Forward and Inverse Models in the Cerebellum

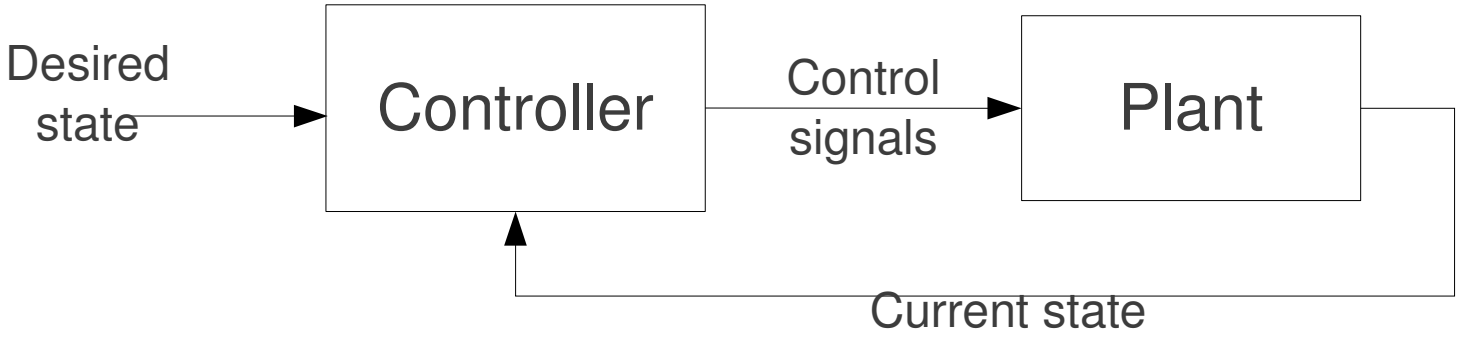
Computational Models of Neural Systems

Lecture 2.3

David S. Touretzky

September, 2013

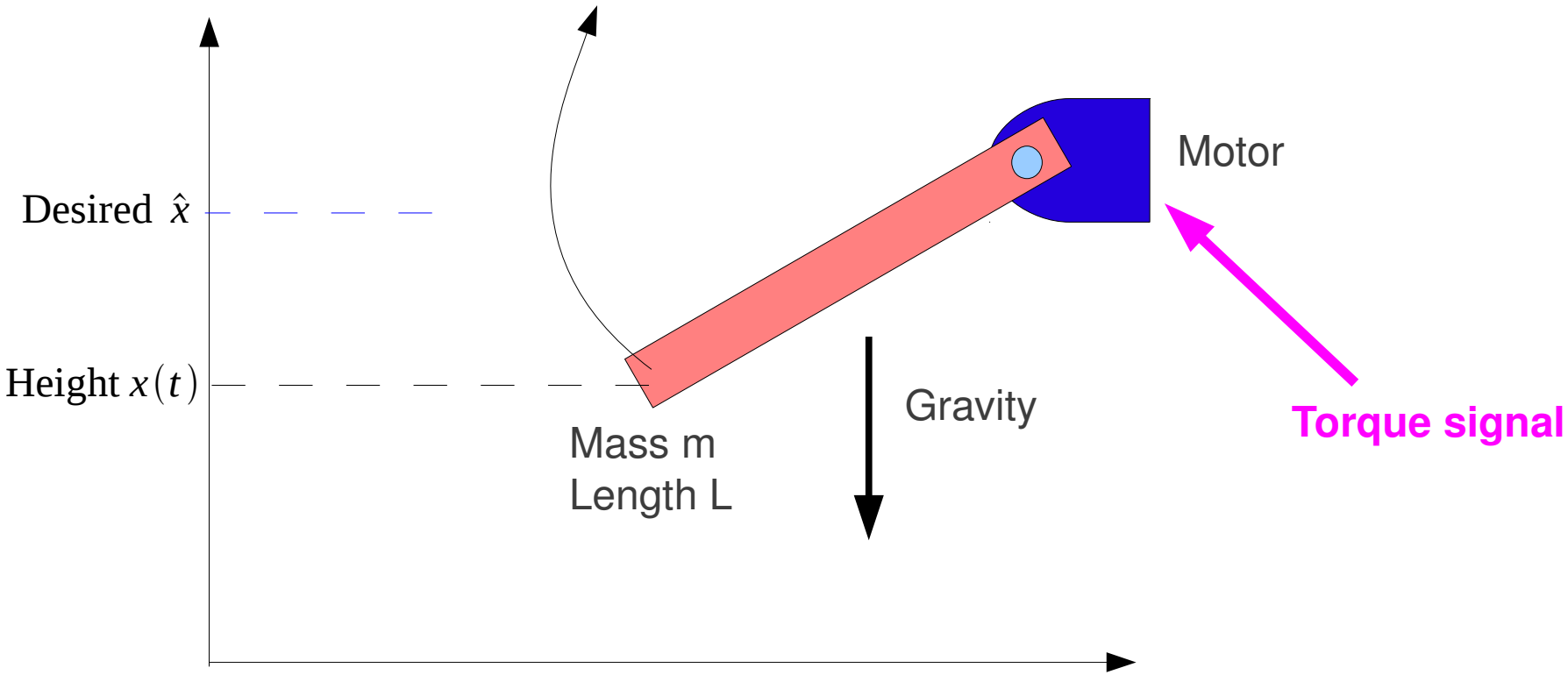
Basics of Control Theory



- The "**plant**" is the thing being controlled.
- The controller translates desired states into **control signals**.
- Control signals might be motor torques or muscle activations.
- The current state could be just the joint positions, or it could include joint velocities, accelerations, load signals, etc.
- Complications: actuators may be slow to respond; feedback may be delayed.

Feedback Control

- A simple way to control a plant is to try to continuously reduce the difference between its current state and the desired state.
- Simple example: control the height of a swinging arm by varying the torque on a motor.

