

ASEN 5070  
Statistical Orbit Determination I  
Fall 2012



Professor George H. Born  
Professor Jeffrey S. Parker

Lecture 12: The Kalman Filter



University of Colorado  
Boulder

# Announcements

---

- ▶ Homework 5 due Today
  - ▶ Exam on 10/11. (Anyone going to miss it?)
    - Eduardo and/or Paul will be reviewing subjects on Tuesday – send them emails with questions/subjects that you'd like them to cover.
    - 1 hour, open book, open notes.
    - Topics: Definitions of variables, Probability/Statistics Observability, Linearization Least squares, Batch processor
- |                          |                      |                         |
|--------------------------|----------------------|-------------------------|
| $\mathbf{X}(t_i),$       | $\mathbf{X}^*(t_i),$ | $\hat{\mathbf{X}}(t_i)$ |
| $\mathbf{x}(t_i),$       | $\mathbf{x}^*(t_i),$ | $\hat{\mathbf{x}}(t_i)$ |
| $\mathbf{Y}(t_i),$       | $\epsilon(t_i),$     | $\hat{\epsilon}(t_i)$   |
| $\Phi(t_i, t_j),$        | $H(t_i),$            | $\tilde{H}(t_i)$        |
| $A(t_i),$                | $P(t_i),$            | $W(t_i)$                |
| $\bar{\mathbf{x}}(t_i),$ | $\bar{P}(t_i),$      | $\bar{W}(t_i)$          |



# Quiz Results

---



If you took the quiz, you scored 100%



University of Colorado  
Boulder

# Quiz Results

---

- ▶ Some feedback:
- ▶ Biggest issues:
  - Moving pretty fast
  - Lectures and HW don't correlate well
- ▶ Review next week:
  - Review of variables
  - Stat OD example from start to finish (something plain and easy to follow)

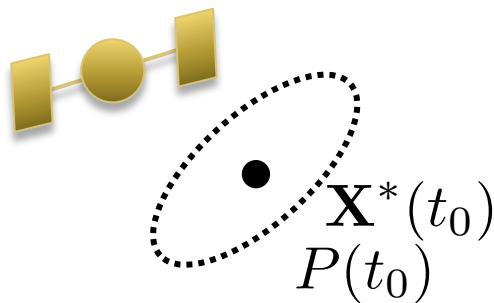


# Review of the Stat OD Process

## ► Setup.

- Given: an initial state  $\mathbf{X}^*(t_0)$
- Optional: an initial covariance  $P(t_0)$

$$\mathbf{X}^*(t_0) = [x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0, c_1, c_2]^T$$



University of Colorado  
Boulder

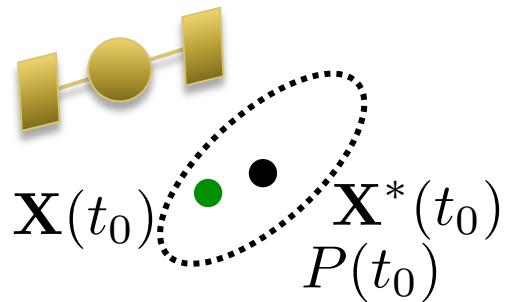
# Review of the Stat OD Process

## ► Setup.

- Given: an initial state  $\mathbf{X}^*(t_0)$
- Optional: an initial covariance  $P(t_0)$

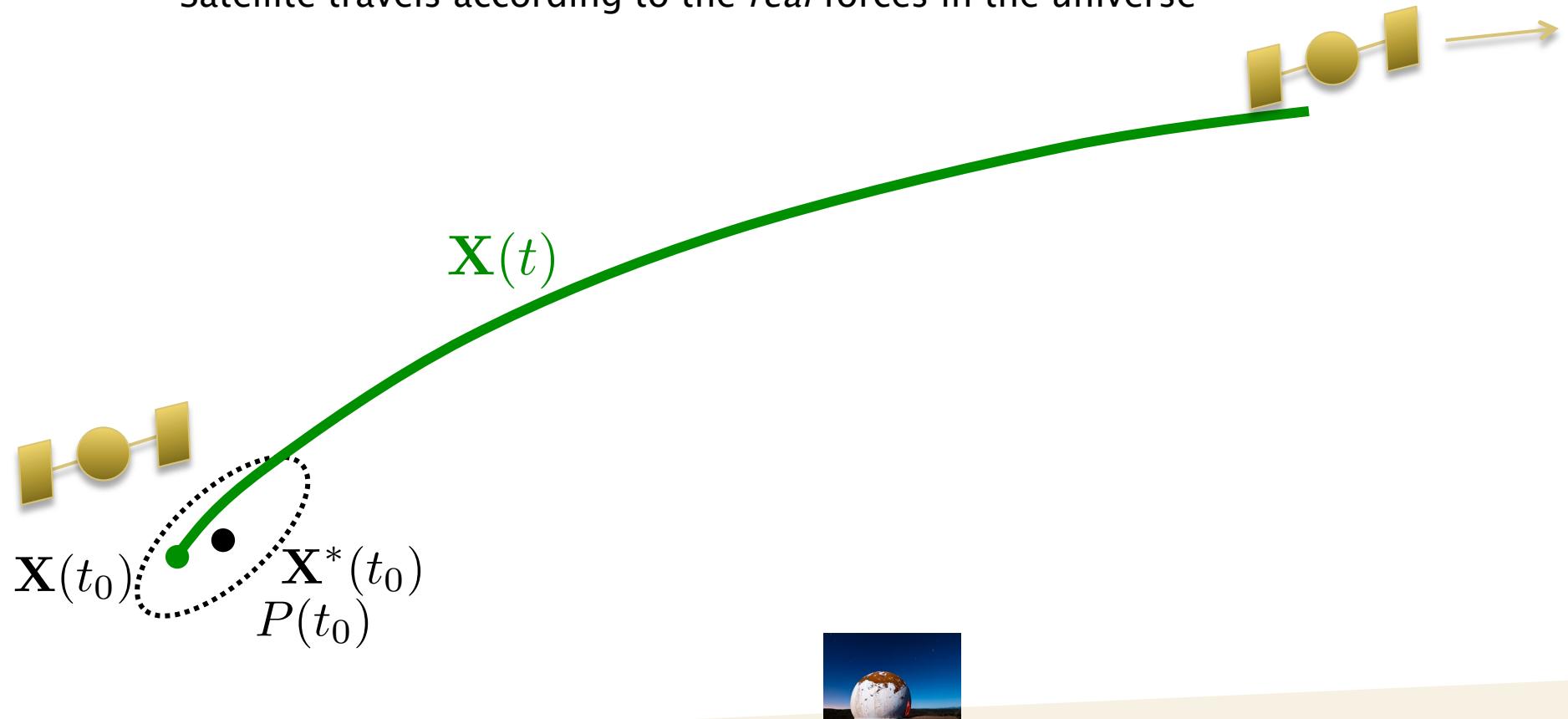
$$\mathbf{X}^*(t_0) = [x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0, c_1, c_2]^T$$

- The satellite will *not* be there, but will (hopefully) be nearby
  - True state =  $\mathbf{X}(t_0)$



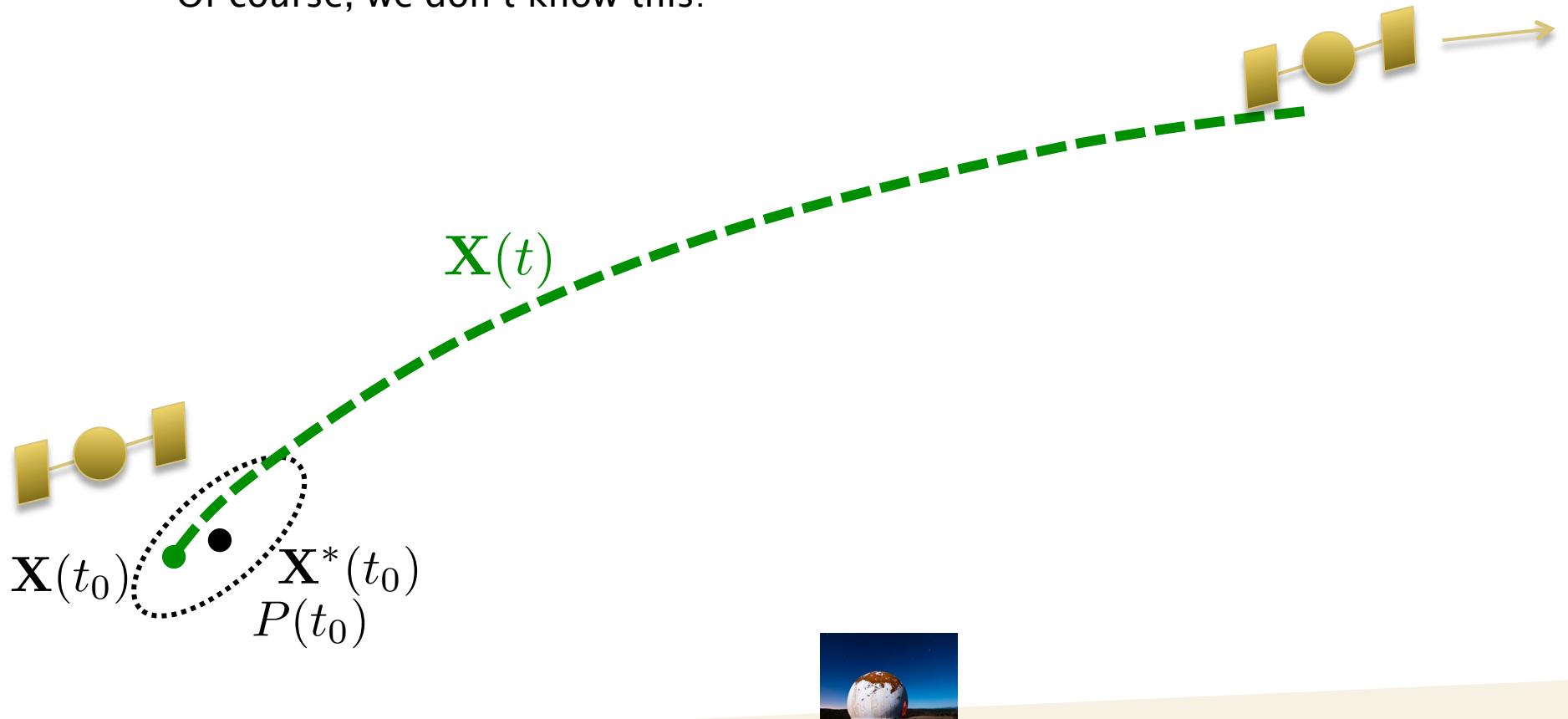
# Review of the Stat OD Process

- ▶ What really happens
  - Satellite travels according to the *real* forces in the universe



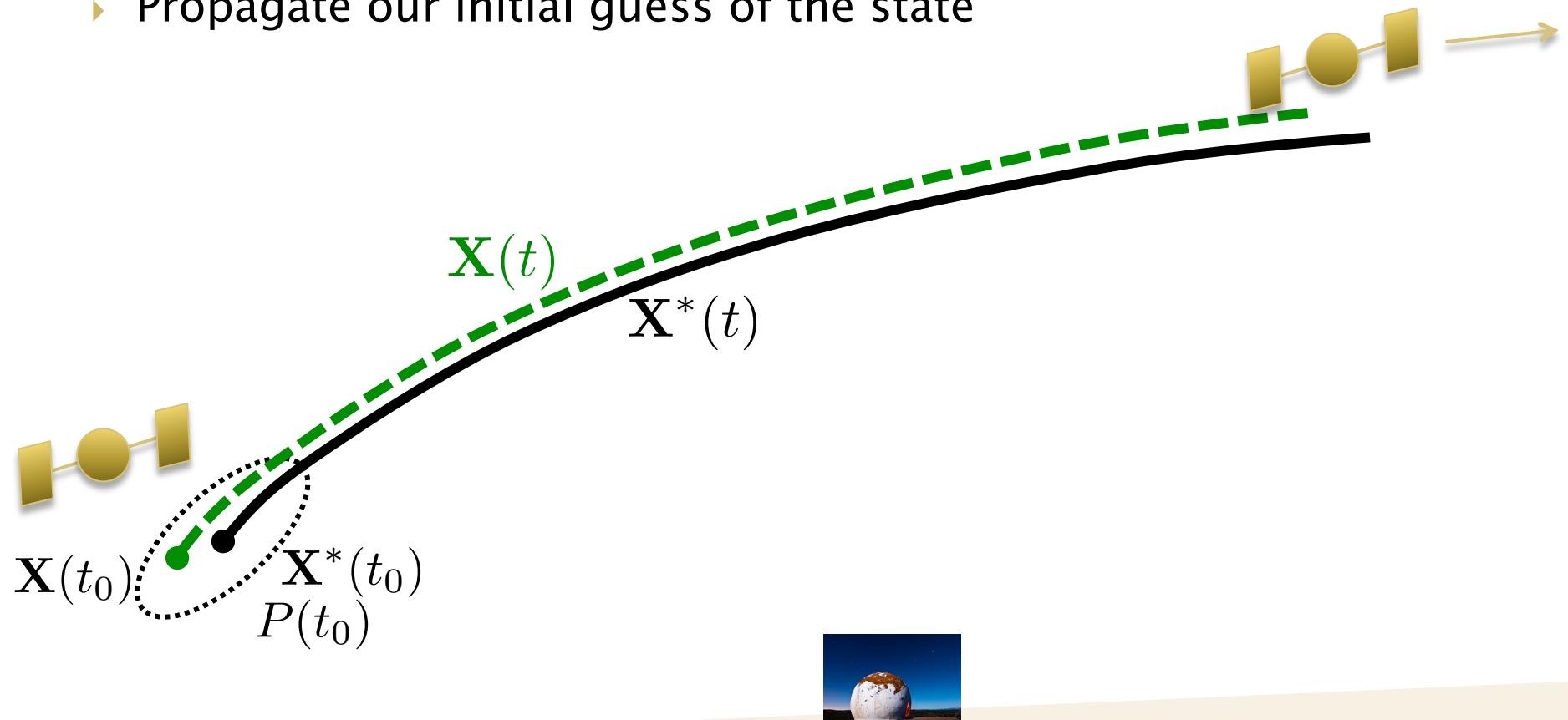
# Review of the Stat OD Process

- ▶ What really happens
  - Of course, we don't know this!



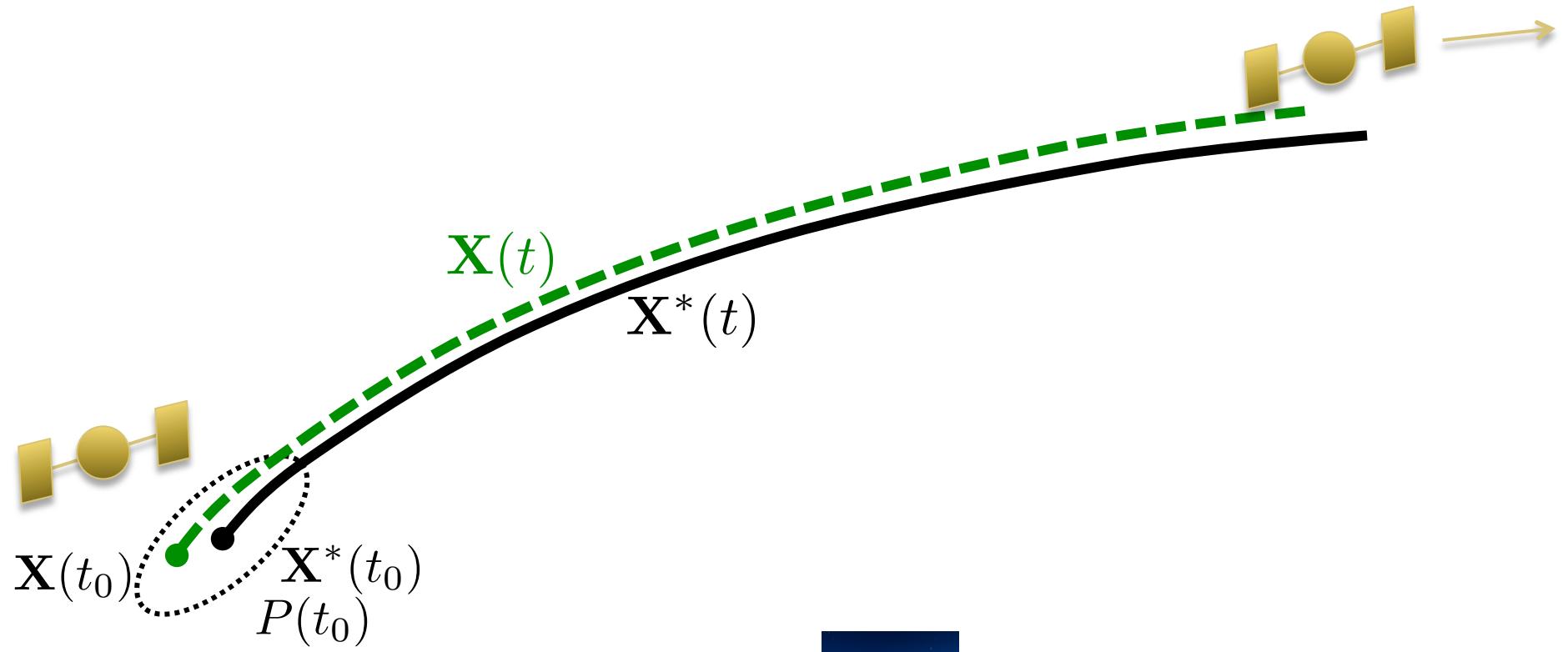
# Review of the Stat OD Process

- ▶ Model reality as best as possible
- ▶ Propagate our initial guess of the state



