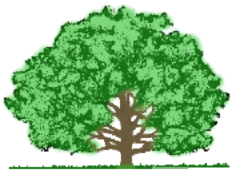
Probabilistic mechanisms in sensorimotor control

Daniel Wolpert, University College London

Q. Why do we have a brain?

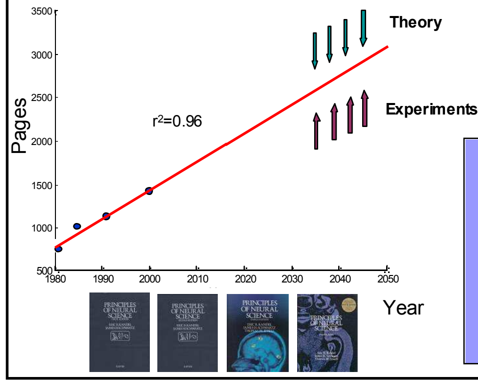
A. To produce adaptable and complex movements

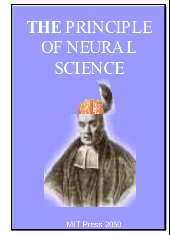
- movement is the only way we have of
  - Interacting with the world
  - Communication: speech, gestures, writing
- sensory, memory and cognitive processes ⇨ future motor outputs



Why study computational sensorimotor control?

Principles of Neural Science, Kandel et al.





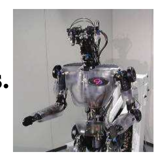
The complexity of motor control

What to move where

vs.

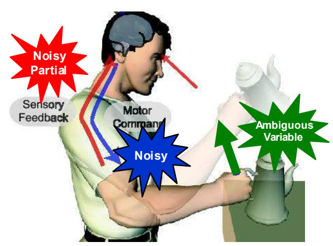
Moving

vs.



Reverse-engineering human sensorimotor control

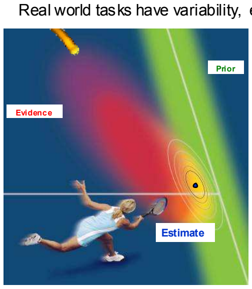
Upper limb
- Multiple degrees of freedom
- Nonlinear & time-varying
- Time delays **& noise**



Outline
- Learning
  - **I. Bayesian learning**
- Prediction
  - **II. Sensory prediction**
- Evaluation
  - **III. Loss function**
- Control
  - **IV. Optimal control**

I. Bayesian Motor Learning

Real world tasks have variability, e.g. estimating ball's bounce location



Sensory feedback (Evidence)
Combine multiple cues
to reduce uncertainty

+

Task statistics (Prior)
Not all locations are
equally likely

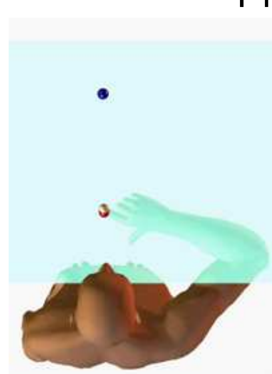
=

Optimal estimate (Posterior)
Bayes rule

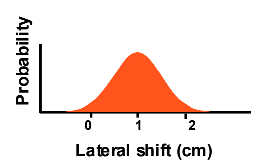
P(state|sensory input) ∝ P(sensory input|state) P(state)

Does sensorimotor learning use Bayes rule?
If so, is it implemented
- Implicitly: mapping sensory inputs to motor outputs to minimize error?
- Explicitly: using separate representations of the statistics of the prior and sensory noise?

Prior

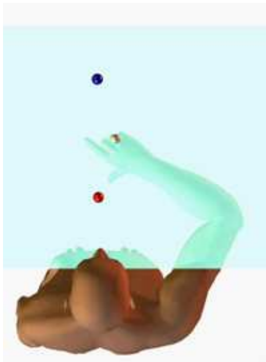


Task in which we control
1) prior statistics of the task
2) sensory uncertainty

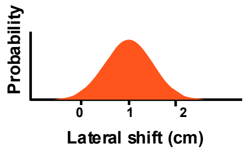


(Körding & Wolpert, Nature, 2004)

