
# From Bayes to Kalman

Instructor: Chris Mavrogiannis

TAs: Kay Ke, Gilwoo Lee, Matt Schmittle

*Slides based on or adapted from Sanjiban Choudhury, Dieter Fox, and Matt Schmittle

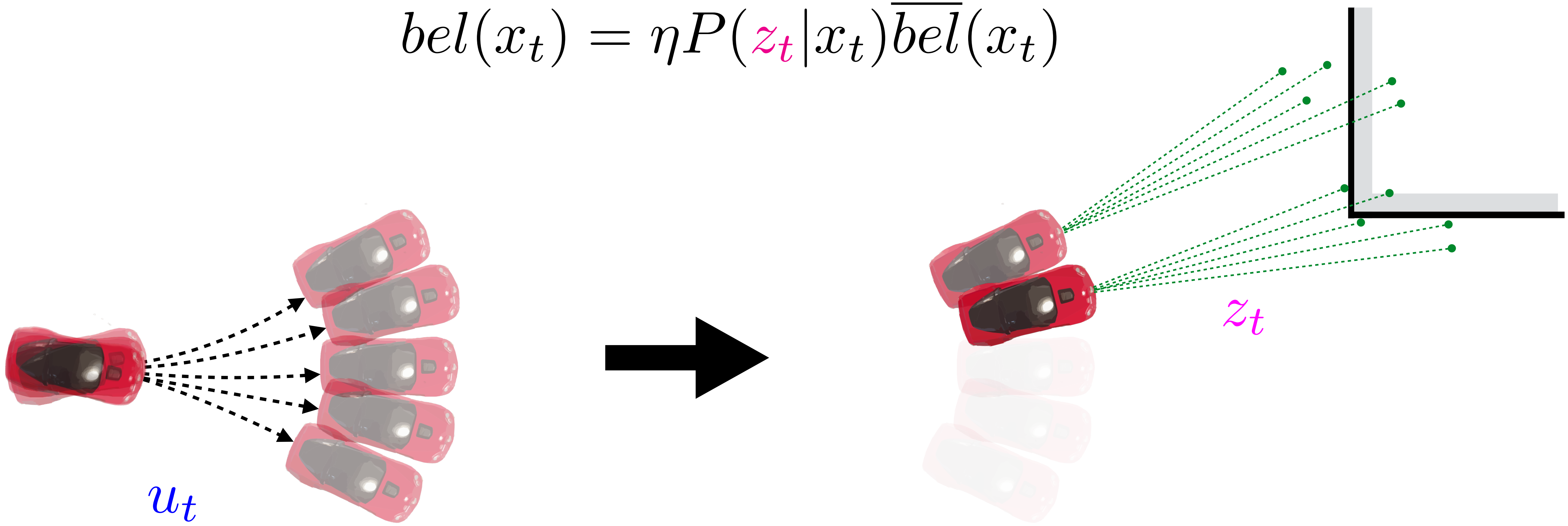
# Bayes filter in a nutshell

Step 1: Prediction - push belief through dynamics given action

$$\overline{bel}(x_t) = \int P(x_t|u_t, x_{t-1})bel(x_{t-1})dx_{t-1}$$

Step 2: Correction - apply Bayes rule given measurement

$$bel(x_t) = \eta P(z_t|x_t)\overline{bel}(x_t)$$

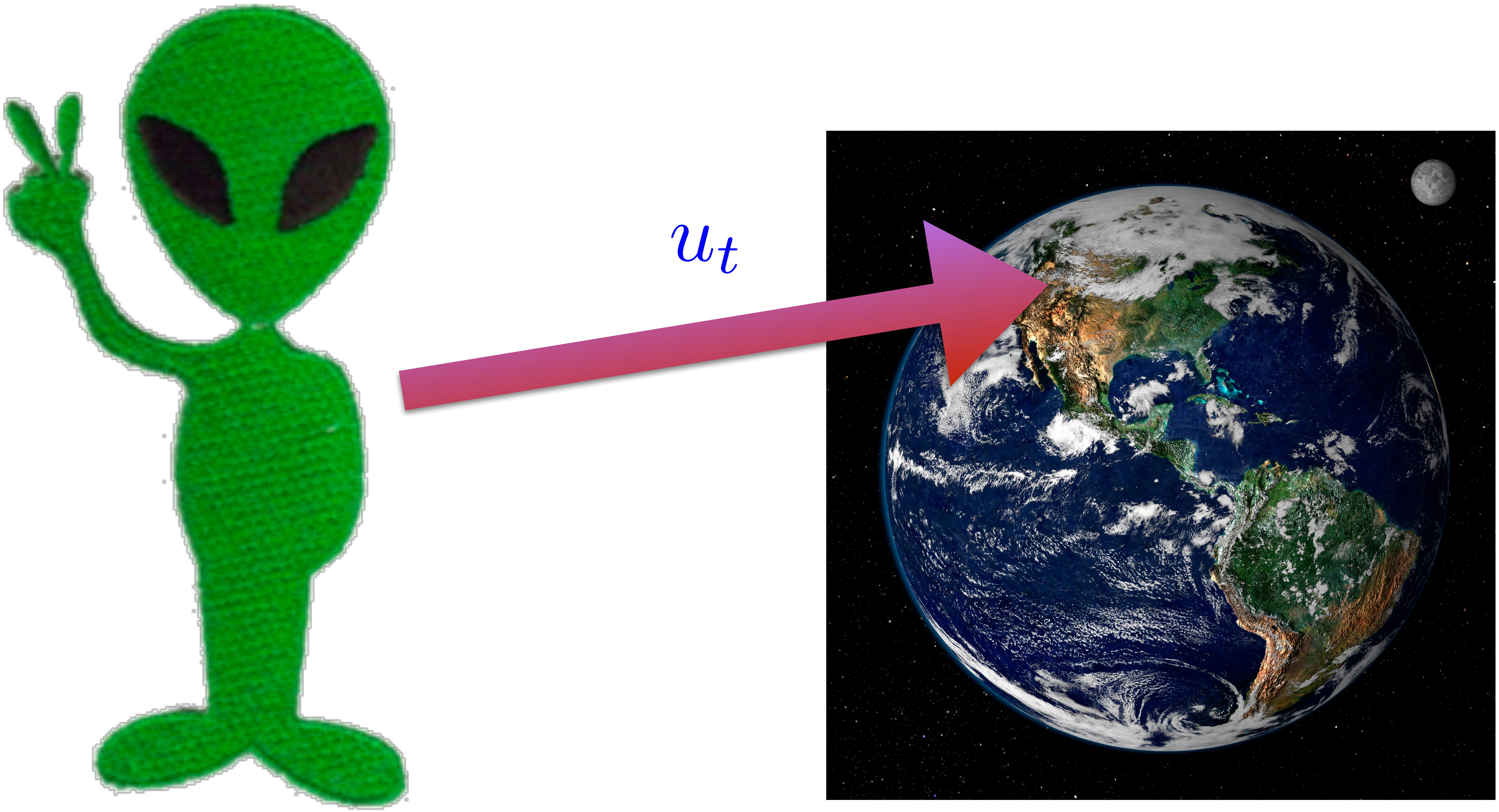



$bel(x_{t-1})$ $\overline{bel}(x_t)$ $bel(x_t)$

# Suppose you are an alien...

$bel(x_{t-1})$

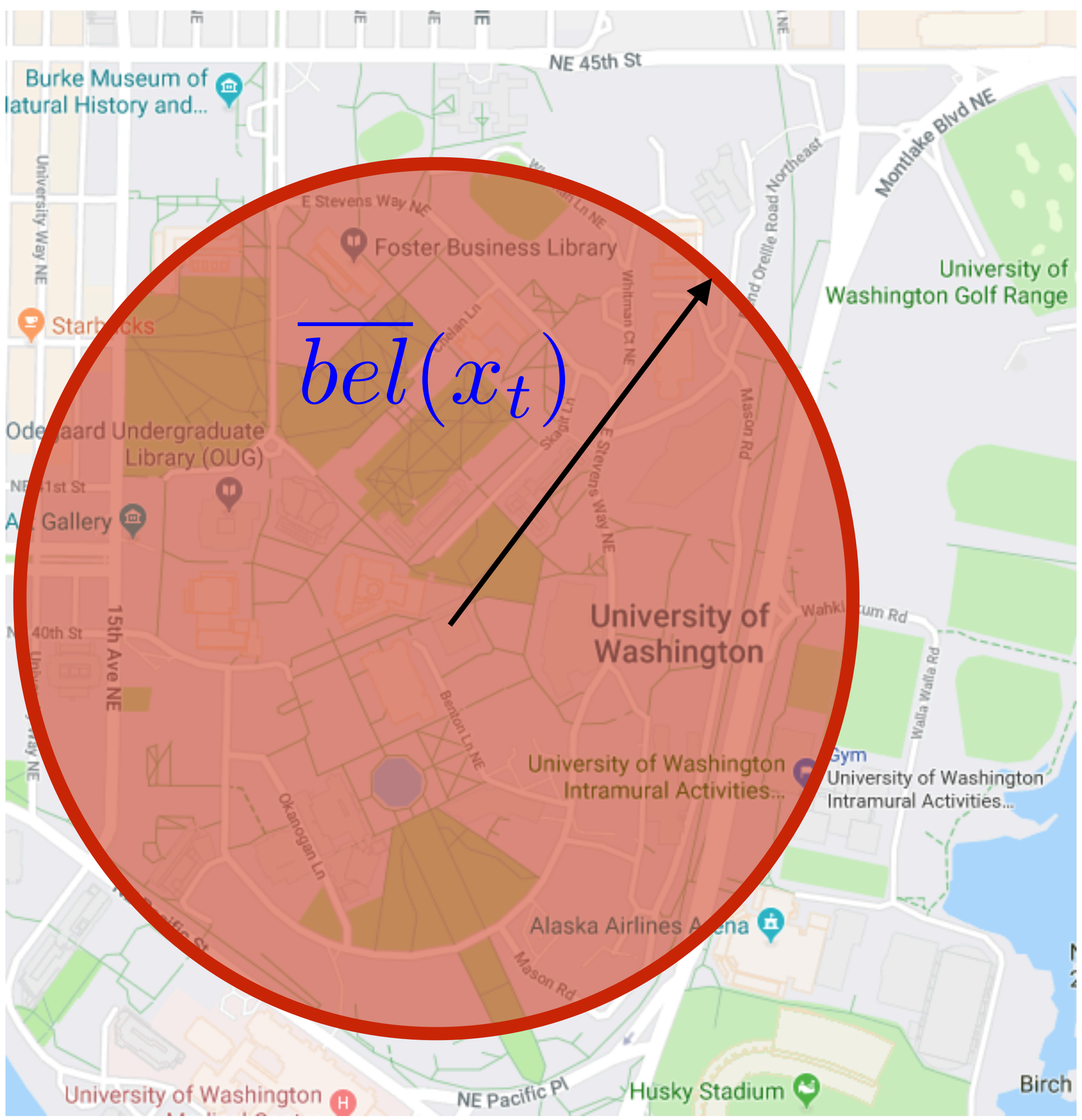
# ...beamed to earth ...



