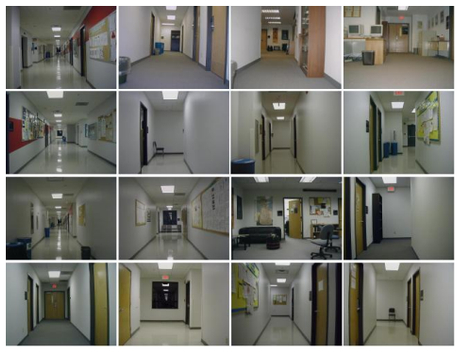
16-311-Q Introduction to Robotics Fall'17

Lecture 18:

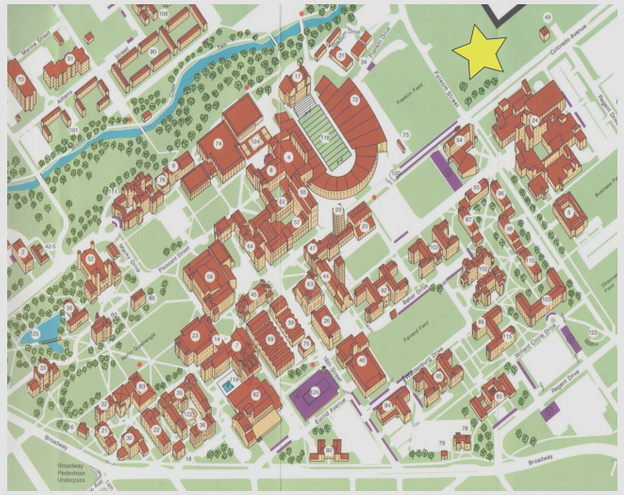
Kalman Filter 1

Instructor:
Gianni A. Di Caro

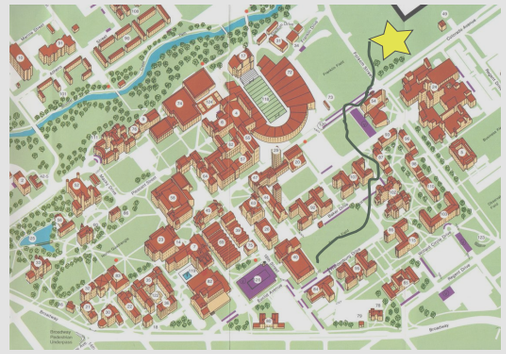
NAVIGATION FOR MOBILE ROBOTS



Where am I?



Where am I going?



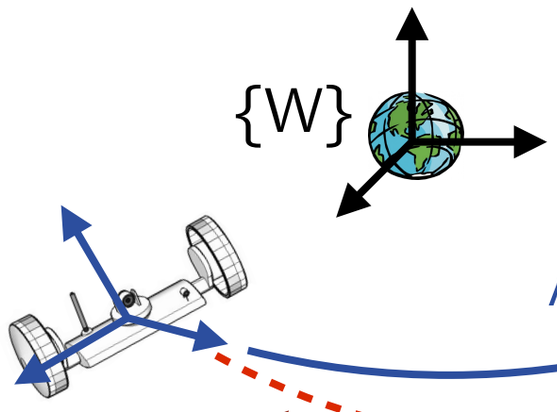
How do I get there?

Localization
(State estimation)

Input from sensors for relative and global localization

Reference system
Map

MOBILE ROBOT LOCALIZATION



Actual trajectory $\xi(t) = [x(t)\ y(t)\ \theta(t)]$

Estimated trajectory using odometry $\hat{\xi}(t) = [\hat{x}(t)\ \hat{y}(t)\ \hat{\theta}(t)]$

$[v\ \omega]_k$ $[v\ \omega]_{k+1}$ ... $[v\ \omega]_{k+n}$ Controls issued at discrete time steps t_k

$$W\begin{bmatrix} x(t+\delta t) \\ y(t+\delta t) \\ \theta(t+\delta t) \end{bmatrix} = \begin{bmatrix} x(t) + \frac{v(t)}{\omega(t)}\left(\sin\left(\theta(t) + \Delta\theta(t+\delta t)\right) - \sin\left(\theta(t)\right)\right) \\ y(t) - \frac{v(t)}{\omega(t)}\left(\cos\left(\theta(t) + \Delta\theta(t+\delta t)\right) - \cos\left(\theta(t)\right)\right) \\ \theta(t) + \omega(t)\delta t \end{bmatrix}$$

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + \Delta S_k \cos\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ y_k + \Delta S_k \sin\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ \theta_k + \Delta\theta_k \end{bmatrix}$$

Runge-Kutta numeric integration

Exact (***ideal***) pose based on kinematics

Estimated pose based on kinematics + measures

$\xi(t)$

vs.

$\hat{\xi}_{t_k}$

Noise / Uncertainties

MOBILE ROBOT = DISCRETE-TIME DYNAMIC SYSTEM

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + \Delta S_k \cos\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ y_k + \Delta S_k \sin\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ \theta_k + \Delta\theta_k \end{bmatrix}$$

Runge-Kutta numeric integration

\+

Noise / Uncertainties

Looking at the mobile robot as a ***discrete-time dynamic system*** with *control inputs* $[v\ \omega]_k$ or, equivalently, with *measured control inputs* ΔS_k and $\Delta\theta_k$, the discrete-time *process dynamics* of system state (the *pose*) can be modeled by the *numeric kinematic equations* with the addition of a *noise* component, aimed to capture all the *random and uncertain* aspects of the actual motion

$$\xi_{k+1} = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} (\Delta S_k + \nu_k^s)\cos(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ (\Delta S_k + \nu_k^s)\sin(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ \Delta\theta_k + \nu_k^\theta \end{bmatrix}$$

A random vector ν_k is added to the control inputs

A STOCHASTIC DYNAMIC SYSTEM

$$\xi_{k+1} = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} (\Delta S_k + \nu_k^s)\cos(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ (\Delta S_k + \nu_k^s)\sin(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ \Delta\theta_k + \nu_k^\theta \end{bmatrix}$$

Stochastic, discrete-time dynamic system

In absence of specific information, a first solution which is "safe" and amenable to analytical computations is to model noise as ***Gaussian white noise:***

$$\nu_k = \begin{bmatrix} \nu_k^s & \nu_k^\theta \end{bmatrix}^T \sim N(0, \mathbf{V}_k)$$

If the two noise components are also *uncorrelated*:

$$\mathbf{V}_k = \begin{bmatrix} \sigma_{ks}^2 & 0 \\ 0 & \sigma_{k\theta}^2 \end{bmatrix}$$