Prior

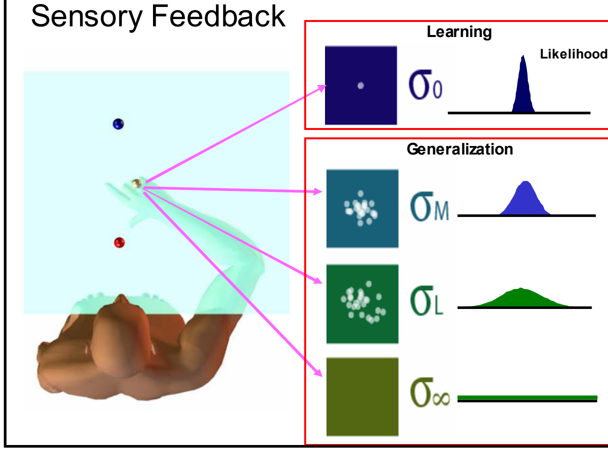


Task in which we control
1) prior statistics of the task
2) sensory uncertainty

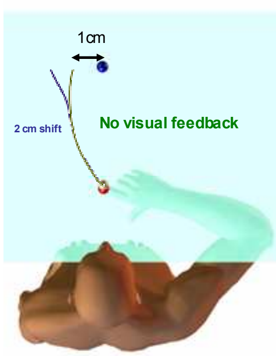


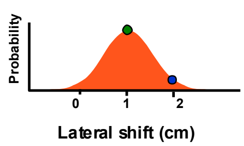
(Körding & Wolpert, Nature, 2004)

Sensory Feedback

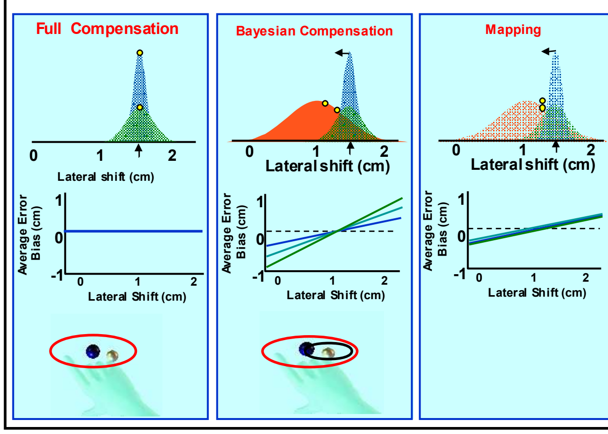


After 1000 trials

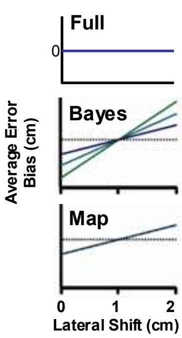


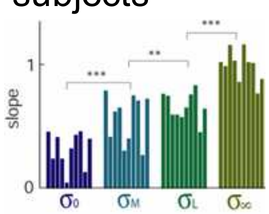


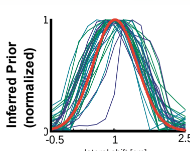
Models



Results: 10 subjects





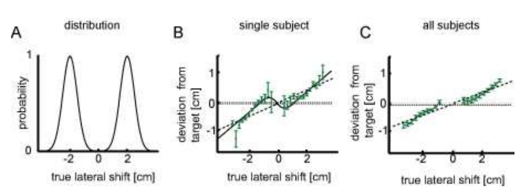


Supports model 2: Bayesian

Bayesian integration

Subjects can learn

- multimodal priors
- priors over forces
- different priors one after the other



(Körding& Wolpert NIPS 2004, Körding, Ku & Wolpert J. Neurophysiol. 2004)

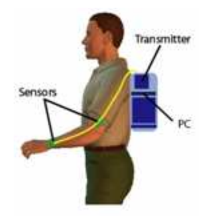
Statistics of the world shape our brain

Objects **Configurations of our body**

- **CNS**
  - **Represents the distribution of tasks**
  - **Estimates its own sensory uncertainty**
  - **Combines these two sources in a Bayesian way**

Statistics of action

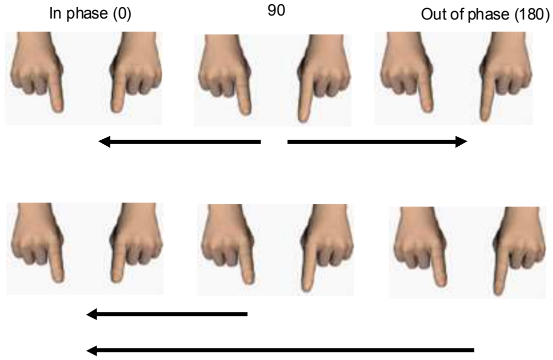
- Statistics of visual/auditory stimuli ⇨ representation visual/auditory cortex
- Statistics of early experience ⇨ what can be perceived in later life (e.g. statistics of spoken language)