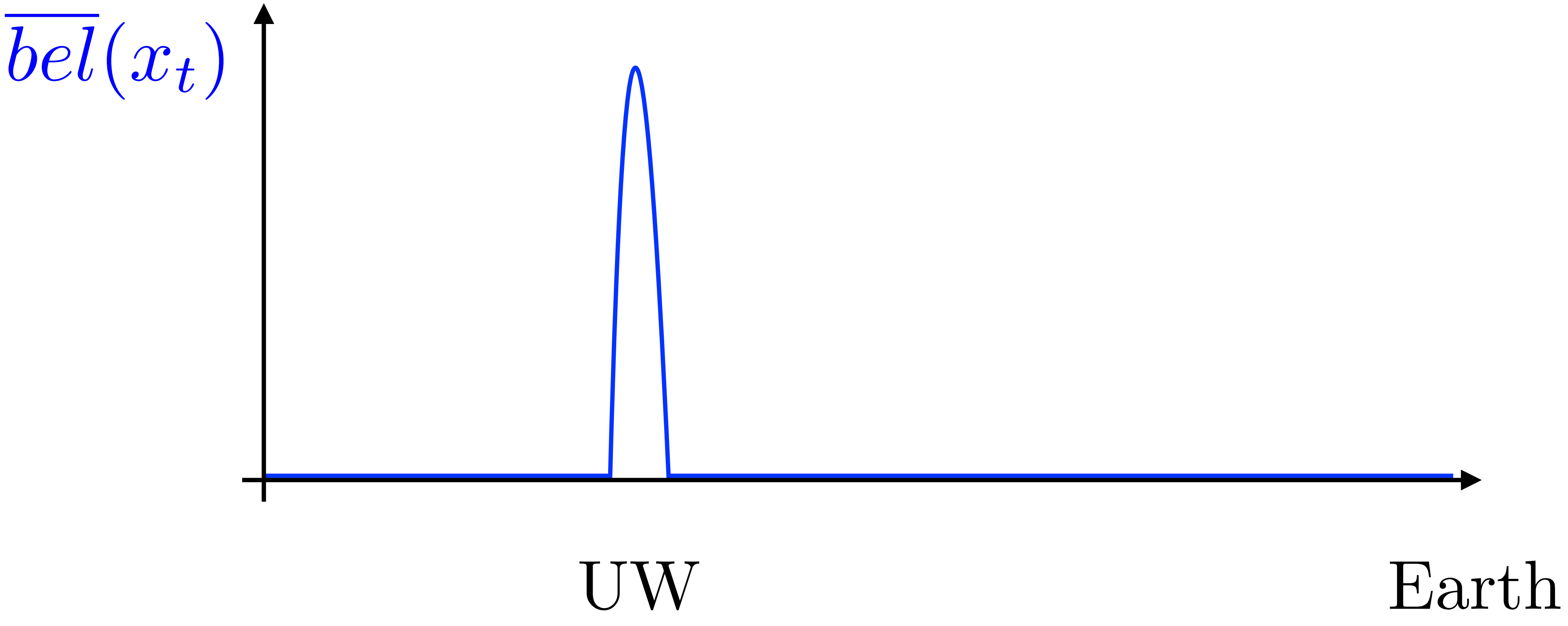
# .. and you predict you landed at UW

Prediction

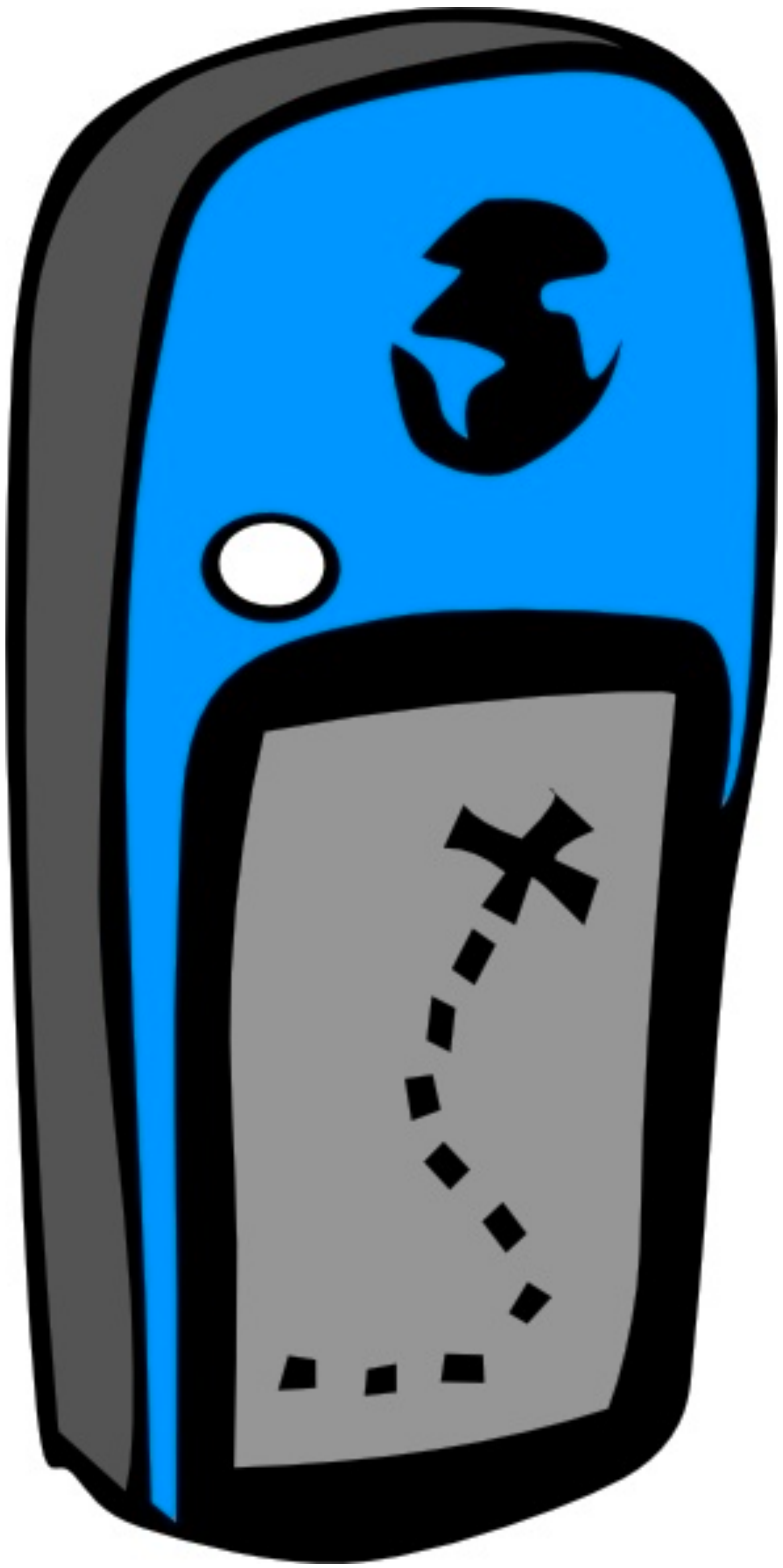
$$\overline{bel}(x_t) = \int P(x_t | u_t, x_{t-1}) bel(x_{t-1}) dx_{t-1}$$



