

# Realizing Reduced & Sparse Biochemical Reaction Networks from Dynamics

**Maurice Filo**

**Mustafa Khammash**

**Department of Biosystems Science and Engineering**

December 10, 2025

**ETH** zürich

**C·T·S·B**  
CONTROL THEORY & SYSTEMS BIOLOGY

# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)

# Motivation & Goal

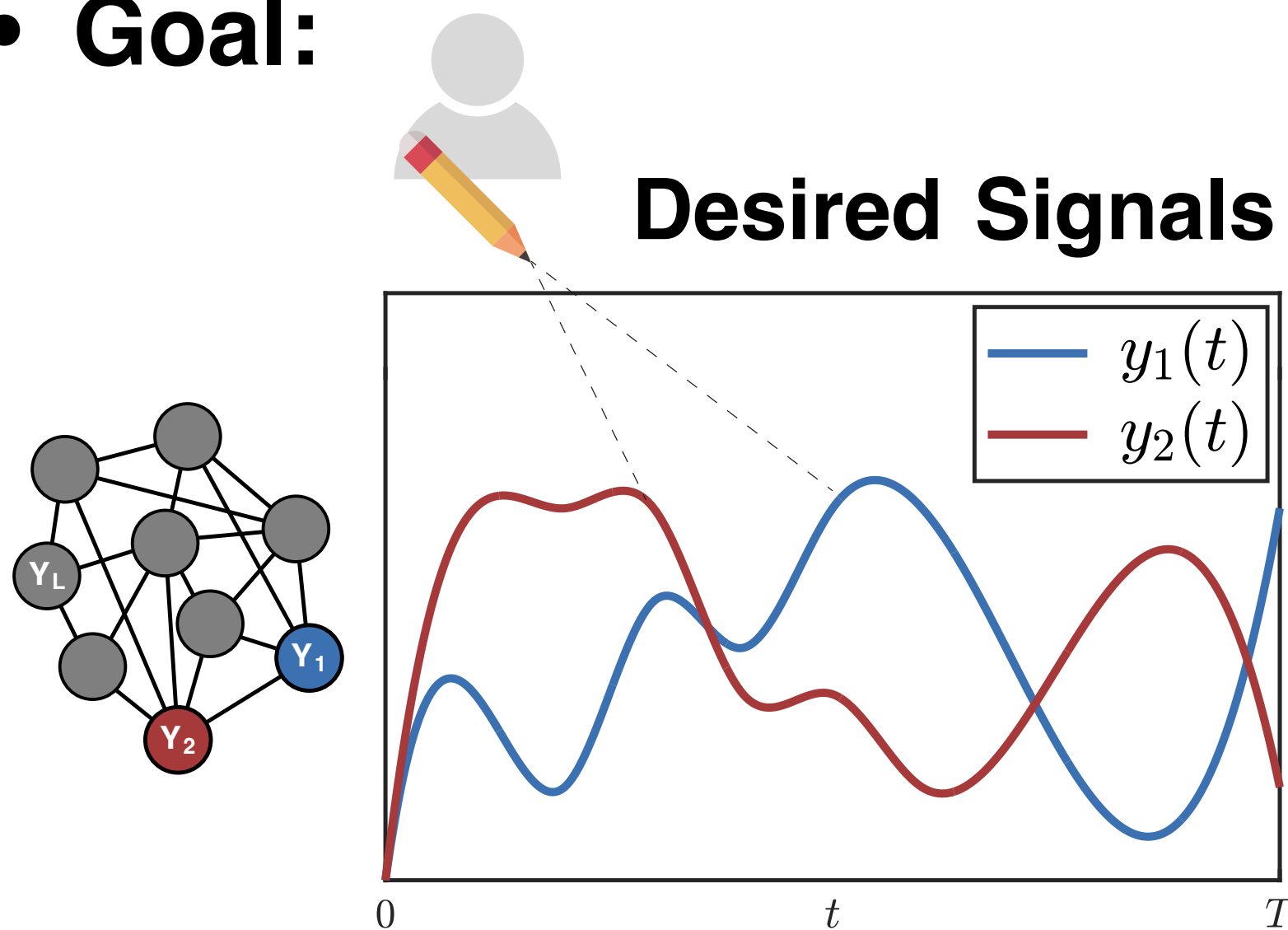
- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available



# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**

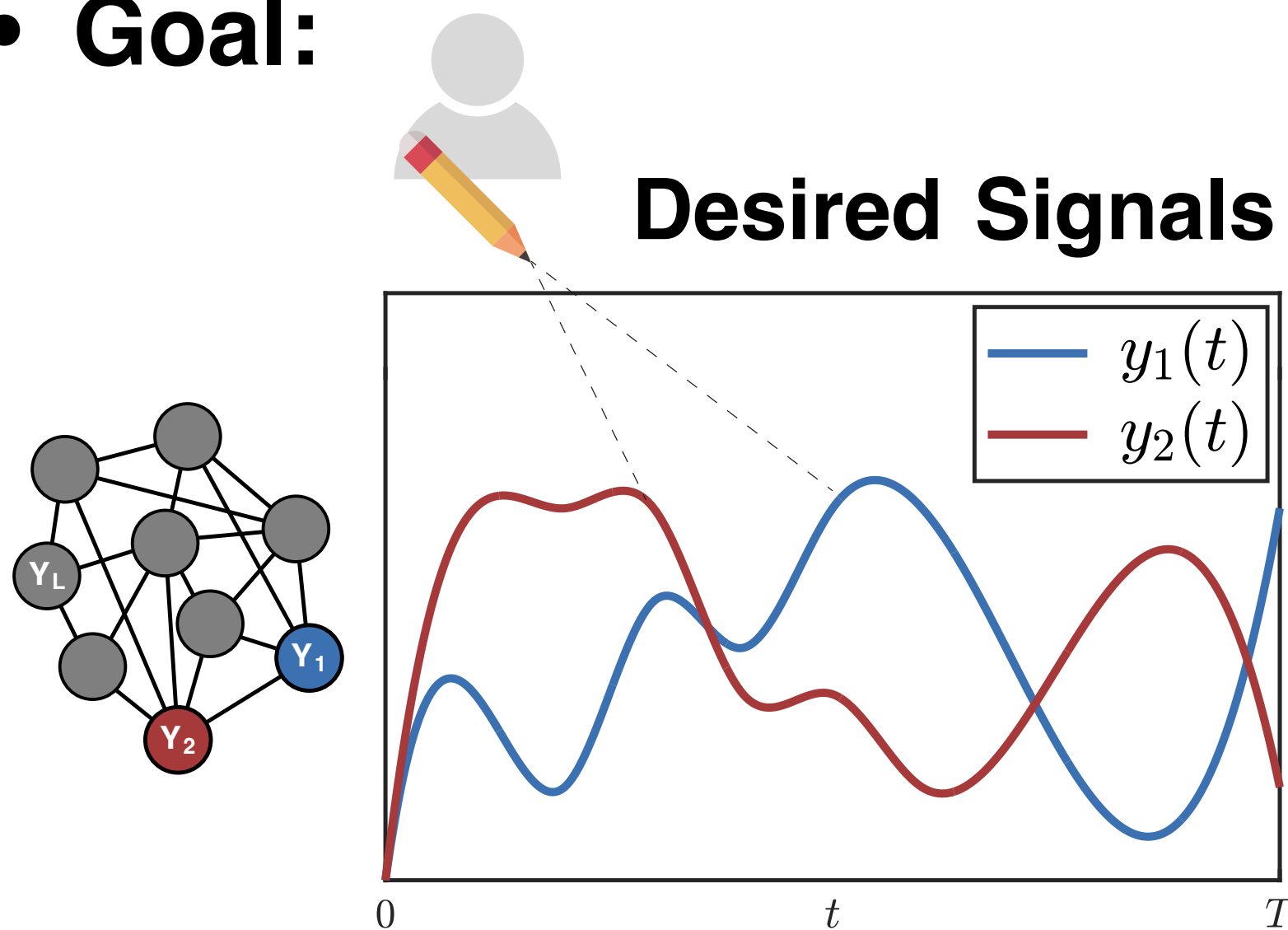




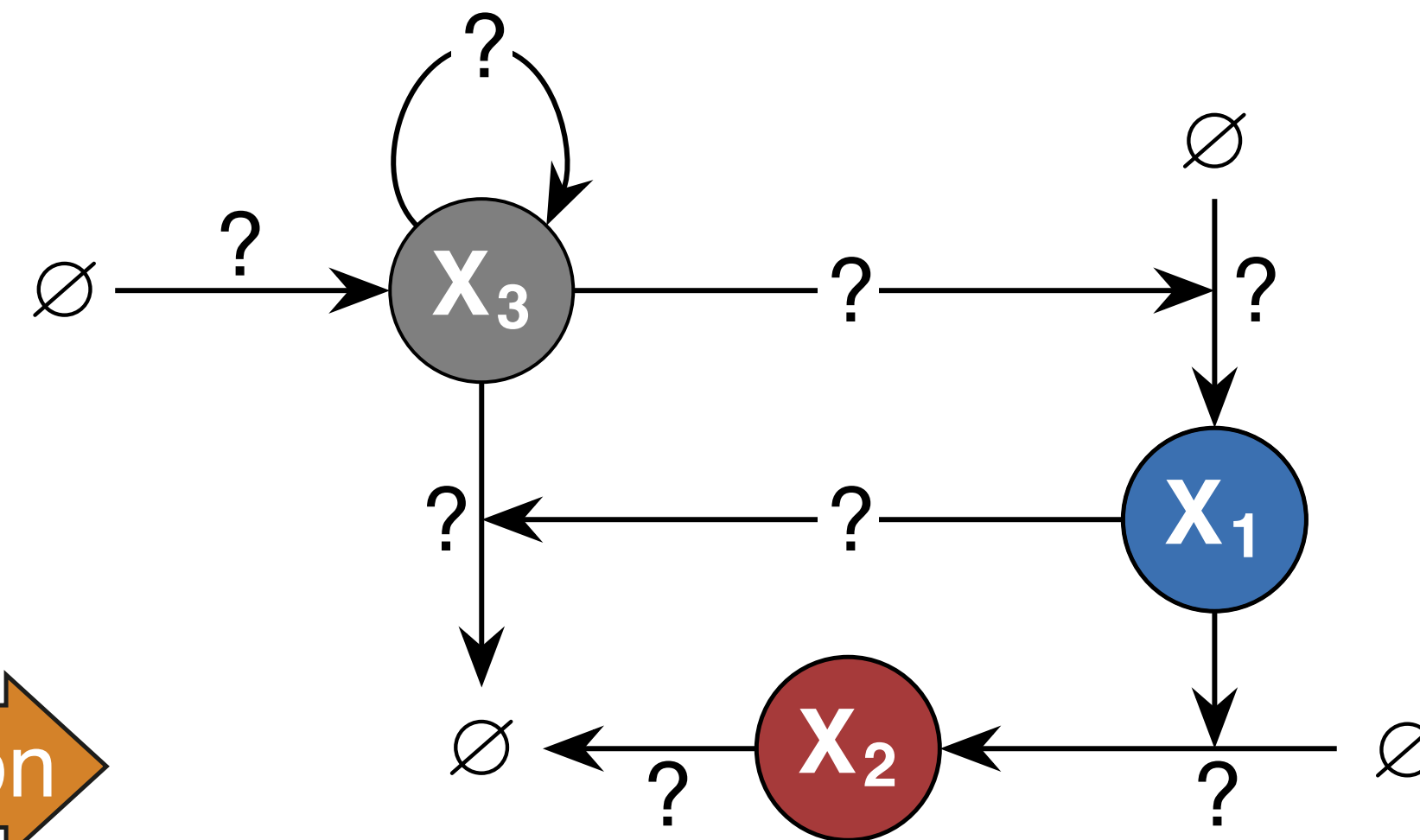
# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**



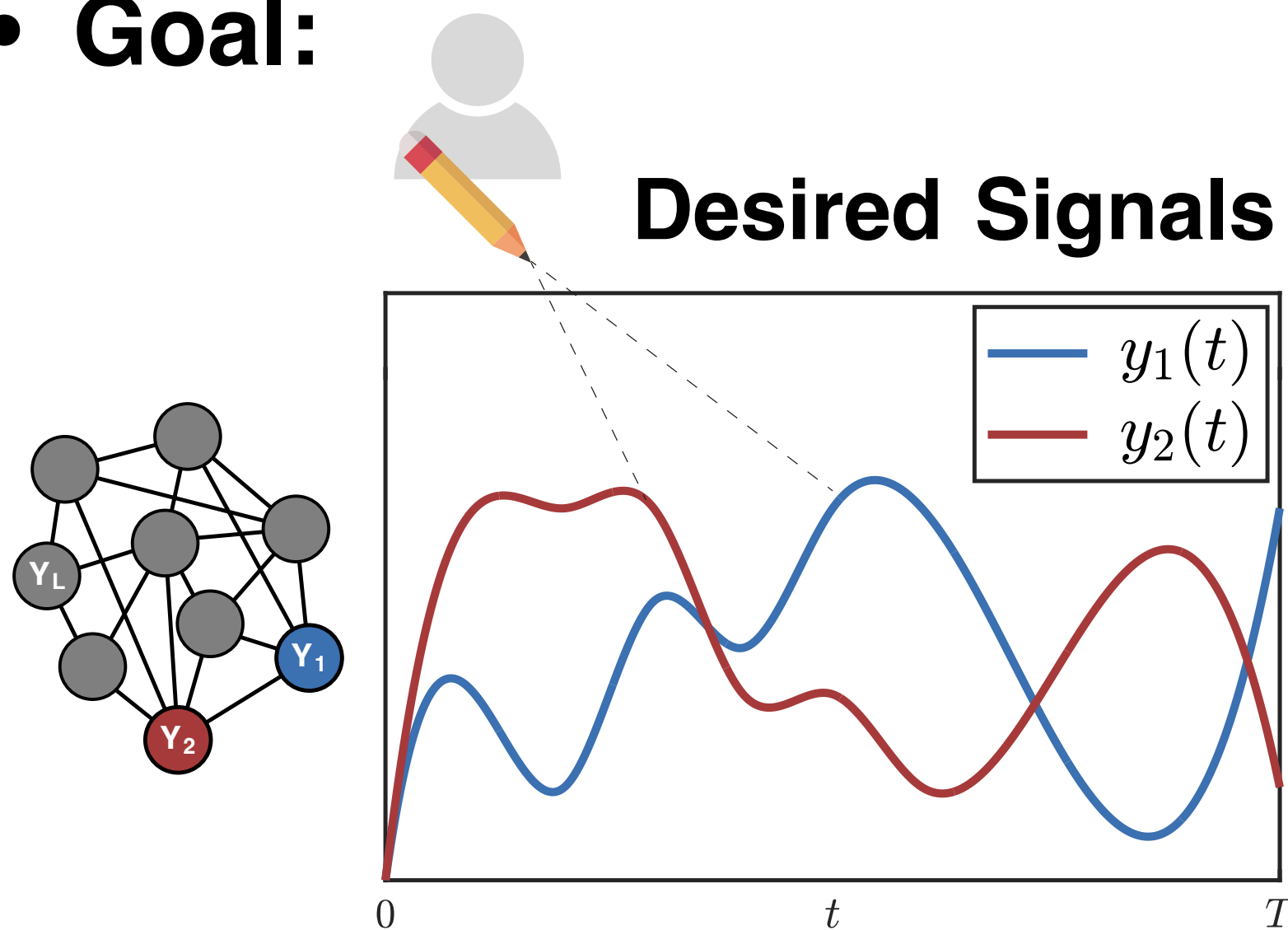
Realization



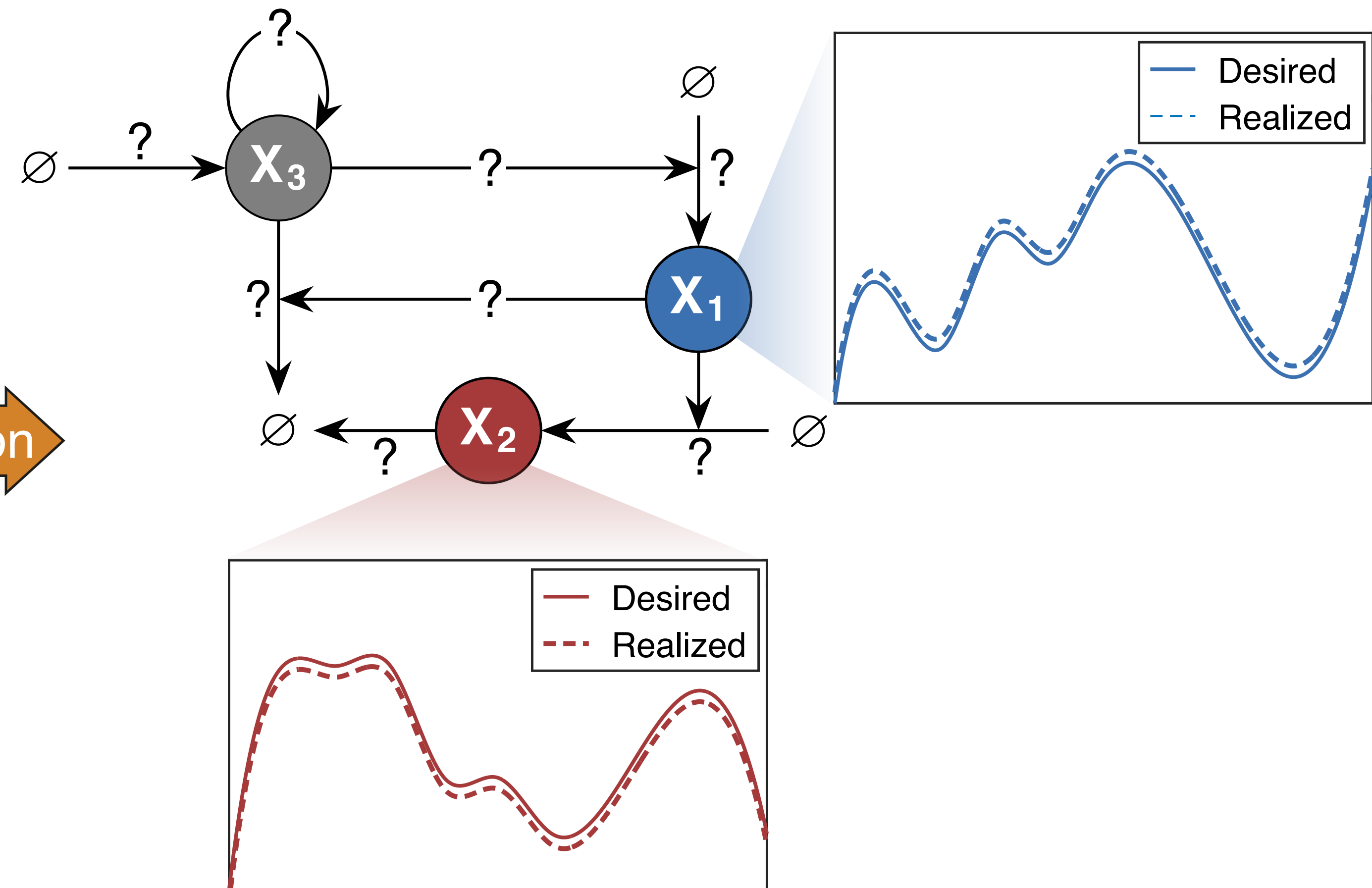
# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**



