KALMAN FILTERING SENSORIMOTOR INTEGRATION

HUMAN MOTION CONTROL

DR. IR. ERWIN R. BOER

TU DELFT – MAY 13, 2019

May 13, 2019

LECTURE ON KALMAN FILTERING: SENSORIMOTOR INTEGRATION

Readings (2 hours)

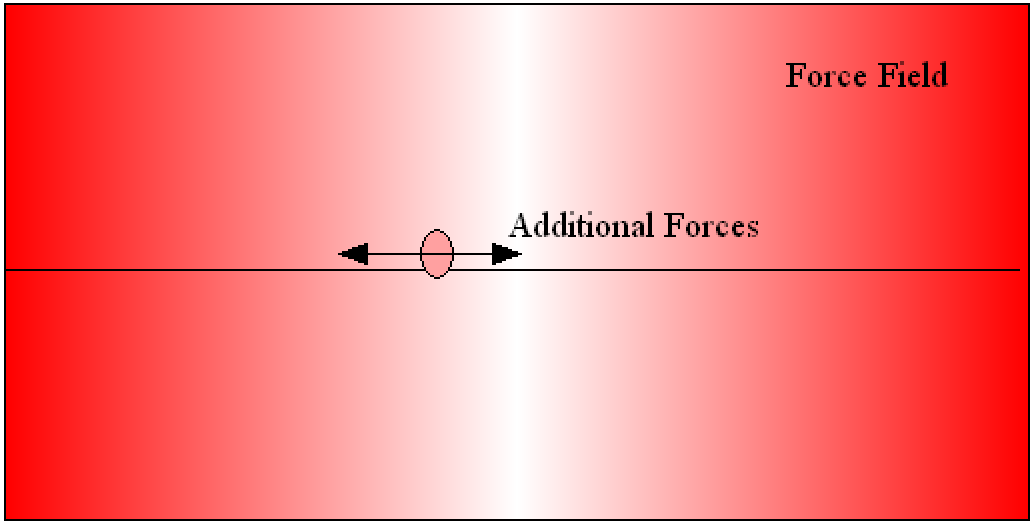
- Wolpert, Ghahramani and Jordan (1995) An internal model for sensorimotor integration. *Science* 269:1880-1882.

Lecture (2 hours)

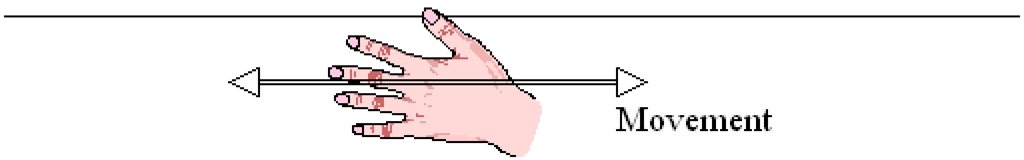
- Wolpert's human movement control experiment
- 'dynamic' sensory integration
- forward model (internal model)
- include efferent copy -> predicted sensory consequences
- mathematics behind Kalman Filter - Bayesian Derivation
- Intuition about Kalman Filter
- Interpretation of Wolpert's experimental results

EXPERIMENT WOLPERT 1995

Wolpert, Ghahramani and Jordan *(1995) An internal model for sensorimotor integration. Science 269:1880-1882.*

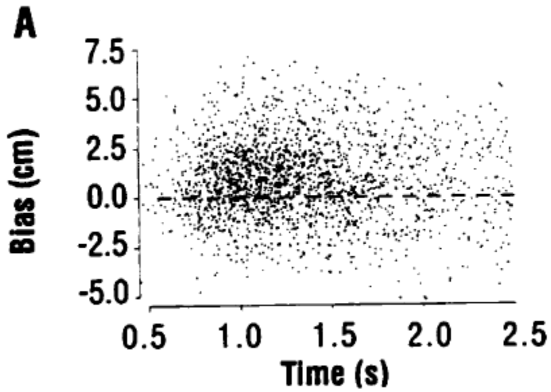


- Visually perceive hand and hand representation on screen
- Turn lights off and only see screen without hand representation
- Move hand for a few seconds along line
- Beep tone
- Stop movement
- Indicate with cursor where hand is along the line



Predict bias and variance of error between true hand position and actual hand position.

EXPERIMENTAL RESULTS



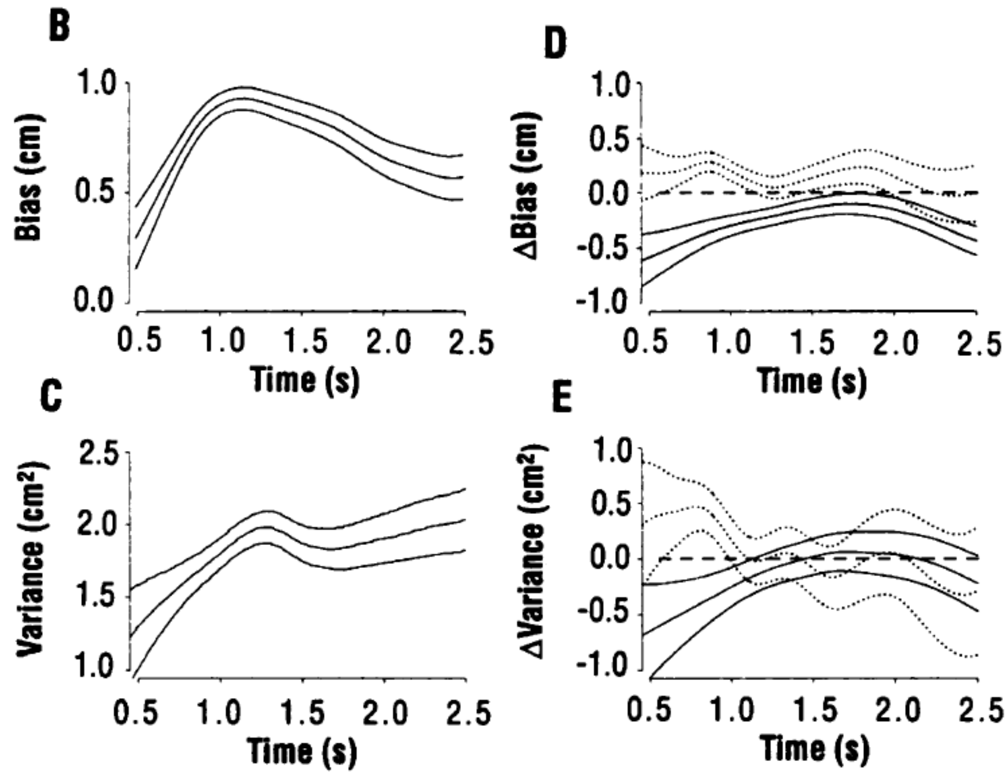


Fig. 1. The raw localization bias against movement duration is shown in (**A**) for all eight participants. There are few data points for short movement durations because of the reaction time of stopping in response to the tone. All graphs are therefore plotted from 0.5 s. (**B** through **E**) The main effect fits of the generalized additive model to the data. The propagation of the (B) bias and (C) variance of the state

How to explain and computational model these results?

The difference in variance propagation between the resistive and assistive fields was not significant over the movement; the difference in bias was significant at the $P = 0.05$ level.

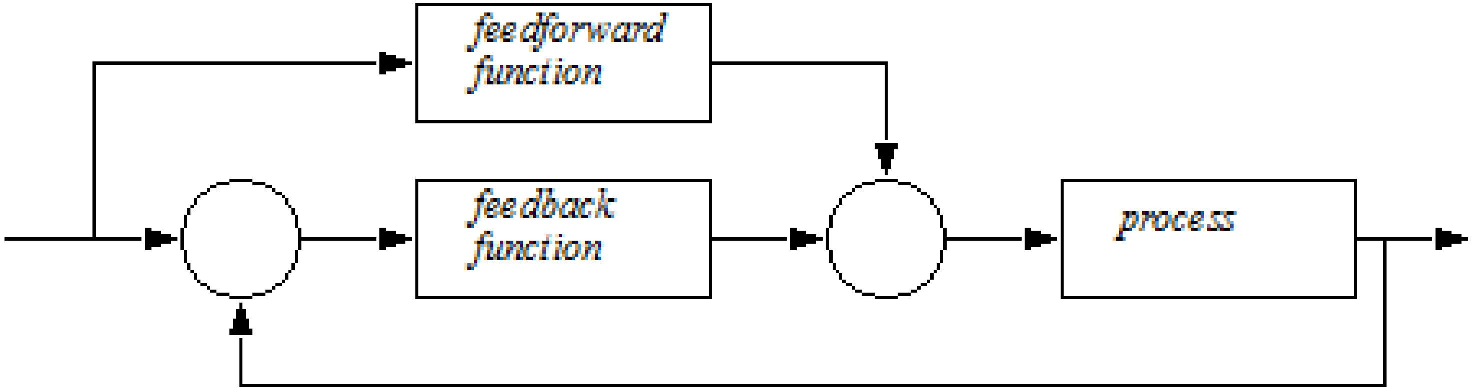
Wolpert_et_al_1995.pdf

An Internal Model for Sensorimotor Integration

Daniel M. Wolpert,* Zoubin Ghahramani, Michael I. Jordan

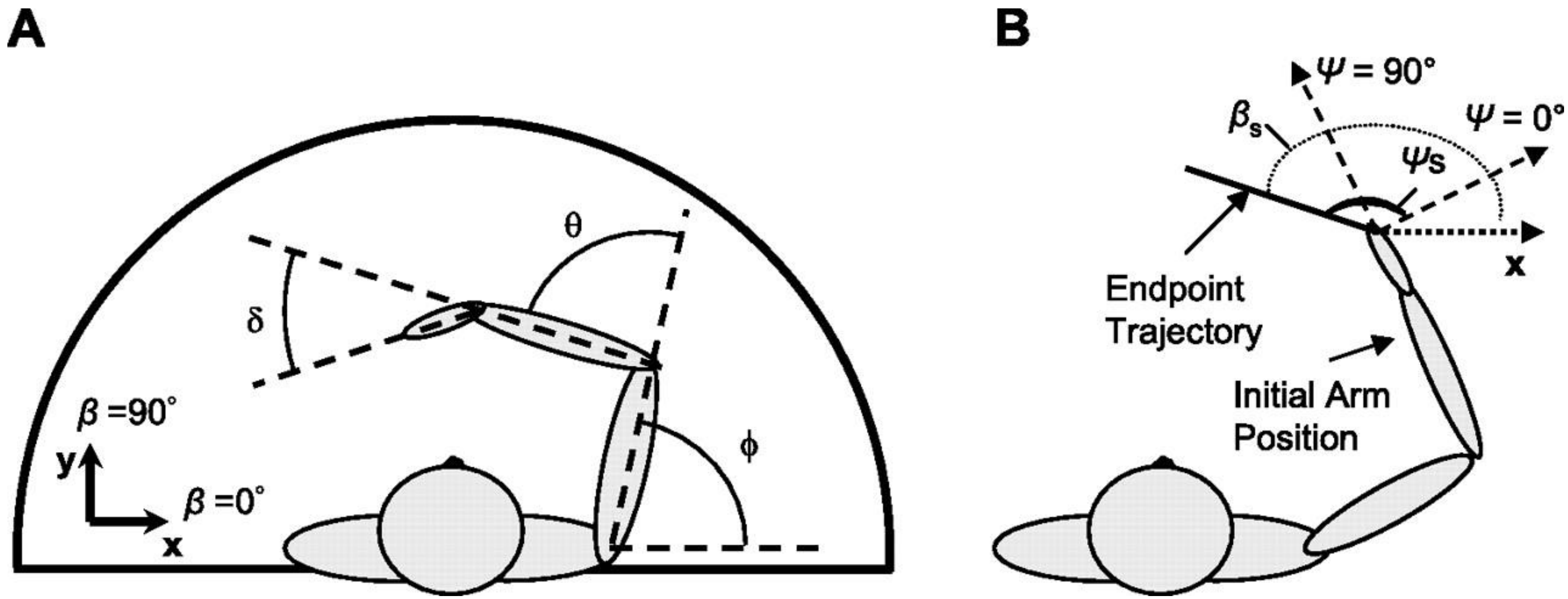
On the basis of computational studies it has been proposed that the central nervous system internally simulates the dynamic behavior of the motor system in planning, control, and learning; the existence and use of such an internal model is still under debate. A sensorimotor integration task was investigated in which participants estimated the location of one of their hands at the end of movements made in the dark and under externally imposed forces. The temporal propagation of errors in this task was analyzed within the theoretical framework of optimal state estimation. These results provide direct support for the existence of an internal model.

Every good controller has a model of the plant it controls!



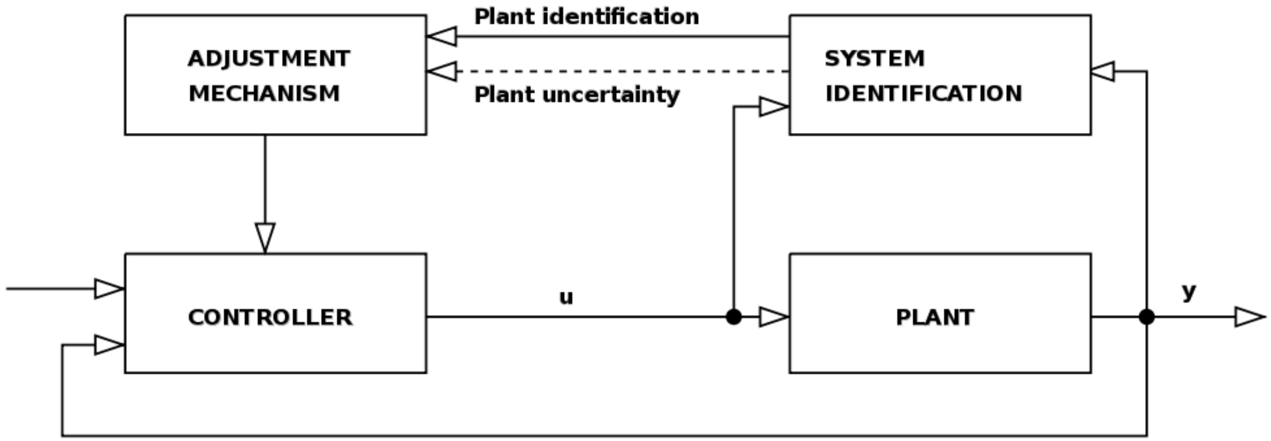
MOTION CONTROL

INTERNAL MODEL OF EFFECT OF MOTOR COMMANDS ON ARM STATE
SENSORY FEEDBACK ABOUT ACTUAL ARM STATE (PROPRIOCEPTION, VISUAL)

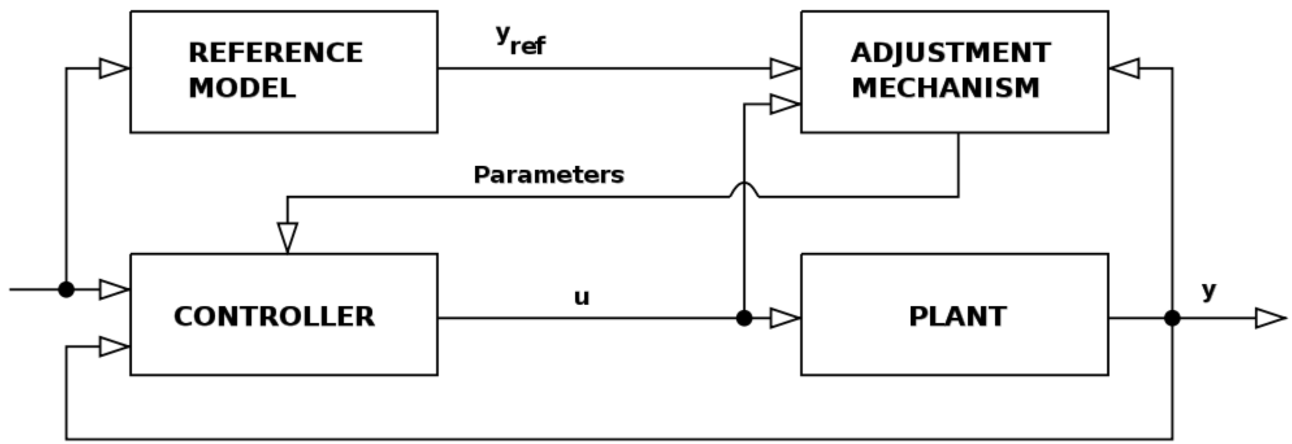


QUALITIES OF A GOOD CONTROLLER

- Control based on prediction of state capable to overcome system lag.
- Prediction requires system state estimates and/or a model of the system if the input is known.
- Kalman filter capable of providing state estimation and model parameter estimation



MODEL IDENTIFICATION ADAPTIVE CONTROL (MIAC)



MODEL REFERENCE ADAPTIVE CONTROL (MRAC)

HUMAN CONTROLLER

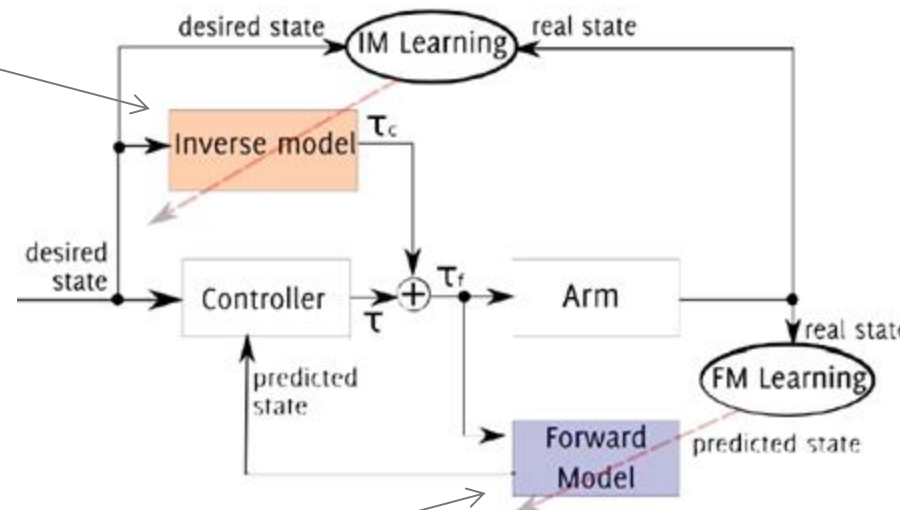
Internal model debate – do we use them or not?

WHAT ARE THE COMPONENTS IN SENSORIMOTOR CONTROL?

There are two varieties of the internal model:

Inverse models, which invert the causal flow by estimating the motor command that caused a particular state transition. Forward models have been shown to be of potential use for solving four fundamental problems in computational motor control.

Forward models, which mimic the causal flow of a process by predicting its next state (for example, position and velocity) given the current state and the motor command; and



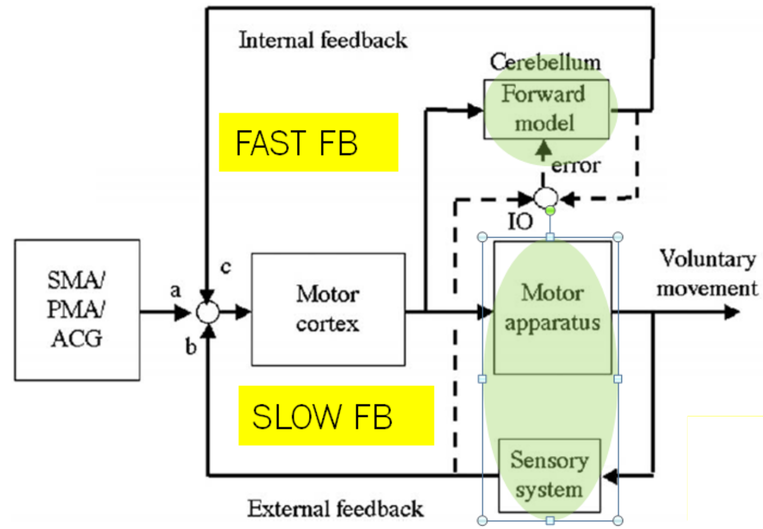
FORWARD MODELS – WHAT IS THEIR PURPOSE?

First, the delays in most sensorimotor loops are large, making feedback control too slow for rapid movements. With the use of a forward model for internal feedback, ***the outcome of an action can be estimated and used before sensory feedback is available.***

Second, a forward model is a key ingredient in a system that uses motor outflow (also called efference copy) to ***anticipate and cancel the sensory effects of movement*** (also called reafference). *Can't Tickle Yourself*

Third, a forward model can be used to transform errors between the desired and actual sensory outcome of a movement into the corresponding errors in the motor command, thereby **providing appropriate signals for *motor learning***. Similarly, by predicting the sensory outcome of the action without actually performing it, a forward model can be used in ***mental practice* to learn** to select between possible actions.

Fourth, a forward model can be used for ***state estimation*** **in which the model's prediction of the next state is combined with a reafferent sensory correction .**

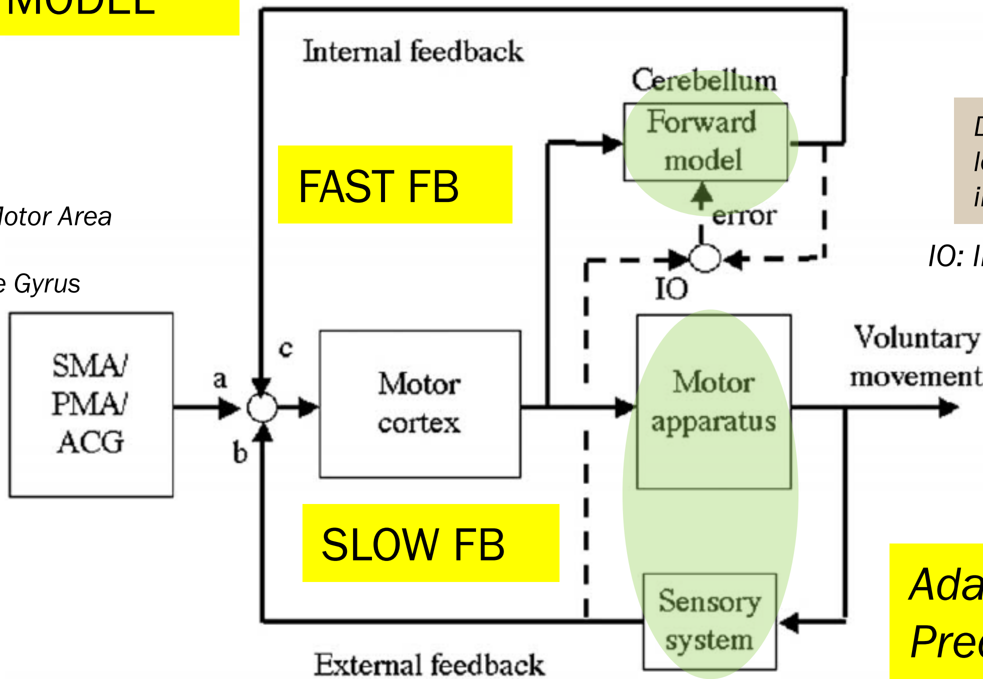


FORWARD MODEL

Sensory Cancellation

Fast Corrective FB

Supplementary Motor Area
Premotor Area
Anterior Cingulate Gyrus



Damage to the cerebellum leads to characteristic slow initiation of movement.

IO: Inferior Olive

Damage to the FM, and thus sensory cancellation, leads to a strong sensory experience from self-initiated movements as if initiated externally (alien control of self – Schizophrenia).

Adaptation on Sensory Prediction Error

Fig. 4. Forward model control system. a, Instruction signal. Note that the output of the forward model (c) functions in two ways. First, c replaces the external feedback b so that the system operates even if the external feedback is not available. Second, c cancels the effect of external feedback b disruptive to the operation of the system. Broken lines indicate how error signals are derived via the inferior olive (IO) from the sensory system monitoring the voluntary movement and the cerebellum.

Mode 1 (Fast Corrective FB): Internal FB replaces external FB

Can perform w/o external FB.

Mode 2 (Sensory Cancellation): Internal FB predicts sensory consequences and acts to cancel it

Enhanced sensitivity to errors even when FB is strong.

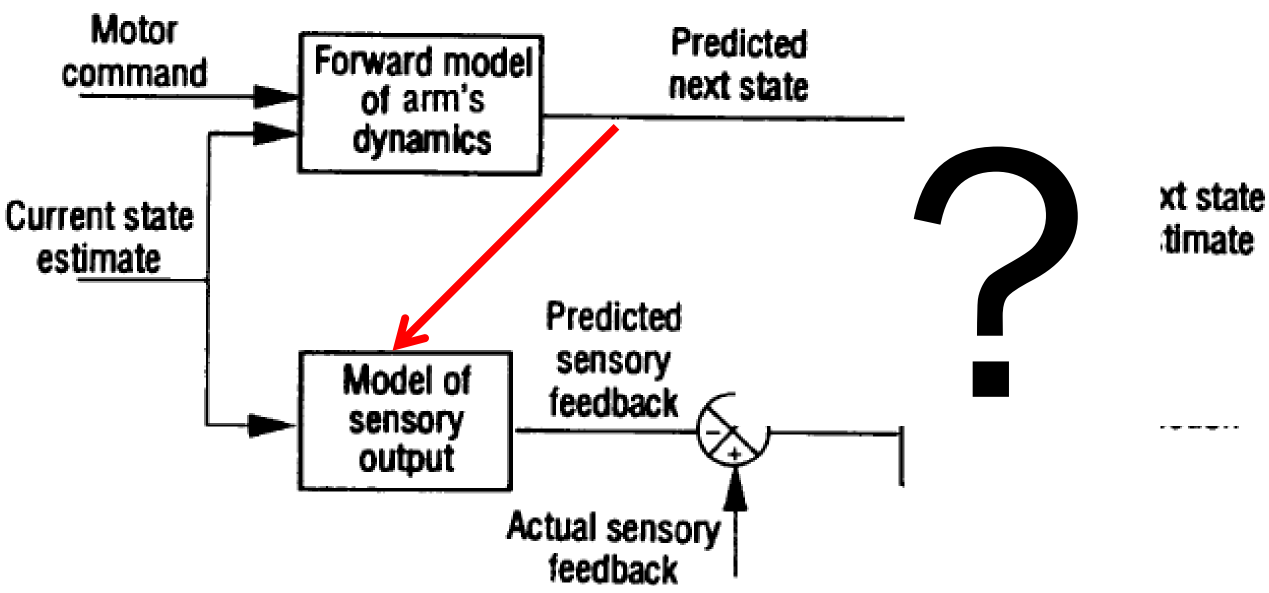
Self generated tactile stimuli experienced as less ticklish.