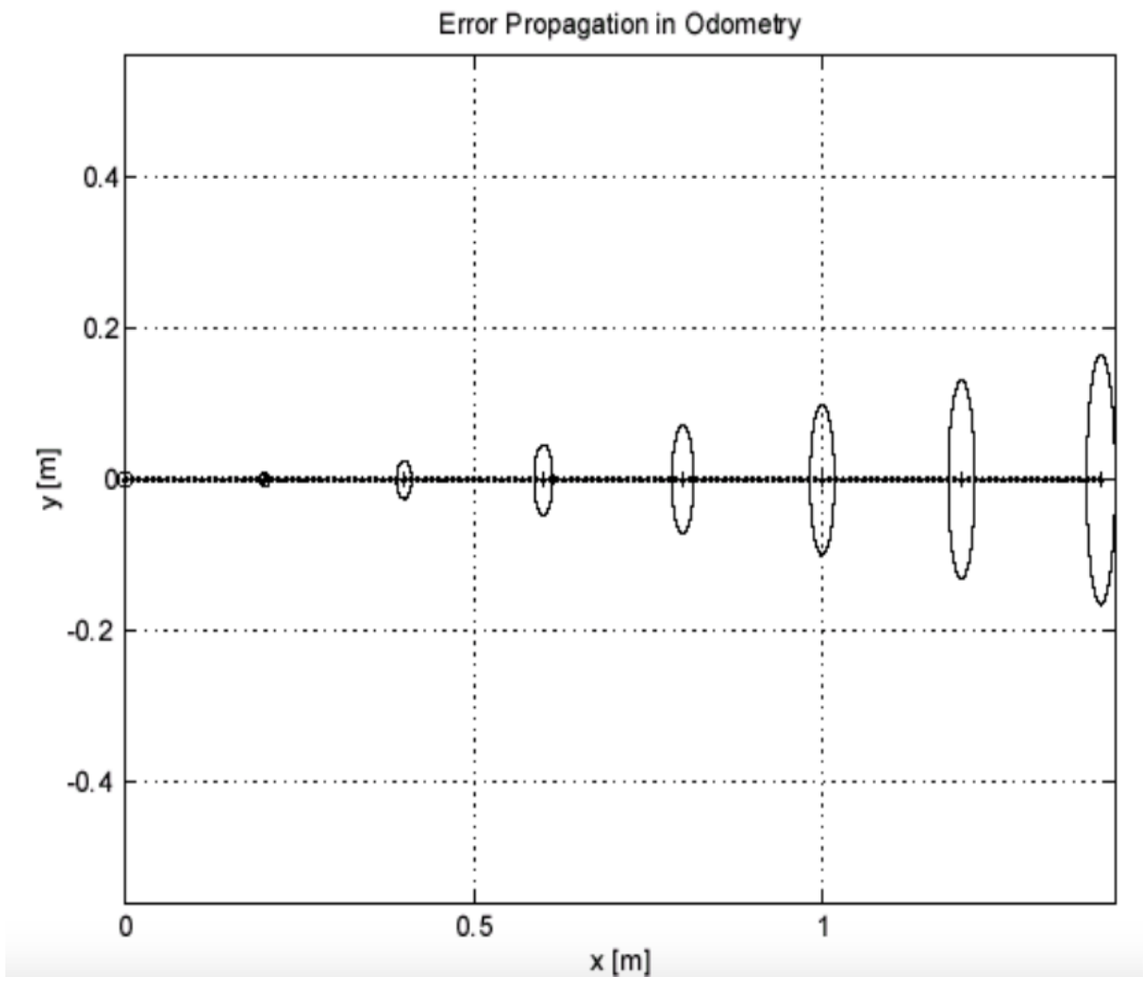
With the introduction of a random vector to model uncertainties, ξ_k, the state of the system—the pose of the robot at time k—becomes a *random variable* that represents a ***probabilistic estimate of the true state of the system***

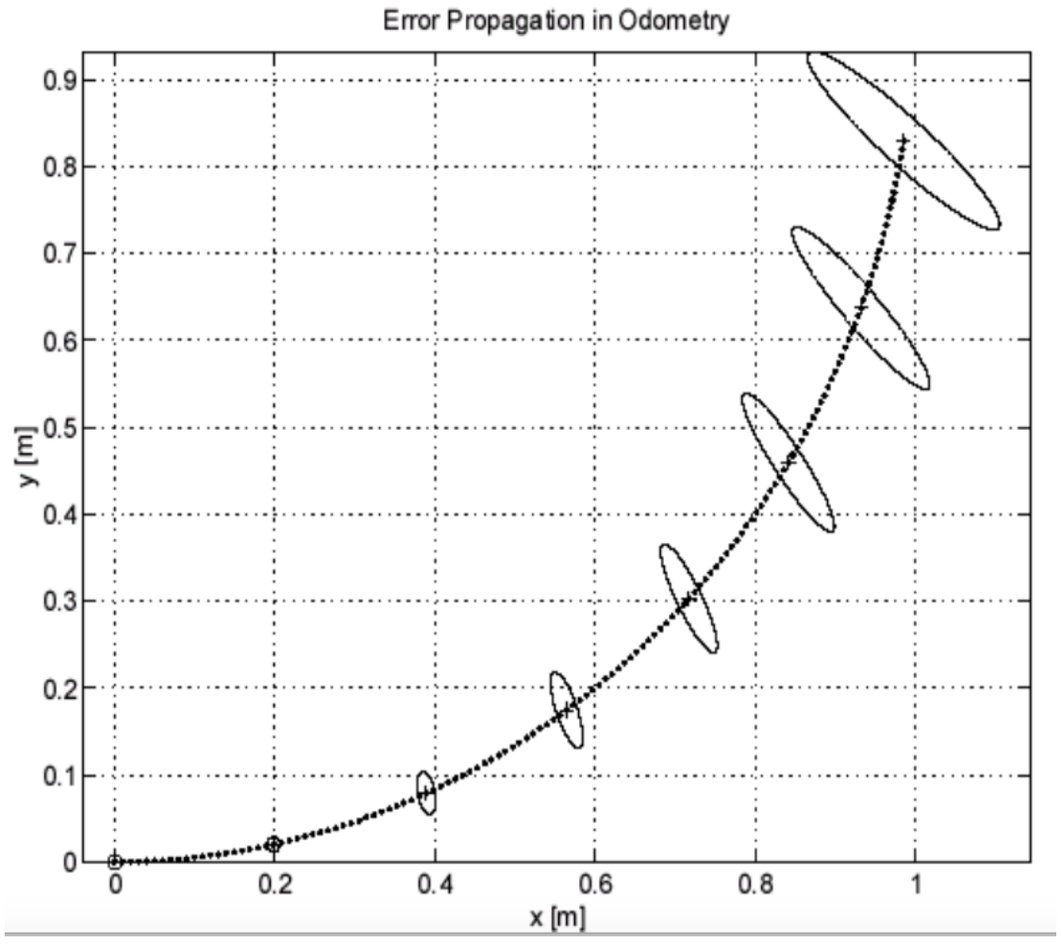
The ***mean*** (expected value) of ξ_k represents the current estimated pose, while its ***covariance*** P_k, a 3x3 matrix, represents the estimated error for the x, y, and θ components