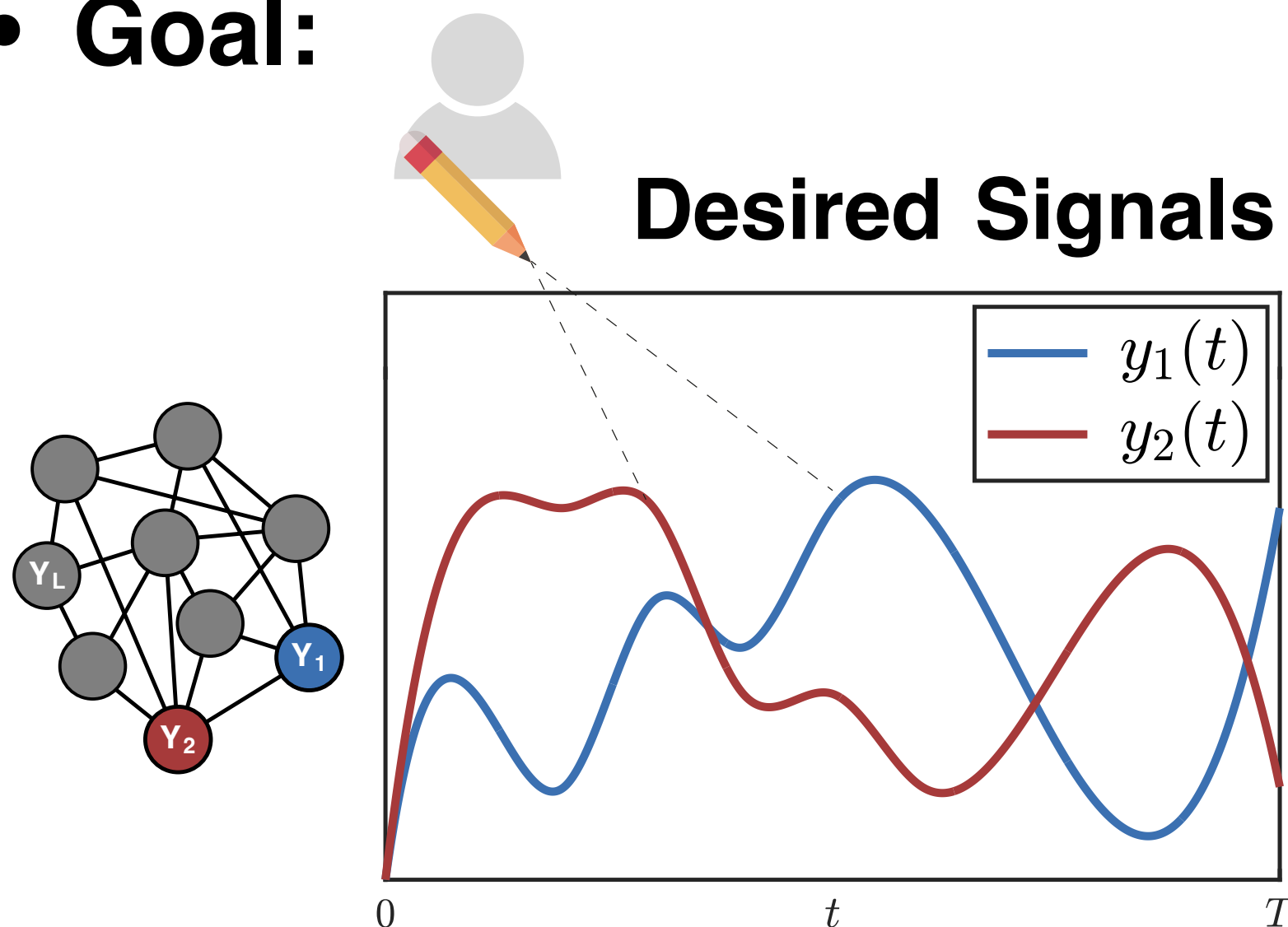
Realization



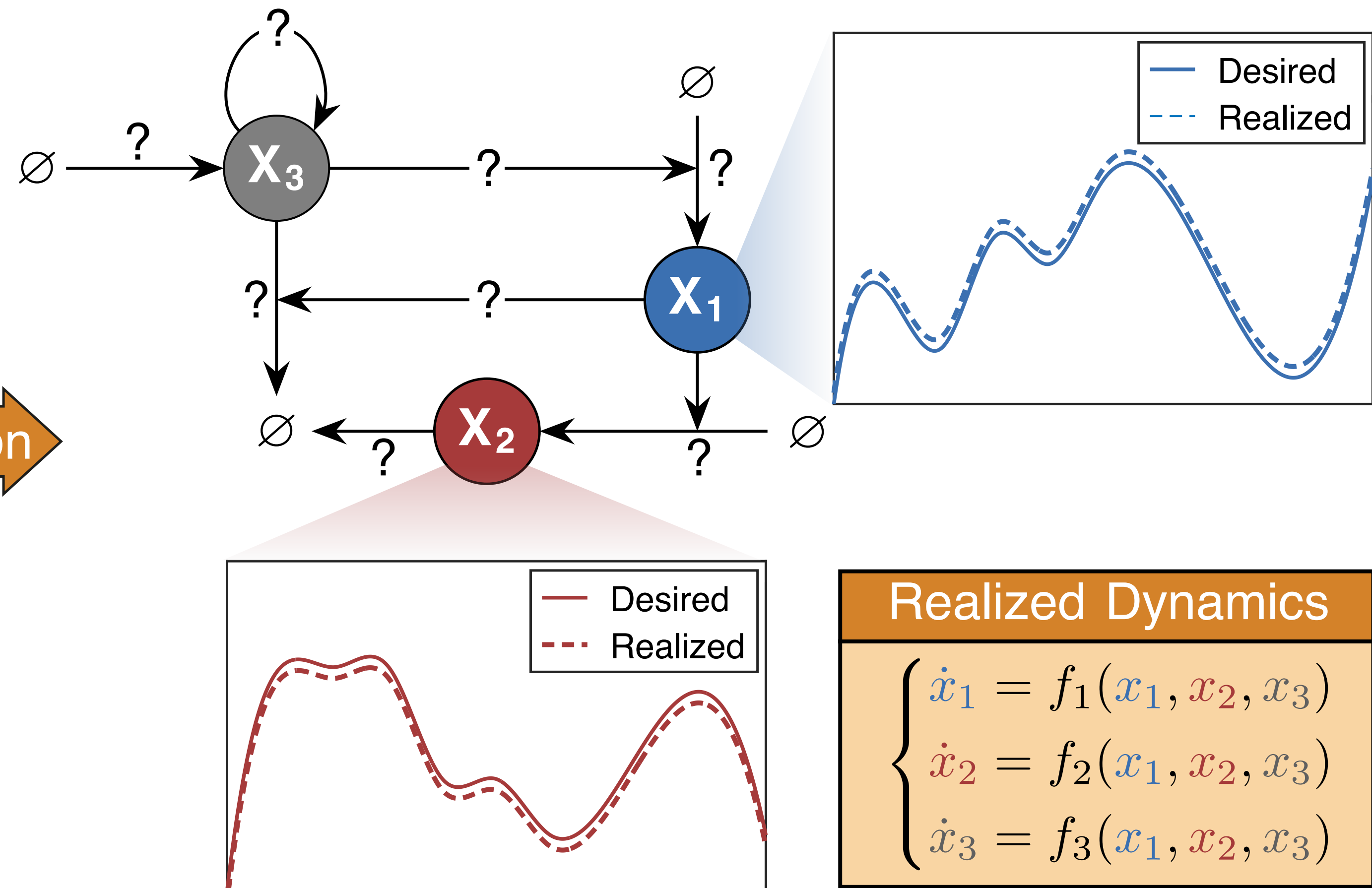
# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**



Realization

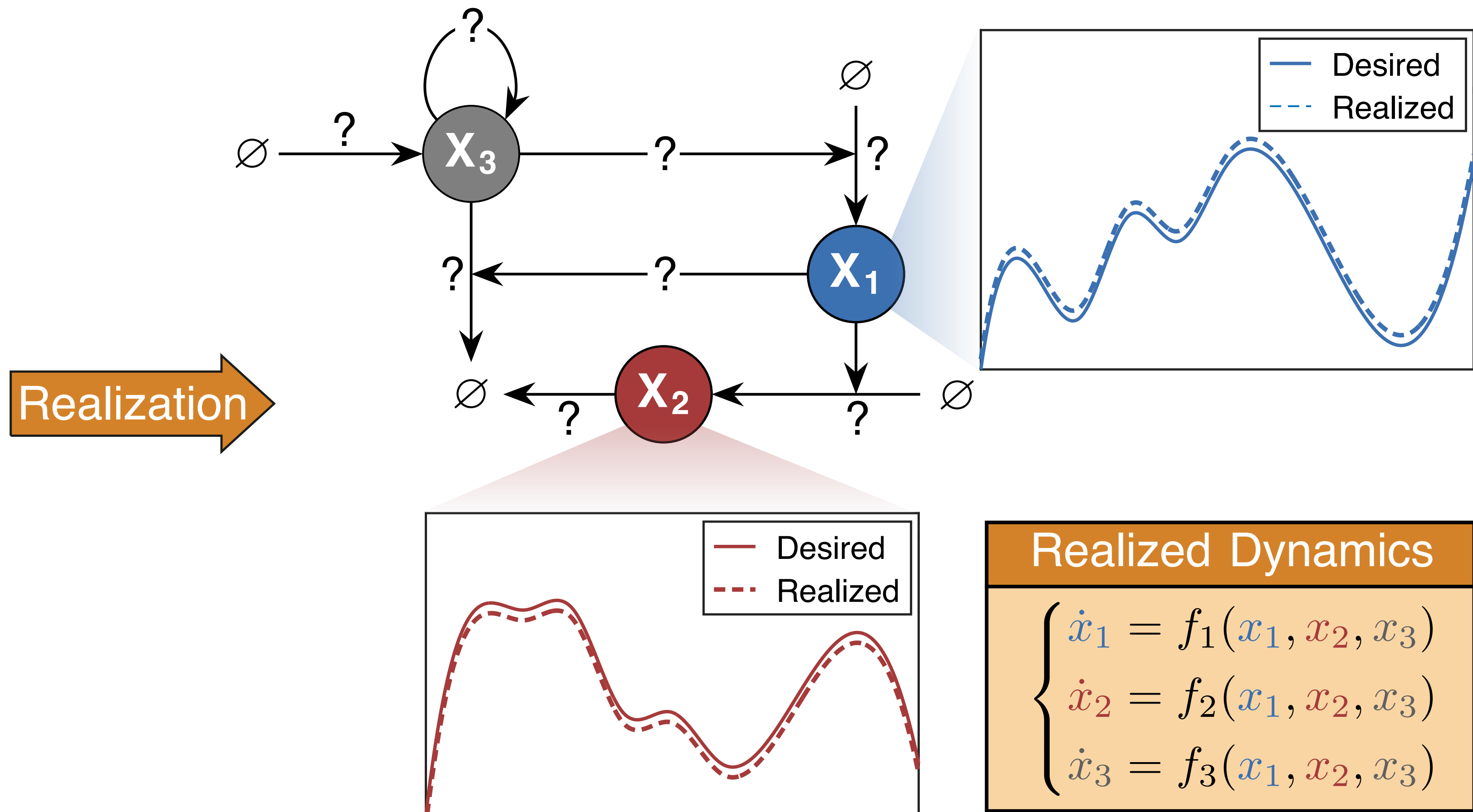
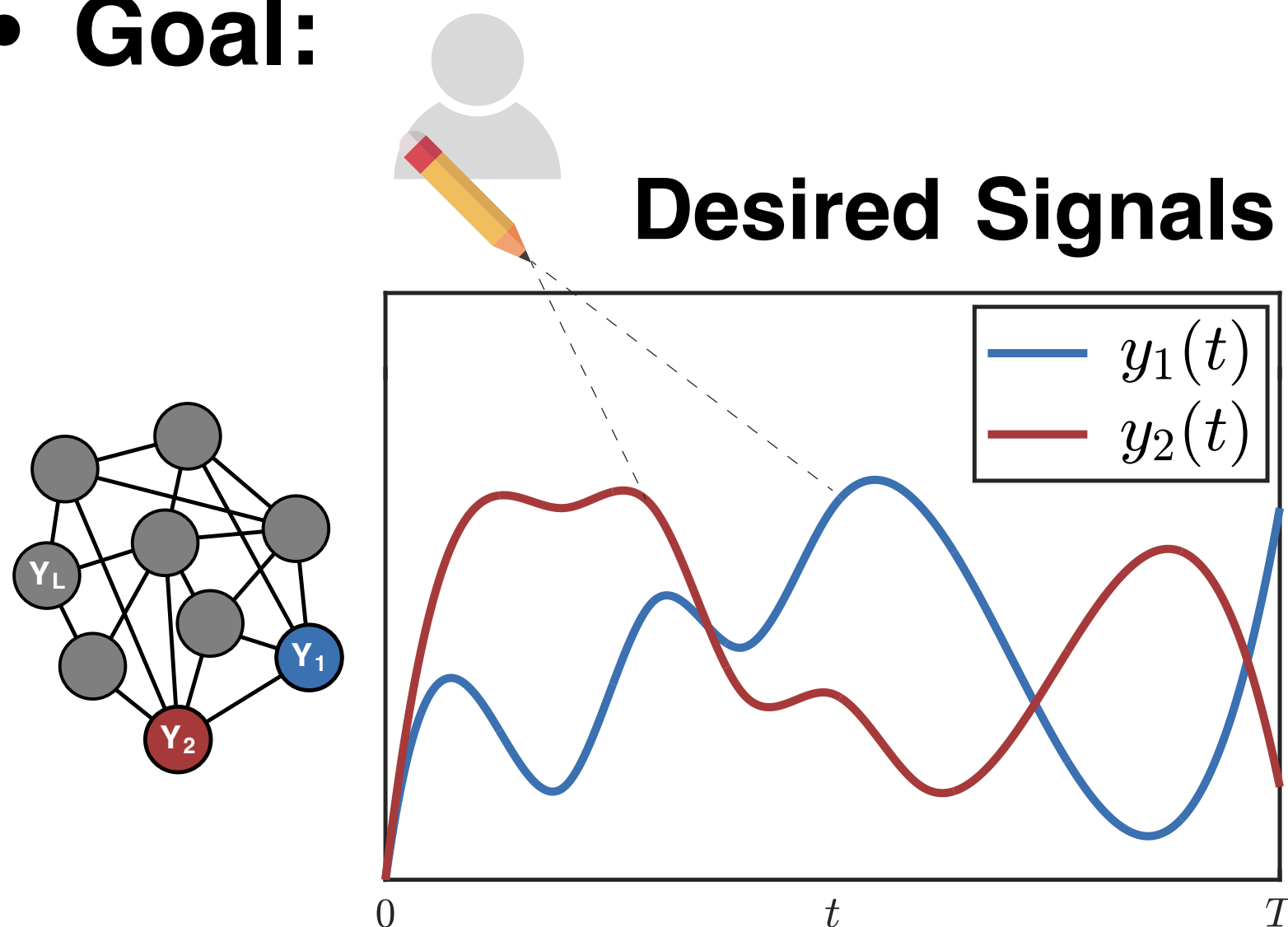




# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**

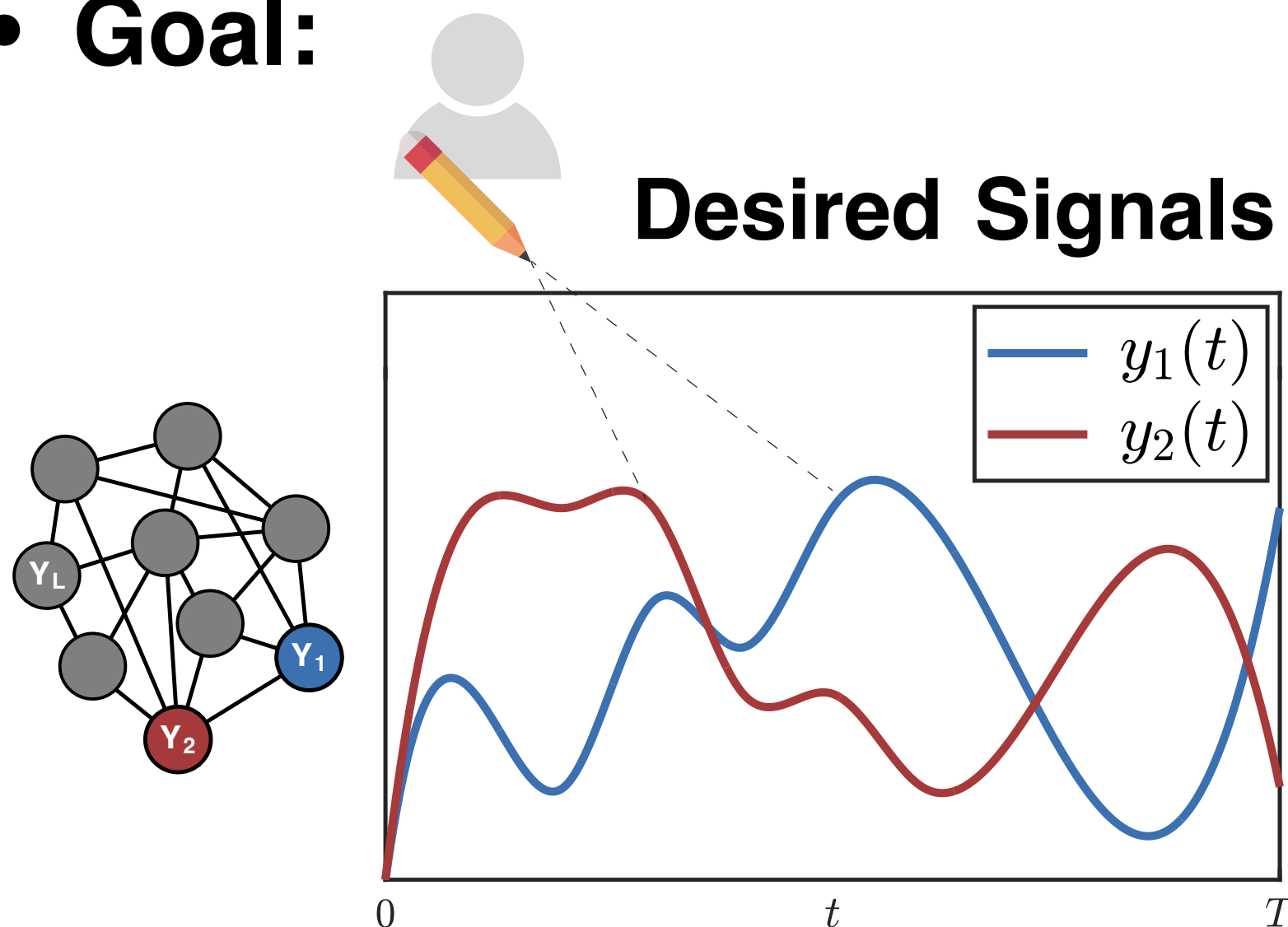


$f_i$  biochemically meaningful

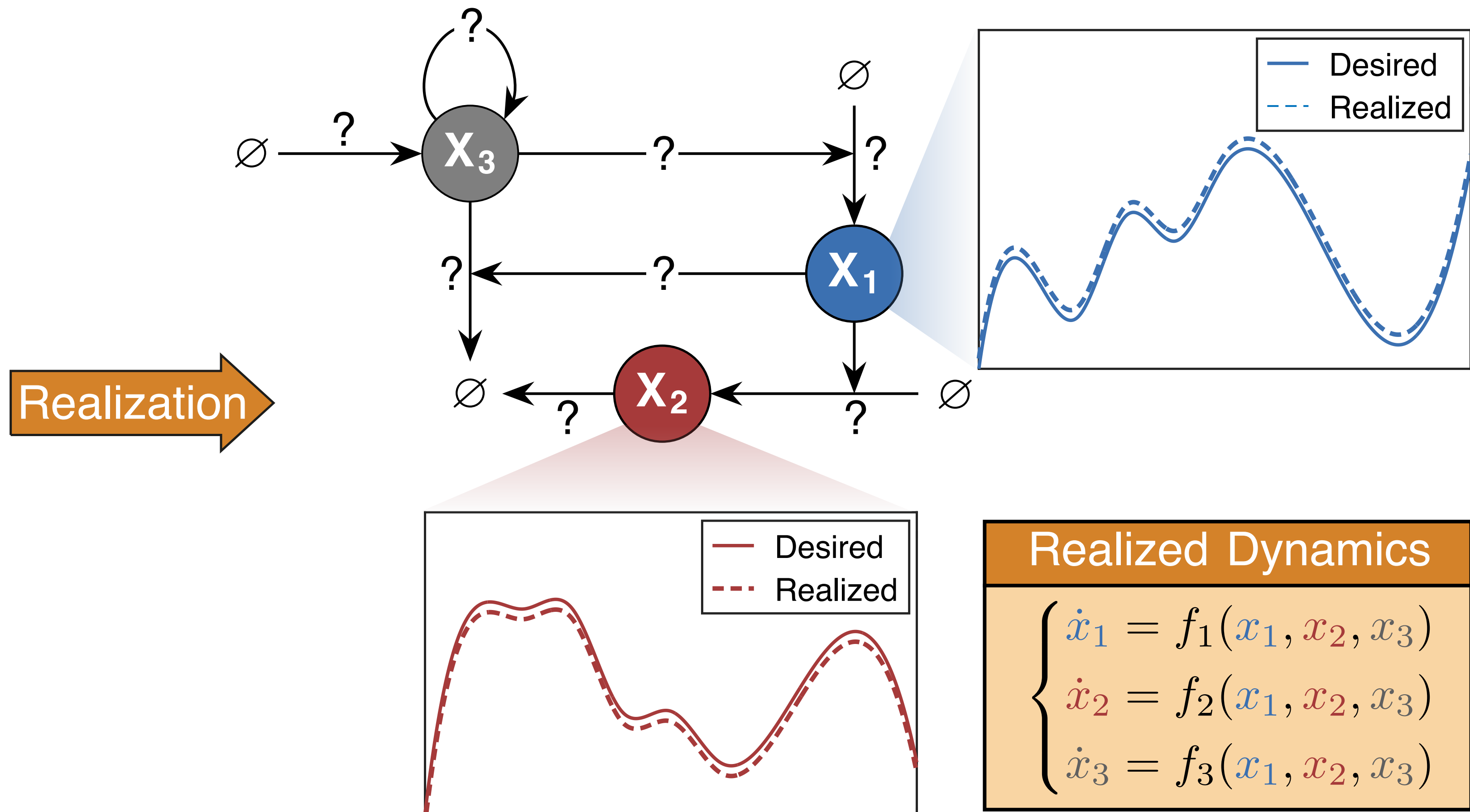
# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- Goal:**



- Reduced order: few species



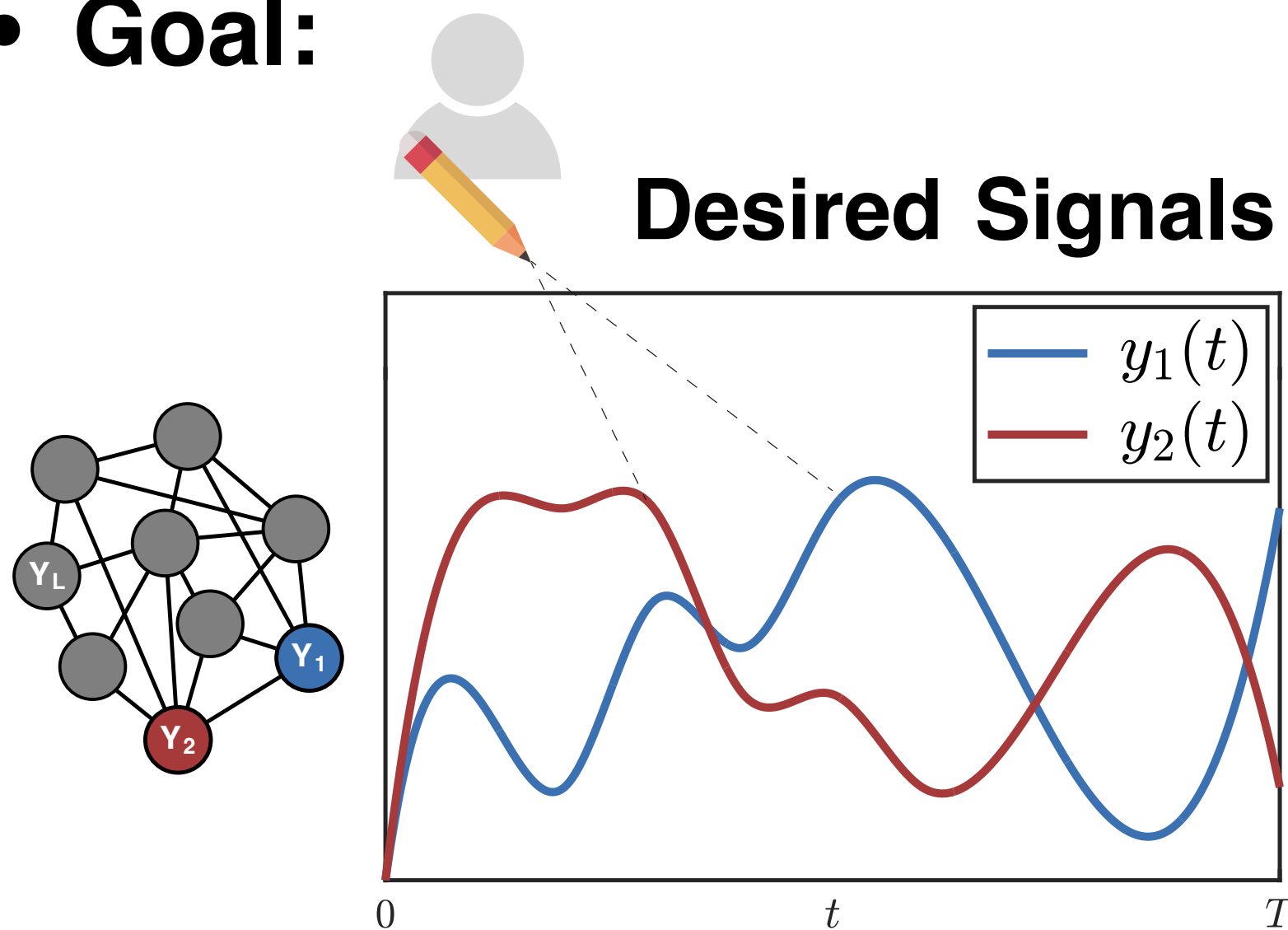
$f_i$  biochemically meaningful



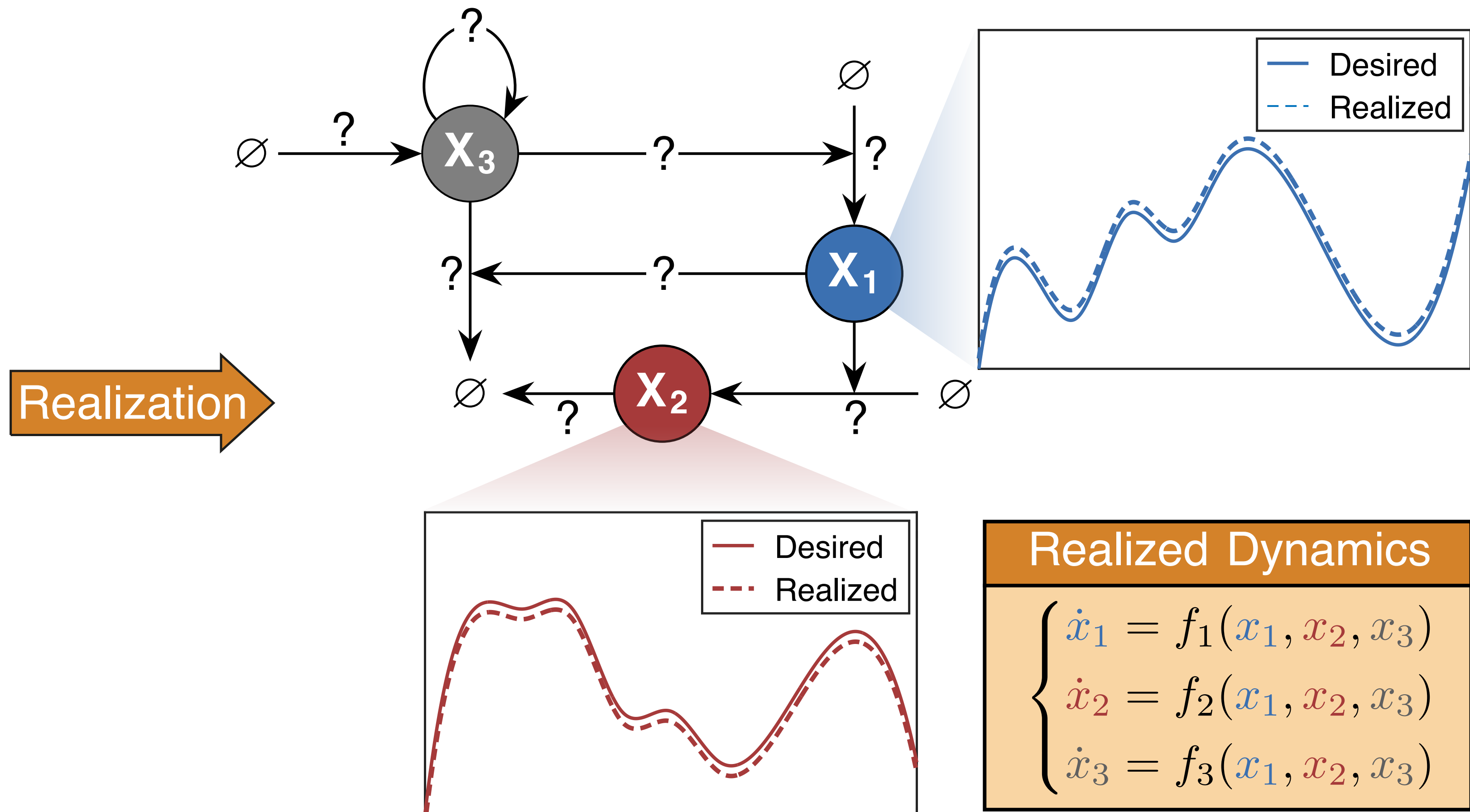
# Motivation & Goal

- CRNs model a wide range of dynamical processes (systems/synthetic biology, epidemiology ...)
- In practice:
  - CRN is unknown
  - time-series of a subset of species is available

- **Goal:**



- Reduced order: few species
- Sparse: few reactions



**Realized Dynamics**

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, x_3) \\ \dot{x}_2 = f_2(x_1, x_2, x_3) \\ \dot{x}_3 = f_3(x_1, x_2, x_3) \end{cases}$$

$f_i$  biochemically meaningful

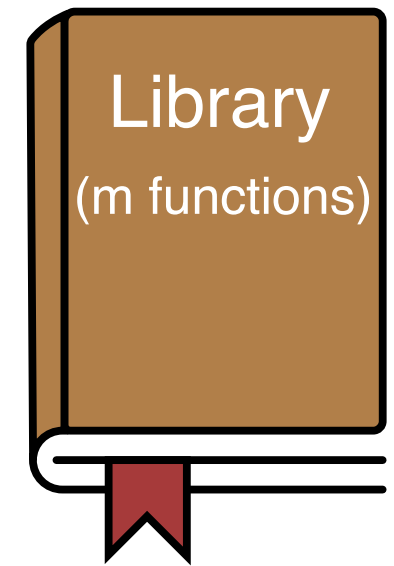
# Indirect (Derivative-based) Approaches