With limited neural resources statistics of motor tasks ⇨ motor performance

Symmetry bias

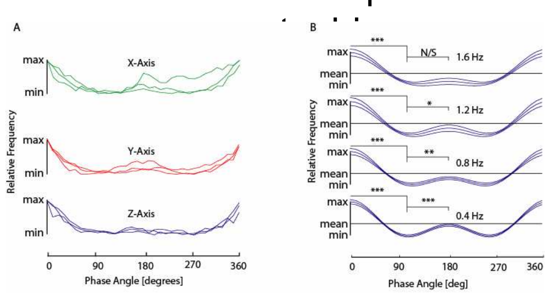
In phase (0) 90 Out of phase (180)



Symmetry bias in extrinsic space (Mechsner et al, Nature, 2001)

System with limited resources
•Allocate more resources to common (important) movements
•If correlations between hands then code with single controller

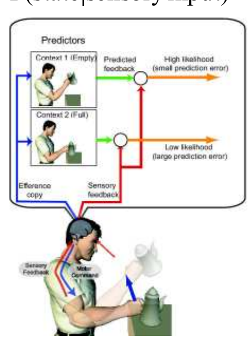
Phase relationships and ...



Symmetry bias may reflect statistics of tasks

II. Sensory likelihood

$$P(\text{state}|\text{sensory input}) \propto P(\text{sensory input}|\text{state})\, P(\text{state})$$



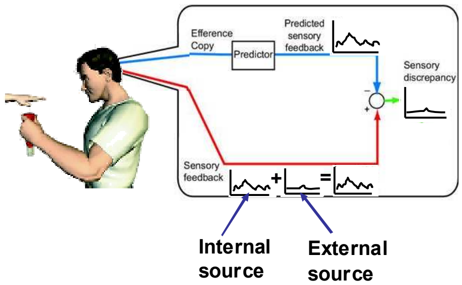
Fundamental for
1. State estimation
2. Control with delays
3. Mental simulation
4. Likelihood estimation
   Wolpert & Kawato, Neural Networks 1998
   Haruno, Wolpert, Kawato, Neural Computation 2001
5. Sensory filtering

(Wolpert et al., Science, 1995)

Sensory prediction

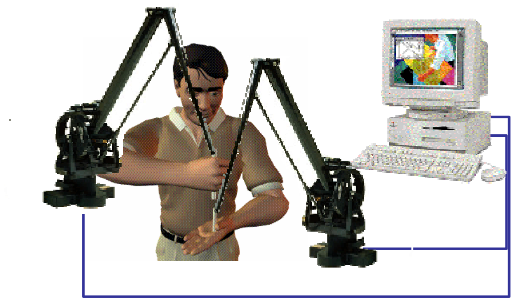
Our sensors report

- Ex-afferent information: changes in outside world
- Re-afferent information: changes we cause

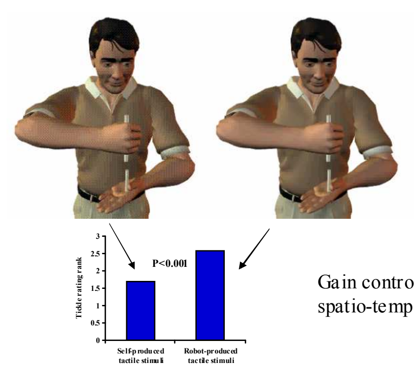


Tickling

Self-administered tactile stimuli rated as less ticklish than externally administered tactile stimuli. (Weiskrantz et al, 1971)

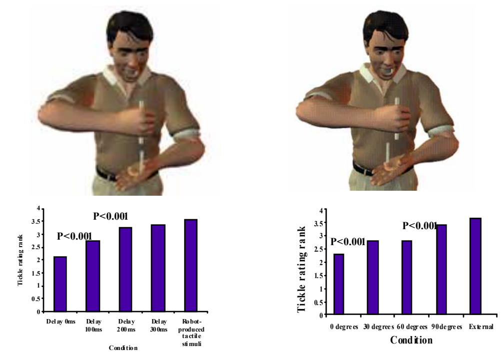


Does prediction underlie tactile cancellation in tickle?