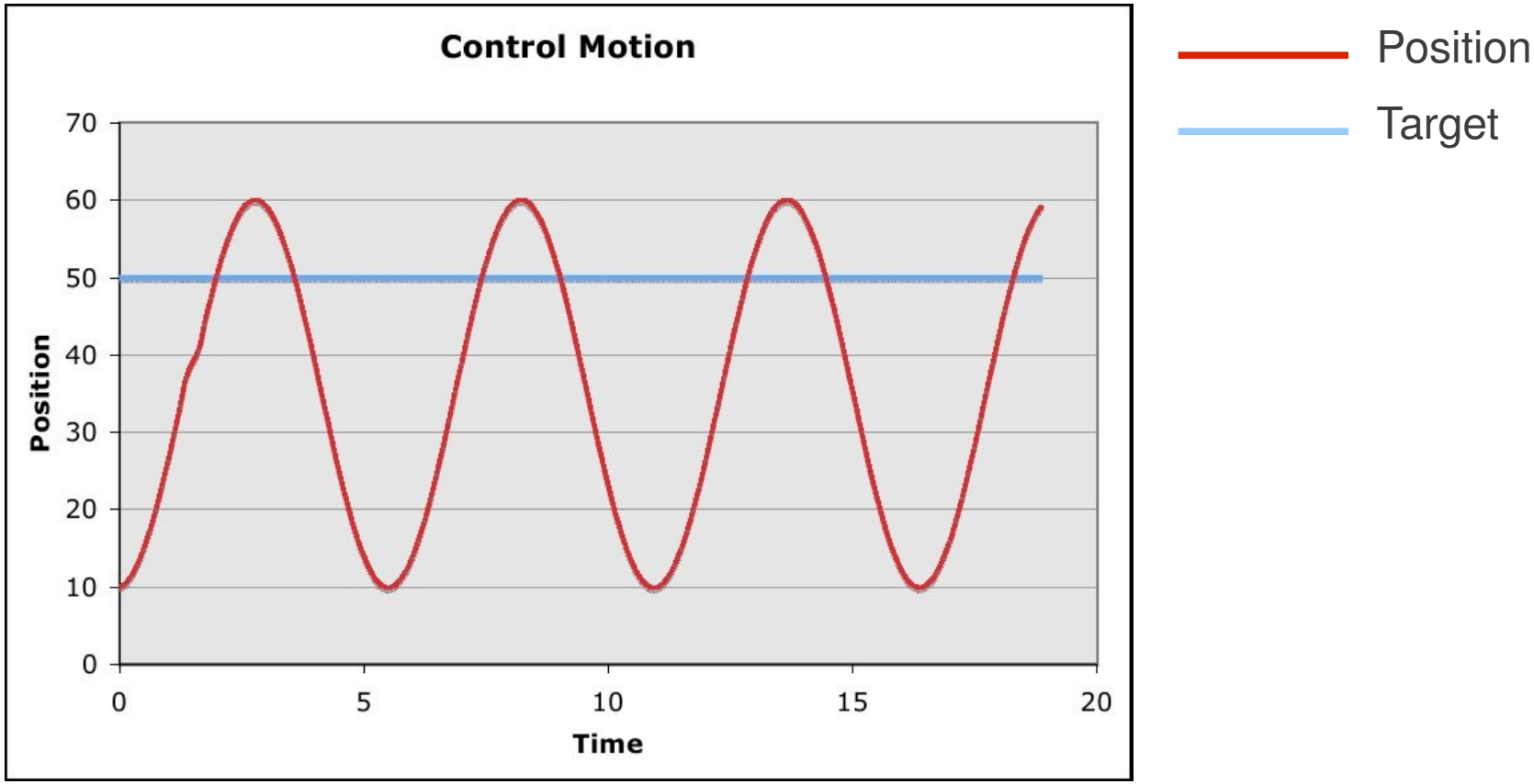


Proportional Control

$$
\begin{aligned}
x(t) &= \text{current position} \\
\hat{x} &= \text{desired position} \\
e(t) &= x(t) - \hat{x} \quad \text{error signal} \\
torque &= -k_p \cdot e(t)
\end{aligned}
$$

- Larger error will generate more torque, proportional to k_p.
- When error is zero, torque is zero.
  - But error won't stay zero due to gravity pulling the arm down.

Proportional Control Is Unstable



- Position oscillates and never converges
- Doesn't even oscillate around the target value.

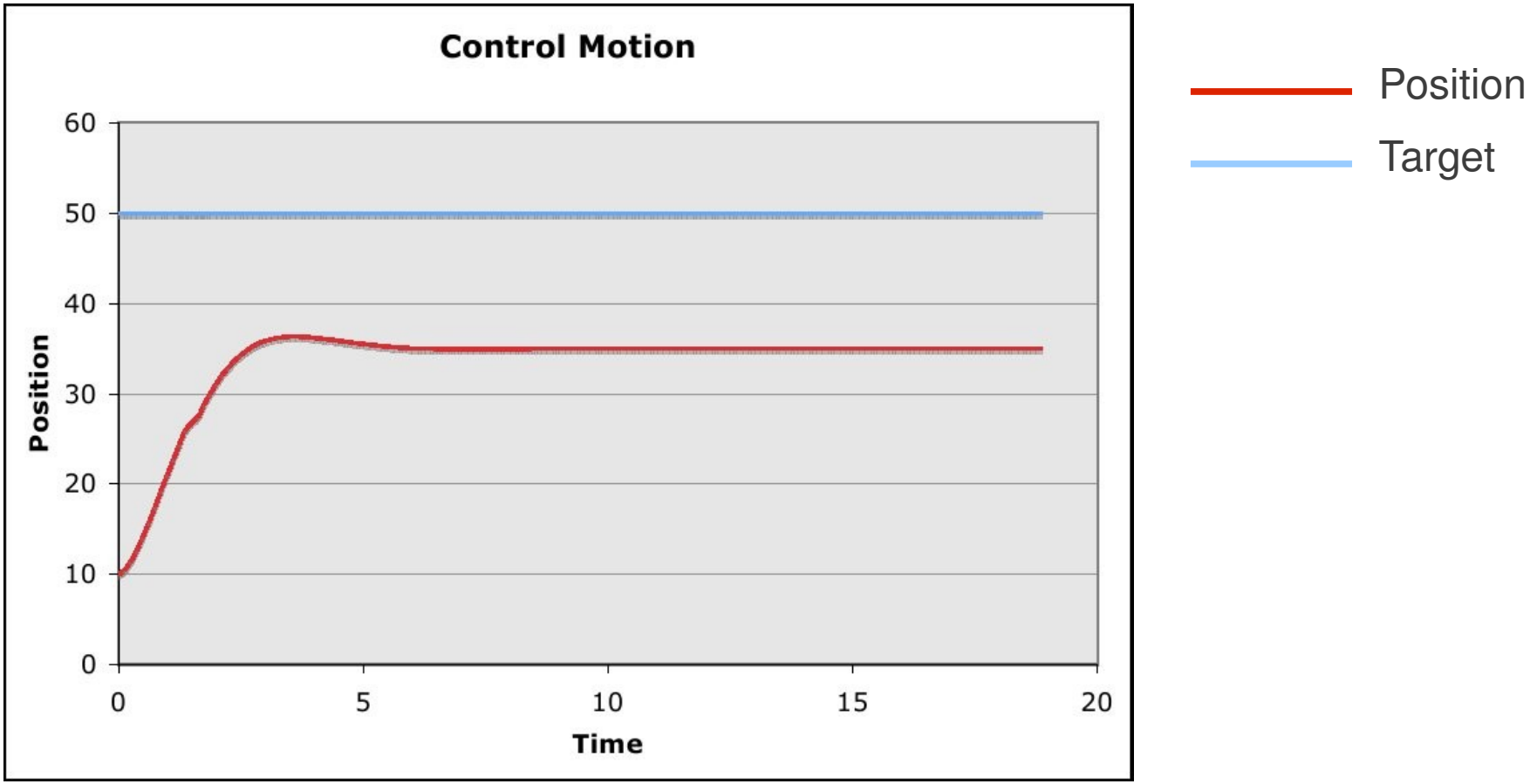
Proportional-Derivative Control

- Oscillation occurs because inertia keeps the arm moving even as the error (and applied torque) are reduced.
- Solution: introduce a braking factor k_d multiplied by the derivative of the error.
  - If error is falling rapidly, apply the brakes so we don't overshoot.

$$torque = -k_p \cdot e(t) - k_d \cdot \frac{\partial e(t)}{\partial t}$$

PD Control Undershoots

- The arm asymptotes at a position where the force of gravity exactly balances the torque from the residual error.