Reduce response to feedback that was expected.
Enhance sensitivity to unexpected responses (climbing fiber).

HUMAN CONTROLLER

WHAT ARE THE COMPONENTS IN SENSORIMOTOR INTEGRATION?

More appropriate – discrete updating



How are these estimates of the arm state combined and is a forward model even used?

WHERE, HOW, WHY KALMAN FILTER?

Wolpert, Ghahramani and Jordan (1995) An internal model for sensorimotor integration. Science 269:1880-1882.

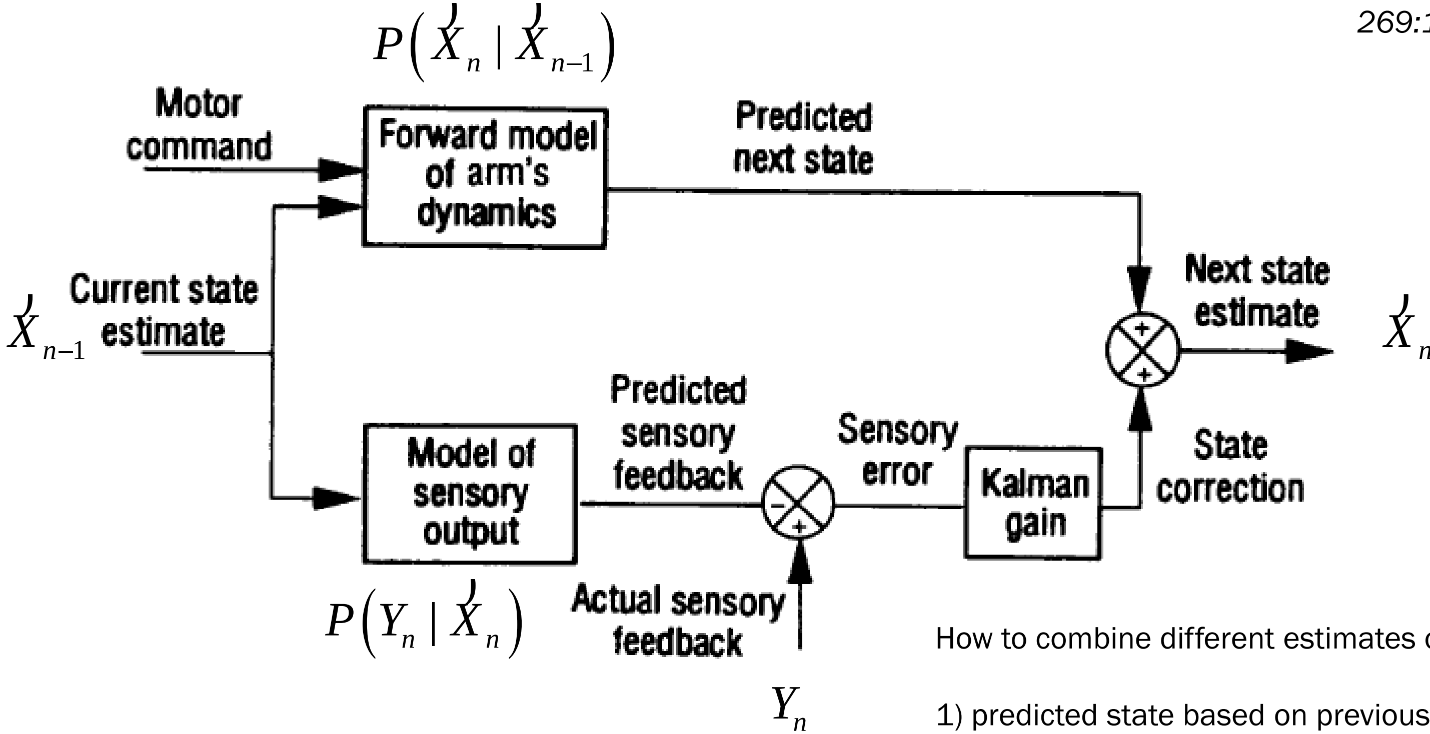


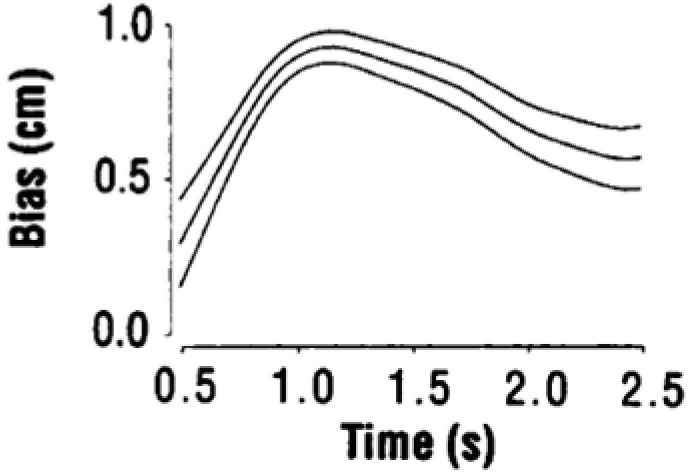
Fig. 2. (**A**) The Kalman filter model is shown schematically, consisting of two processes. The first (upper part) uses the motor command and the current state estimate to achieve a state estimate using the forward model to simulate the arm's dynamics. The second process (lower part) uses the difference between expected and actual sensory feedback to correct the forward model state estimate. The relative weighting of these two processes is mediated through the Kalman gain. (**B** through **E**) Simulated bias and variance propagation, in the same representation and scale as Fig. 1, B through E, from the Kalman filter model of the sensorimotor integration process.

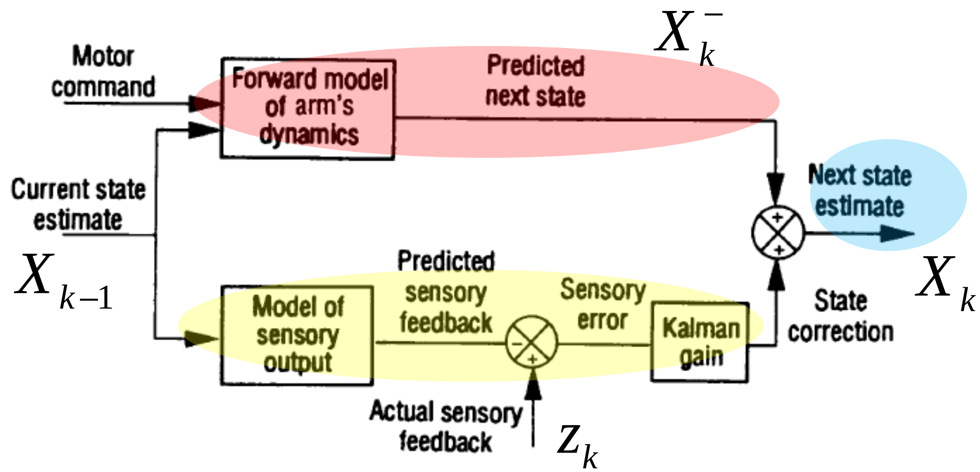
How to combine different estimates of state:

1) predicted state based on previous state plus motor command and
2) Estimated state based in observation

to arrive at a better estimate?

May 25, 2018





Therefore, during the early part of the movement, when the current state estimate is accurate, the sensorimotor integration process weights heavily the contribution of the forward model to the final estimate. However, in the later stages of the movement, when the current state estimate is less accurate, the sensory feedback must be relied on to correct for inaccuracies in the forward model. In the Kalman filter, the relative weighting shifts from the forward model toward sensory feedback over the first second of movement and then remains approximately constant, resulting in the asymptote of the variance propagation.

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

OTHER APPLICATION OF KALMAN FILTER

Automatica, Vol. 6, pp. 357–369. Pergamon Press, 1970. Printed in Great Britain.

An Optimal Control Model of Human Response
Part I: Theory and Validation*

Un modèle à commande optimale de la reponse humaine
1ere partie: Théorie et justification

Optimales Regelungsmodell der menschlichen Reaktion Teil I:
Theorie und Geltungsbereich

Оптимальнo-управляющая модель человеческой реакции
I. Часть: Теория и доказательство

D. L. KLEINMAN,† S. BARON† and W. H. LEVISON†

A mathematical model of the human as a feedback controller is developed using optimal control and estimation theory. The model is verified by comparison with experimental data.

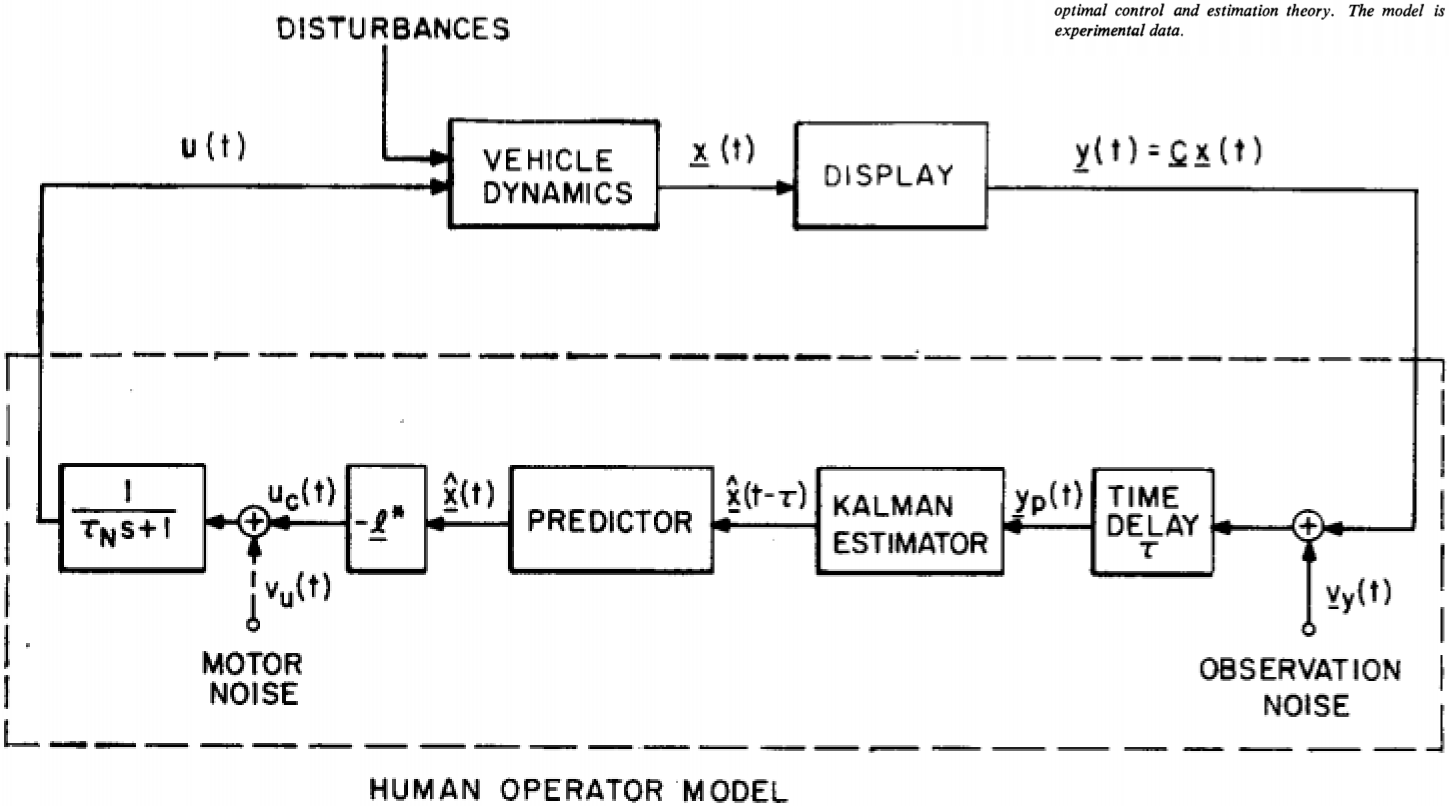
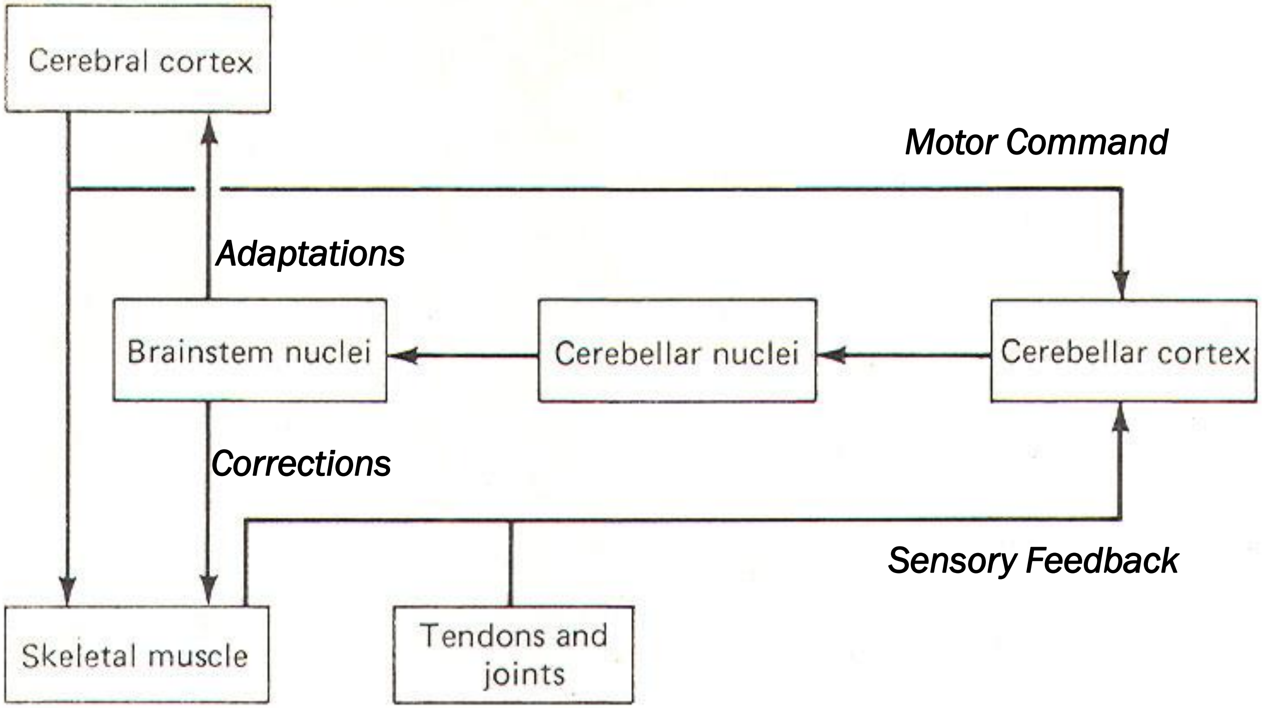
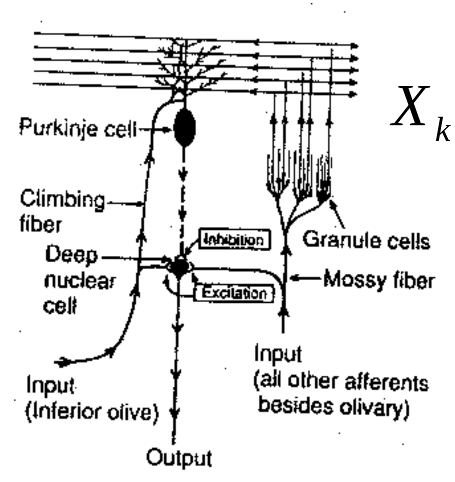


FIG. 2. Control-theoretic model of optimal human behavior.

HOW DOES THE CEREBELLUM WORK?

THE CEREBELLUM AS A COMPARATOR & COORDINATOR



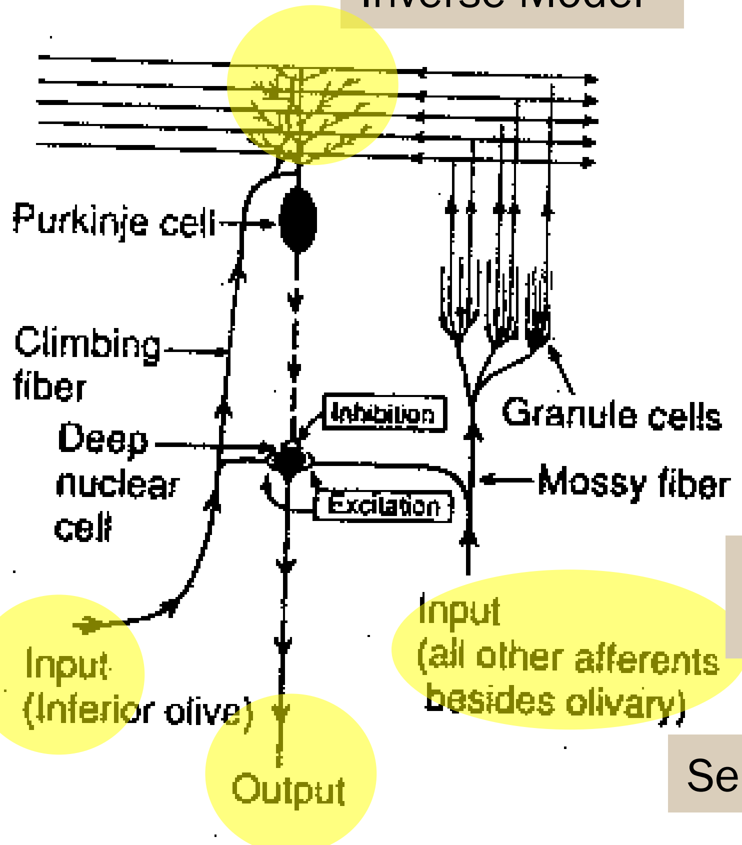


The cerebellum appears to function as a comparator, at least with respect to its role in muscle control. A sample of the motor command from the cerebral cortex to the skeletal muscles is relayed to the cerebellar cortex for evaluation. Once the motor act begins, the cerebellar cortex begins to receive input (via spinocerebellar tracts) from the proprioceptors in those muscles, tendons, and joints involved in the movement. In this way the cerebellum is in a position to compare the actual performance of a given movement with the original "intent" of the brain. Of course this comparing only has functional value if the cerebellum is capable of making adjustments when the actual performance doesn't equal the intent. As illustrated, the cerebellar cortex, through the cerebellar and brainstem nuclei, can direct corrective action both at the cortical source through ascending pathways (adaptation), as well as at the spinal cord level through descending pathways (corrections).

CEREBELLAR CIRCUIT

Map sensory state to motor corrections

Inverse Model



Weaken effect of sensory states that produce an error (reinforcement learning)

Adaptation

Corrective Command
EXCITATION

Error Signal

Intended minus observed or predicted limb control.
OR
Corrective action was needed to reach goal state. → next slide

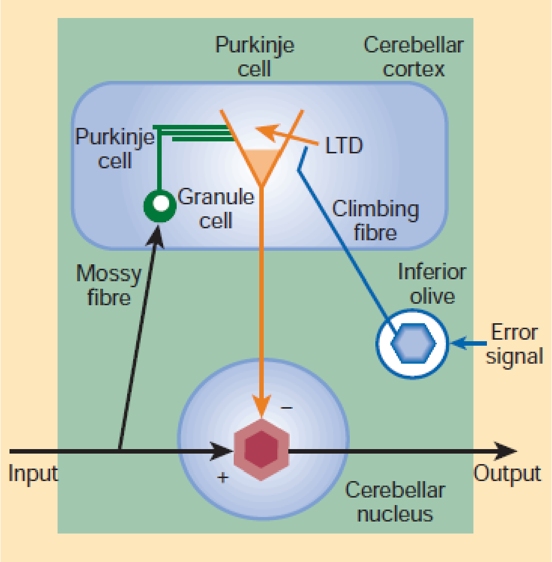
Motor Correction

Signal to muscles as well as motor cortex.

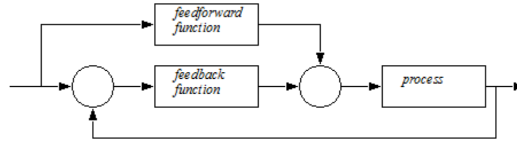
Motor Cortex Command
EXCITATION

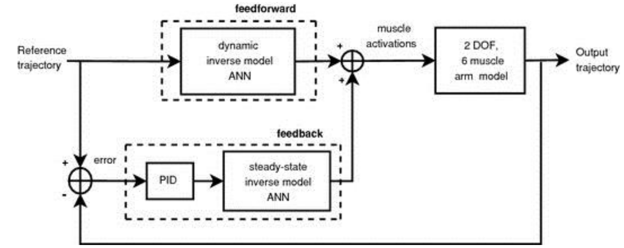
Sensory State / Context

State of environment (context) and limbs.



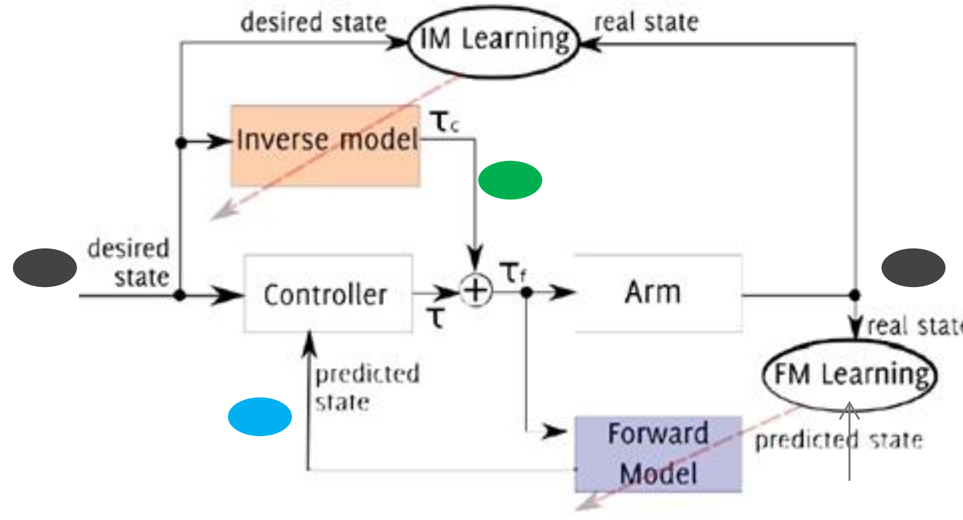
CEREBELLUM AS A LEARNING MACHINE





<u>Predictive Control</u>

- **Inverse Model**
  - From predicted body state to required control
  - From perceived target (desired) state to fast forward control
  - Act on sensory perceived future goal
- **Feed Forward Model**
  - From issued control to predicted sensory inputs
  - Efferent copies of motor commands to generate fast feedback control
  - React to neurologically computed expected error



Sensory State

Motor Command / Correction

Cortex Command / Correction

Combining FM and IM

A NEW COUPLING SCHEME OF CEREBELLAR INTERNAL MODELS: ONLINE AND OFFLINE ADAPTATION IN PROCEDURAL TASKS

Jean-Baptiste Passot[1] and Angelo Arleo[1]
[1]CNRS - UPMC Univ Paris 6, UMR 7102, F75005, Paris, France
Corresponding author: jbpassot@snv.jussieu.fr