COVARIANCE GROWS UNBOUNDED....

$$\xi_{k+1} = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} (\Delta S_k + \nu_k^s)\cos(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ (\Delta S_k + \nu_k^s)\sin(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ \Delta\theta_k + \nu_k^\theta \end{bmatrix}$$

Example of evolution of the covariance for straight and circular motions (and equation linearization)





Errors are ellipses, because of the Gaussian assumption for the noise

HOW DO WE COMPUTE MEAN AND COVARIANCE?

$$\xi_{k+1} = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} (\Delta S_k + \nu_k^s)\cos(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ (\Delta S_k + \nu_k^s)\sin(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ \Delta\theta_k + \nu_k^\theta \end{bmatrix}$$

$$\boldsymbol{\xi}_{k+1} = f(\boldsymbol{\xi}_k, \Delta S_k, \Delta\theta_k, \boldsymbol{\nu}_k), \quad \boldsymbol{\nu}_k = \begin{bmatrix} \nu_k^s & \nu_k^\theta \end{bmatrix}^T \sim N(0, \mathbf{V}_k)$$

***f* is a complex, highly non linear vector function**

Linearization
is a feasible strategy!

Expected value of ξ_{k+1}?

Covariance of ξ_{k+1}?

From probability theory: ***Linear transformations of Gaussian RVs are Gaussian***

$$\mathbf{X} \sim N(\boldsymbol{\mu}_x, \Sigma_x), \quad \mathbf{Y} = A\mathbf{X} + \mathbf{b} \quad \Longrightarrow \quad \mathbf{Y} \sim N(A\boldsymbol{\mu}_x + \mathbf{b},\ A\Sigma_x A^T)$$

From probability theory: Sums (and Difference) of Gaussian RVs are Gaussian

$$X \sim N(\mu_x, \Sigma_x),\ Y \sim N(\mu_y, \Sigma_y), \quad \begin{array}{lll} Z = X+Y & \Longrightarrow & \mu_Z = \mu_X + \mu_Y, \quad \sigma_Z = \sigma_X + \sigma_Y + \sigma_{XY} \\ Z = X-Y & \Longrightarrow & \mu_Z = \mu_X - \mu_Y, \quad \sigma_Z = \sigma_X + \sigma_Y - \sigma_{XY} \end{array}$$

LINEARIZATION + EXTEROCEPTIVE MEASURES

$$\xi_{k+1} = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} (\Delta S_k + \nu_k^s)\cos(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ (\Delta S_k + \nu_k^s)\sin(\theta_k + \frac{\Delta\theta_k}{2} + \nu_k^\theta) \\ \Delta\theta_k + \nu_k^\theta \end{bmatrix}$$

Proprioceptive measures / Relative localization

Linearization of equations → Mean and Covariance of ξ_{k+1}

... the error / covariance of the state estimate will keep growing over time ...!

Exteroceptive measures / Global localization

Shrink the error!

Mean accuracy ↑
Covariance error ↓

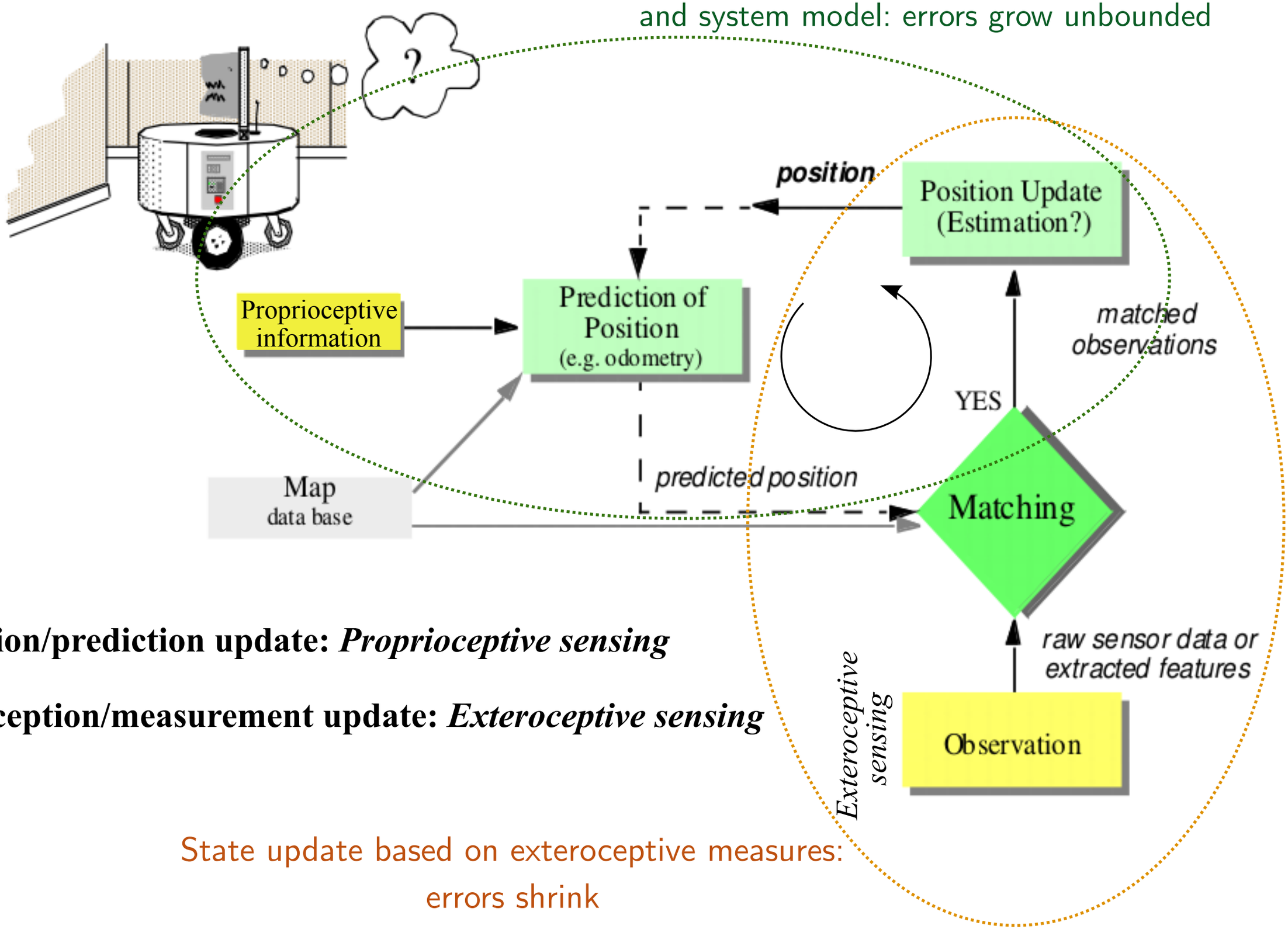
Coordinates from GPS

Detection of ground landmark with know position on the map → External (possibly noisy, but reliable) measures of the current state ξ_{k+1}

...