# .. and you predict you landed at UW

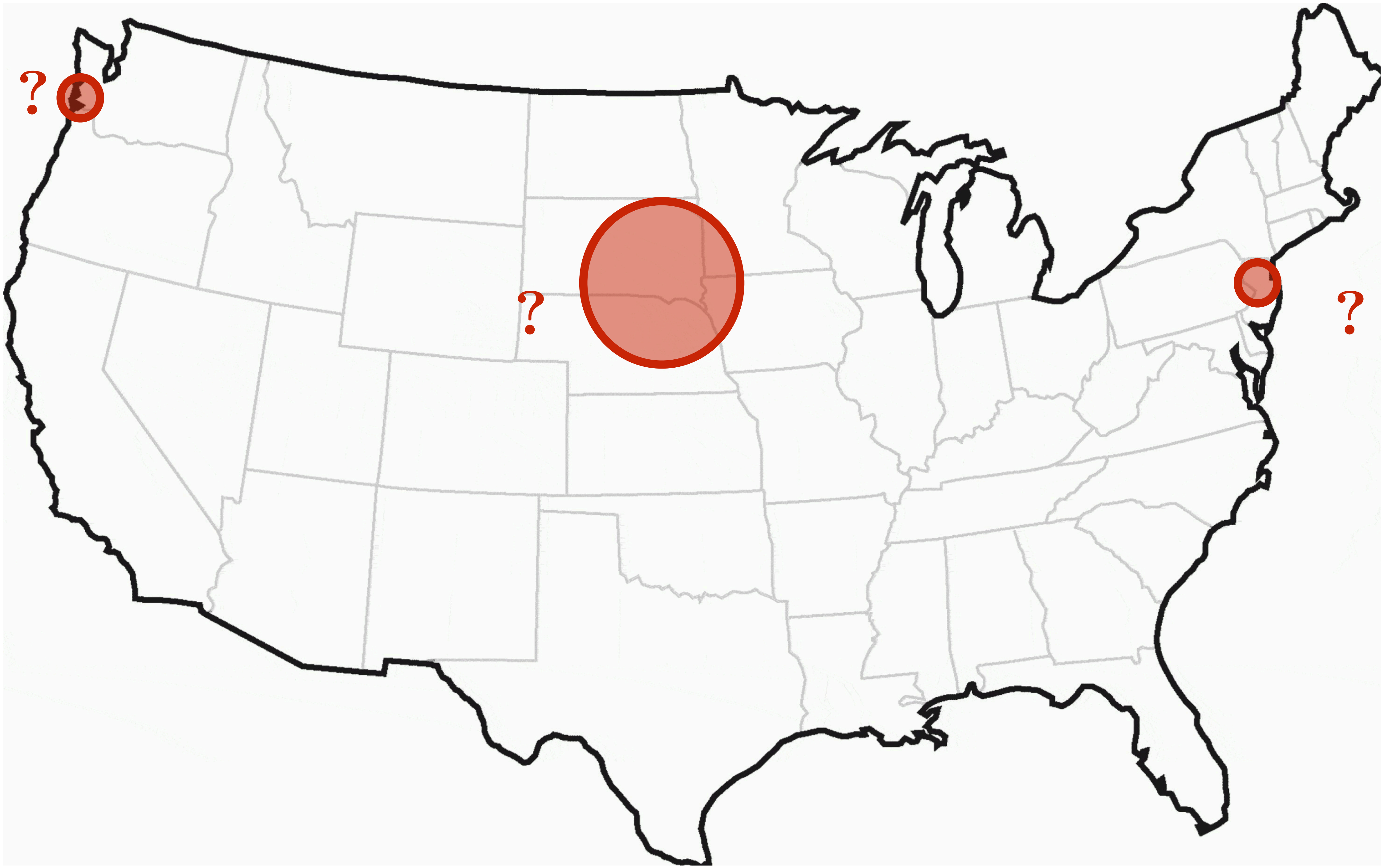


# Eventually GPS measurement comes in...



# … and says you are in New York

# What should we set as our new belief?



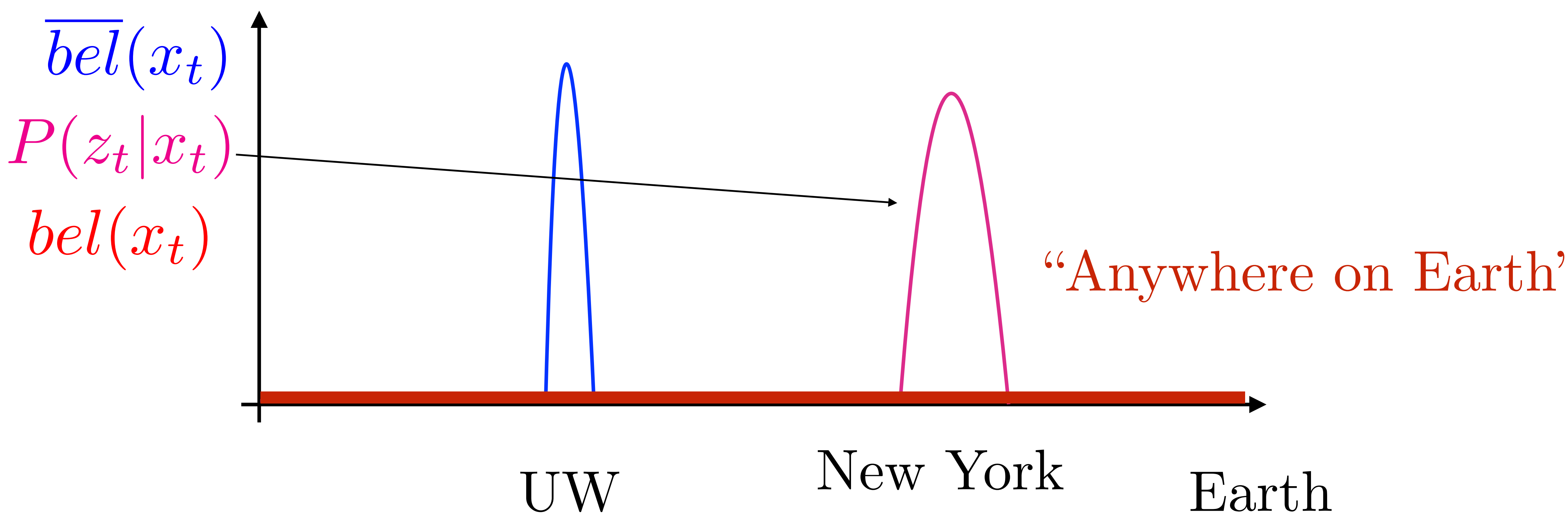

Depends on measurement uncertainty

# Case A: Measurement uncertainty is “small”

# Case A: Measurement uncertainty is "small"

Correction

$$bel(x_t) = \eta P(z_t|x_t)\overline{bel}(x_t)$$

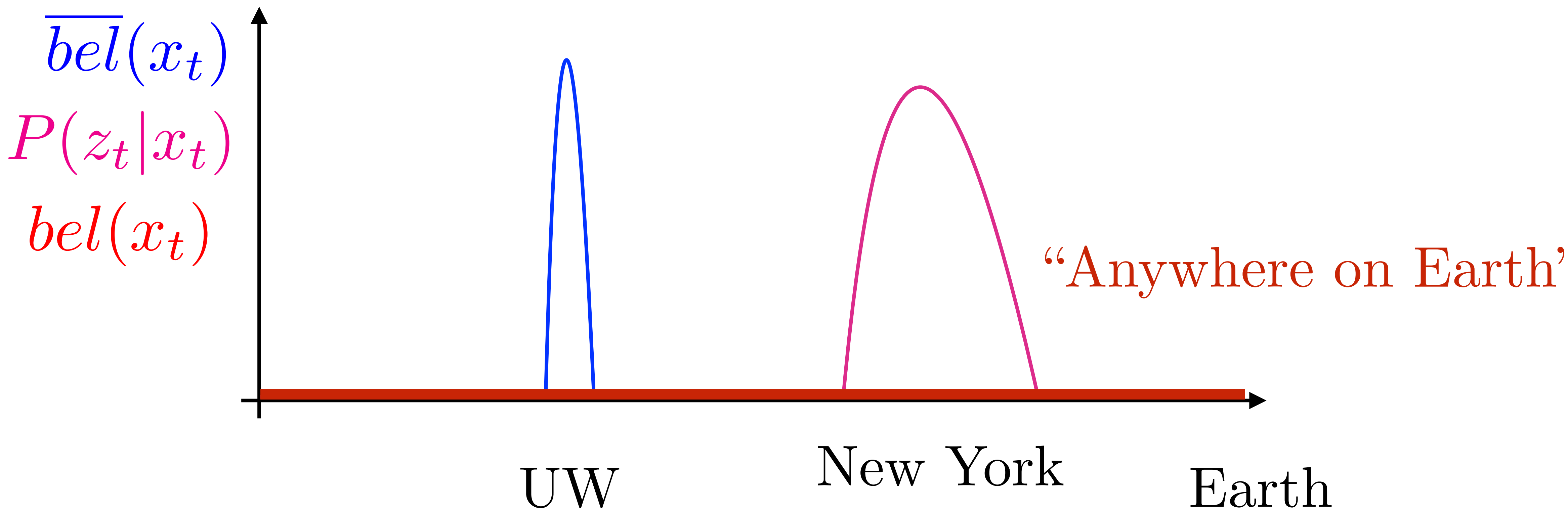


# Case B: Uncertainty is "medium"

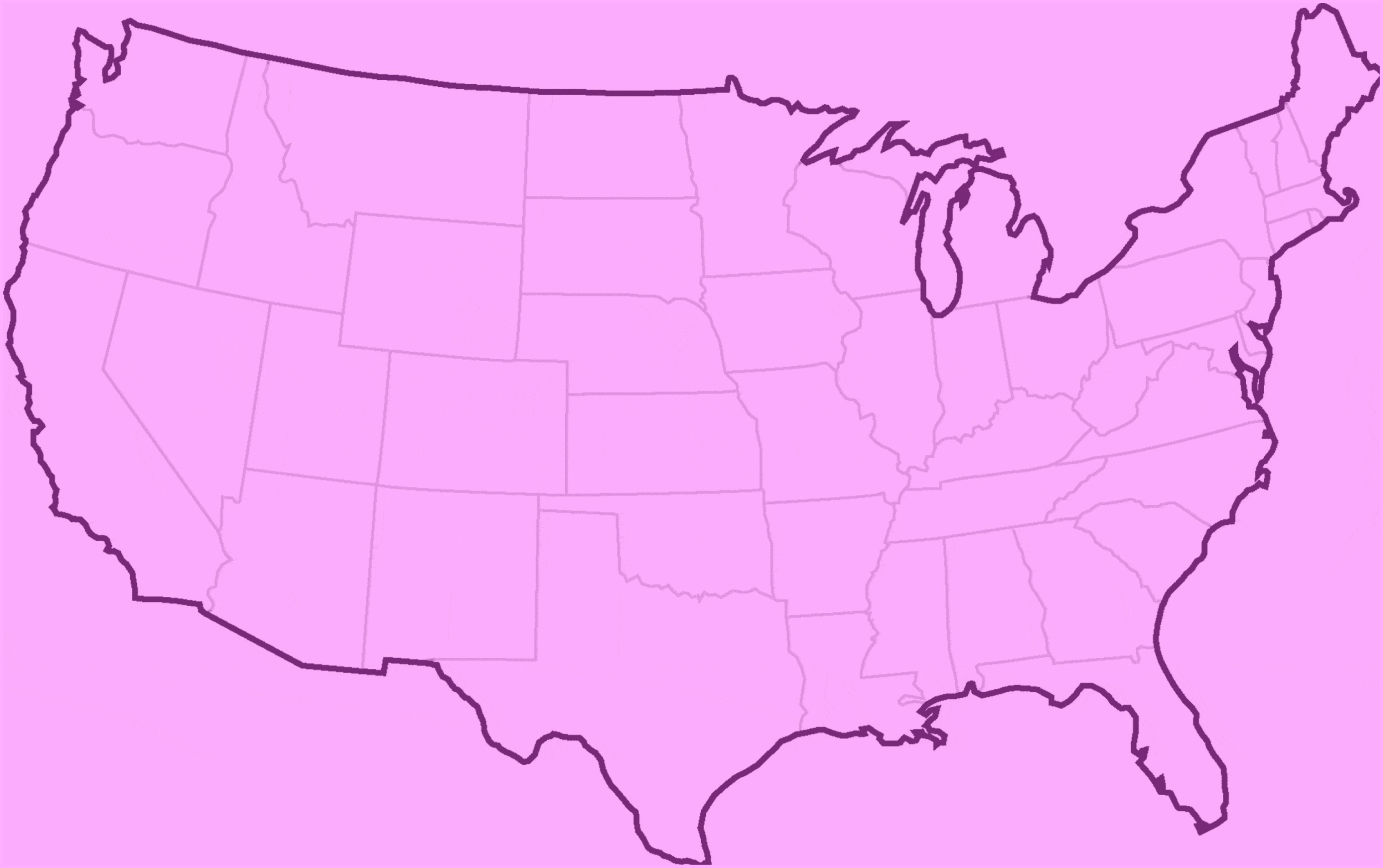
# Case B: Uncertainty is "medium"

Correction

$$bel(x_t) = \eta P(z_t|x_t)\overline{bel}(x_t)$$



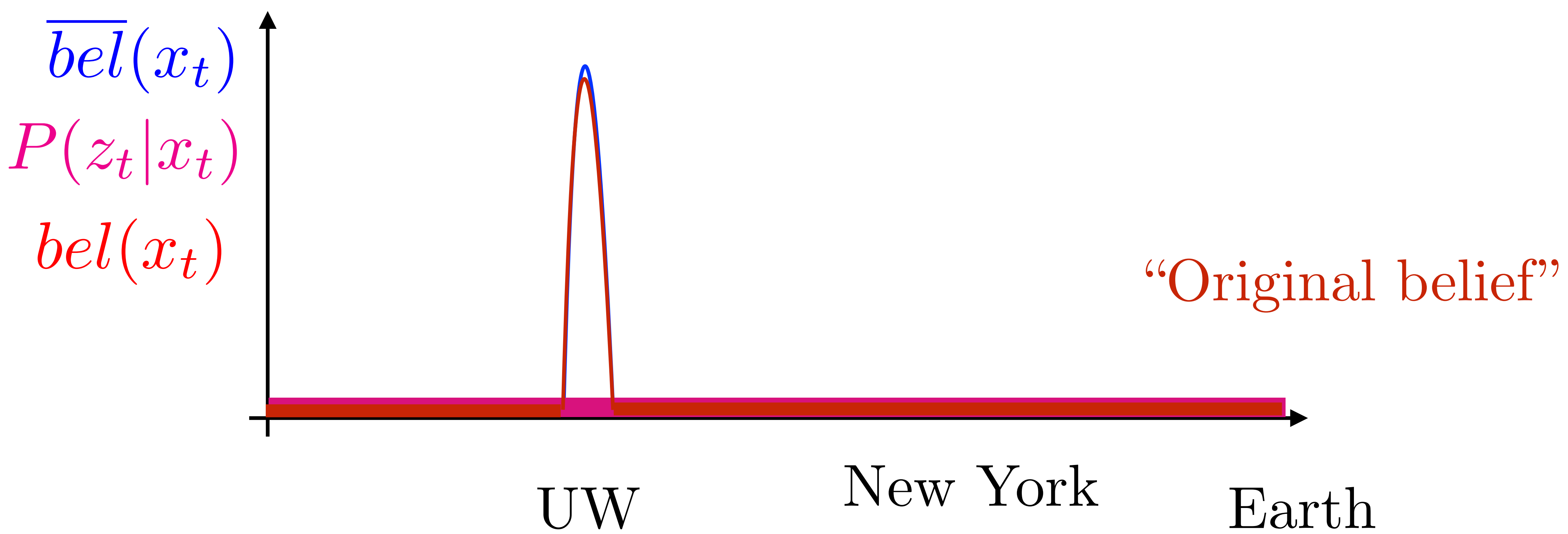
# Case C: Uncertainty is “large”



