate only the output networks
- These output networks are called “Microzones”

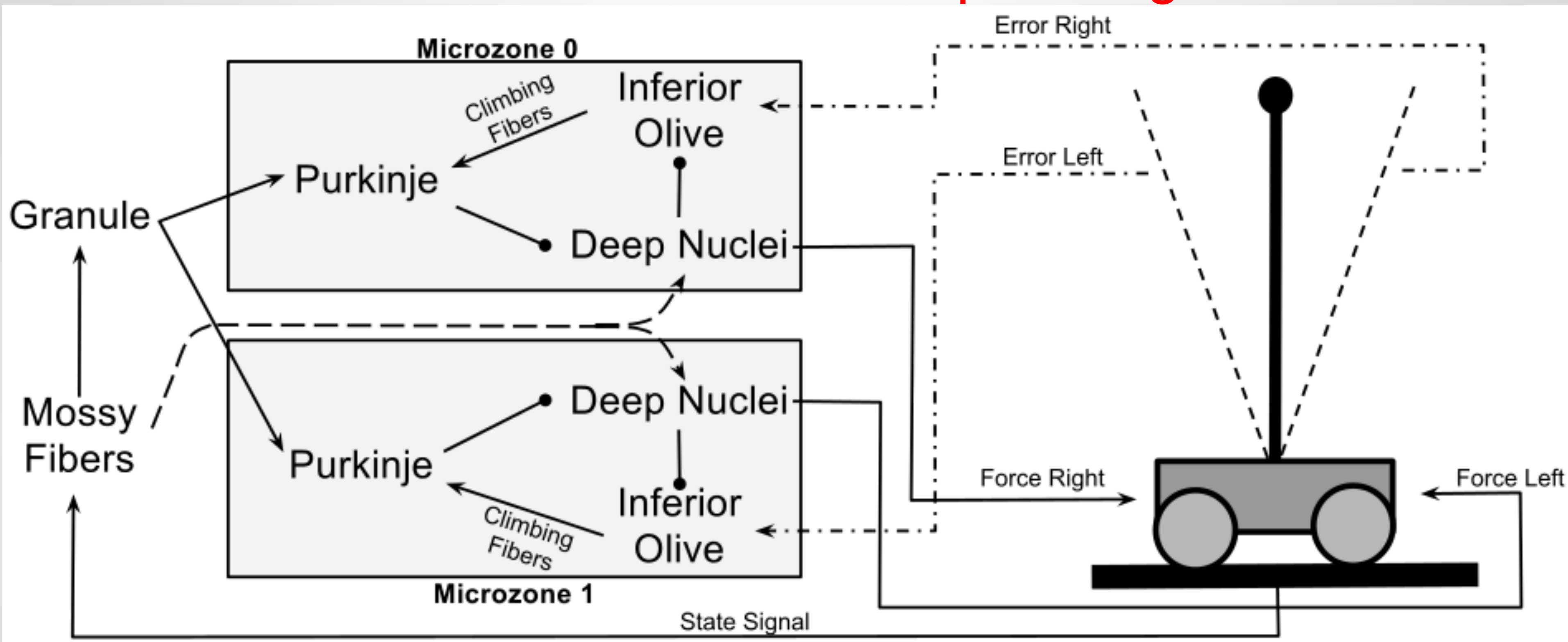
Full Cerebellum-Cartpole Interface

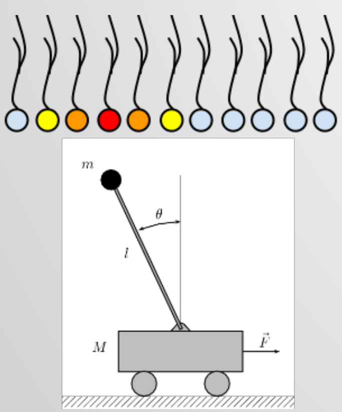


- Directional error signals are delivered to corresponding Microzones, encouraging greater force output.

Interface Summary

- Errors proportional to pole angle

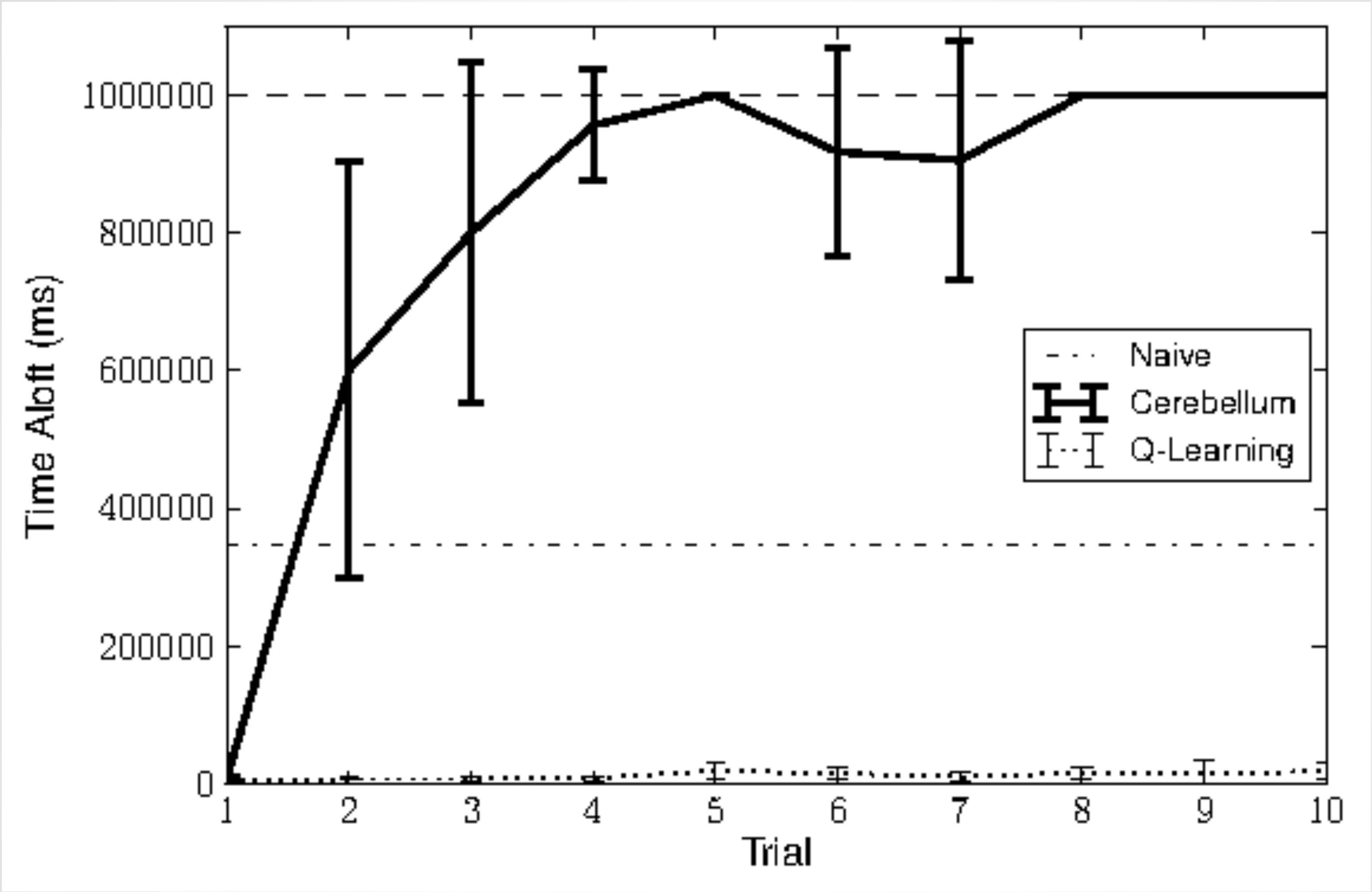




- Forces are real [0,1] values = NumFiringNC / 8.

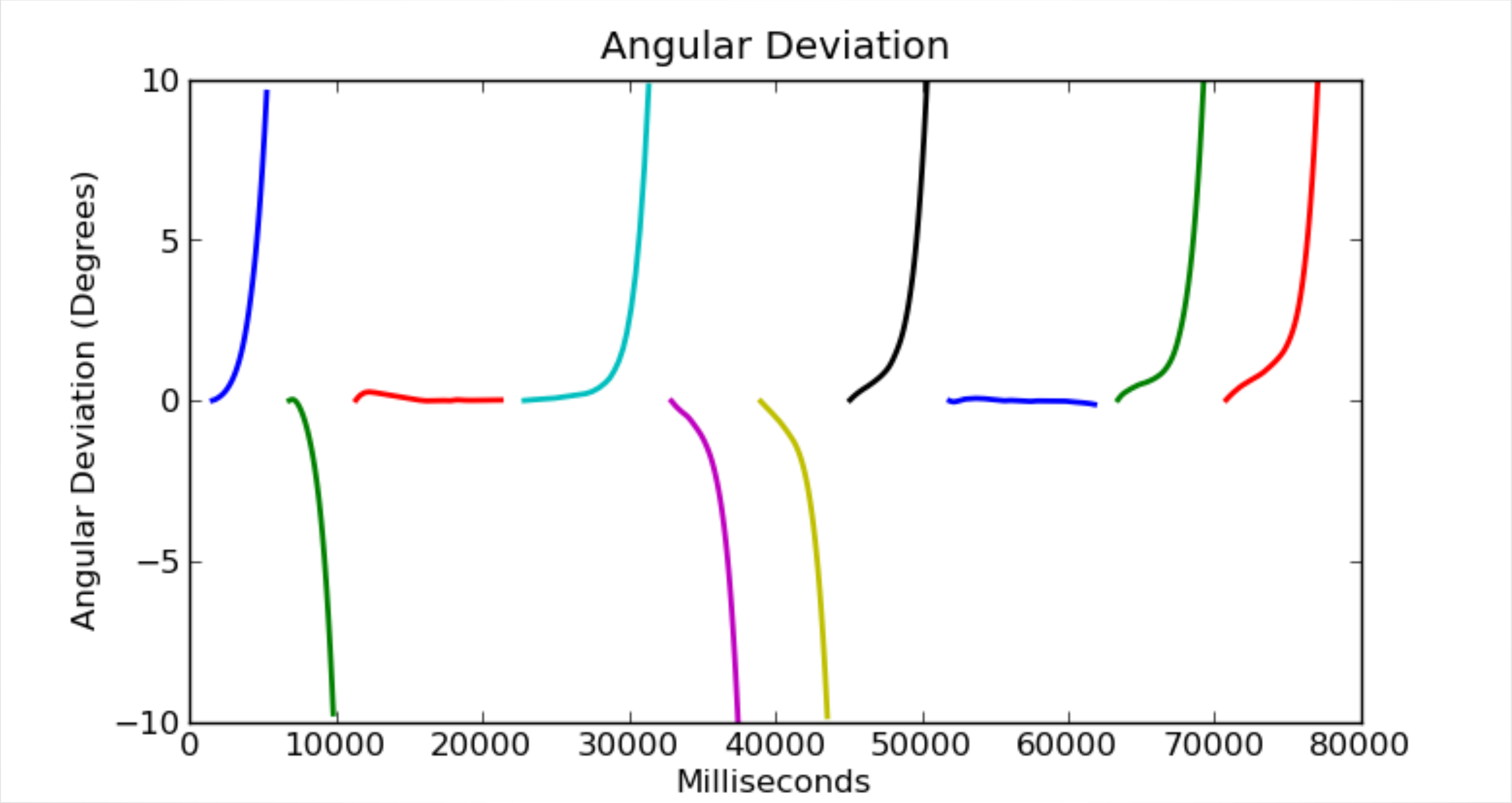
- Gaussian MF Encoding

Q-Learning Comparison



- Q-Learning uses same state & error encoding.
- Requires 1,000-10,000 trials before comparative performance is achieved.