Gain control or precise spatio-temporal prediction?

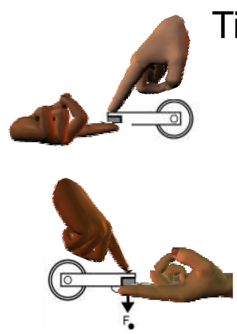
Spatio-temporal prediction

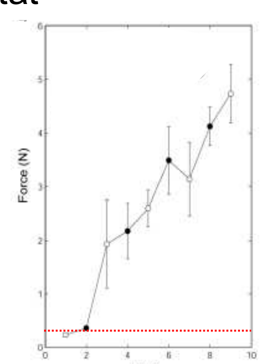


(Blakemore, Frith & Wolpert. J. Cog. Neurosci. 1999)

The escalation of force

Tit-for-tat

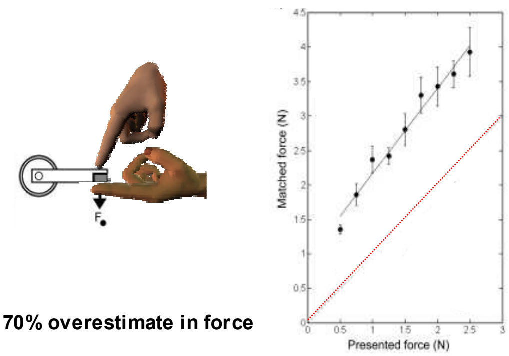




Force escalates under rules designed to achieve parity: Increase by ~40% per turn

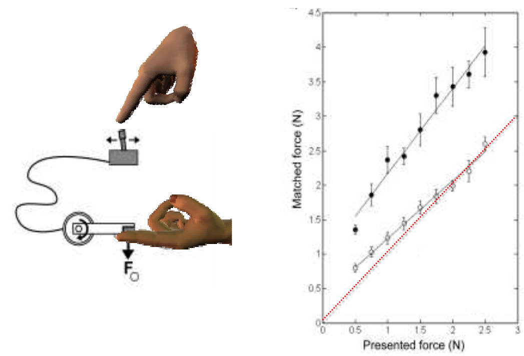
(Shergill, Bays, Frith & Wolpert, Science, 2003)

Perception of force

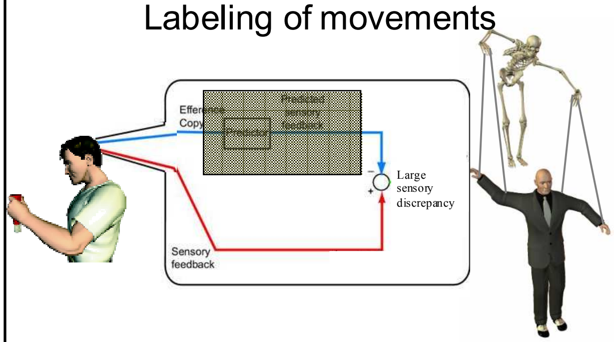


70% overestimate in force

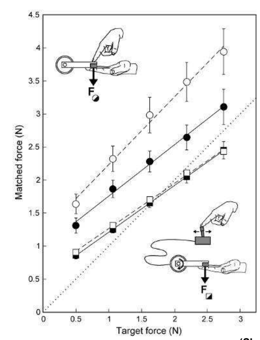
Perception of force



Labeling of movements



Defective prediction in patients with schizophrenic



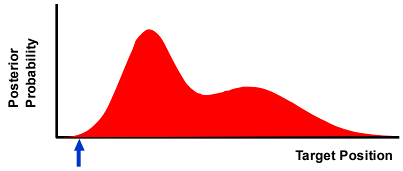
Patients ■ ●
Controls □ ○

- **The CNS predicts sensory consequences**
- **Sensory cancellation in Force production**
- **Defects may be related to delusions of control**

(Shergill, Samson, Bays, Frith & Wolpert, Am J. Psychiatry, In Press)