SCENARIO FOR MAP-BASED LOCALIZATION

State update/prediction based on proprioceptive sensing and system model: errors grow unbounded



1. Action/prediction update: *Proprioceptive sensing*

2. Perception/measurement update: *Exteroceptive sensing*

State update based on exteroceptive measures: errors shrink

LINEAR+GAUSSIAN STATE ESTIMATION: KALMAN FILTER (1960)

Given models:

- A linear discrete-time dynamic system, whose ***state*** $\boldsymbol{\xi}_k$ is not directly observable, is used to represent the ***evolution dynamics*** of the system of interest
- A linear control model relates how issued ***control inputs*** $\mathbf{u}_k$ affects system's dynamics
- A linear observation model relates ***sensor measurements*** $\mathbf{z}_{k+1}$ with the system state
- Gaussian white noise models $\boldsymbol{\nu}_k$, $\mathbf{w}_k$ (with associated covariance matrix $\mathbf{V}_k$ and $\mathbf{W}_k$ respectively) are used to express ***uncertainties for controls and measures*** respectively

Given parameters and initial conditions:

$$\boldsymbol{\xi}_k \sim N(\widehat{\boldsymbol{\xi}}_k, \boldsymbol{P}_k),\ \boldsymbol{\nu}_k \sim N(0, \boldsymbol{V}_k),\ \boldsymbol{W}_k \sim N(0, \boldsymbol{W}_k),\ \boldsymbol{u}_k,\ \boldsymbol{\xi}_0, \boldsymbol{P}_0, \boldsymbol{V}, \boldsymbol{W}$$

System equations

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{A}_k \boldsymbol{\xi}_k + \boldsymbol{B}_k \boldsymbol{u}_k + \boldsymbol{\nu}_k$$

$$\boldsymbol{z}_{k+1} = \boldsymbol{C}_k \boldsymbol{\xi}_k + \boldsymbol{w}_k$$

← ***Process/motion dynamics model*** (proprioceptive):
Predict next state based on current state, inputs, and noise

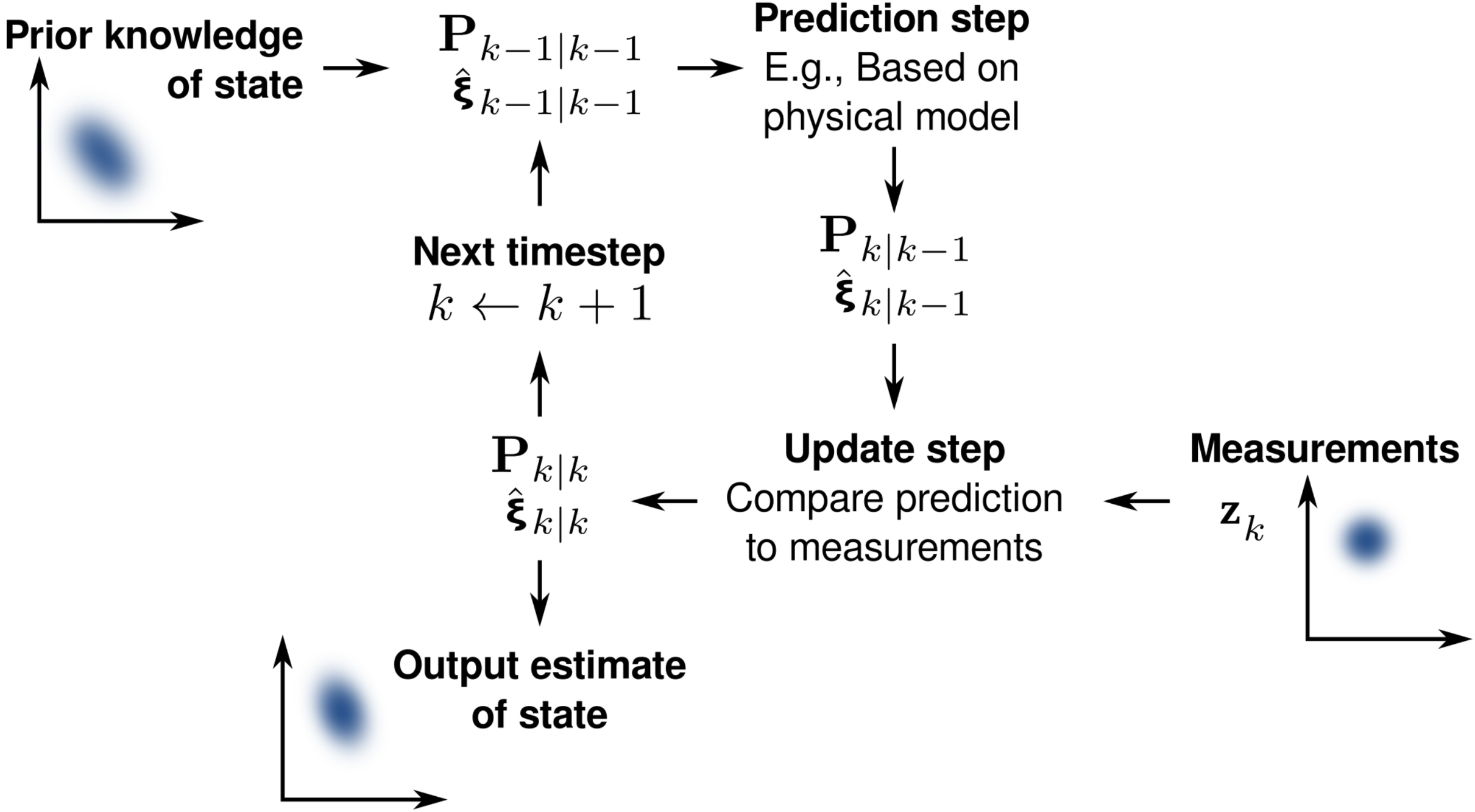
← ***Measurement/observation model*** (exteroceptive):
Predict next observation based on current state and noise

THE COMPONENTS OF THE DYNAMIC SYSTEM

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{A}_k\boldsymbol{\xi}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{\nu}_k \qquad \boldsymbol{z}_{k+1} = \boldsymbol{C}_k\boldsymbol{\xi}_k + \boldsymbol{w}_k$$