Extinction



- Error signals delivered at end of trial result in cycles of learning & unlearning (extinction)
- Reliable performance requires regular error signals even if performance is good

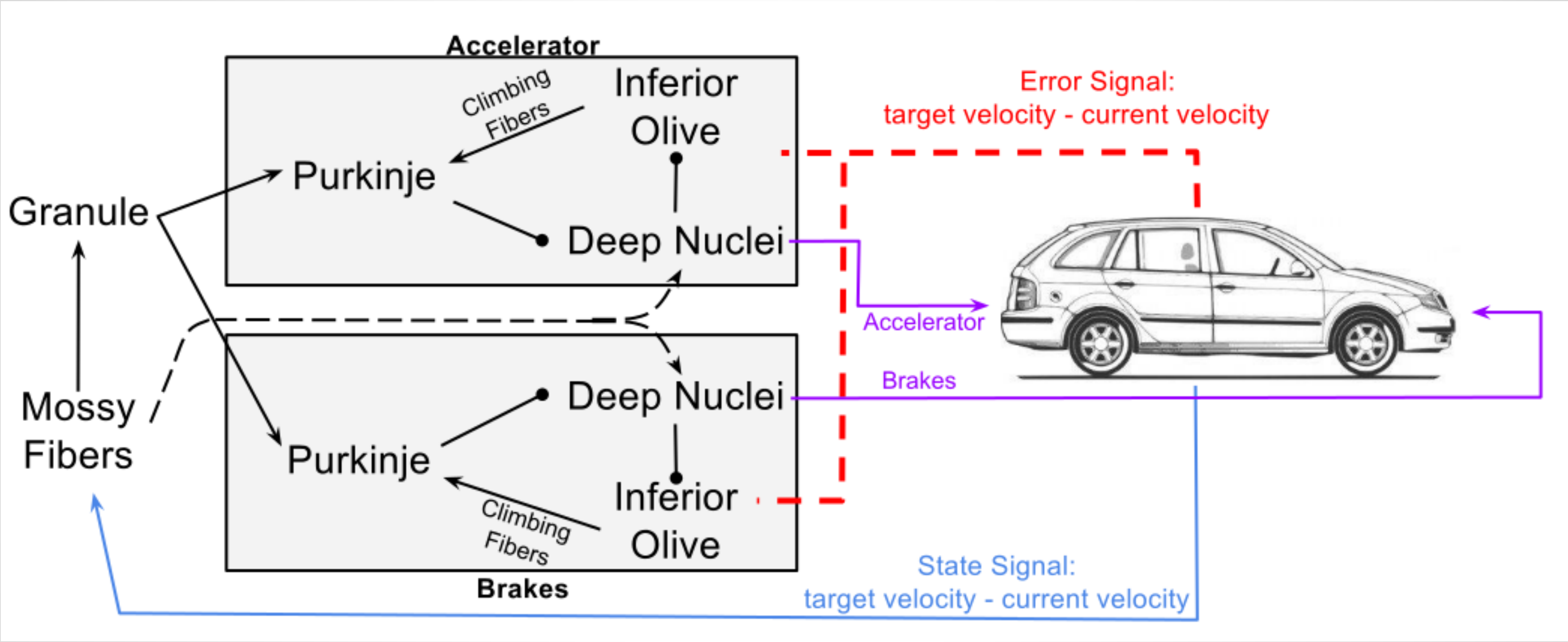
Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition
- Conclusions

PID Control

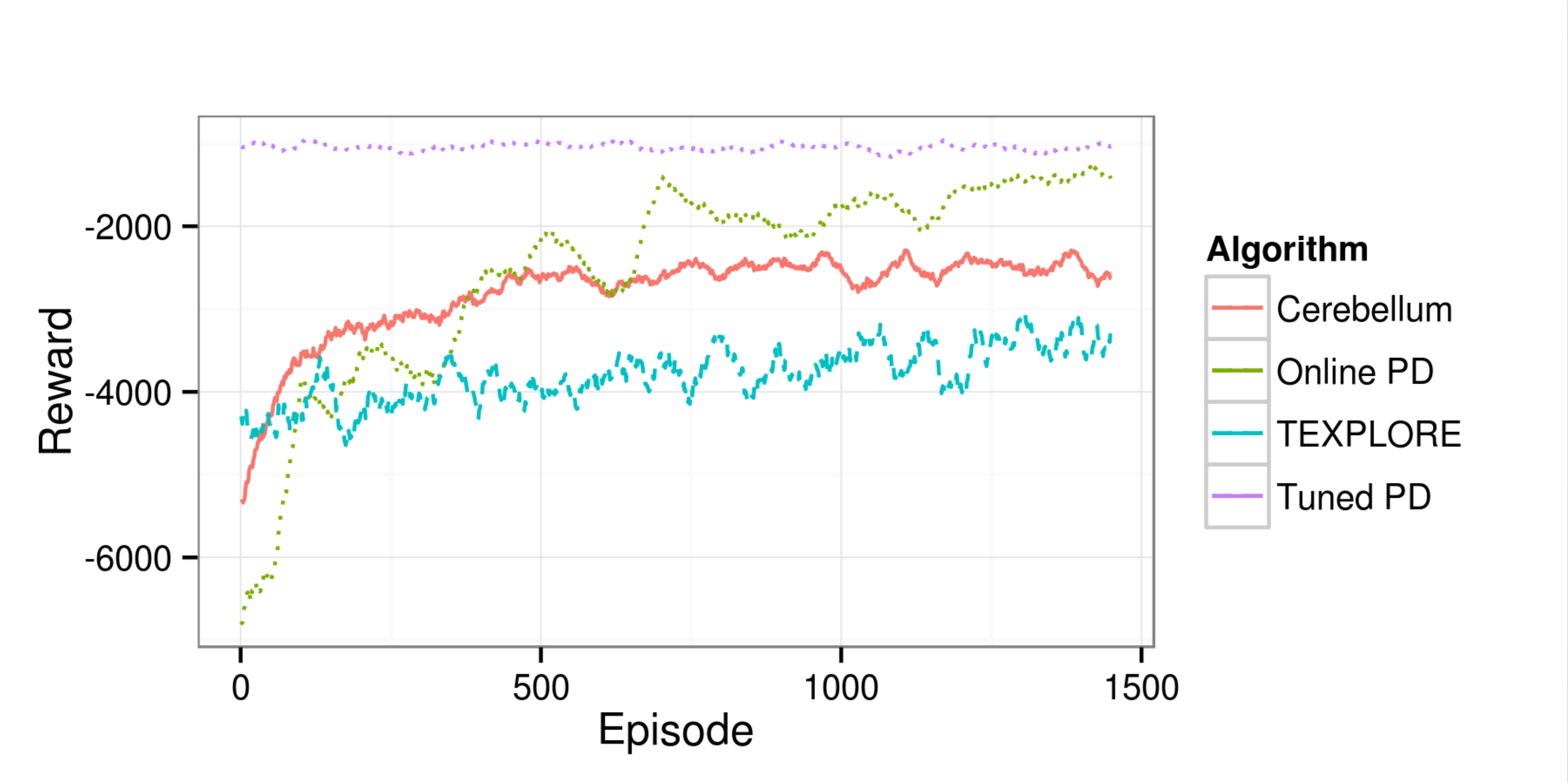
- Setpoint control generalizes the pendulum balancing domain (vertical setpoint)
- Typically setpoint control tasks solved by PID controllers
- Focus on simulated autonomous vehicle acceleration control

Velocity Control Architecture



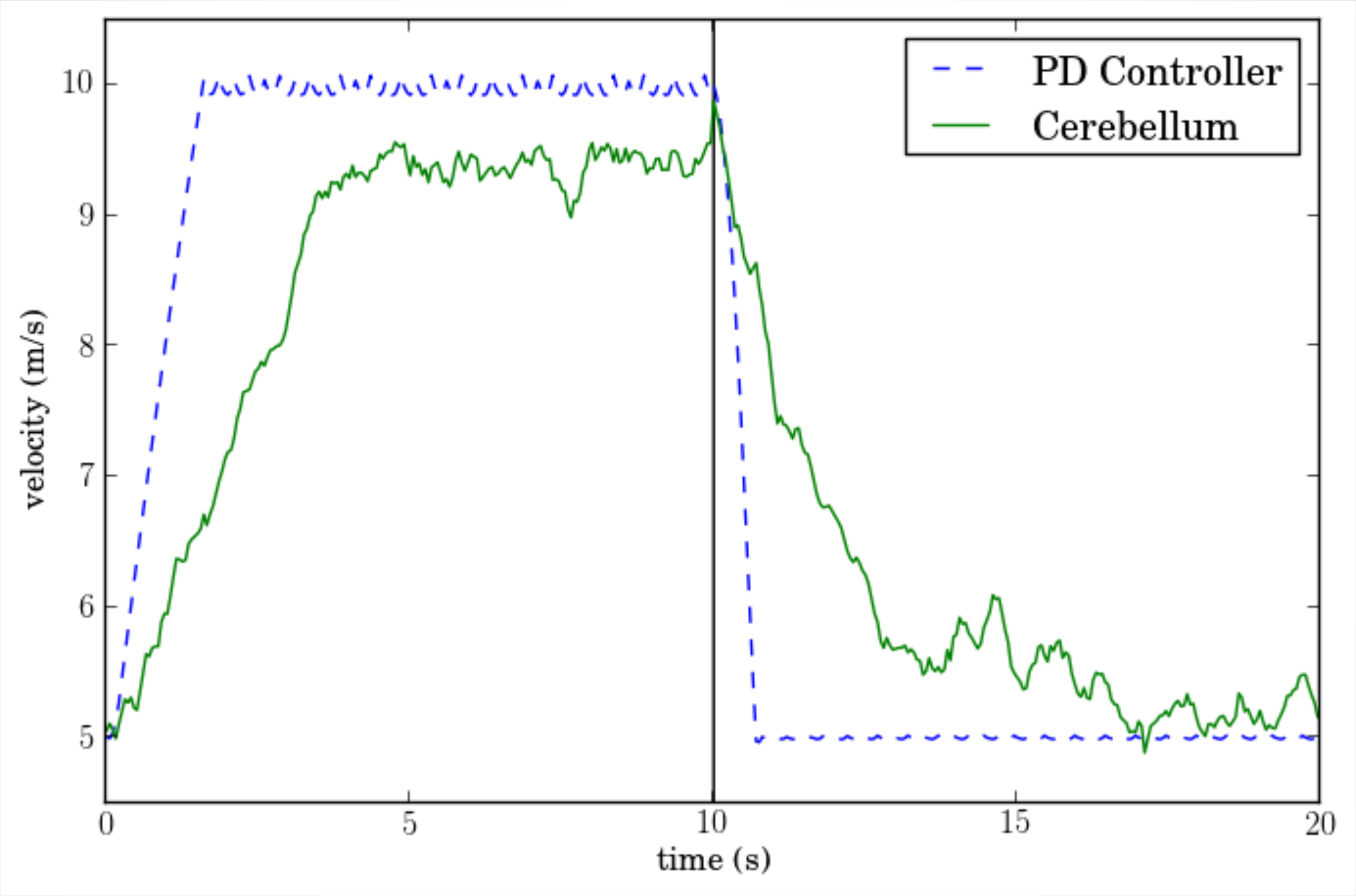
- Randomly generated current/target velocity in range [0,11] m/s
- Each trial lasts 10 seconds simulated time
- Reward = 10 * Sum(abs(target velocity - current velocity))

Velocity Control Results



Results averaged over 10 trials and smoothed with a 50 episode sliding window.

Velocity Control Analysis



Cerebellum is slower than PD controller to reach the target point.

Velocity Control Conclusions

- Cerebellum can perform PID/setpoint control tasks to some degree of precision

- These tasks feature supervised error signals which occur regularly

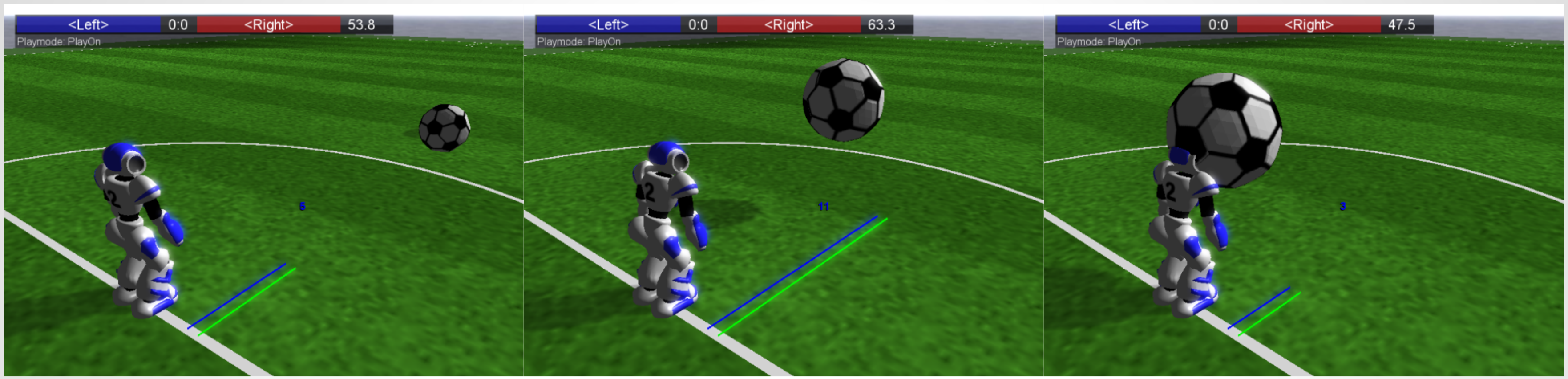
Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition
- Conclusions

Simulated Robocup Balance

- Domain: Robocup 3D Simulator
- Objective: Dynamic Balance
- Difference from previous domains: Delayed error signals

Task Specifics



- Large Soccer Ball - 10x mass, 6x size, 10m/s
- Objective: Don’t fall after impact!
- Control: Hip Joints - allow the robot to lean forwards & backwards
- Sensing: Timer counting down to the shot