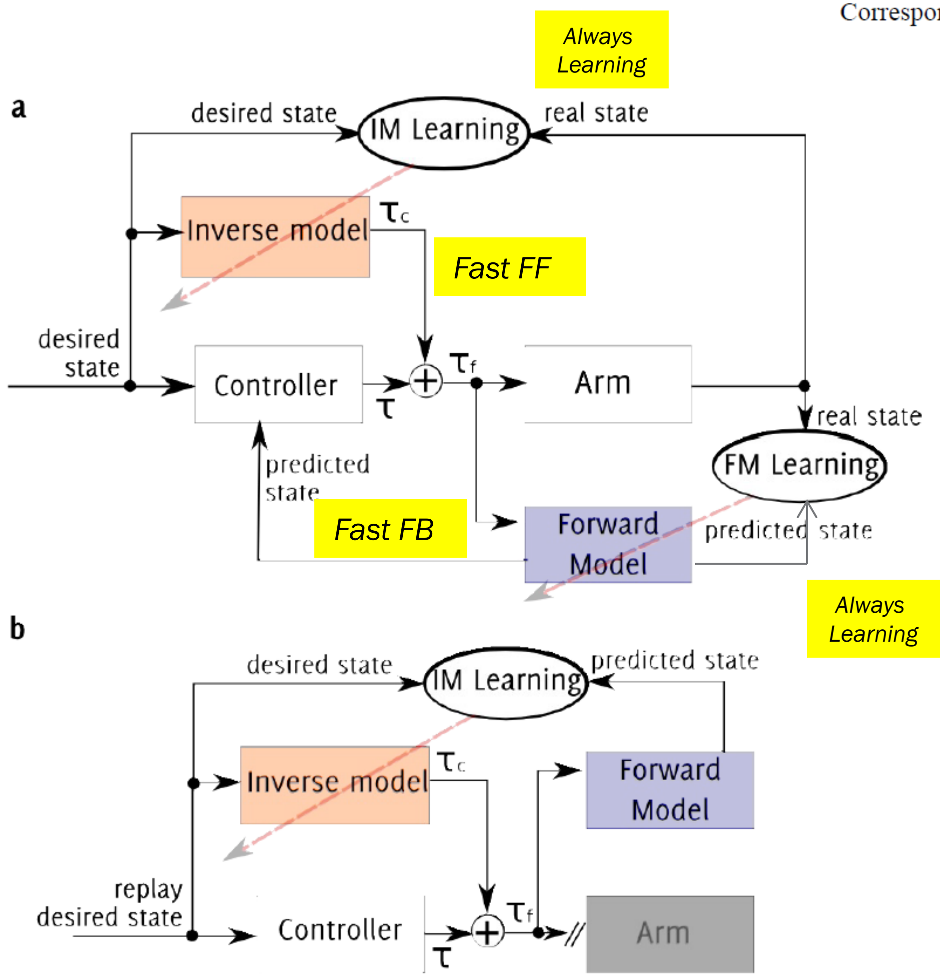


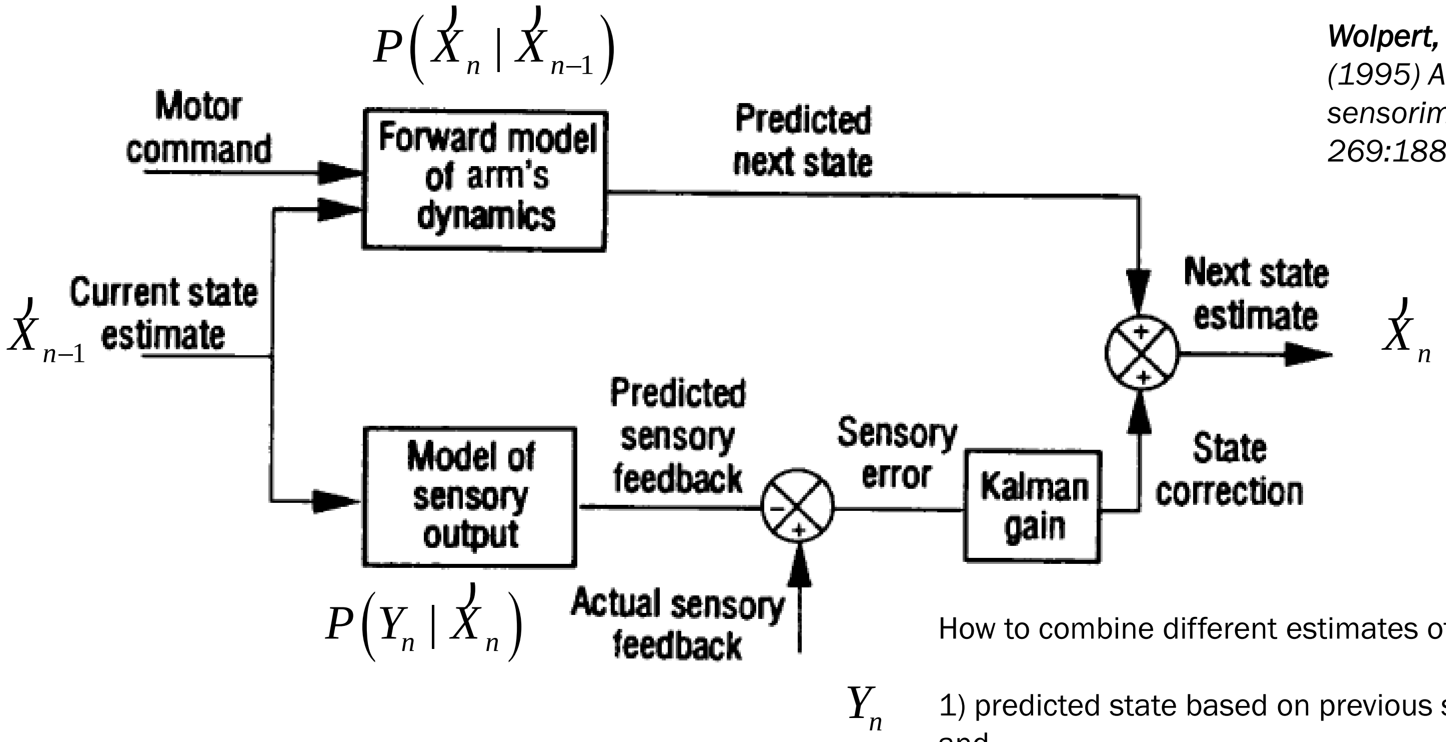
Fig. 1. *Coupling scheme for online and offline motor learning.* (a) Online adaptation. The arm controller receives the desired state and maps it onto a motor command (τ). The desired state is also sent to the inverse model that acts as a feed-forward corrector and calculates the motor correction (τ_c). The resulting command (τ_f) is then sent to the arm actuators. By comparing the desired state against the sensed real state, the inverse model learns to reduce the error between desired and real arm positions. While the motor command τ_f is being sent to the arm, an efference copy of the order is also conveyed to the forward model that learns to predict the consequent future position of the arm. The predicted state is then sent to the arm controller that can recalculate a new trajectory if the expected position in the trajectory differs from the predicted one. Finally, the real state is used to adapt the forward model to mimic the motor apparatus of the arm. (b) Offline adaptation. During offline processing, sensory feedbacks (i.e. the real state signals driving forward and inverse model learning) are not available. Yet, if the forward model is at least partially learnt, the predicted state signals can be used to continue to train the inverse model.

Cerebellum always adapting.

Arthur Rubinstein

Learn new maneuvers lying in bed.

WHERE, HOW, WHY KALMAN FILTER?



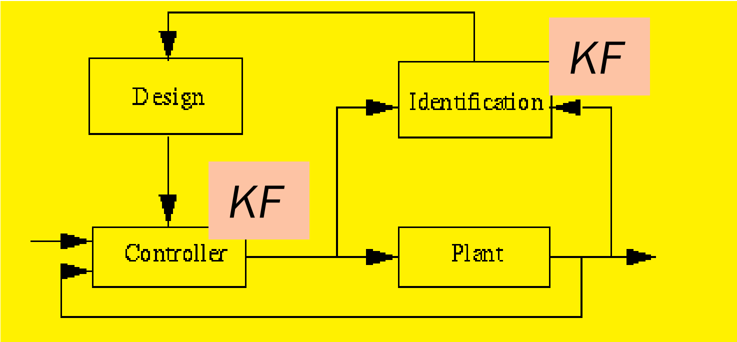
Wolpert, Ghahramani and Jordan (1995) An internal model for sensorimotor integration. Science 269:1880-1882.

How to combine different estimates of state:

1) predicted state based on previous state plus motor command and
2) Estimated state based in observation

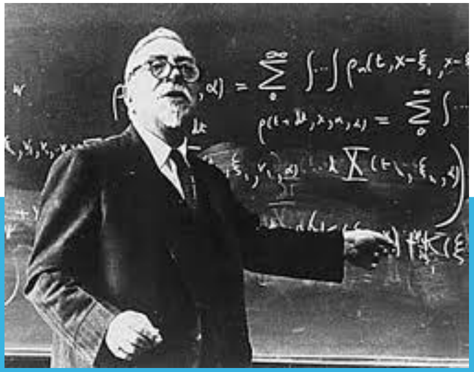
to arrive at a better estimate?

ORIGIN 1960



- Introduction of digital computers creates practical engineering needs for:
  - Separation of random signals from noise
  - Prediction of random signals → clean estimate of derivatives
  - Estimation of system state → state can be hidden
  - Estimation of system parameters → from input and output signals
  - Estimation of time varying system parameters → systems change as a function of conditions as well as state (i.e. nonlinear)
- The Kalman filter offers discrete design solutions to all; developed by Rudy Kalman in response to Norbert Wiener's filter.

Estimation & Prediction for Control and System Identification



LETS DERIVE KALMAN FILTER FOR SCALAR CASE FIRST THEN ALSO FOR THE FULL MATRIX CASE

R. E. KALMAN
Research Institute for Advanced Study,[2]
Baltimore, Md.

A New Approach to Linear Filtering and Prediction Problems[1]

The classical filtering and prediction problem is re-examined using the Bode-Shannon representation of random processes and the "state transition" method of analysis of dynamic systems. New results are:

(1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to growing-memory and infinite-memory filters.

(2) A nonlinear difference (or differential) equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the coefficients of the difference (or differential) equation of the optimal linear filter are obtained without further calculations.

(3) The filtering problem is shown to be the dual of the noise-free regulator problem.

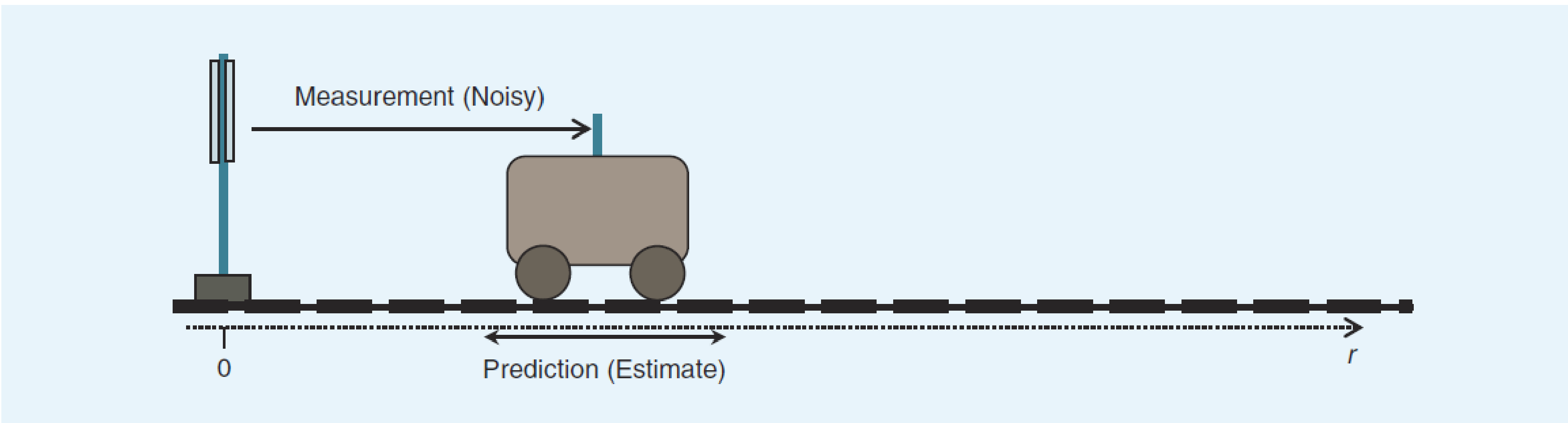
The new method developed here is applied to two well-known problems, confirming and extending earlier results.

The discussion is largely self-contained and proceeds from first principles; basic concepts of the theory of random processes are reviewed in the Appendix.

Transactions of the ASME–Journal of Basic Engineering, 82 (Series D): 35-45. Copyright © 1960 by ASME

HOW TO ACCURATELY ESTIMATE WHERE THE TRAIN / HAND IS?

State X
Measurement / Observation Y



[FIG1] This figure shows the one-dimensional system under consideration.

LETS START TO SEND CONTROL SIGNALS

Control U

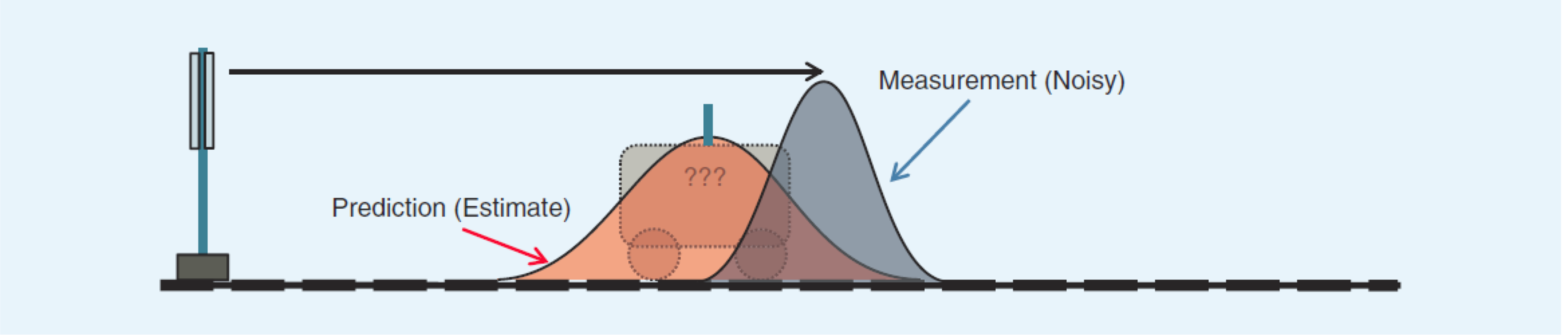
[FIG2] The initial knowledge of the system at time $t = 0$. The red Gaussian distribution represents the pdf providing the initial confidence in the estimate of the position of the train. The arrow pointing to the right represents the known initial velocity of the train.

OBTAINING

PREDICTION (BASED ON CONTROL)

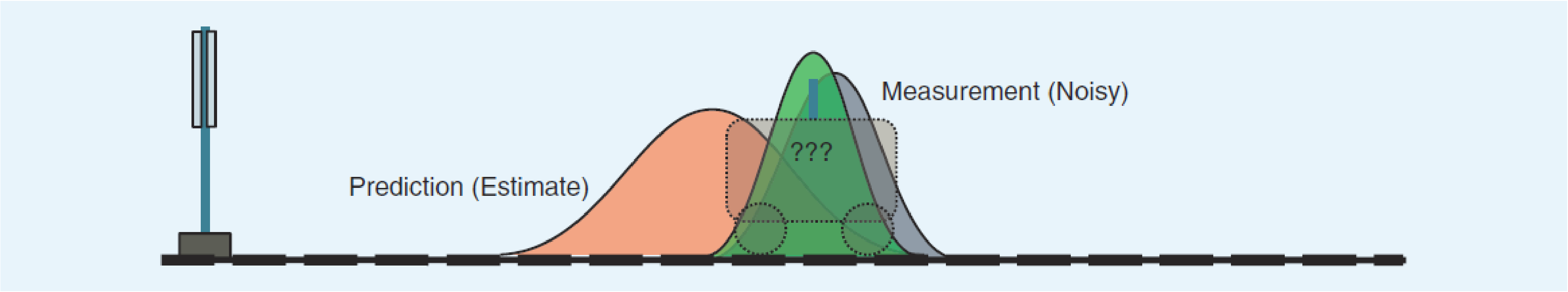
&

MEASUREMENT (NOISY)



[FIG4] Shows the measurement of the location of the train at time $t = 1$ and the level of uncertainty in that noisy measurement, represented by the blue Gaussian pdf. The combined knowledge of this system is provided by multiplying these two pdfs together.

COMBINING TWO "ESTIMATES": PREDICTION & MEASUREMENT



[FIG5] Shows the new pdf (green) generated by multiplying the pdfs associated with the prediction and measurement of the train's location at time $t = 1$. This new pdf provides the best estimate of the location of the train, by fusing the data from the prediction and the measurement.

CONVOLUTION: PROBABILITY AT EACH POSSIBLE VALUE OF RANDOM VARIABLE 1 IS MULTIPLIED BY DISTRIBUTION OF RANDOM VARIABLE 2. ALL THESE PROBABILITIES ARE ADDED TOGETHER.

The general formula for the distribution of the sum $Z = X + Y$ of two independent discrete variables is

$$P(Z=z) = \sum_{k=-\infty}^{\infty} P(X=k)P(Y=z-k).:$$

The counterpart for independent continuous variables with density functions $f(x), g(y)$ is

$$h(z) = (f*g)(z) = \int_{-\infty}^{\infty} f(z-t)g(t)dt = \int_{-\infty}^{\infty} f(t)g(z-t)dt.$$

If X and Y are independent random variables that are normally distributed (and therefore also jointly so), then their sum is also normally distributed. i.e., if

$$X \sim N(\mu_X, \sigma_X^2)$$
$$Y \sim N(\mu_Y, \sigma_Y^2)$$
$$Z = X + Y,$$

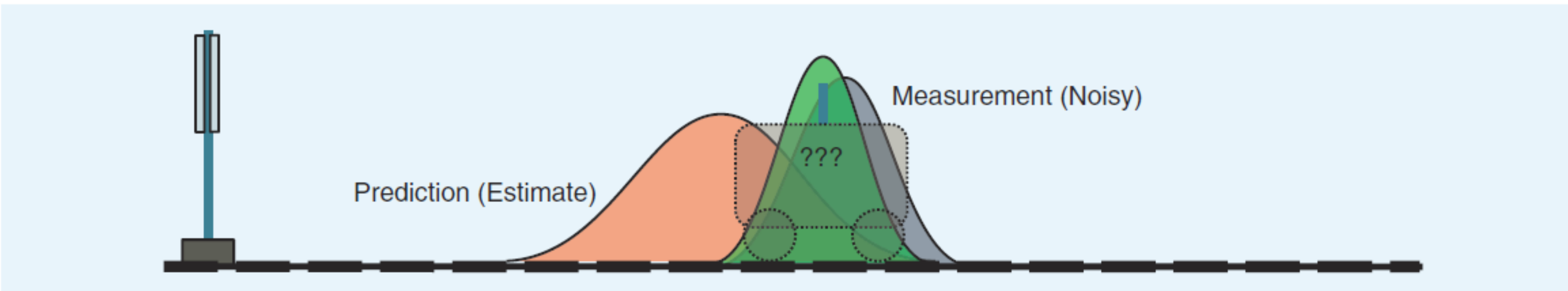
then

$$Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

Is this what we want?

NO BECAUSE THE VARIANCE KEEPS INCREASING – NO USE COMBINING HE TWO TO IMPROVE ESTIMATE

HOW TO COMBINE



[FIG5] Shows the new pdf (green) generated by multiplying the pdfs associated with the prediction and measurement of the train's location at time $t = 1$. This new pdf provides the best estimate of the location of the train, by fusing the data from the prediction and the measurement.

$$N_1(\mu_1, \sigma_1) \qquad N_2(\mu_2, \sigma_2)$$

WEIGHTED SUM

$$N_{new}\left(\mu_{new},\sigma_{new}\right)=w_1 N_1\left(\mu_1,\sigma_1\right)+w_2 N_2\left(\mu_2,\sigma_2\right)$$

What should w1 and w2 be?

COMBINING NOISY MEASUREMENTS OF SAME VARIABLE

Measurements x_1 with std σ_1 and x_2 with std σ_2 are independent measurements of the same variable x_3.

$$x_3 = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}, \text{ multiply numerator and demonimator by } \sigma_1^2\sigma_2^2$$

$$x_3 = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}$$

$$\operatorname{cov}(x_3) = E\left\{x_3^2\right\} = E\left\{\left(\frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}\right)^2\right\} = E\left\{\frac{\sigma_2^4 x_1^2 + \sigma_1^4 x_2^2 + 2\sigma_2^2 x_1 \sigma_1^2 x_2}{\left(\sigma_1^2 + \sigma_2^2\right)^2}\right\}$$

$$= \frac{\sigma_2^4\sigma_1^2 + \sigma_1^4\sigma_2^2}{\left(\sigma_1^2 + \sigma_2^2\right)^2} = \frac{\sigma^2\sigma_1^2\left(\sigma_1^2 + \sigma_2^2\right)}{\left(\sigma_1^2 + \sigma_2^2\right)^2} = \frac{\sigma_2^2\sigma_1^2}{\left(\sigma_1^2 + \sigma_2^2\right)}$$

Handling multiple uncertain measurements of system:

- *Measurements can be optimally combined prior to feeding to the Kalman Filter → Centralized Kalman Filter*
- *A Kalman filter can be created on each measurement and Kalman filter estimates can be combined optimally. → Decentralized Kalman Filter*

HOW TO COMBINE TWO GAUSSIAN DISTRIBUTIONS

$$y_1(r;\mu_1,\sigma_1) \triangleq \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(r-\mu_1)^2}{2\sigma_1^2}}.$$

$$y_2(r;\mu_2,\sigma_2) \triangleq \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(r-\mu_2)^2}{2\sigma_2^2}}.$$

$$\begin{aligned} & y_{\text{fused}}(r;\mu_1,\sigma_1,\mu_2,\sigma_2) \\ &= \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(r-\mu_1)^2}{2\sigma_1^2}} \times \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(r-\mu_2)^2}{2\sigma_2^2}} \\ &= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}}\, e^{-\left(\frac{(r-\mu_1)^2}{2\sigma_1^2}+\frac{(r-\mu_2)^2}{2\sigma_2^2}\right)}. \qquad (10) \end{aligned}$$

Combination

$$y_{\text{fused}}(r;\mu_1,\sigma_1,\mu_2,\sigma_2)$$
$$= \frac{1}{\sqrt{2\pi\sigma_1^2}}e^{-\frac{(r-\mu_1)^2}{2\sigma_1^2}} \times \frac{1}{\sqrt{2\pi\sigma_2^2}}e^{-\frac{(r-\mu_2)^2}{2\sigma_2^2}}$$
$$= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}}\, e^{-\left(\frac{(r-\mu_1)^2}{2\sigma_1^2}+\frac{(r-\mu_2)^2}{2\sigma_2^2}\right)}. \qquad (10)$$

Target Distribution Formulation

$$y_{\text{fused}}(r;\mu_{\text{fused}},\sigma_{\text{fused}})$$
$$= \frac{1}{\sqrt{2\pi\sigma_{\text{fused}}^2}}e^{-\frac{(r-\mu_{\text{fused}})^2}{2\sigma_{\text{fused}}^2}}, \qquad (11)$$

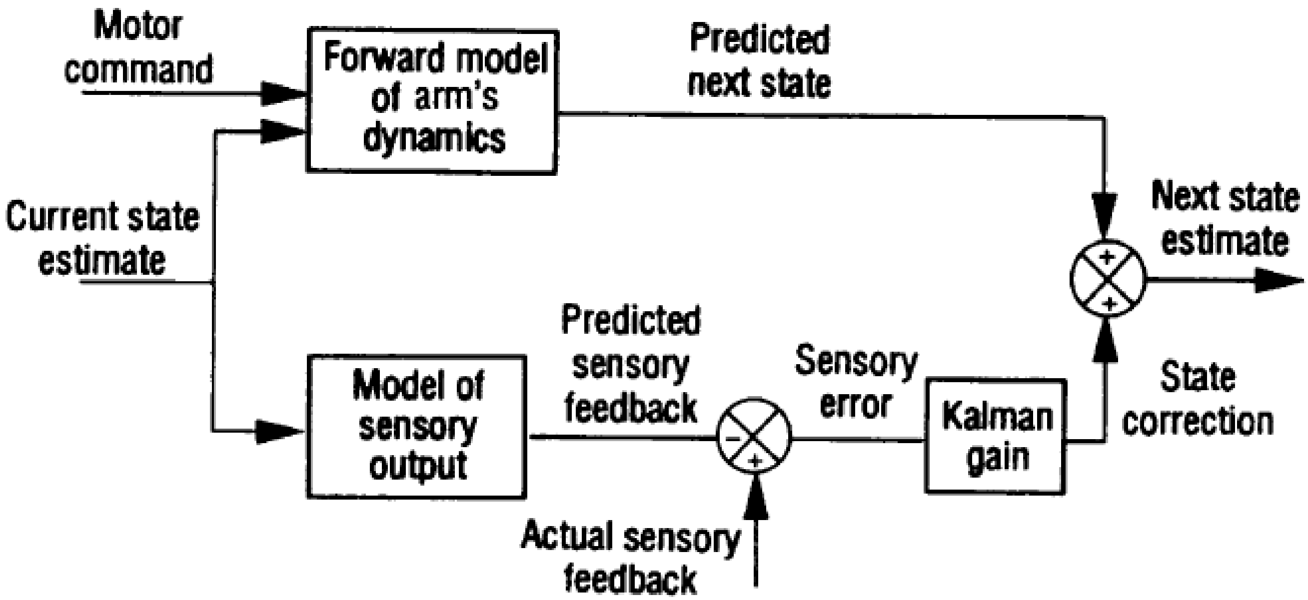
where *Lengthy Derivation*

$$\mu_{\text{fused}} = \frac{\mu_1\sigma_2^2+\mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}$$
$$= \mu_1 + \frac{\sigma_1^2(\mu_2-\mu_1)}{\sigma_1^2+\sigma_2^2} \qquad (12)$$

and

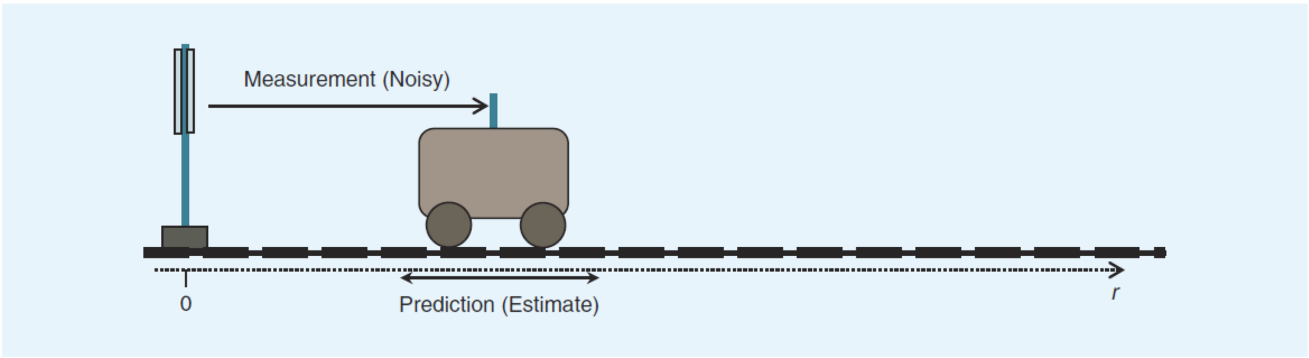
$$\sigma_{\text{fused}}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} = \sigma_1^2 - \frac{\sigma_1^4}{\sigma_1^2+\sigma_2^2}.$$

WHAT ARE THESE $N_1(\mu_1, \sigma_1)$ AND $N_2(\mu_2, \sigma_2)$



Depends on how we represent the system we are controlling!