- $\boldsymbol{\xi}_k \in \mathbb{R}^n$ is the state vector of the system (at step k) to be estimated together with its covariance (uncertainty) matrix $\boldsymbol{P}_k$
- $\boldsymbol{A}_k \in n \times n$, is the state transition matrix, that relates system's states in the absence of a *forcing* function (expressed by the controls $\boldsymbol{u}_k$)
- $\boldsymbol{B}_k \in n \times q$, is the matrix that relates how controls affect state transitions, it transforms the control vector into the coordinates of the state
- $\boldsymbol{\nu}_k \in \mathbb{R}^n$, is a vector of white Gaussian noise with zero mean and covariance matrix $\boldsymbol{V}_k$, $\boldsymbol{\nu}_k \sim N(0, \boldsymbol{V}_k)$, it expresses process noise
- $\boldsymbol{u}_k \in \mathbb{R}^q$, is the vector of control inputs issued at step k
- $\boldsymbol{z}_k \in \mathbb{R}^m$ is the measurement vector at step k, commonly $m > n$
- $\boldsymbol{C}_k \in m \times n$, is the measurement matrix, or observation matrix, that relates $\boldsymbol{\xi}_k$ to $\boldsymbol{z}_k$ in the absence of measurement noise
- $\boldsymbol{w}_k \in \mathbb{R}^m$, is a vector of white Gaussian noise with zero mean and covariance matrix $\boldsymbol{W}_k$, $\boldsymbol{w}_k \sim N(0, \boldsymbol{W}_k)$, it expresses measurement noise

DIAGRAM OF THE UPDATE SCHEME OF THE KF



KALMAN FILTER RECURSIVE EQUATIONS

Recursively build an estimate $\hat{\boldsymbol{\xi}}_{k+1}$ of the true state of the system, $\boldsymbol{\xi}_{k+1}$, in two steps:

1. ***Generate a prediction*** $\hat{\boldsymbol{\xi}}_{k+1|k}$ by propagating $\hat{\boldsymbol{\xi}}_k$ based on *process dynamics* and *controls*
2. ***Correct state prediction*** by minimizing the expected value of the *innovation*: the difference between the *predicted state* $\hat{\boldsymbol{\xi}}_{k+1|k}$ and the *observed measures* $\mathbf{z}_{k+1}$

At every time step k:

$$\widehat{\boldsymbol{\xi}}_{k+1|k} = \boldsymbol{A}_k \widehat{\boldsymbol{\xi}}_k + \boldsymbol{B}_k \boldsymbol{u}_k \qquad \text{(State prediction)}$$

$$\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{V}_k \qquad \text{(Covariance prediction)}$$

At every time step $k+1$ when an *observation* $\boldsymbol{z}_{k+1}$ is available:

$$\widehat{\boldsymbol{\xi}}_{k+1} = \widehat{\boldsymbol{\xi}}_{k+1|k} + \boldsymbol{G}_{k+1}(\boldsymbol{z}_{k+1} - \boldsymbol{C}_{k+1}\widehat{\boldsymbol{\xi}}_{k+1|k}) \qquad \text{(Correction for innovation} \rightarrow \text{State update)}$$

$$\boldsymbol{P}_{k+1} = \boldsymbol{P}_{k+1|k} - \boldsymbol{G}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k} \qquad \text{(Covariance correction} \rightarrow \text{Covariance update)}$$

$$\boldsymbol{G}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^T(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^T + \boldsymbol{W}_{k+1})^{-1} = \boldsymbol{C}_{k+1,M}^{+} + \boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^T\boldsymbol{W}_{k+1}^{-1} \text{ (Gain)}$$

A COMPUTATIONAL VIEW OF THE RECURSIVE EQUATIONS

State prediction
(Predict where the system state will be)

$$\widehat{\xi}_{k+1|k} = A_k \widehat{\xi}_k + B_k u_k$$

Covariance prediction
(Predict the amount of error in state prediction)

$$P_{k+1|k} = A_k P_k A_k^T + V_k$$

Innovation
(Compare reality against prediction)

$$\epsilon_{k+1} = z_{k+1} - C_{k+1} \widehat{\xi}_{k+1|k}$$

Innovation covariance
(Compare real error against prediction)

$$\Sigma_{\epsilon_{k+1}} = C_{k+1} P_{k+1|k} C_{k+1}^T + W_{k+1}$$

Note: z_{k+1} and $C_{k+1}\widehat{\xi}_{k+1|k}$ are independent RVs $\rightarrow \text{Cov}(\epsilon_{k+1}) = \text{Cov}(z_{k+1}) + \text{Cov}(C_{k+1}\widehat{\xi}_{k+1|k})$

Kalman gain
(Rescale/weight the prediction)

$$G_{k+1} = P_{k+1|k} C_{k+1}^T \Sigma_{\epsilon_{k+1}}^{-1}$$

State update
(New estimate of the system state)

$$\widehat{\xi}_{k+1} = \widehat{\xi}_{k+1|k} + G_{k+1} \epsilon_{k+1}$$

Covariance update
(New estimate of error)

$$P_{k+1} = (I - G_{k+1} C_{k+1}) P_{k+1|k}$$

Blue = Inputs, Red = Outputs, **Black = Constant Parameters**, Gray = Working variables

OBSERVATIONS ON THE KALMAN FILTER

- At each time step k, the ***Kalman Gain*** matrix $\mathbf{G}_k$ weighs the accuracy of the prediction based on the process model vs. the accuracy of the (exteroceptive) measurements, such that the correction to the state is done accordingly:
  - **G** large, means that measurements are more reliable than model predictions
  - **G** small, means that model predictions are more reliable than measurements

$\rightarrow$ State estimate correction can be more or less large based on how reliable are sensor measurements

- The Kalman filter is ***optimal***, in the sense that for each k, $E[\hat{\boldsymbol{\xi}}_{k+1} - \boldsymbol{\xi}_{k+1}]$ is *minimized*

- The Kalman filter is ***correct***, in the sense that the estimates are the conditional expectations of the true state values:

$$\widehat{\boldsymbol{\xi}}_k = E\left[\boldsymbol{\xi}_k | z_k\right], \quad \widehat{\boldsymbol{\xi}}_{k|k-1} = E\left[\boldsymbol{\xi}_k | z_{k-1}\right]$$

IF EVERYTHING IS GAUSSIAN, IT STAYS GAUSSIAN

- State Belief - Belief distribution under Gaussian assumptions:
  Belief distributions are **posterior probabilities** over the state variables $\boldsymbol{\xi}_k$ conditioned on the available data:

$$bel(\boldsymbol{\xi}_k) = p(\boldsymbol{\xi}_k \mid z_{1:k}, \boldsymbol{u}_{1:k})$$

  Giving an initial belief as a Gaussian distribution, the linearity and Gaussian noise of the Kalman filter assumptions guarantee that the state belief stays Gaussian

- Prediction - Linear **motion** model under Gaussian noise:

$$p(\boldsymbol{\xi}_k \mid \boldsymbol{u}_k, \boldsymbol{\xi}_{k-1}) = \frac{1}{\det(2\pi \boldsymbol{V}_k)^{1/2}} exp\Big[-\frac{1}{2}(\boldsymbol{\xi}_k - \boldsymbol{A}_k\boldsymbol{\xi}_{k-1} - \boldsymbol{B}_K\boldsymbol{u}_k)\boldsymbol{V}_k^{-1}(\boldsymbol{\xi}_k - \boldsymbol{A}_k\boldsymbol{\xi}_{k-1} - \boldsymbol{B}_K\boldsymbol{u}_k)^T\Big]$$

  The covariance matrix $\boldsymbol{V}_k$ describes motion uncertainty at step k.