# Indirect (Derivative-based) Approaches

**Model class:**

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \dots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x)$$

$$\{\phi_1(\cdot), \dots, \phi_m(\cdot)\}$$



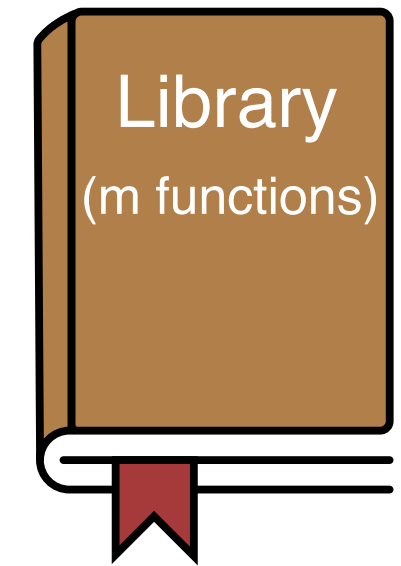


# Indirect (Derivative-based) Approaches

**Model class:**

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix}$$

$\{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$

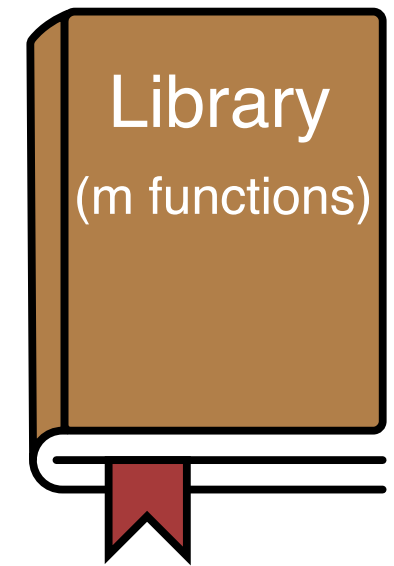


# Indirect (Derivative-based) Approaches

**Model class:**

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \dots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \dots, \phi_m(\cdot)\}$$

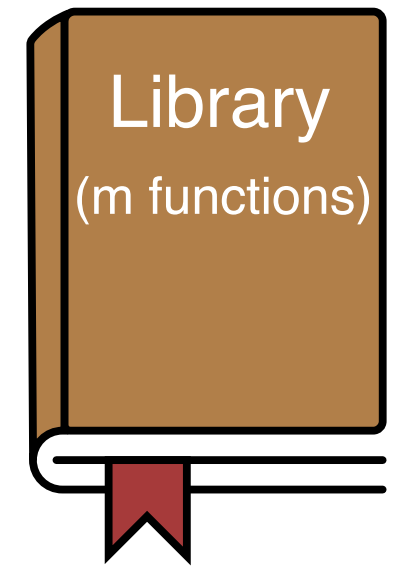
**Pipeline:**



# Indirect (Derivative-based) Approaches

**Model class:**

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



**Pipeline:**

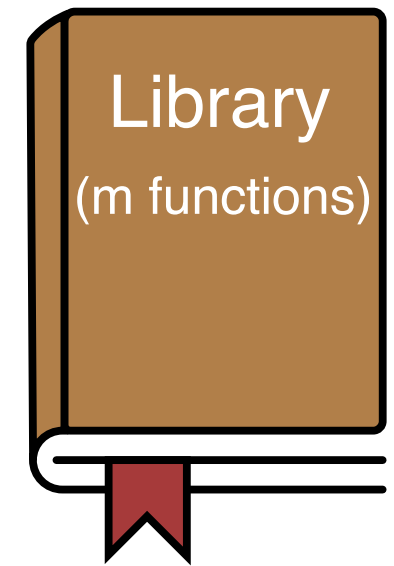
1. Desired trajectory:  $x_T(t); \quad \text{for } t \in [0, T]$



# Indirect (Derivative-based) Approaches

## Model class:

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



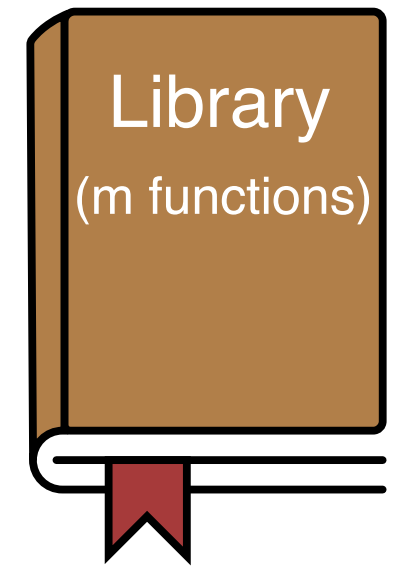
## Pipeline:

1. Desired trajectory:  $x_T(t)$ ; for  $t \in [0, T]$
2. Numerically estimate:  $\dot{x}_T(t)$

# Indirect (Derivative-based) Approaches

## Model class:

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



## Pipeline:

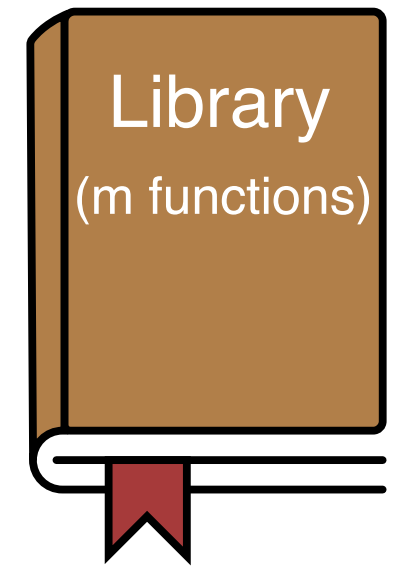
1. Desired trajectory:  $x_T(t)$ ; for  $t \in [0, T]$
2. Numerically estimate:  $\dot{x}_T(t)$
3. Solve:  $\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt$

Bamieh et al. "On discovering low order models in biochemical reaction kinetics." 2007 ACC.

# Indirect (Derivative-based) Approaches

## Model class:

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



## Pipeline:

1. Desired trajectory:  $x_T(t)$ ; for  $t \in [0, T]$

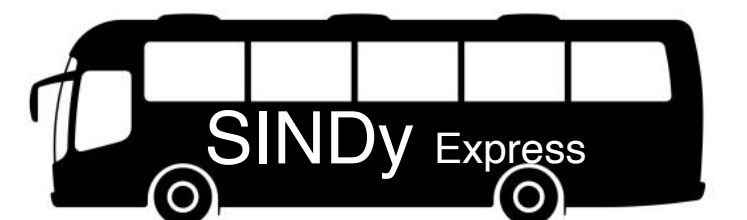
2. Numerically estimate:  $\dot{x}_T(t)$

3. Solve:  $\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt$

Bamieh et al. "On discovering low order models in biochemical reaction kinetics." *2007 ACC*.

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1$$

Brunton et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." *PNAS* (2016).

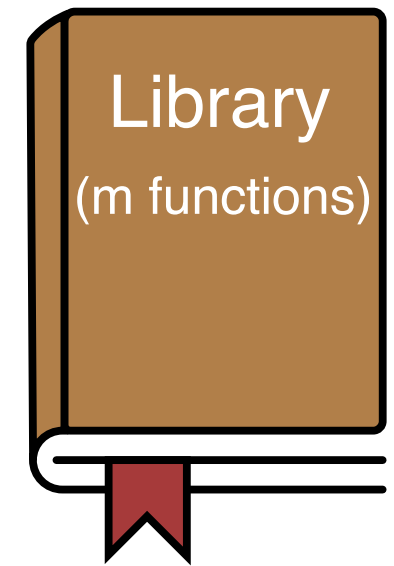




# Indirect (Derivative-based) Approaches

## Model class:

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



## Pipeline:

1. Desired trajectory:  $x_T(t)$ ; for  $t \in [0, T]$

2. Numerically estimate:  $\dot{x}_T(t)$

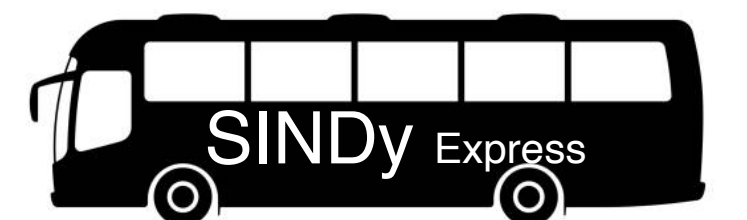
3. Solve:  $\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt$

Bamieh et al. "On discovering low order models in biochemical reaction kinetics." *2007 ACC*.

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1$$

Brunton et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." *PNAS* (2016).

**Pros:** Convex — exploits linearity in parameters

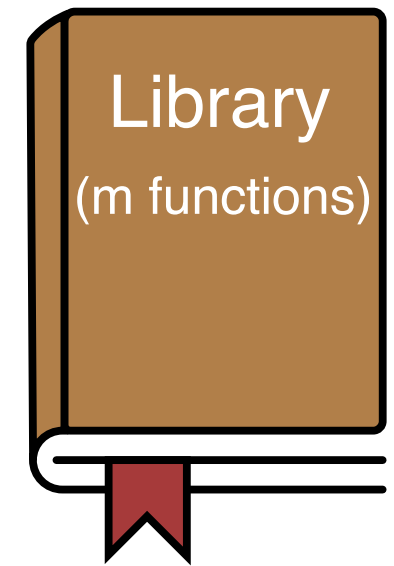




# Indirect (Derivative-based) Approaches

## Model class:

$$\begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} \theta_1 \end{bmatrix} \phi_1(x) + \cdots + \begin{bmatrix} \theta_m \end{bmatrix} \phi_m(x) = \begin{bmatrix} \boldsymbol{\theta} \end{bmatrix} \begin{bmatrix} \Phi(x) \end{bmatrix} \quad \{\phi_1(\cdot), \cdots, \phi_m(\cdot)\}$$



## Pipeline:

1. Desired trajectory:  $x_T(t)$ ; for  $t \in [0, T]$

2. Numerically estimate:  $\dot{x}_T(t)$

3. Solve:  $\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt$

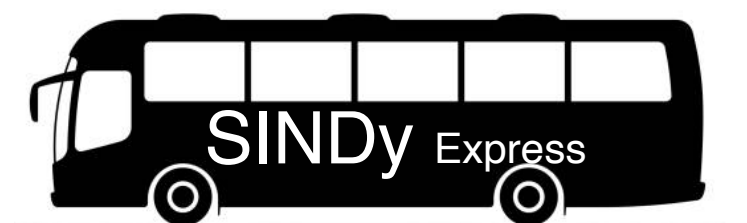
Bamieh et al. "On discovering low order models in biochemical reaction kinetics." *2007 ACC*.

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \boldsymbol{\theta} \Phi(x_T(t)) \right\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1$$

Brunton et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." *PNAS* (2016).

**Pros:** Convex — exploits linearity in parameters

**Cons:** Fits derivatives not trajectories



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta\Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta\Phi(x_T(t)) - \dot{x}_T(t)$



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:  
$$e_D(t; \theta) \triangleq \underset{\text{solution of ODE}}{\overset{\uparrow}{x(t; \theta)}} - \underset{\text{desired trajectory}}{\overset{\uparrow}{x_T(t)}}$$

# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:  $e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$   

solution of ODE  
↑

desired trajectory  
↑

But ... how are they related?

# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:  
$$e_D(t; \theta) \triangleq \overset{\text{solution of ODE}}{\underset{\uparrow}{x(t; \theta)}} - \overset{\text{desired trajectory}}{\underset{\uparrow}{x_T(t)}}$$

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_{\theta}x + b_{\theta}, \quad x(0) = x_T(0)$$



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:  
$$e_D(t; \theta) \triangleq \overset{\text{solution of ODE}}{\underset{\uparrow}{x(t; \theta)}} - \overset{\text{desired trajectory}}{\underset{\uparrow}{x_T(t)}}$$

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta\Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta\Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

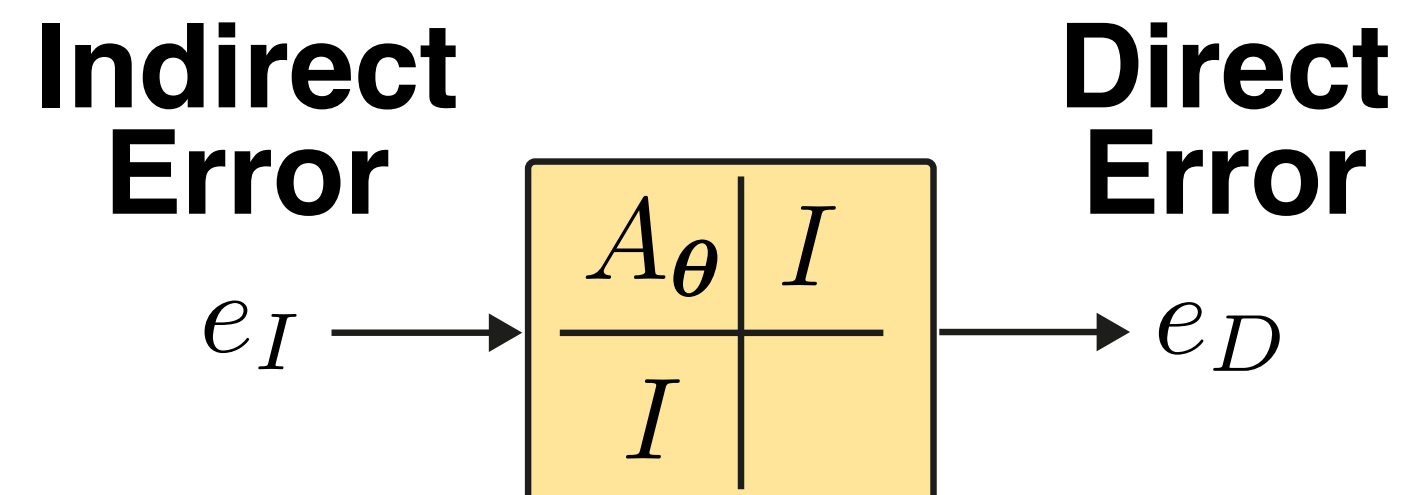
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\xrightarrow{\text{red arrow}}$  desired trajectory  $\xleftarrow{\text{blue arrow}}$

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta\Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta\Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

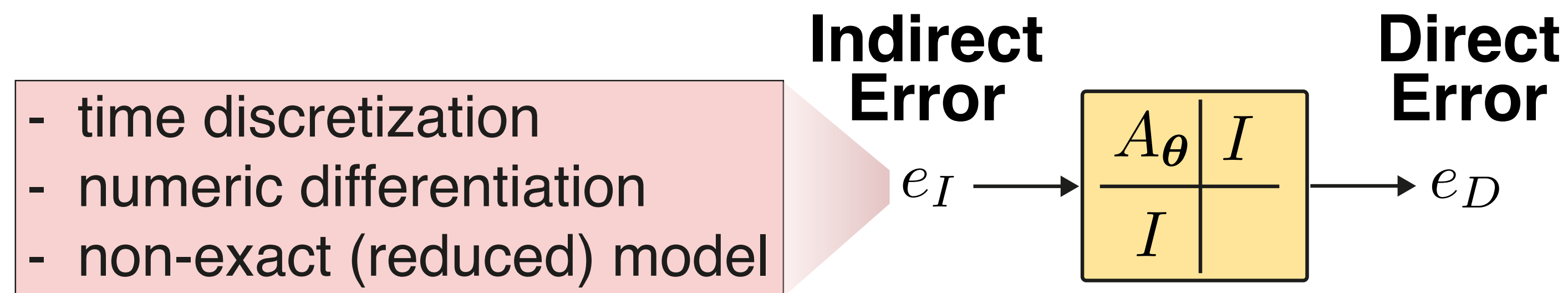
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\nearrow$  desired trajectory

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

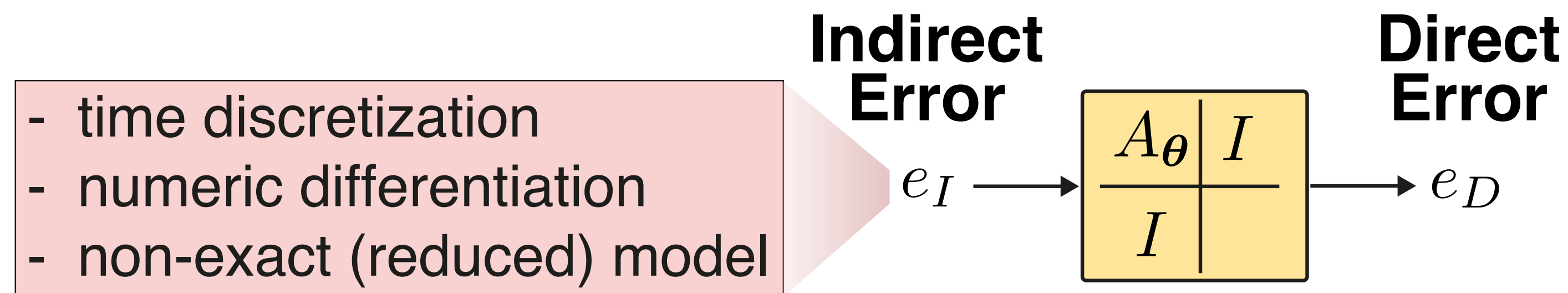
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\nearrow$  desired trajectory