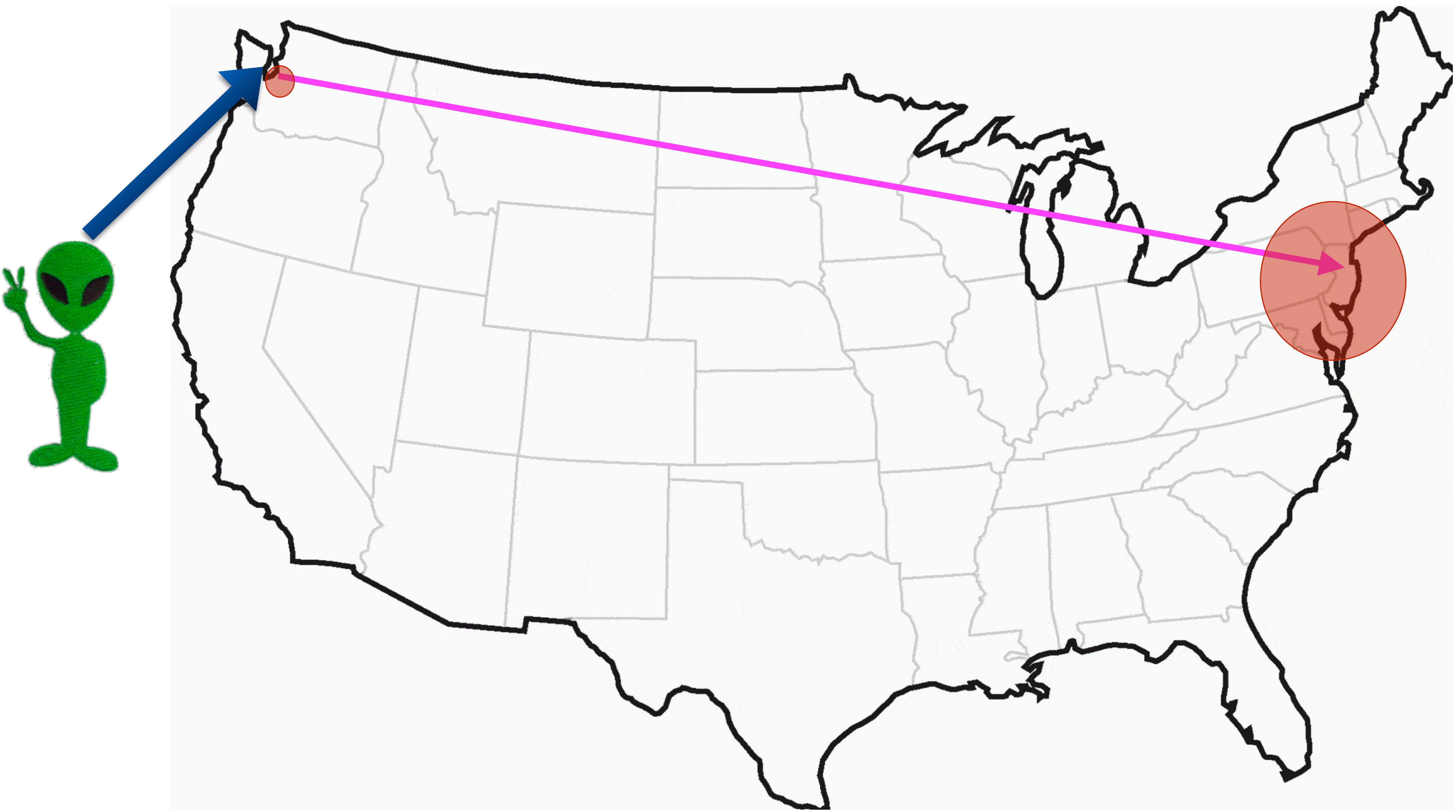
# Case C: Uncertainty is "large"

Correction

$$bel(x_t) = \eta P(z_t|x_t)\overline{bel}(x_t)$$



# Recap of the scenario



# What should we set as our new belief?

If we were to do Bayes filtering in our head ...

| **Measurement Uncertainty** | **Updated belief** |
|---|---|
| Small | |
| Medium | |
| Large | |

# The Kalman Filter

(Bayes filter with
Gaussian beliefs and linear models)

# A bit of history...


**Rudolf Emil Kálmán**
**1960**


**Peter Swerling**
**1959**



[1] R. Kalman, "A new approach to Linear Filtering and Prediction Problems", Journal of Basic Engineering. 82: 35–45.

[2] P. Swerling, "First-Order Error Propagation in a Stagewise Smoothing Procedure for Satellite Observations", Research Memoranda. RM-2329.

# 1-D Kalman Filtering

Belief is a Gaussian

$$bel(x_t) = P(x_t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{(x_t - \mu_t)^2}{2\sigma_t^2}}$$

$$= \mathcal{N}(\mu_t, \sigma_t^2)$$

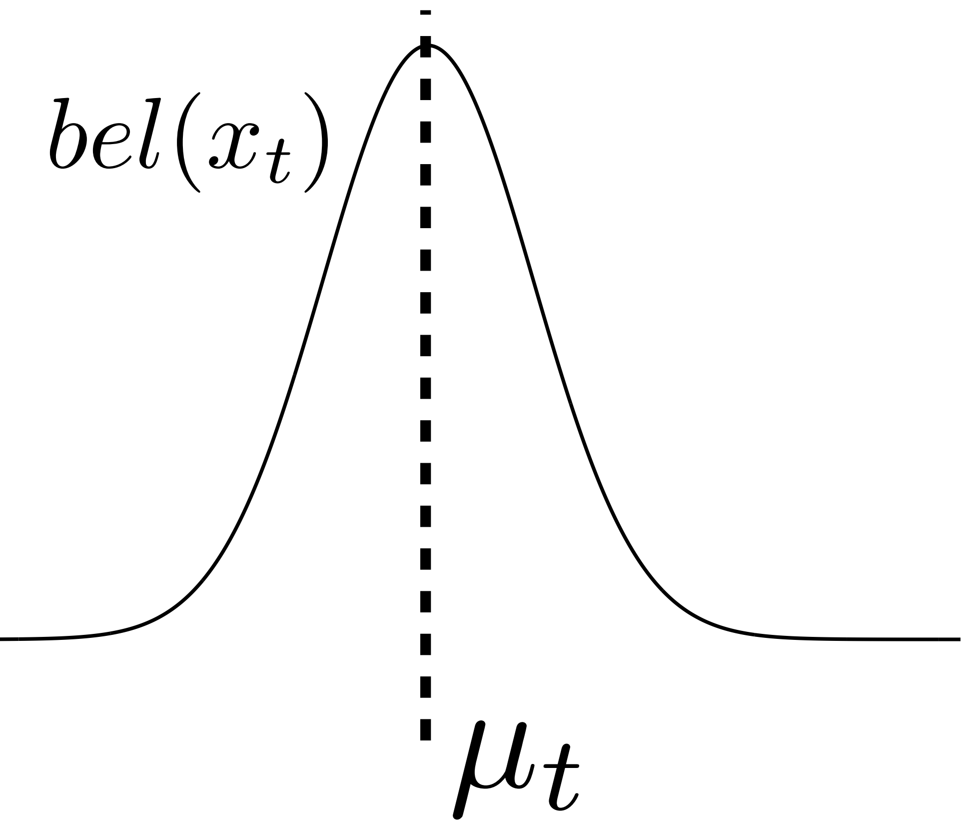
Motion model is linear with Gaussian noise

$$x_{t+1} = x_t + u_{t+1} + \mathcal{N}(0, \sigma_u^2)$$

Observation model is linear with Gaussian noise

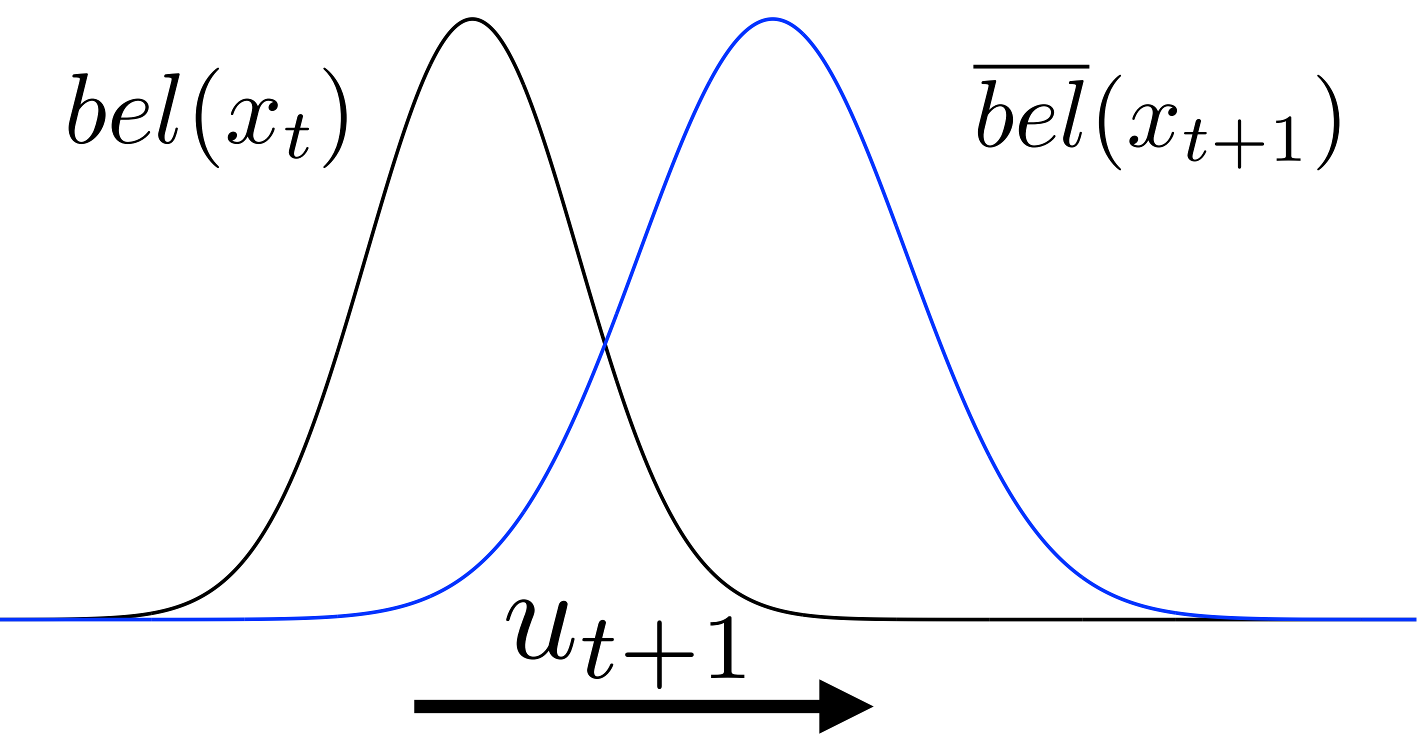
$$z_{t+1} = x_{t+1} + \mathcal{N}(0, \sigma_z^2)$$