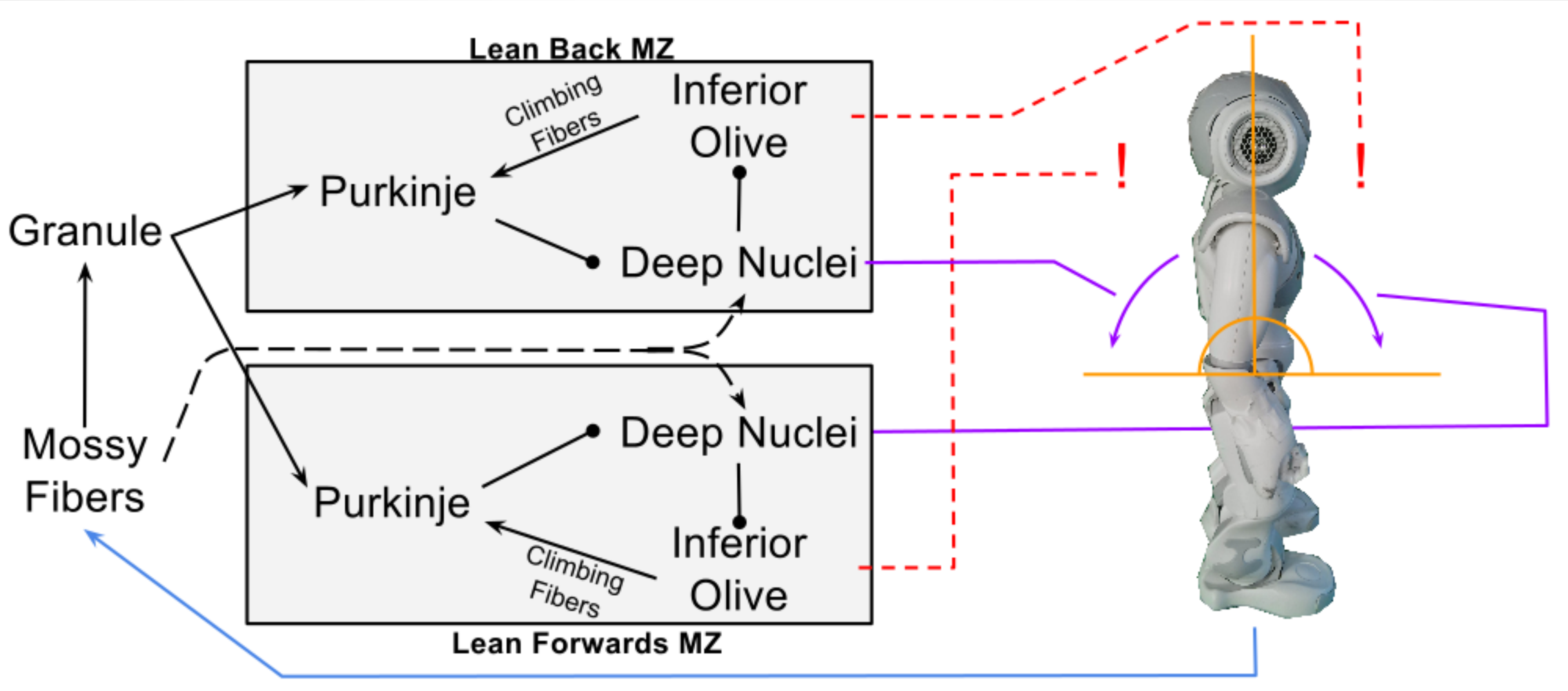
Complexity

- Task requires the robot to lean forwards in anticipation of impact, then lean backwards shortly thereafter.

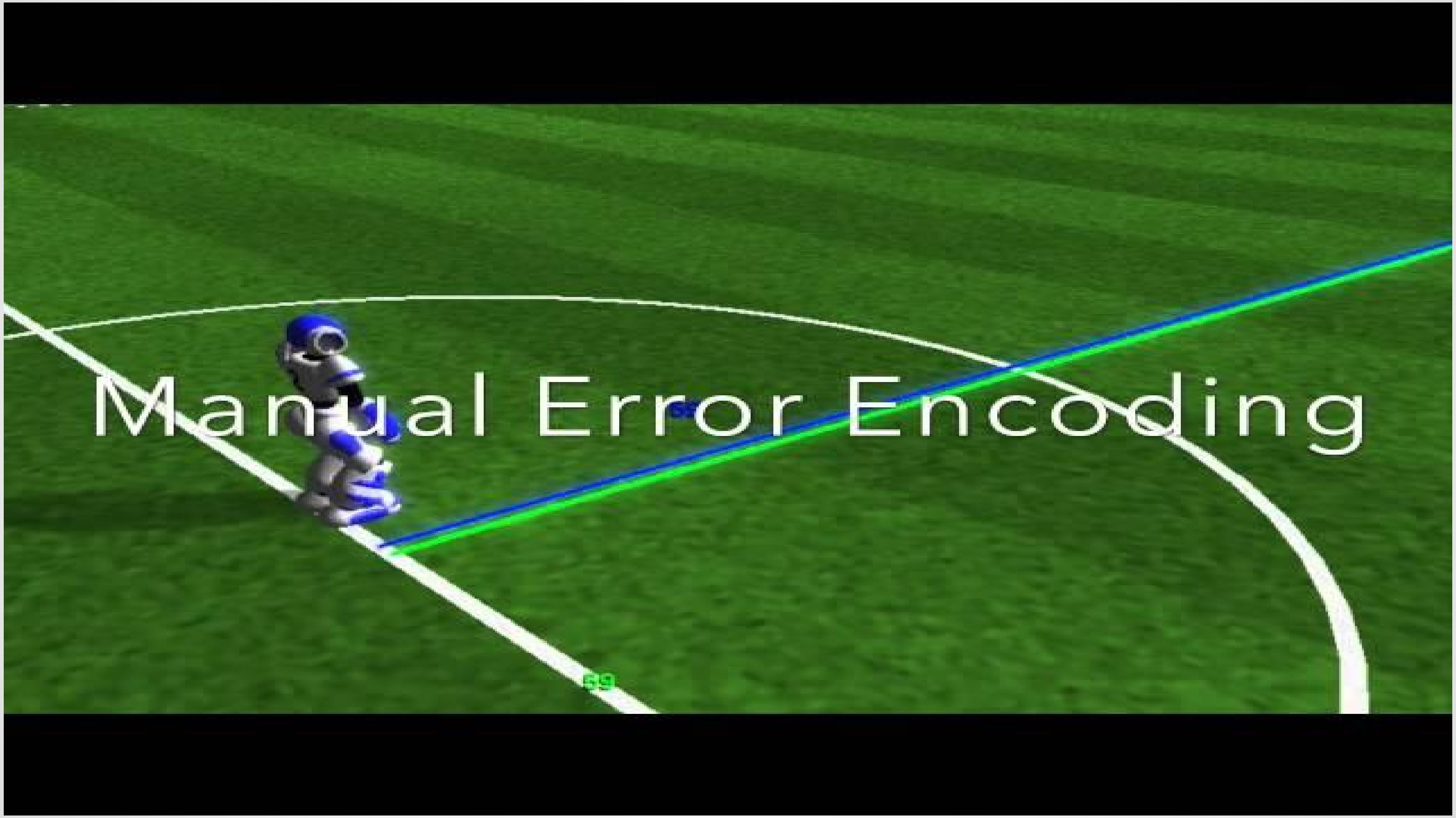
- Failure to do either will result in a fall.

- Simple policy can solve this task: Lean forwards .5 seconds before impact, then return to neutral.

Robocup Balance Architecture



- Experiments run with 3 different Error Signals:
  - Difference from known solution (Manual Encoding)
  - Gyroscope errors
  - Accelerometer errors


Manual Error Encoding

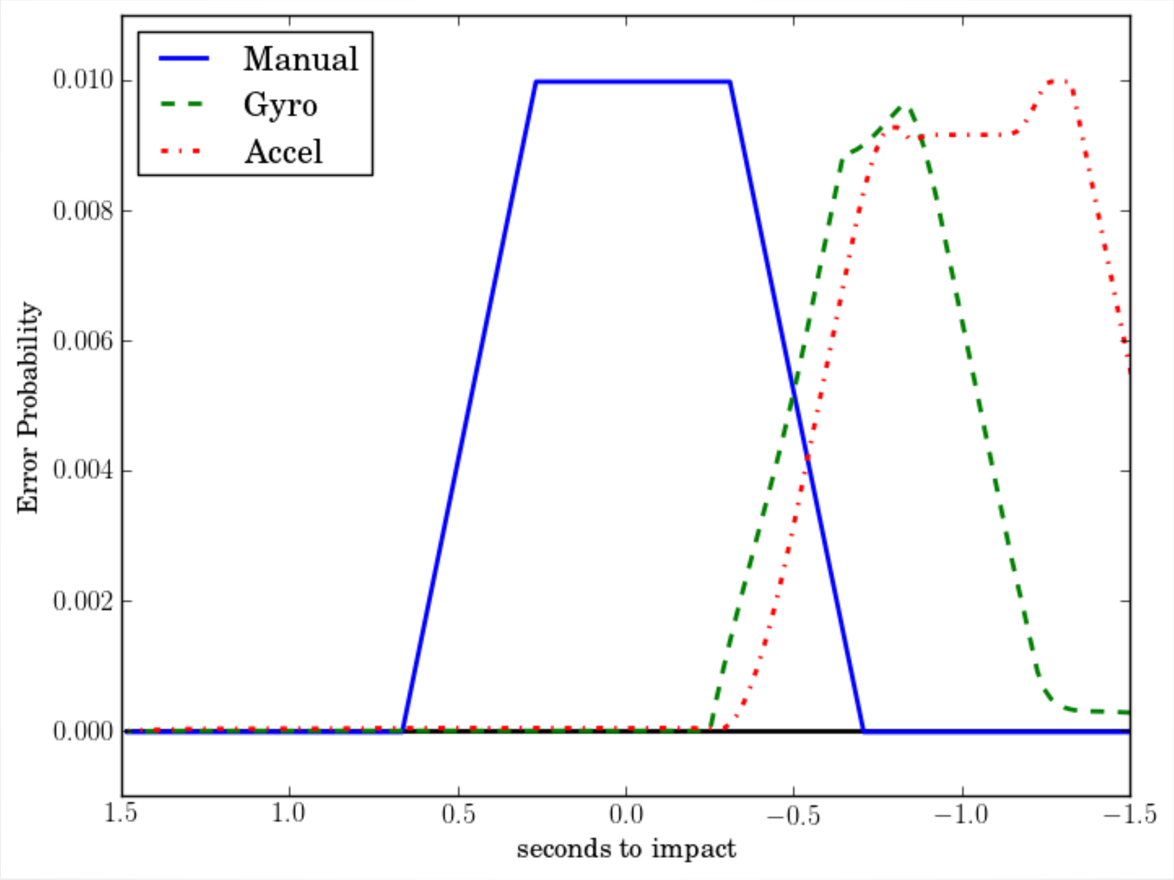
Balance Results

Error Encoding	Manual	Gyro	Accelerometer
No Fall	40.4%	.4%	2.4%
Fall Back	52.4%	95.2%	87.2%
Fall Forwards	7.2%	4.4%	10.4%

Experiments run up to 250 trials. Single run per result.

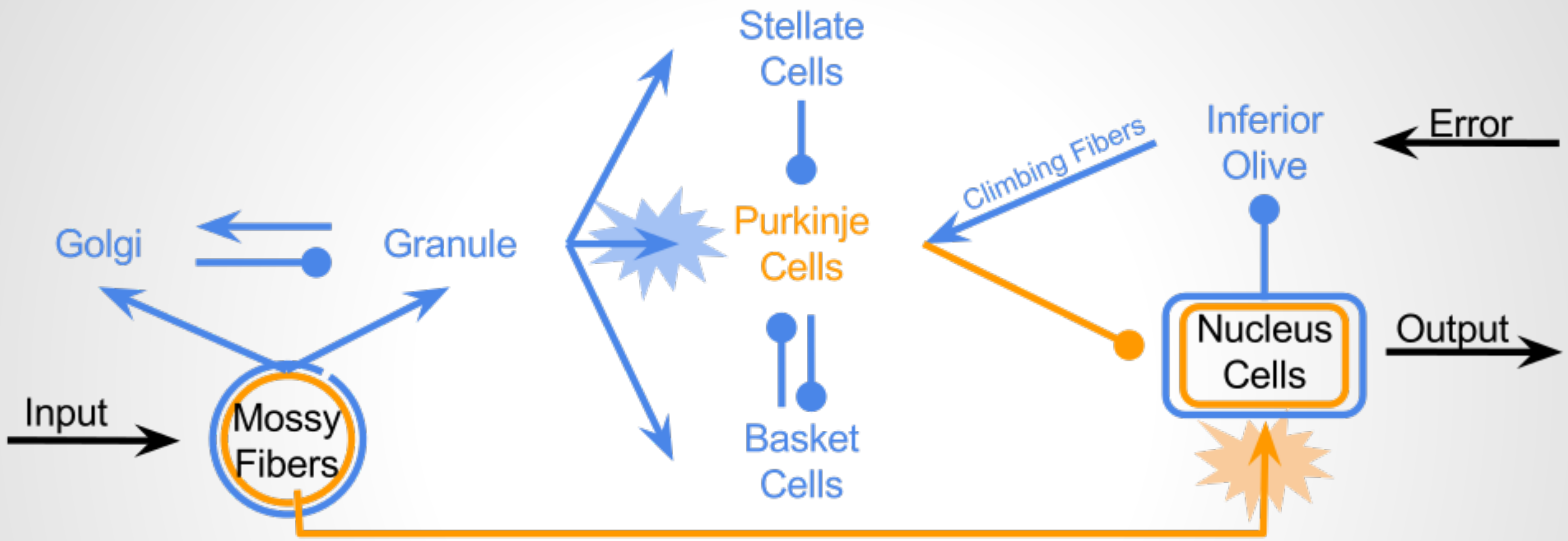
- Why do the Gyro and Accelerometer-based error signals perform so much worse than Manual?