  The next state is computed as the one that maximizes the posterior probability, $\rightarrow$ The vector $\hat{\xi}_k$ that minimizes the quadratic form (Weighted Least Squares problem)

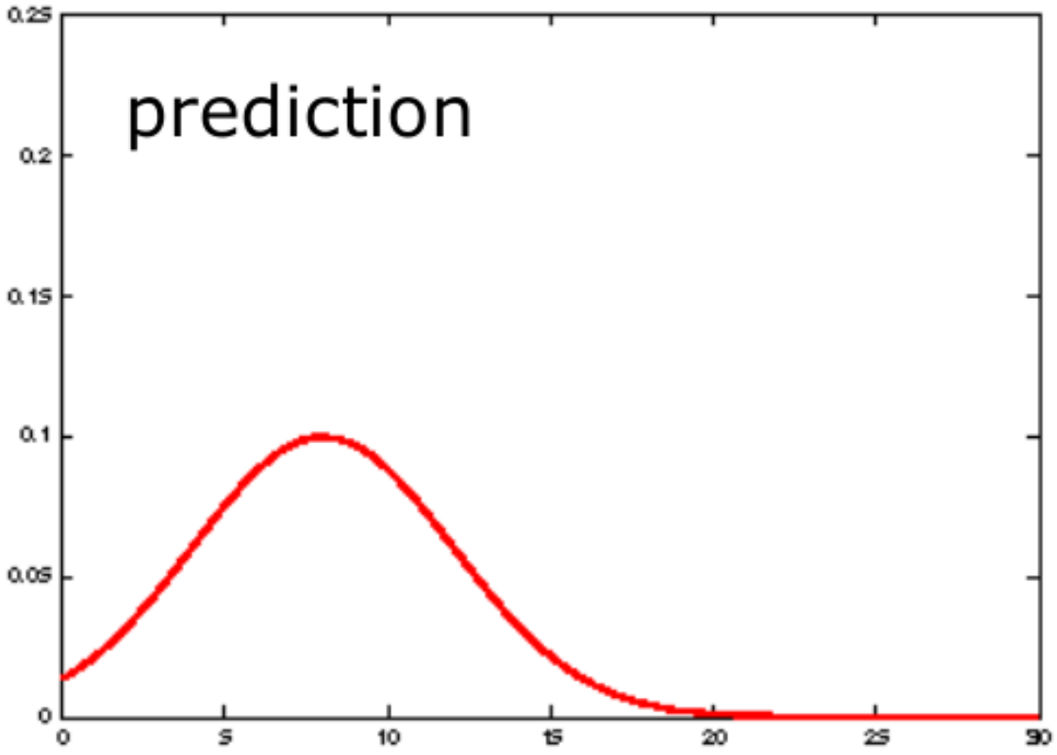
- Measurement - Linear **observation** model under Gaussian noise:

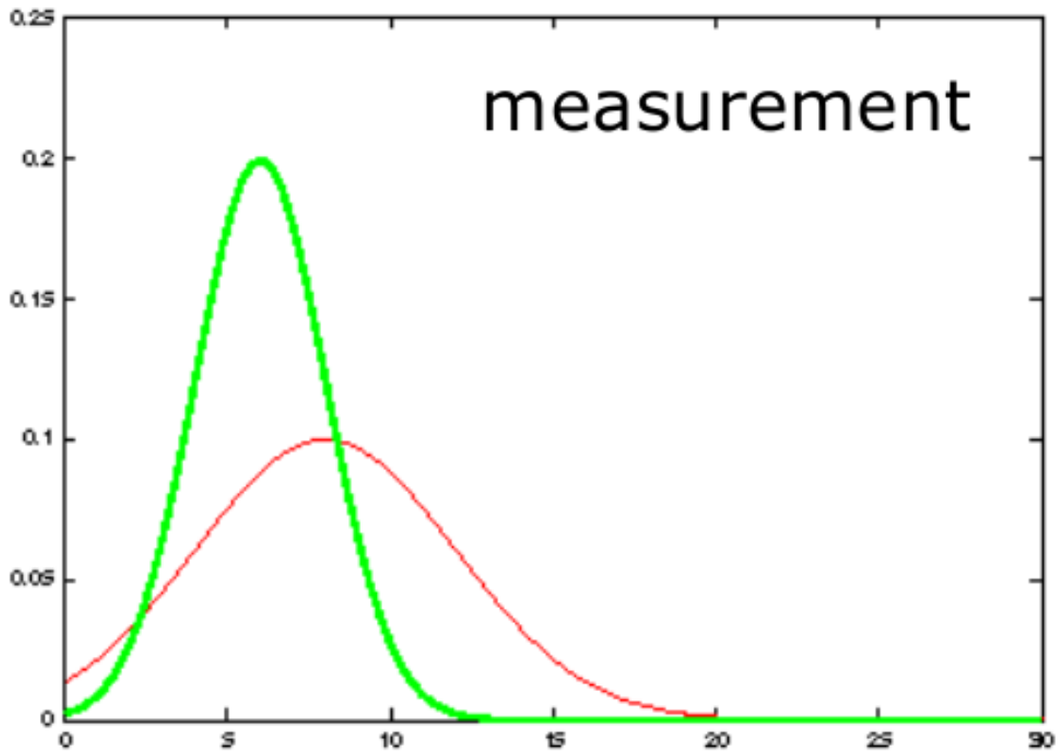
$$p(z_k \mid \boldsymbol{\xi}_k) = \frac{1}{\det(2\pi \boldsymbol{W}_k)^{1/2}} exp\Big[-\frac{1}{2}(z_k - \boldsymbol{C}_k\boldsymbol{\xi}_k)\boldsymbol{W}_k^{-1}(z_k - \boldsymbol{C}_k\boldsymbol{\xi}_k)^T\Big]$$