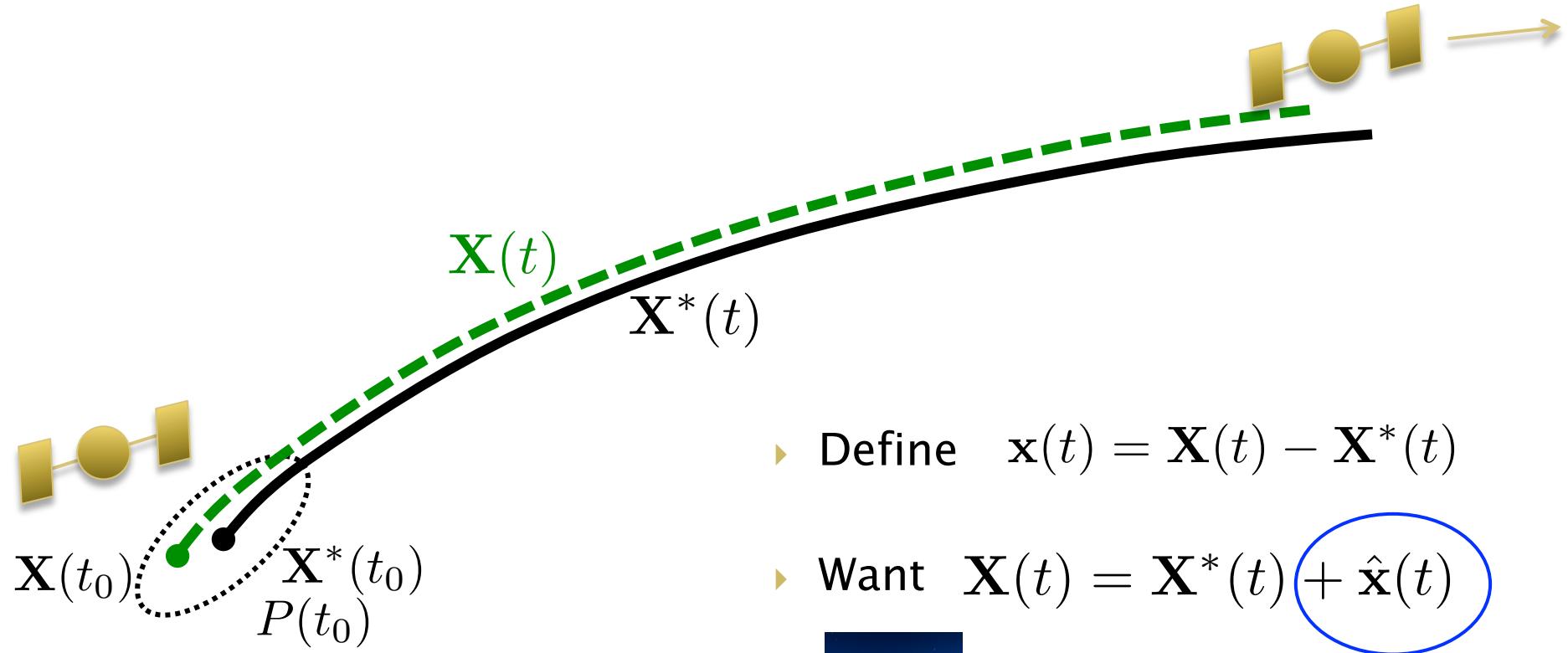
# Review of the Stat OD Process

- ▶ Goal: Determine how to modify  $\mathbf{X}^*(t)$  to match  $\mathbf{X}(t)$



# Review of the Stat OD Process

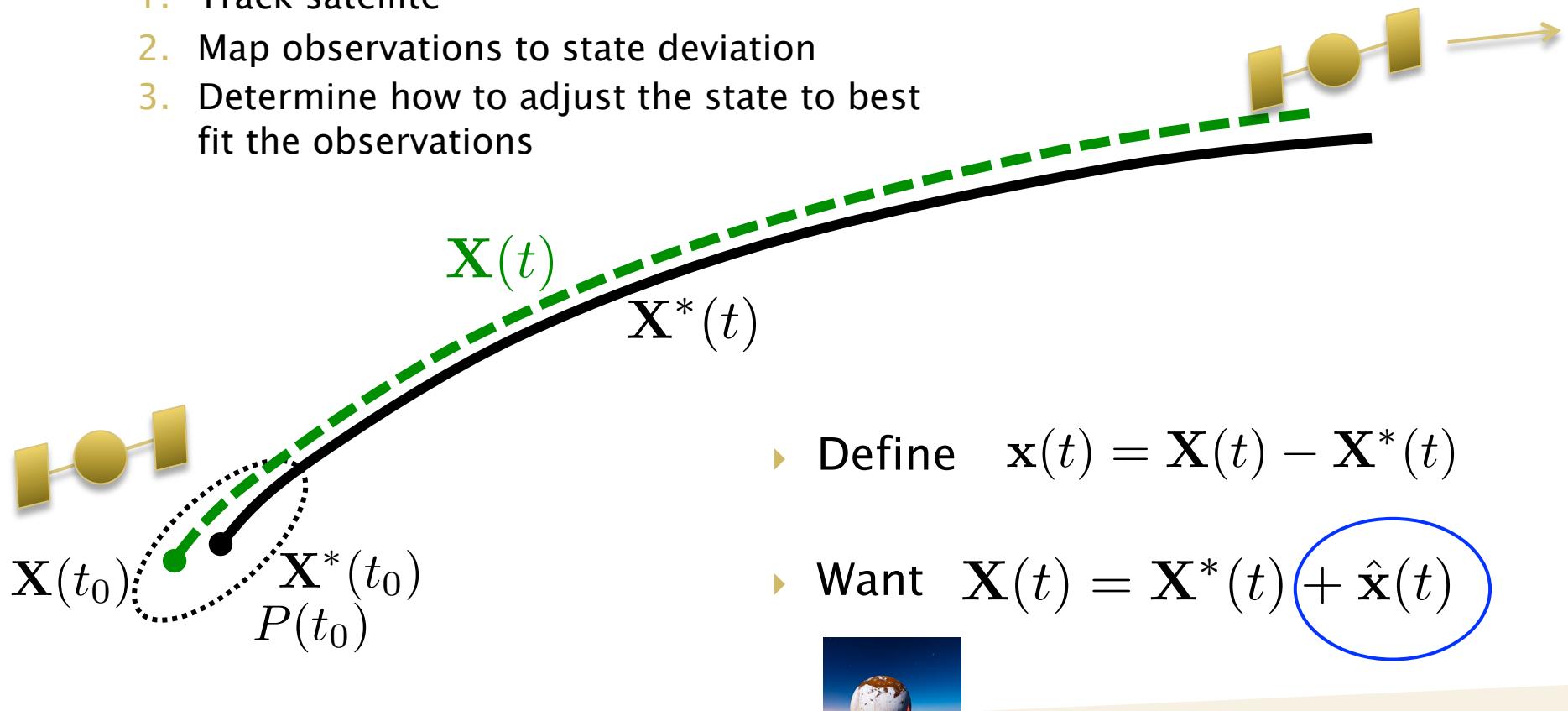
- ▶ Goal: Determine how to modify  $\mathbf{X}^*(t)$  to match  $\mathbf{X}(t)$