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



$\implies$  Indirect error propagates through the dynamics supposed to be realized



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

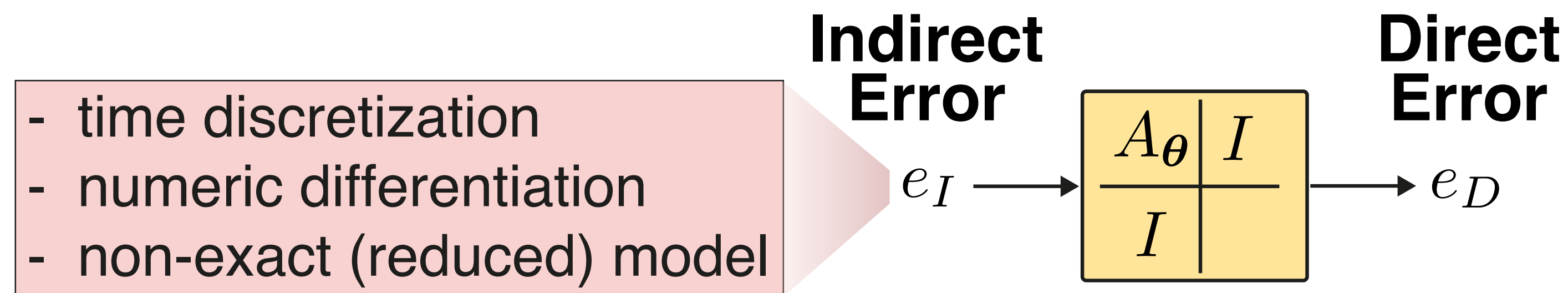
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\nearrow$  desired trajectory

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



careful when realizing  
(near) unstable dynamics

$\Rightarrow$  Indirect error propagates through the dynamics supposed to be realized

# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

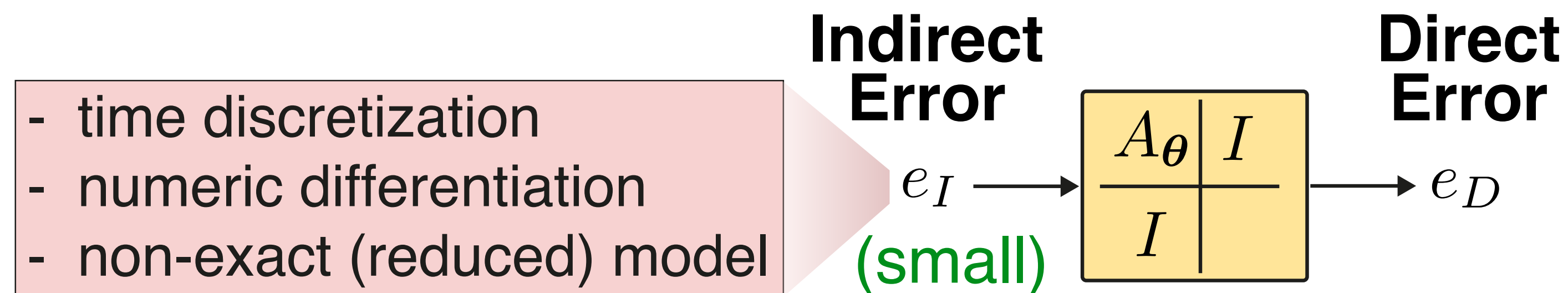
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\xrightarrow{\text{red arrow}}$  desired trajectory  $\xleftarrow{\text{blue arrow}}$

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



careful when realizing  
(near) unstable dynamics

$\Rightarrow$  Indirect error propagates through the dynamics supposed to be realized



# Price of Convexity: Mismatch between Objective & Goal

**Model class:**  $\dot{x} = \theta \Phi(x)$

**Desired trajectory:**  $\{x_T(t)\}_{0 \leq t \leq T}$

- Indirect objective minimizes ODE residuals @  $\{x_T(t)\}_{0 \leq t \leq T}$ :  $e_I(t; \theta) \triangleq \theta \Phi(x_T(t)) - \dot{x}_T(t)$

- But we really care about is direct trajectory error:

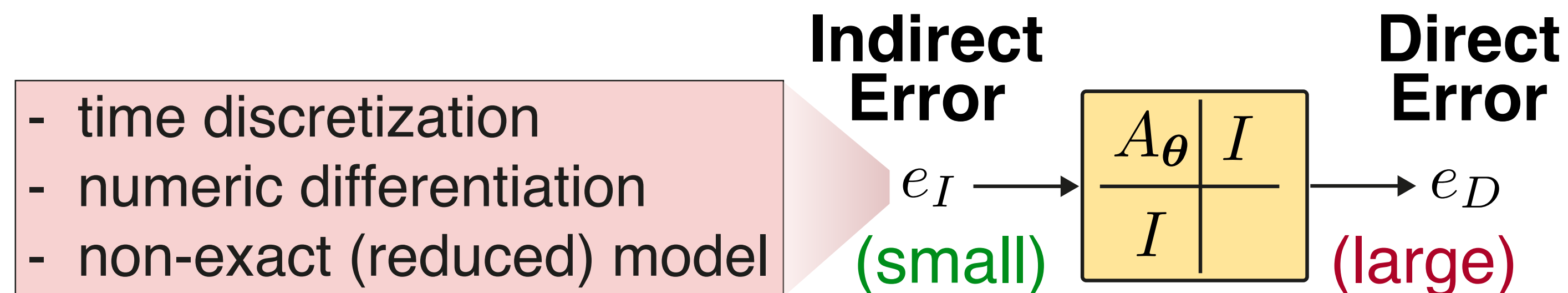
$$e_D(t; \theta) \triangleq x(t; \theta) - x_T(t)$$

solution of ODE  $\xrightarrow{\text{red arrow}}$  desired trajectory  $\xleftarrow{\text{blue arrow}}$

**Example class:** linear-affine dynamics (unimolecular CRNs)

$$\dot{x} = A_\theta x + b_\theta, \quad x(0) = x_T(0)$$

can show that:  $\dot{e}_D = A_\theta e_D + e_I$



careful when realizing  
(near) unstable dynamics

$\Rightarrow$  Indirect error propagates through the dynamics supposed to be realized

# Direct Approach: Target Trajectories, Not Derivatives



# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

↑ Allows adding hidden species to enhance expressiveness

# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

**Direct optimization problem:**

$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$

# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

**Direct optimization problem:**

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \quad \frac{1}{2} \int_0^T \boxed{\|y(t) - y_T(t)\|_2^2 dt} + \boxed{\lambda \|\boldsymbol{\theta}\|_1}$$

direct trajectory error sparsity penalty

subject to 
$$\begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases}$$



# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

**Direct optimization problem:**

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \quad \frac{1}{2} \int_0^T \boxed{\|y(t) - y_T(t)\|_2^2 dt} + \boxed{\lambda \|\boldsymbol{\theta}\|_1}$$

direct trajectory error sparsity penalty

subject to  $\begin{cases} \boxed{\begin{aligned} \dot{x} &= \boldsymbol{\theta} \Phi(x) \\ y &= Cx \end{aligned}} \quad \text{dynamics constraint} \\ \boxed{S \circ \boldsymbol{\theta} \geq 0.} \quad \text{positive dynamics constraint} \end{cases}$

# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

**Direct optimization problem:**

$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{2} \int_0^T \boxed{\|y(t) - y_T(t)\|_2^2 dt} + \boxed{\lambda \|\theta\|_1}$$

direct trajectory error sparsity penalty

subject to  $\begin{cases} \boxed{\begin{aligned} \dot{x} &= \theta \Phi(x) \\ y &= Cx \end{aligned}} \quad \text{dynamics constraint} \\ \boxed{S \circ \theta \geq 0.} \quad \text{positive dynamics constraint} \end{cases}$

**Advantages:**

- No numerical differentiation
- Hidden species enrich modeling capacity
- Directly optimizes what we care about

# Direct Approach: Target Trajectories, Not Derivatives

**Model class:** 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \end{cases}$$

**Desired trajectory:**  $\{y_T(t)\}_{0 \leq t \leq T}$

**Direct optimization problem:**

$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{2} \int_0^T \boxed{\|y(t) - y_T(t)\|_2^2 dt} + \boxed{\lambda \|\theta\|_1}$$

direct trajectory error sparsity penalty

subject to  $\begin{cases} \boxed{\begin{aligned} \dot{x} &= \theta \Phi(x) \\ y &= Cx \end{aligned}} & \text{dynamics constraint} \\ \boxed{S \circ \theta \geq 0.} & \text{positive dynamics constraint} \end{cases}$

**Advantages:**

- No numerical differentiation
- Hidden species enrich modeling capacity
- Directly optimizes what we care about

**Disadvantages:**

- Nonconvex



Bayesian perspectives of  
the direct and indirect approaches



# Convert to Unconstrained Optimization

## Constrained Optimization

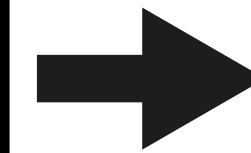
$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$



# Convert to Unconstrained Optimization

## Constrained Optimization

$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$



## Parameter-to-Output Operator

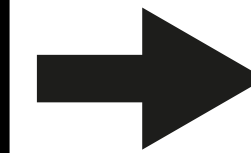
$$\mathcal{M} : \mathbb{R}^{n \times m} \rightarrow \mathbb{L}_n^2[0, T]$$

$$y = \mathcal{M}(\boldsymbol{\theta}) \iff \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx. \end{cases}$$

# Convert to Unconstrained Optimization

## Constrained Optimization

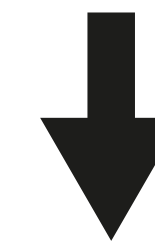
$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$



## Parameter-to-Output Operator

$$\mathcal{M} : \mathbb{R}^{n \times m} \rightarrow \mathbb{L}_n^2[0, T]$$

$$y = \mathcal{M}(\boldsymbol{\theta}) \iff \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx. \end{cases}$$



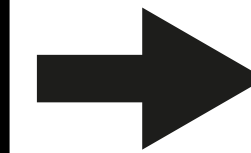
## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})$$

# Convert to Unconstrained Optimization

## Constrained Optimization

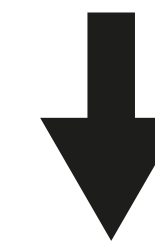
$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$



## Parameter-to-Output Operator

$$\mathcal{M} : \mathbb{R}^{n \times m} \rightarrow \mathbb{L}_n^2[0, T]$$

$$y = \mathcal{M}(\boldsymbol{\theta}) \iff \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx. \end{cases}$$



## Unconstrained Optimization

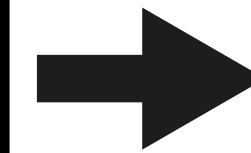
$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad \frac{1}{2} \int_0^T \left\| \overbrace{[\mathcal{M}(\boldsymbol{\theta})](t)}^{\text{smooth } J(\boldsymbol{\theta})} - y_T(t) \right\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})$$



# Convert to Unconstrained Optimization

## Constrained Optimization

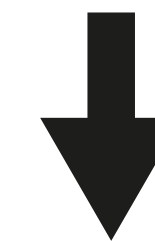
$$\begin{aligned} \boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad & \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1 \\ \text{subject to} \quad & \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases} \end{aligned}$$



## Parameter-to-Output Operator

$$\mathcal{M} : \mathbb{R}^{n \times m} \rightarrow \mathbb{L}_n^2[0, T]$$

$$y = \mathcal{M}(\boldsymbol{\theta}) \iff \begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx. \end{cases}$$



## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \quad \frac{1}{2} \int_0^T \underbrace{\left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2}_{\text{smooth } J(\boldsymbol{\theta})} dt + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\boldsymbol{\theta})} + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\boldsymbol{\theta})} + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \alpha \nabla J_{\boldsymbol{\theta}^{(k)}}$$



# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\boldsymbol{\theta})} + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

$$\boldsymbol{\theta}^{(k+1)} = \operatorname{prox}_{h,\alpha} \left( \boldsymbol{\theta}^{(k)} - \alpha \nabla J_{\boldsymbol{\theta}^{(k)}} \right)$$

$$\operatorname{prox}_{h,\alpha}(\boldsymbol{\theta}) \triangleq \arg \min_{\tilde{\boldsymbol{\theta}}} \left\{ \frac{1}{2\alpha} \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_F^2 + h(\tilde{\boldsymbol{\theta}}) \right\} \text{ (closest parameter that minimizes h)}$$



# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\boldsymbol{\theta})} + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

$$\boldsymbol{\theta}^{(k+1)} = \operatorname{prox}_{h,\alpha} \left( \boldsymbol{\theta}^{(k)} - \alpha \nabla J_{\boldsymbol{\theta}^{(k)}} \right)$$

$$\operatorname{prox}_{h,\alpha}(\boldsymbol{\theta}) \triangleq \arg \min_{\tilde{\boldsymbol{\theta}}} \left\{ \frac{1}{2\alpha} \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_F^2 + h(\tilde{\boldsymbol{\theta}}) \right\} \text{ (closest parameter that minimizes h)}$$

$$= \max \{ S \circ \boldsymbol{\theta} - \lambda\alpha, 0 \} + \operatorname{sign}(\bar{S} \circ \boldsymbol{\theta}) \circ \max \{ |\bar{S} \circ \boldsymbol{\theta}| - \lambda\alpha, 0 \}$$

projected soft-thresholding onto  $\mathbb{R}_{\geq 0}$

# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\boldsymbol{\theta})](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\boldsymbol{\theta})} + \underbrace{\lambda \|\boldsymbol{\theta}\|_1 + \mathcal{I}_+(S \circ \boldsymbol{\theta})}_{h(\boldsymbol{\theta}) \text{ non-smooth but convex}}$$

$$\boldsymbol{\theta}^{(k+1)} = \operatorname{prox}_{h,\alpha} \left( \boldsymbol{\theta}^{(k)} - \alpha \nabla J_{\boldsymbol{\theta}^{(k)}} \right)$$

$$\operatorname{prox}_{h,\alpha}(\boldsymbol{\theta}) \triangleq \operatorname{argmin}_{\tilde{\boldsymbol{\theta}}} \left\{ \frac{1}{2\alpha} \|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_F^2 + h(\tilde{\boldsymbol{\theta}}) \right\} \text{ (closest parameter that minimizes h)}$$

$$= \max \{ S \circ \boldsymbol{\theta} - \lambda\alpha, 0 \} + \operatorname{sign}(\bar{S} \circ \boldsymbol{\theta}) \circ \max \{ |\bar{S} \circ \boldsymbol{\theta}| - \lambda\alpha, 0 \}$$

projected soft-thresholding onto  $\mathbb{R}_{\geq 0}$

$$\nabla J_{\boldsymbol{\theta}^{(k)}} = \int_0^T \lambda^{(k)}(t) [b^{(k)}(t)]^T dt.$$

comes from backward simulation  
of a co-state equation

comes from forward simulation  
state equation

# Solution Strategy: Proximal Gradient Descent

## Unconstrained Optimization

$$\theta^* = \operatorname{argmin}_{\theta} \underbrace{\frac{1}{2} \int_0^T \left\| [\mathcal{M}(\theta)](t) - y_T(t) \right\|_2^2 dt}_{\text{smooth } J(\theta)} + \underbrace{\lambda \|\theta\|_1 + \mathcal{I}_+(S \circ \theta)}_{h(\theta) \text{ non-smooth but convex}}$$

$$\theta^{(k+1)} = \operatorname{prox}_{h,\alpha} \left( \theta^{(k)} - \alpha \nabla J_{\theta^{(k)}} \right) \quad \text{Also modified to obtain and accelerated method}$$

$$\operatorname{prox}_{h,\alpha}(\theta) \triangleq \operatorname{argmin}_{\tilde{\theta}} \left\{ \frac{1}{2\alpha} \|\theta - \tilde{\theta}\|_F^2 + h(\tilde{\theta}) \right\} \quad \text{(closest parameter that minimizes h)}$$

$$= \max \{ S \circ \theta - \lambda\alpha, 0 \} + \operatorname{sign}(\bar{S} \circ \theta) \circ \max \{ |\bar{S} \circ \theta| - \lambda\alpha, 0 \}$$

projected soft-thresholding onto  $\mathbb{R}_{\geq 0}$

$$\nabla J_{\theta^{(k)}} = \int_0^T \lambda^{(k)}(t) [b^{(k)}(t)]^T dt.$$

comes from backward simulation  
of a co-state equation

comes from forward simulation  
state equation



# Example 1: Drosophila Circadian Oscillator

$$\left\{ \begin{array}{l} \dot{x}_1 = \frac{K_{ip}^n v_{sp}}{K_{ip}^n + x_{10}^n} - k_d x_1 - \frac{v_{mp} x_1}{K_{mp} + x_1} \\ \dot{x}_2 = K_{sp} x_1 - k_d x_2 - \frac{v_{1p} x_2}{K_{1p} + x_2} + \frac{v_{2p} x_3}{K_{2p} + x_3} \\ \dot{x}_3 = \frac{v_{1p} x_2}{K_{1p} + x_2} - k_d x_3 - \frac{v_{2p} x_3}{K_{2p} + x_3} - \frac{v_{3p} x_3}{k_{3p} + x_3} + \frac{v_{4p} x_4}{k_{4p} + x_4} \\ \dot{x}_4 = k_4 x_9 - k_d x_4 - k_3 x_4 x_8 - \frac{v_{dp} x_4}{K_{dp} + x_4} + \frac{v_{3p} x_3}{k_{3p} + x_3} - \frac{v_{4p} x_4}{k_{4p} + x_4} \\ \dot{x}_5 = \frac{K_{iT}^n v_{st}}{K_{iT}^n + x_{10}^n} - k_d x_5 - \frac{v_{mt} x_5}{k_{mt} + x_5} \\ \dot{x}_6 = k_{st} x_5 - k_d x_6 - \frac{v_{1t} x_6}{k_{1t} + x_6} + \frac{v_{2t} x_7}{k_{2t} + x_7} \\ \dot{x}_7 = \frac{v_{1t} x_6}{k_{1t} + x_6} - k_d x_7 - \frac{v_{2t} x_7}{k_{2t} + x_7} - \frac{v_{3t} x_7}{k_{3t} + x_7} + \frac{v_{4t} x_8}{k_{4t} + x_8} \\ \dot{x}_8 = k_4 x_9 - k_d x_8 - k_3 x_4 x_8 + \frac{v_{3t} x_7}{k_{3t} + x_7} - \frac{v_{4t} x_8}{k_{4t} + x_8} - \frac{v_{dt} x_8}{k_{dt} + x_8} \\ \dot{x}_9 = k_2 x_{10} - k_1 x_9 - k_4 x_9 - k_d x_9 + k_3 x_4 x_8 \\ \dot{x}_{10} = k_1 x_9 - k_2 x_{10} - k_d x_{10}. \end{array} \right.$$

- 10-species model describing PER/TIM circadian dynamics
- Mass-action and Hill-type nonlinearities
- Exhibits oscillations, chaos, birhythmicity

Leloup et al. "Chaos and birhythmicity in a model for circadian oscillations of the PER and TIM proteins in Drosophila." *Journal of theoretical biology* (1999).