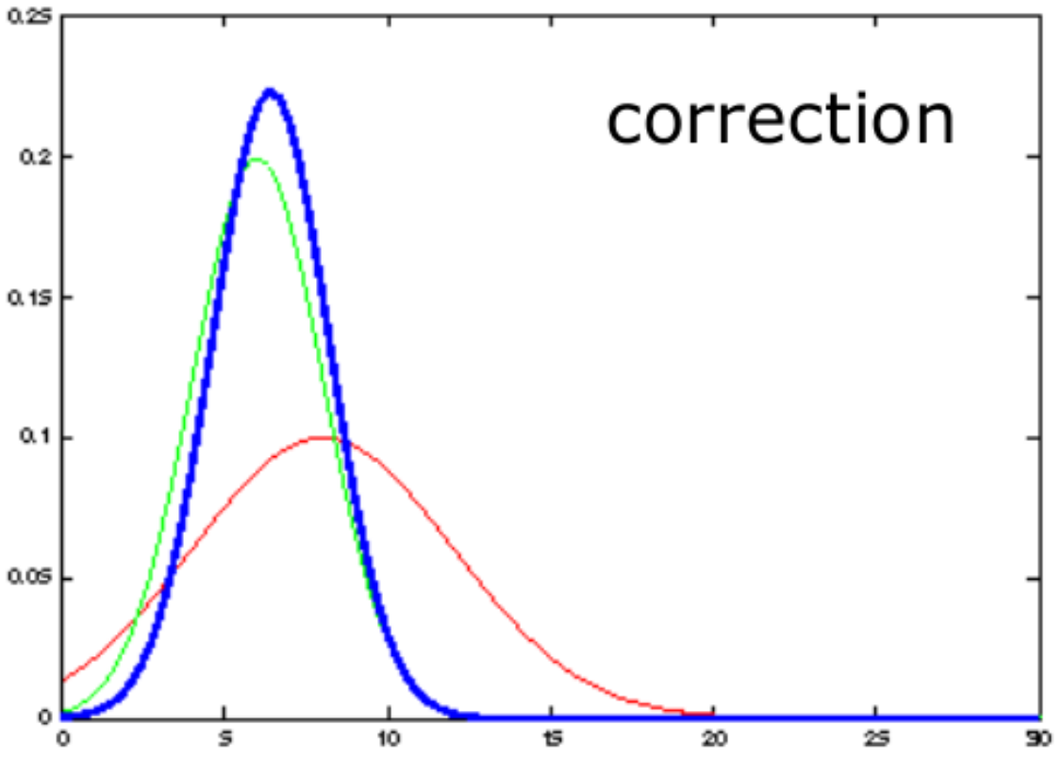
  The covariance matrix $\boldsymbol{W}_k$ describes measurement uncertainty at step k.