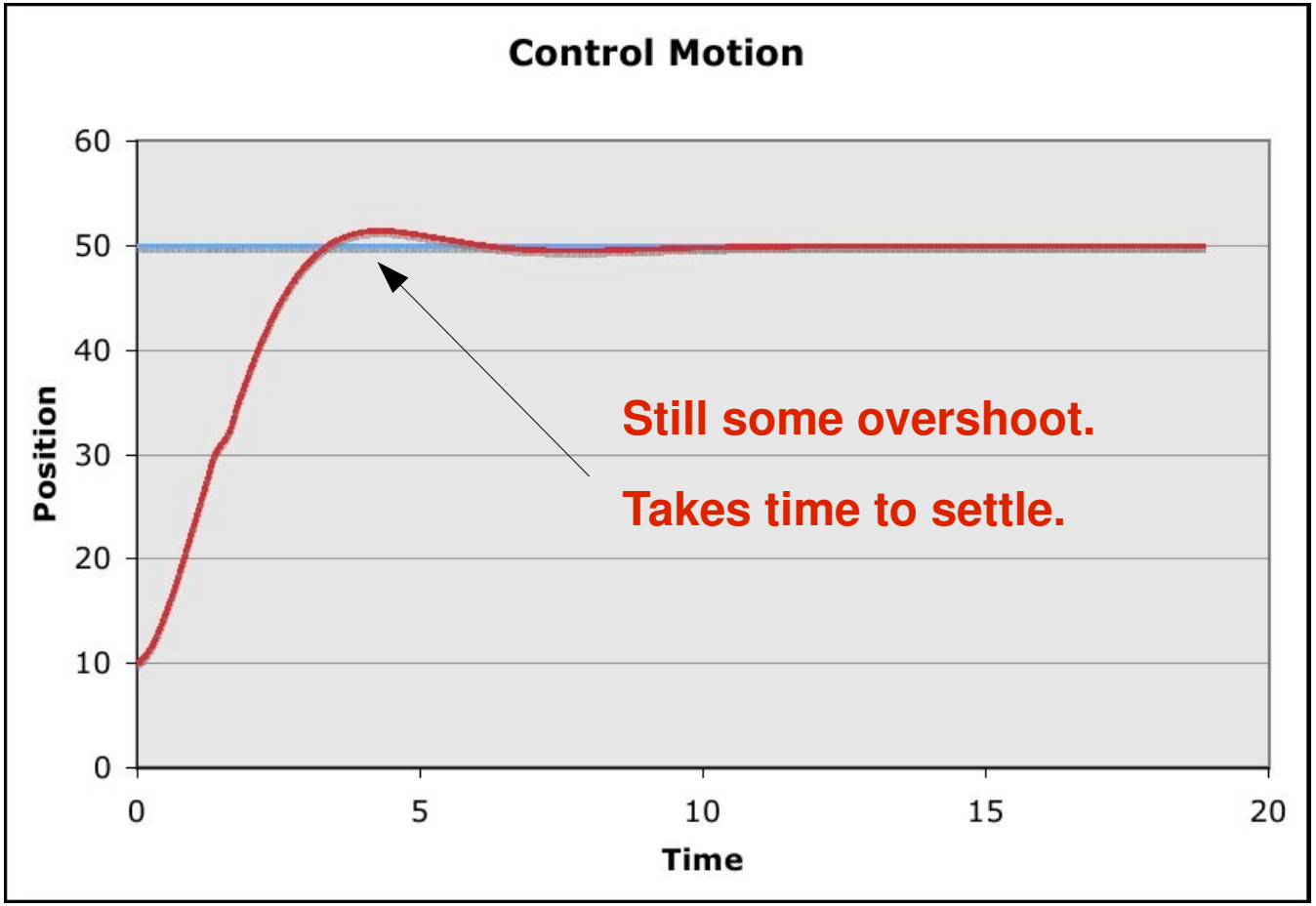


Proportional-Integral-Derivative Control

- Need another term to counteract constant inputs to the system, such as gravity pulling the arm down.
- Use an integral of the error term, so persistent error will gradually be met with increasing force.

$$torque = -k_p \cdot e(t) - k_i \cdot \int e(t)dt - k_d \cdot \frac{\partial e(t)}{\partial t}$$

PID Control Works Better

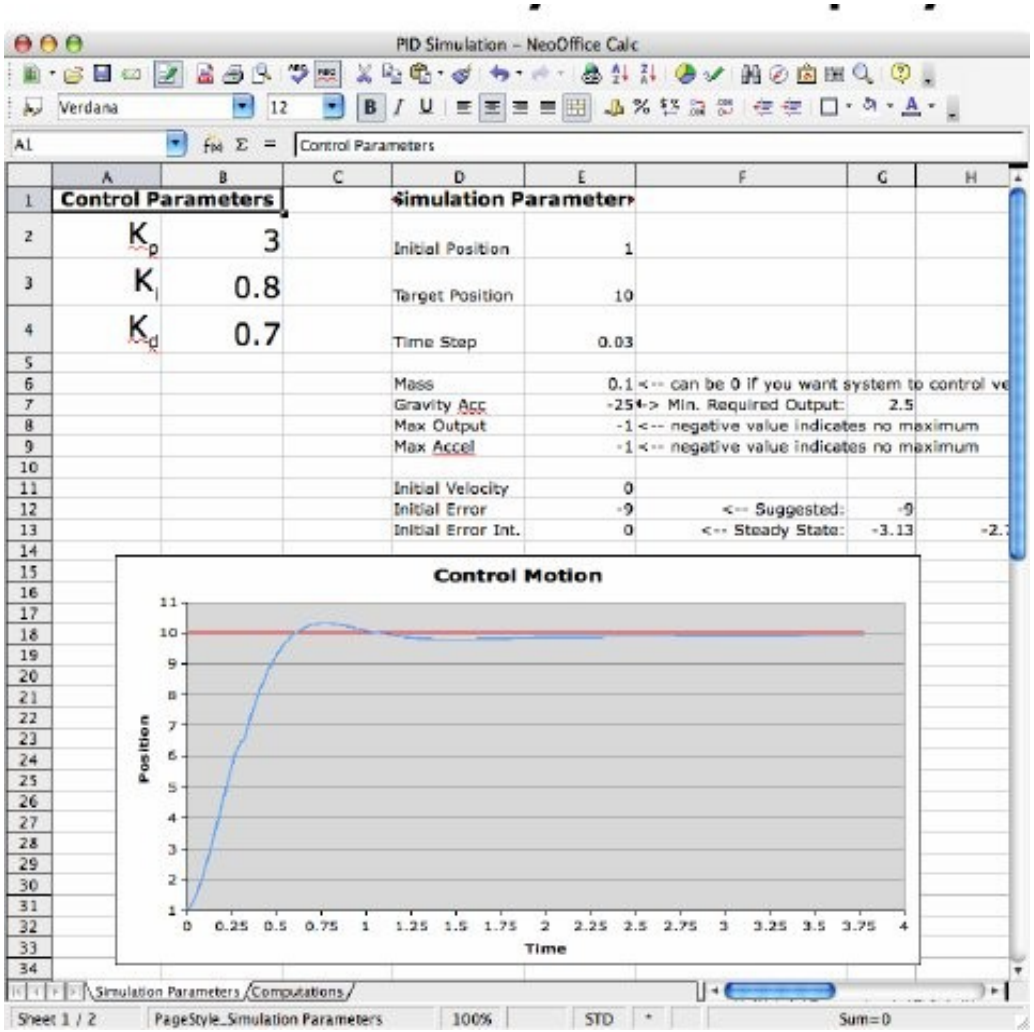


Position

Target

Demos

- Excel spreadsheet for PID control:



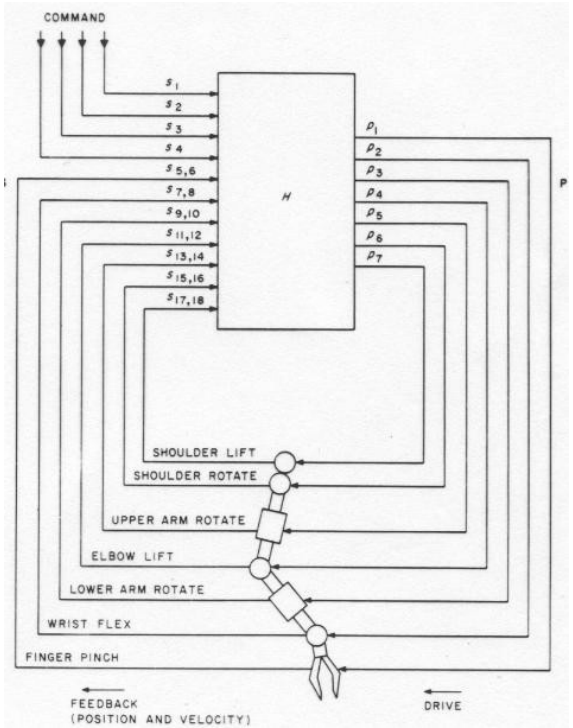
- Video of P vs. PID control of a wheeled cart

Control Theory: General

- Branch of engineering and mathematics dealing with dynamical systems.
- If we have a complete description of the system (mass distribution, torques, friction) we can derive controllers for it mathematically.
  - Differential equations describe the system.
  - Many control strategies possible: linear, nonlinear, adaptive, ...
- Model identification: *learning* the system description through observation.
- Machine learning can be used to learn an efficient controller from experience.

Plants With Complex Dynamics

Simple PID controllers won't work well for plants where the the actuators can interact and the dynamics are complex.



Instead, we need a model of the plant that captures these complex dynamics.

Forward model: maps control signals to predicted plant behavior.

Inverse model: maps desired behavior to control signals that will produce that behavior.

Wolpert et al.

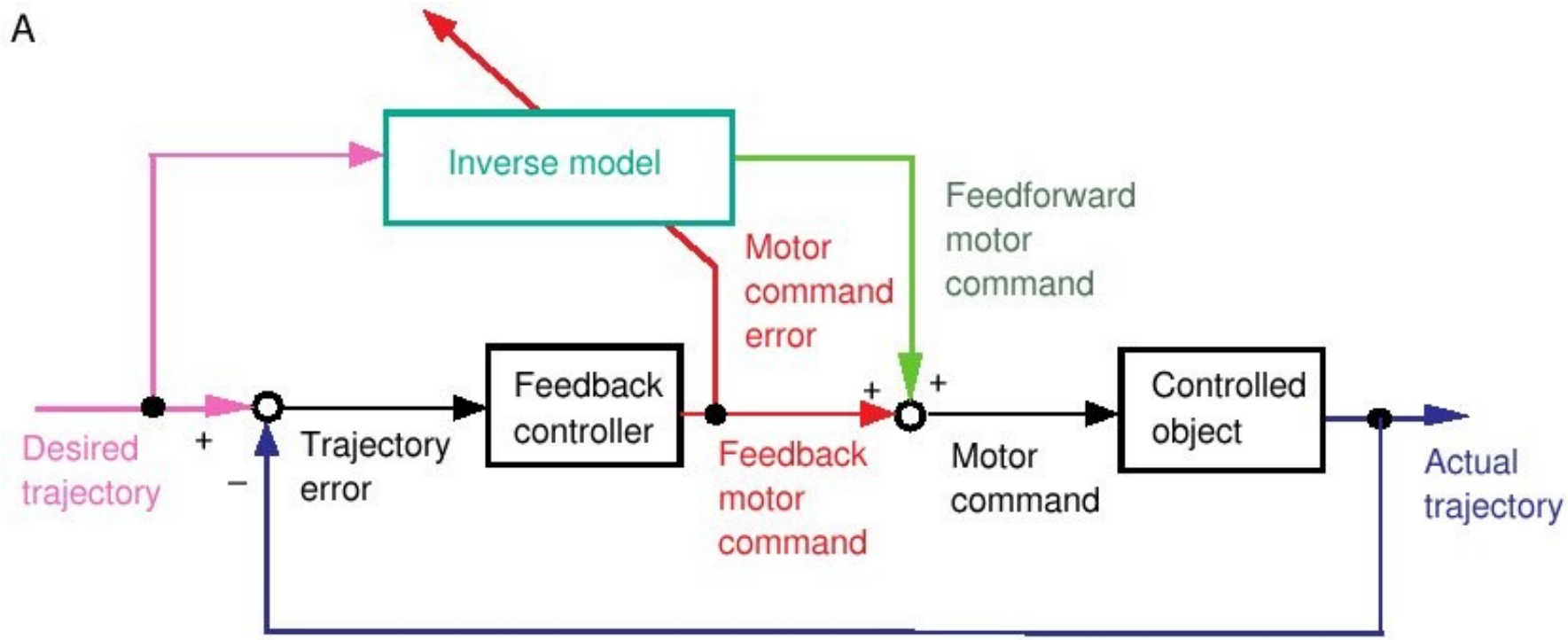
- Simple feedback controllers won't work for animals because biological feedback loops are slow and have small gains.

- Proposal: use an inverse model to anticipate what the plant will do and generate appropriate control signals.

- But how do we train such a model?
  - We don't know the correct control signals to start with.
  - So how do we correct errors in the inverse model's output?