# Example 1: Drosophila Circadian Oscillator

$$\begin{aligned}
 \dot{x}_1 &= \frac{K_{ip}^n v_{sp}}{K_{ip}^n + x_{10}^n} - k_d x_1 - \frac{v_{mp} x_1}{K_{mp} + x_1} \\
 \dot{x}_2 &= K_{sp} x_1 - k_d x_2 - \frac{v_{1p} x_2}{K_{1p} + x_2} + \frac{v_{2p} x_3}{K_{2p} + x_3} \\
 \dot{x}_3 &= \frac{v_{1p} x_2}{K_{1p} + x_2} - k_d x_3 - \frac{v_{2p} x_3}{K_{2p} + x_3} - \frac{v_{3p} x_3}{k_{3p} + x_3} + \frac{v_{4p} x_4}{k_{4p} + x_4} \\
 \dot{x}_4 &= k_4 x_9 - k_d x_4 - k_3 x_4 x_8 - \frac{v_{dp} x_4}{K_{dp} + x_4} + \frac{v_{3p} x_3}{k_{3p} + x_3} - \frac{v_{4p} x_4}{k_{4p} + x_4} \\
 \dot{x}_5 &= \frac{K_{iT}^n v_{st}}{K_{iT}^n + x_{10}^n} - k_d x_5 - \frac{v_{mt} x_5}{k_{mt} + x_5} \\
 \dot{x}_6 &= k_{st} x_5 - k_d x_6 - \frac{v_{1t} x_6}{k_{1t} + x_6} + \frac{v_{2t} x_7}{k_{2t} + x_7} \\
 \dot{x}_7 &= \frac{v_{1t} x_6}{k_{1t} + x_6} - k_d x_7 - \frac{v_{2t} x_7}{k_{2t} + x_7} - \frac{v_{3t} x_7}{k_{3t} + x_7} + \frac{v_{4t} x_8}{k_{4t} + x_8} \\
 \dot{x}_8 &= k_4 x_9 - k_d x_8 - k_3 x_4 x_8 + \frac{v_{3t} x_7}{k_{3t} + x_7} - \frac{v_{4t} x_8}{k_{4t} + x_8} - \frac{v_{dt} x_8}{k_{dt} + x_8} \\
 \dot{x}_9 &= k_2 x_{10} - k_1 x_9 - k_4 x_9 - k_d x_9 + k_3 x_4 x_8 \\
 \dot{x}_{10} &= k_1 x_9 - k_2 x_{10} - k_d x_{10}.
 \end{aligned}$$

- 10-species model describing PER/TIM circadian dynamics
- Mass-action and Hill-type nonlinearities
- Exhibits oscillations, chaos, birhythmicity

Leloup et al. "Chaos and birhythmicity in a model for circadian oscillations of the PER and TIM proteins in Drosophila." *Journal of theoretical biology* (1999).



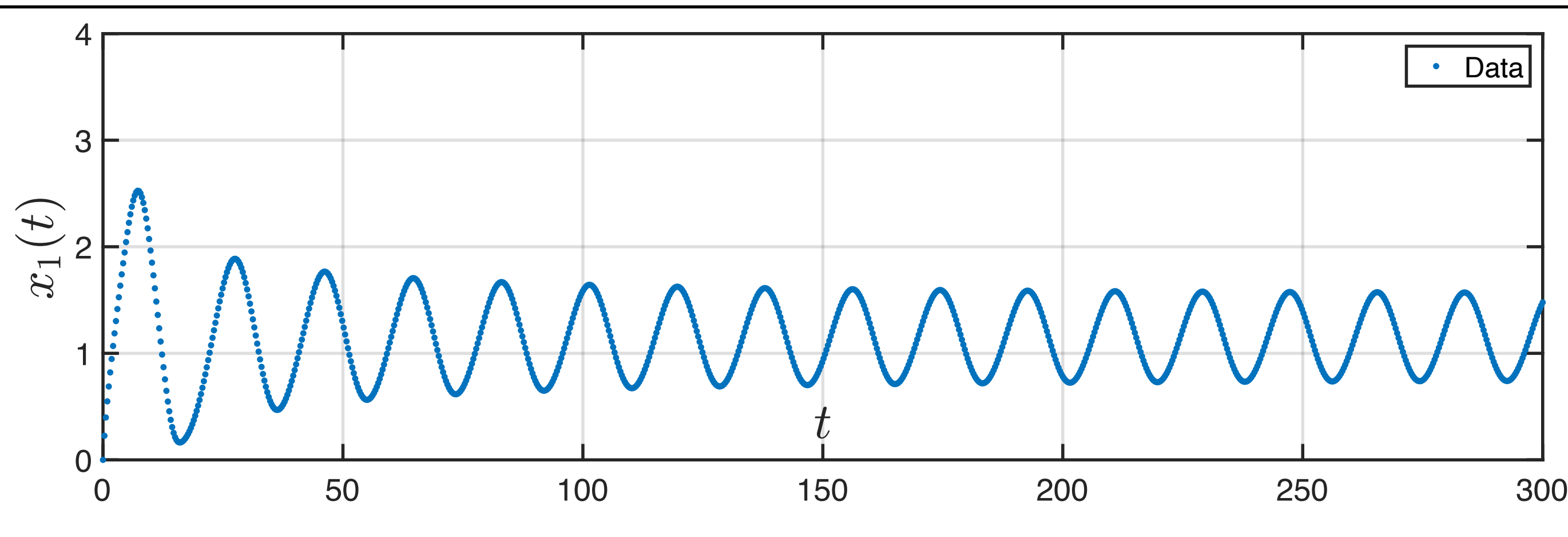
# Example 1: Drosophila Circadian Oscillator

# Example 1: Drosophila Circadian Oscillator



# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \boldsymbol{\theta} \Phi(x)$

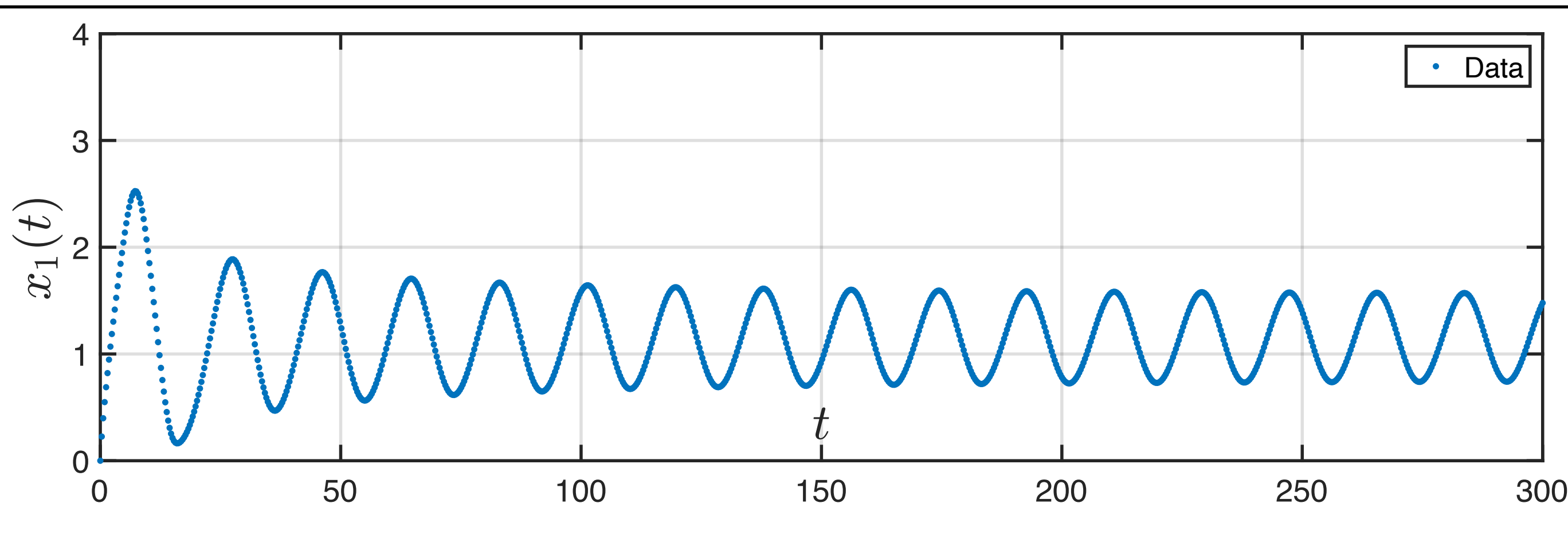
$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\boldsymbol{\theta}\|_1$$

subject to

$$\begin{cases} \dot{x} = \boldsymbol{\theta} \Phi(x) \\ y = Cx \\ S \circ \boldsymbol{\theta} \geq 0. \end{cases}$$

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

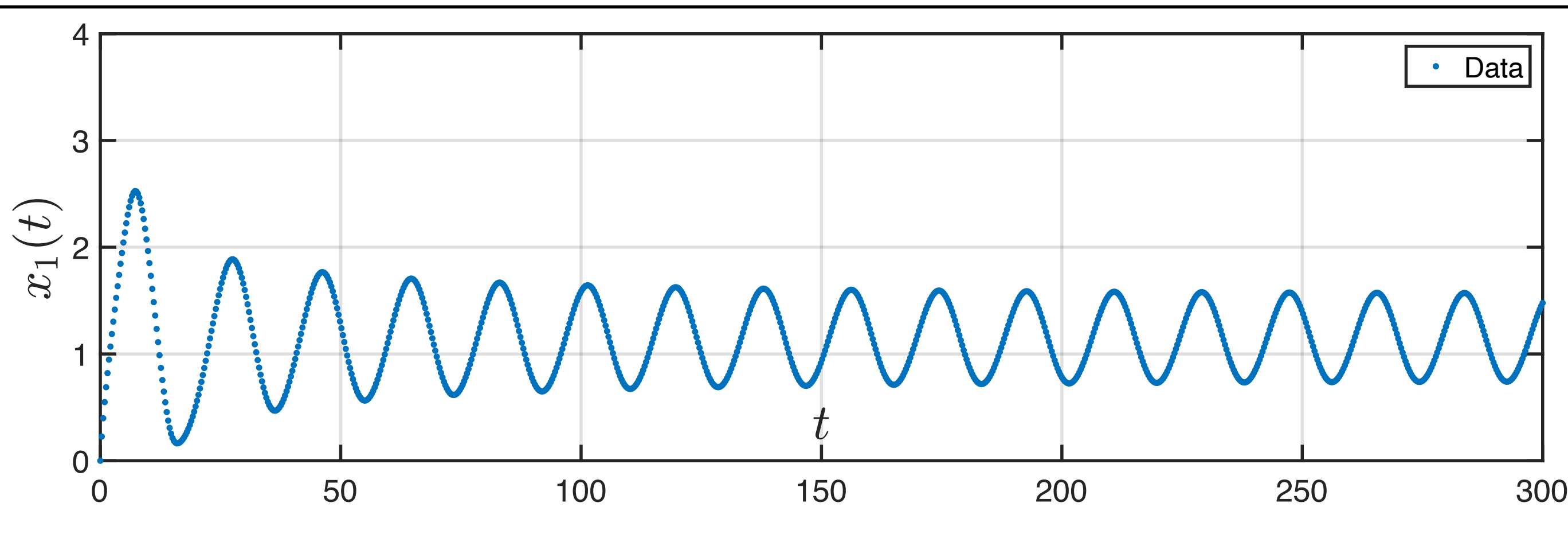
- $n = 2$  species (1 observed + 1 hidden)

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

subject to 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

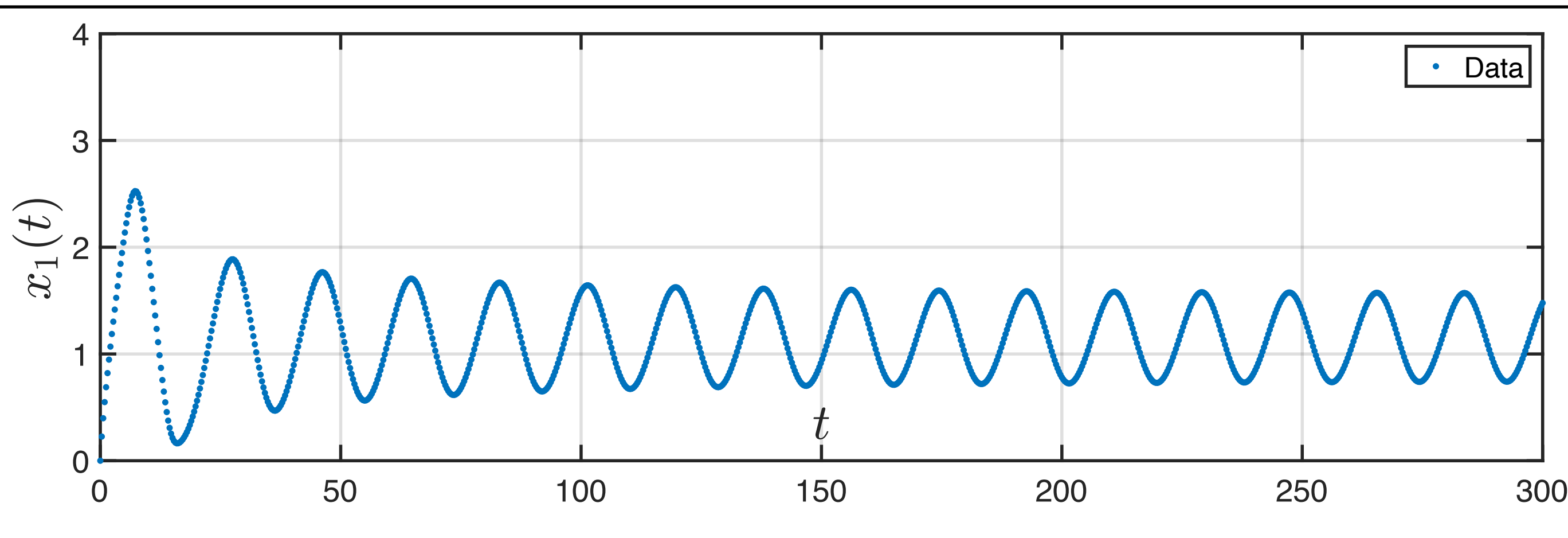
$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

subject to 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$



# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)