# Step 0: Start with belief at time t

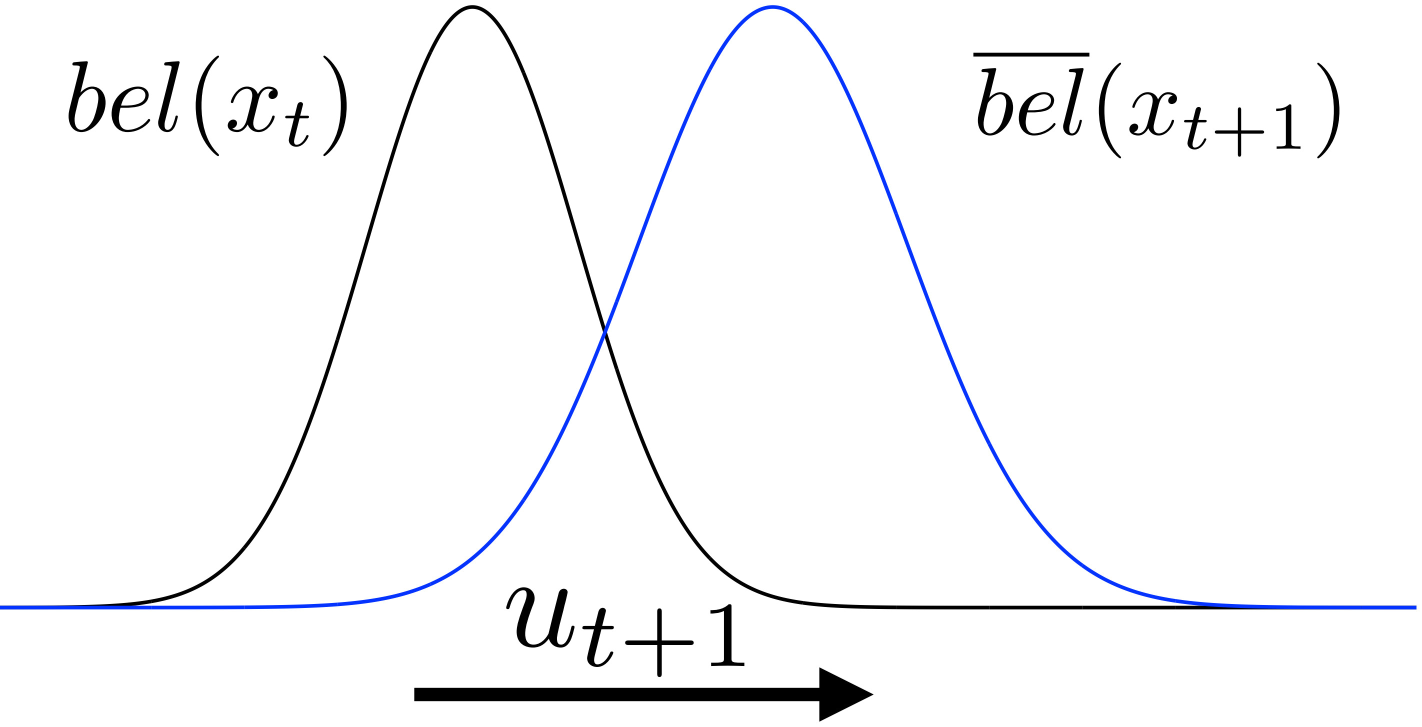



$$bel(x_t) = \mathcal{N}(\mu_t, \sigma_t^2)$$

# Execute control action



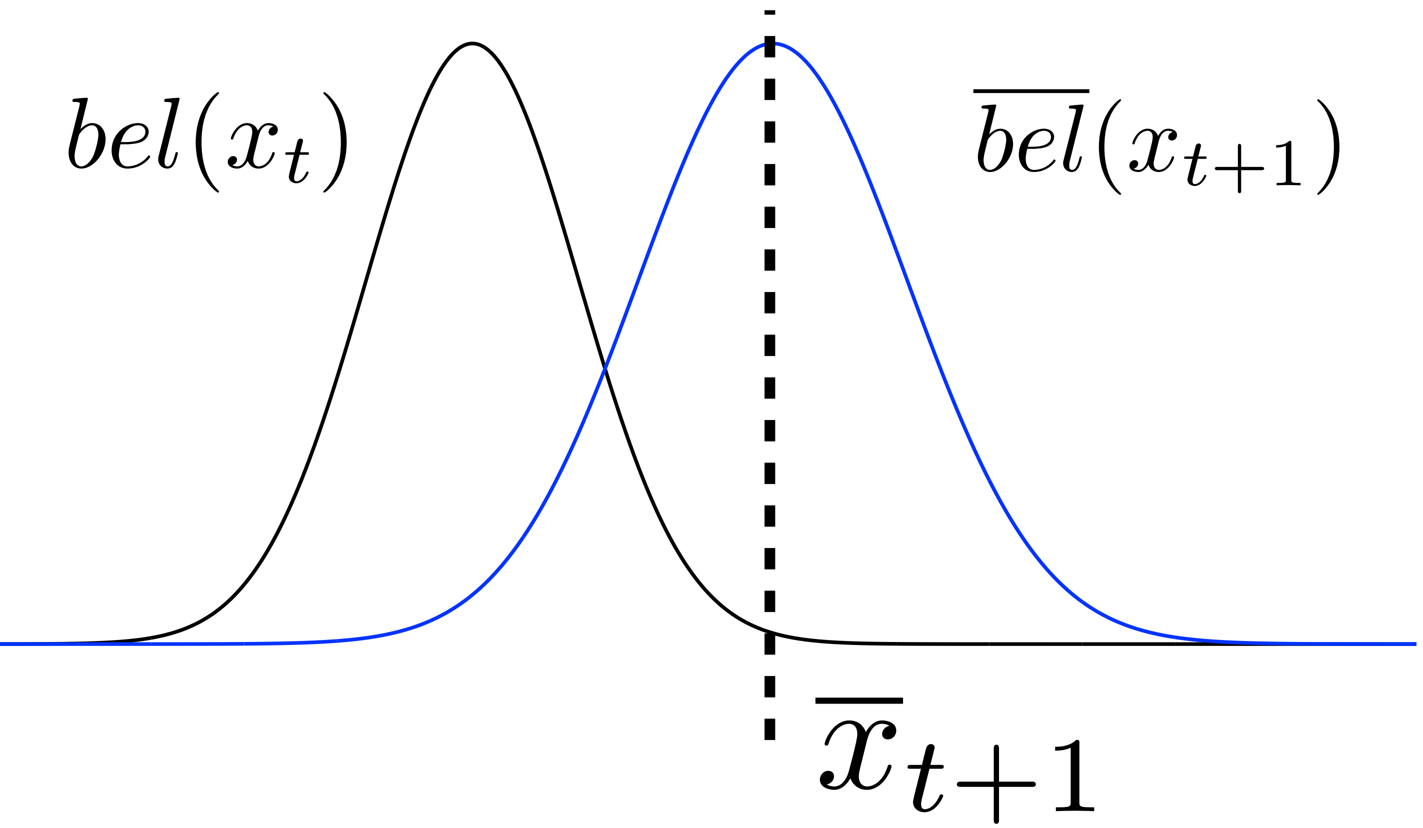
# Step 1: Prediction




$$\overline{bel}(x_{t+1}) = \int_{-\infty}^{\infty} P(x_{t+1}|x_t, {\color{blue}u_{t+1}})\ bel(x_t)\ dx_t$$