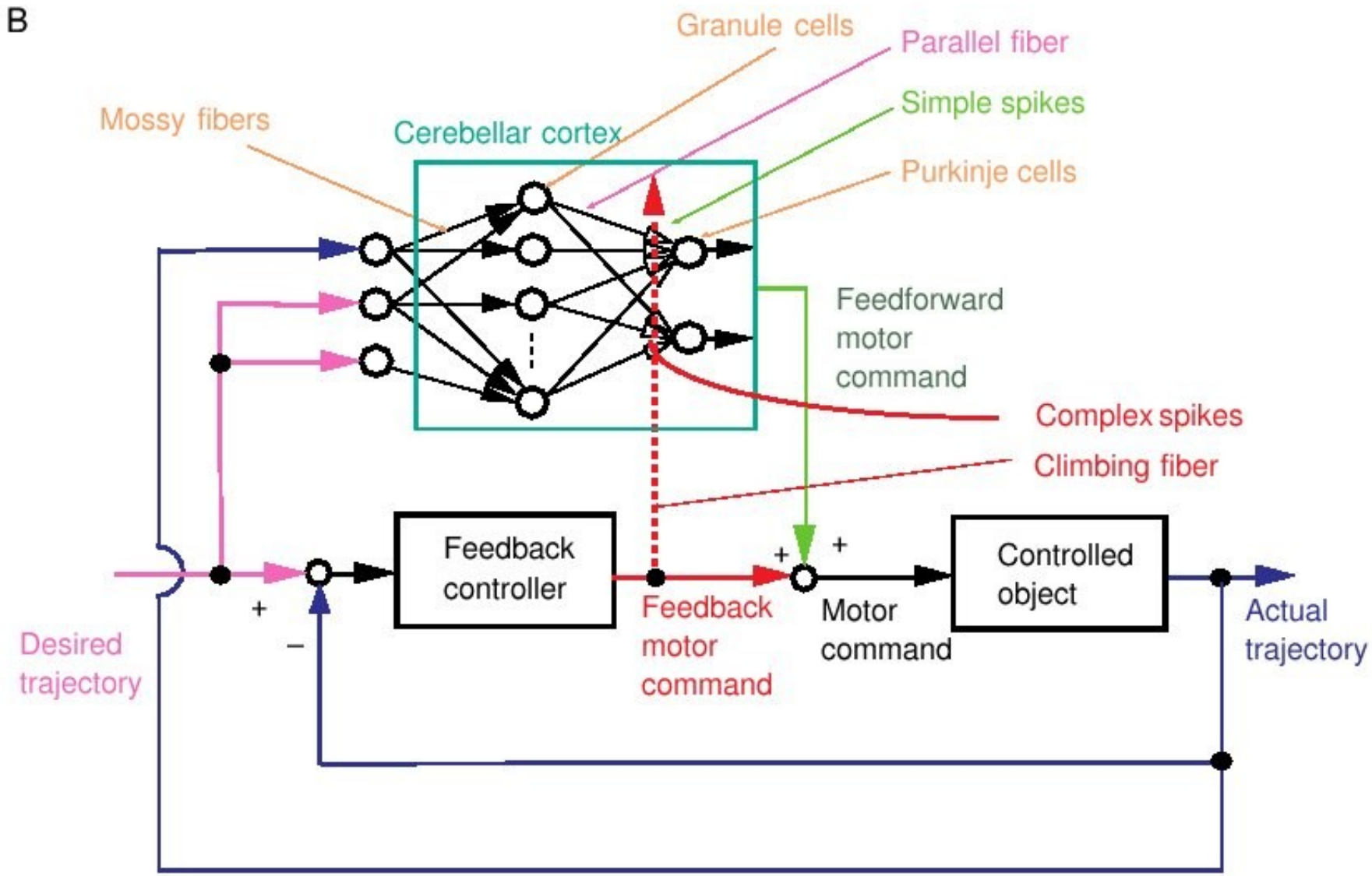
Training the Inverse Model

- Assume a feedback controller that can convert sensory signals to control signal error.
- Use this error to train the inverse model.



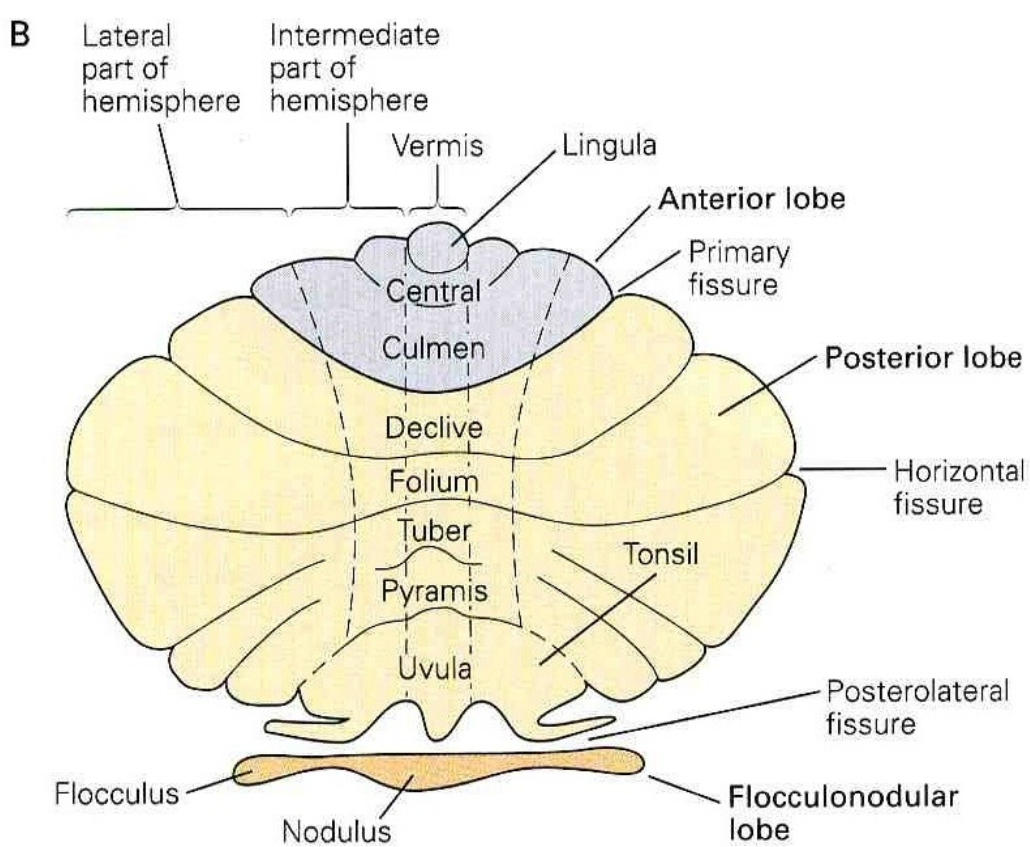
Does the Cerebellum Contain Inverse Models?

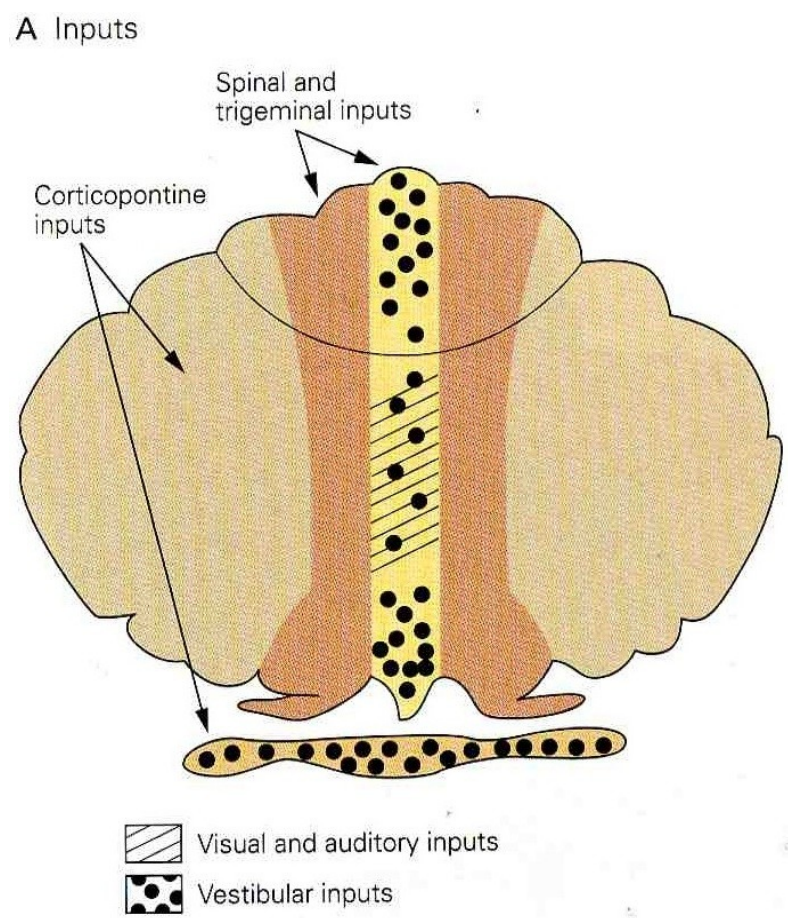
Kawato's CBEFLM (Cerebellar Feedback-Error Learning Model)