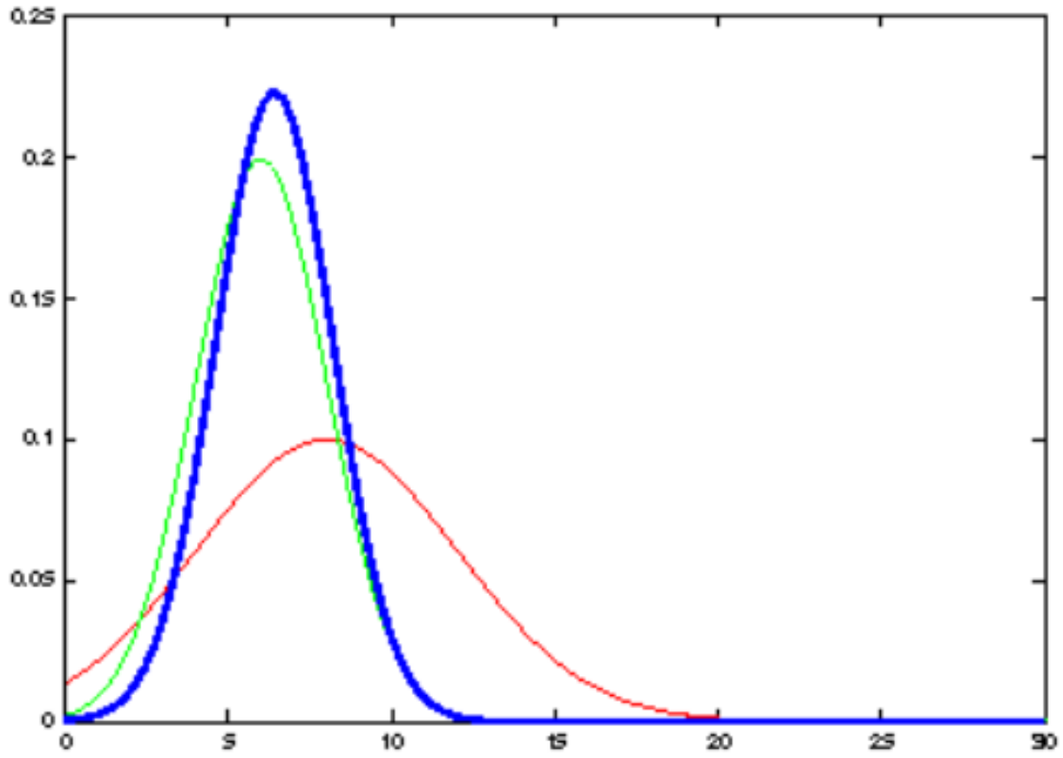
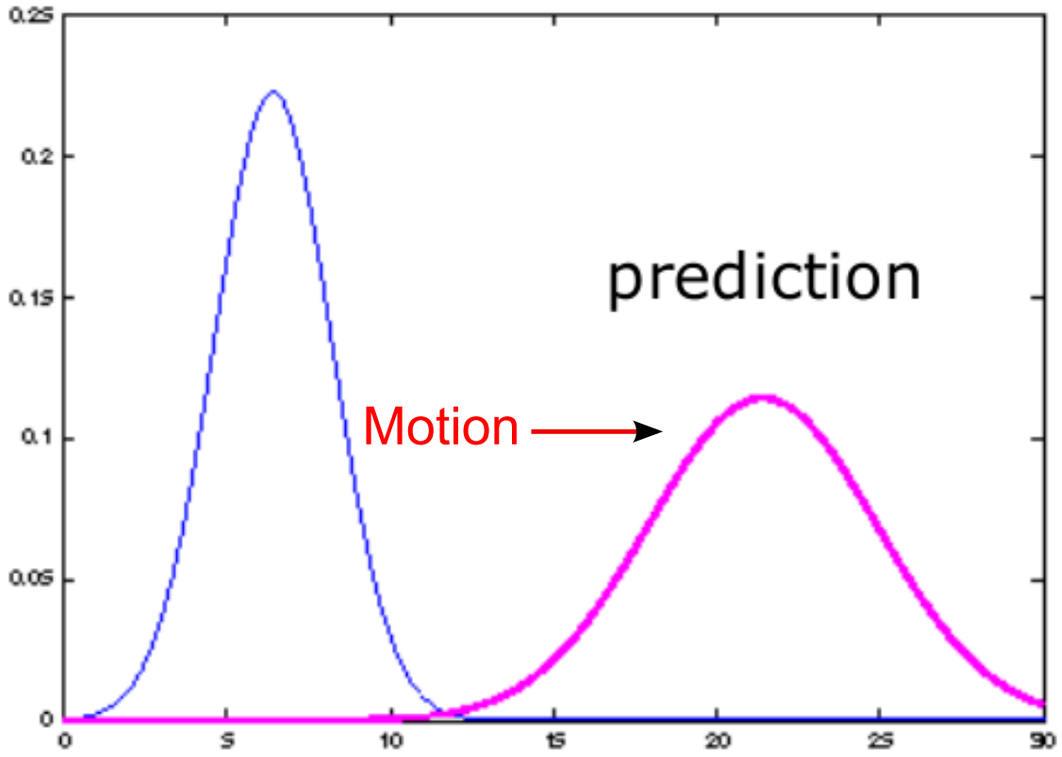
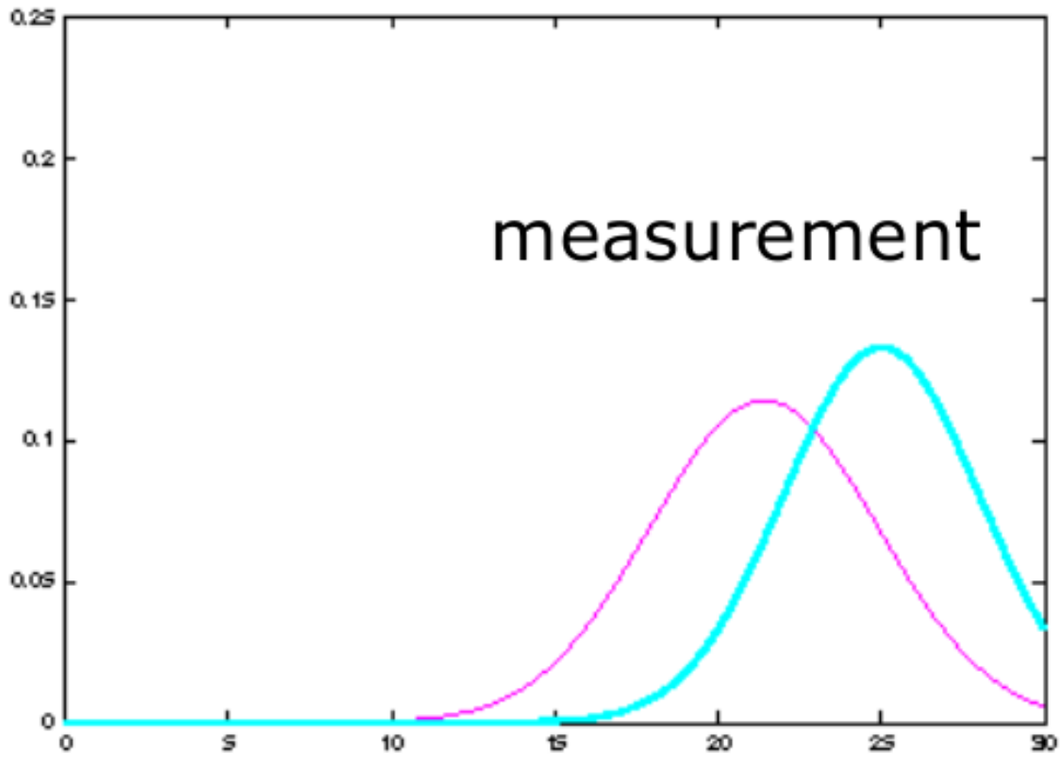
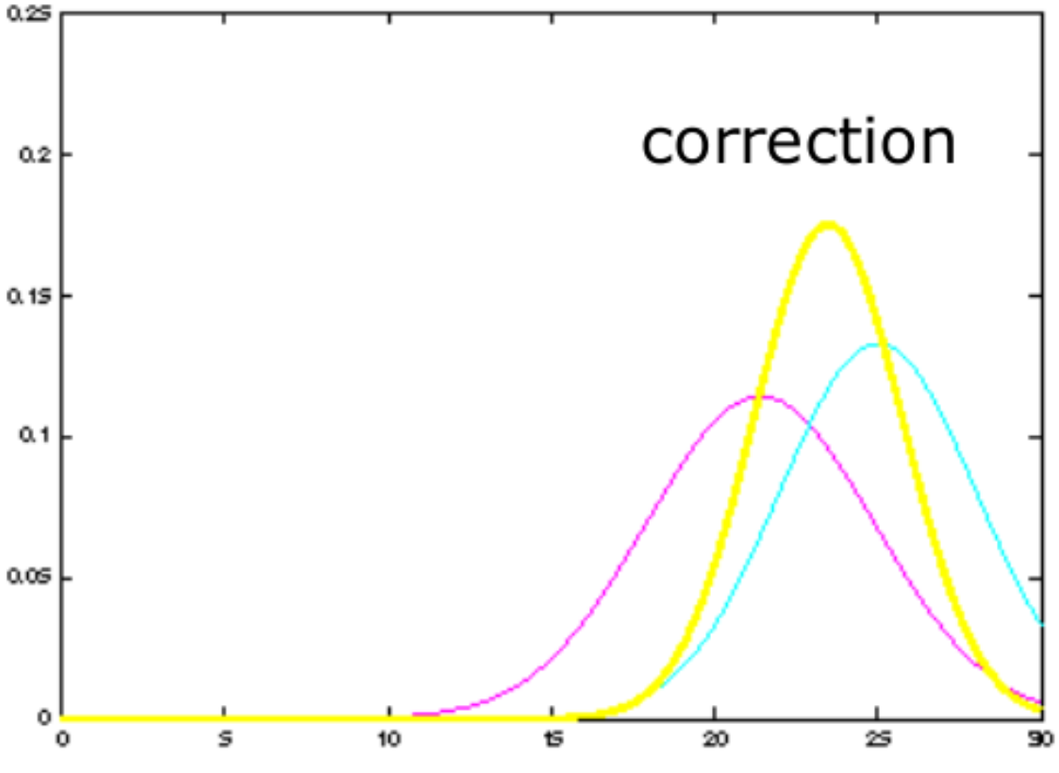
  The next state is computed as the one that maximizes the posterior probability, $\rightarrow$ The vector $\hat{\xi}_k$ that minimizes the quadratic form (Weighted Least Squares problem)

EXAMPLE OF KF BELIEF EVOLUTION IN A SCALAR CASE







It's a weighted mean!

EXAMPLE OF KF BELIEF EVOLUTION IN A SCALAR CASE