Gaussian Gaussian Gaussian

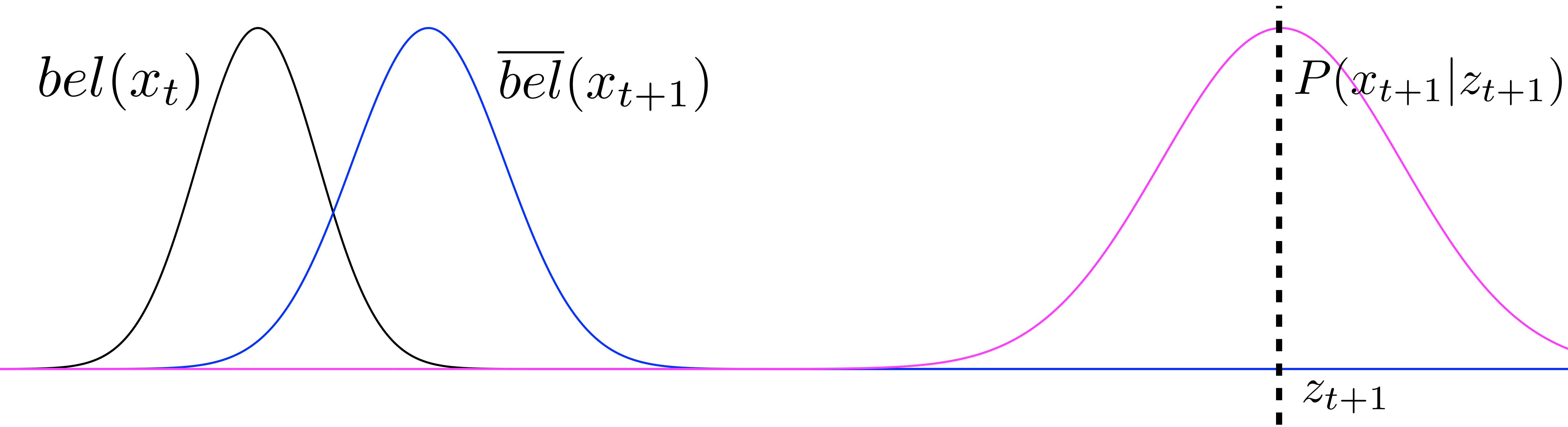
# Step 1: Prediction



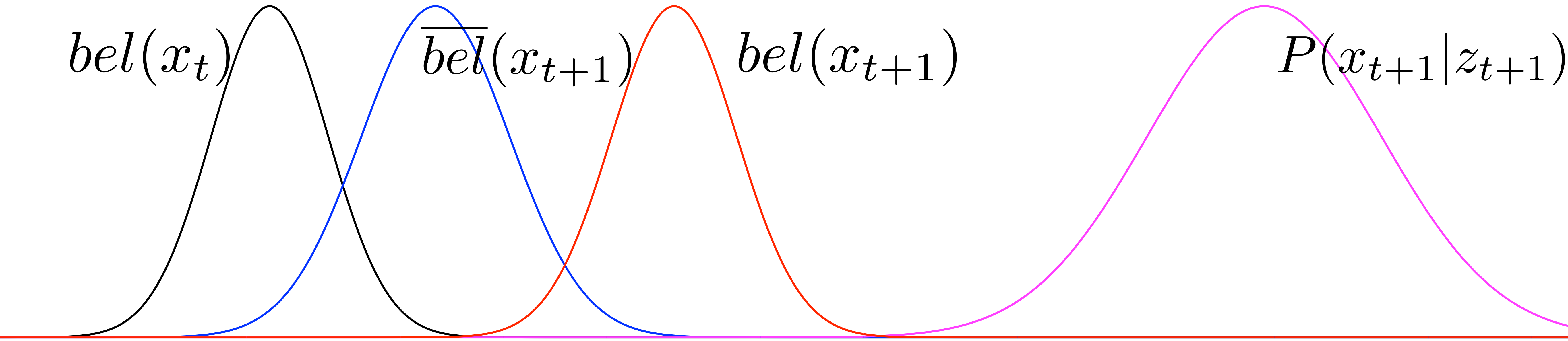

$$\overline{bel}(x_{t+1})$$

$$= \mathcal{N}(\overline{x}_{t+1}, \overline{\sigma}^2_{t+1})$$

# Receive a measurement



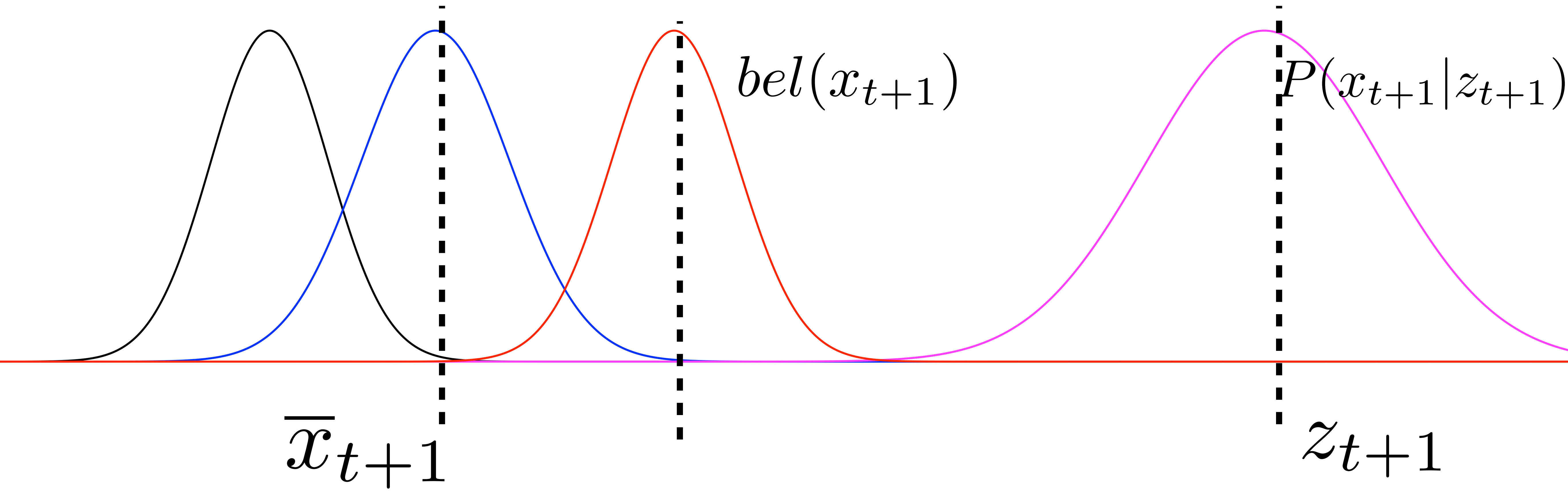
# Step 2: Correction




$$bel(x_{t+1}) = \eta \, P(z_{t+1}|x_{t+1}) \, \overline{bel}(x_{t+1})$$

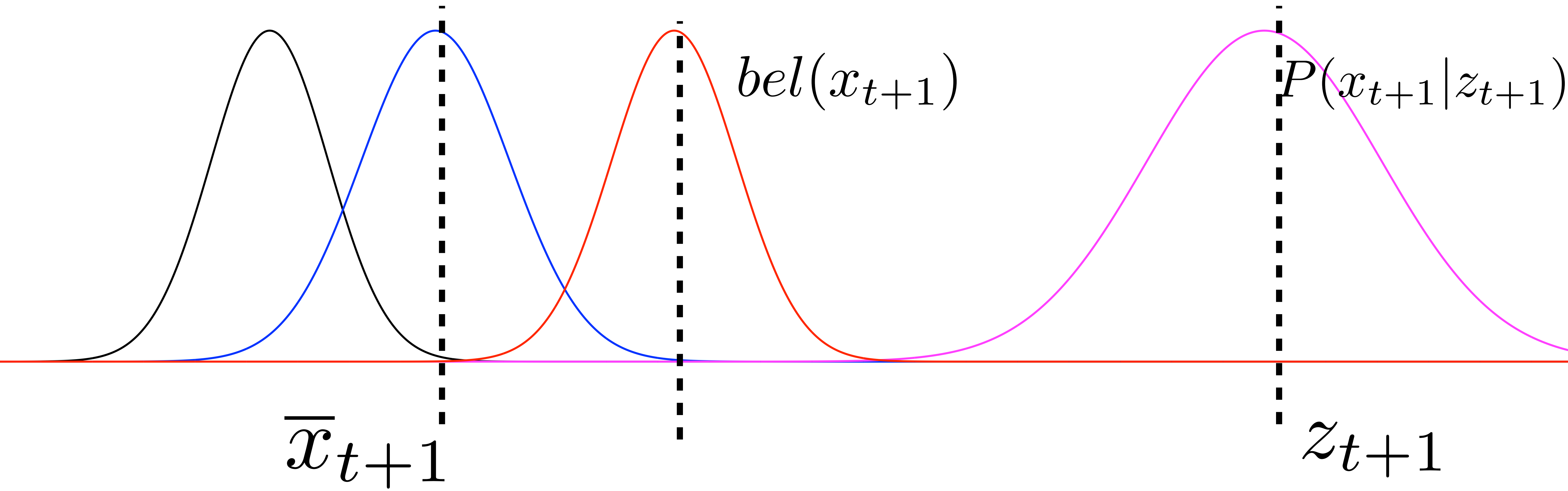
Gaussian Gaussian Gaussian

# Updated belief also a Gaussian!



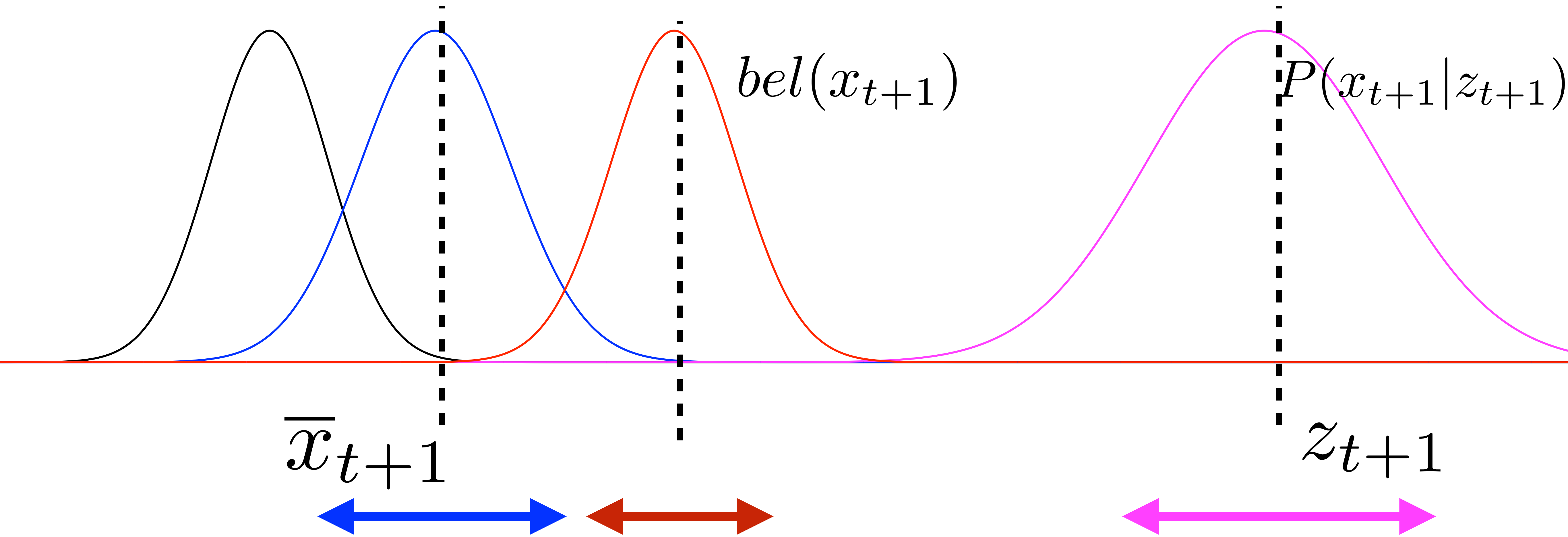

$$bel(x_{t+1}) = \mathcal{N}\left(\frac{\frac{1}{\overline{\sigma}^2_{t+1}}\overline{x}_{t+1} + \frac{1}{\sigma^2_z} z_{t+1}}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma^2_z}}, \frac{1}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma^2_z}}\right)$$

# Linearly interpolate prediction and measurement




$$bel(x_{t+1}) = \mathcal{N}\left(\frac{\frac{1}{\overline{\sigma}^2_{t+1}}\overline{x}_{t+1} + \frac{1}{\sigma_z^2}z_{t+1}}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma_z^2}}, \frac{1}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma_z^2}}\right)$$