Delayed Rewards



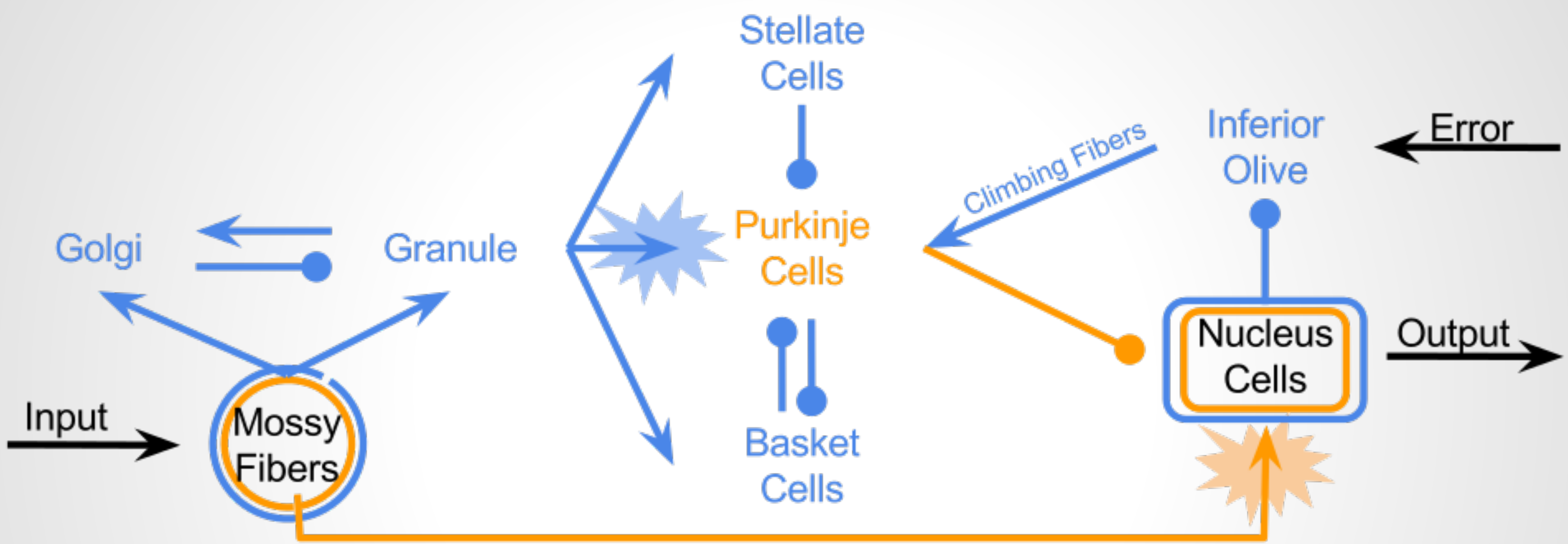
- How to analyze cerebellar learning with these different encodings?

Granule Weight Measure



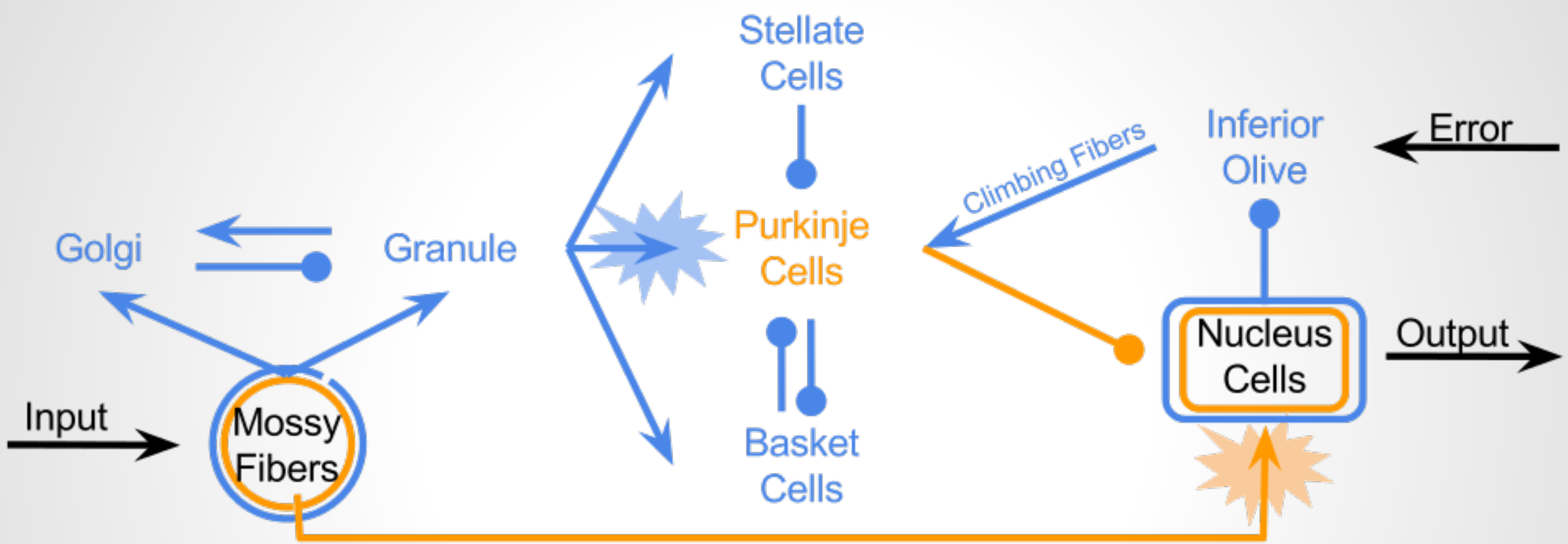
- Analyzes how each MF affects output forces by examining the weights of connected Granule Cells

Granule Weight Measure



- Each MF connected to 1024 Granule Cells
- Initial MF→GR Connection weights ~= 1
- Expected Sum Connected GR weights ~= 1000
- Weights change as the cerebellum learns

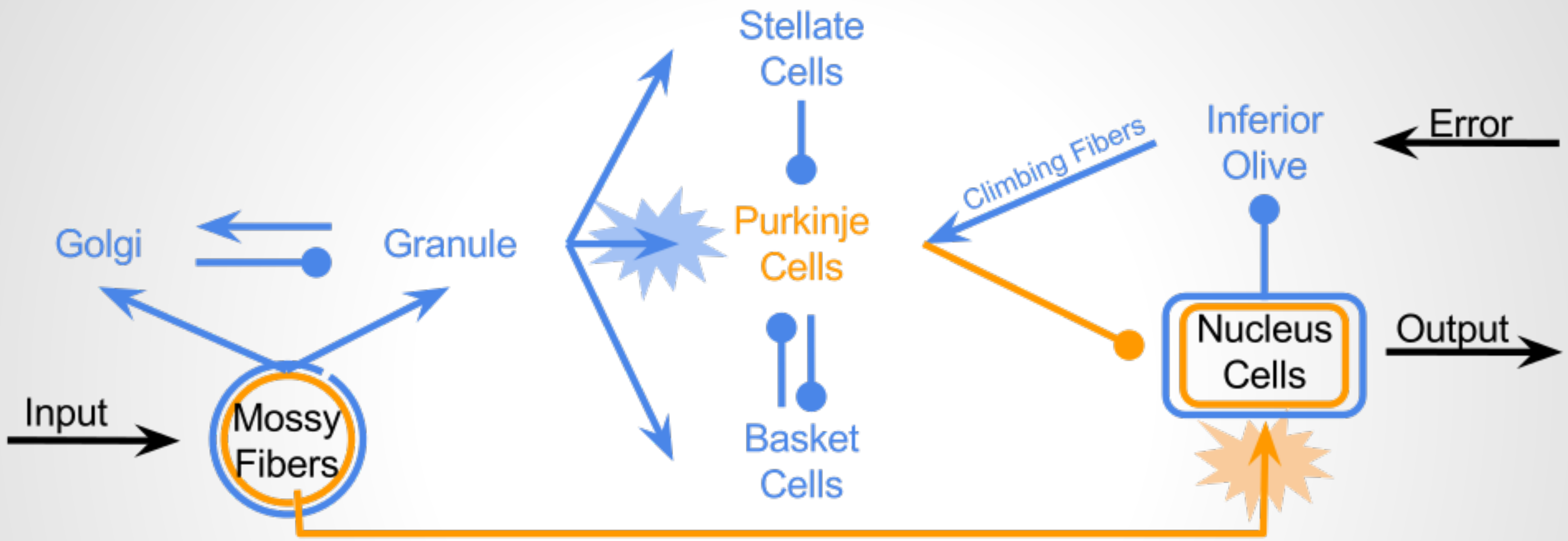
Granule Weight Measure



GWM (Mossy Fiber *m*) =

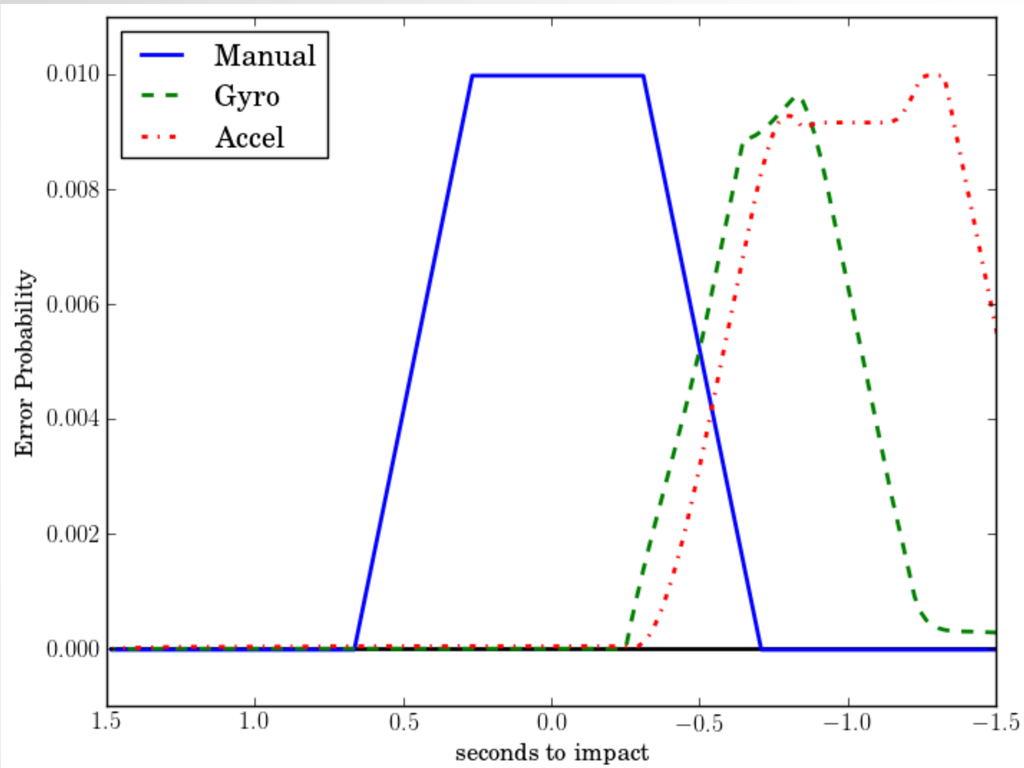
- Sum over connected granule cells *g*:
  - weight(g)
- Minus expected sum of granule weights (~1000)

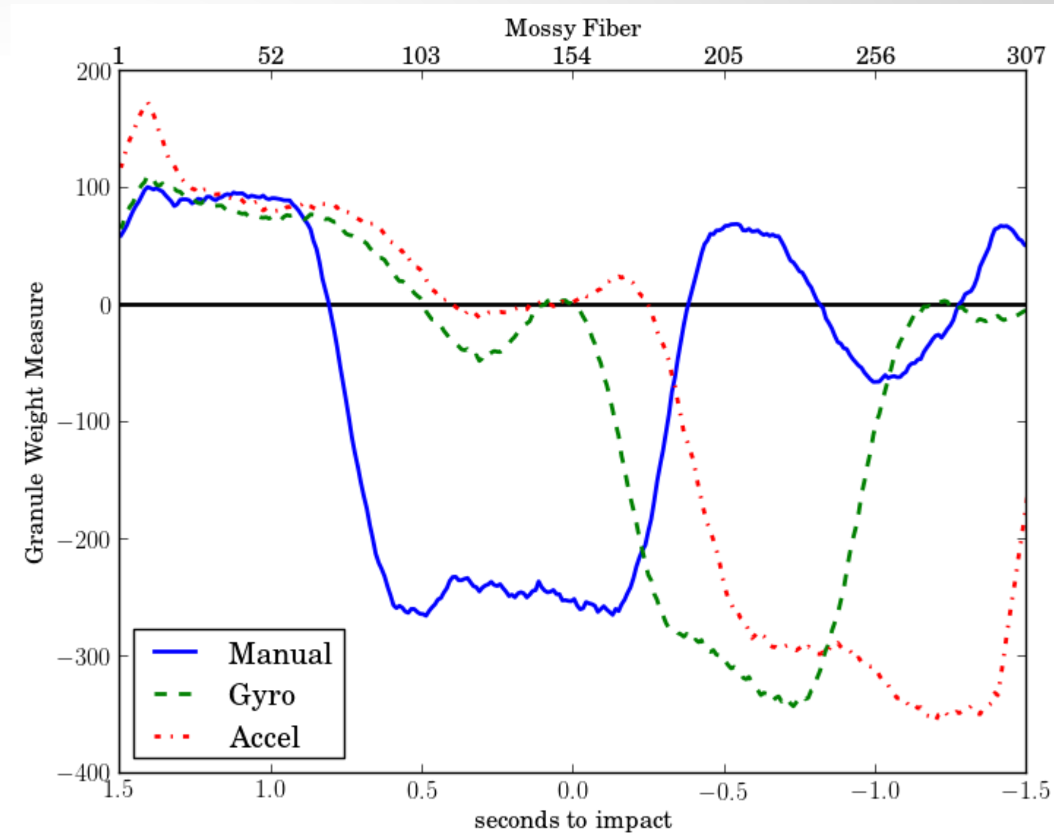
Granule Weight Measure



- High GWM indicates that whenever *m* is active, output will be low
- Low GWM predicts high cerebellar output forces for associated MF input *m*

Dynamic Balance Analysis





- GWM corresponds with error signal
- No temporal credit assignment!