STATE SPACE EQUATIONS
WE WANT TO ESTIMATE HIDDEN STATES



[FIG1] This figure shows the one-dimensional system under consideration.

State X
Measurement / Observation Z

$$x_n = Ax_{n-1} + Bu_n + w_n$$

$$y_n = Cx_n + v_n$$

SCALAR KALMAN FILTER DERIVATION

$$X_n = AX_{n-1} + BU_n + w_n \qquad Y_n = CX_n + v_n$$

$$x_n = ax_{n-1} + w_n$$

$$y_n = cx_n + v_n$$

Estimate of x based on prior state information $\tilde{x}_n^{pred} = a\hat{x}_{n-1}$

Estimate of x based on observation $\tilde{x}_n^{obs} = \frac{y_n}{c}$

$$x_n = ax_{n-1} + w_n$$

$$y_n = cx_n + v_n$$

Given :

$$\tilde{x}_n^{pred} = a\hat{x}_{n-1}$$

$$\tilde{x}_n^{obs} = \frac{y_n}{c}$$

Variances :

$$\tilde{P}_n^{pred} = aP_n + Q$$

$$\tilde{P}_n^{obs} = \frac{R}{c^2}$$

Combining

$\tilde{x}_n^{pred}$ *and* $\tilde{x}_n^{obs}$ *to* $\hat{x}_n$

Variance weighted combination of independent estimates of x_n

Given:

$$\tilde{x}_n^{pred} = a\hat{x}_{n-1}$$

$$\tilde{x}_n^{obs} = \frac{y_n}{c}$$

Variances:

$$\tilde{P}_n^{pred} = aP_n + Q$$

$$\tilde{P}_n^{obs} = \frac{R}{c^2}$$

$$\hat{x}_n = \frac{\dfrac{\tilde{x}_n^{pred}}{\tilde{P}_n^{pred}} + \dfrac{\tilde{x}_n^{obs}}{\tilde{P}_n^{obs}}}{\dfrac{1}{\tilde{P}_n^{pred}} + \dfrac{1}{\tilde{P}_n^{obs}}} = \frac{\tilde{P}_n^{obs}\tilde{x}_n^{pred} + \tilde{P}_n^{pred}\tilde{x}_n^{obs}}{\tilde{P}_n^{obs} + \tilde{P}_n^{pred}}$$

$$\hat{x}_n = \frac{\dfrac{R}{c^2}a\hat{x}_{n-1} + \tilde{P}_n^{pred}\dfrac{y_n}{c}}{\dfrac{R}{c^2} + \tilde{P}_n^{pred}} = \frac{Ra\hat{x}_{n-1} + \tilde{P}_n^{pred}cy_n}{R + c^2\tilde{P}_n^{pred}} = \frac{R}{R + c^2\tilde{P}_n^{pred}}a\hat{x}_{n-1} + \frac{\tilde{P}_n^{pred}cy_n}{R + c^2\tilde{P}_n^{pred}}$$

$$\hat{x}_n = \left(\frac{R + c^2\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}} - \frac{c^2\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}}\right)a\hat{x}_{n-1} + \frac{\tilde{P}_n^{pred}cy_n}{R + c^2\tilde{P}_n^{pred}} = \left(1 - \frac{c^2\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}}\right)a\hat{x}_{n-1} + \frac{\tilde{P}_n^{pred}c^2}{R + c^2\tilde{P}_n^{pred}}\frac{y_n}{c}$$

$$\hat{x}_n = \left(1 - K'\right)a\hat{x}_{n-1} + K'\frac{y_n}{c} = a\hat{x}_{n-1} + K'\left(\frac{y_n}{c} - a\hat{x}_{n-1}\right) = a\hat{x}_{n-1} + \frac{K}{c}\left(y_n - ca\hat{x}_{n-1}\right)$$

$$K' = \frac{c^2\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}}, K = \frac{c\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}}$$

Variance weighted combination of variances of the independent estimates of x_n

$$P_n = \frac{\dfrac{\tilde{P}_n^{pred}}{\tilde{P}_n^{pred}} + \dfrac{\tilde{P}_n^{obs}}{\tilde{P}_n^{obs}}}{\dfrac{1}{\tilde{P}_n^{pred}} + \dfrac{1}{\tilde{P}_n^{obs}}} = \frac{\dfrac{R}{c^2}\tilde{P}_n^{pred}}{\dfrac{R}{c^2} + \tilde{P}_n^{pred}} = \frac{R\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}} = \frac{R}{R + c^2\tilde{P}_n^{pred}}\tilde{P}_n^{pred}$$

$$P_n = \frac{R}{R + c^2\tilde{P}_n^{pred}}\tilde{P}_n^{pred} = \left(1 - \frac{c^2\tilde{P}_n^{pred}}{R + c^2\tilde{P}_n^{pred}}\right)\tilde{P}_n^{pred} = \left(1 - K'\right)\tilde{P}_n^{pred} = \left(1 - Kc\right)\tilde{P}_n^{pred}$$

$Given:$

$\tilde{x}_n^{pred} = a\hat{x}_{n-1}$

$\tilde{x}_n^{obs} = \frac{y_n}{c}$

$$\hat{x}_n = a\hat{x}_{n-1} + K\left(\frac{y_n}{c} - a\hat{x}_{n-1}\right)$$

$Variances:$

$\tilde{P}_n^{pred} = aP_n + Q$

$$P_n = (1-K)\tilde{P}_n^{pred}$$

$\tilde{P}_n^{obs} = \frac{R}{c^2}$

KALMAN FILTER EQUATIONS

State Space Representation $$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$ *Cov(w) = Q*

Known A,B,H

$$z_k = Hx_k + v_k$$ *Cov(v) = R*

Prediction $$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$ *Grows with Q*

Innovation $$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$ ***Target Kalman Filter Form***

OPTIMALLY COMBINING NOISY MEASUREMENTS

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$y_{\text{fused}}(r; \mu_{\text{fused}}, \sigma_{\text{fused}}) = \frac{1}{\sqrt{2\pi\sigma_{\text{fused}}^2}} e^{-\frac{(r-\mu_{\text{fused}})^2}{2\sigma_{\text{fused}}^2}}, \qquad (11)$$

where

$$\mu_{\text{fused}} = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2} = \mu_1 + \frac{\sigma_1^2(\mu_2 - \mu_1)}{\sigma_1^2 + \sigma_2^2} \qquad (12)$$

and

$$\sigma_{\text{fused}}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \sigma_1^2 - \frac{\sigma_1^4}{\sigma_1^2 + \sigma_2^2}.$$

COMBINING TWO MEASUREMENTS

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

- Simplify KF with scalar unit A and H and no B.
- The available signals are a current best estimate and a new measurement.

$\hat{x}$ with covariance matrix P

$\hat{x}^-$ with covariance matrix P^- (a posteriori estimate of $\hat{x}$ and P)

z with covariance matrix R, the measurement has a noise at each time step with cov(v) = R

where

$P^- = P + Q$, the state can change each time step with cov(w) = Q

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

- The new state estimate is a weighted sum of $\hat{x}^-$ and z

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$\hat{x} = \frac{R}{R+P^-}\hat{x}^- + \frac{P^-}{R+P^-}z = \left(1 - \frac{P^-}{R+P^-}\right)\hat{x}^- + \frac{P^-}{R+P^-}z$$

$$= \hat{x}^- + \frac{P^-}{R+P^-}\left(z - \hat{x}^-\right) = \hat{x}^- + K\left(z - \hat{x}^-\right)$$

Prediction *Kalman Gain* *Innovation*

$$x_3 = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}, \text{ multiply numerator and demonimator by } \sigma_1^2\sigma_2^2$$

$$x_3 = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}$$

$$\operatorname{cov}(x_3) = E\left\{x_3^2\right\} = E\left\{\left(\frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}\right)^2\right\} = E\left\{\frac{\sigma_2^4 x_1^2 + \sigma_1^4 x_2^2 + 2\sigma_2^2 x_1 \sigma_1^2 x_2}{\left(\sigma_1^2 + \sigma_2^2\right)^2}\right\}$$

$$= \frac{\sigma_2^4\sigma_1^2 + \sigma_1^4\sigma_2^2}{\left(\sigma_1^2 + \sigma_2^2\right)^2} = \frac{\sigma^2\sigma_1^2\left(\sigma_1^2 + \sigma_2^2\right)}{\left(\sigma_1^2 + \sigma_2^2\right)^2} = \frac{\sigma_2^2\sigma_1^2}{\left(\sigma_1^2 + \sigma_2^2\right)}$$

SOME INTUITIONS

Q

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

R

$$z_k = Hx_k + v_k$$

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

P

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P = \mathrm{cov}(\hat{x}) = \frac{P^- R}{P^- + R}$$

with

$$\mathrm{K} = \frac{P^-}{P^- + R}$$

we obtain

$$\mathrm{P} = \frac{R}{P^- + R} P^- = \left(1 - \frac{P^-}{P^- + R}\right) P^-$$

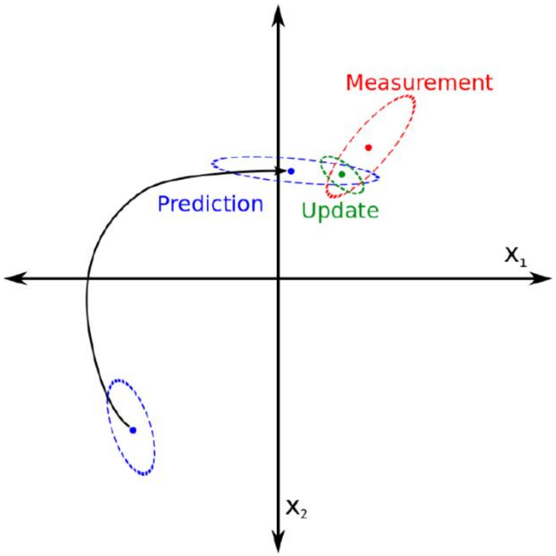
$$= (1 - K) P^-$$

$$P^- = P + Q$$

DERIVATION OF FULL KALMAN FILTER FROM BAYESIAN PRINCIPLES

State Equations:

$$\begin{aligned} \boldsymbol{x}_{n+1} &= \boldsymbol{F}_n \boldsymbol{x}_n + \boldsymbol{\nu}_{o,n+1} \\ \boldsymbol{y}_n &= \boldsymbol{\Phi} \boldsymbol{x}_n + \boldsymbol{\nu}_{d,n}. \end{aligned}$$



Bayes Principles:

$$p(x,y) = p(x\,|\,y)\,p(y) = p(y\,|\,x)\,p(x)$$

Joint probability of state and observation same is
Likelihood of state given observation times prior probability of observation
Likelihood of observation given state times prior probability of state

$$\hat{\boldsymbol{x}}_n = \arg\max_{\boldsymbol{x}_n} \left[p(\boldsymbol{y}_n | \boldsymbol{x}_n) p_{\hat{\boldsymbol{x}}_{n-1}}(\boldsymbol{x}_n) \right]$$

FORMULA OF BAYES

Likelihood
Probability of collecting this data when our hypothesis is true

Bill Howe, UW

Prior
The probability of the hypothesis being true before collecting data

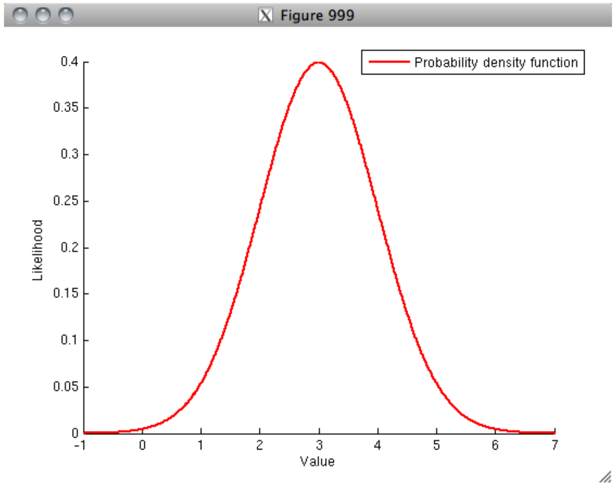
$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)}$$

Posterior
The probability of our hypothesis being true given the data collected

Marginal
What is the probability of collecting this data under all possible hypotheses?

DERIVATION OF PROBABILITY JOINT PROBABILITY OF STATE AND OBSERVATION

$$\begin{aligned} \boldsymbol{x}_{n+1} &= \boldsymbol{F}_n \boldsymbol{x}_n + \boldsymbol{\nu}_{o,n+1} \\ \boldsymbol{y}_n &= \boldsymbol{\Phi} \boldsymbol{x}_n + \boldsymbol{\nu}_{d,n}. \end{aligned}$$

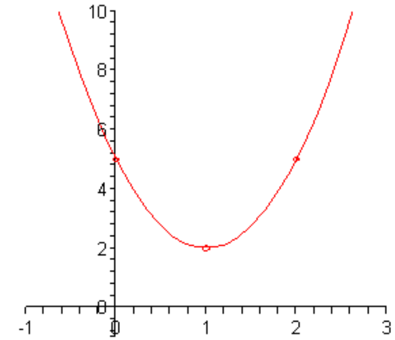


$$\arg\max_{\hat{\boldsymbol{x}}_n} p(\hat{\boldsymbol{x}}_n, \boldsymbol{y}_n) = \arg\max_{\hat{\boldsymbol{x}}_n} p(\boldsymbol{y}_n | \hat{\boldsymbol{x}}_n) p(\hat{\boldsymbol{x}}_n)$$

$$= \arg\max_{\hat{\boldsymbol{x}}_n} e^{-(\boldsymbol{y}_n - \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_n)^H \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_n)} e^{-(\hat{\boldsymbol{x}}_n - \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1})^H (\boldsymbol{F}_n^H \boldsymbol{P}_{n-1} \boldsymbol{F}_n + \boldsymbol{Q}_n)^{-1} (\hat{\boldsymbol{x}}_n - \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1})}$$

$$= \arg\min_{\hat{\boldsymbol{x}}_n} (\boldsymbol{y}_n - \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_n)^H \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_n) + (\hat{\boldsymbol{x}}_n - \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1})^H (\boldsymbol{F}_n \boldsymbol{P}_{n-1} \boldsymbol{F}_n^H + \boldsymbol{Q}_n)^{-1} (\hat{\boldsymbol{x}}_n - \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1})$$

MINIMUM IN QUADRATIC FUNCTION IS WHERE DERIVATIVE IS ZERO



$$0 = \frac{\partial}{\partial \hat{x}_n}\left((y_n - \Phi_n \hat{x}_n)^H R^{-1}(y_n - \Phi_n \hat{x}_n) + (\hat{x}_n - F_n \hat{x}_{n-1})^H (F_n P_{n-1} F_n^H + Q_n)^{-1}(\hat{x}_n - F_n \hat{x}_{n-1})\right)$$

$$= \frac{\partial}{\partial \hat{x}_n}(\hat{x}_n^H(\Phi_n^H R_n^{-1}\Phi_n + (F_n P_{n-1} F_n^H + Q_n)^{-1})\hat{x}_n - \hat{x}_n^H(\Phi_n^H R_n^{-1} y_n + (F_n P_{n-1} F_n^H + Q_n)^{-1} F_n \hat{x}_{n-1})$$

$$-(y_n^H R_n^{-1}\Phi_n + \hat{x}_{n-1}^H F_n^H (F_n P_{n-1} F_n^H + Q_n)^{-1})\hat{x}_n)$$

$$= 2(\Phi_n^H R_n^{-1}\Phi + (F_n P_{n-1} F_n^H + Q_n)^{-1})\hat{x}_n - 2(\Phi_n^H R_n^{-1} y_n + (F_n P_{n-1} F_n^H + Q_n)^{-1} F_n \hat{x}_{n-1})$$

$$\hat{x}_{n|n-1} = \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1}$$

$$\boldsymbol{P}_{n|n-1} = \boldsymbol{F}_n \boldsymbol{P}_{n-1} \boldsymbol{F}_n^H + \boldsymbol{Q}_n$$

$$\begin{aligned}
\hat{x}_n &= \left[\boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1} \boldsymbol{\Phi}_n + \boldsymbol{P}_{n|n-1}^{-1}\right]^{-1} \left[\boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1} \boldsymbol{y}_n + \boldsymbol{P}_{n|n-1}^{-1} \hat{\boldsymbol{x}}_{n|n-1}\right] \\
&= \left[\boldsymbol{P}_{n|n-1} - \boldsymbol{P}_{n|n-1} \phi_n^H (\boldsymbol{R}_n + \boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H)^{-1} \boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1}\right] \left[\boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1} \boldsymbol{y}_n + \boldsymbol{P}_{n|n-1}^{-1} \hat{\boldsymbol{x}}_{nn|n-1}\right] \\
&= \hat{x}_{n|n-1} - \boldsymbol{K}_n \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_{n|n-1} + \left[\boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1} - \boldsymbol{P}_{n|n-1} \phi_n^H (\boldsymbol{R}_n + \boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H)^{-1} \boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1}\right] \boldsymbol{y}_n
\end{aligned}$$

$$\begin{aligned}
&= \hat{x}_{n|n-1} - \boldsymbol{K}_n \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \left[(\boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H + \boldsymbol{R}_n) \boldsymbol{R}_n^{-1} - \boldsymbol{\Phi} \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H \boldsymbol{R}_n^{-1}\right] \boldsymbol{y}_n \\
&= \hat{x}_{n|n-1} - \boldsymbol{K}_n \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \boldsymbol{y}_n \\
&= \hat{x}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{y}_n - \boldsymbol{\Phi}_n \hat{\boldsymbol{x}}_{n|n-1})
\end{aligned}$$

$$\boldsymbol{K}_n := \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}^H \left[\boldsymbol{R}_n + \boldsymbol{\Phi}_n \boldsymbol{P}_{n|n-1} \boldsymbol{\Phi}_n^H\right]^{-1}$$

ALL STEPS SIMPLE ALGEBRAIC COMBINING EXCEPT

$$\hat{x}_n = \left[\Phi_n^H R_n^{-1}\Phi_n + P_{n|n-1}^{-1}\right]^{-1}\left[\Phi_n^H R_n^{-1} y_n + P_{n|n-1}^{-1}\hat{x}_{n|n-1}\right]$$

$$= \left[P_{n|n-1} - P_{n|n-1}\phi_n^H(R_n + \Phi_n P_{n|n-1}\Phi_n^H)^{-1}\Phi_n P_{n|n-1}\right]\left[\Phi_n^H R_n^{-1} y_n + P_{n|n-1}^{-1}\hat{x}_{nn|n-1}\right]$$

Develop derivations as if all vectors and matrices and square invertible matrices. Then the goa is to derive a final set of equations that no longer depend on this mathematical convenience assumption.

$$\left[\phi^H R^{-1}\phi + P^{-1}\right]^{-1}$$

$$\left[\left(\phi^H R^{-1}\phi\right)\left(\left(\phi^H R^{-1}\phi\right)^{-1} + \left(P^{-1}\right)^{-1}\right)P^{-1}\right]^{-1}$$

$$\left[\phi^H R^{-1}\phi\left(\phi^{-1}R\phi^{-H} + P\right)P^{-1}\right]^{-1}$$

$$\left[\phi^H R^{-1}\phi\left(\phi^{-1}R\phi^{-H} + P\right)\phi^H\phi^{-H}P^{-1}\right]^{-1}$$

$$\left[\phi^H R^{-1}\phi\phi^{-1}R\phi^{-H}\phi^H\phi^{-H}P^{-1} + \phi^H R^{-1}\phi P\phi^H\phi^{-H}P^{-1}\right]^{-1}$$

$$\left[\phi^H R^{-1}\left(R + \phi P\phi^H\right)\phi^{-H}P^{-1}\right]^{-1}$$

$$P\phi^H\left(R + \phi P\phi^H\right)^{-1}R\phi^{-H}$$

$$P\phi^H\left(R + \phi P\phi^H\right)^{-1}R\phi^{-H} + P\phi^H\left(R + \phi P\phi^H\right)^{-1}\phi P\phi^H\phi^{-H} - P\phi^H\left(R + \phi P\phi^H\right)^{-1}\phi P\phi^H\phi^{-H}$$

$$P\phi^H\left(R + \phi P\phi^H\right)^{-1}\left(R + \phi P\phi^H\right)\phi^{-H} - P\phi^H\left(R + \phi P\phi^H\right)^{-1}\phi P\phi^H\phi^{-H}$$

$$P - P\phi^H\left(R + \phi P\phi^H\right)^{-1}\phi P$$

Using:

$$A^{-1} + B^{-1} = B^{-1}(A+B)A^{-1}$$

$$\left(A^{-1} + B^{-1}\right)^{-1} = \left(B^{-1}(A+B)A^{-1}\right)^{-1} = A(A+B)^{-1}B$$

Using:

$$A(A+B)^{-1}B =$$

$$A(A+B)^{-1}B + B(A+B)^{-1}B - B(A+B)^{-1}B$$

$$(A+B)(A+B)^{-1}B - B(A+B)^{-1}B$$

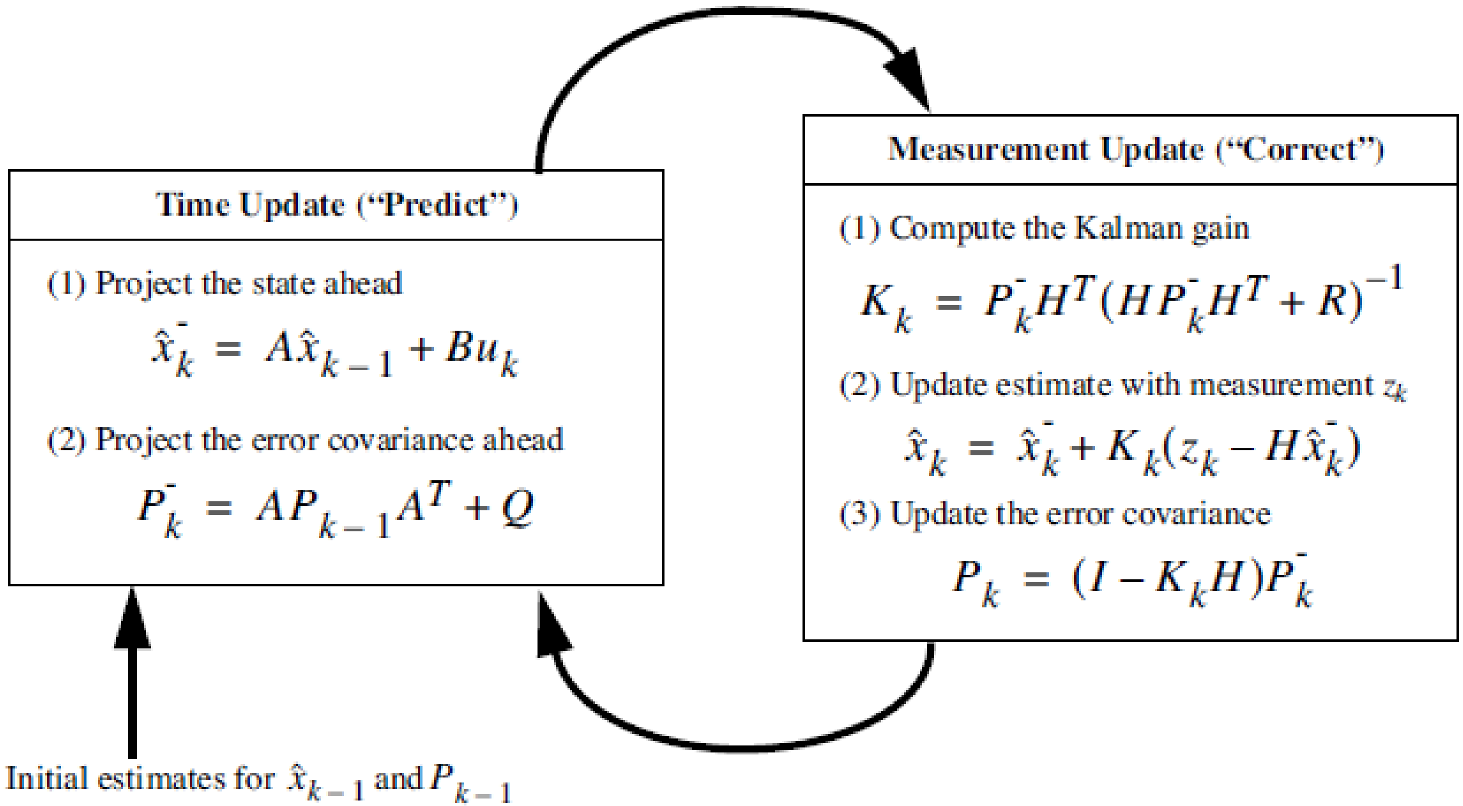
$$B - B(A+B)^{-1}B$$

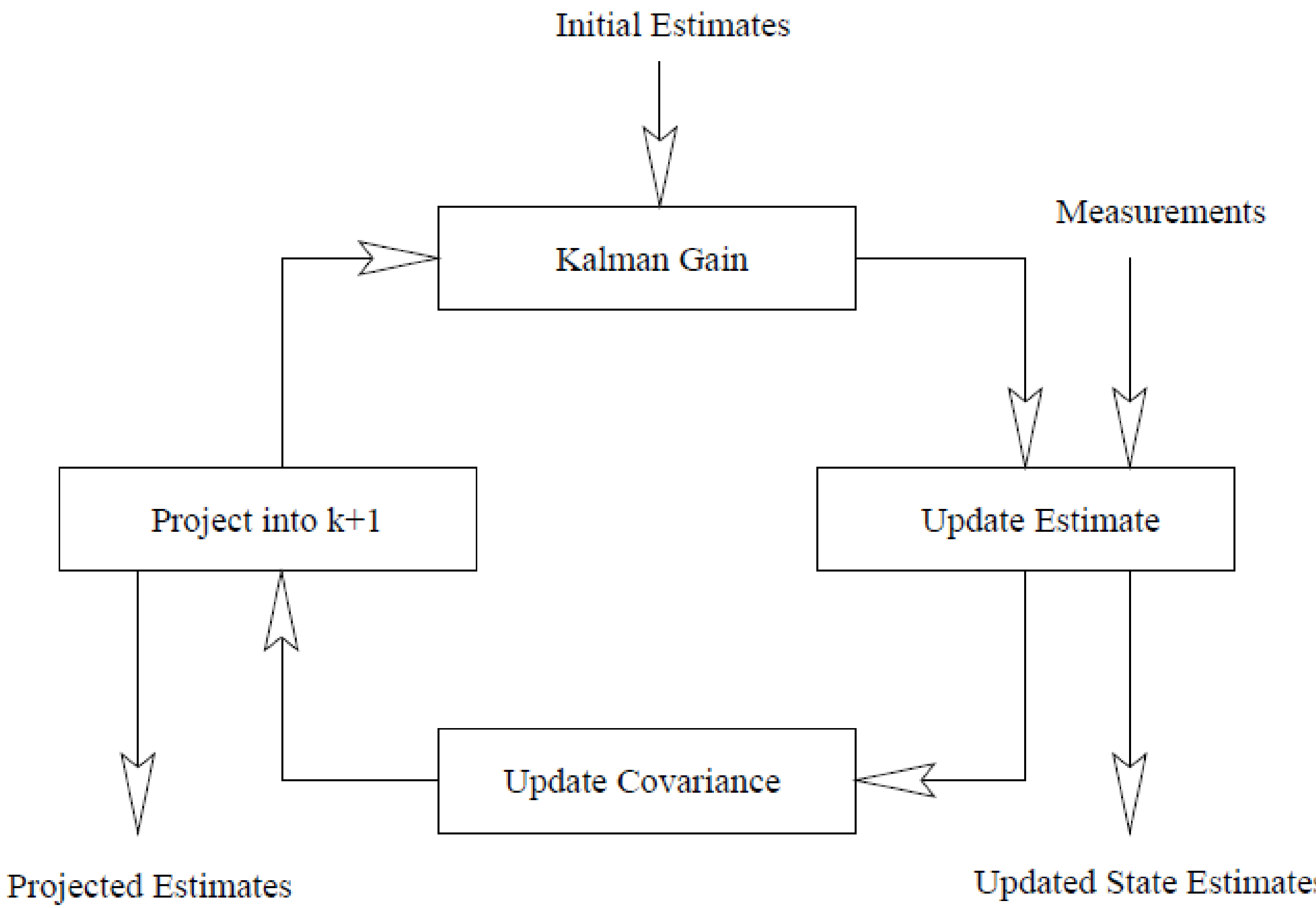
Calculation of Covariance $\boldsymbol{P}_n$ of Estimated State equals: $E[\hat{x}_{n+1}\hat{x}_{n+1}^H]$