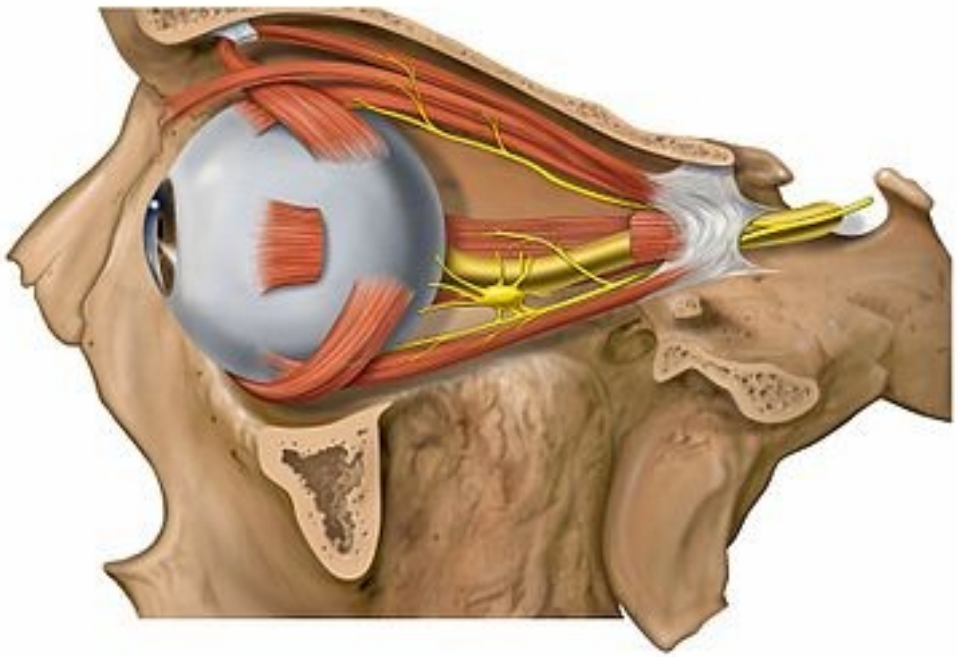
Cerebellar Control of Eye Movements

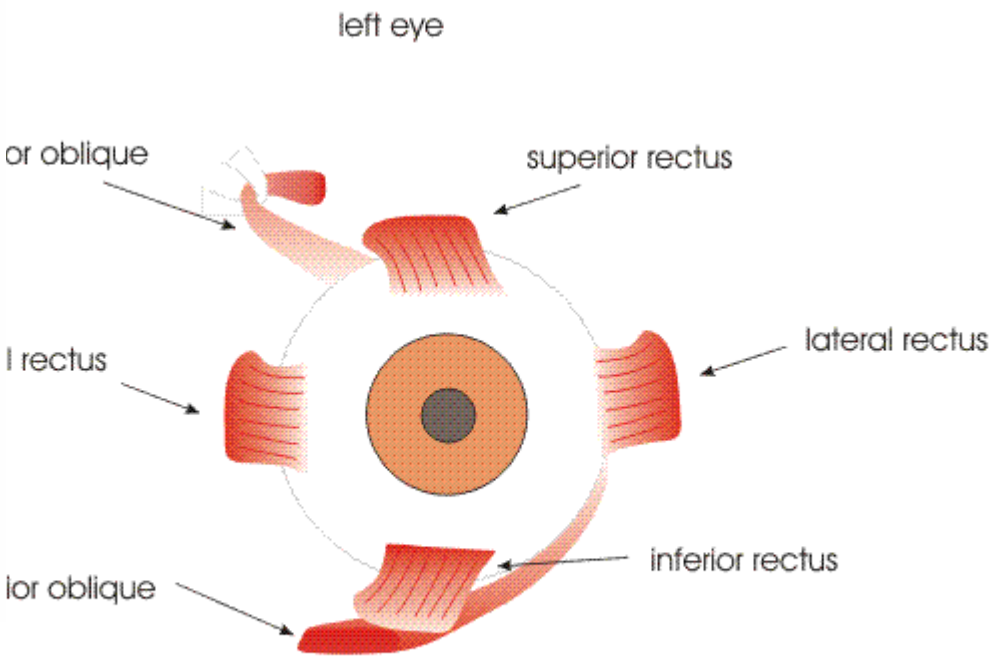
- Assume each cerebellar "microzone" contains a separate inverse model for some part of the body.
- Optical following response (OFR) generated in ventral paraflocculus.