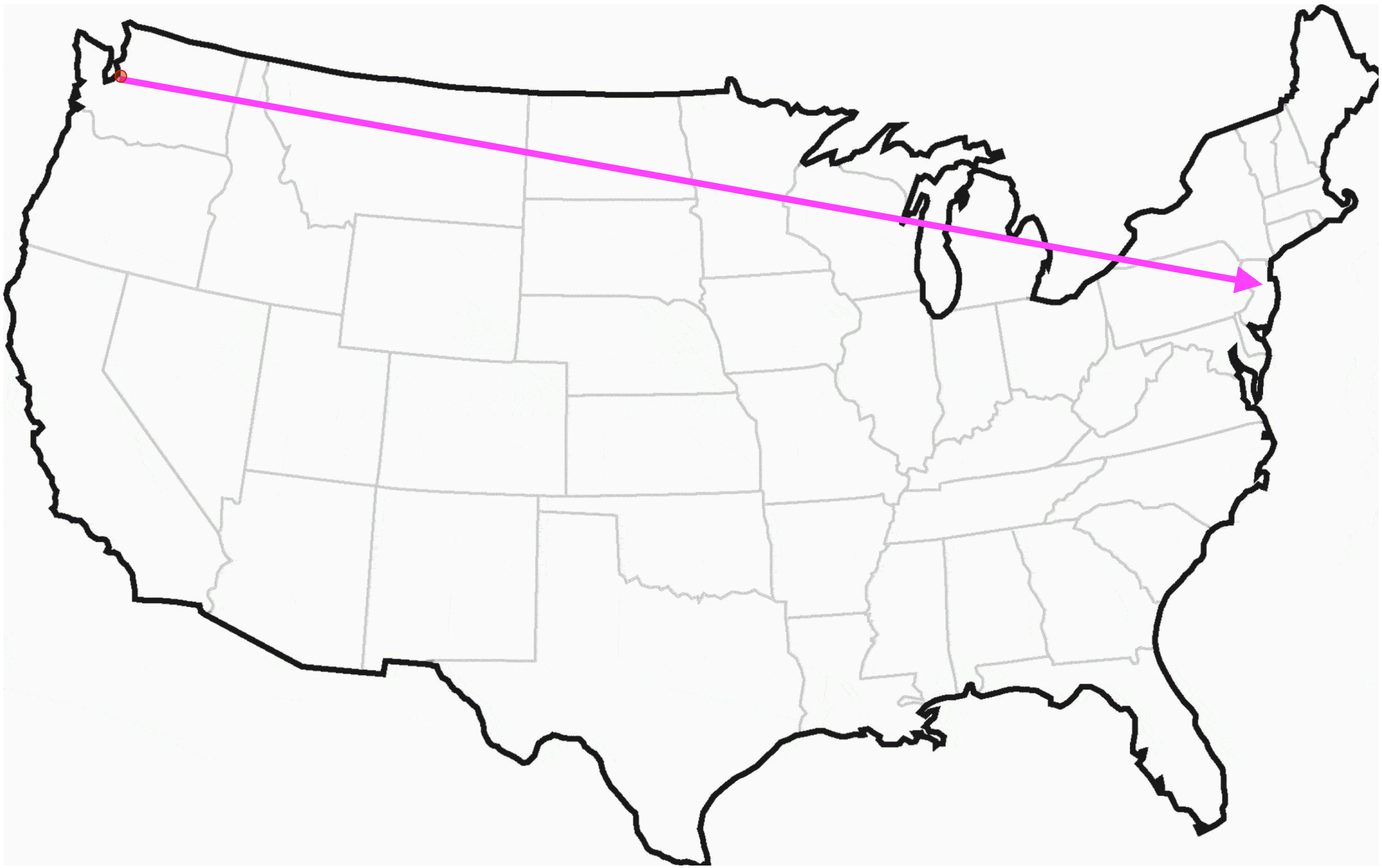
# Problem: Variance ALWAYS decreases!




$$bel(x_{t+1}) = \mathcal{N}\left( \frac{\frac{1}{\overline{\sigma}^2_{t+1}}\overline{x}_t + \frac{1}{\sigma_z^2} z_t}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma_z^2}}, \frac{1}{\frac{1}{\overline{\sigma}^2_{t+1}} + \frac{1}{\sigma_z^2}} \right)$$

... no matter what the measurement values are!

# Back to example ...



# What should we set as our new belief?

| **Measurement Uncertainty** | **Our reasonable guess** | **Kalman Filter** |
|---|---|---|
| Small | Anywhere on earth | midpoint; small uncertainty |
| Medium | Anywhere on earth | close to UW; small uncertainty |
| Large | UW; 500m | UW; 500m |

# What is broken ?!?

Is the
linear model
broken?

Is the
Gaussian
assumption
broken?

Is the
Bayes
filtering
broken?

# Problem: Overconfidence

- KF works best when $\overline{\sigma}_{t+1}, \sigma_z$ comparable.
- $\overline{\sigma}_{t+1}$ may become unrealistically low (overconfidence) by:
  - taking long time steps
  - accumulating incomplete/noisy measurements
- Gaussian update "ignores" measurements
- Fix: inflate variance of model uncertainty, e.g. add noise!

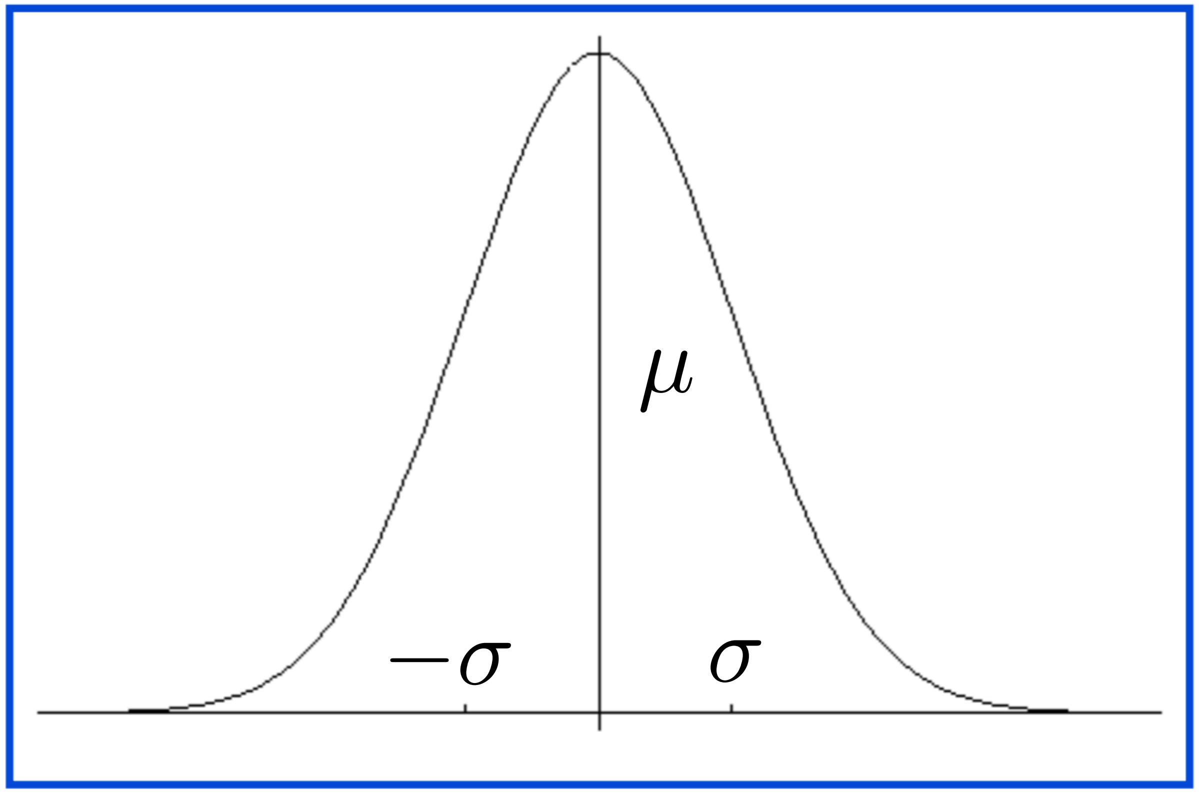
# Going Deeper on Kalman Filters

# Aside: Gaussians

Univariate

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

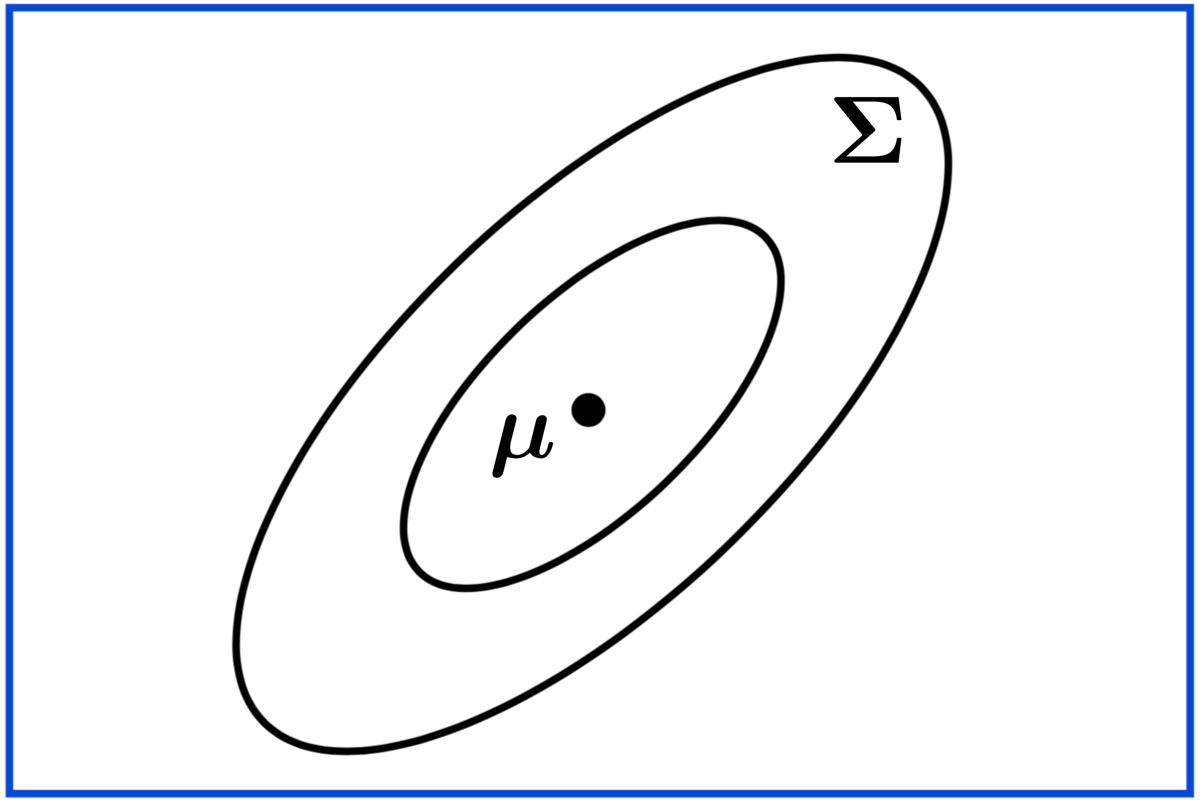
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$




Multivariate

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$



# Aside: Gaussians have nice properties

$$\left.\begin{array}{l} x \sim \mathcal{N}(\mu, \sigma^2) \\ y = ax + b \end{array}\right\} \implies y \sim \mathcal{N}\left(a\mu + b, a^2\sigma^2\right)$$

$$\left.\begin{array}{l} x_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \\ x_2 \sim \mathcal{N}(\mu_2, \sigma_2^2) \end{array}\right\} \implies p(x_1)p(x_2) \sim \mathcal{N}\left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\mu_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\mu_2, \frac{1}{\sigma_1^{-2} + \sigma_2^{-2}}\right)$$