# Review of the Stat OD Process

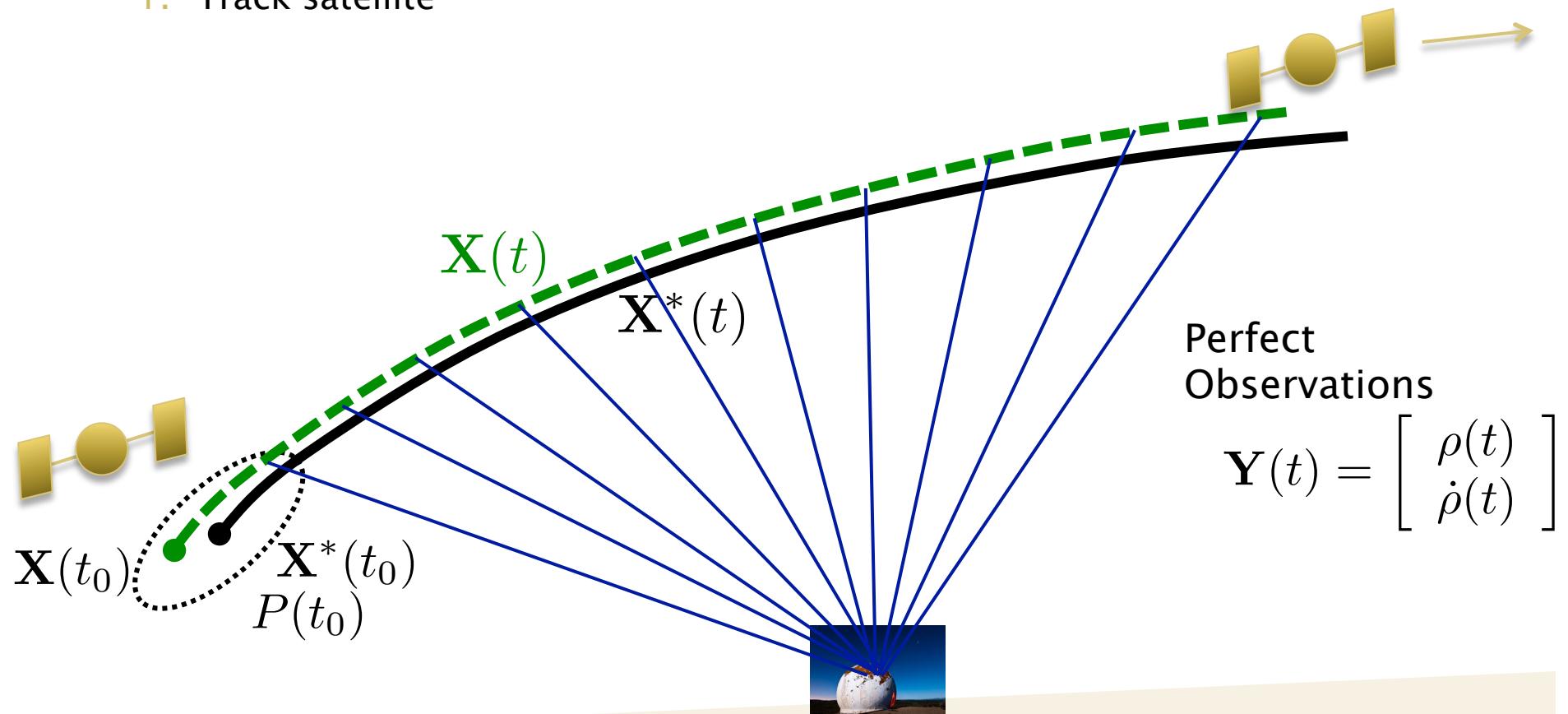
- ▶ **Process:**

1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations



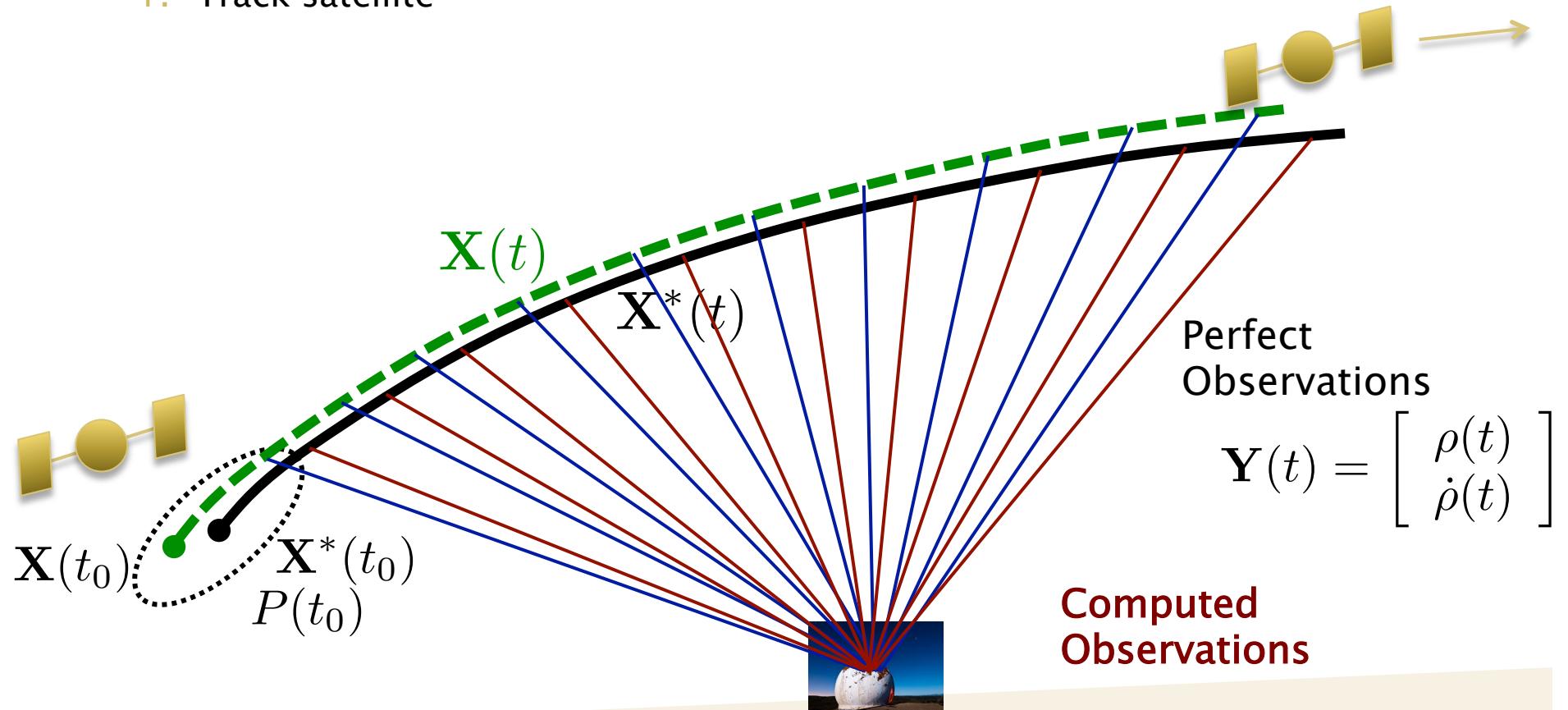
# Review of the Stat OD Process

- ▶ Process:
  1. Track satellite



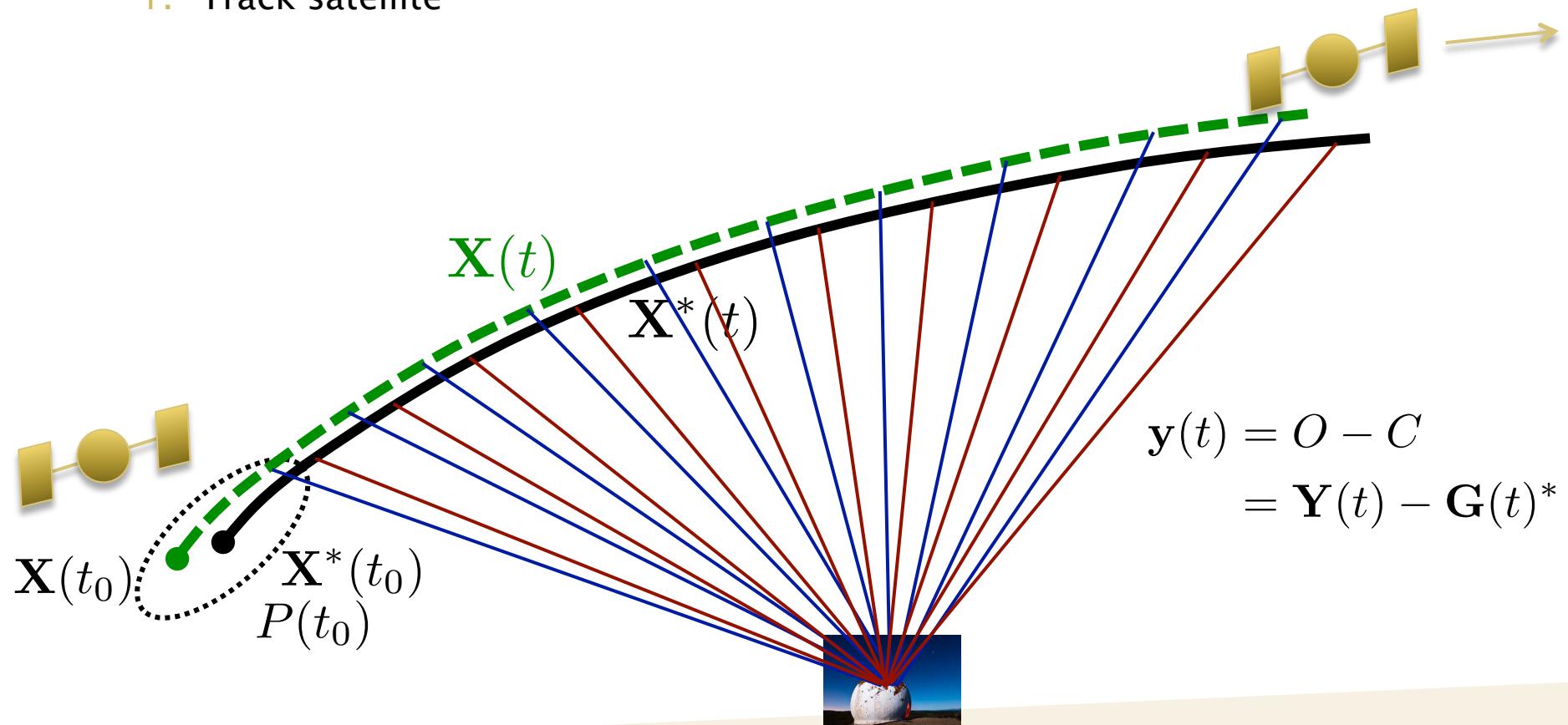
# Review of the Stat OD Process

- ▶ Process:
  1. Track satellite



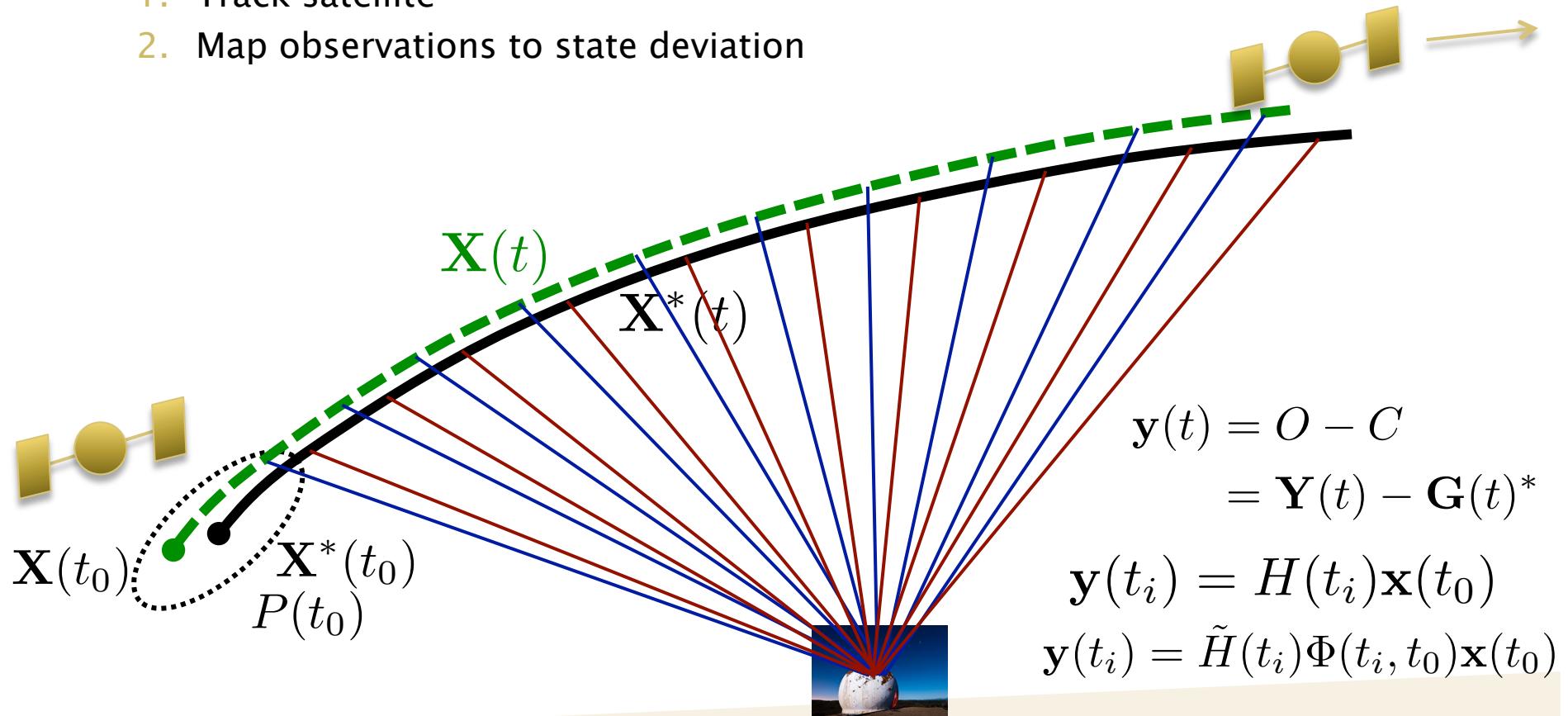
# Review of the Stat OD Process

- ▶ Process:
  1. Track satellite



# Review of the Stat OD Process

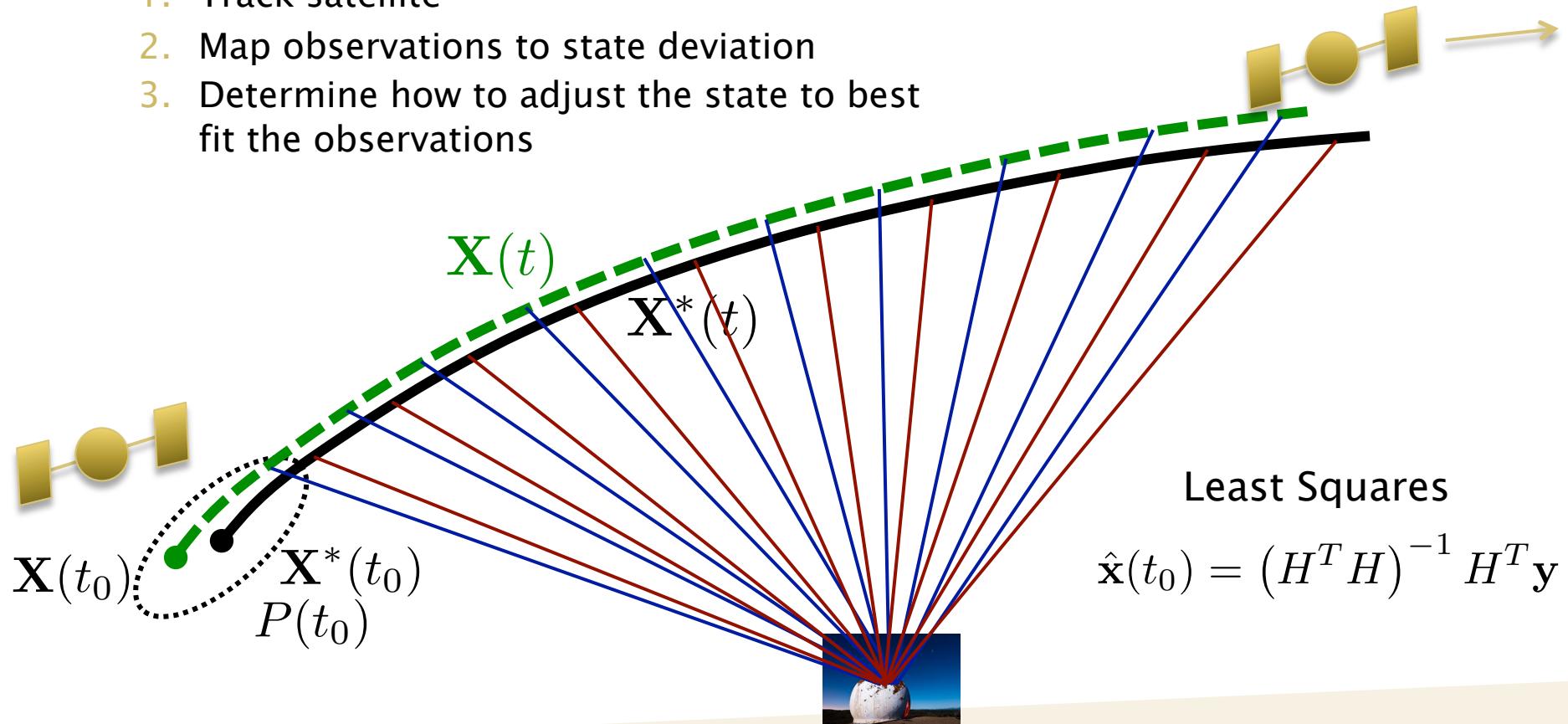
- ▶ Process:
  1. Track satellite
  2. Map observations to state deviation



# Review of the Stat OD Process

## Process:

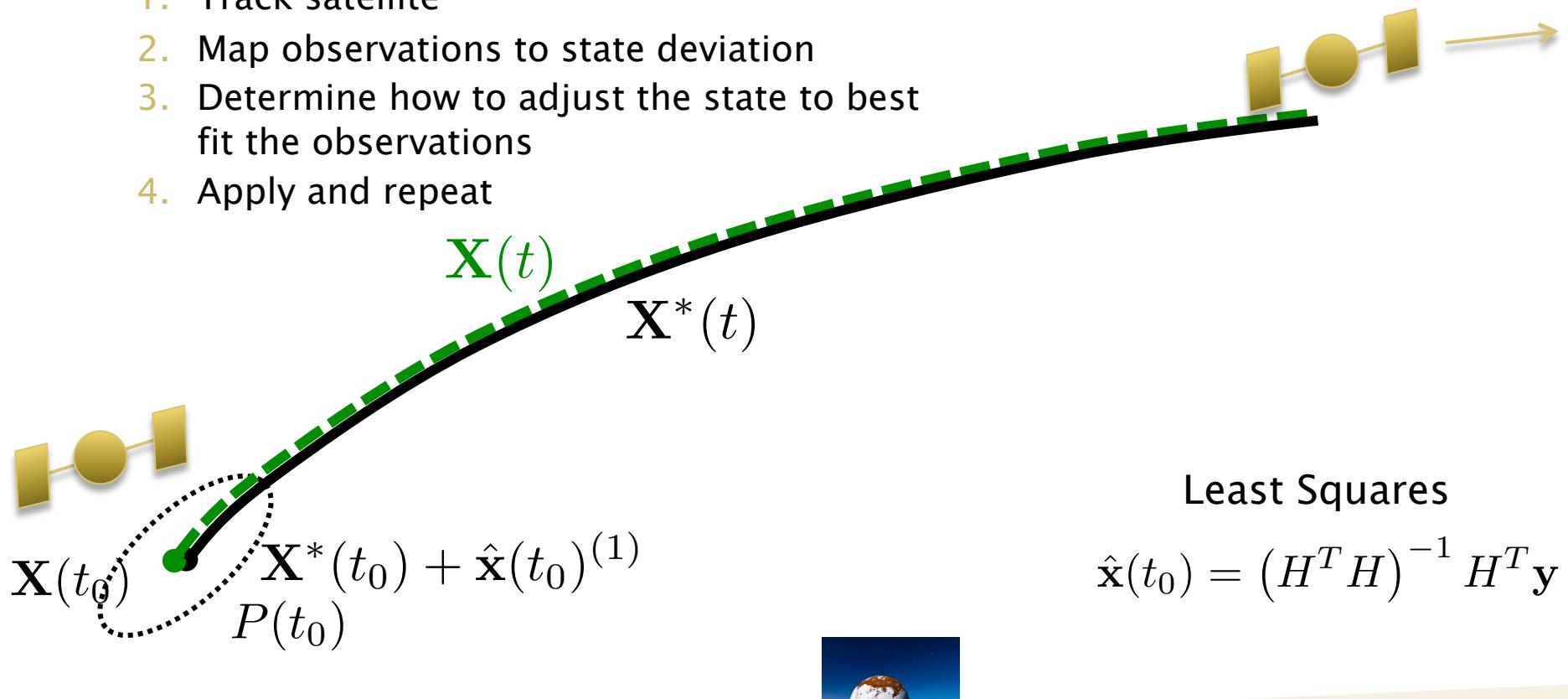
1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations



# Review of the Stat OD Process

## Process:

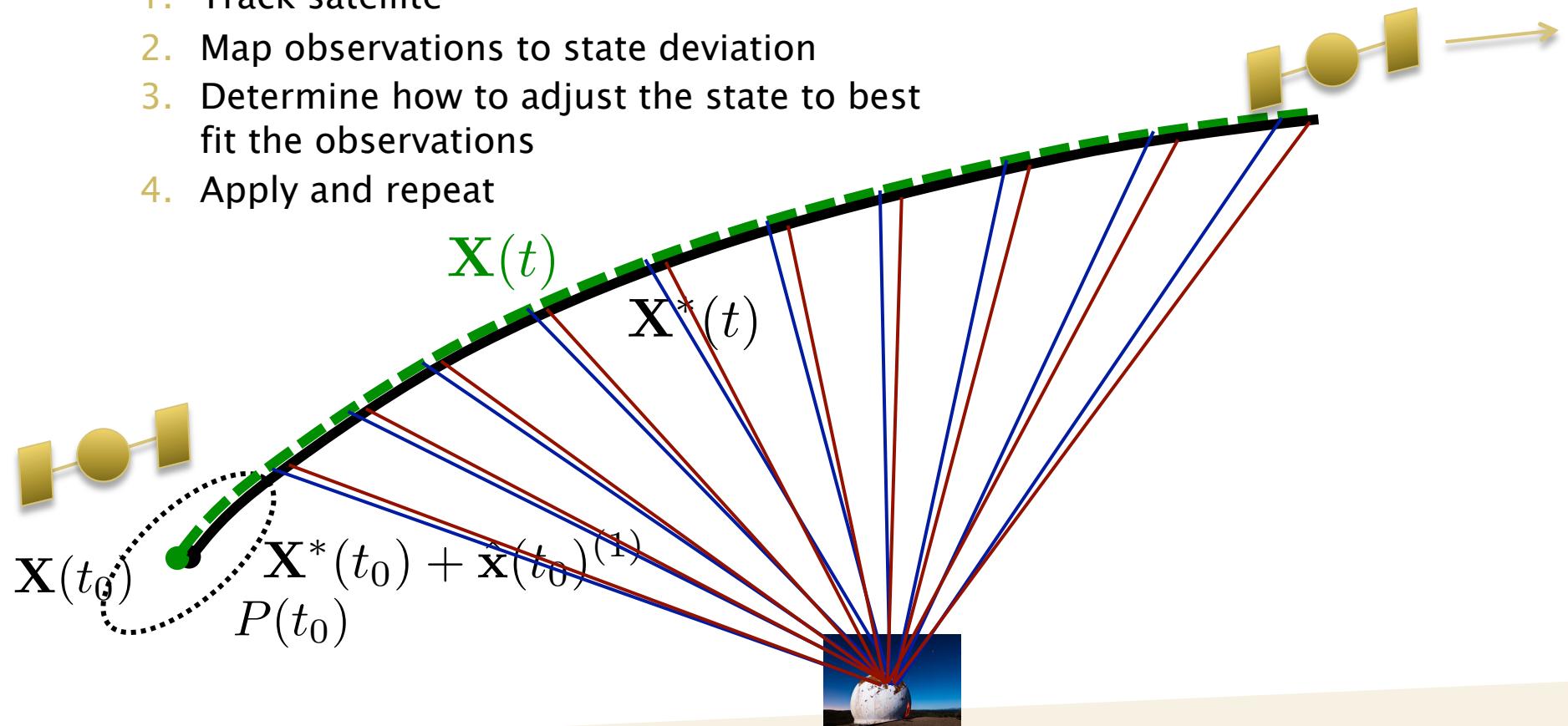
1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat



# Review of the Stat OD Process

## Process:

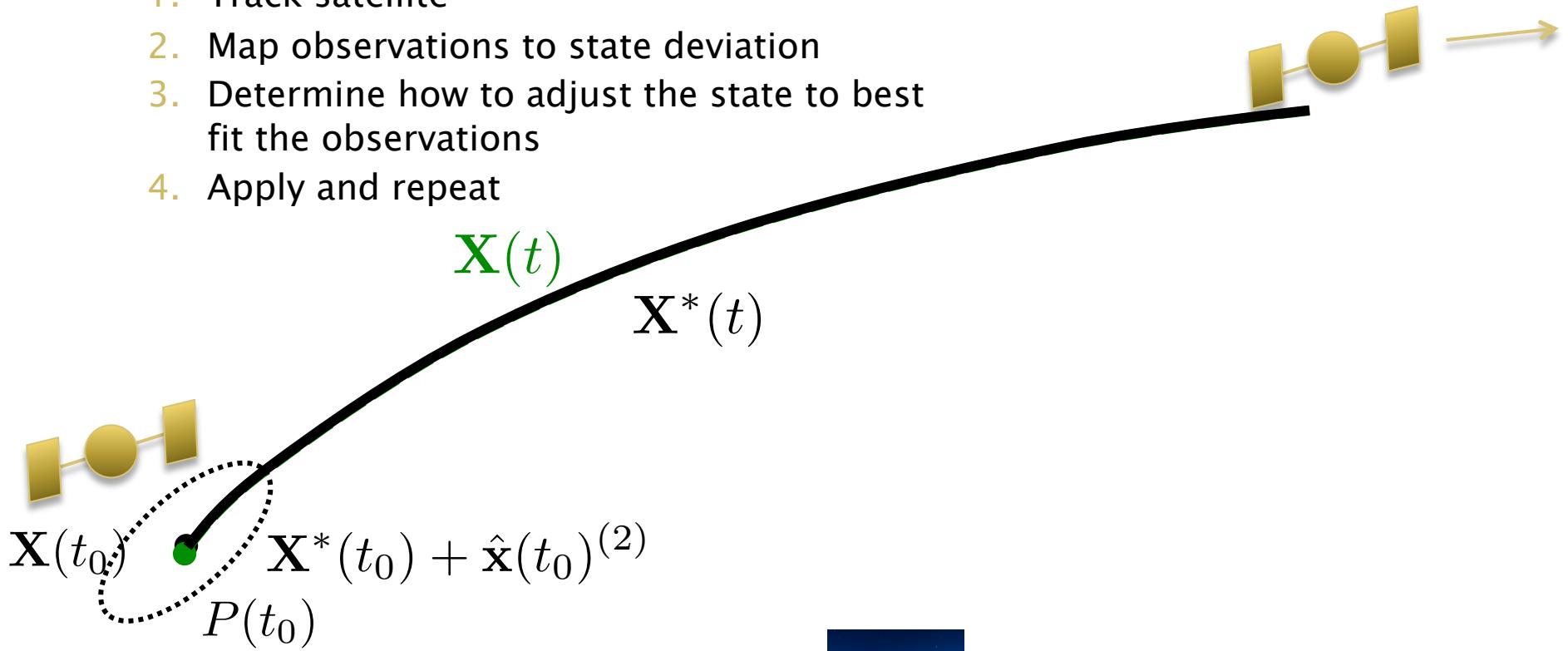
1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat



# Review of the Stat OD Process

## ▶ Process:

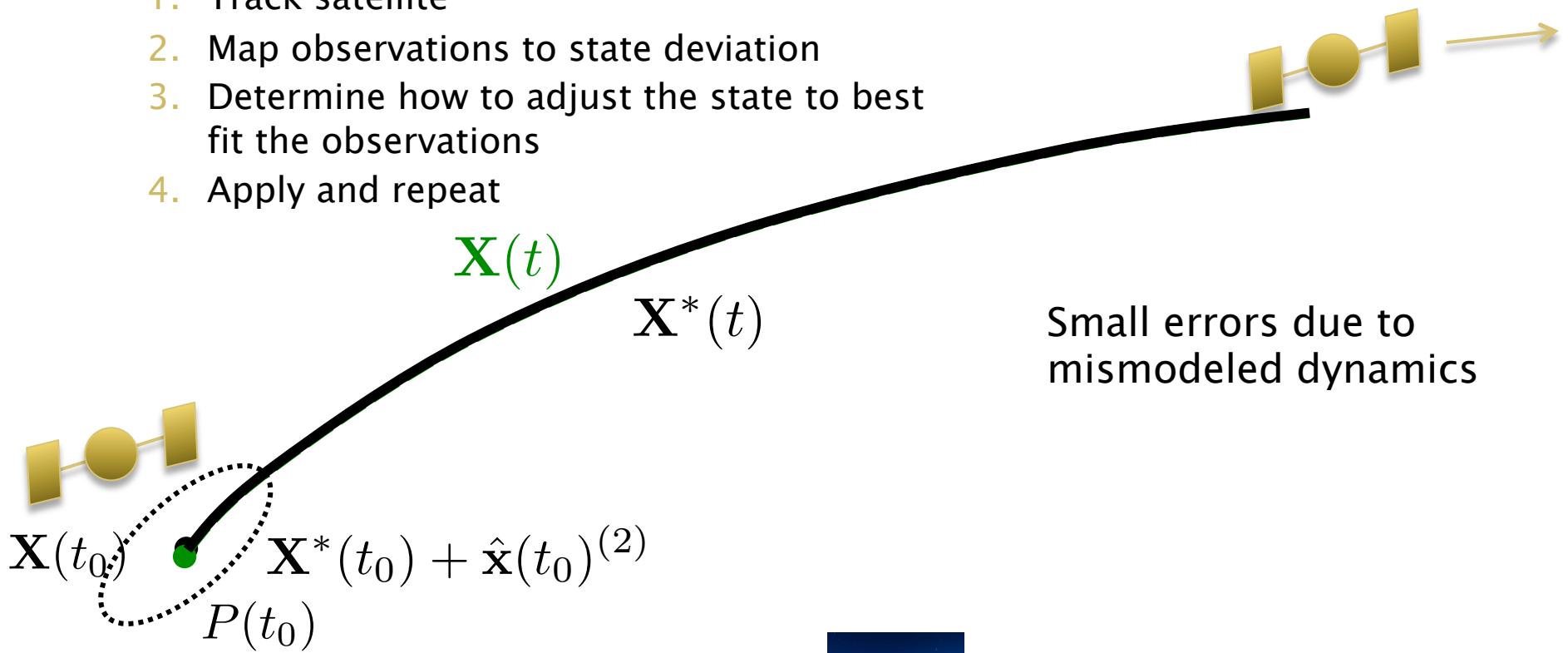
1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat



# Review of the Stat OD Process

- ▶ Process:

1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat

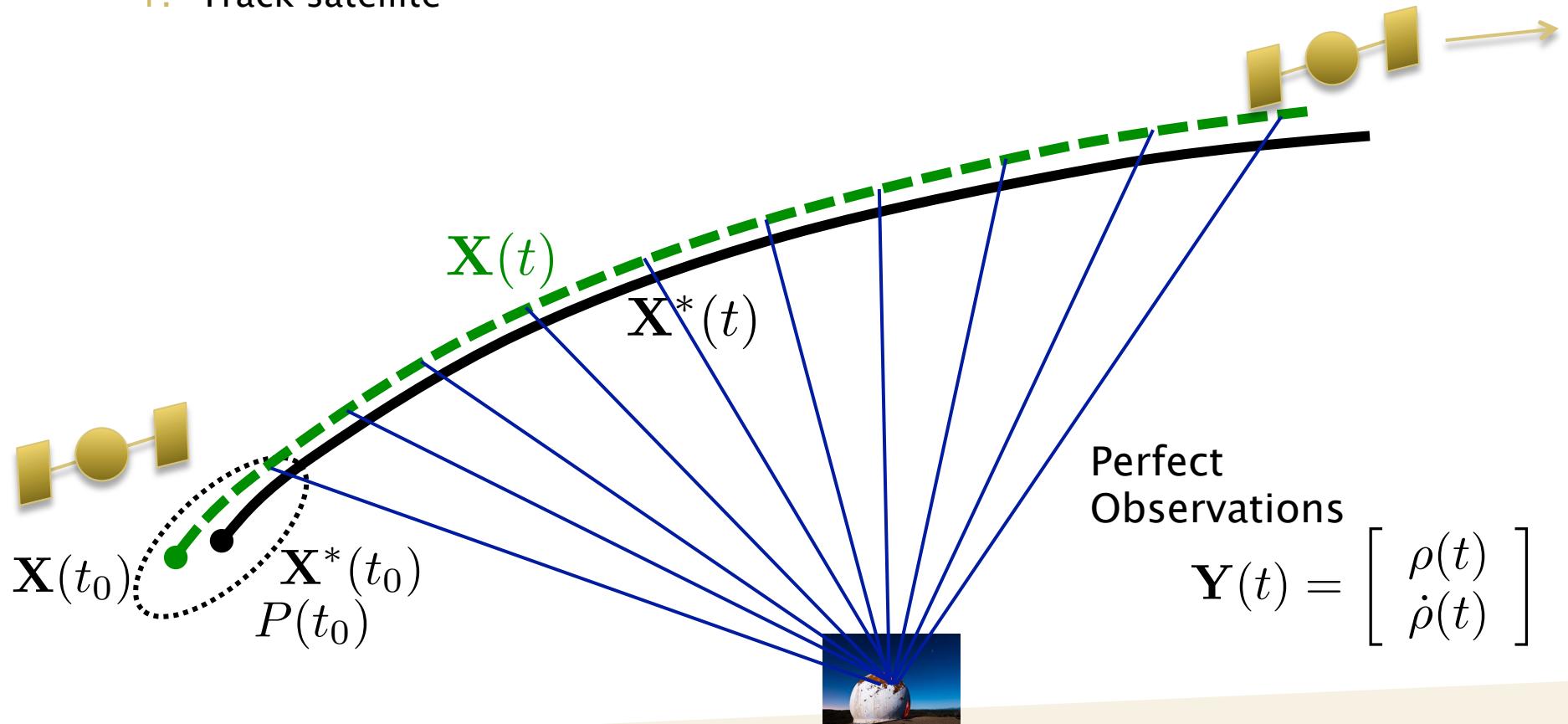


University of Colorado  
Boulder

# Review of the Stat OD Process

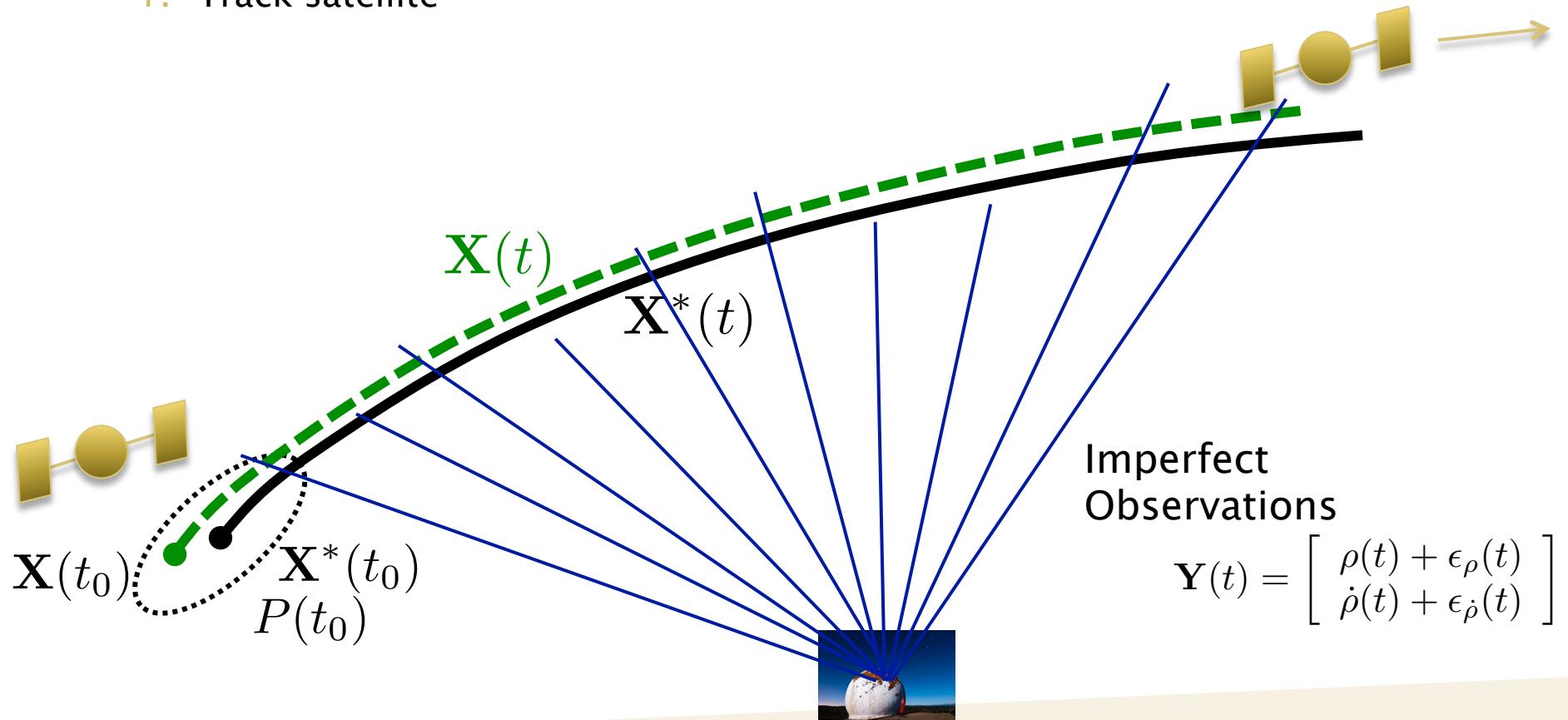
- ▶ Process:

1. Track satellite



# Review of the Stat OD Process

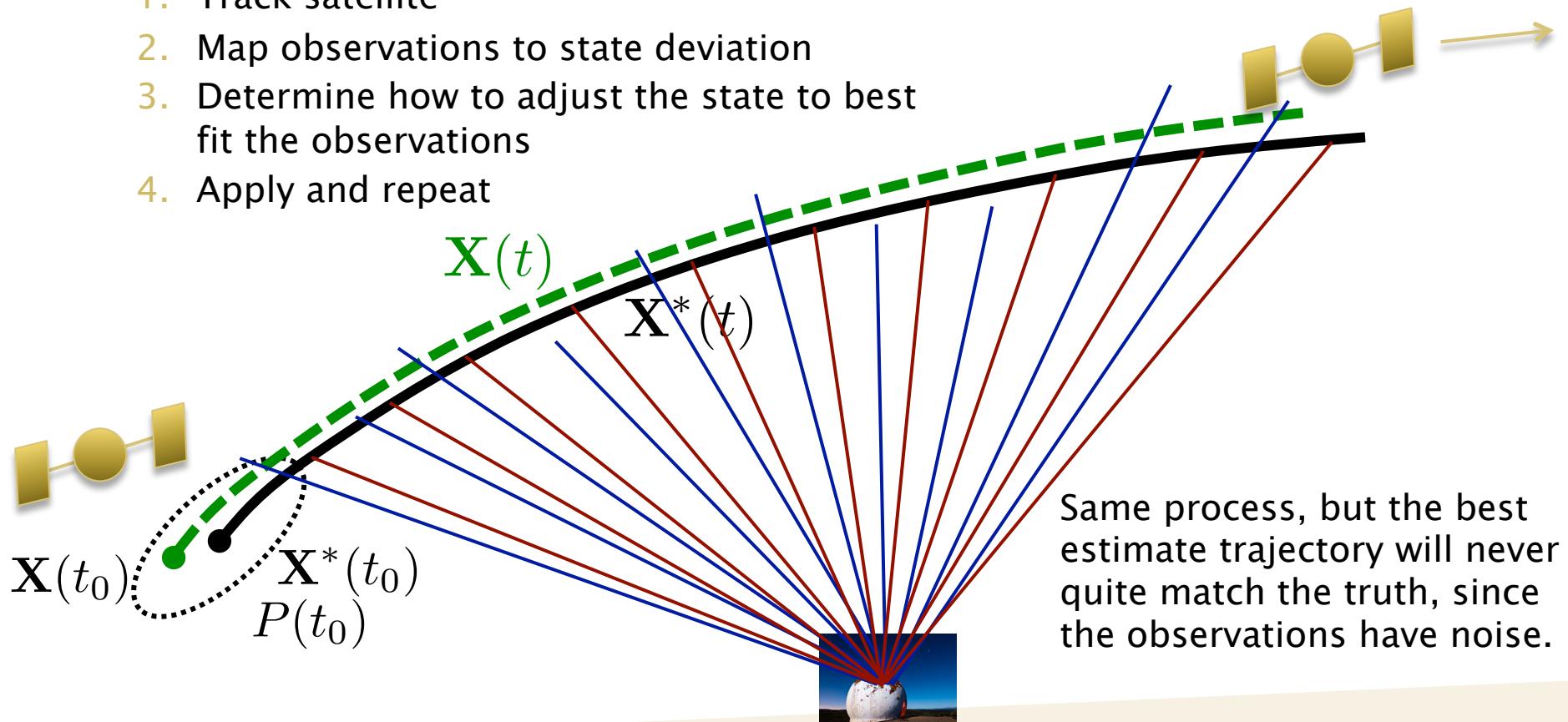
- ▶ Process:
  1. Track satellite



# Review of the Stat OD Process

## Process:

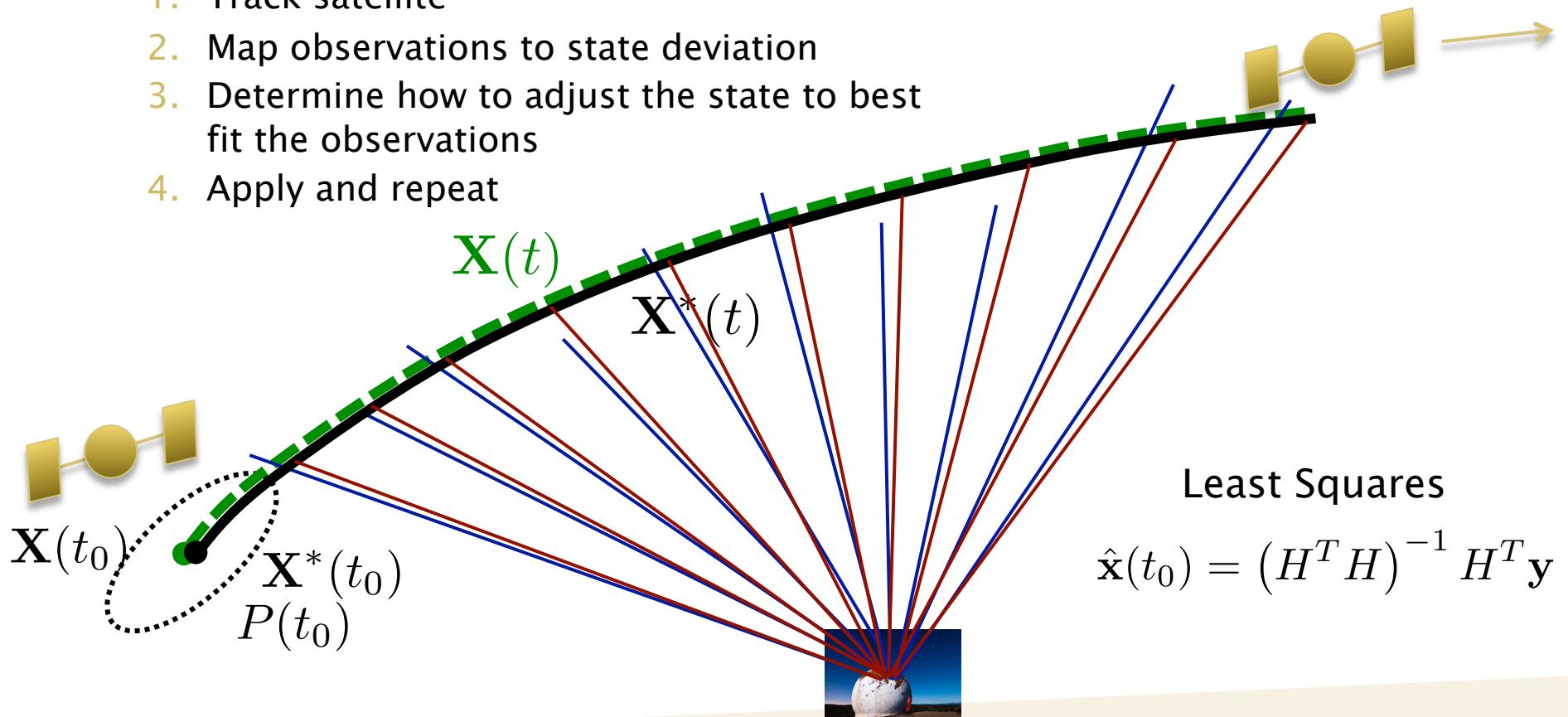
1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat



# Review of the Stat OD Process

## Process:

1. Track satellite
2. Map observations to state deviation
3. Determine how to adjust the state to best fit the observations
4. Apply and repeat



# Least Squares Options

---

- ▶ Least Squares

$$\hat{\mathbf{x}}_k = (H^T H)^{-1} H^T \mathbf{y}$$

- ▶ Weighted Least Squares

$$\hat{\mathbf{x}}_k = (H^T W H)^{-1} H^T W \mathbf{y}$$

- ▶ Least Squares with *a priori*

$$\hat{\mathbf{x}}_k = (H^T W H + \bar{W}_k)^{-1} (H^T W \mathbf{y} + \bar{W}_k \bar{\mathbf{x}}_k)$$

- ▶ Min Variance

$$\hat{\mathbf{x}}_k = (H^T R^{-1} H)^{-1} H^T R^{-1} \mathbf{y}$$

- ▶ Min Variance with *a priori*

$$\hat{\mathbf{x}}_k = (\tilde{H}_k^T R^{-1} \tilde{H}_k + \bar{P}_k^{-1})^{-1} (\tilde{H}_k^T R^{-1} \mathbf{y}_k + \bar{P}_k^{-1} \bar{\mathbf{x}}_k)$$



# Algorithm Options

---

## ▶ Batch

- Process all observations at once

$$\hat{\mathbf{x}}_0 = \left( \sum_{i=1}^p (H_i^T R^{-1} H_i) + \bar{P}_0^{-1} \right)^{-1} \left( \sum_{i=1}^p (H_i^T R^{-1} \mathbf{y}_i) + \bar{P}_0^{-1} \bar{\mathbf{x}}_0 \right)$$