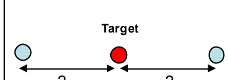
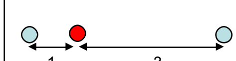
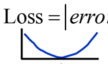
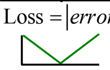
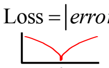
III. Loss Functions in Sensorimotor system

$$\underbrace{P(\text{state}|\text{sensory input})}_{\text{Posterior}} \propto \underbrace{P(\text{sensory input}|\text{state})}_{\text{Likelihood}} \underbrace{P(\text{state})}_{\text{Prior}}$$

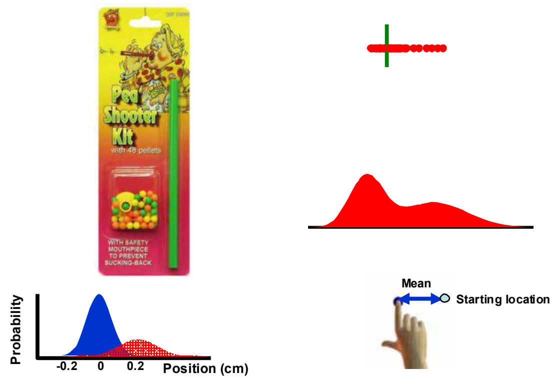


- What is the performance criteria (loss, cost, utility, reward)?
  - How bad is 2cm vs. 1cm error?
- Often assumed in statistics & machine learning
  - that we wish to minimize squared error
  - for analytic or algorithmic tractability
- What measure of error does the brain care about?

Loss function

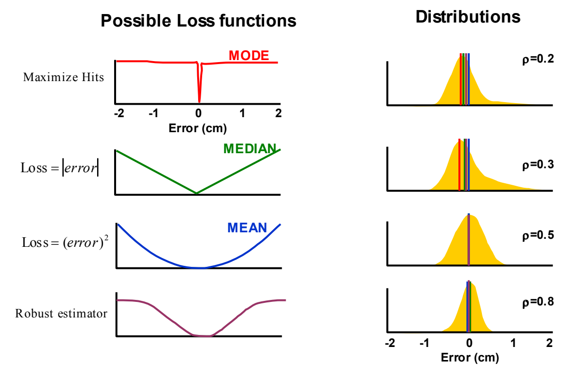
	Scenario 1	Scenario 2
		
$\text{Loss} = \lvert error \rvert^2$ 	Loss=4+4=8 ☑	Loss=1+9=10 ☒
$\text{Loss} = \lvert error \rvert$ 	Loss=2+2=4 ☑	Loss=1+3=4 ☑
$\text{Loss} = \lvert error \rvert^{\frac{1}{2}}$ 	Loss=1.4+1.4=2.8 ☒	Loss=1+1.7=2.7 ☑

Virtual pea shooter

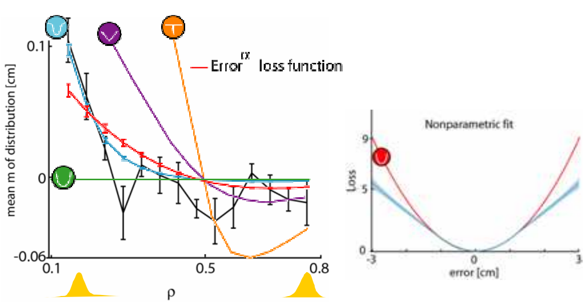


(Körding & Wolpert, PNAS, 2004)

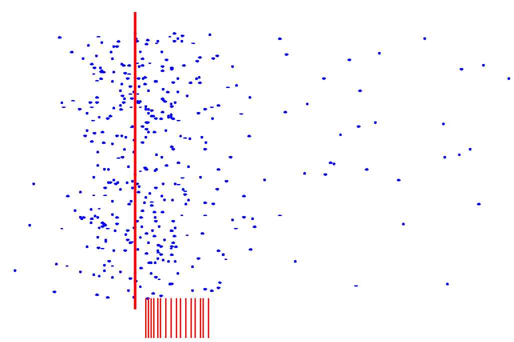
Probed distributions and optimal means



Shift of mean against asymmetry (n=8)