# Aside: Gaussians have nice properties

$$\left.\begin{array}{l}\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ \mathbf{y} = A\mathbf{x} + B\end{array}\right\} \Longrightarrow \mathbf{y} \sim \mathcal{N}\left(A\boldsymbol{\mu} + B, A\boldsymbol{\Sigma}A^{\mathsf{T}}\right)$$

$$\left.\begin{array}{l}\mathbf{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \\ \mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)\end{array}\right\} \Longrightarrow p(\mathbf{x}_1)p(\mathbf{x}_2) \sim \mathcal{N}\left(\frac{\boldsymbol{\Sigma}_2}{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}\boldsymbol{\mu}_1 + \frac{\boldsymbol{\Sigma}_1}{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}\boldsymbol{\mu}_2, \frac{1}{\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}}\right)$$

*As long as we start from a* ***Gaussian*** *and perform only* ***linear*** *transformations, we remain in the Gaussian world.*

# Kalman Filter: Motion & Sensor Models

m-vector

nxn nxm

n-vector

Covariance nxn

Dynamics

$$\mathbf{x}_t = A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t + \epsilon_t$$

$$\boldsymbol{\epsilon}_t \sim \mathcal{N}\left(\mathbf{0}, R_t\right)$$

Measurement

k-vector

Covariance kxk

n-vector

$$\mathbf{z}_t = C_t \mathbf{x}_t + \boldsymbol{\delta}_t$$

$$\boldsymbol{\delta}_t \sim \mathcal{N}\left(\mathbf{0}, Q_t\right)$$

k-vector

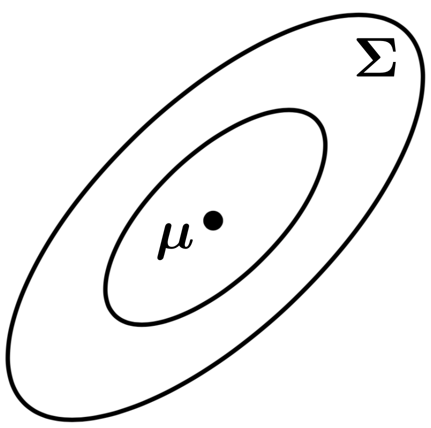
Discrete-time Linear Systems with Gaussian noise

kxn

# Kalman Filter Assumptions

1. Initial belief is a gaussian distribution

$$\mathbf{x}_t \sim \mathcal{N}\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$$



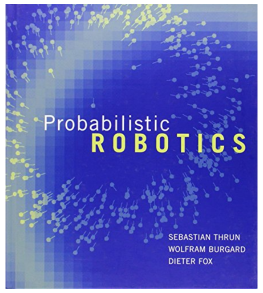

2. Linear Dynamics

$$\mathbf{x}_t = A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t + \epsilon_t$$

3. Linear Measurement Model

$$\mathbf{z}_t = C_t \mathbf{x}_t + \boldsymbol{\delta}_t$$

# The Kalman Filter Algorithm




Sec. 3.2.4

1. Algorithm **Kalman_filter**($\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$):

2. Prediction:
3. $\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$
4. $\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$

5. Correction:
6. $K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$
7. $\mu_t = \bar{\mu}_t + K_t(\underline{z_t - C_t \bar{\mu}_t})$
8. $\Sigma_t = (I - K_t C_t)\bar{\Sigma}_t$

9. Return $\mu_t, \Sigma_t$

Kalman Gain: degree at which observation factors into belief

"Innovation"

*Slide by Dieter Fox

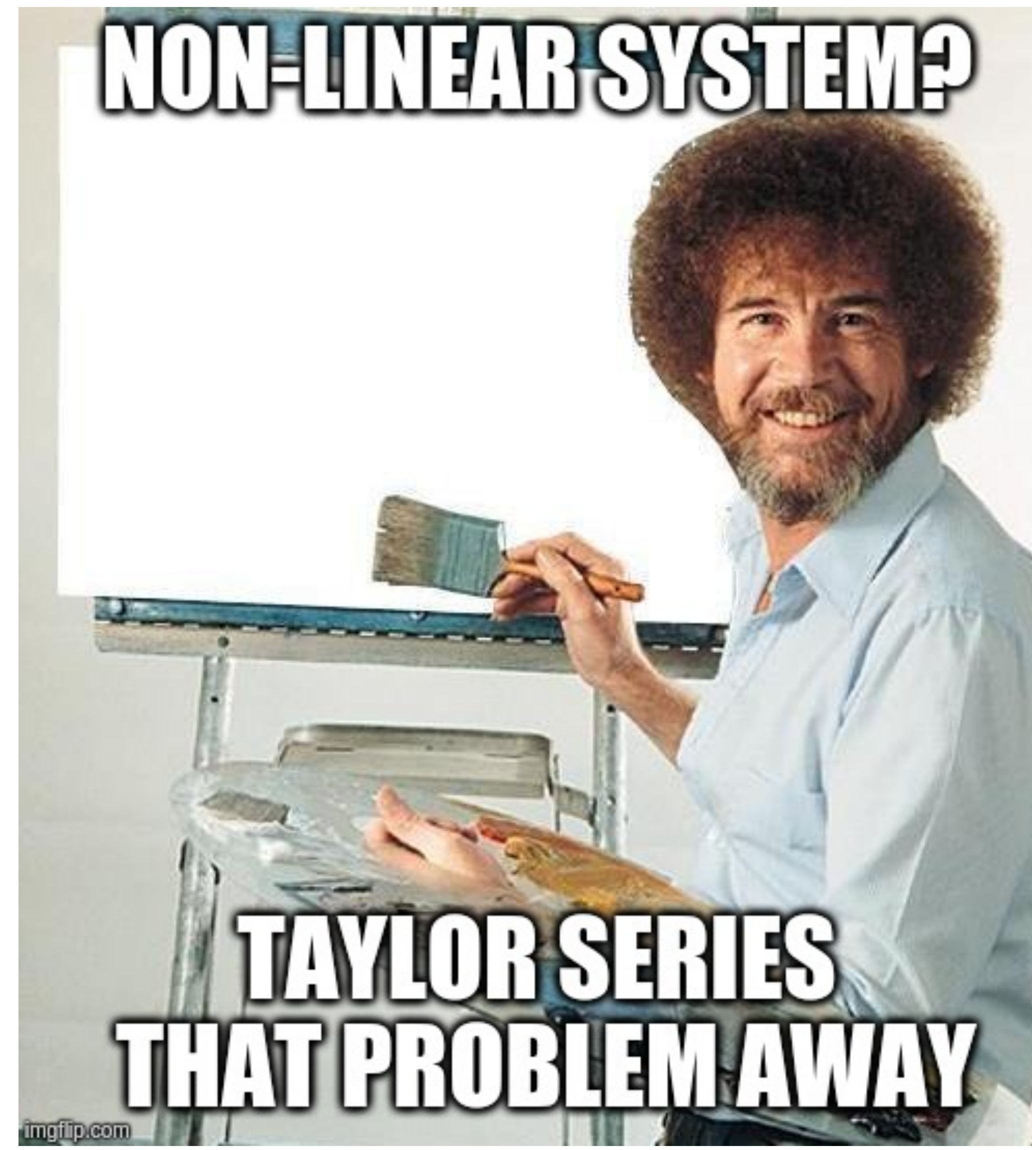
# Summary

- Highly efficient: $O(k^{2.376} + n^2)$
- Optimal for linear Gaussian systems (minimizes variance)
- Requires linear motion and observation model
- Overconfidence

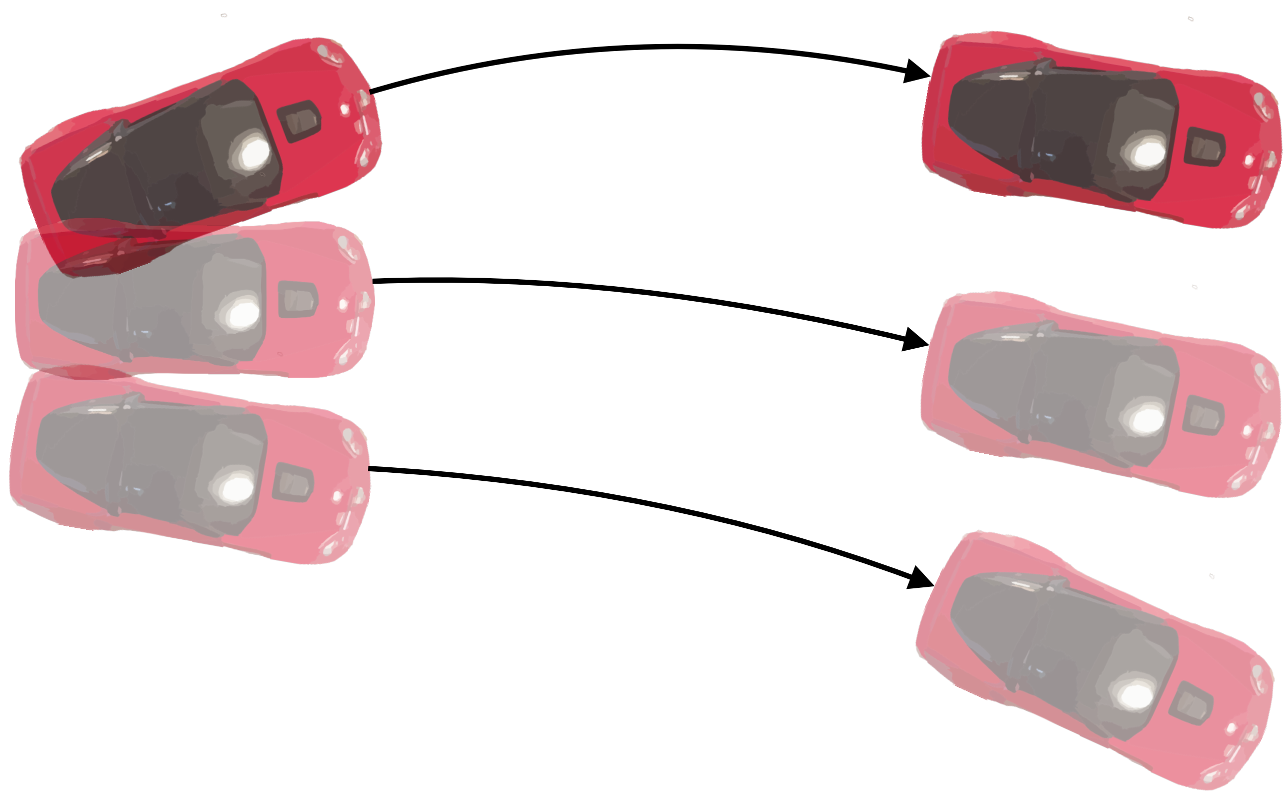
# Plot Twist: Most Robotic Systems are Nonlinear...

**Extended Kalman Filter (EKF)**

- Linearize Motion/Sensor models
- 1st order Taylor Series expansion
- Sec. 3.3 of *Probabilistic Robotics*



# Coming up next...



Particle Filters