## ▶ Sequential

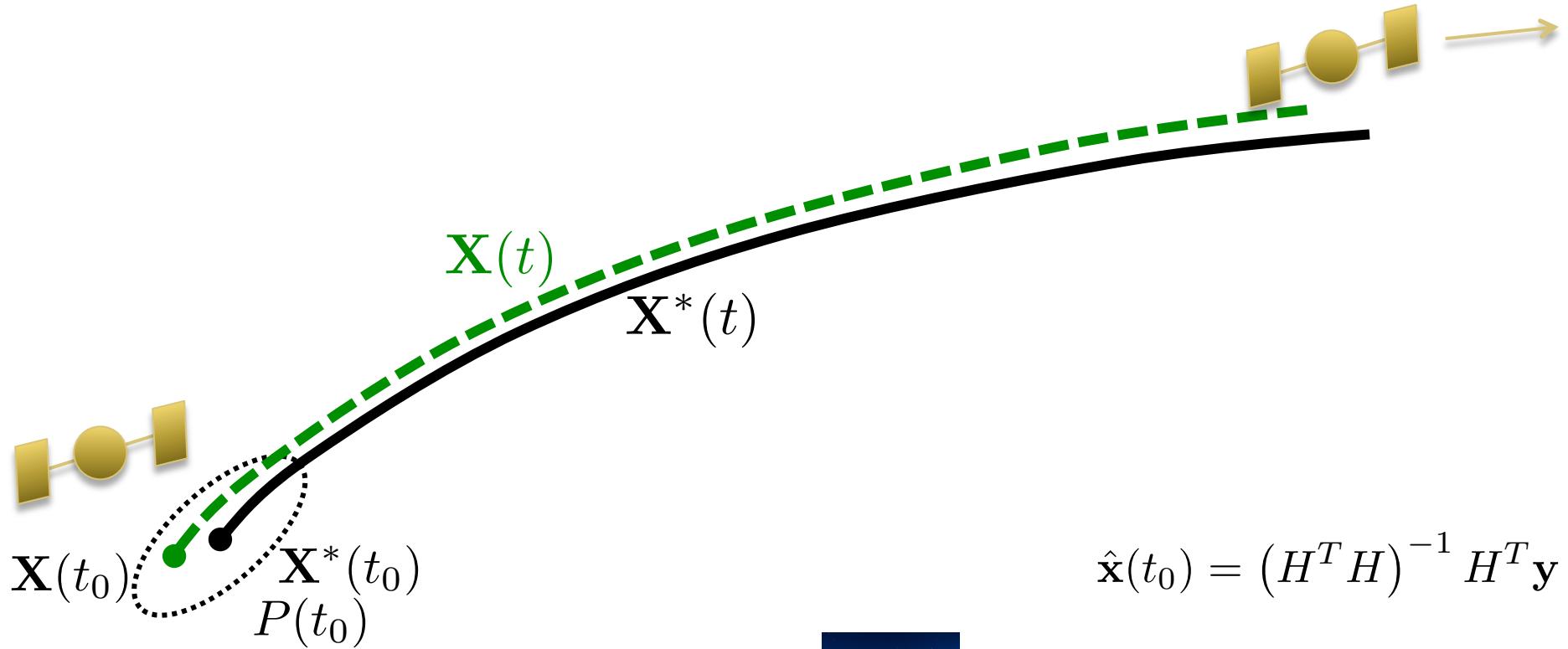
- Process one observation at a time

$$\hat{\mathbf{x}}_k = \left( \tilde{H}_k^T R^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1} \left( \tilde{H}_k^T R^{-1} \mathbf{y}_k + \bar{P}_k^{-1} \bar{\mathbf{x}}_k \right)$$



# Batch Processor

- ▶ Collect mapped information



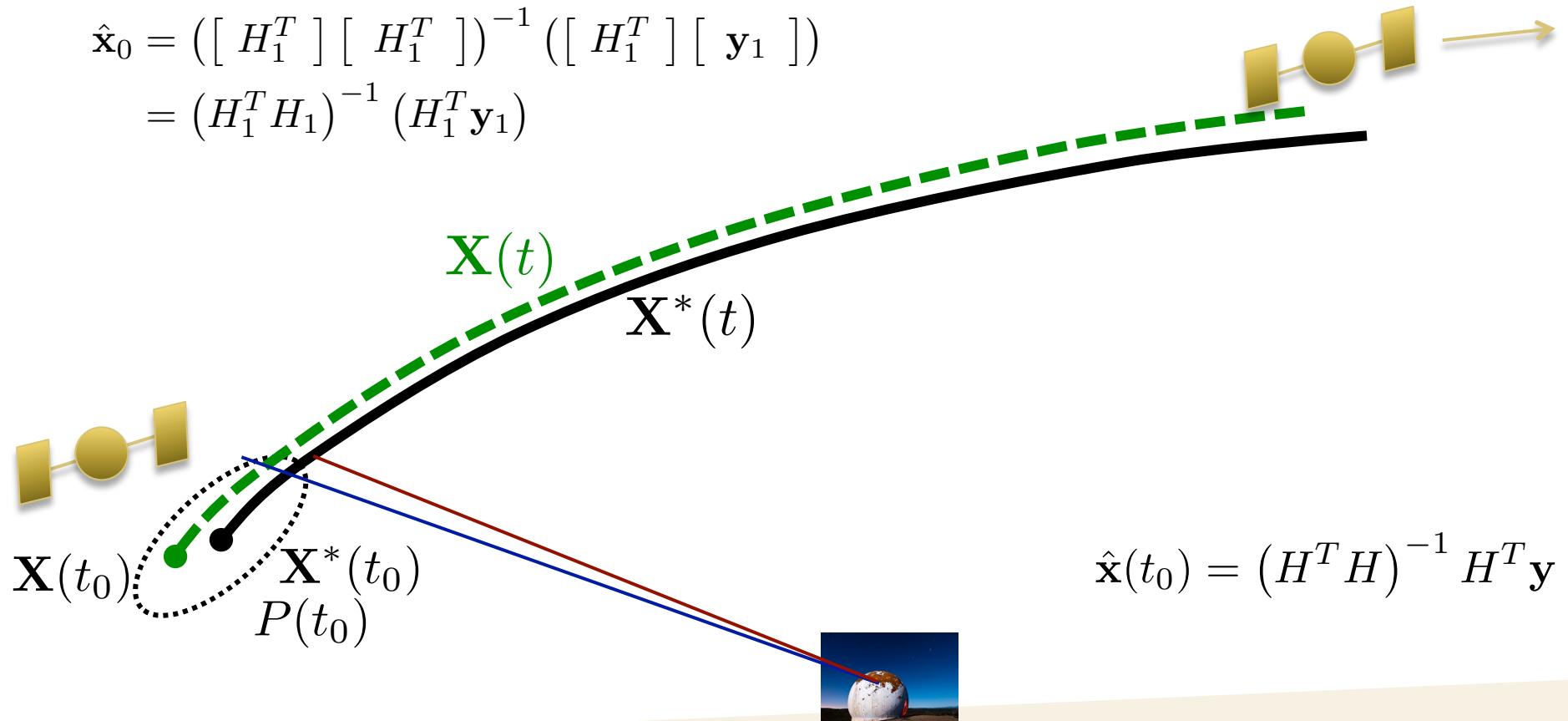
$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Batch Processor

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_0 = ([H_1^T] [H_1^T])^{-1} ([H_1^T] [\mathbf{y}_1]) \\ = (H_1^T H_1)^{-1} (H_1^T \mathbf{y}_1)$$



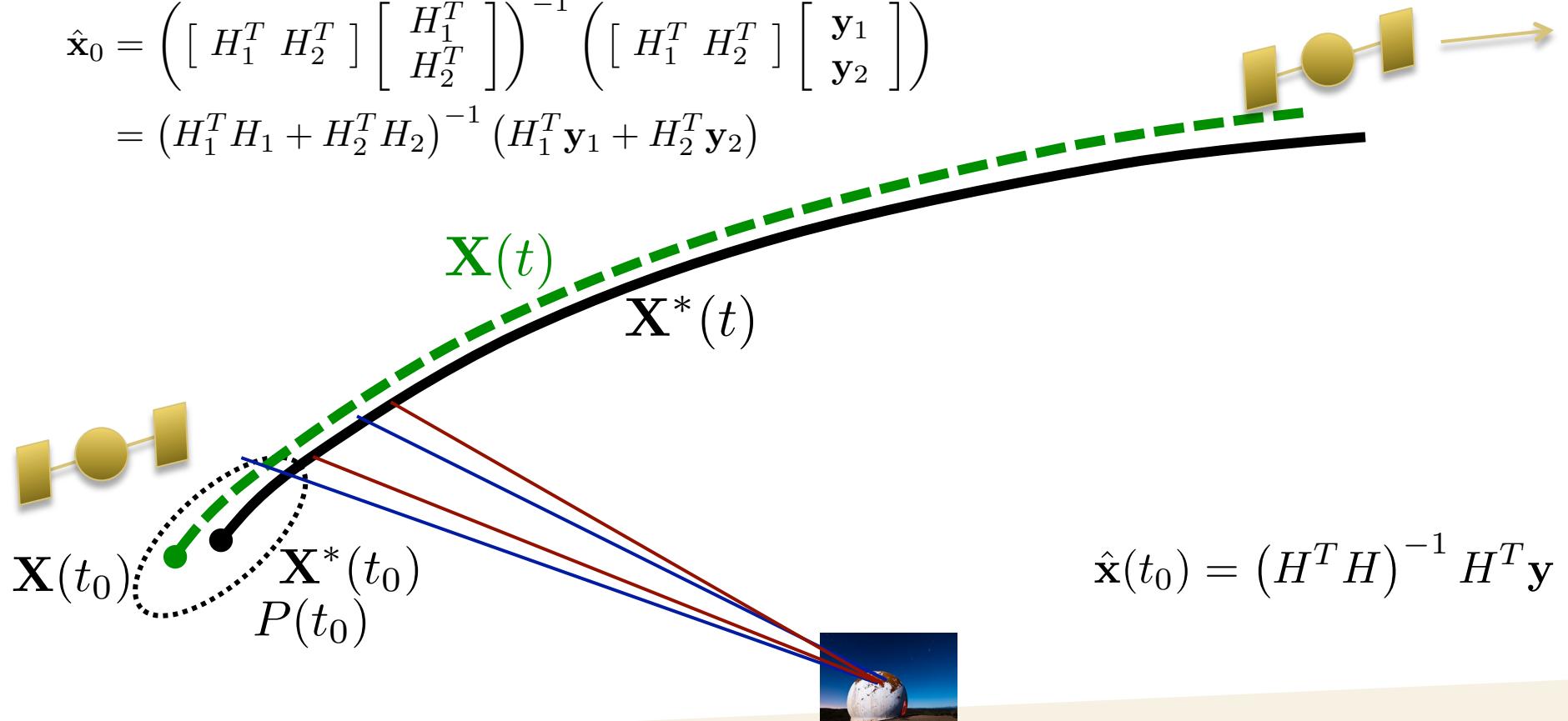
$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Batch Processor

- ▶ Collect mapped information

$$\begin{aligned}\hat{\mathbf{x}}_0 &= \left( \begin{bmatrix} H_1^T & H_2^T \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} H_1^T & H_2^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \right) \\ &= (H_1^T H_1 + H_2^T H_2)^{-1} (H_1^T \mathbf{y}_1 + H_2^T \mathbf{y}_2)\end{aligned}$$



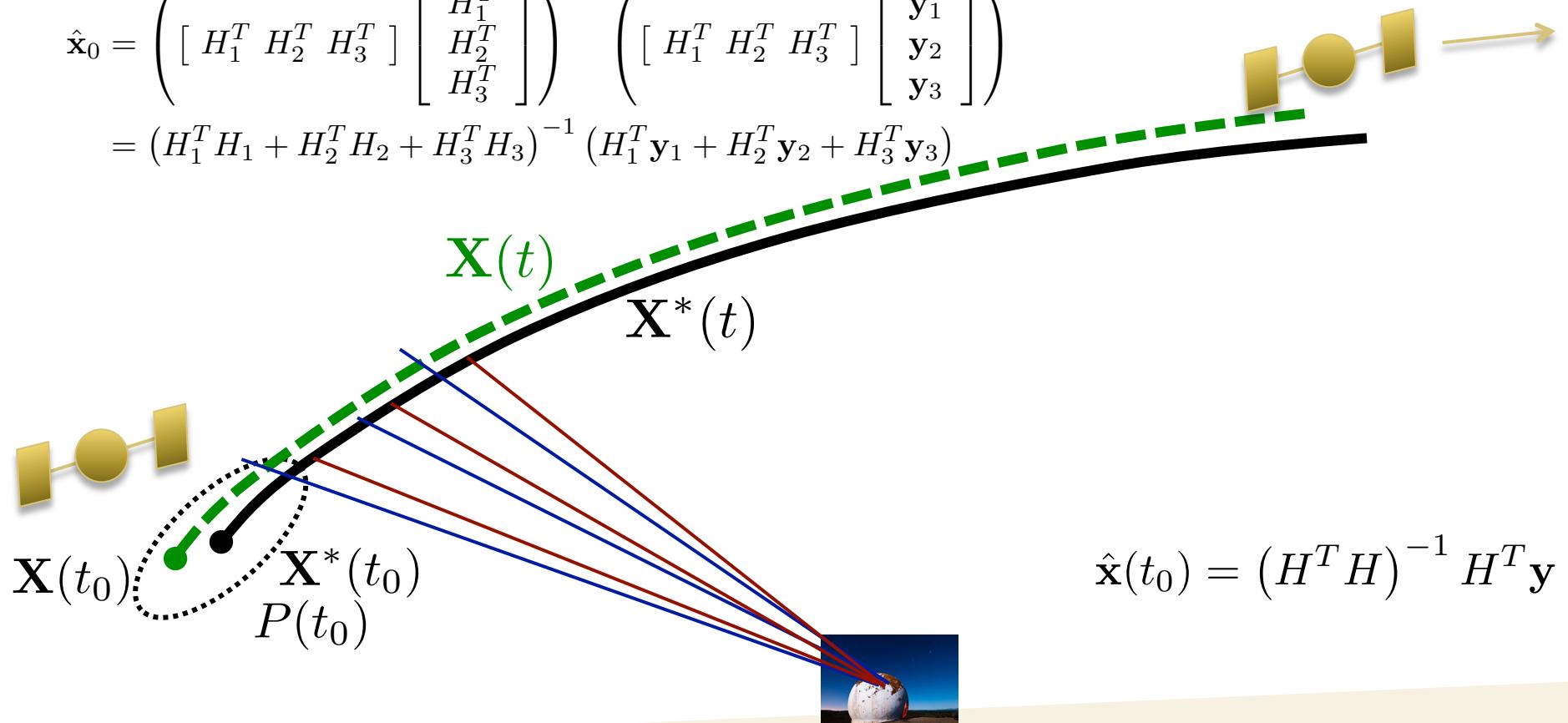
$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Batch Processor

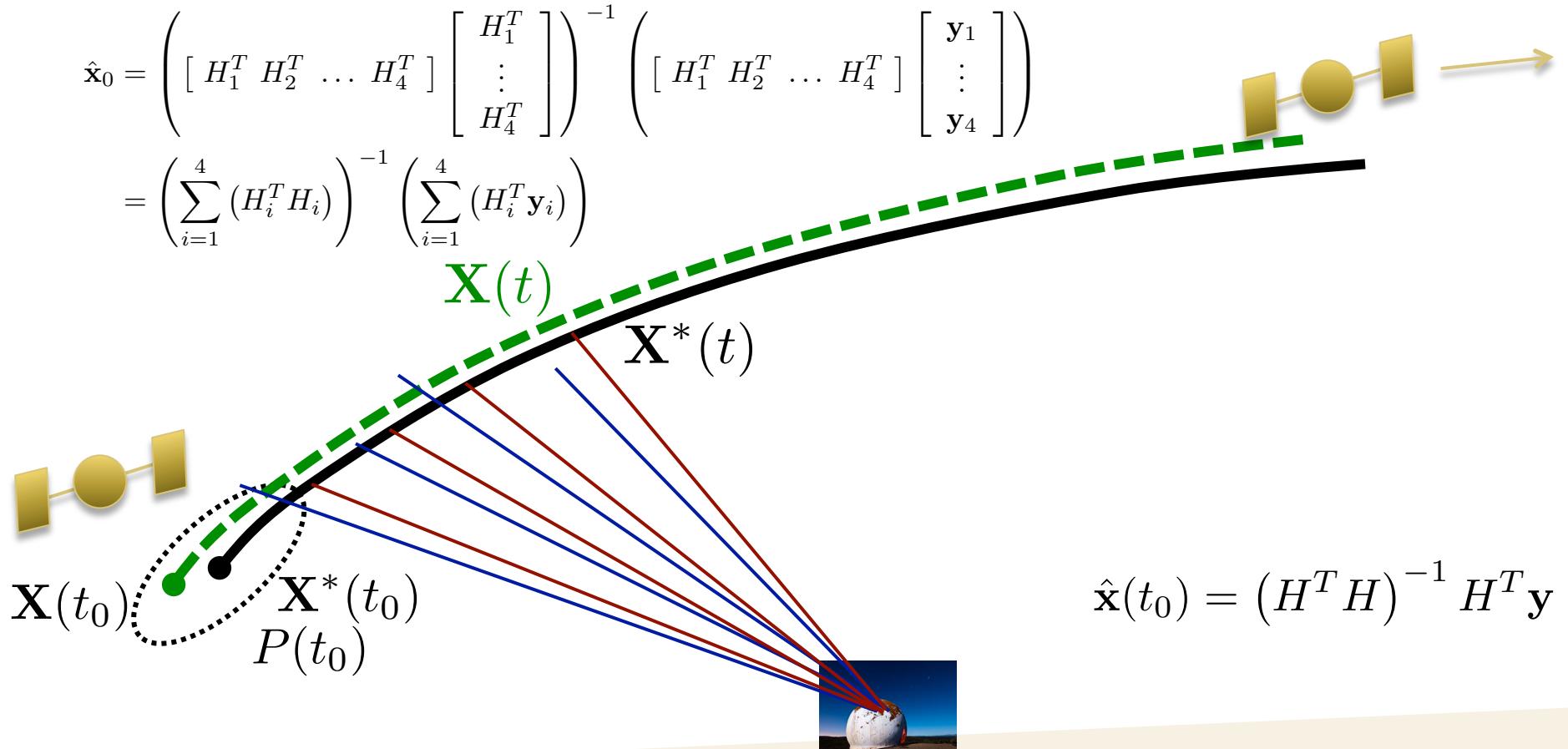
- ▶ Collect mapped information

$$\hat{\mathbf{x}}_0 = \left( \begin{bmatrix} H_1^T & H_2^T & H_3^T \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \\ H_3^T \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} H_1^T & H_2^T & H_3^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix} \right)$$
$$= (H_1^T H_1 + H_2^T H_2 + H_3^T H_3)^{-1} (H_1^T \mathbf{y}_1 + H_2^T \mathbf{y}_2 + H_3^T \mathbf{y}_3)$$