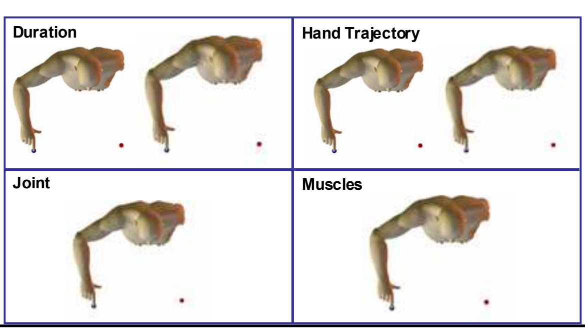
Mean squared error with robustness to outliers



$$\text{Loss} = \sum error_i^{\alpha}$$

IV. Decisions in a redundant motor systems

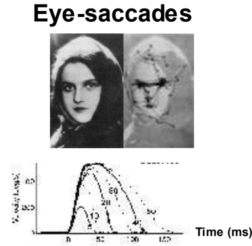
- Tasks are usually specified at a symbolic level
- Motor system works at a detailed level, specifying muscle activations
- Gap between high and low-level specification



Motor evolution/learning results in stereotypy

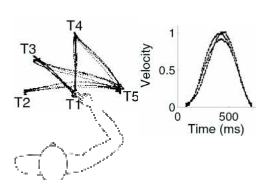
Stereotypy between repetitions and individuals

Eye-saccades



- Main sequence
- Donder's law
- Listings Law

Arm- movements

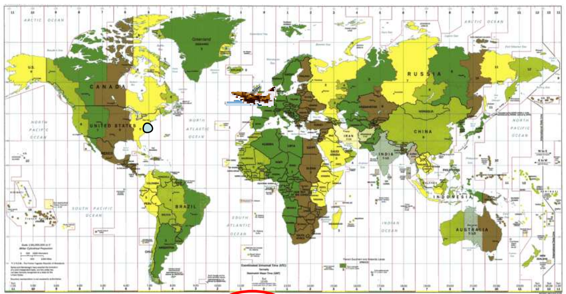


- 2/3 power law
- Fitts' law
- Isochrony

The Assumption of Optimality

Movements have evolved to maximize fitness

- improve through evolution/learning
- every possible movement which can achieve a task has a cost
- we select movement with the lowest **cost**



Overall cost = $cost_1$ + $cost_2$ + $cost_3$ ....

Previous (incorrect) costs

Saccadic eye movements

- little vision over 4°/s
- saccades >200°/s
- frequent 2-3 /sec each ~50ms
- deprives of vision ~90 min/day

⇒Minimize time

Arm Movements

- Are smooth

⇒ Minimum jerk (rate of change of acceleration)

⇒ Minimum torque change

Criteria for cost for goal-directed movement

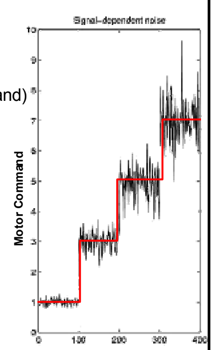
- Makes sense in terms of advantage for evolution & learning
- Simple for CNS to measure
- Generalizes to different systems e.g. eye, head, arm
- Generalizes to different tasks e.g. pointing, grasping, drawing

→ Reproduce & predict behaviour

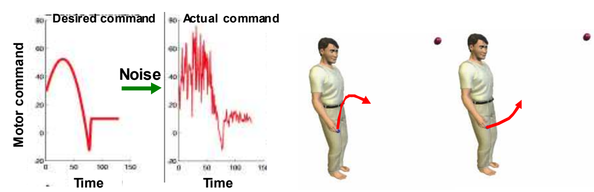
Fundamental constraint=Signal-dependent noise

- Signal-dependent noise:
  - Constant coefficient of variation
  - SD (motor command) ~ Mean (motor command)
- Evidence from
  - Experiments: SD (Force) ~ Mean (Force)
  - Modelling
    - Spikes drawn from a renewal process
    - Recruitment properties of motor units

(Jones, Hamilton & Wolpert , J. Neurophysiol., 2002)



Signal-dependent noise and task achievement



Sequence of motor commands ⇒ probability distribution (statistics) of movement.