$$
\begin{aligned}
E[\hat{\boldsymbol{x}}_n\hat{\boldsymbol{x}}_n^H] &= E[(\hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n\boldsymbol{y}_n - \boldsymbol{K}_n\boldsymbol{\Phi}_n\hat{\boldsymbol{x}}_{n|n-1})(\hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n\boldsymbol{y}_n - \boldsymbol{K}_n\boldsymbol{\Phi}_n\hat{\boldsymbol{x}}_{n|n-1})^H] \\
&= (\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{\Phi}_n)\boldsymbol{P}_{n|n-1}(\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{\Phi}_n)^H + \boldsymbol{K}_n\boldsymbol{R}_n\boldsymbol{K}_n^H \\
&= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n\boldsymbol{\Phi}\boldsymbol{P}_{n|n-1} - \boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}^H\boldsymbol{K}_n^H + \boldsymbol{K}_n(\boldsymbol{\Phi}_n\boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}_n^H + \boldsymbol{R}_n)\boldsymbol{K}_n^H \\
&= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n\boldsymbol{\Phi}_n\boldsymbol{P}_{n|n-1} - \boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}_n^H\boldsymbol{K}_n^H + \boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}_n^H\left[\boldsymbol{R}_n + \boldsymbol{\Phi}_n\boldsymbol{P}_{n|n_n}\boldsymbol{\Phi}_n^H\right]^{-1}(\boldsymbol{\Phi}_n\boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}_n^H + \boldsymbol{R}_n)\boldsymbol{K}_n^H \\
&= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n\boldsymbol{\Phi}_n\boldsymbol{P}_{n|n-1} - \boldsymbol{P}_{n|n-1}\boldsymbol{\Phi}_n^H\boldsymbol{K}_n^H + \boldsymbol{P}_{n|n_n}\boldsymbol{\Phi}_n^H\boldsymbol{K}_n^H \\
&= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n\boldsymbol{\Phi}\boldsymbol{P}_{n|n-1}
\end{aligned}
$$

KALMAN FILTER EQUATIONS

$$z_k = Hx_k$$





Description	Equation
Kalman Gain	$K_k = P'_k H^T \left(H P'_k H^T + R \right)^{-1}$
Update Estimate	$\hat{x}_k = \hat{x}'_k + K_k (z_k - H \hat{x}'_k)$
Update Covariance	$P_k = (I - K_k H) P'_k$
Project into $k+1$	$\hat{x}'_{k+1} = \Phi \hat{x}_k$ $P_{k+1} = \Phi P_k \Phi^T + Q$

Figure 11.1: Kalman Filter Recursive Algorithm

TIME VARYING MODEL COEFFICIENTS

- Requires online algorithm → recursion

$$M1_N = M1_{N-1} + x_N$$
$$M2_N = M2_{N-1} + x_N^2$$

Example of Recursion (mean and variance)

$$\mu_N = \frac{M1_N}{N} = \frac{M1_{N-1} + x_N}{N} = \frac{\frac{N-1}{N-1} M1_{N-1} - x_N}{N} = \frac{N-1}{N}\frac{M1_{N-1}}{N-1} + \frac{x_N}{N} = \frac{(N-1)\mu_{N-1}}{N} + \frac{x_N}{N}$$
$$= \mu_{N-1} + \frac{1}{N}\left(x_n - \mu_{N-1}\right)$$

$$\sigma_N^2 = \frac{M2_N}{N} - \left(\frac{M1_N}{N}\right)^2 = \frac{M2_N}{N} - \mu_N^2 = \frac{M2_{N-1} + x_N^2}{N} - \mu_N^2 = \frac{\frac{N-1}{N-1} M2_{N-1} + x_N^2}{N} - \mu_N^2$$
$$= \frac{N-1}{N}\frac{M2_{N-1}}{N-1} + \frac{x_N^2}{N} - \mu_N^2 = \frac{N-1}{N}\frac{M2_{N-1}}{N-1} + \left\{-\frac{N-1}{N}\mu_{N-1}^2 + \frac{N-1}{N}\mu_{N-1}^2\right\} + \frac{x_N^2}{N} - \mu_N^2$$
$$= \frac{N-1}{N}\left(\frac{M2_{N-1}}{N-1} - \mu_{N-1}^2\right) + \frac{N-1}{N}\mu_{N-1}^2 + \frac{x_N^2}{N} - \mu_N^2 = \frac{N-1}{N}\sigma_{N-1}^2 + \frac{N-1}{N}\mu_{N-1}^2 + \frac{x_N^2}{N} - \mu_N^2$$
$$= \frac{N-1}{N}\sigma_{N-1}^2 + \left(\frac{N\left(\mu_{N-1}^2 - \mu_N^2\right) + x_N^2 - \mu_{N-1}^2}{N}\right)$$
$$= \sigma_{N-1}^2 + \frac{N\left(\mu_{N-1}^2 - \mu_N^2\right) + x_N^2 - \mu_{N-1}^2 - \sigma_{N-1}^2}{N}$$

If mean is zero the equation is clearly correct. If mean is non-zero but not changing, then the variance is updated with a simple equation.

RECURSIVE LEAST SQUARES WITH FORGETTING

$$\mu_N = \mu_{N-1} + \frac{1}{N}\left(x_n - \mu_{N-1}\right)$$

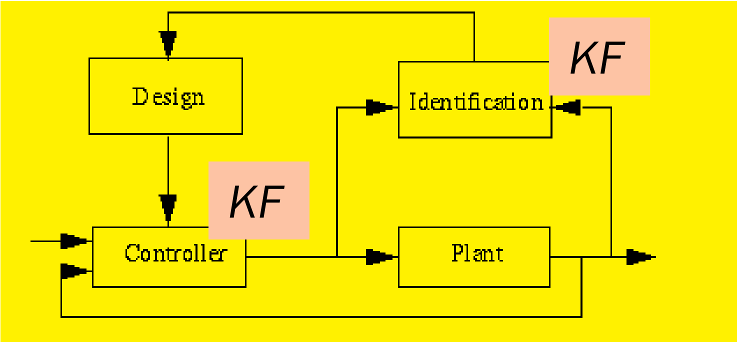
We want to forget the past because the mean may change over time!

Introduce a forgetting factor lambda!

$$\mu_N = \mu_{N-1} + \lambda\left(x_n - \mu_K\right)$$

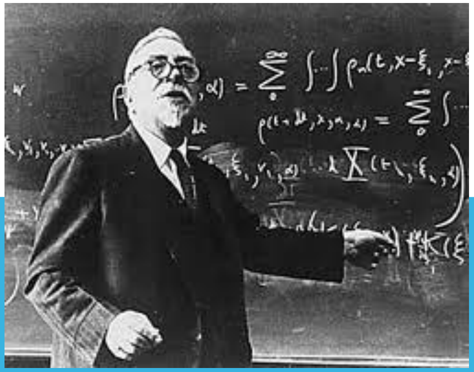
The question is: How to chose lambda?

ORIGIN 1960



- Introduction of digital computers creates practical engineering needs for:
  - Separation of random signals from noise
  - Prediction of random signals → clean estimate of derivatives
  - Estimation of system state → state can be hidden
  - Estimation of system parameters → from input and output signals
  - Estimation of time varying system parameters → systems change as a function of conditions as well as state (i.e. nonlinear)
- The Kalman filter offers discrete design solutions to all; developed by Rudy Kalman in response to Norbert Wiener's filter.

Estimation & Prediction for Control and System Identification



STATE SPACE EQUATION FOR MODEL IDENTIFICATION

ARMA model representation of system

State a $y_n = a_1 y_{n-1} + a_2 y_{n-2} + b_0 r_n + b_1 r_{n-1} + v_n + c_1 v_{n-1} + c_2 v_{n-2}$

State space equations for the model coefficients:

$$\theta_n = \theta_{n-1} + w_n$$
$$y_n = \theta_n^T \phi_n + v_n$$

$$\theta_n = [a_1, a_2, b_0, b_1, c_1, c_2]$$
$$\Phi_n = [y_{n-1}, y_{n-2}, r_n, r_{n-1}, v_{n-1}, v_{n-2}].$$

$X = \theta$ {ARMAcoefficients}

$A = I$

$B = 0$

$Y = Y$ {observable system output signal}

$H = \phi$ {observable past outputs and past and present inputs}

STANDARD MODEL IDENTIFICATION APPROACH

ARMA model representation

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + b_0 r_n + b_1 r_{n-1} + v_n + c_1 v_{n-1} + c_2 v_{n-2}$$

$$y_n = \theta_n^T \phi_n + v_n$$

$y_n = \phi_n^T \theta_n$, where θ_n is unknown.

$$\phi_n y_n = \phi_n \phi_n^T \theta_n$$

$$\left(\phi_n \phi_n^T\right)^{-1} \phi_n y_n = \theta_n$$

Can be converted to a recursive form:

$$\boxed{\begin{aligned} B_k &= B_{k-1} + \gamma_k P_{k-1} U_k^T \left(y_k - U_k B_{k-1}\right) \\ P_k &= P_{k-1} - \gamma_k P_{k-1} U_k^T U_k P_{k-1} \\ \gamma_k &= \frac{1}{U_k P_{k-1} U_k^T + \sigma_k^2}, \quad \sigma_k^2 = \text{variance of the } k^{th} \text{ noise element} \end{aligned}}$$

INTERPRETATION OF Q AND R

Cov(w) = Q

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

Kalman filter combines 2 representations of the noise signal:

Prediction & Observation

$$z_k = Hx_k + v_k$$

Cov(v) = R

$$P_n^- = P_{n-1} + Q$$

$$x_n = x_{n-1} + \frac{P_n^-}{P_n^- + R}\left(z_n - x_{n-1}\right)$$

$$x_n = \left(1 - \frac{P_n^-}{P_n^- + R}\right)x_{n-1} + \frac{P_n^-}{P_n^- + R}z_n$$

$$x_n = \left(\frac{R}{P_n^- + R}\right)x_{n-1} + \frac{P_n^-}{P_n^- + R}z_n$$

$$x_n = \left(\frac{R}{P_{n-1} + Q + R}\right)x_{n-1} + \frac{P_{n-1} + Q}{P_{n-1} + Q + R}z_n$$

Cases:

Large R, Small Q → Hold on to state
Small R, Large Q → Trust observation

Small R, Small Q → Trust observation,
but conservative on state changes.

Initial Condition (P_0) → Uncertain state value;
equivalent to high initial Q

EFFECT OF Q AND R ON KALMAN FILTER TIME CONSTANT OR LAG

$$z = e^{sT_s}$$

$$P_n^- = P_{n-1} + Q$$

$$x_n = x_{n-1} + \frac{P_n^-}{P_n^- + R}\left(z_n - x_{n-1}\right)$$

$$x_n = \left(1 - \frac{P_n^-}{P_n^- + R}\right) x_{n-1} + \frac{P_n^-}{P_n^- + R} z_n$$

$$x_n = \left(\frac{R}{P_n^- + R}\right) x_{n-1} + \frac{P_n^-}{P_n^- + R} z_n$$

$$x_n = \left(\frac{R}{P_{n-1} + Q + R}\right) x_{n-1} + \frac{P_{n-1} + Q}{P_{n-1} + Q + R} z_n$$

ARMA(1,0) : $\left(1 + a_1 z^{-1}\right) x_n = b_0 z_n$

Discrete time pole is at:

$$a_1^d = \frac{R}{P_{n-1} + Q + R}$$

Equivalent continuous time pole is at:

$$a_1^d = e^{-T_s a_1^c} = e^{-T_s 2\pi\omega} = e^{-\frac{T_s}{T_c}}$$

$$-\log\left(a_1^d\right) = T_s 2\pi\omega = \frac{T_s}{T_c}$$

Tc: Time Constant

: Cut off Frequency

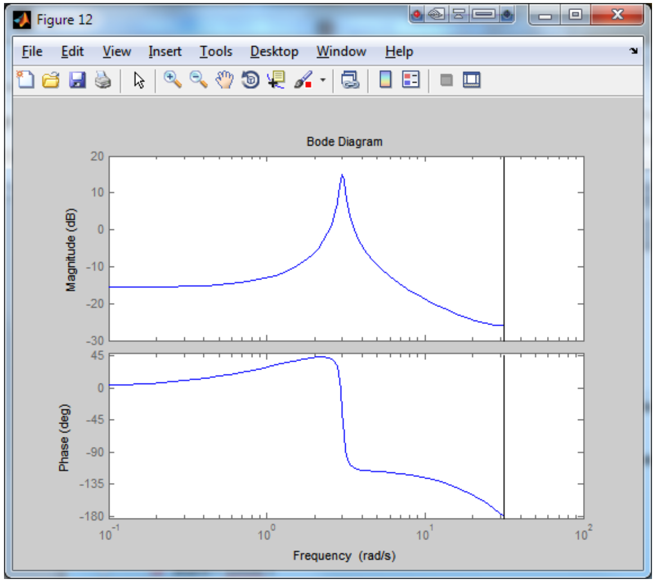
Cases:

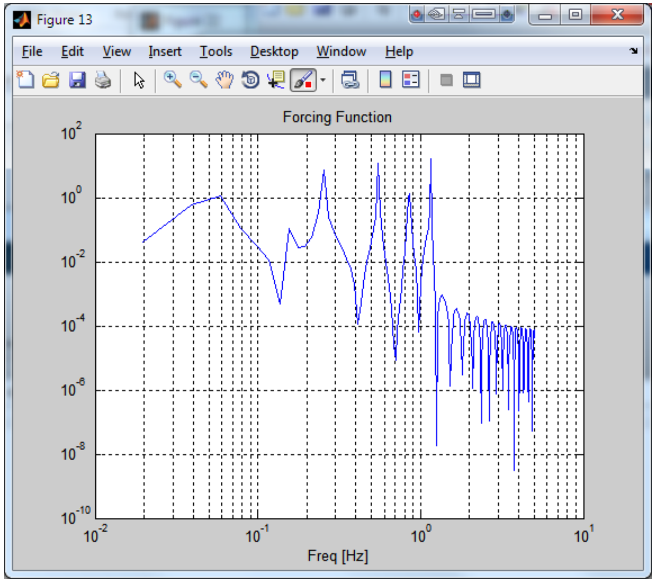
- R is large → -log$\left(a_1^d\right)$=small→ cof=low, tc=long
- R is small → -log$\left(a_1^d\right)$=large→ cof=high, tc=short
- P is large → -log$\left(a_1^d\right)$=large→ cof=high, tc=short

KALMAN FILTER AS A STATE ESTIMATOR

- 2nd order system with Known model A,B,H + Known input u → No need for Q
- Bandwidth 2nd order system approximately 7rad/s
- Forcing Function is a sum of sinusoids with a Freq Range [0.05,1.15] Hz or [0.3,7. 2] rad/s and a flat amplitude (normalized std) → excites entire system bandwidth (i.e. all poles and zeros).
- Sampling frequency set at 10Hz.
- Signals:
  - Green= True Y
  - Red = Noisy Y
  - Black = Estim Y
  - Yellow = input

MATLAB





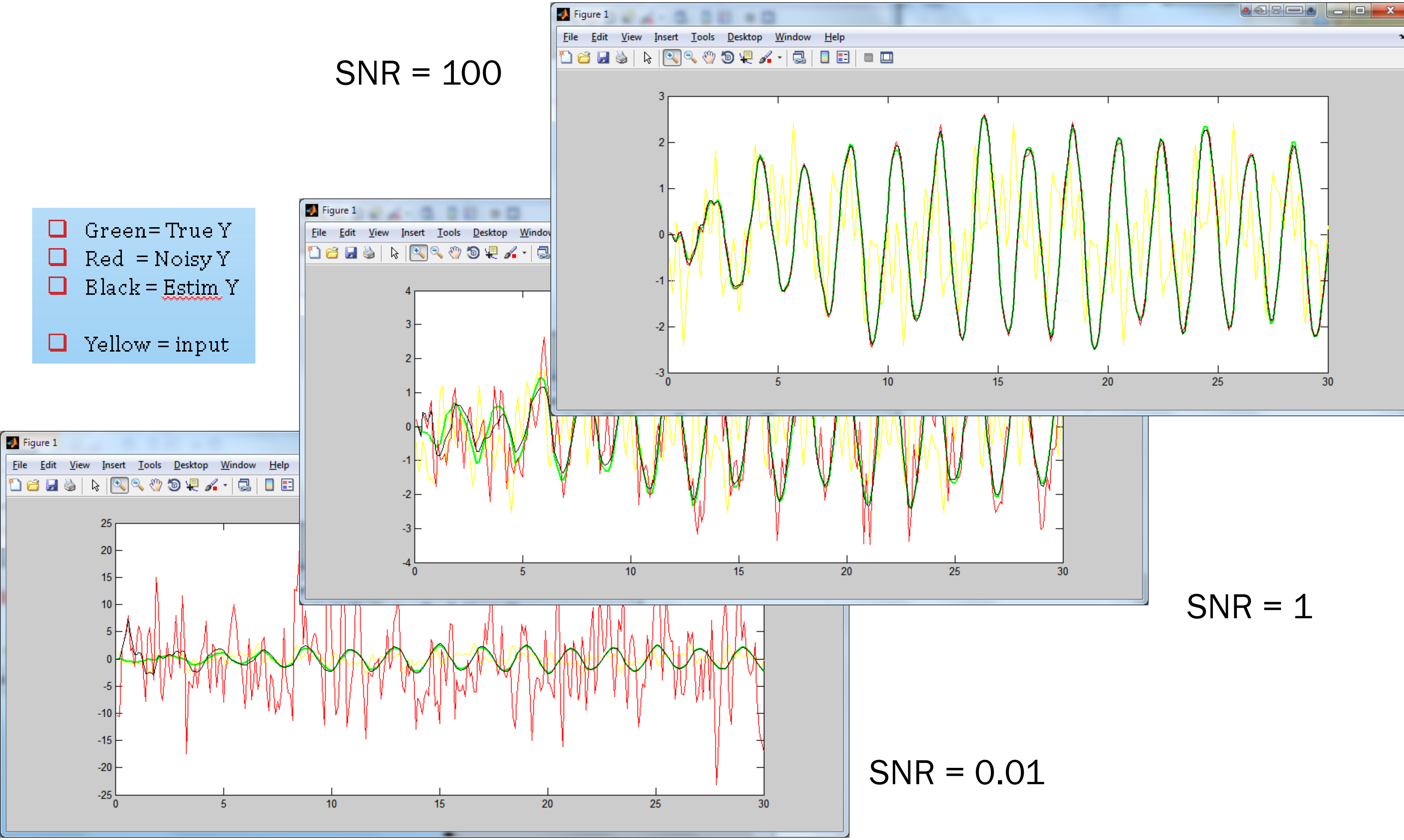
Noise is in measurement but is not fed back into system; no noise assumed to be on the input because we apply the input.

EFFECT OF OBSERVATION NOISE – KNOWN INPUT (NO NEED FOR Q=0.01)

SNR = 100

- Green= True Y
- Red = Noisy Y
- Black = Estim Y

- Yellow = input



SNR = 1

SNR = 0.01

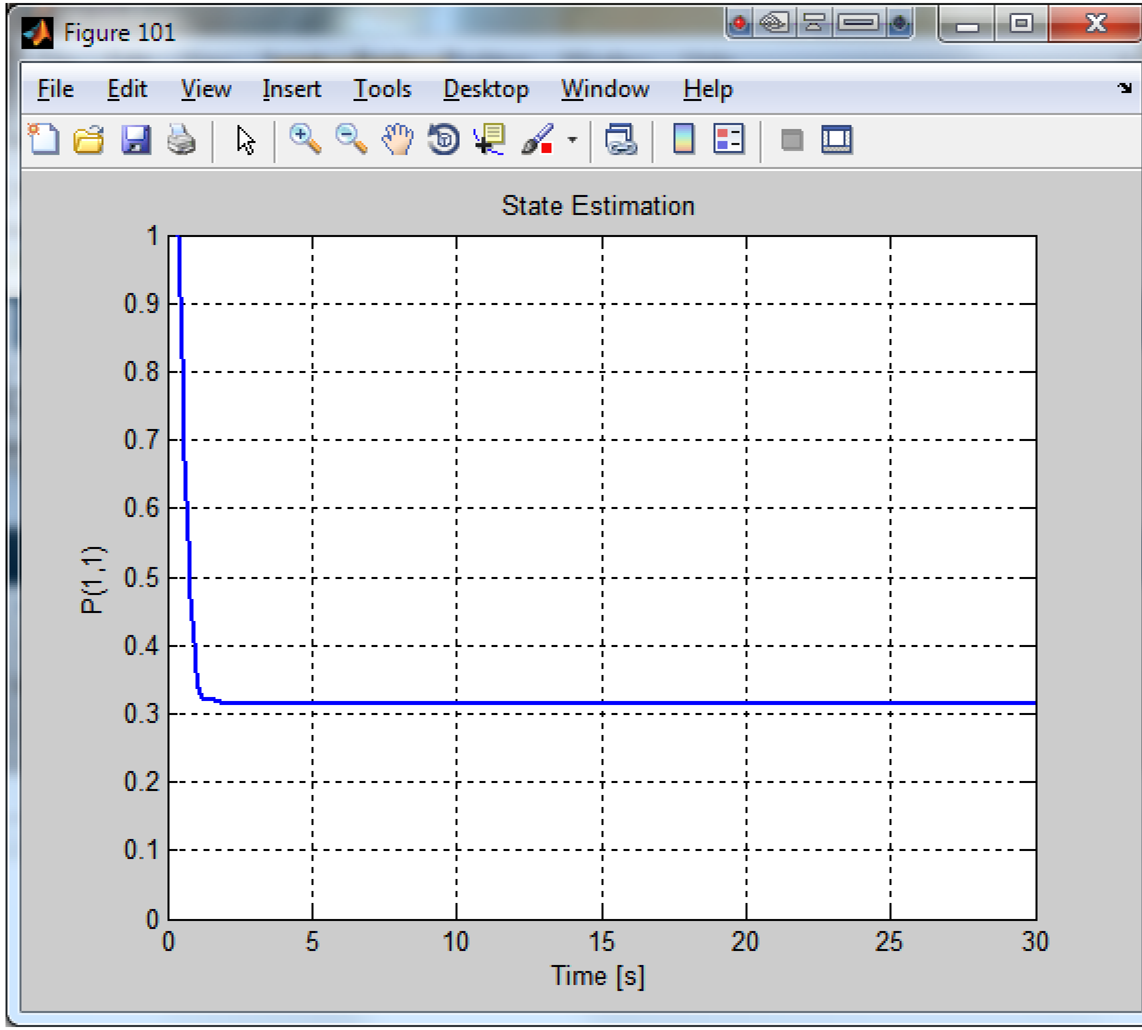
$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

CHANGE IN ERROR COVARIANCE MATRIX

P converges to a steady state in dependent of input.

When a change has been detected, P should be reset to a large value to assure quick re-convergence.