# Batch Processor

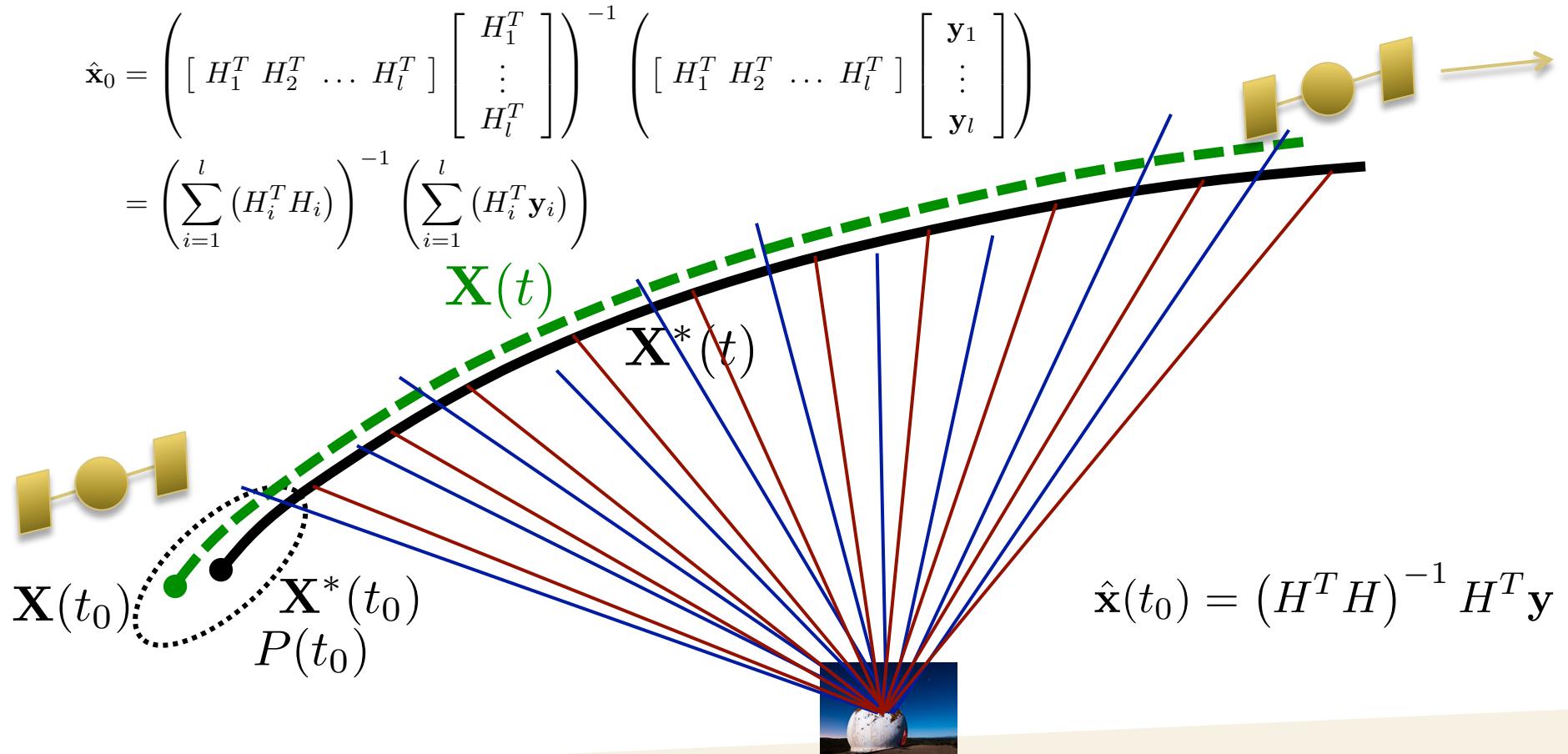
- ▶ Collect mapped information



University of Colorado  
Boulder

# Batch Processor

- ▶ Collect mapped information



University of Colorado  
Boulder

► (Break)

- TAs will go through an end-to-end Batch run to demo equations, etc.
- Next up: Kalman Filter



University of Colorado  
Boulder

# Sequential Processor

- ▶ Consider

$$\hat{\mathbf{x}}_k = \left( \tilde{H}_k^T R^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1} \left( \tilde{H}_k^T R^{-1} \mathbf{y}_k + \bar{P}_k^{-1} \bar{\mathbf{x}}_k \right)$$

- ▶ Rather than mapping all observations to one epoch and processing them simultaneously, what if we processed each separately and mapped the best estimate through each?



# Sequential Processor

- ▶ Consider

$$\hat{\mathbf{x}}_k = \left( \tilde{H}_k^T R^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1} \left( \tilde{H}_k^T R^{-1} \mathbf{y}_k + \bar{P}_k^{-1} \bar{\mathbf{x}}_k \right)$$

- ▶ Rather than mapping all observations to one epoch and processing them simultaneously, what if we processed each separately and mapped the best estimate through each?

$$\bar{\mathbf{x}}_k = \Phi(t_k, t_j) \hat{\mathbf{x}}_j$$

$$\bar{P}_k = \Phi(t_k, t_j) P_j \Phi^T(t_k, t_j)$$



# Sequential Processor

---

- Given an *a priori* state and covariance, we know how to generate a new estimate of the state:

$$\hat{\mathbf{x}}_k = \left( \tilde{H}_k^T R^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1} \left( \tilde{H}_k^T R^{-1} \mathbf{y}_k + \bar{P}_k^{-1} \bar{\mathbf{x}}_k \right)$$

- Need a way to generate the *a posteriori* covariance matrix as well.
- Recall

$$P_k = \Lambda_k^{-1} = \left( \tilde{H}_k^T R_k^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1}$$

- The trouble is inverting the  $n \times n$  matrix.



# Sequential Processor

- ▶ We can use the Schur Identity and a bunch of math (see Section 4.7) and obtain:

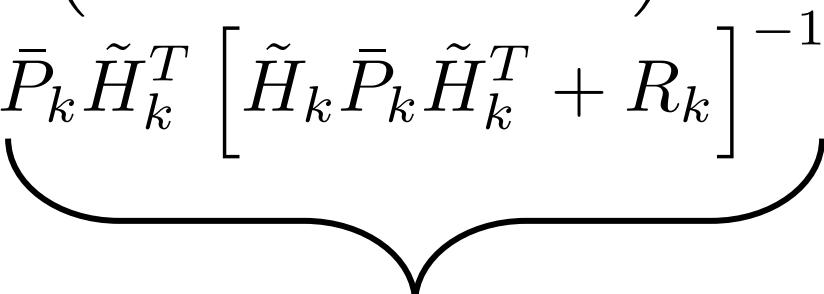
$$P_k = \Lambda_k^{-1} = \left( \tilde{H}_k^T R_k^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1}$$

$$P_k = \bar{P}_k - \bar{P}_k \tilde{H}_k^T \left[ \tilde{H}_k \bar{P}_k \tilde{H}_k^T + R_k \right]^{-1} \tilde{H}_k \bar{P}_k$$



# Sequential Processor

- ▶ We can use the Schur Identity and a bunch of math (see Section 4.7) and obtain:

$$P_k = \Lambda_k^{-1} = \left( \tilde{H}_k^T R_k^{-1} \tilde{H}_k + \bar{P}_k^{-1} \right)^{-1}$$
$$P_k = \bar{P}_k - \bar{P}_k \tilde{H}_k^T \left[ \tilde{H}_k \bar{P}_k \tilde{H}_k^T + R_k \right]^{-1} \tilde{H}_k \bar{P}_k$$

$$K_k$$

Kalman Gain



University of Colorado  
Boulder

# Sequential Processor

- ▶ After some more math, we can simplify to obtain:

$$P_k = \left[ I - K_k \tilde{H}_k^T \right] \bar{P}_k$$

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + K_k \left[ \mathbf{y}_k - \tilde{H}_k \bar{\mathbf{x}}_k \right]$$



# Sequential Algorithm

- ▶ 1. Initialize the first run
- ▶ 2. Start at the reference epoch
- ▶ 3. Time Update
  - Integrate from the current time to the next time of interest
  - Map the state estimate and the covariance to the new time
- ▶ 4. Measurement Update
  - If there is a new measurement, process it.
  - Update the state estimate and covariance with this new information
- ▶ Repeat 3–4 until all measurements have been processed and all times of interest have been recorded.
- ▶ Optional: Map the estimate and covariance back to the reference epoch and iterate the whole process.



# Sequential Algorithm

## ► Initialization

*Given:*  $\hat{\mathbf{x}}_{k-1}$ ,  $P_{k-1}$ ,  $\mathbf{X}_{k-1}^*$ , and  $R_k$ , and the observation  $\mathbf{Y}_k$ , at  $t_k$  (at the initial time  $t_0$ , these would be  $\mathbf{X}_0^*$ ,  $\hat{\mathbf{x}}_0$ , and  $P_0$ ).



# Sequential Algorithm

## ► Time Update

- Integration

$$\begin{aligned}\dot{\mathbf{X}}^* &= F(\mathbf{X}^*, t), \\ \dot{\Phi}(t, t_{k-1}) &= A(t)\Phi(t, t_{k-1}),\end{aligned}$$

$$\begin{aligned}\mathbf{X}^*(t_{k-1}) &= \mathbf{X}_{k-1}^* \\ \Phi(t_{k-1}, t_{k-1}) &= I.\end{aligned}$$

- Mapping

$$\bar{\mathbf{x}}_k = \Phi(t_k, t_{k-1})\hat{\mathbf{x}}_{k-1} \quad \bar{P}_k = \Phi(t_k, t_{k-1})P_{k-1}\Phi^T(t_k, t_{k-1})$$



University of Colorado  
Boulder

# Sequential Algorithm

## ▶ Measurement Update

- Collect measurement

$$\mathbf{y}_k = \mathbf{Y}_k - G(\mathbf{X}_k^*, t_k) \quad \tilde{H}_k = \frac{\partial G(\mathbf{X}_k^*, t_k)}{\partial \mathbf{X}}$$

- Compute update

$$K_k = \bar{P}_k \tilde{H}_k^T [\tilde{H}_k \bar{P}_k \tilde{H}_k^T + R_k]^{-1}$$

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + K_k [\mathbf{y}_k - \tilde{H}_k \bar{\mathbf{x}}_k]$$

$$P_k = [I - K_k \tilde{H}_k] \bar{P}_k.$$



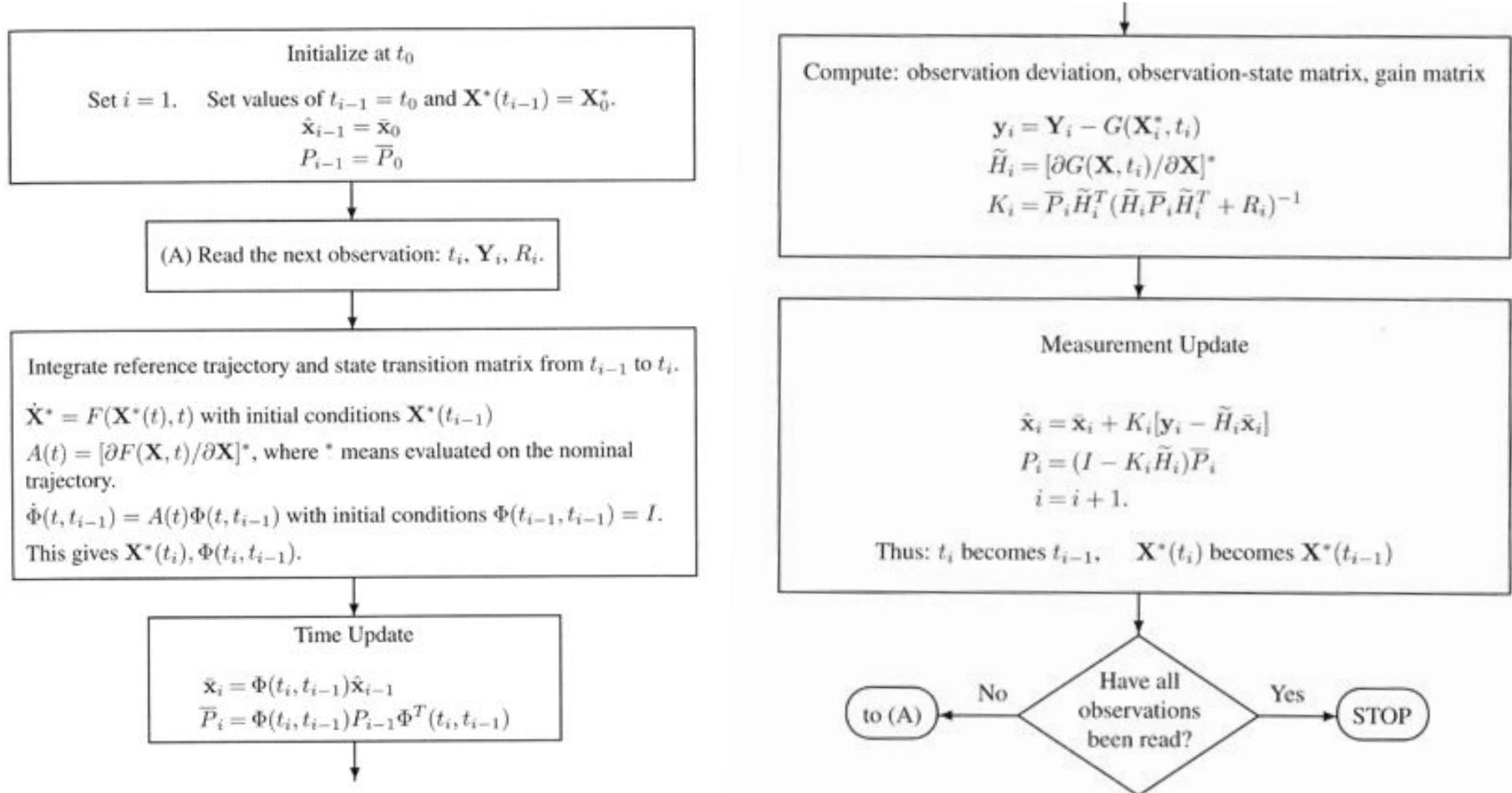
# Sequential Processor

## ► Repeat just like the Batch

- Replace the reference trajectory with the new best estimate.
- Make sure to update the *a priori* state deviation vector to retain any information.
- Recompute all observation residuals