Dynamic Balance Conclusions

- Simulated Cerebellar balance pretty shoddy
- Shouldn’t be this way… Something Missing?
- Cerebellum alone cannot perform credit assignment
- Cerebellum needs supervised error signals - it is not a Reinforcement Learner
- Basal Ganglia hypothesized to do RL

**Complementary roles of basal ganglia and cerebellum in learning and motor control.* Doya ‘00. Opinion in Neurobiology.

Outline

- Introduction: Biology of the cerebellum

- Cerebellum Simulator

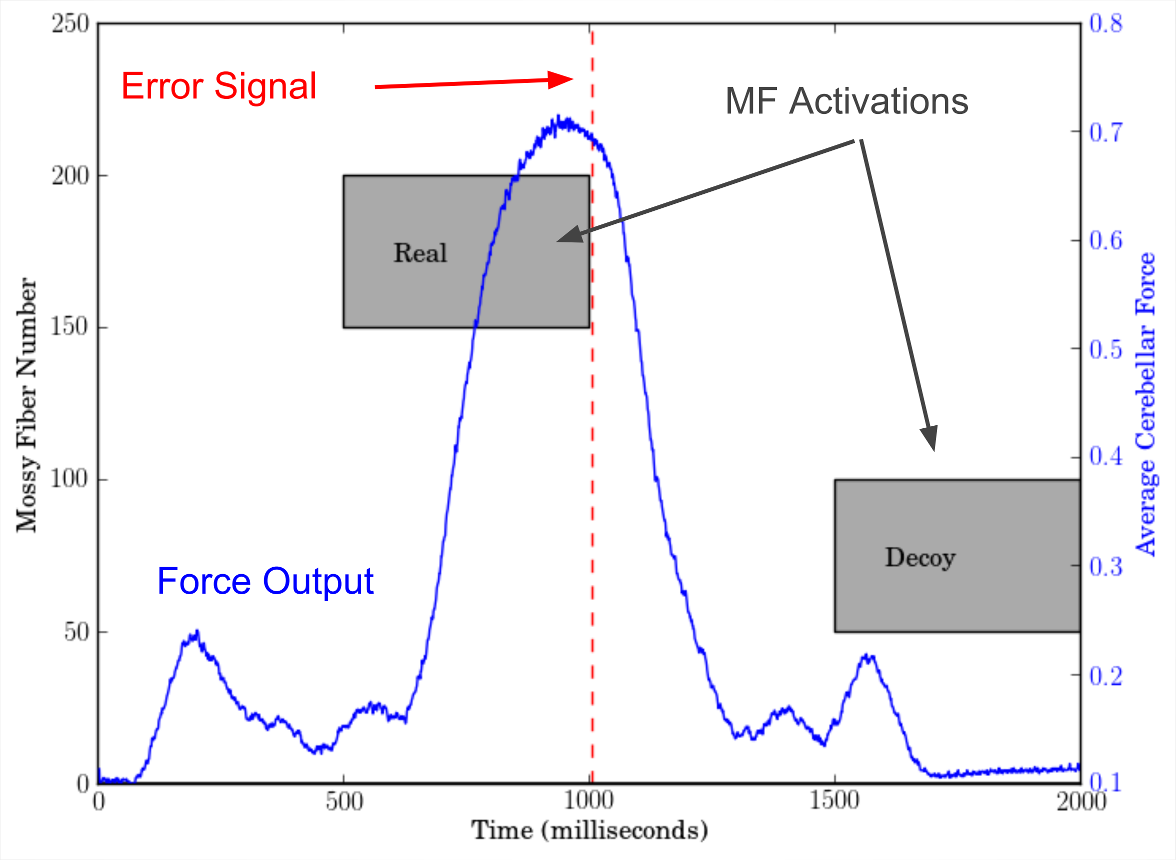
- Experimental Domains
  - Eyelid Conditioning
  - Cartpole
  - PID Control
  - Robocup Balance
  - Pattern Recognition
  - Audio Recognition

- Conclusions

Pattern Recognition

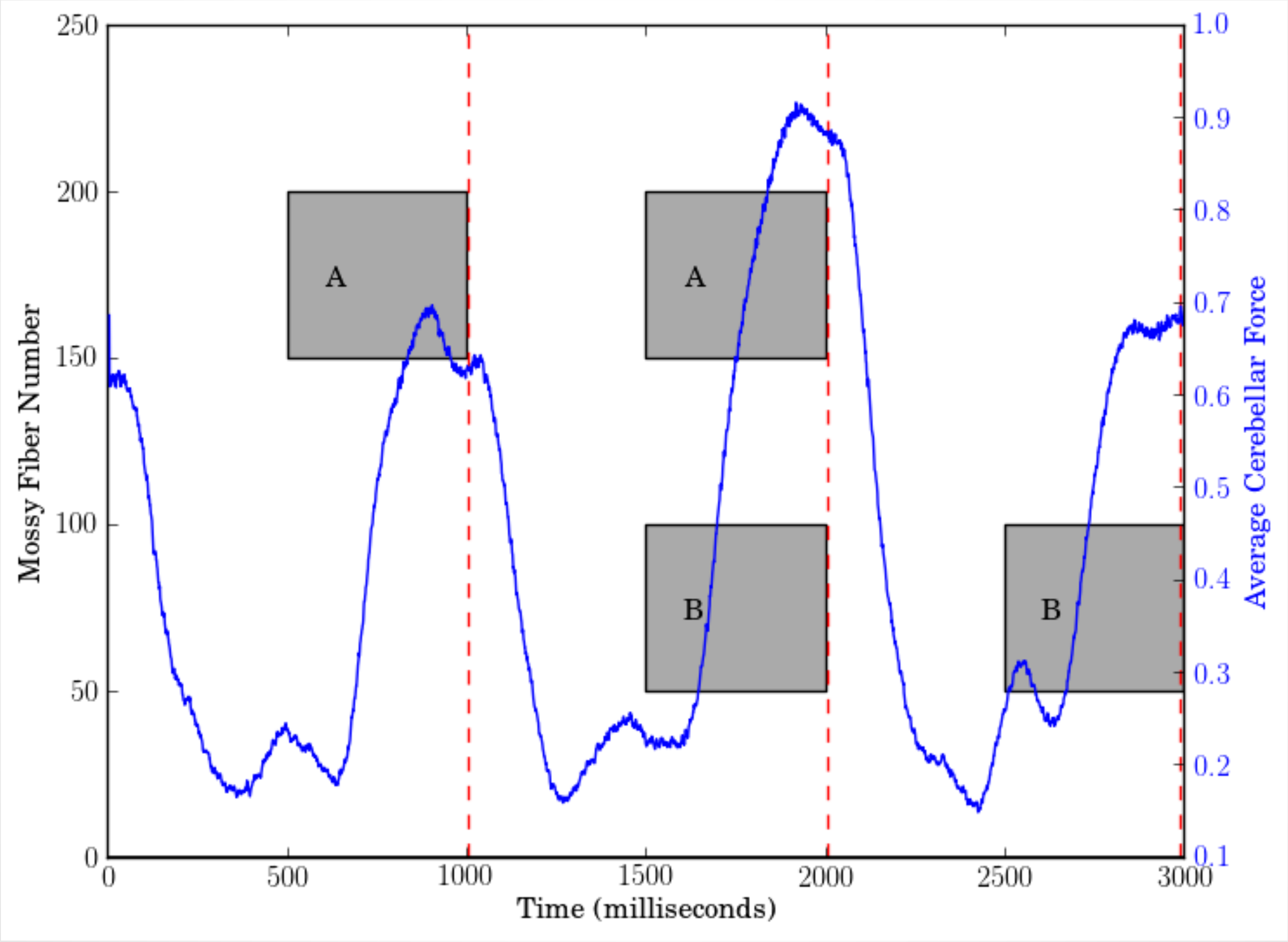
- Alright, the cerebellum is a supervised learner
- What types of patterns (functions) of state input can it identify?
- Start with static patterns and next move to temporal patterns

Static Pattern Recognition: Identity



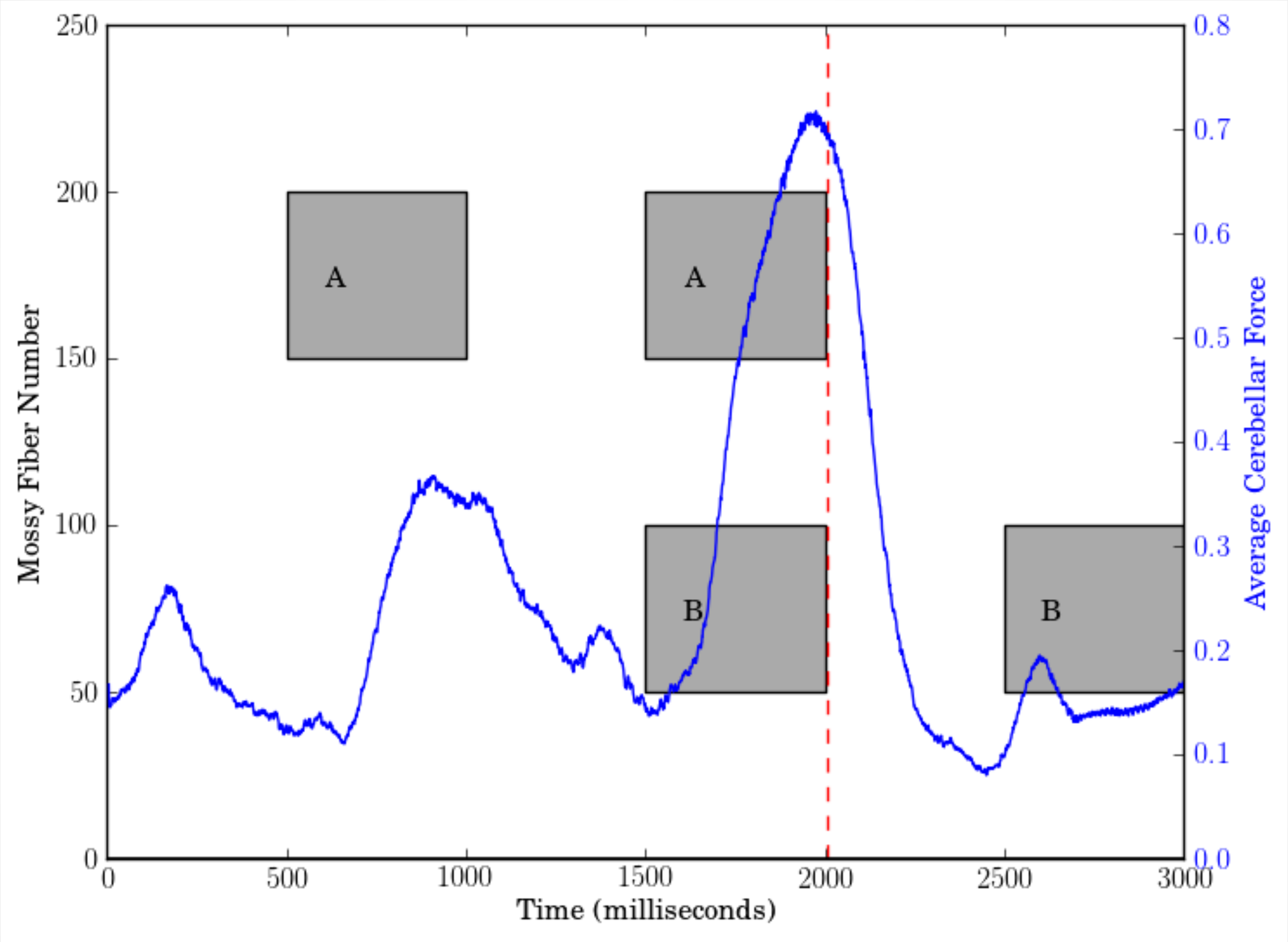
Objective: High force output preceding error signal(s)