Kalman Gain	$K_k = P'_k H^T \left(H P'_k H^T + R \right)^{-1}$
Update Estimate	$\hat{x}_k = \hat{x}'_k + K_k (z_k - H \hat{x}'_k)$
Update Covariance	$P_k = (I - K_k H) P'_k$
Project into $k+1$	$\hat{x}'_{k+1} = \Phi \hat{x}_k$ $P_{k+1} = \Phi P_k \Phi^T + Q$

$$K \propto \frac{P}{P+R} \approx \frac{Q}{Q+R}$$

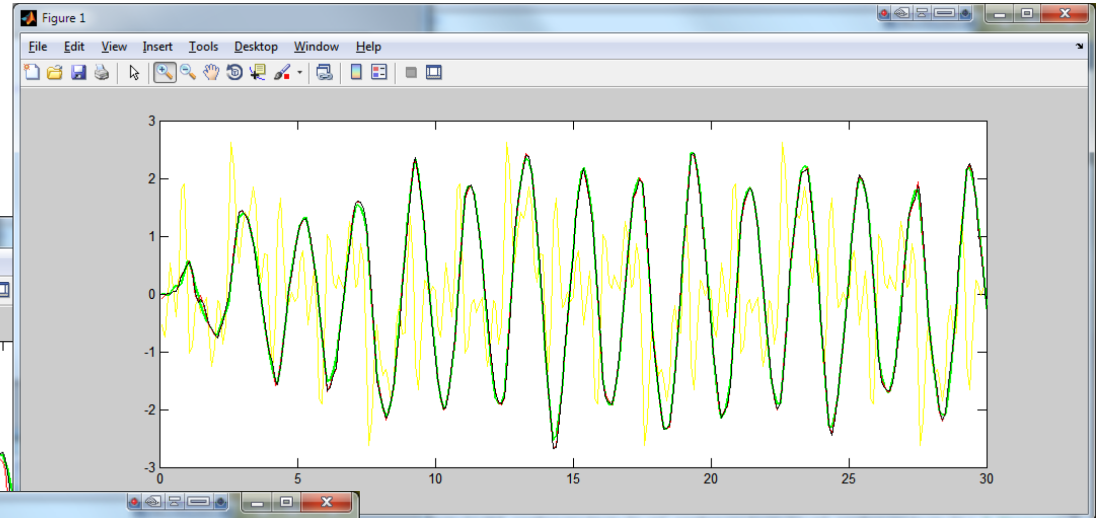
EFFECT OF PROCESS NOISE – UNKNOWN INPUT (NEED FOR Q(1,1)=B*VAR(U))

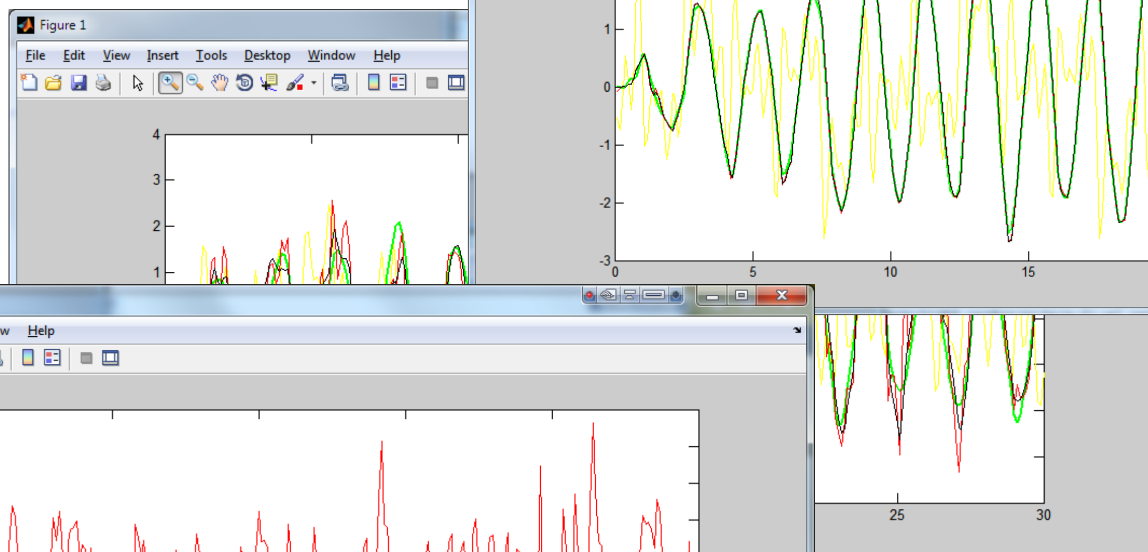
$$K \propto \frac{P}{P+R} \approx \frac{Q}{Q+R}$$

- Green = True Y
- Red = Noisy Y
- Black = Estim Y

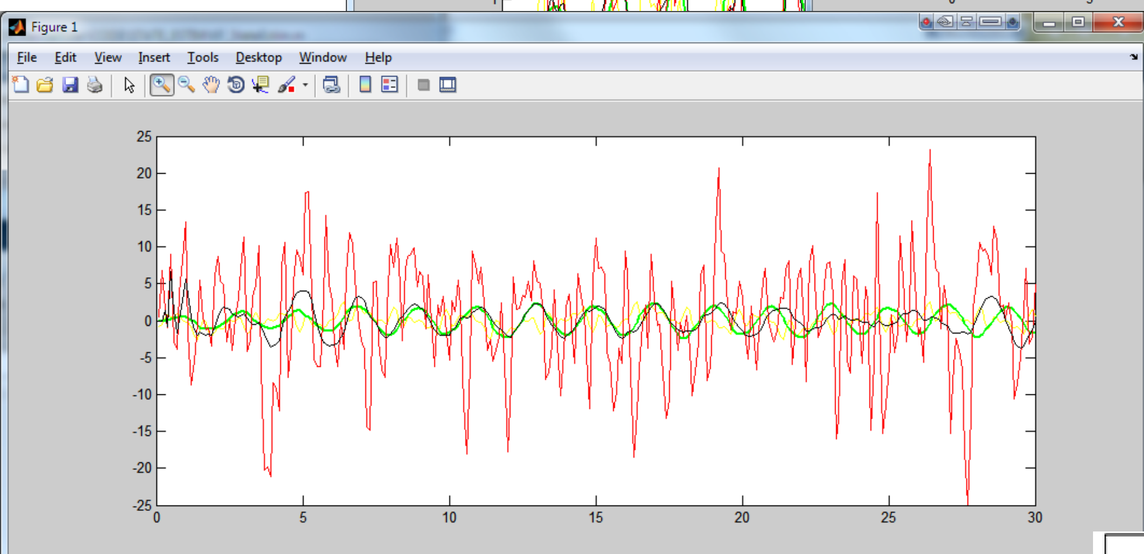
- Yellow = input

SNR = 100





SNR = 1



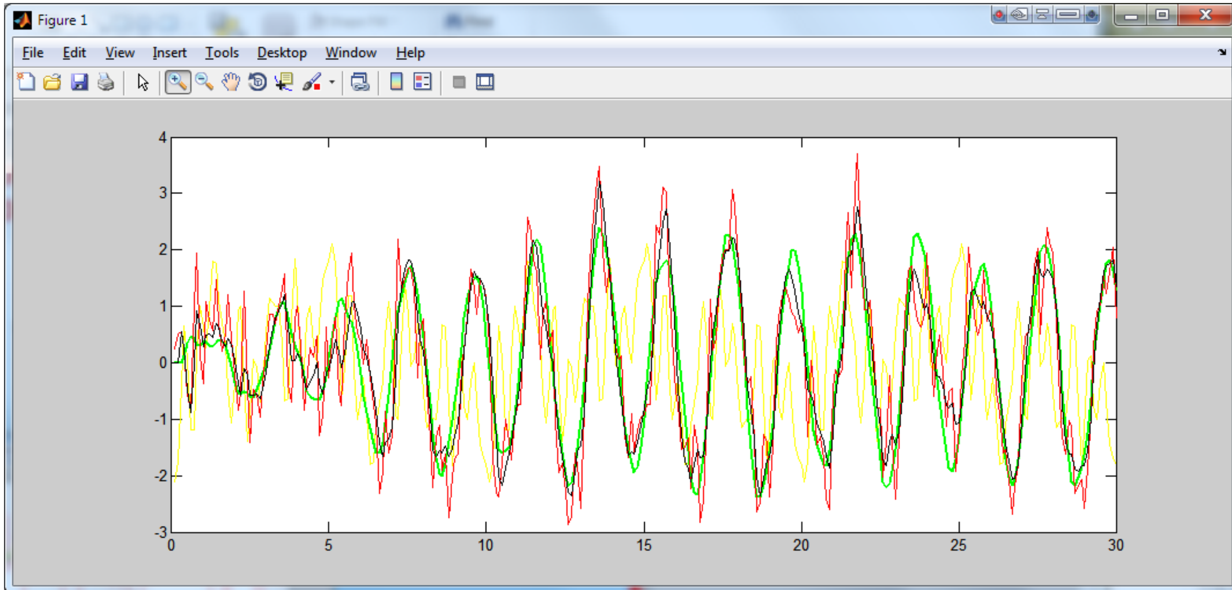
SNR = 0.01

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \qquad z_k = Hx_k + v_k$$

Kalman Gain	$K_k = P'_k H^T \left(H P'_k H^T + R\right)^{-1}$
Update Estimate	$\hat{x}_k = \hat{x}'_k + K_k (z_k - H\hat{x}'_k)$
Update Covariance	$P_k = (I - K_k H) P'_k$
Project into $k+1$	$\hat{x}'_{k+1} = \Phi \hat{x}_k$ $P_{k+1} = \Phi P_k \Phi^T + Q$

EFFECT OF SMALL Q WITH SNR = 1

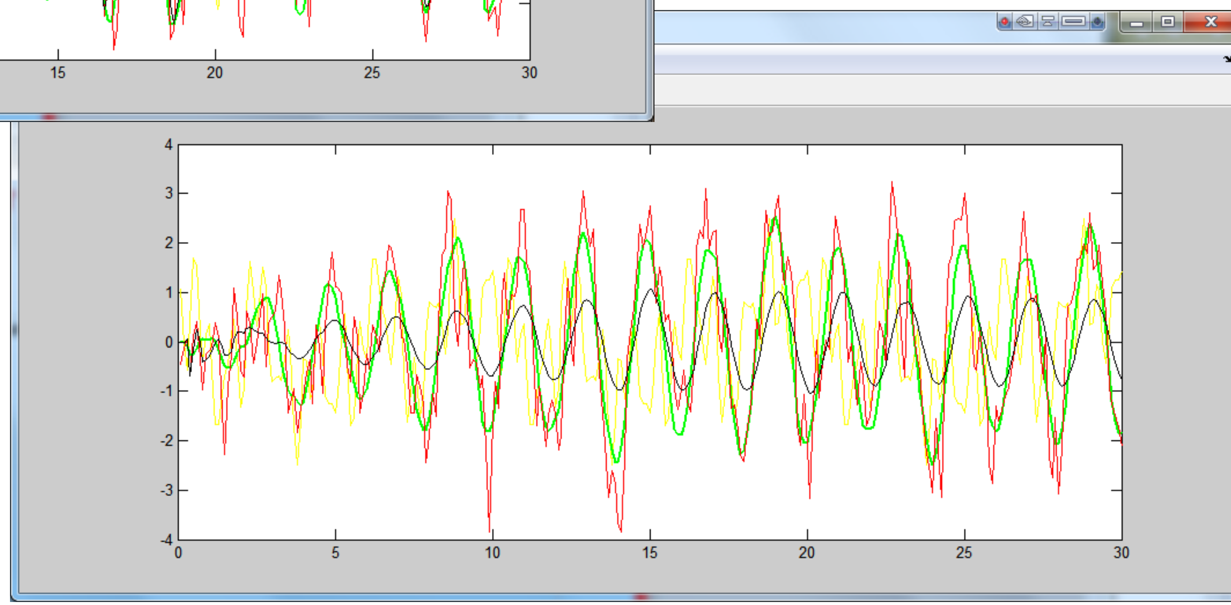
$$K \propto \frac{P}{P+R} \approx \frac{Q}{Q+R}$$



Q correct

- Green = True Y
- Red = Noisy Y
- Black = Estim Y
- Yellow = input

Q too small – Sluggish



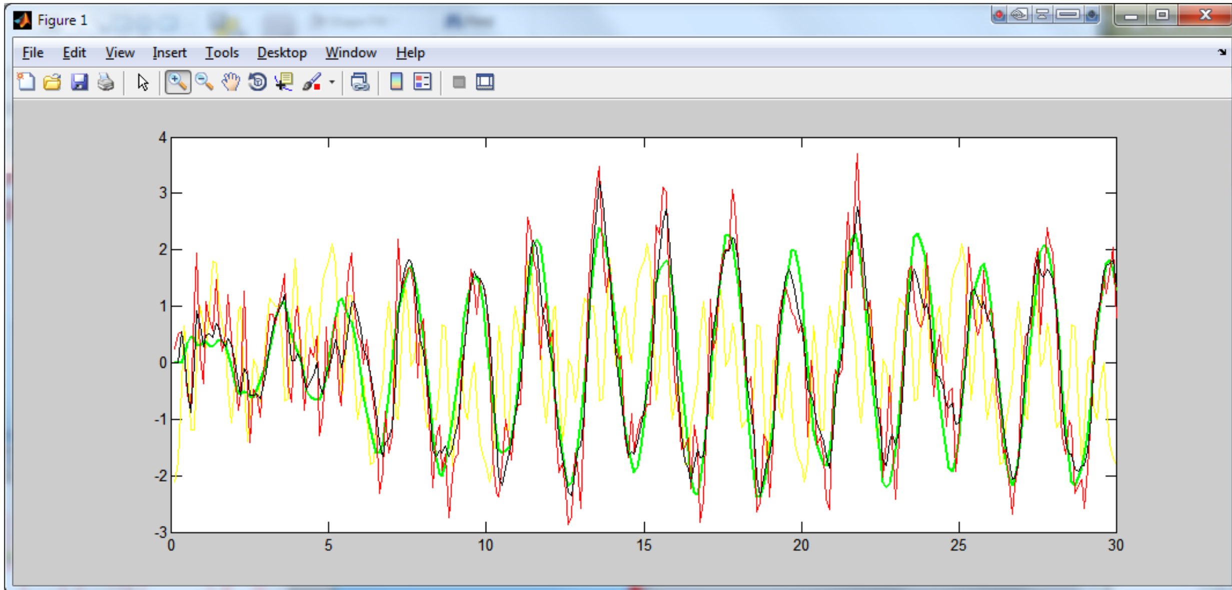
$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

Kalman Gain	$K_k = P'_k H^T \left(H P'_k H^T + R \right)^{-1}$
Update Estimate	$\hat{x}_k = \hat{x}'_k + K_k (z_k - H\hat{x}'_k)$
Update Covariance	$P_k = (I - K_k H) P'_k$
Project into $k+1$	$\hat{x}'_{k+1} = \Phi \hat{x}_k$ $P_{k+1} = \Phi P_k \Phi^T + Q$

EFFECT WITH LARGE Q WITH SNR = 1

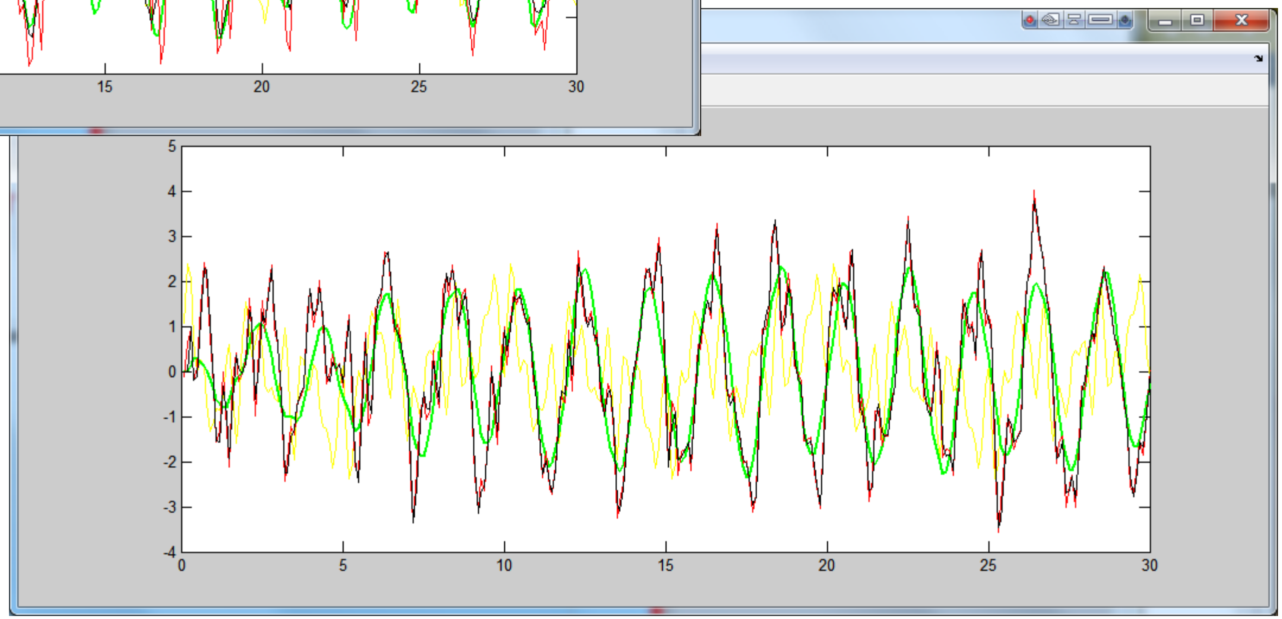
$$K \propto \frac{P}{P+R} \approx \frac{Q}{Q+R}$$



Q correct

- Green = True Y
- Red = Noisy Y
- Black = Estim Y
- Yellow = input

Q too large – Follow Noise

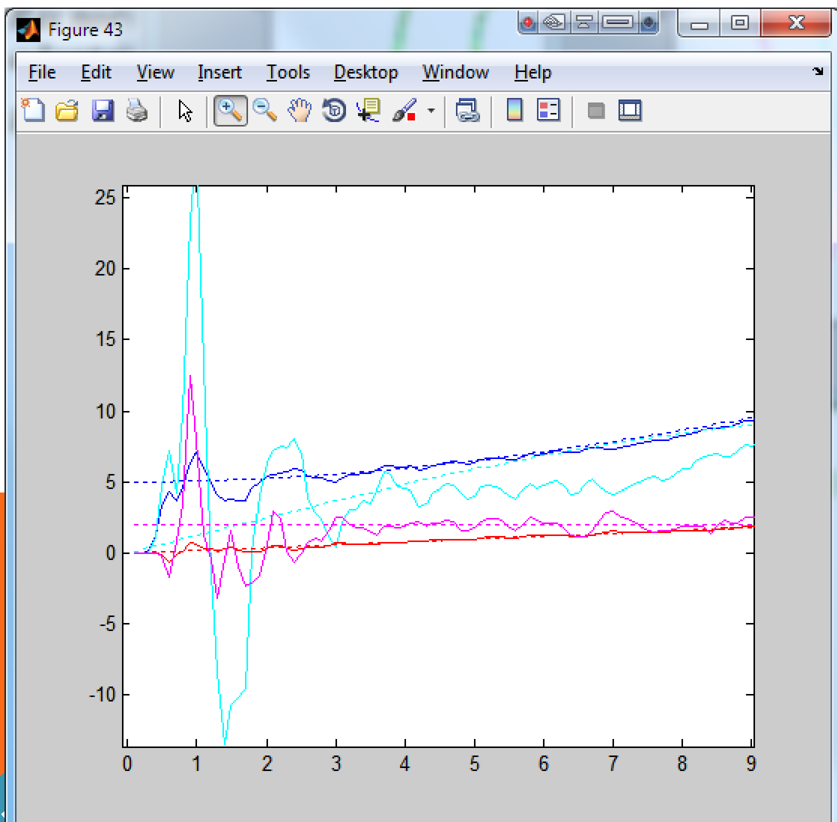


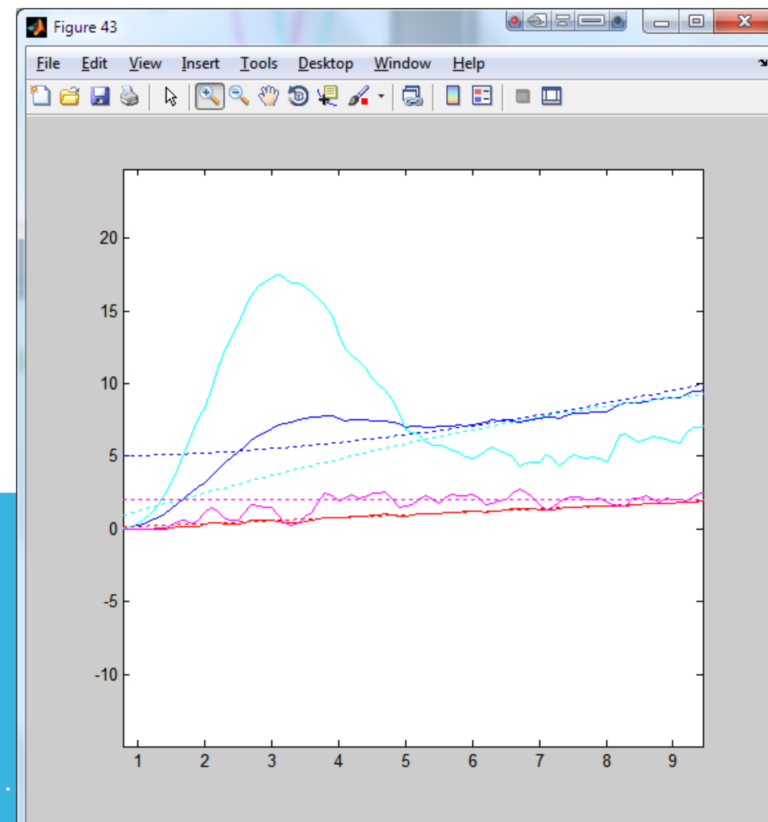
Kalman Gain	$K_k = P'_k H^T \left(H P'_k H^T + R \right)^{-1}$
Update Estimate	$\hat{x}_k = \hat{x}'_k + K_k (z_k - H \hat{x}'_k)$
Update Covariance	$P_k = (I - K_k H) P'_k$
Project into $k+1$	$\hat{x}'_{k+1} = \Phi \hat{x}_k$ $P_{k+1} = \Phi P_k \Phi^T + Q$

EFFECT OF INITIAL P

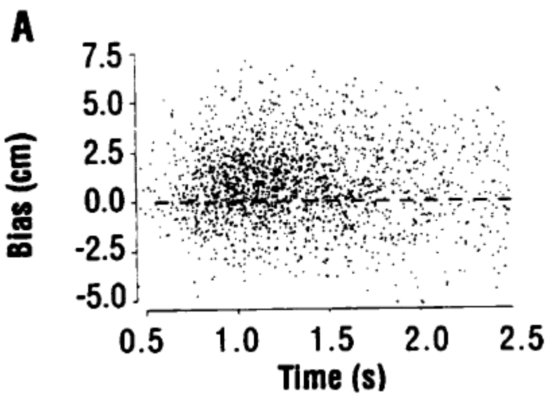
P is normally 1.0 to indicate lack of confidence in initial state guess. A quick adaptation takes place as P converges to about Q.

If P is set to nearly zero, then initial convergence is slow.