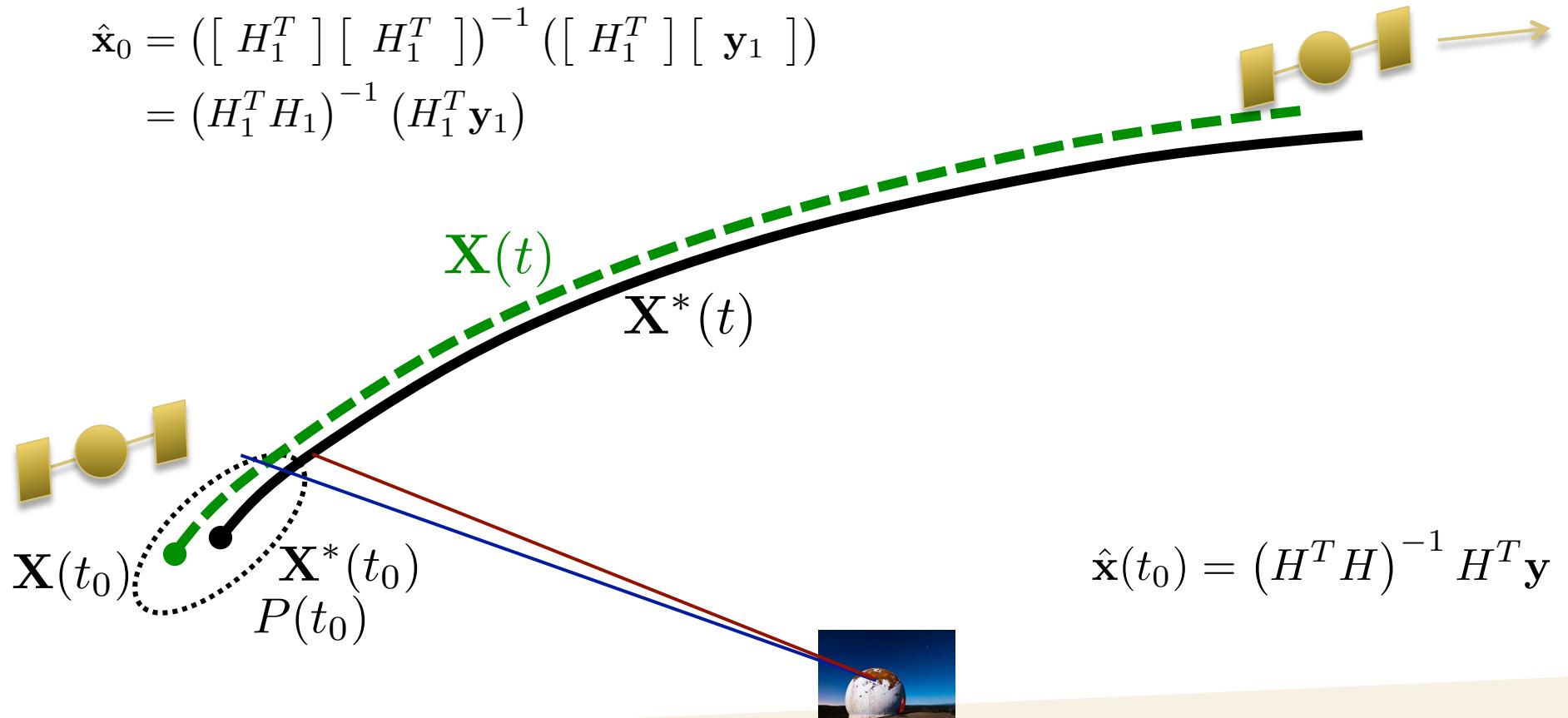
# Sequential Flow-Chart



# Batch Processor

- ▶ Collect mapped information

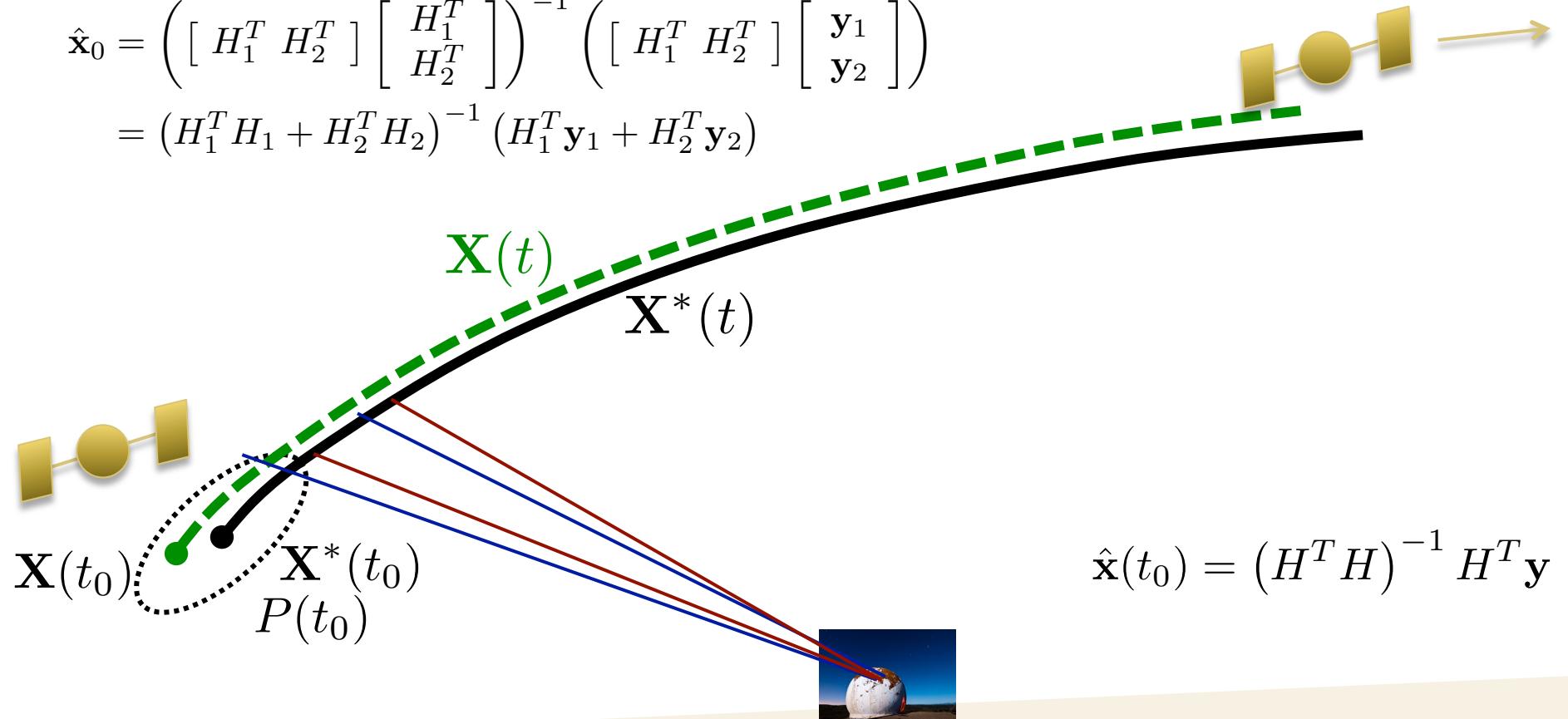
$$\hat{\mathbf{x}}_0 = ([H_1^T] [H_1^T])^{-1} ([H_1^T] [\mathbf{y}_1]) \\ = (H_1^T H_1)^{-1} (H_1^T \mathbf{y}_1)$$



# Batch Processor

- ▶ Collect mapped information

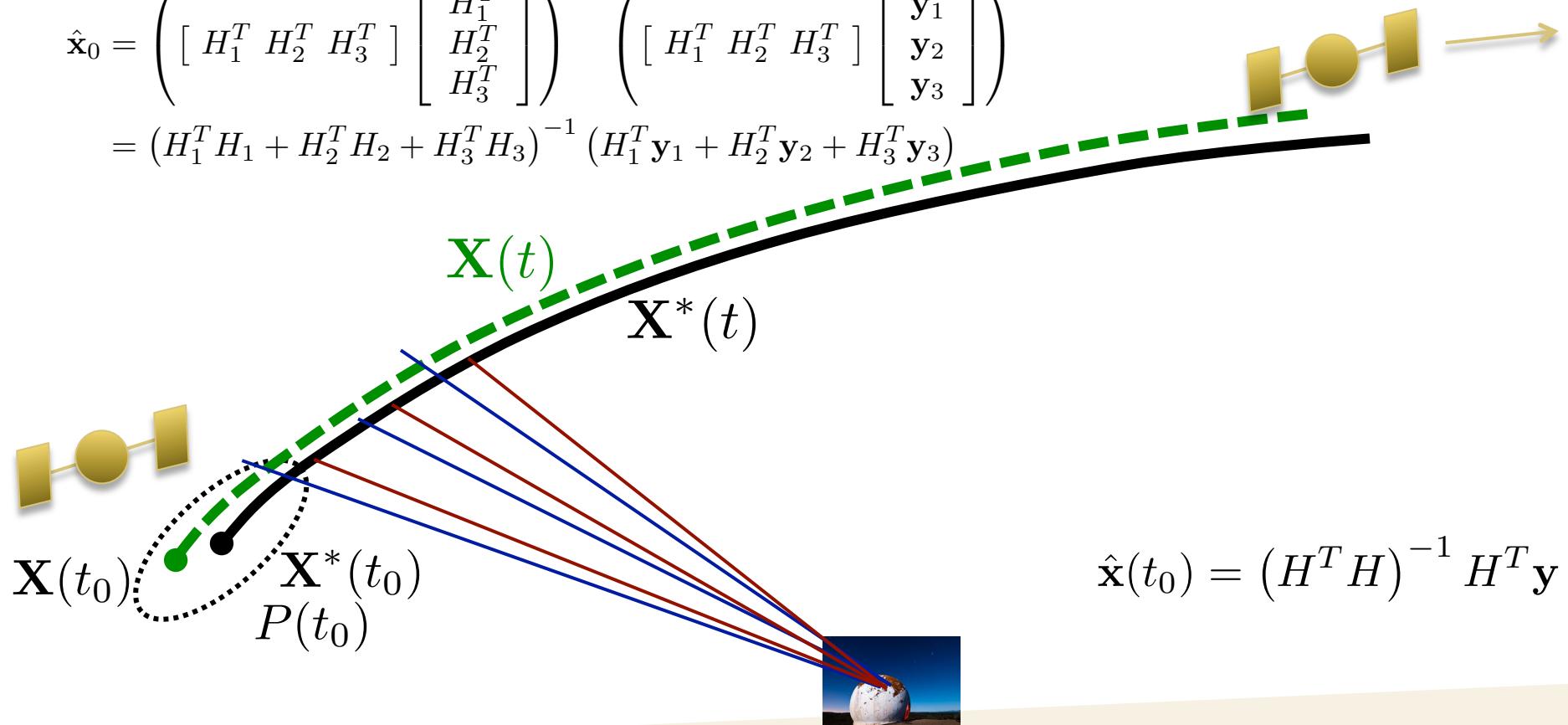
$$\begin{aligned}\hat{\mathbf{x}}_0 &= \left( \begin{bmatrix} H_1^T & H_2^T \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} H_1^T & H_2^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \right) \\ &= (H_1^T H_1 + H_2^T H_2)^{-1} (H_1^T \mathbf{y}_1 + H_2^T \mathbf{y}_2)\end{aligned}$$



# Batch Processor

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_0 = \left( \begin{bmatrix} H_1^T & H_2^T & H_3^T \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \\ H_3^T \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} H_1^T & H_2^T & H_3^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix} \right)$$
$$= (H_1^T H_1 + H_2^T H_2 + H_3^T H_3)^{-1} (H_1^T \mathbf{y}_1 + H_2^T \mathbf{y}_2 + H_3^T \mathbf{y}_3)$$

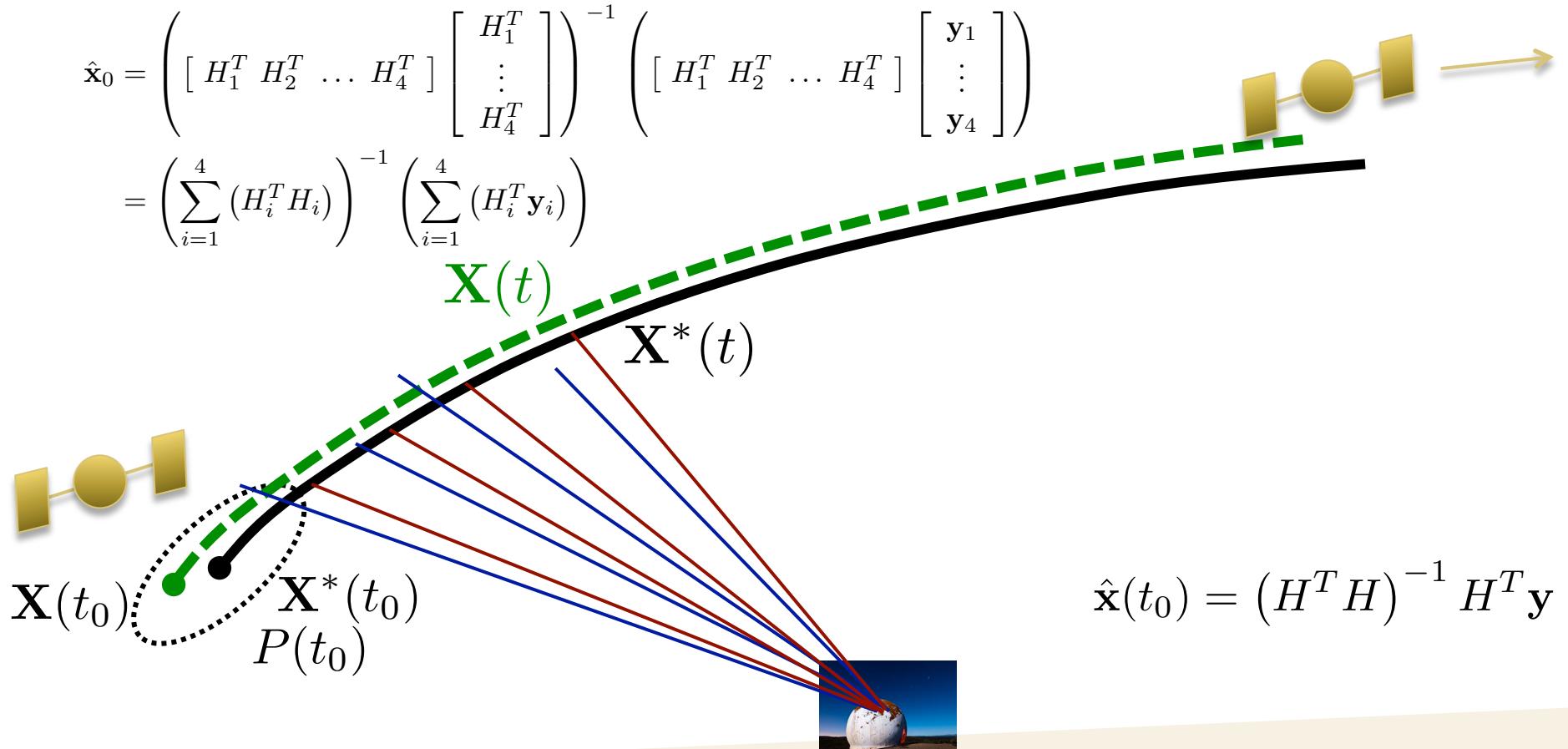


$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Batch Processor

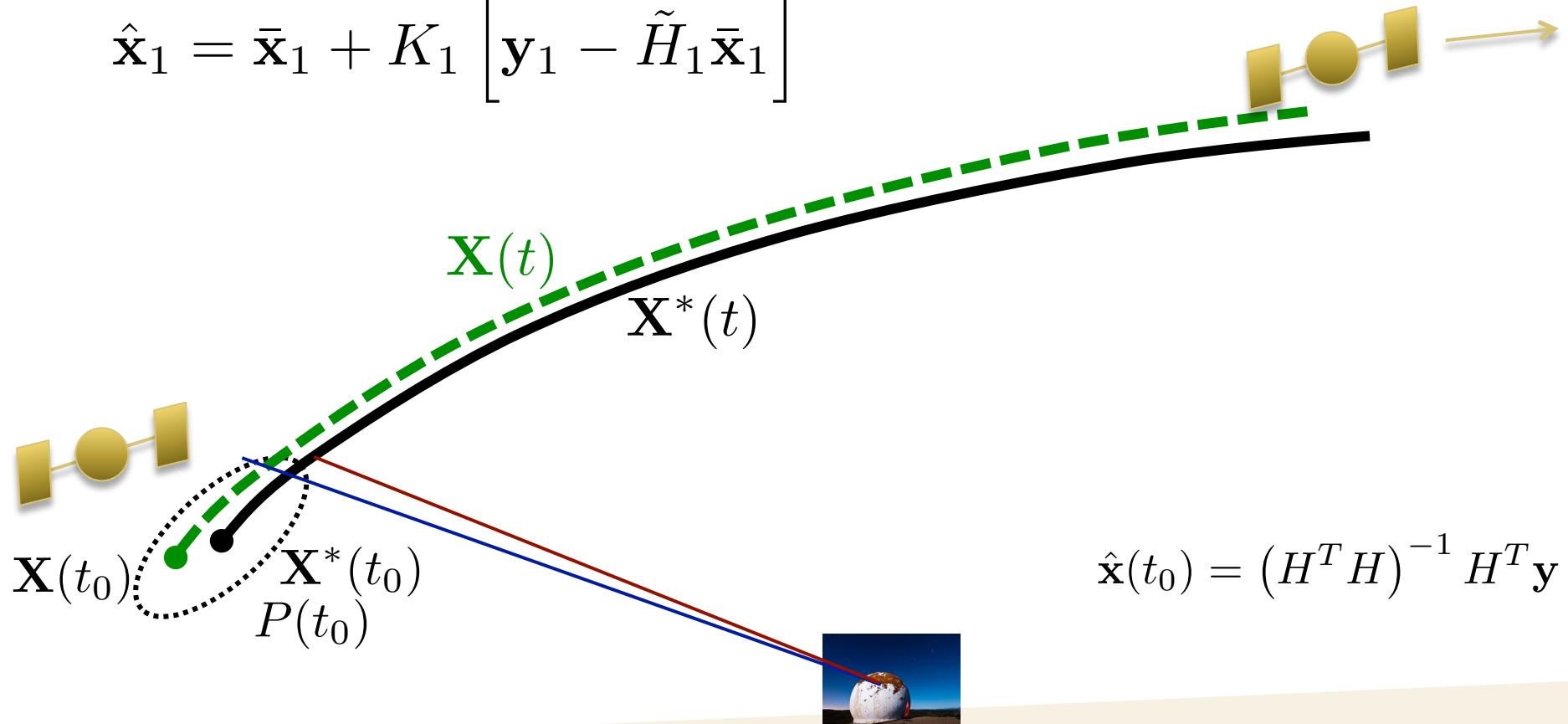
- ▶ Collect mapped information



# Kalman Filter

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_1 = \bar{\mathbf{x}}_1 + K_1 \left[ \mathbf{y}_1 - \tilde{H}_1 \bar{\mathbf{x}}_1 \right]$$



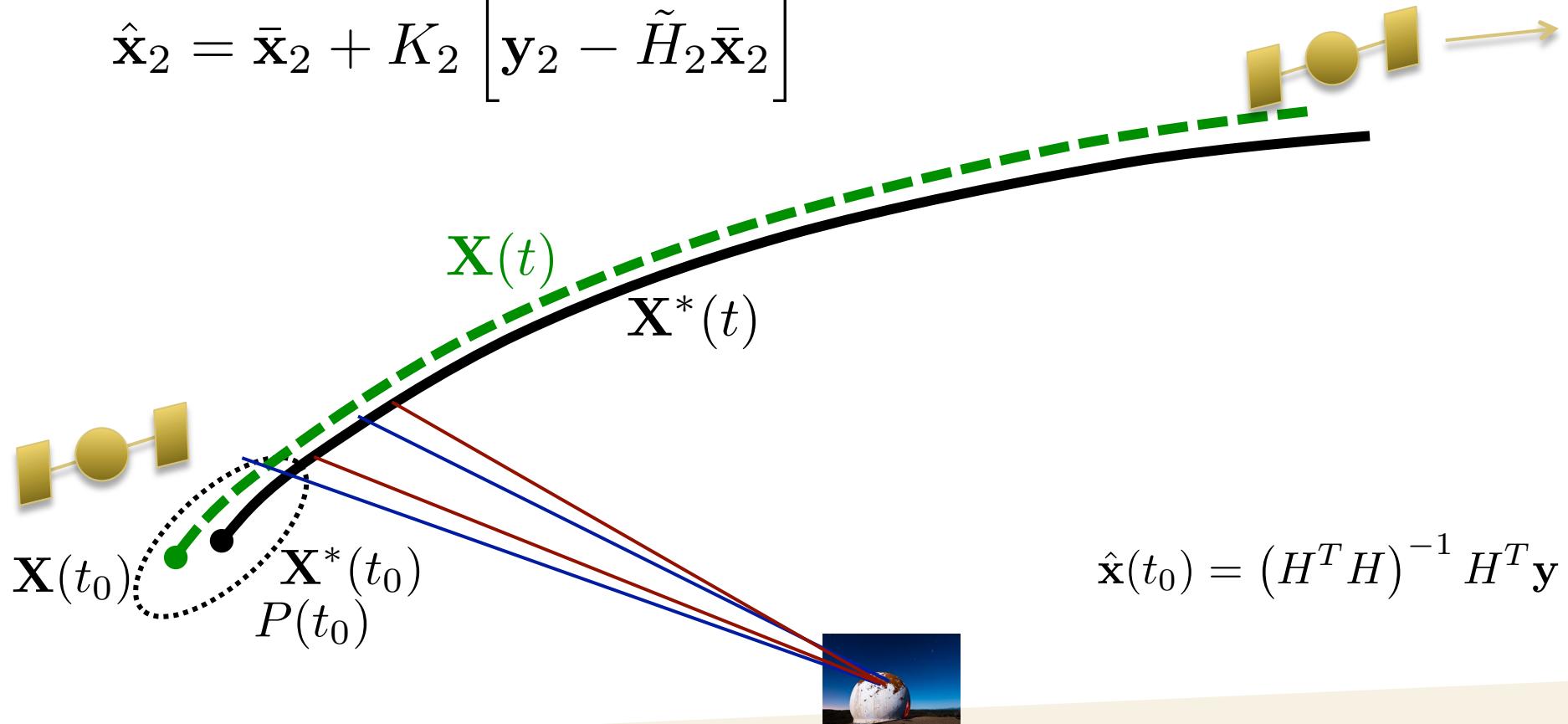
$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Kalman Filter