Static Pattern Recognition: Disjunction



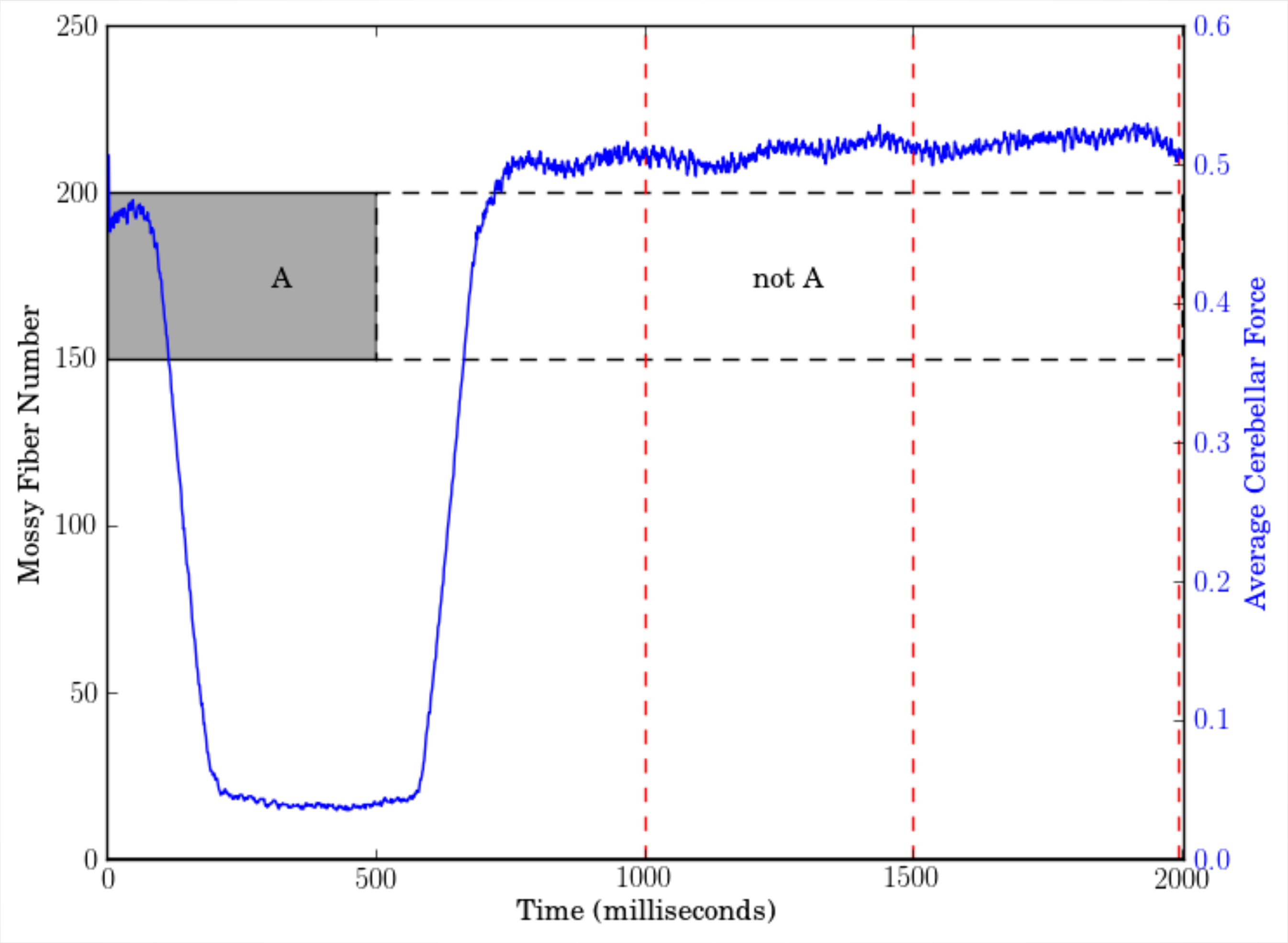
Successfully Recognized

Static Pattern Recognition: Conjunction



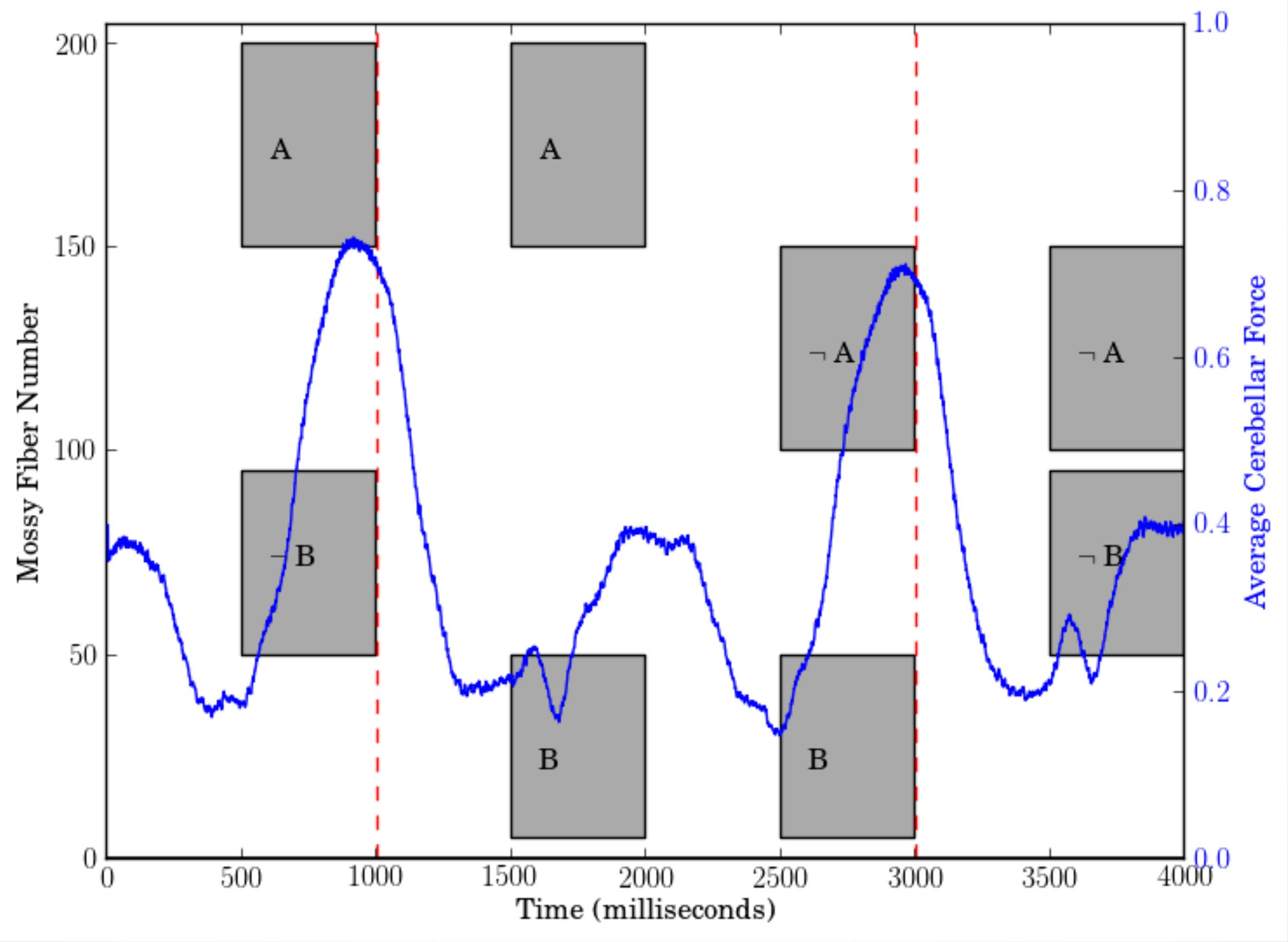
Successfully Recognized

Static Pattern Recognition: Negation



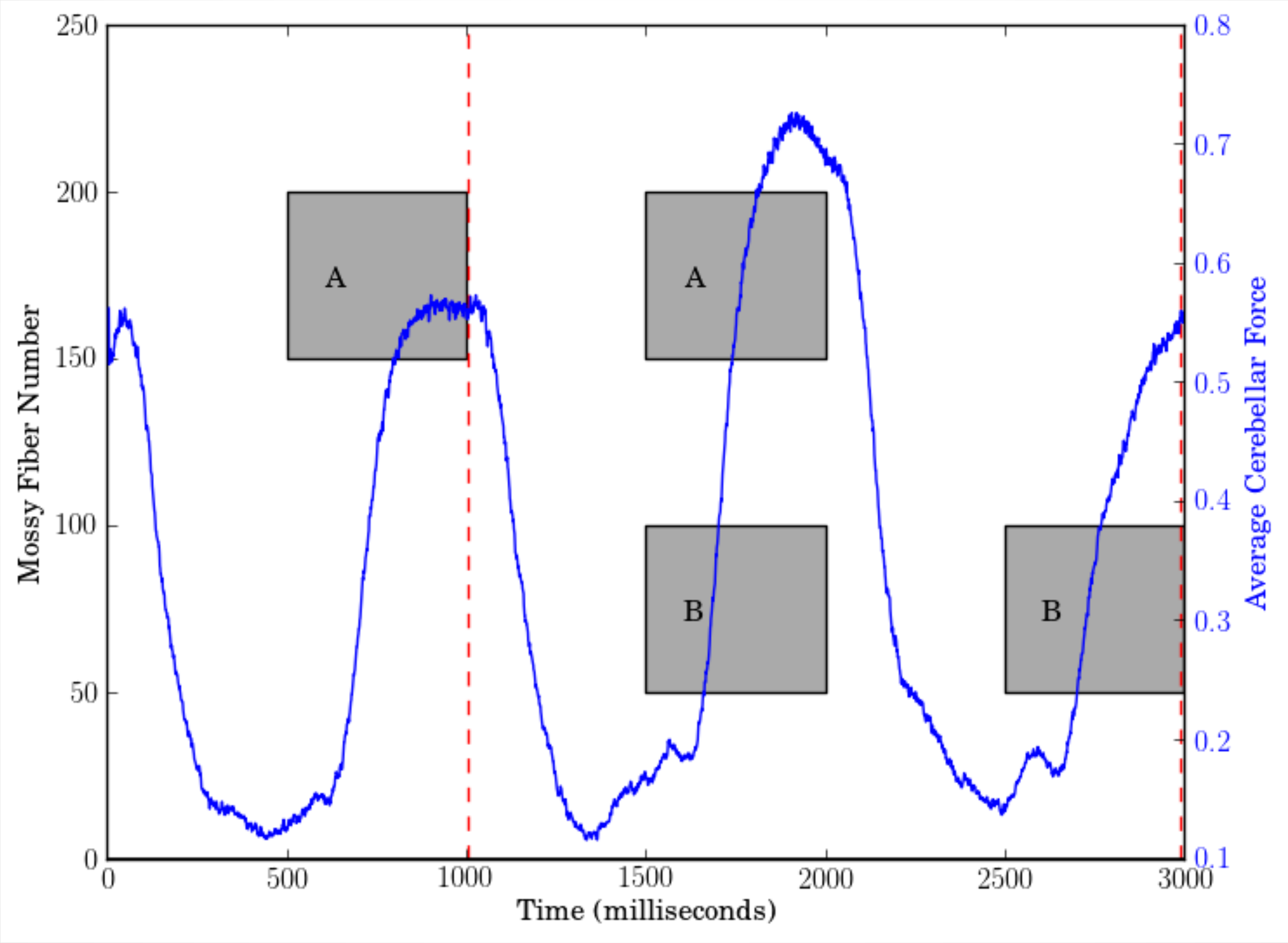
Successfully Recognized

Static Pattern Recognition: XOR



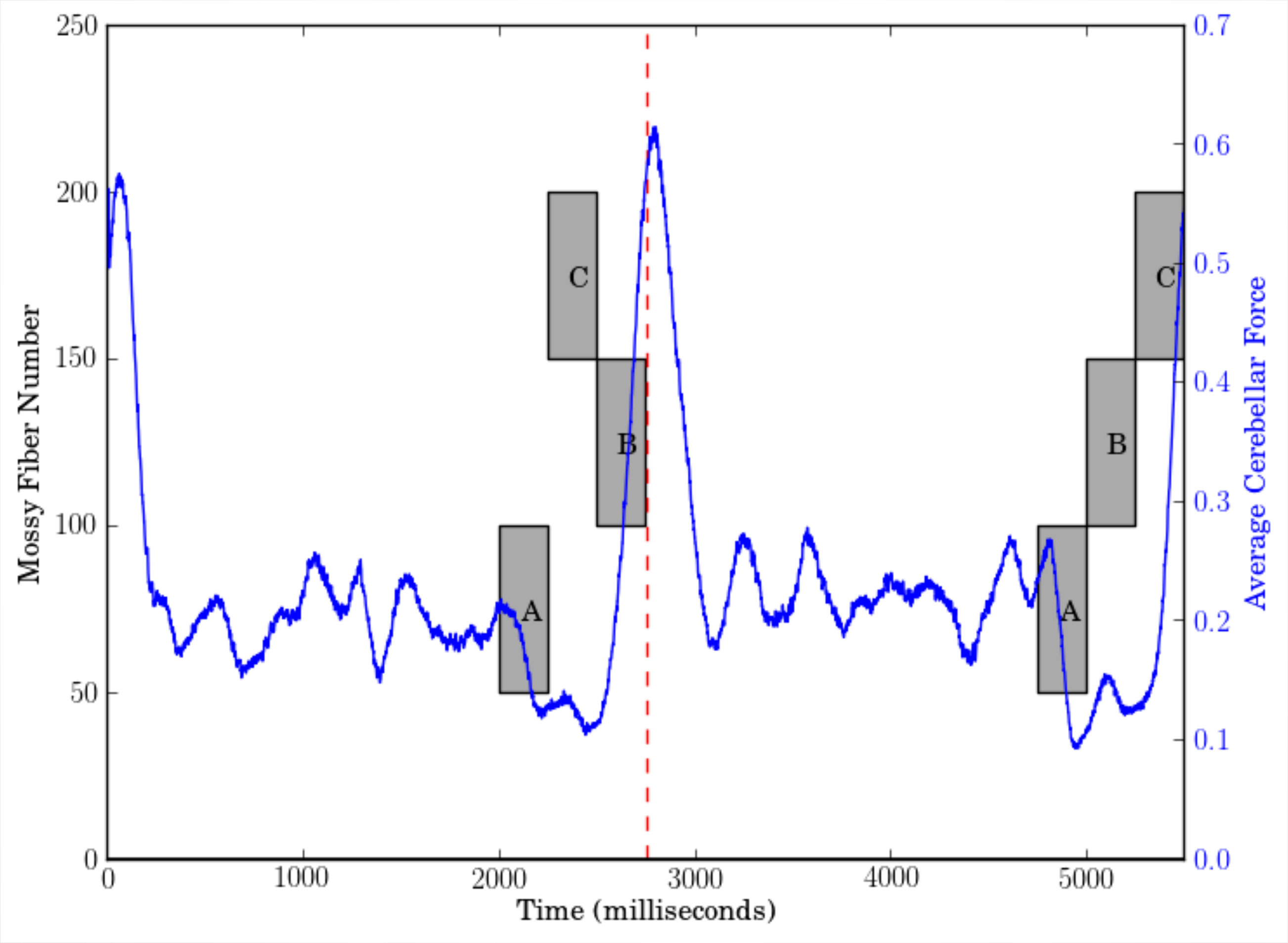
Successfully Recognized

Static Pattern Recognition: NAND



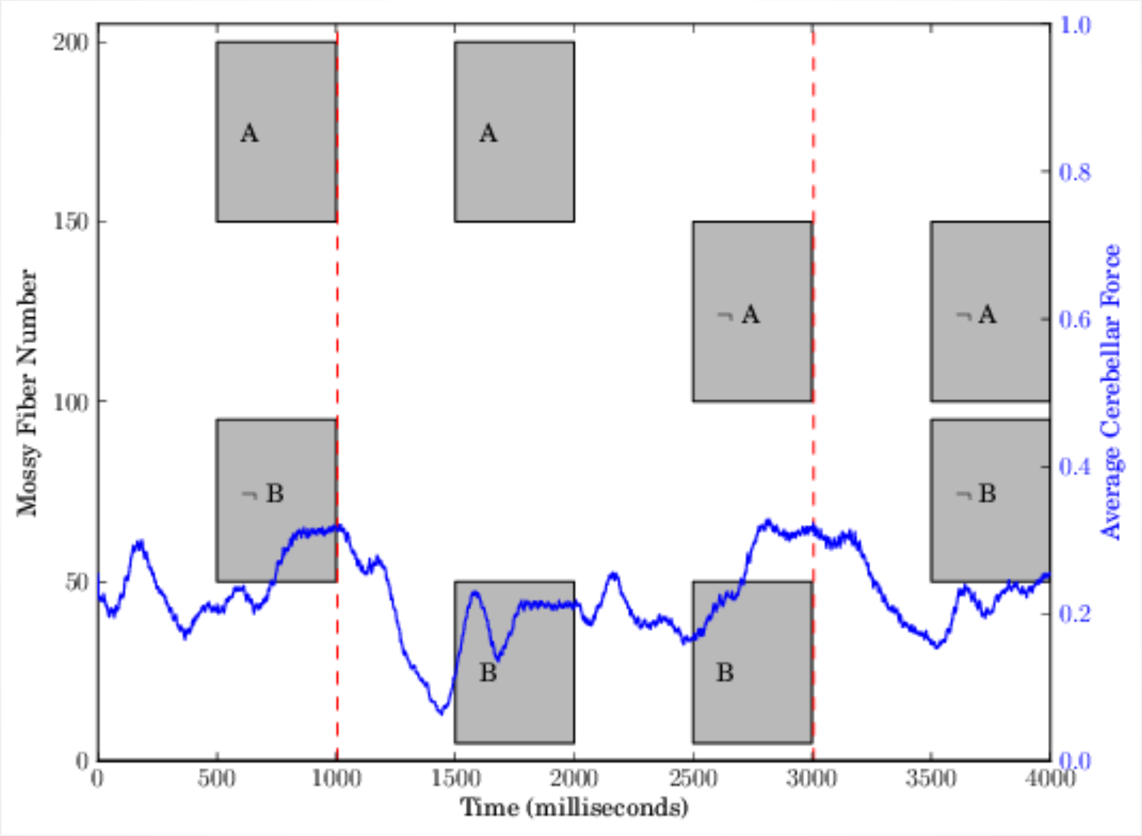
Not Recognized

Temporal Pattern Recognition



Not Recognized

Alternating XOR



When tones are played in alternating timesteps, recognition is lost

Pattern Recognition Conclusions

- Cerebellum can recognize all boolean functions of 1-2 variables except NAND

- Temporal pattern recognition is extremely limited

Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
    - Eyelid Conditioning
    - Cartpole
    - PID Control