WOLPERT EXPERIMENTAL DATA



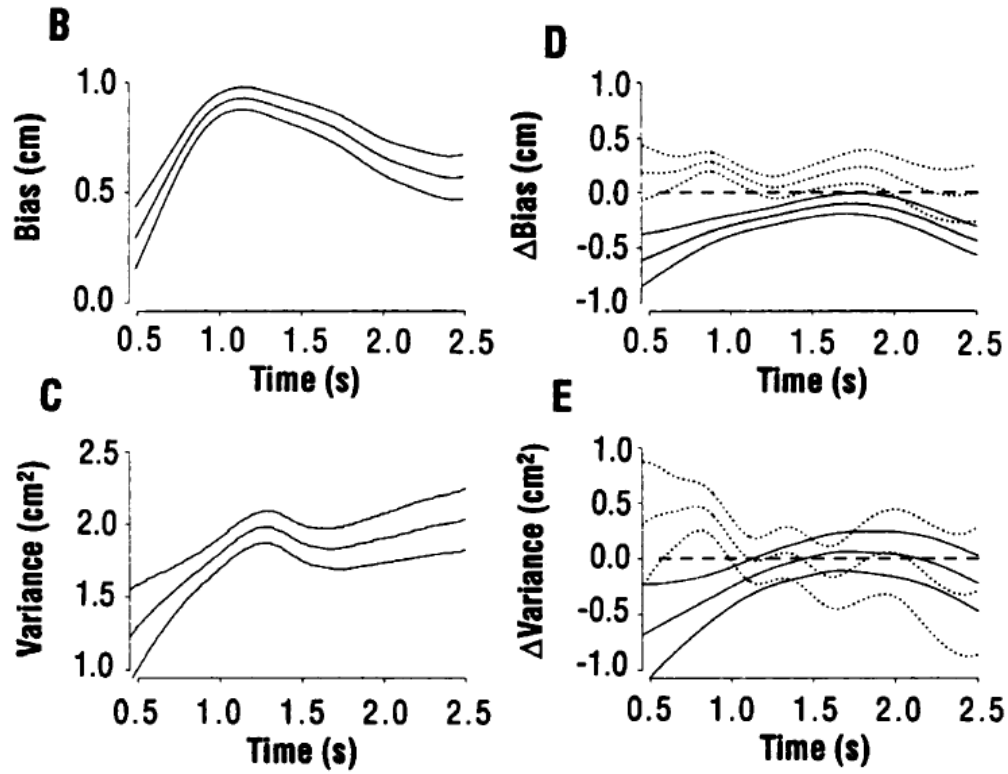
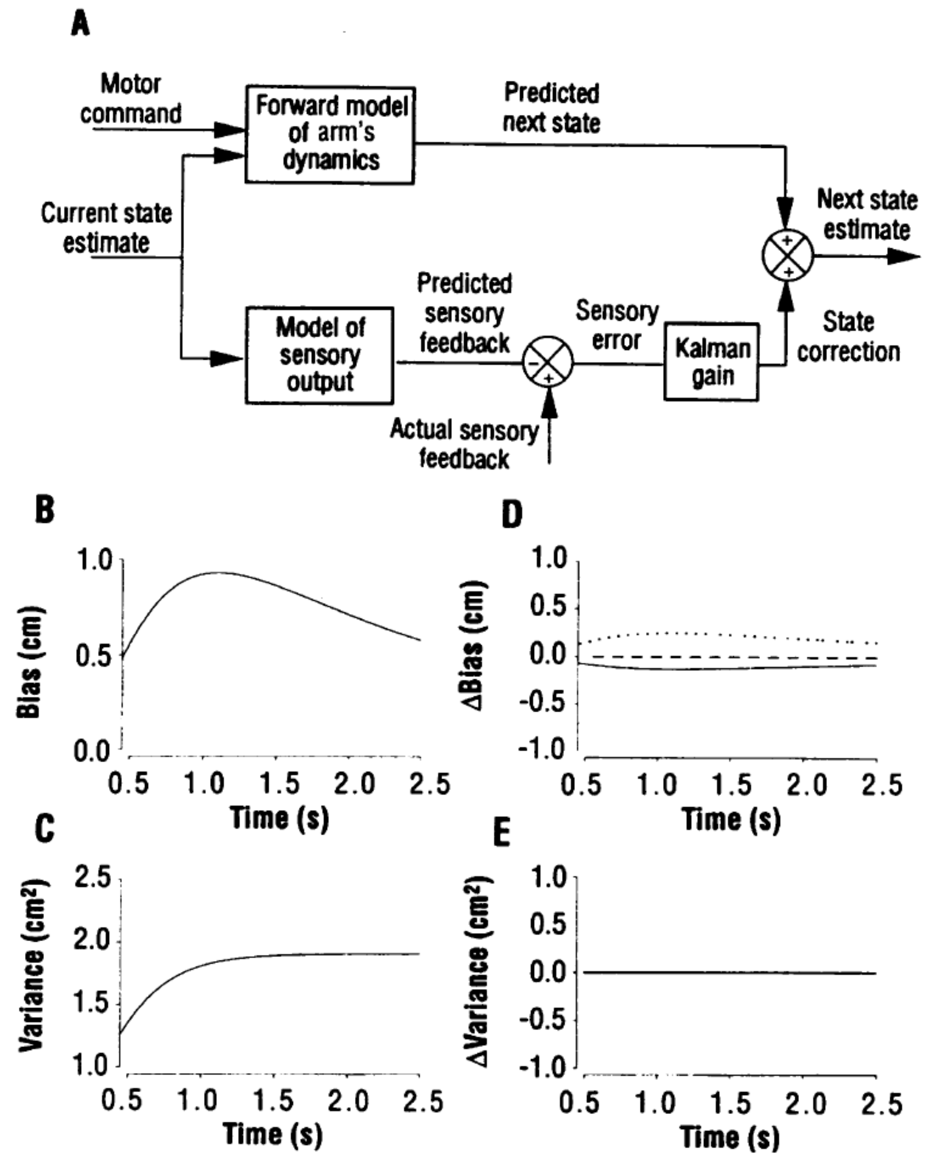


Fig. 1. The raw localization bias against movement duration is shown in (**A**) for all eight participants. There are few data points for short movement durations because of the reaction time of stopping in response to the tone. All graphs are therefore plotted from 0.5 s. (**B** through **E**) The main effect fits of the generalized additive model to the data. The propagation of the (B) bias and (C) variance of the state estimate is shown, with outer standard error lines, against movement duration. The differential effects on (D) bias and (E) variance of the external force, assistive (dotted lines) and resistive (solid lines), are also shown relative to zero (dashed line). A positive bias represents an overestimation of the distance moved. The difference in variance propagation between the resistive and assistive fields was not significant over the movement; the difference in bias was significant at the $P = 0.05$ level.

KALMAN FILTER PREDICTIONS

Fig. 2. (**A**) The Kalman filter model is shown schematically, consisting of two processes. The first (upper part) uses the motor command and the current state estimate to achieve a state estimate using the forward model to simulate the arm's dynamics. The second process (lower part) uses the difference between expected and actual sensory feedback to correct the forward model state estimate. The relative weighting of these two processes is mediated through the Kalman gain. (**B** through **E**) Simulated bias and variance propagation, in the same representation and scale as Fig. 1, B through E, from the Kalman filter model of the sensorimotor integration process.



KALMAN FILTER AS A TIME INVARIANT SYSTEM IDENTIFIER

MATLAB

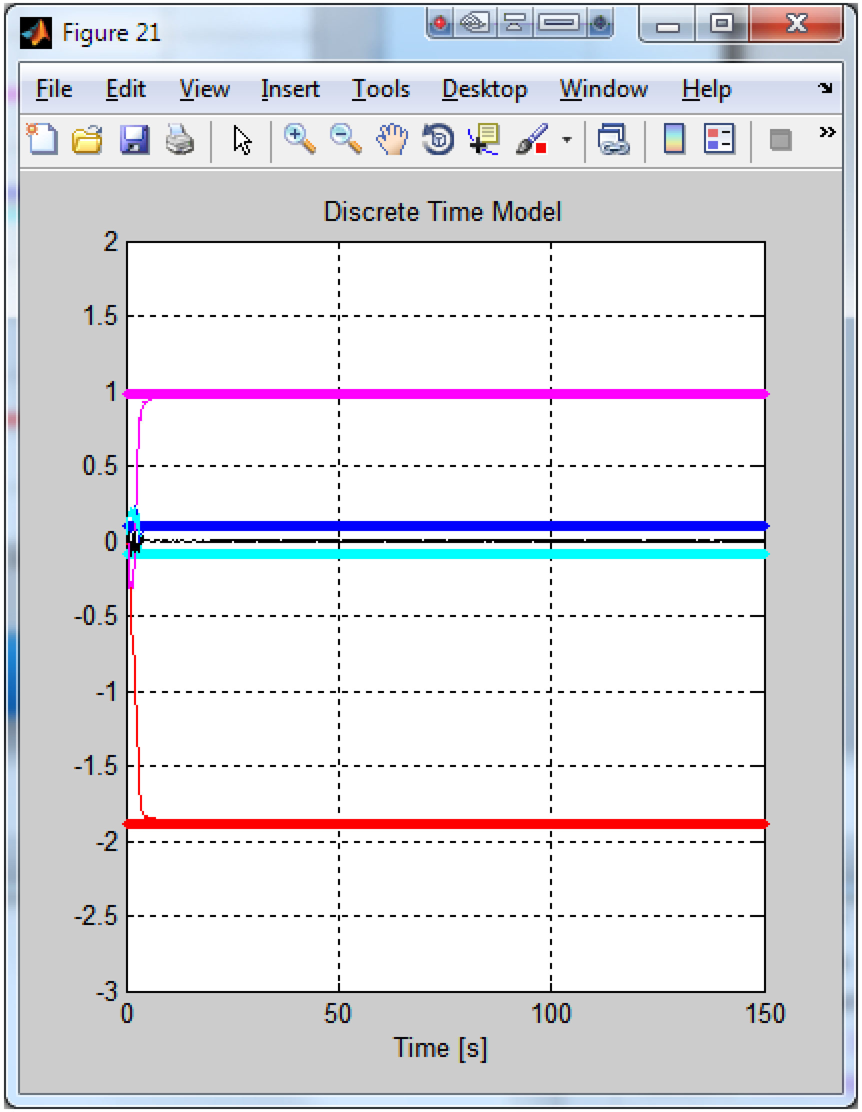
```
% Zeros
 sigma_z = 1.5;
 omega_z = 0.0;
 % Poles
 if (TimeVaryingSystem)
     sigma_p = 0.6 + 0.0*sin(2*pi*0.01*t(n));
     if (n>N/2)
         sigma_p = 0.6 - 0.5;%*sin(2*pi*0.01*t(n));
     end
 else
     sigma_p = 0.1;
 end
 omega_p = 3.0; % 0.5Hz
```

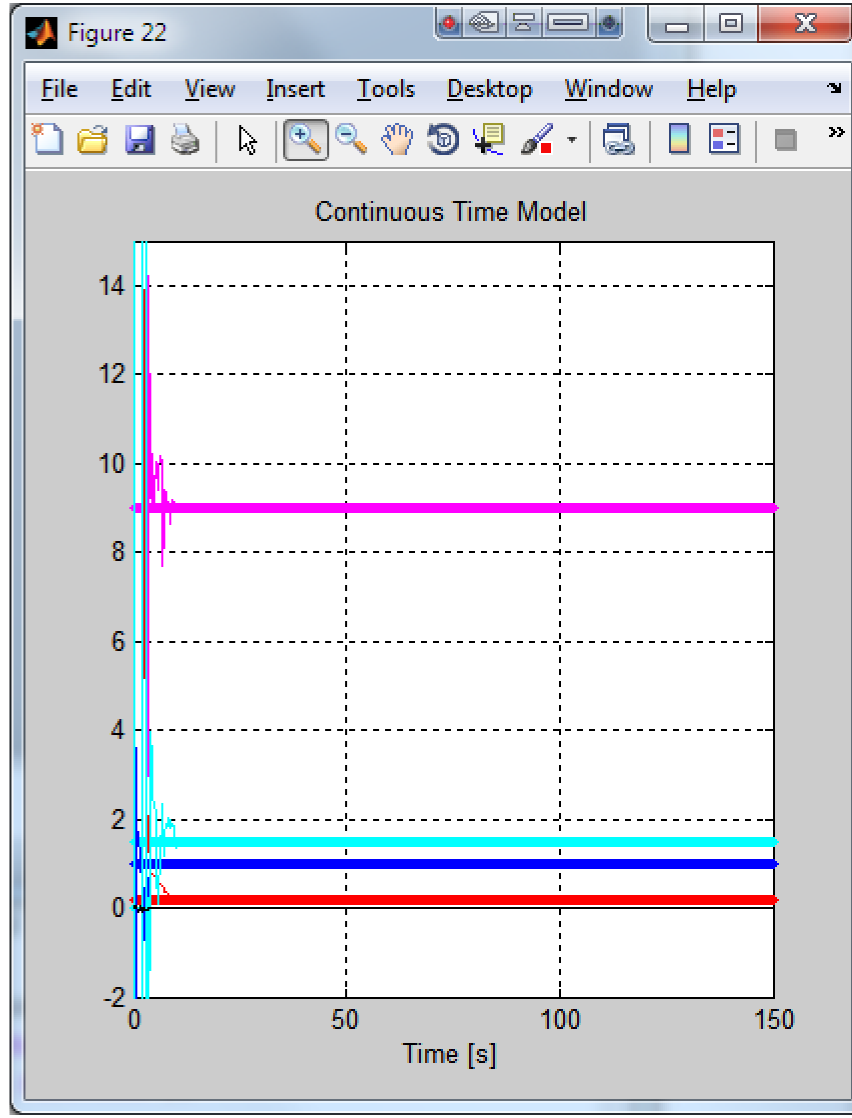
$$\frac{s+\sigma_z}{s^2+2\sigma_p s+\sigma_p^2+\omega^2}$$

```
      s + 1.5
  -----------------
s^2 + 0.2 s + 9.01
```

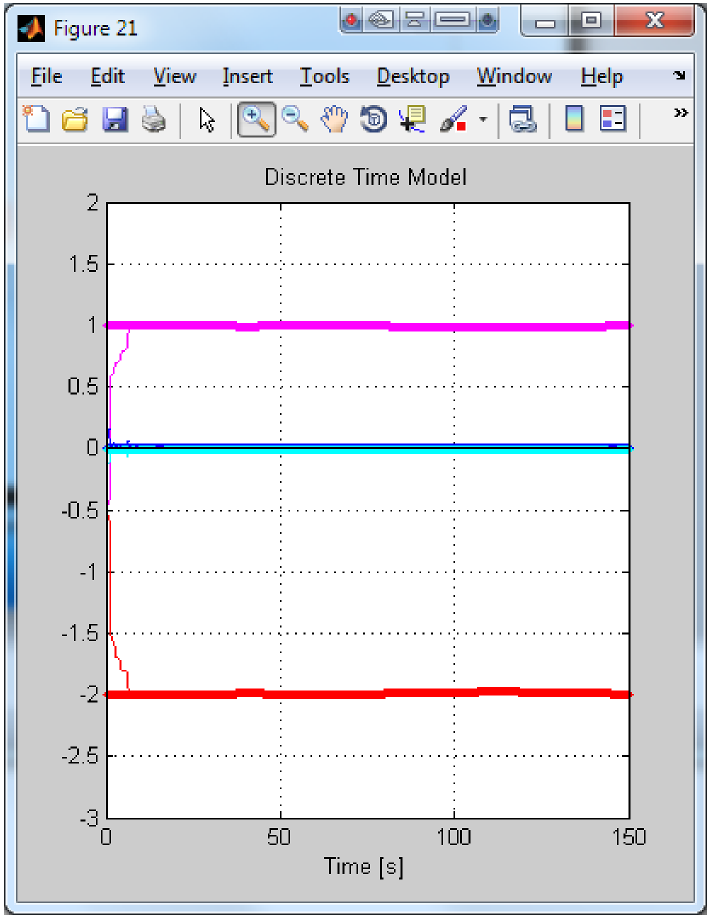
TIME INVARIANT SYSTEM: ZERO NOISE

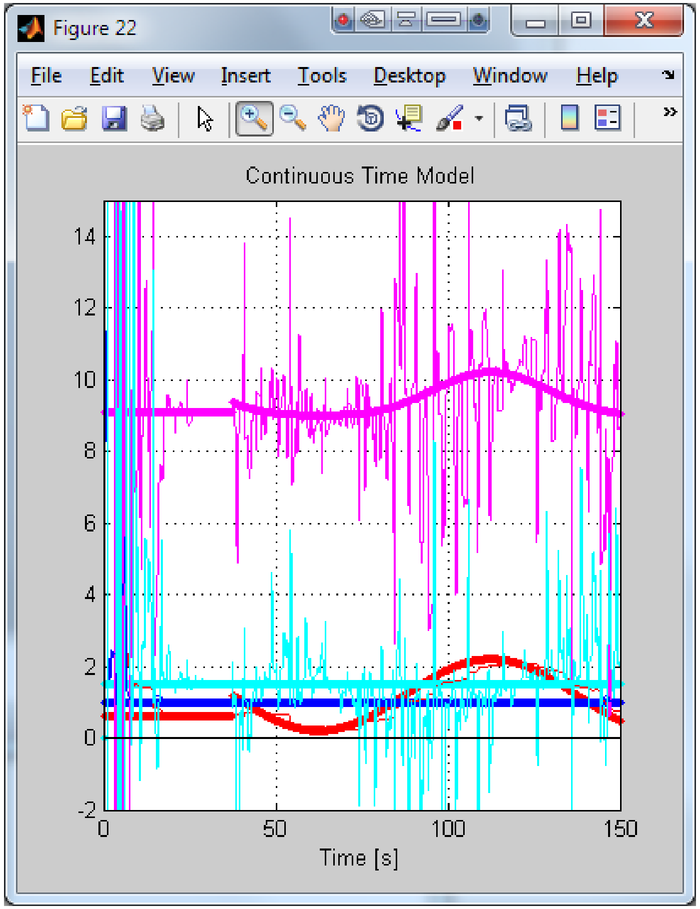
Small P_0





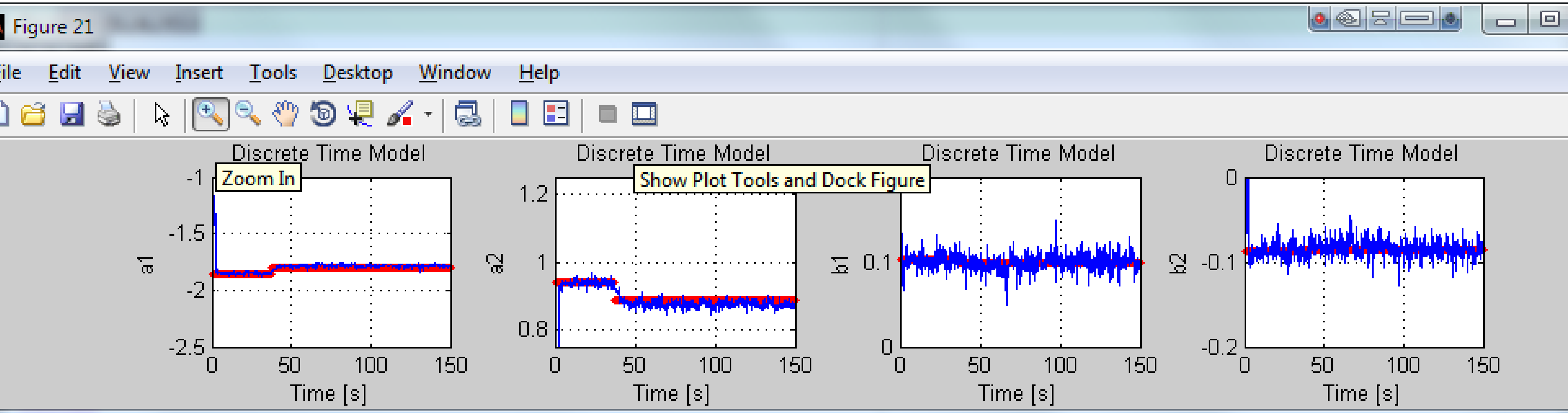
EFFECT OF SAMPLING FREQUENCY → ESTIMATING NOISE / NOT ENOUGH CHANGE PER TIME STEP FS=100HZ INSTEAD OF 10HZ

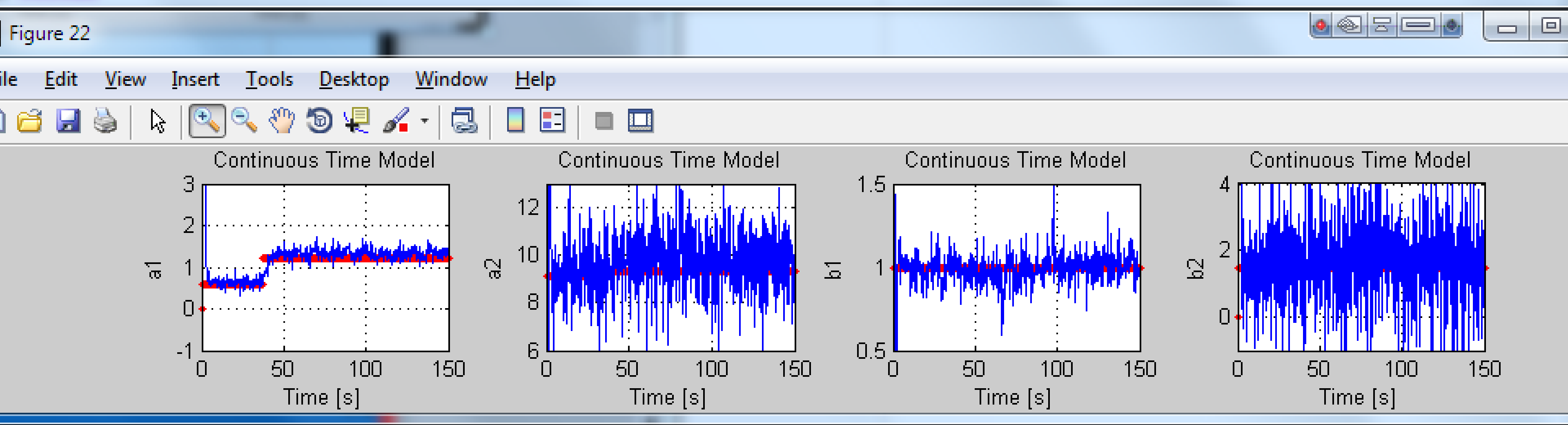




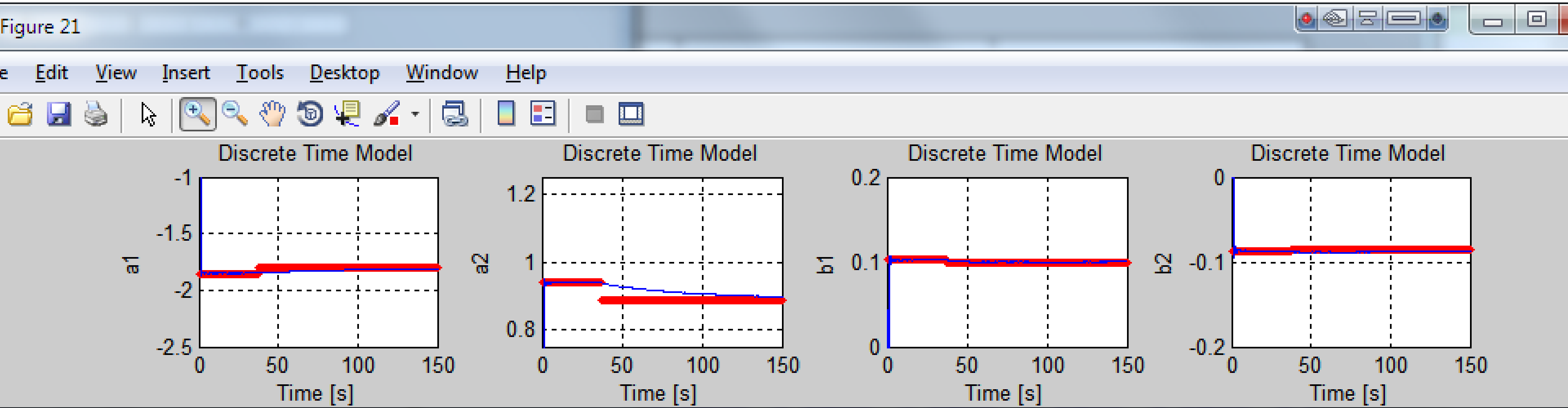
STATIONARY SYSTEM WITH SWITCH TO OTHER SYSTEM SIGMA-V=0.01, SIGMA-W=0.01

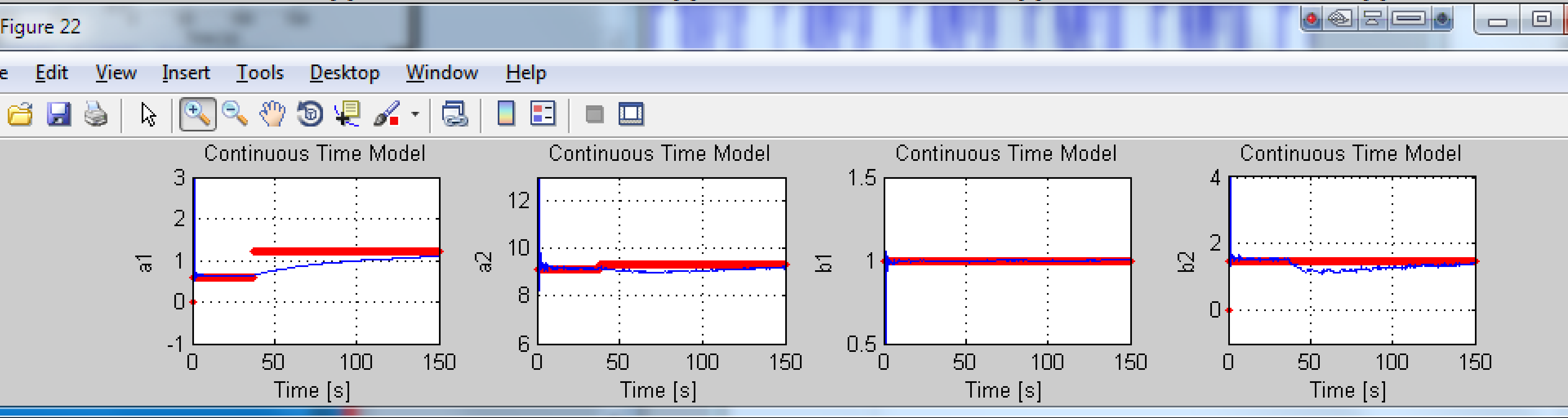
Q correct-ish





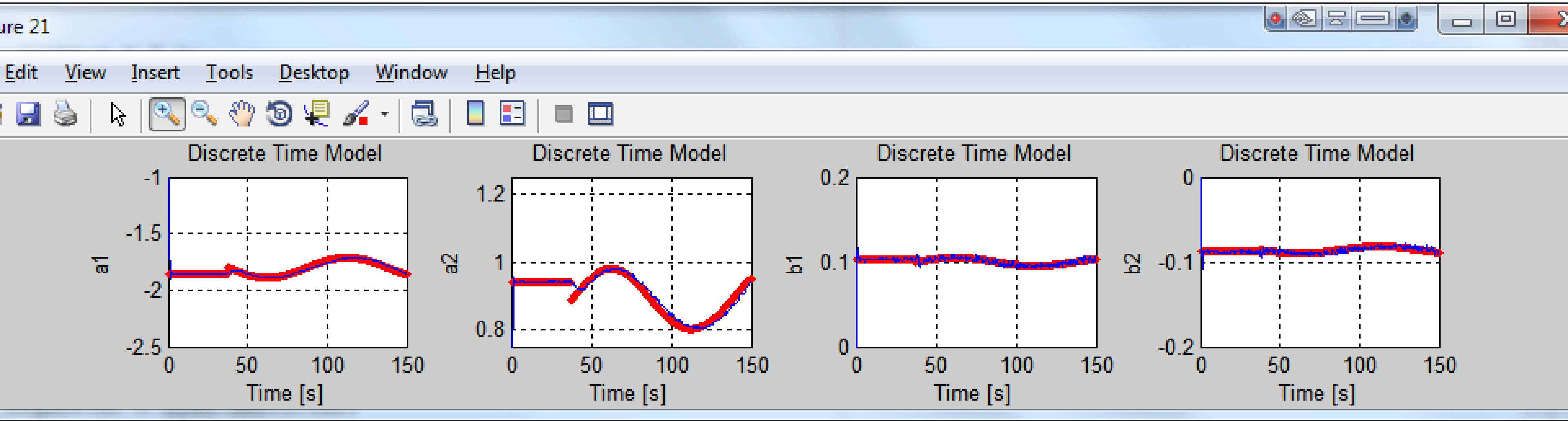
STATIONARY SYSTEM WITH SWITCH TO OTHER SYSTEM
SIGMA-V=0.01, SIGMA-W=0.0001 Q too small

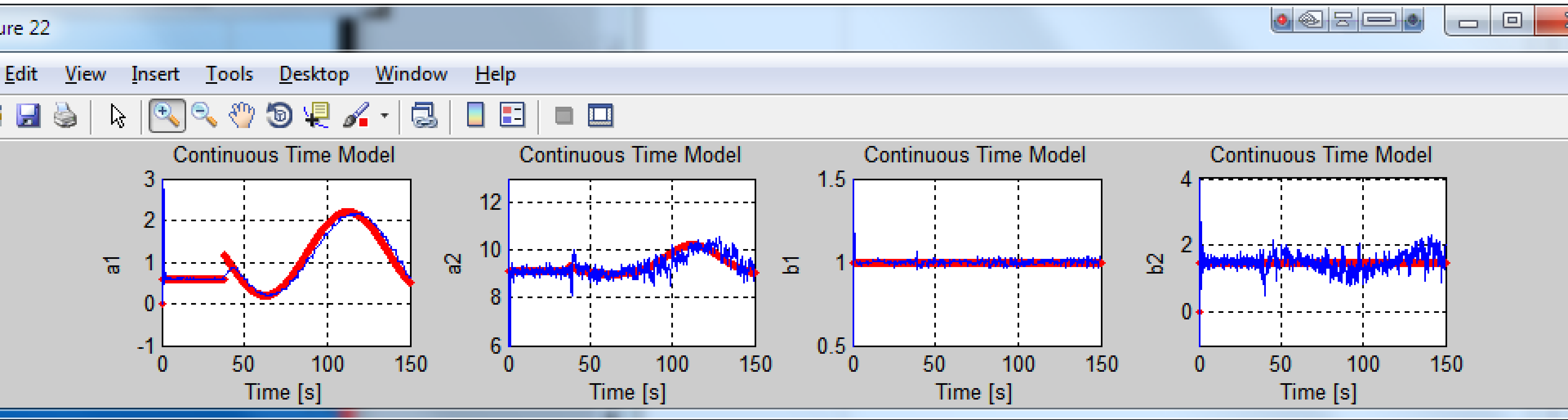




Tradeoff between accurate step following and stable steady state estimates;
this highlights one reason why models of env effect on plant model is important
(e.g. sudden load on muscles).