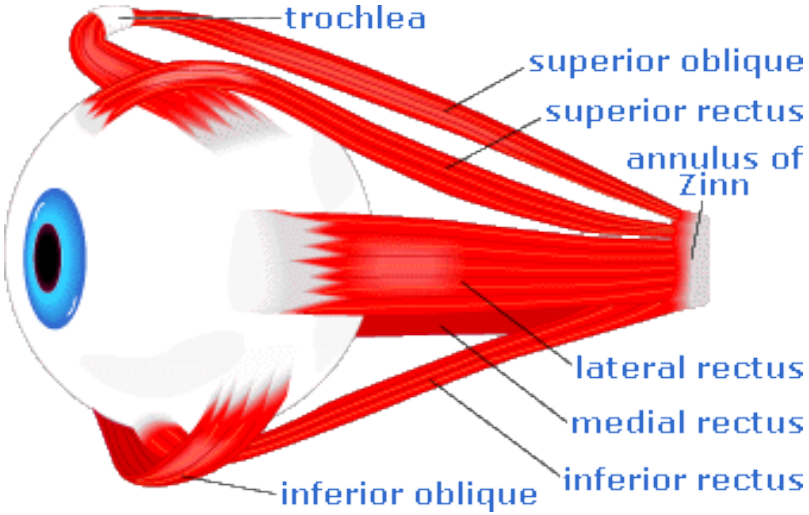




Musculature of the Eye

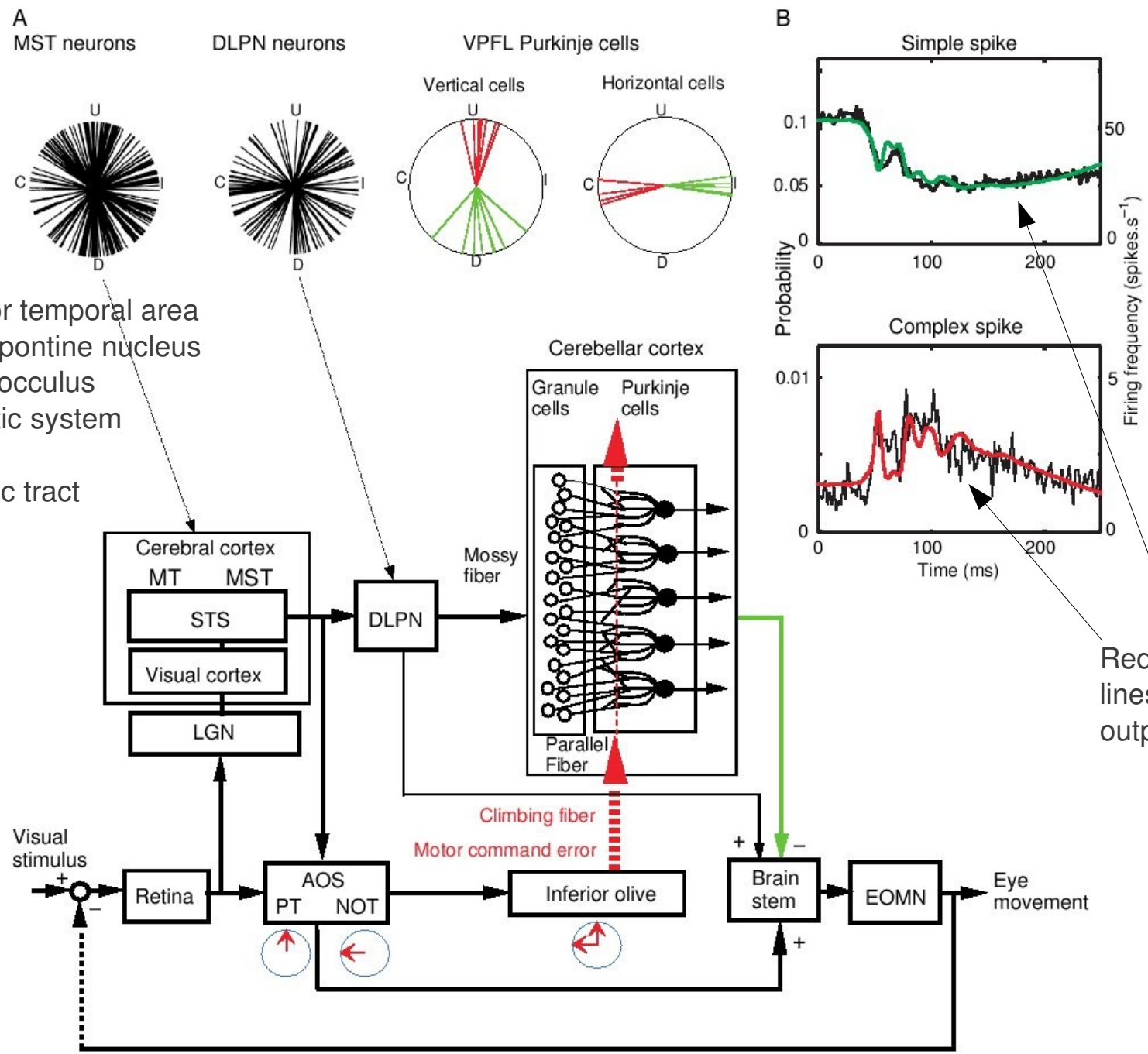






Ocular Following Response (OFR)

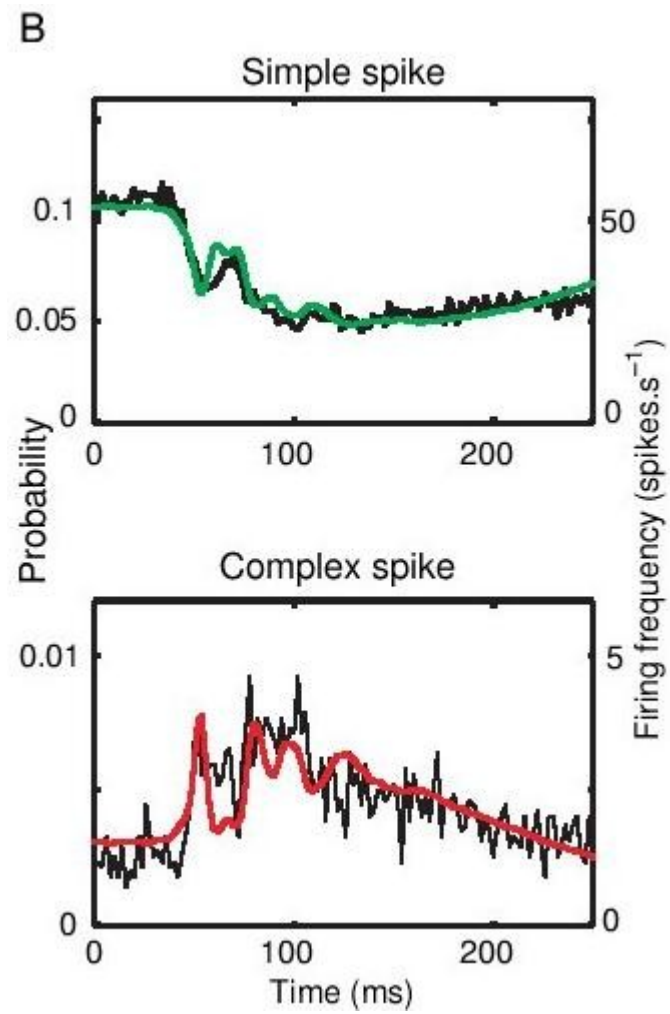
MST: Medial superior temporal area
DLPN: Dorsolateral pontine nucleus
VPFL: ventral paraflocculus
AOS: Accessory optic system
PT: Pretectum
NOT: nucelus of optic tract
EOMN: extra-ocular motor neurons



Red and green lines = model output

Modeling Purkinje Cell Responses

- Model used linear combination of eye acceleration, velocity, and position.
- Quantities were measured 10 ms after simple spike measurement (accounts for conduction delay).
- Good fit for Purkinje cells in VPFL.
- So VPFL may be the inverse model for ocular following response.
- Not so good fit for neurons in MST or DLPN, which provide the input to VPFL. Do they encode trajectories (input to inverse model)?



What Do The Input Fibers Encode?

Parallel fibers:

- Eye movements : motor representation
- Retinal slip: sensory representation

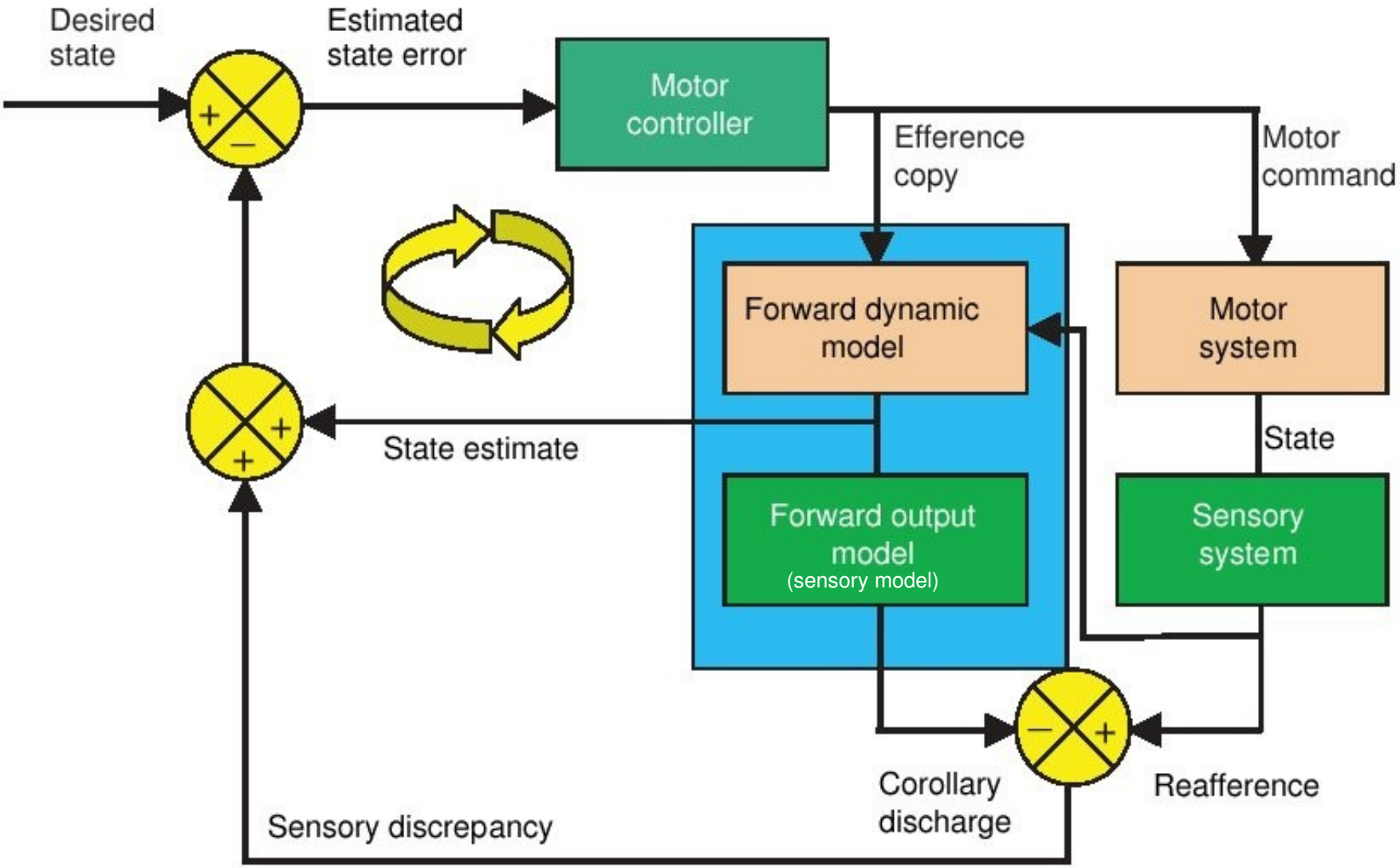
Climbing fibers

- Motor error?

Forward Models in the Cerebellum?

- Forward dynamics model predicts the future state of the plant.
- Forward sensory model can predict future delayed sensory inputs.
- Why are forward models useful here?
  - Sensory feedback has long time delays.
  - Forward model can provide for much faster corrections.
- A Smith predictor is a type of forward model useful when there are delays in:
  - Sensory processing
  - Sensory-motor coupling
  - Motor execution

Smith Predictor Model



Arguments for Multiple Controllers

1. Human motor behavior is rich and complex.
   - Unreasonable to expect everything to be captured by a single inverse or forward model.
2. Assigning different behaviors to different modules allows them to be learned independently, avoiding mutual interference.
3. If we have multiple controllers, we can take weighted combinations of them to synthesize new control regimes.
   - Controllers could serve as motor primitives.
4. Prism glasses deadaptation and readaptation are faster than adaptation, suggesting that there is switching going on.

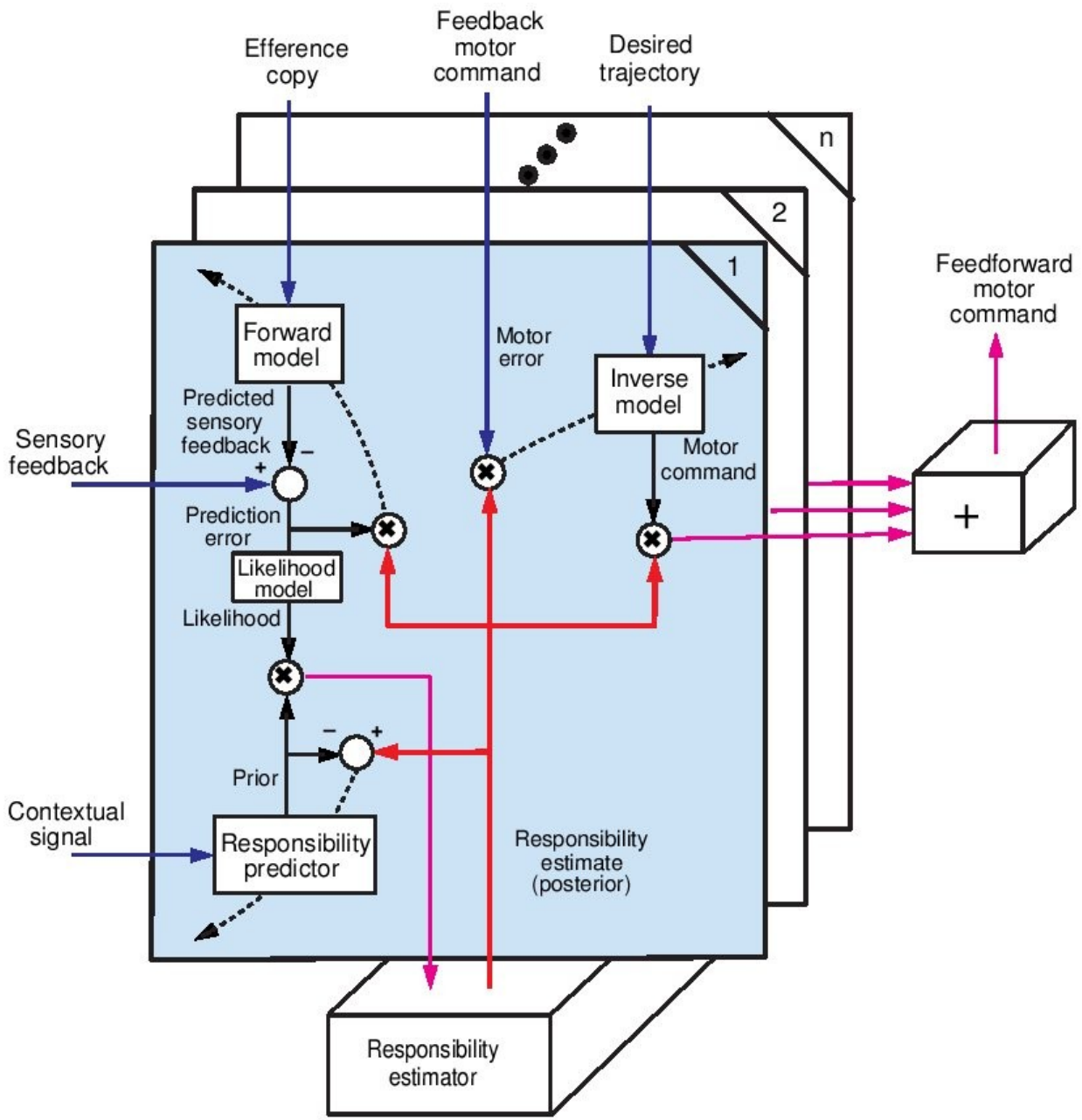
But how do we decide which model(s) to apply?

Multiple Paired Forward and Inverse Models?

Inverse model specialized for a particular behavioral context.

Forward models help determine "responsibility" for their associated inverse model in the current context, based on the goodness of their sensory predictions.

Prior estimate comes from a separate responsibility predictor.



Summary

- Biological motor control is difficult due to sensory and motor delays, and complex dynamics of the plant.
- Eye movement is a good control problem to study because it's relatively simple compared to reaching tasks.
  - But there are actually several types of eye movements: OFR, VOR, saccades, ...
- We know that cerebellum learns, but <u>what</u> is it learning?
  - Inverse model? Forward model? Something else?
- Cerebellar circuitry appears to be uniform throughout. So how does this theory account for cerebellar contributions to:
  - Motion planning (cerebrocerebellum)
  - Classical conditioning (timing of responses)
  - Cognitive phenomena, including language tasks