The statistics of action can be controlled by changing the motor command

Task ≡ Optimizing f(statistics)

(Harris & Wolpert, Nature, 1998)

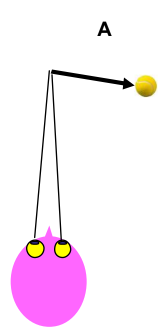
Pointing movements: (f=terminal error)

Given:

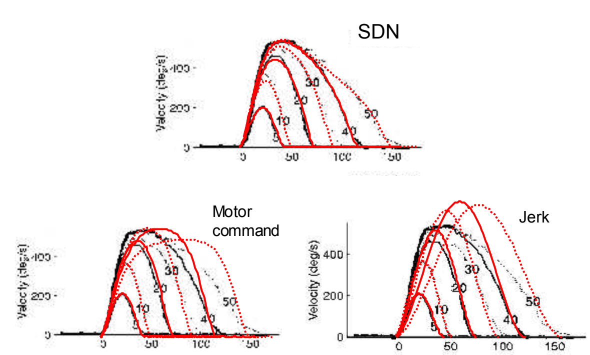
- Amplitude
- Duration
- Eye model: 3rd order linear system
- Signal-dependent noise $\sigma_u \propto u$

Unique optimal trajectory

$$u(t) = \frac{k_1 e^{\frac{(t-M)}{\tau_1}} + k_2 e^{\frac{(t-M)}{\tau_2}} + k_3 e^{\frac{(t-M)}{\tau_3}}}{(\tau_2-\tau_3)\tau_1 e^{\frac{t-T}{\tau_1}} + (\tau_3-\tau_1)\tau_2 e^{\frac{t-T}{\tau_2}} + (\tau_1-\tau_2)\tau_3 e^{\frac{t-T}{\tau_3}}}$$

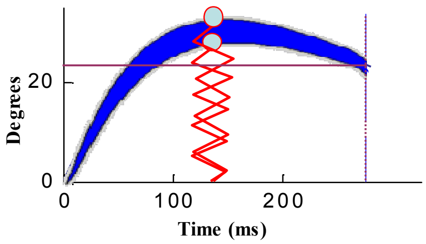


Saccade predictions

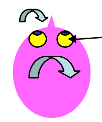


Prediction: very slow saccade

22 degree saccade in 270 ms (normally ~ 70 ms)



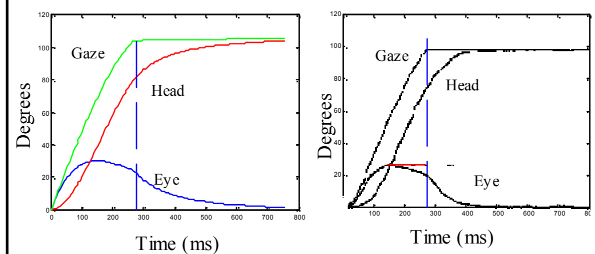
Head free saccade



$T_{1=0.15}$ $T_{2=0.08}$

$T_{1=0.3}$ $T_{2=0.3}$

Free parameter, eye:head noise



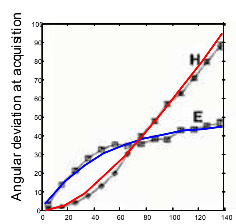
(Tomlinson & Bahra, 1986)

Coordination: Sharing reduces variance

Eye only — Var=k A^2

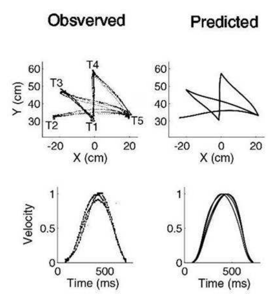
Head only — Var=k A^2

Eye + Head — Var= k $(A/2)^2$ + k $(A/2)^2$ = k $A^2/2$



Gaze amplitude

Arm movements

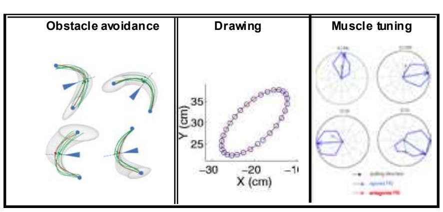


Smoothness

Non smooth movement
⇒ requires abrupt change in velocity
⇒ given inertial system
⇒ large motor command
⇒ increased noise

Smoothness ⇒ accuracy

Arm movements



- **Biologically plausible underpinning for eye, head, arm and wrist movements**
- **Noise lead to statistics of movement**
- **We can control the statistics by choosing different ways to move**

Summary

CNS
- **Minimizes uncertainty through Bayesian estimation**
- **Predict consequences of actions**
- **Penalize squared errors but robust to outliers**
- **Controls statistics through planning**

www.wolpertlab.com