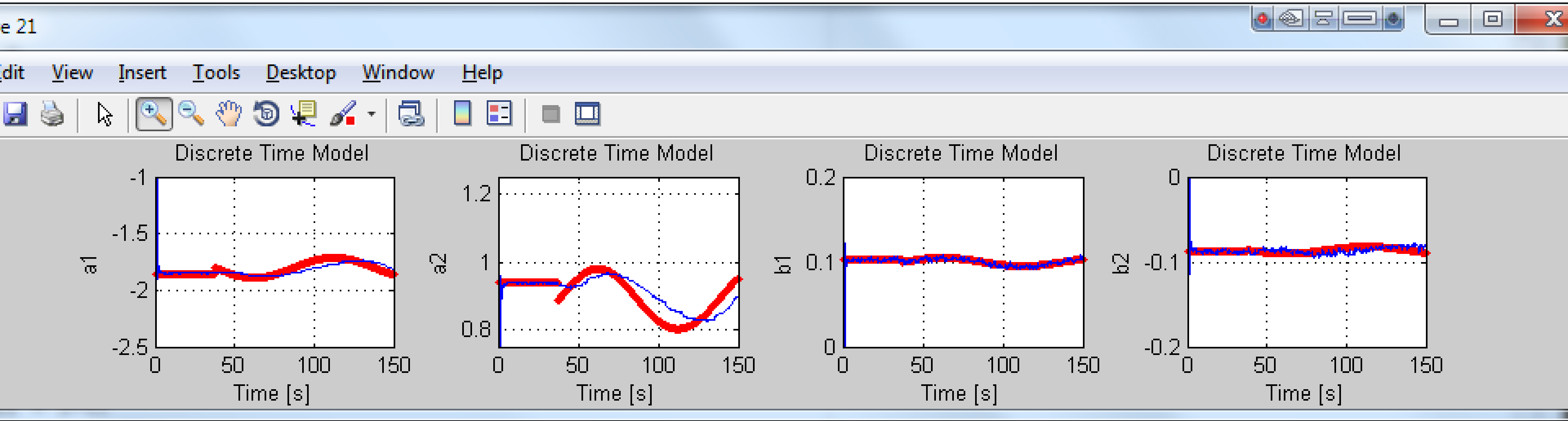
SIGMA-V = 0.001 & SIGMA-W = 0.001 Q correct-ish

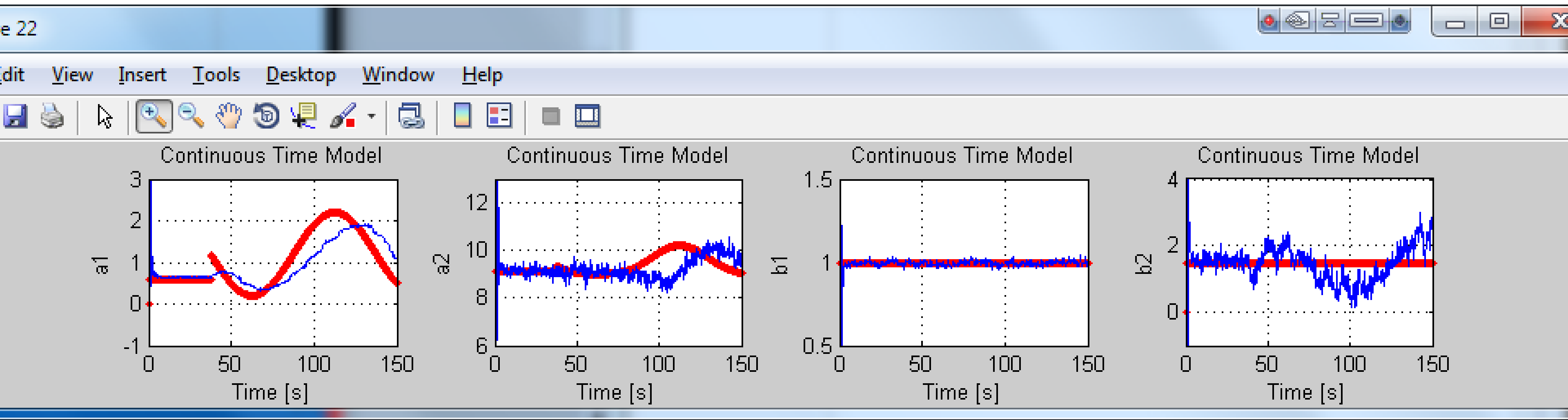




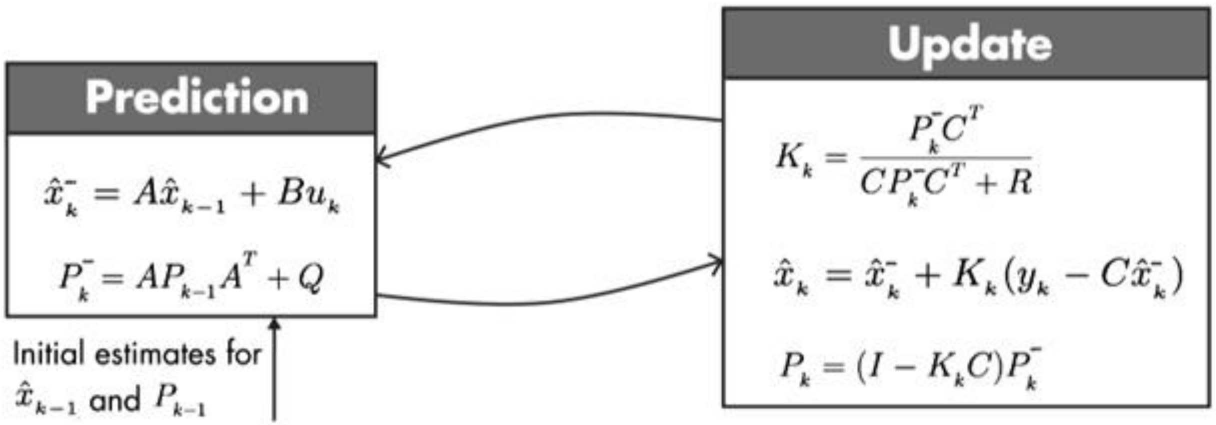
SIGMA-V = 0.01 & SIGMA-W = 0.001

Q too small

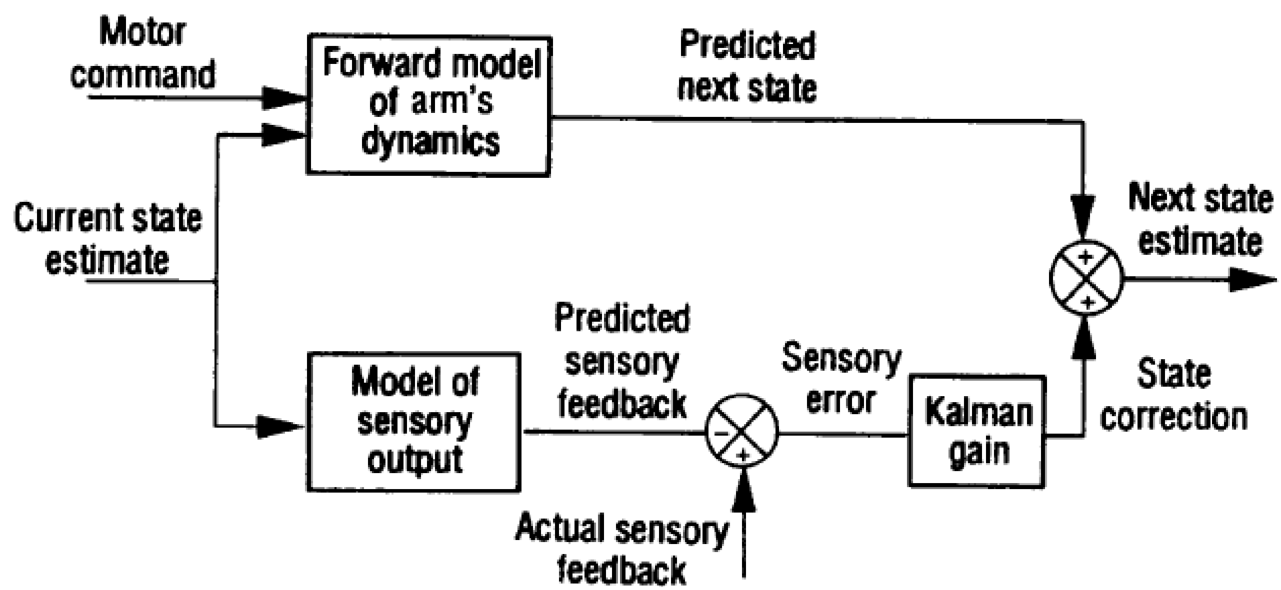




Compare against higher bandwidth forcing function.



Prediction
$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$
$P_k^- = AP_{k-1}A^T + Q$
Initial estimates for $\hat{x}_{k-1}$ and P_{k-1}
Update
$K_k = \dfrac{P_k^- C^T}{CP_k^- C^T + R}$
$\hat{x}_k = \hat{x}_k^- + K_k(y_k - C\hat{x}_k^-)$
$P_k = (I - K_k C)P_k^-$

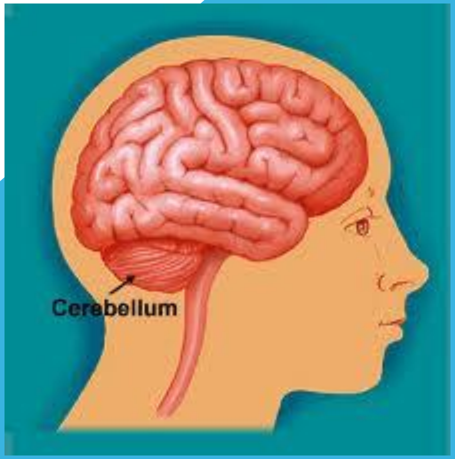


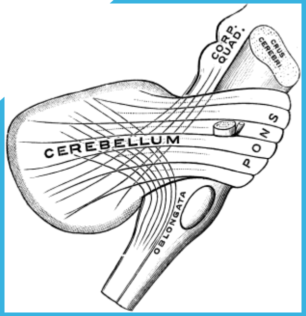
Motor command
Forward model of arm's dynamics
Predicted next state
Current state estimate
Next state estimate
Model of sensory output
Predicted sensory feedback
Sensory error
Kalman gain
State correction
Actual sensory feedback

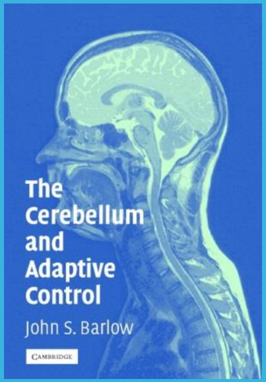
THE CEREBELLUM

SEAT OF

ADAPTIVE CONTROL AND INTERNAL MODEL DEVELOPMENT





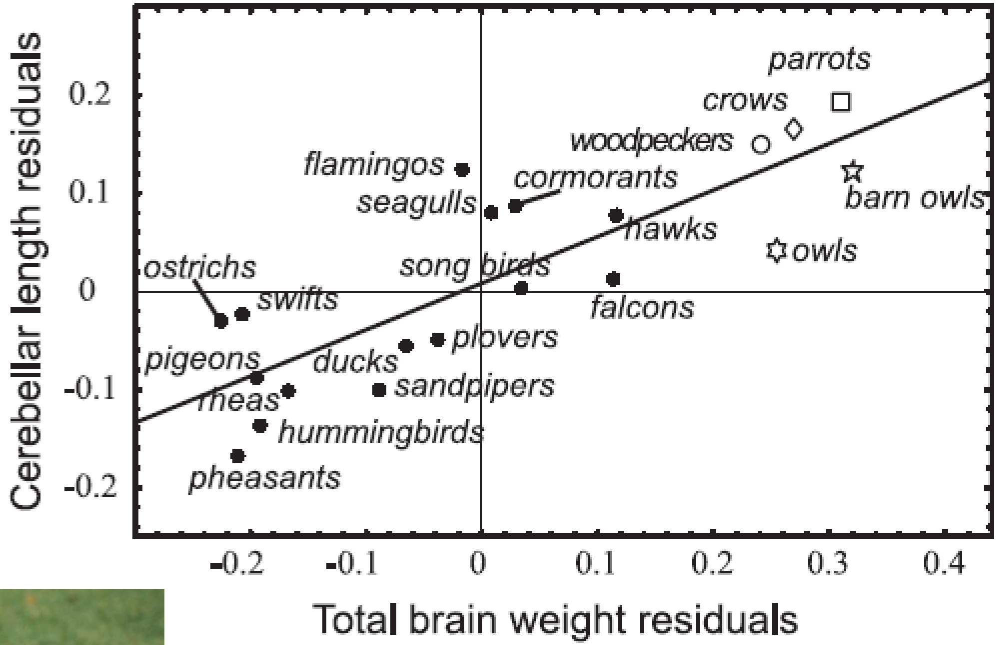


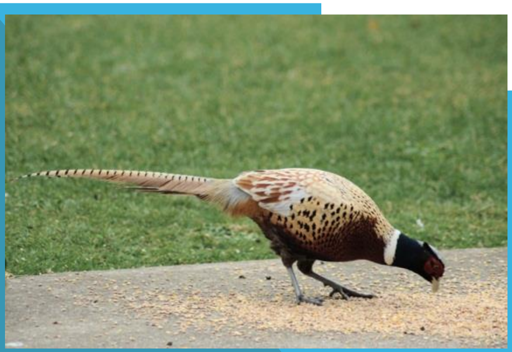
CEREBELLUM:
ALL ANIMALS HAVE IT BUT IN DIFFERENT SIZES

Hypothesis:

- Greater sensory input (esp. visual)
- Greater dexterity

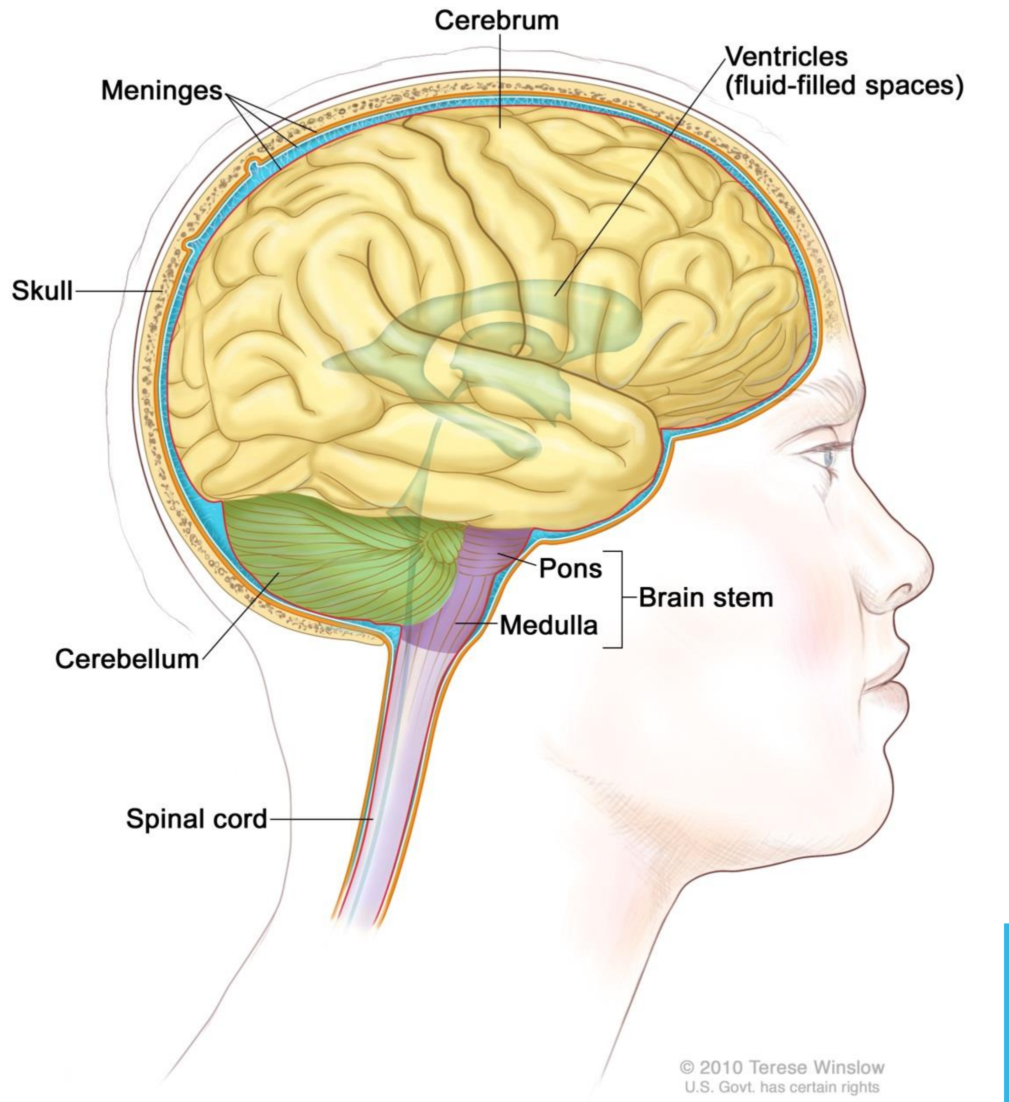
leads to a greater cerebellum because more complex coordinated movements can be made.





CEREBELLUM

- Contains
  - ~70% of all the brain's neurons
  - ~10% of the brain’s volume
- Complete removal produces:
  - No muscle weakness or loss of perception
  - Reduced coordination of movement

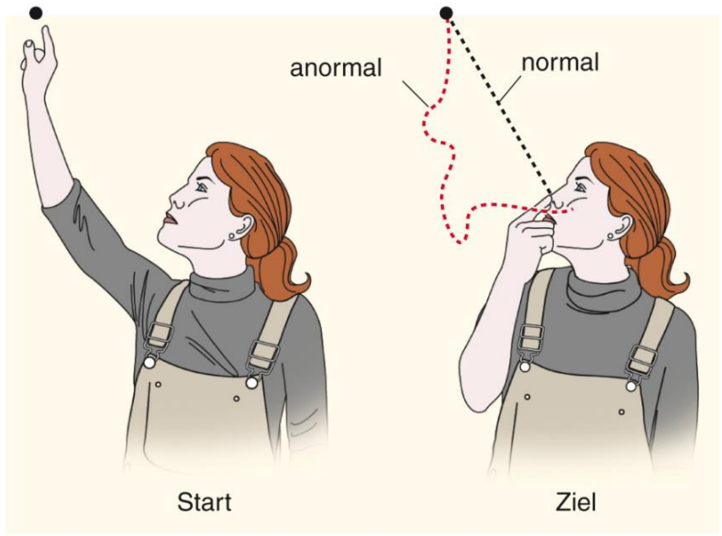


CEREBELLUM
ROLE & PURPOSE & FUNCTION

- **Motor control**
  - Equilibrium
  - Posture
  - Reaching
- **Coordination – not initiation**
  - Timing
  - Precision and accuracy
  - Reflex adjustment

- **Motor learning**
  - Learning motor skills
  - Learning to adjust sensorimotor relationships (sensorimotor calibration)
  - Sensory control of movement motor calibration

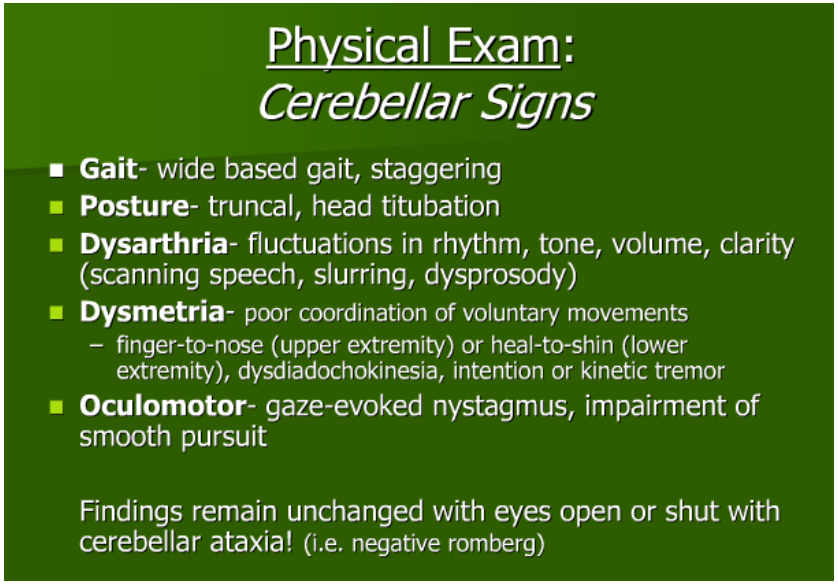
CEREBELLAR LESIONS

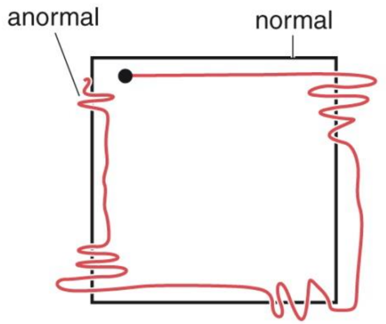


- Ataxia: inability to make
  - Smooth
  - Coordination
  - Accuracy

movements.

- Inability to learn new motor control tasks.





Aus: Bear et al., *Neurowissenschaften, 3. Aufl.*
© Spektrum Akademischer Verlag GmbH 2009

2.3 Assignments

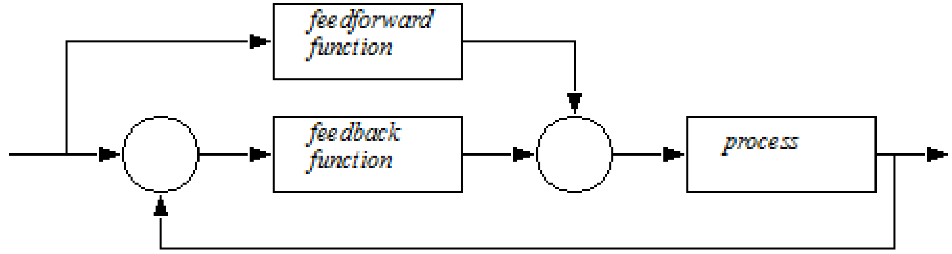
Assignment 4 Adaptation of Intrinsic & Reflexive Feedback During Postural Control of the Shoulder (8 hours, including practical and feedback lecture)

Assignment 5 Kalman filtering in state estimation during reaching (12 hours, including feedback lecture)

An Internal Model for Sensorimotor Integration

Daniel M. Wolpert,* Zoubin Ghahramani, Michael I. Jordan

On the basis of computational studies it has been proposed that the central nervous system internally simulates the dynamic behavior of the motor system in planning, control, and learning; the existence and use of such an internal model is still under debate. A sensorimotor integration task was investigated in which participants estimated the location of one of their hands at the end of movements made in the dark and under externally imposed forces. The temporal propagation of errors in this task was analyzed within the theoretical framework of optimal state estimation. These results provide direct support for the existence of an internal model.



WHAT YOU LEARNED

Kalman Filter as a Powerful Mechanism for

State Estimation

Internal Model Updating (Identification)

Kalman Filter as Possible Mechanism in

Human Control (esp. rapid control)

Human Adaptation (what to do with errors)