- ▶ Collect mapped information

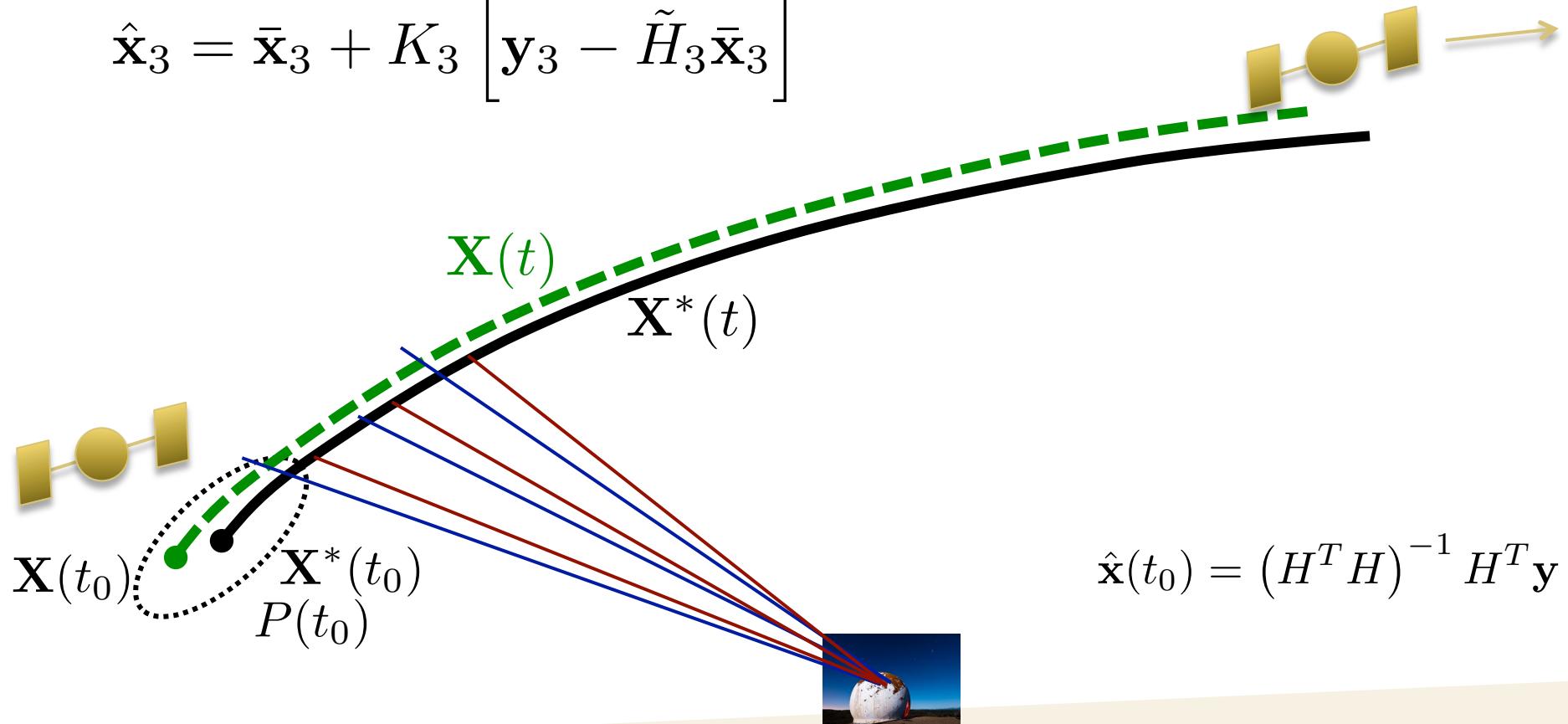
$$\hat{\mathbf{x}}_2 = \bar{\mathbf{x}}_2 + K_2 \left[ \mathbf{y}_2 - \tilde{H}_2 \bar{\mathbf{x}}_2 \right]$$



# Kalman Filter

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_3 = \bar{\mathbf{x}}_3 + K_3 \left[ \mathbf{y}_3 - \tilde{H}_3 \bar{\mathbf{x}}_3 \right]$$



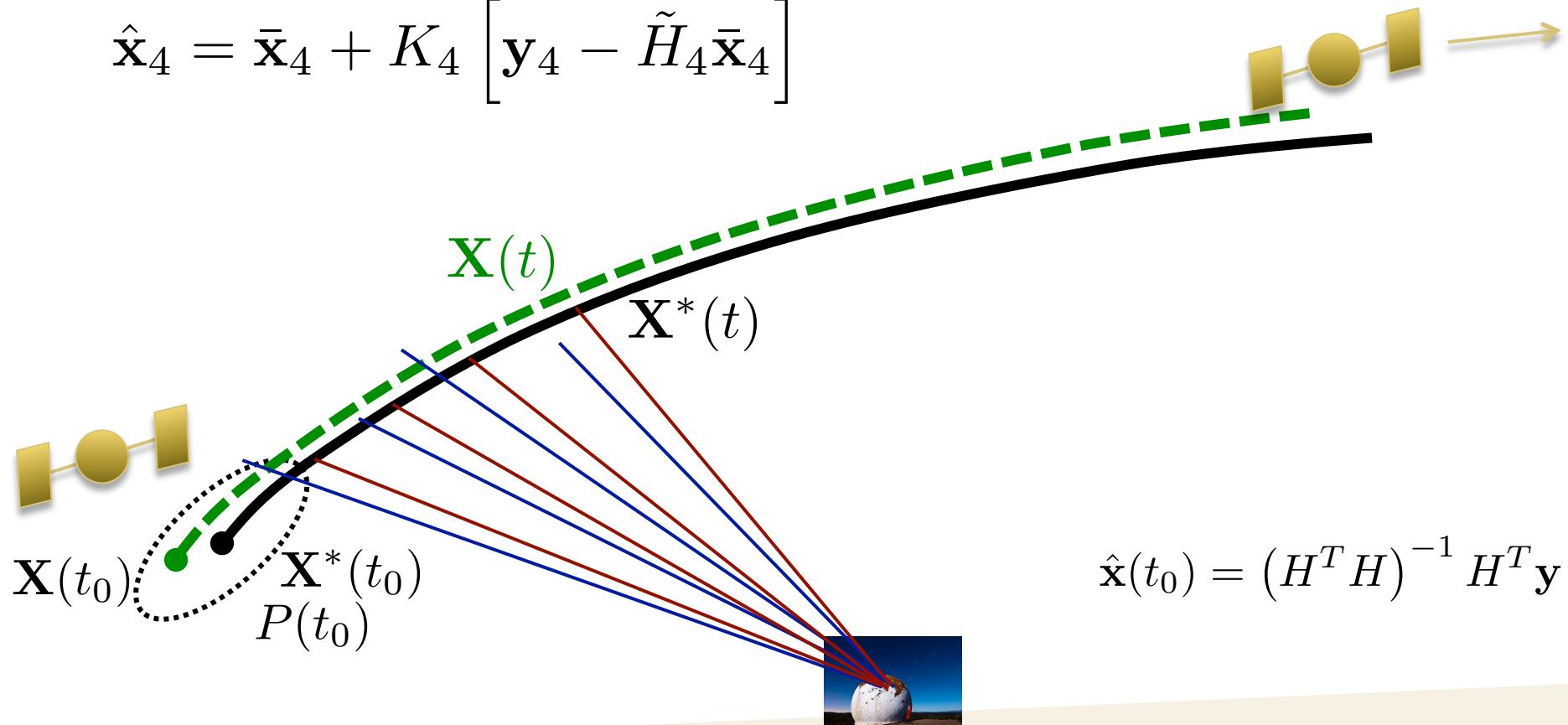
$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



# Kalman Filter

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_4 = \bar{\mathbf{x}}_4 + K_4 \left[ \mathbf{y}_4 - \tilde{H}_4 \bar{\mathbf{x}}_4 \right]$$

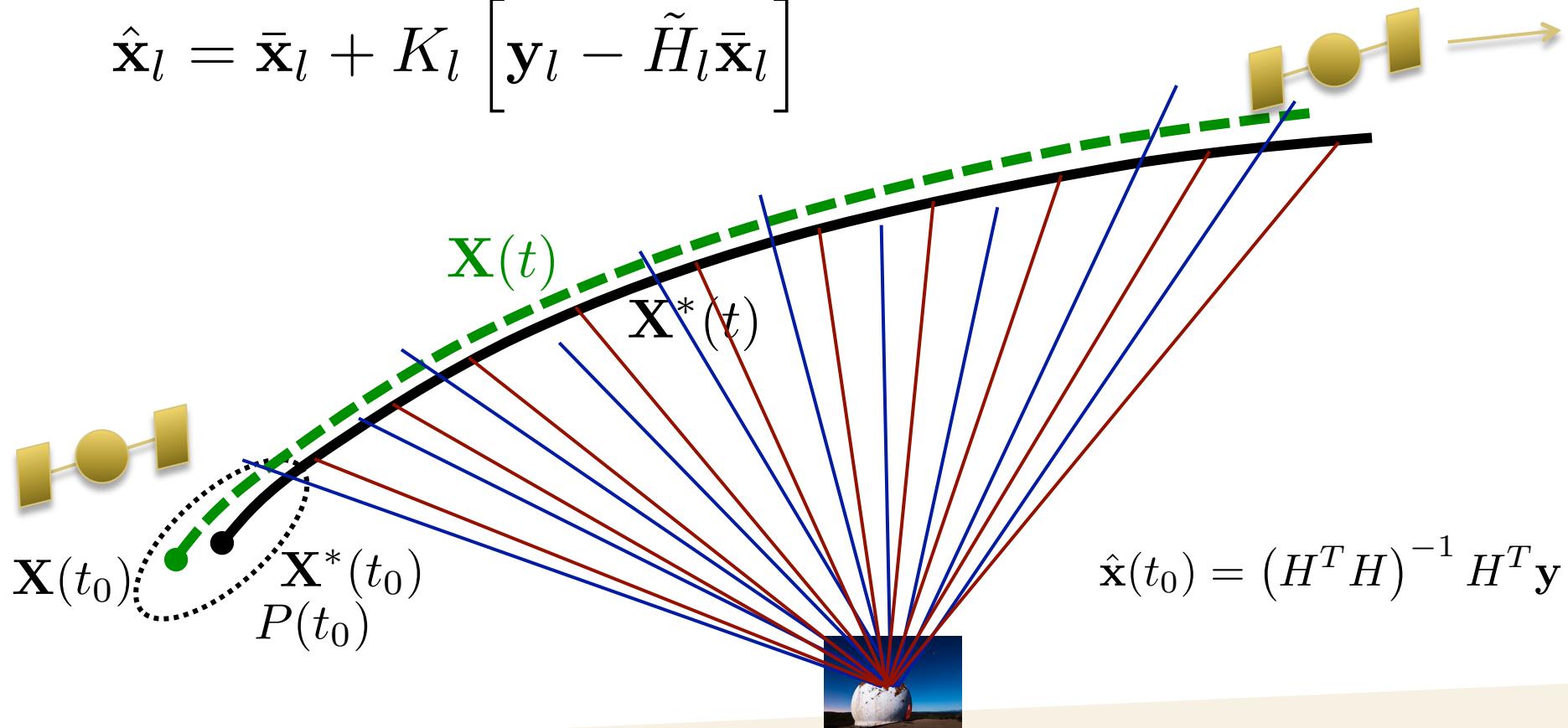


$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$

# Kalman Filter

- ▶ Collect mapped information

$$\hat{\mathbf{x}}_l = \bar{\mathbf{x}}_l + K_l \left[ \mathbf{y}_l - \tilde{H}_l \bar{\mathbf{x}}_l \right]$$



$$\hat{\mathbf{x}}(t_0) = (H^T H)^{-1} H^T \mathbf{y}$$



University of Colorado  
Boulder

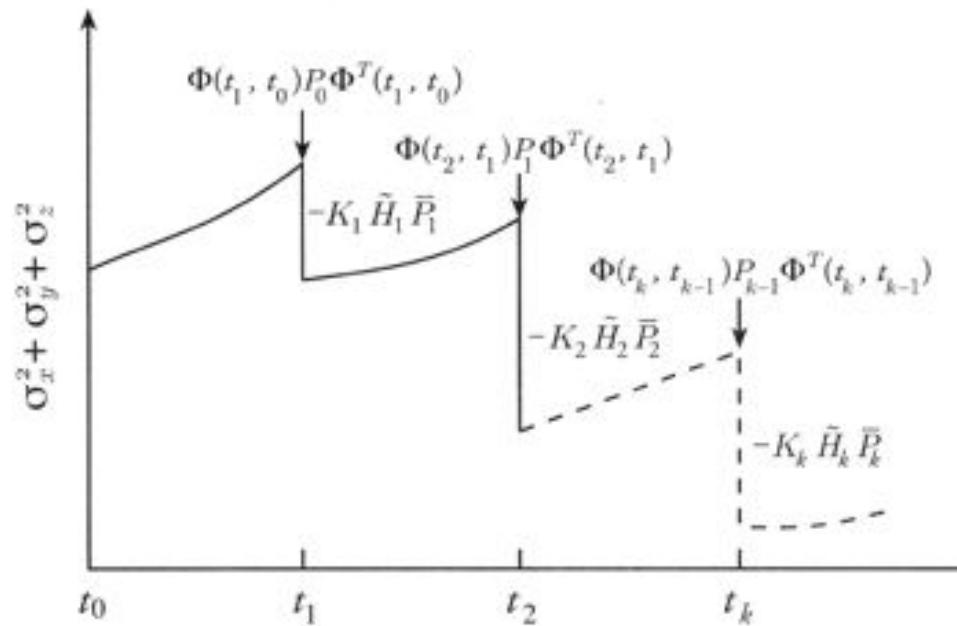
# Kalman Filter

- ▶ Evolution of the covariance matrix as observations are processed.
  - Q: How do you imagine it would change?
  - Q: What would cause it to shrink? To grow?



# Kalman Filter

- ▶ Evolution of the covariance matrix as observations are processed.



# Kalman Filter History

- ▶ Least squares estimation began with Gauss
- ▶ 1963: Kalman's sequential approach
  - Introduced minimum variance
  - Introduced process noise
  - Permitted covariance analyses without data
- ▶ Schmidt proposed a linearization method that would work for OD problems
  - Supposed that linearizing around the best estimate trajectory is better than linearizing around the nominal trajectory
- ▶ 1970: Extended Kalman Filter
- ▶ Gradually, researchers identified problems.
  - (a) Divergence due to the use of incorrect *a priori* statistics and unmodeled parameters.
  - (b) Divergence due to the presence of nonlinearities.
  - (c) Divergence due to the effects of computer round-off.



# Kalman Filter History

- ▶ Numerical issues cause the covariance matrix to lose their symmetry and nonnegativity
- ▶ Possible corrections:
  - (a) Compute only the upper (or lower) triangular entries and force symmetry
  - (b) Compute the entire matrix and then average the upper and lower fields
  - (c) Periodically test and reset the matrix
  - (d) Replace the optimal Kalman measurement update by other expressions (Joseph, Potter, etc)
  - (e) Use larger process noise and measurement noise covariances.



# Kalman Filter History

- ▶ Potter is credited with introducing square root factorization.
  - Worked for the Apollo missions!
- ▶ 1968: Andrews extended Potter's algorithms to include process noise and correlated measurements.
- ▶ 1965 – 1969: Development of the Householder transformation
  - Worked for Mariner 9 in 1971!
- ▶ 1969: Dyer–McReynolds filter added additional process noise effects.
  - Worked for Mariner 10 in 1973 for Venus and Mercury!



# Final Statements

---

- ▶ Homework 5 due Today
  - ▶ Exam on 10/11.
    - Eduardo and/or Paul will be reviewing subjects on Tuesday – send them emails with questions/subjects that you'd like them to cover.
    - 1 hour, open book, open notes.
    - Topics: Definitions of variables, Probability/Statistics Observability, Linearization Least squares, Batch processor
- |                          |                      |                         |
|--------------------------|----------------------|-------------------------|
| $\mathbf{X}(t_i),$       | $\mathbf{X}^*(t_i),$ | $\hat{\mathbf{X}}(t_i)$ |
| $\mathbf{x}(t_i),$       | $\mathbf{x}^*(t_i),$ | $\hat{\mathbf{x}}(t_i)$ |
| $\mathbf{Y}(t_i),$       | $\epsilon(t_i),$     | $\hat{\epsilon}(t_i)$   |
| $\Phi(t_i, t_j),$        | $H(t_i),$            | $\tilde{H}(t_i)$        |
| $A(t_i),$                | $P(t_i),$            | $W(t_i)$                |
| $\bar{\mathbf{x}}(t_i),$ | $\bar{P}(t_i),$      | $\bar{W}(t_i)$          |