    - Robocup Balance
    - Pattern Recognition
    - Audio Recognition
- Conclusions

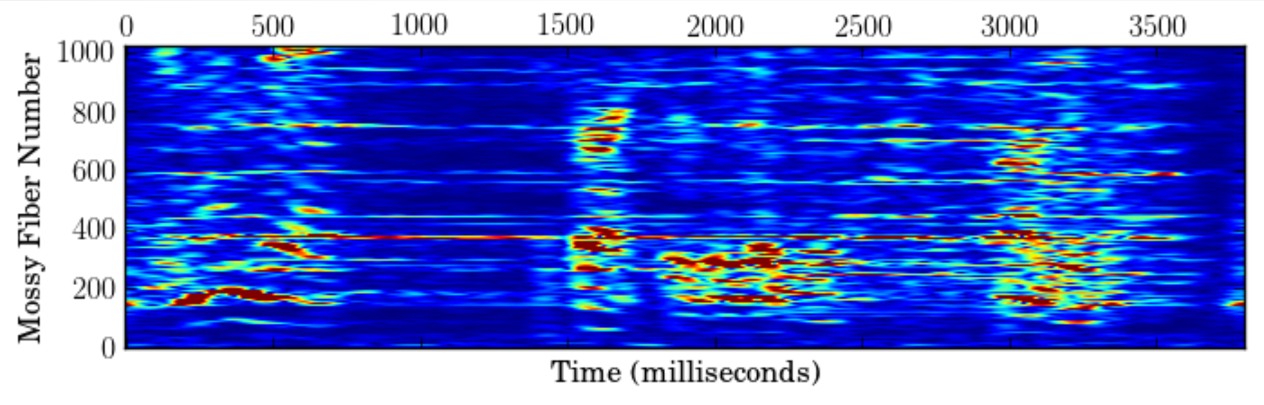
Audio Recognition

- Test cerebellum’s pattern recognition capabilities in a real world domain

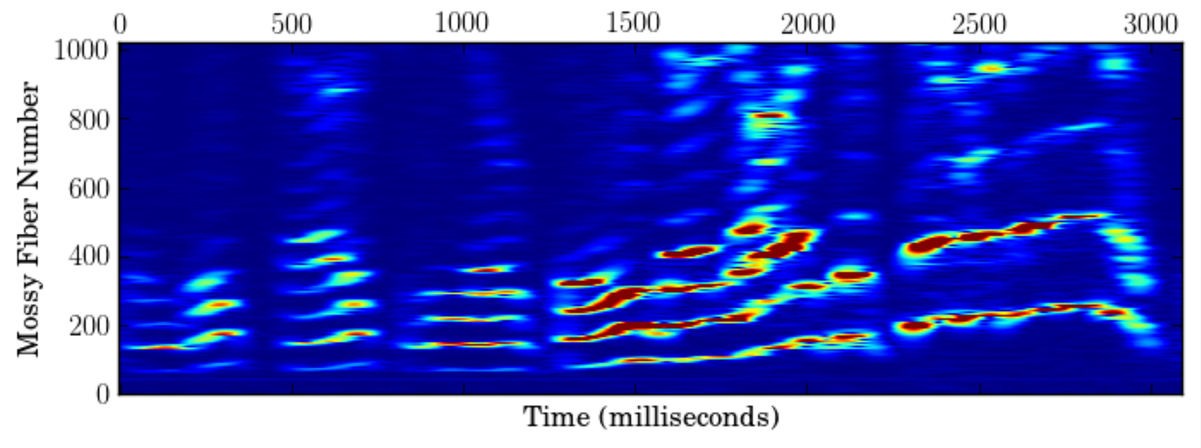
- Objective: distinguish between two different audio clips

- Clips are transformed by FFT and then converted to MF activations

Audio Preparation



Force: "The force will be with you, always." - Obi Wan Kenobi



Thermo: "In this house we obey the laws of thermodynamics!" - Homer Simpson

Training

- Audio clips were played in alternation

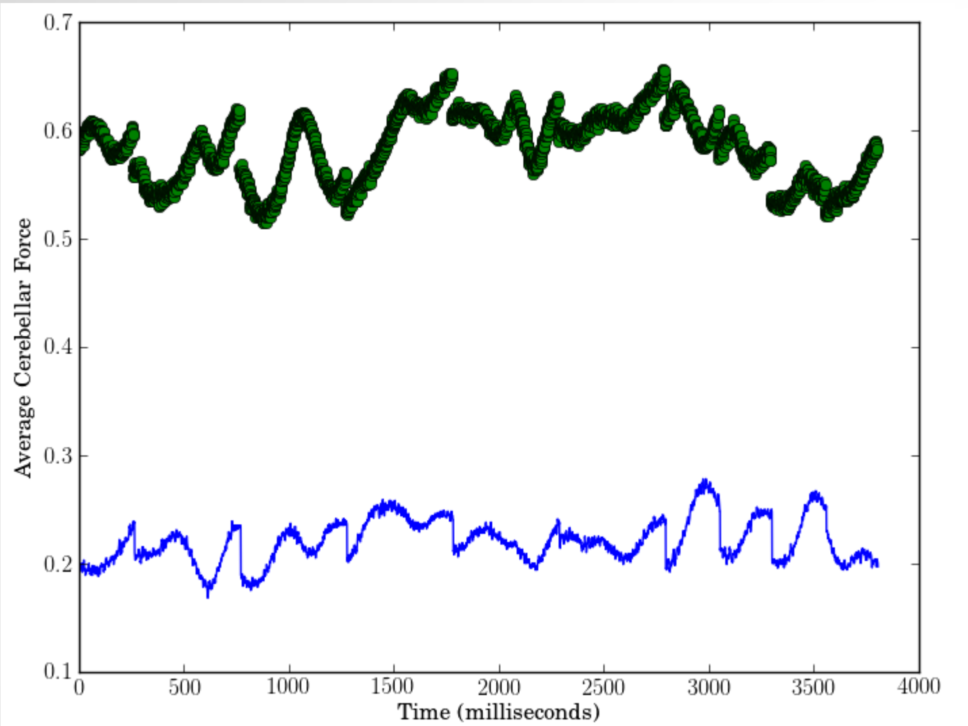
- Two Microzones trained - one to recognize each different clip

- Training: While a clip is playing, the associated MZ gets periodic error signals

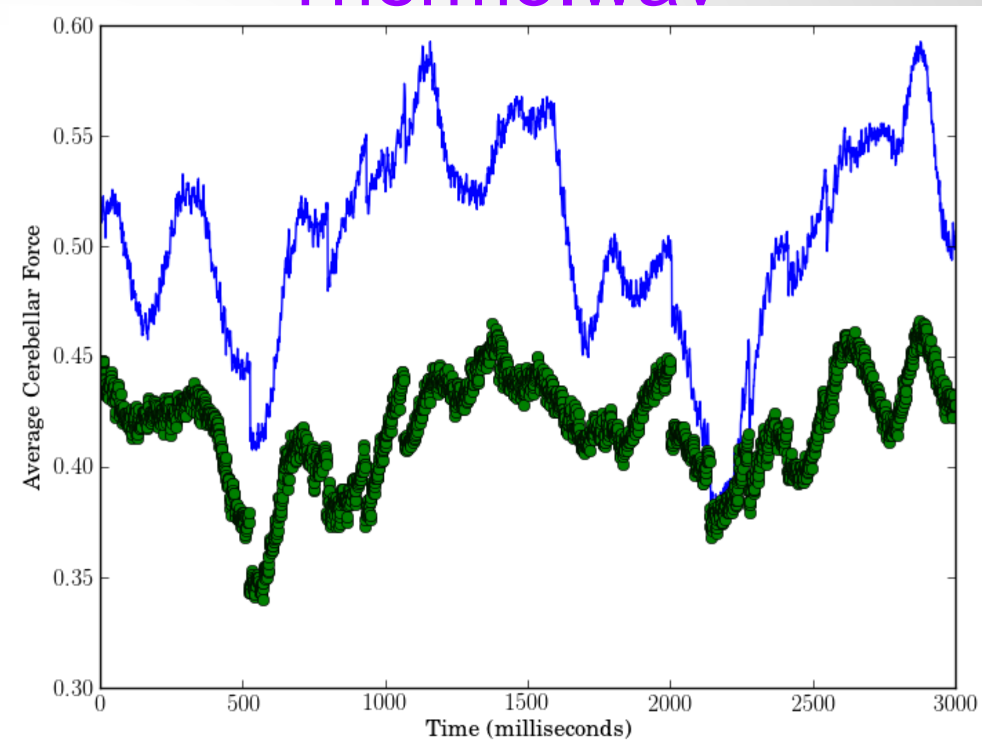
- Test: A clip is played back and the associated MZ should exhibit high force output

Audio Recognition Results

Force.wav



Thermo.wav

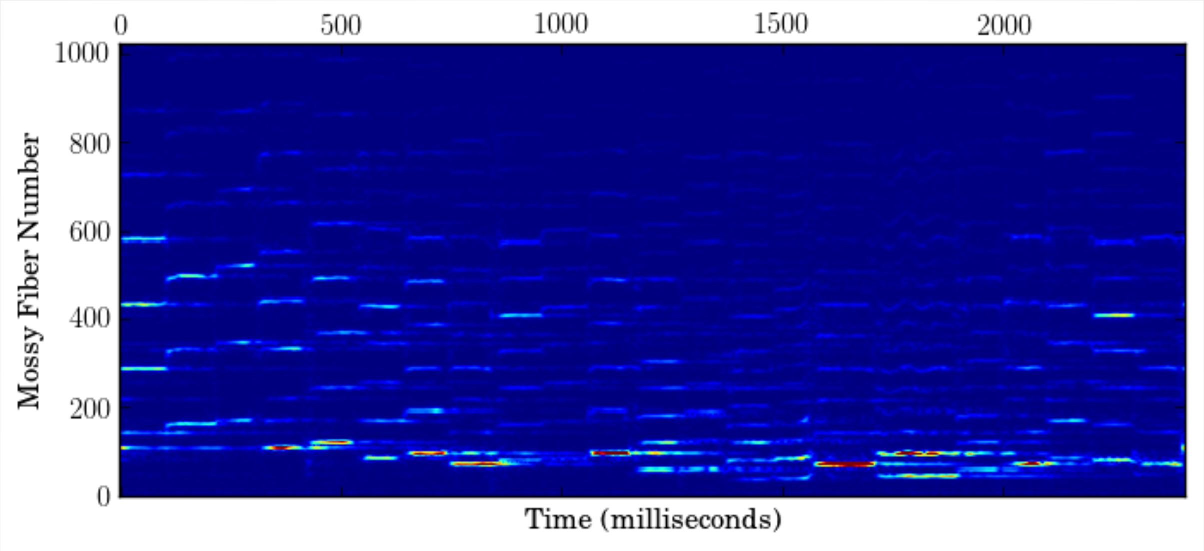


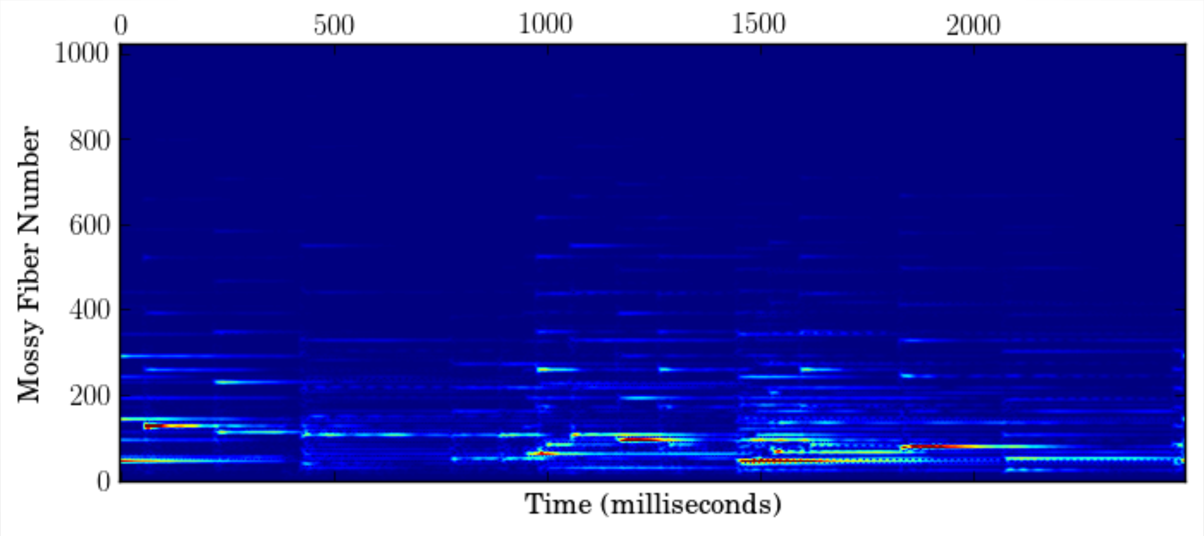
Green: Output from MZ trained on Force Clip

Blue: Output from MZ trained on Thermo Clip

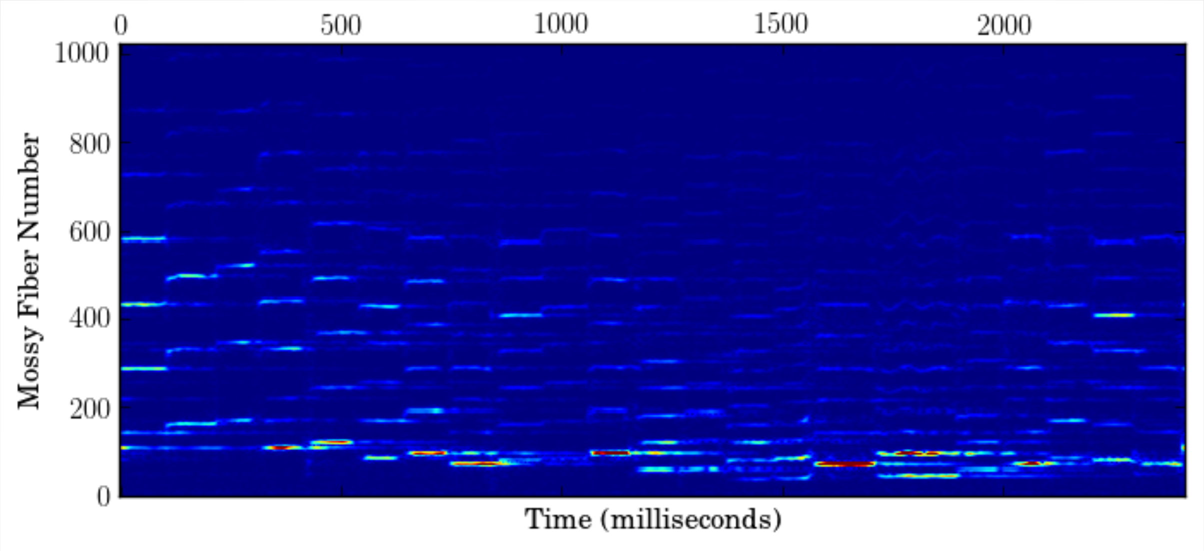
Conclusion: Successful recognition!

Can you identify piano/violin?

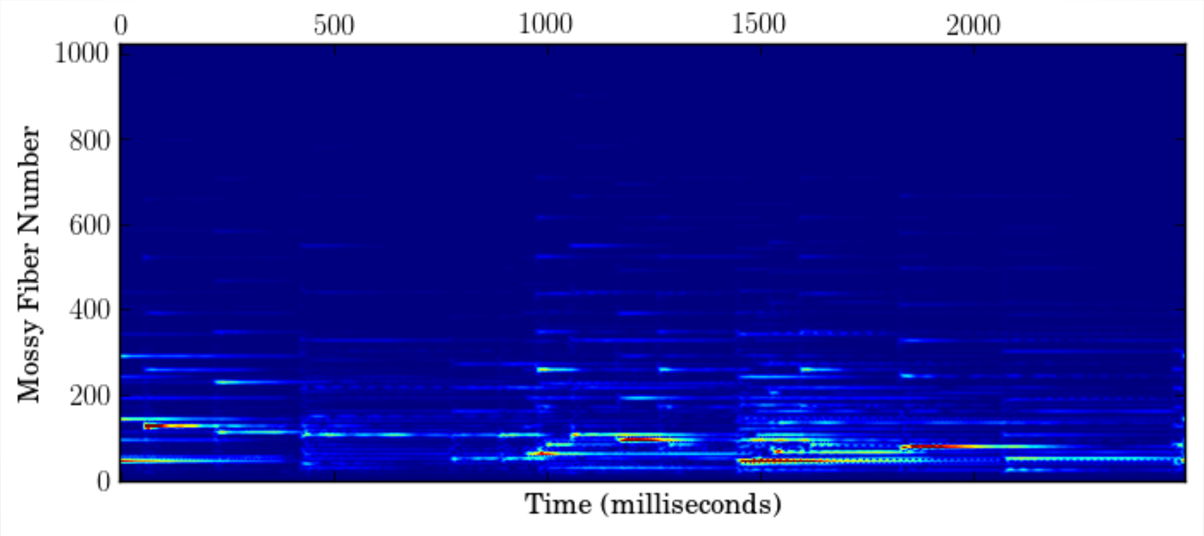




Harder Audio Recognition



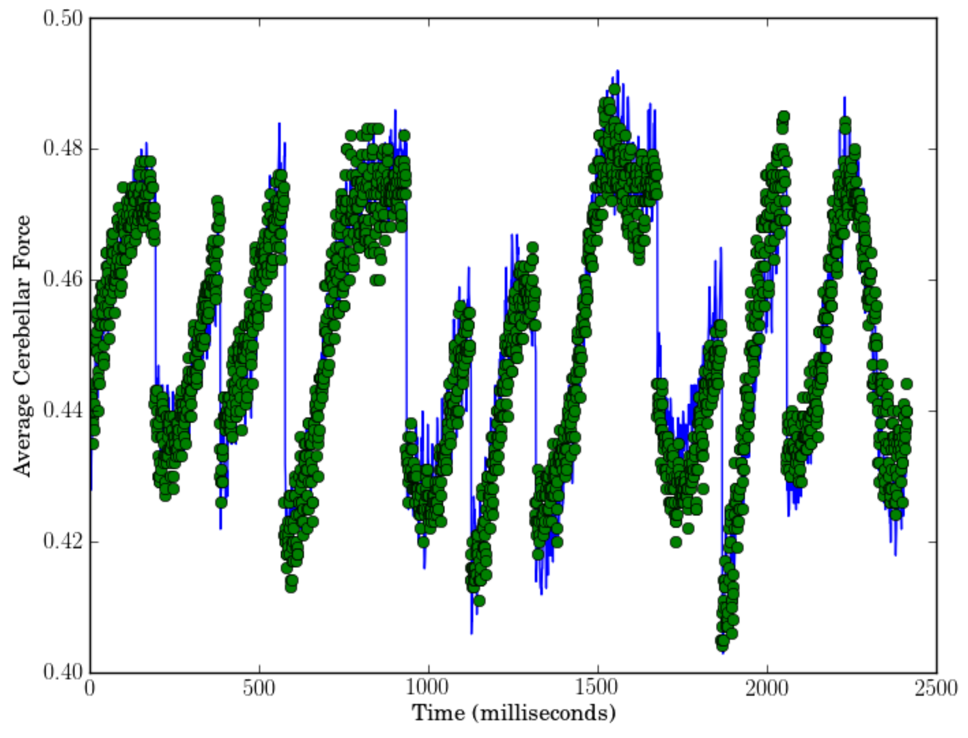
Violin



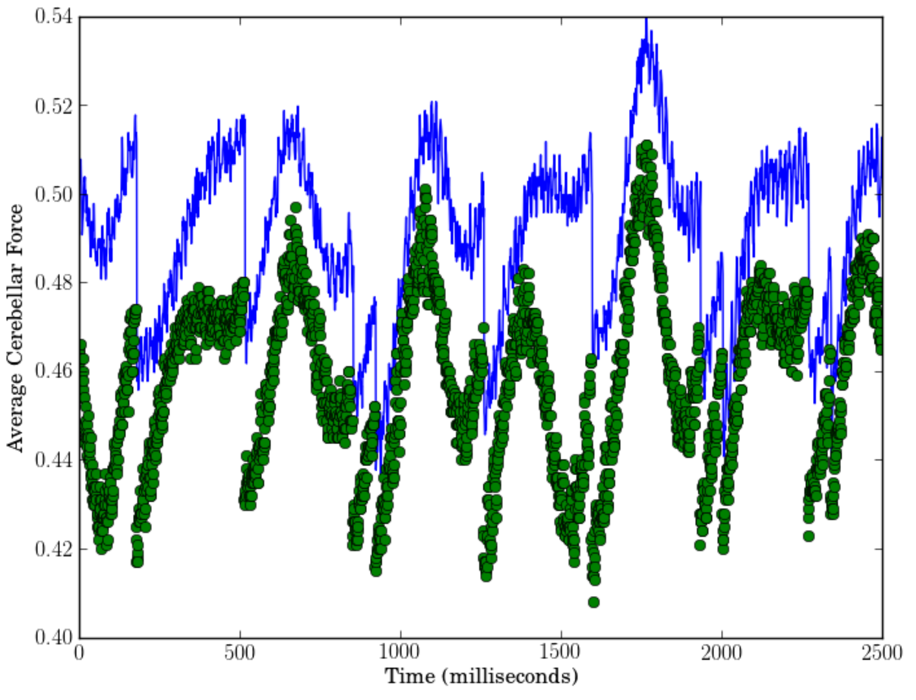
Piano

Audio Recognition Results

Violin.wav



Piano.wav



Green: Output from MZ trained on Violin.wav

Blue: Output from MZ trained on Piano.wav

Conclusion: Differences not robust!

Audio Recognition Conclusions

- Cerebellum can identify different audio signals provided their frequencies are sufficiently separated (e.g. different static patterns)

- More advanced audio recognition requires temporal pattern recognition and proves difficult for the cerebellum

Outline

- Introduction: Biology of the cerebellum
- Cerebellum Simulator
- Experimental Domains
    - Eyelid Conditioning
    - Cartpole
    - PID Control
    - Robocup Balance
    - Pattern Recognition
    - Audio Recognition
- Conclusions

Guidelines for Cerebellar Tasks

- Tasks need supervised error signals that occur regularly regardless of performance.
- Nearly all static patterns of state input are recognized (except NAND). Temporal patterns generally not recognized.
- Overcoming limitations of cerebellar learning likely requires integration of additional brain regions.