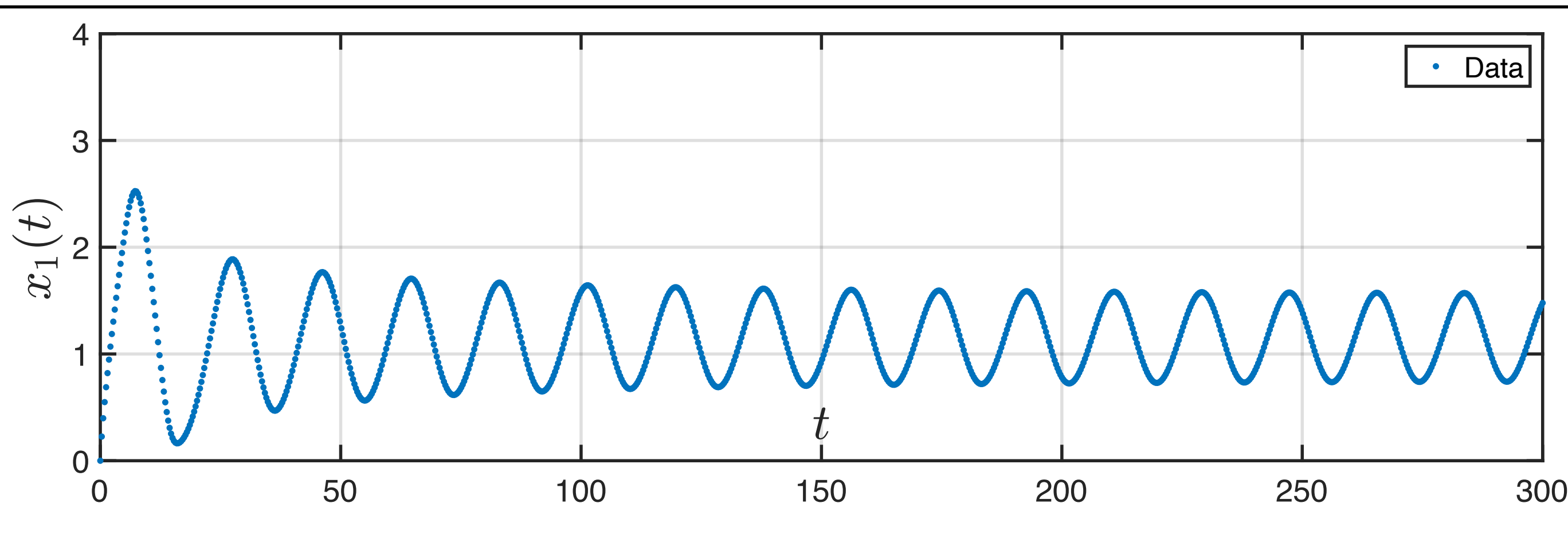
$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

subject to 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

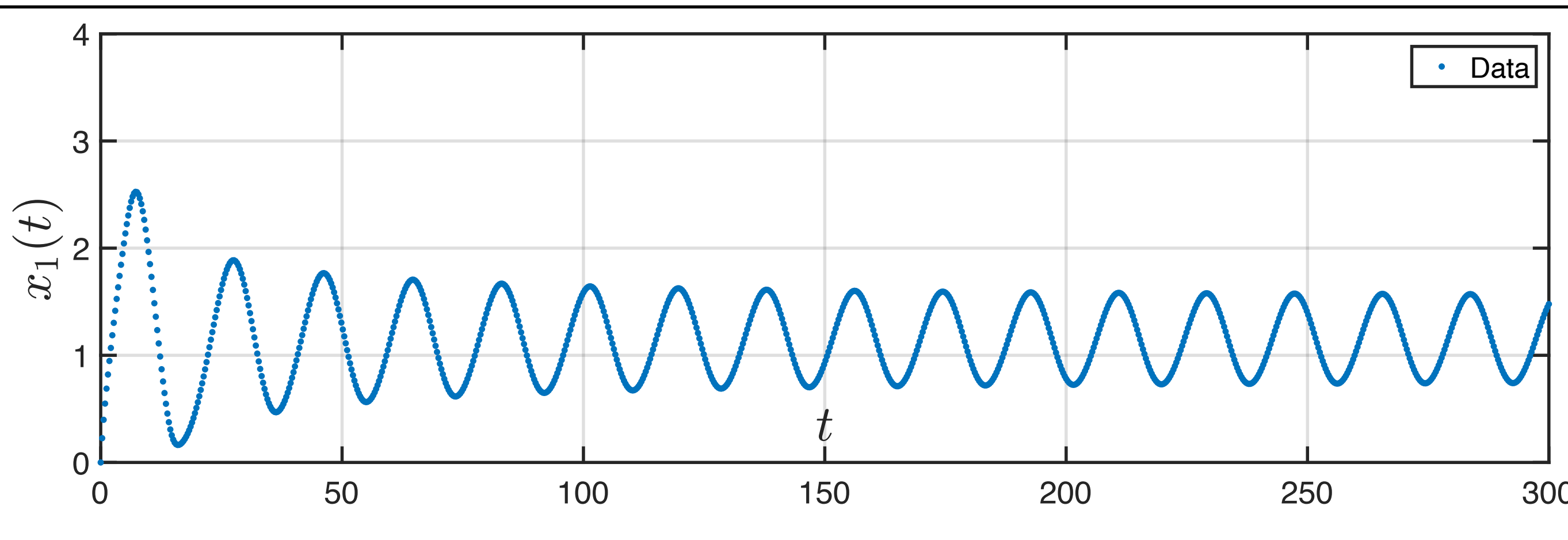
$$\Phi(x) = \begin{matrix} \text{Library} \\ \text{(m functions)} \end{matrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

subject to 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix} \quad \Phi(x) = \begin{matrix} \text{Library} \\ \text{(m functions)} \end{matrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

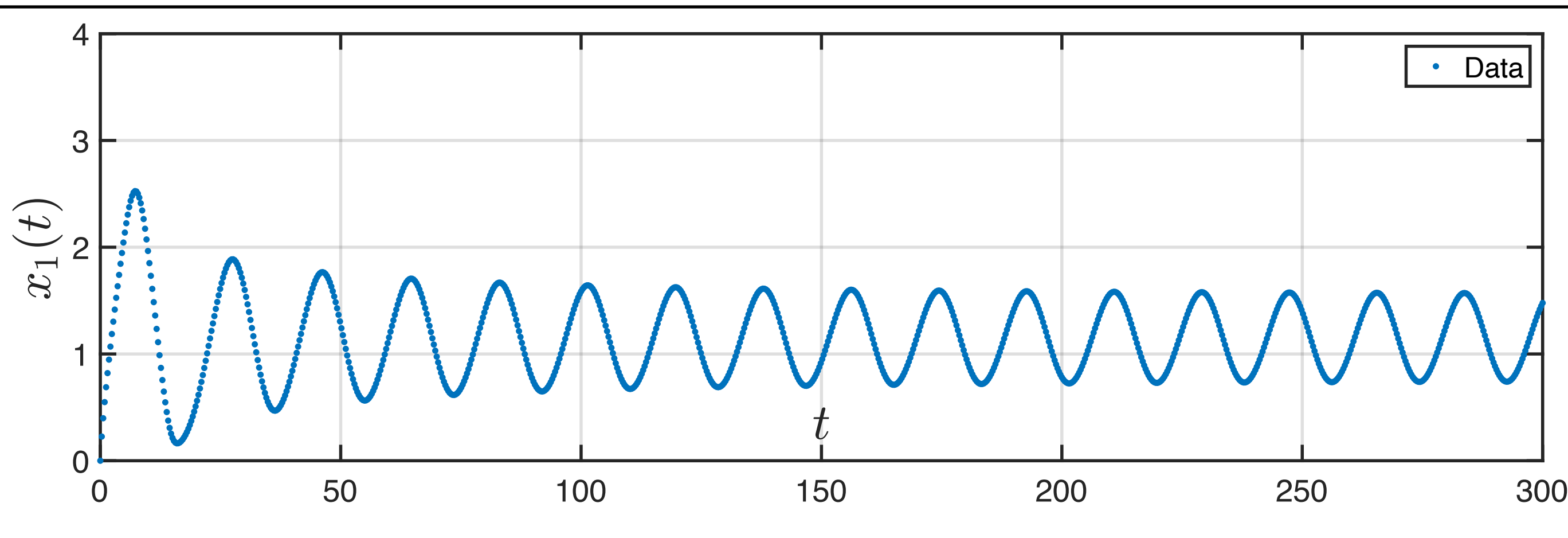
$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

subject to 
$$\begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$



# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix} \quad \Phi(x) = \begin{matrix} \text{Library} \\ \text{(m functions)} \end{matrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

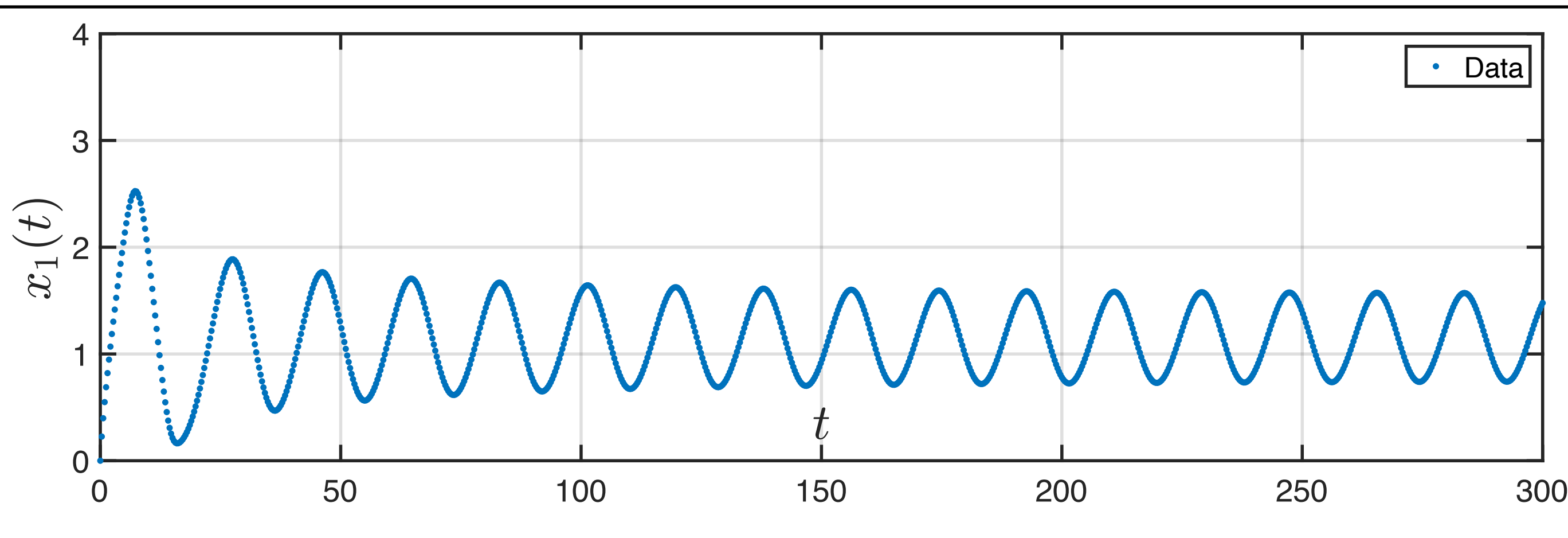
$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

$$\text{subject to } \begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$



# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

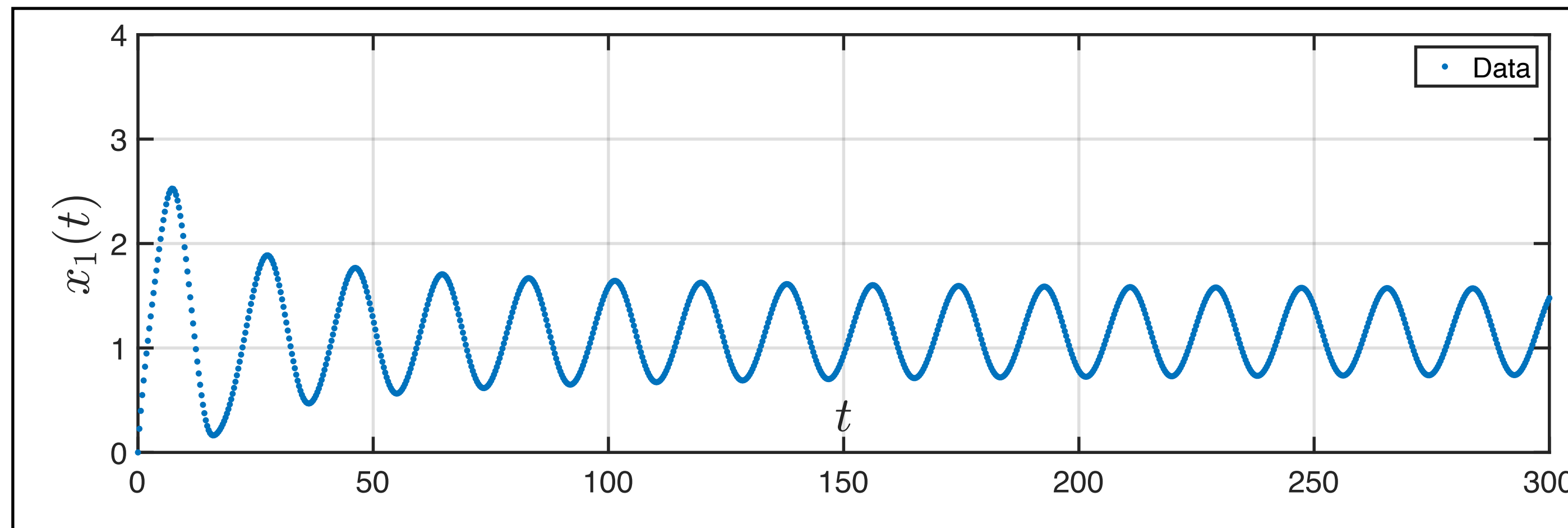
$$\text{subject to } \begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

Library  
(m functions)

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

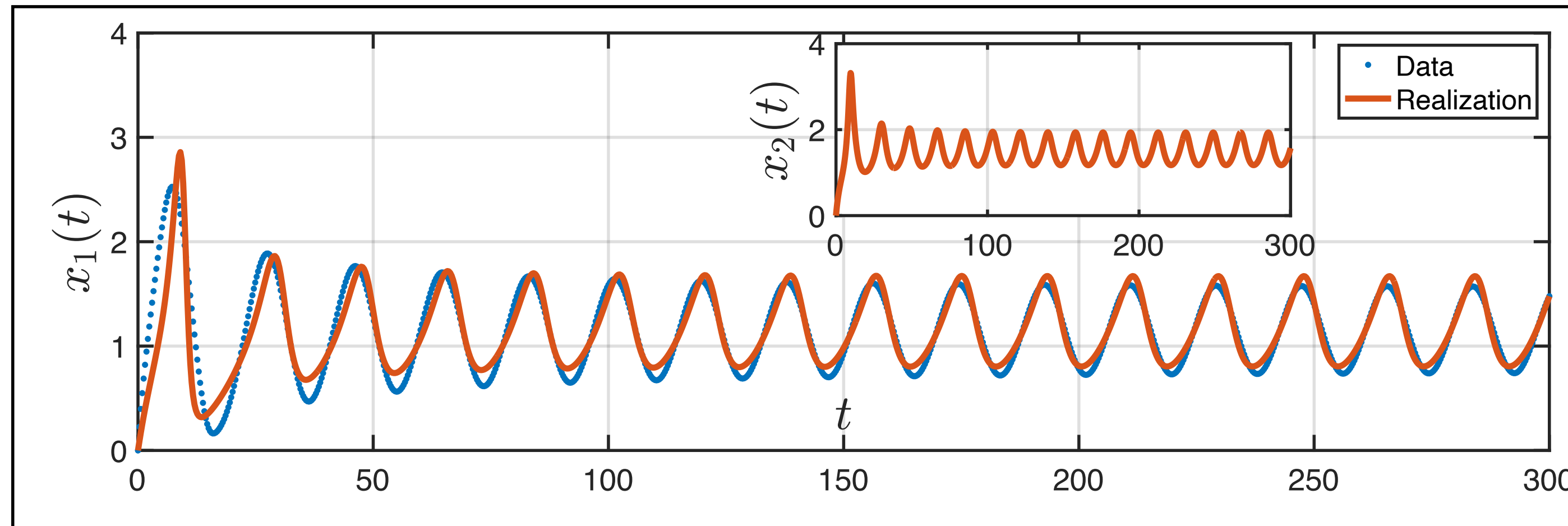
$$\text{subject to } \begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

Library  
(m functions)

# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

$$\text{subject to } \begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$

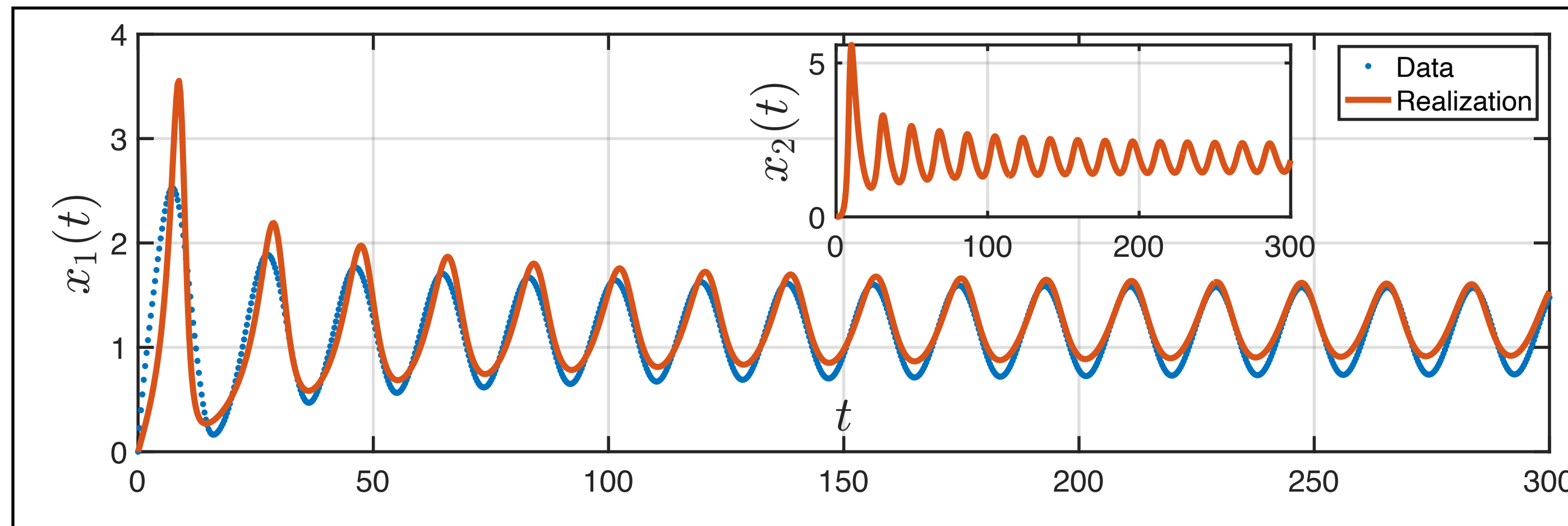
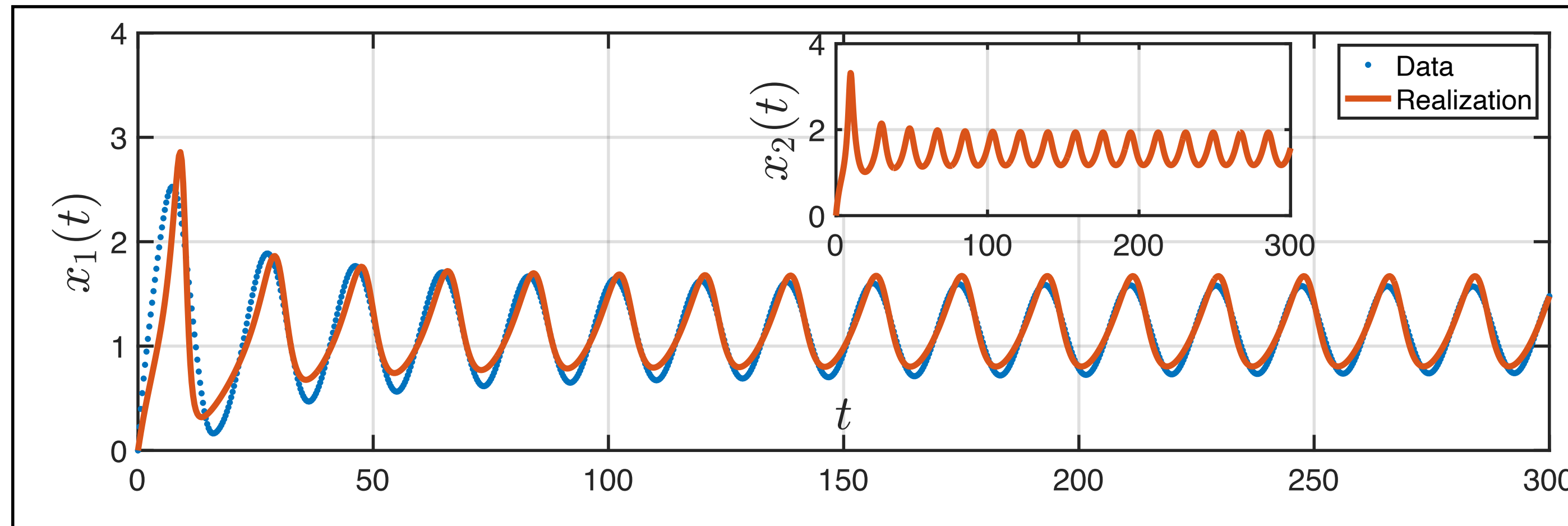
$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

Library  
(m functions)



# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$

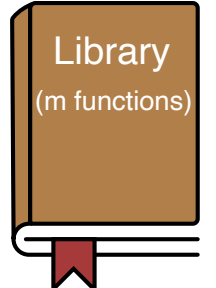


**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )

$$\theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$



$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

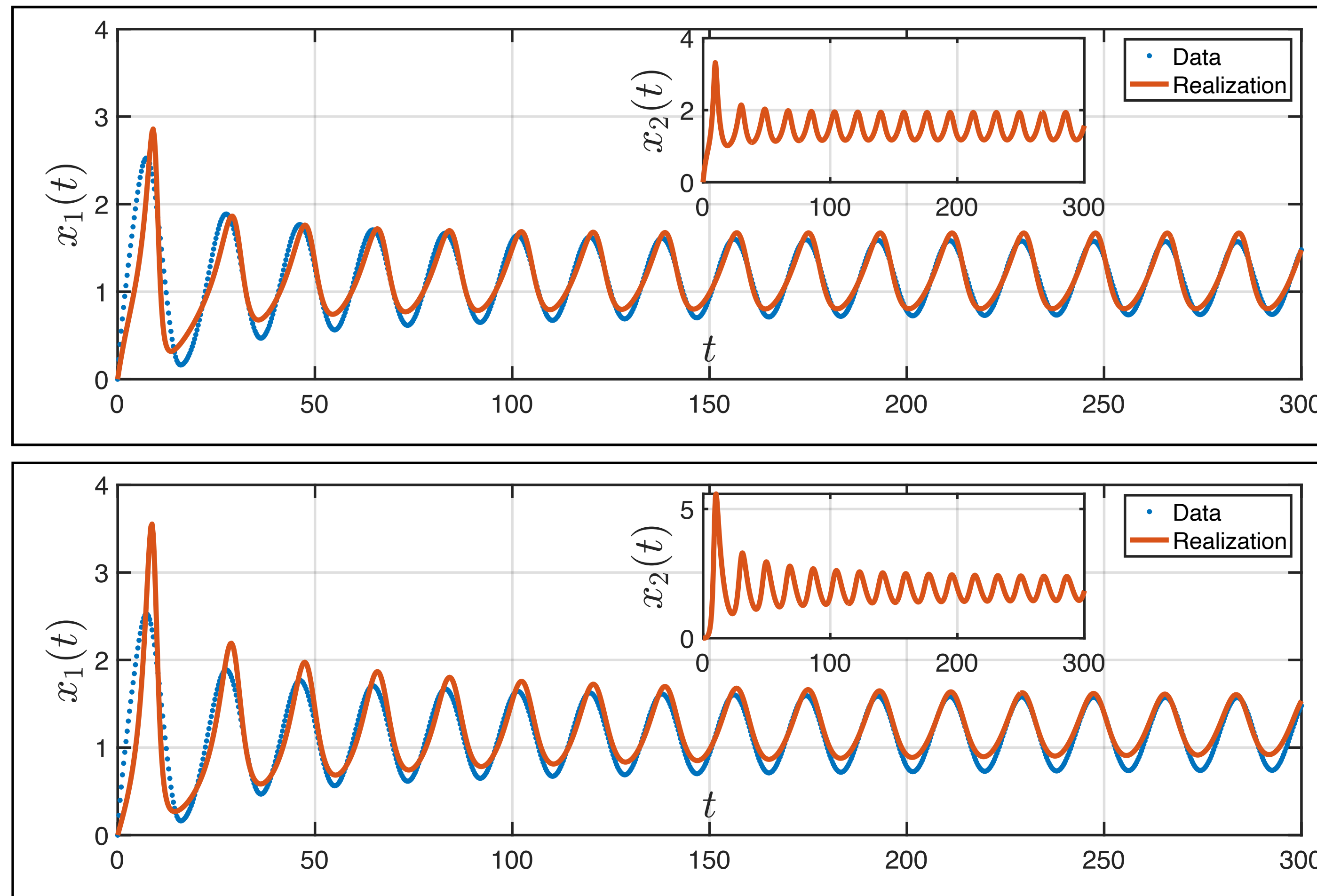
$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{2} \int_0^T \|y(t) - y_T(t)\|_2^2 dt + \lambda \|\theta\|_1$$

$$\text{subject to } \begin{cases} \dot{x} = \theta \Phi(x) \\ y = Cx \\ S \circ \theta \geq 0. \end{cases}$$



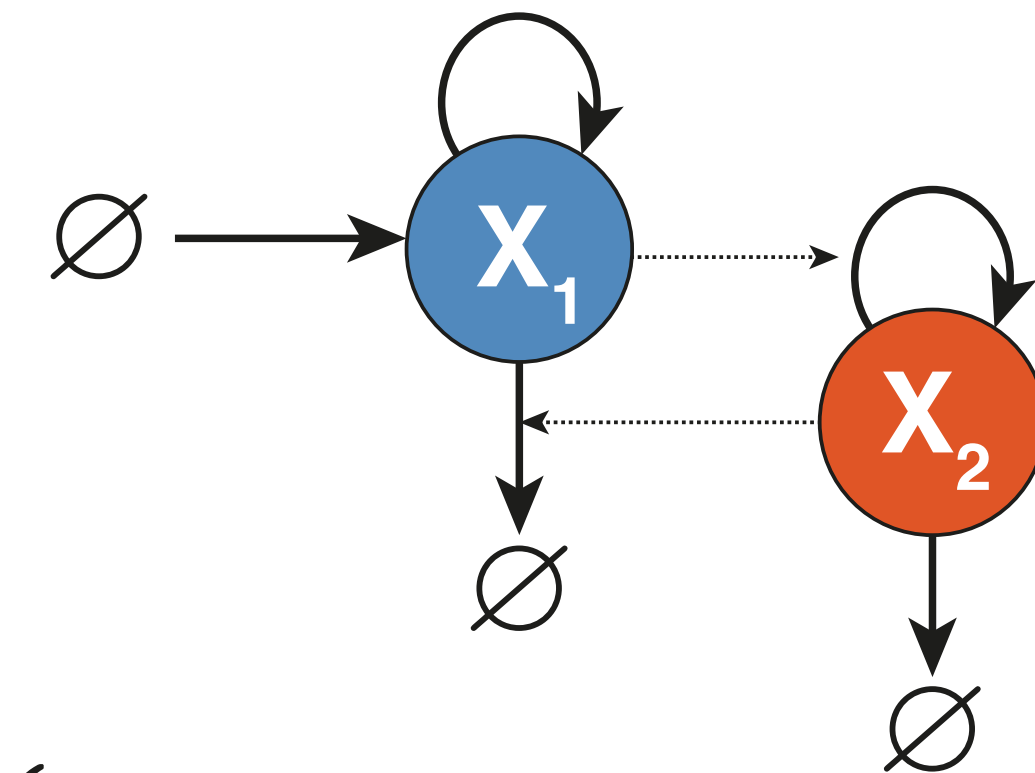
# Example 1: Drosophila Circadian Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species (1 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )



$$\begin{cases} \dot{x}_1 = \mu_0 + kx_1 + \eta_1 x_1^2 - \gamma_0 x_1 x_2 \\ \dot{x}_2 = \eta x_1 x_2 + \eta_2 x_1^2 - \gamma_2 x_2 \end{cases}$$

|          |        |            |        |
|----------|--------|------------|--------|
| $\mu_0$  | 0.1626 | $k$        | 0.087  |
| $\eta$   | 0.0274 | $\gamma_0$ | 0.2297 |
| $\eta_1$ | 0.2338 | $\gamma_2$ | 0.1892 |
| $\eta_2$ | 0.1882 |            |        |

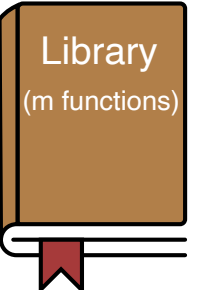
# Example 1: Indirect method

**Desired trajectory**  $\{x_T(t)\}_{0 \leq t \leq T}$

**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 2$  species
- Bimolecular reactions (up to quadratic)
- No sparsity (full model)

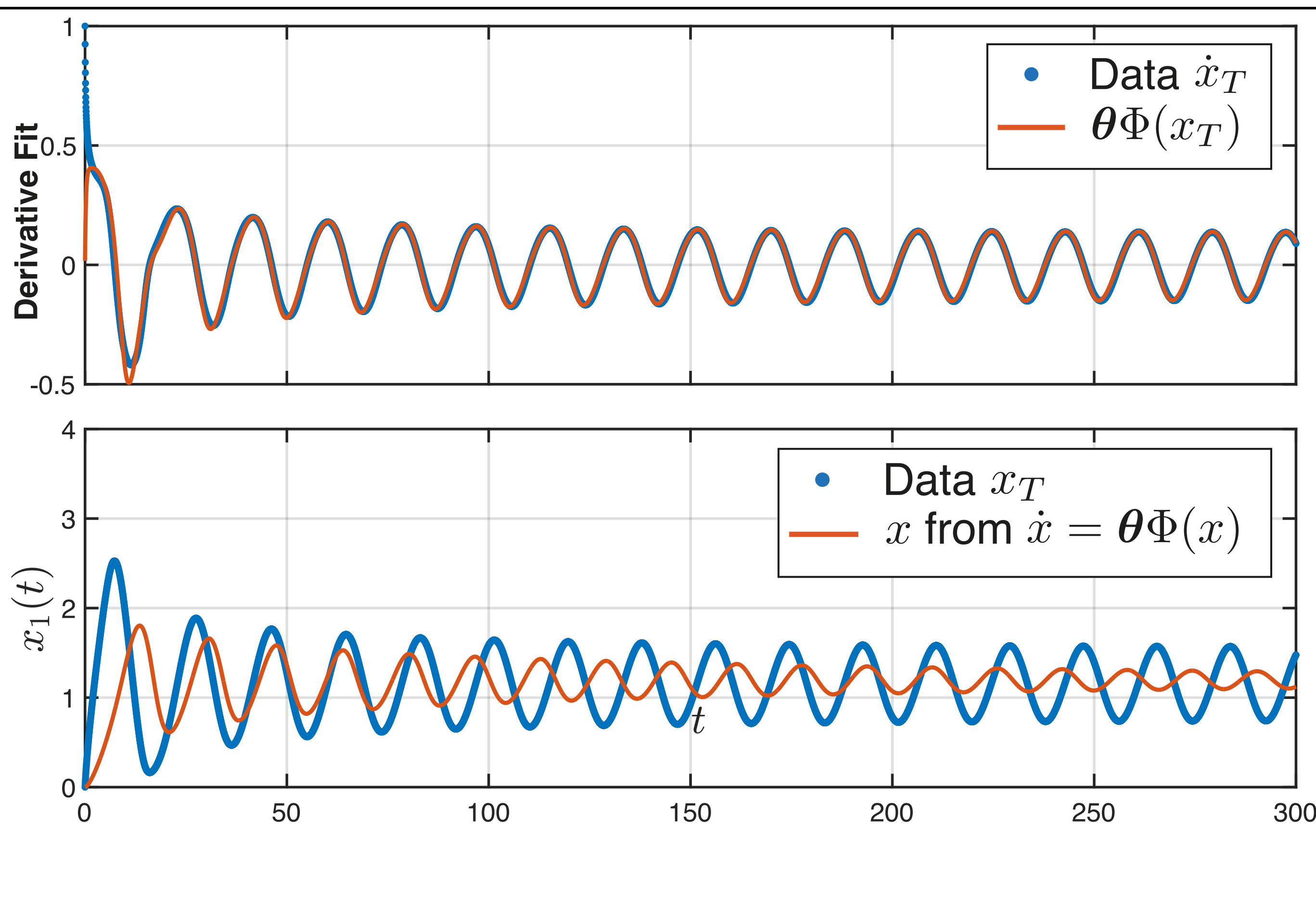
$$\theta = \begin{bmatrix} \theta_{11} & \cdots & \theta_{16} \\ \theta_{21} & \cdots & \theta_{26} \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$


$$\min_{\theta} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \theta \Phi(x_T(t)) \right\|_2^2 dt$$

# Example 1: Indirect method

**Desired trajectory**  $\{x_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta\Phi(x)$

- $n = 2$  species
- Bimolecular reactions (up to quadratic)
- No sparsity (full model)

$$\theta = \begin{bmatrix} \theta_{11} & \cdots & \theta_{16} \\ \theta_{21} & \cdots & \theta_{26} \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix}$$

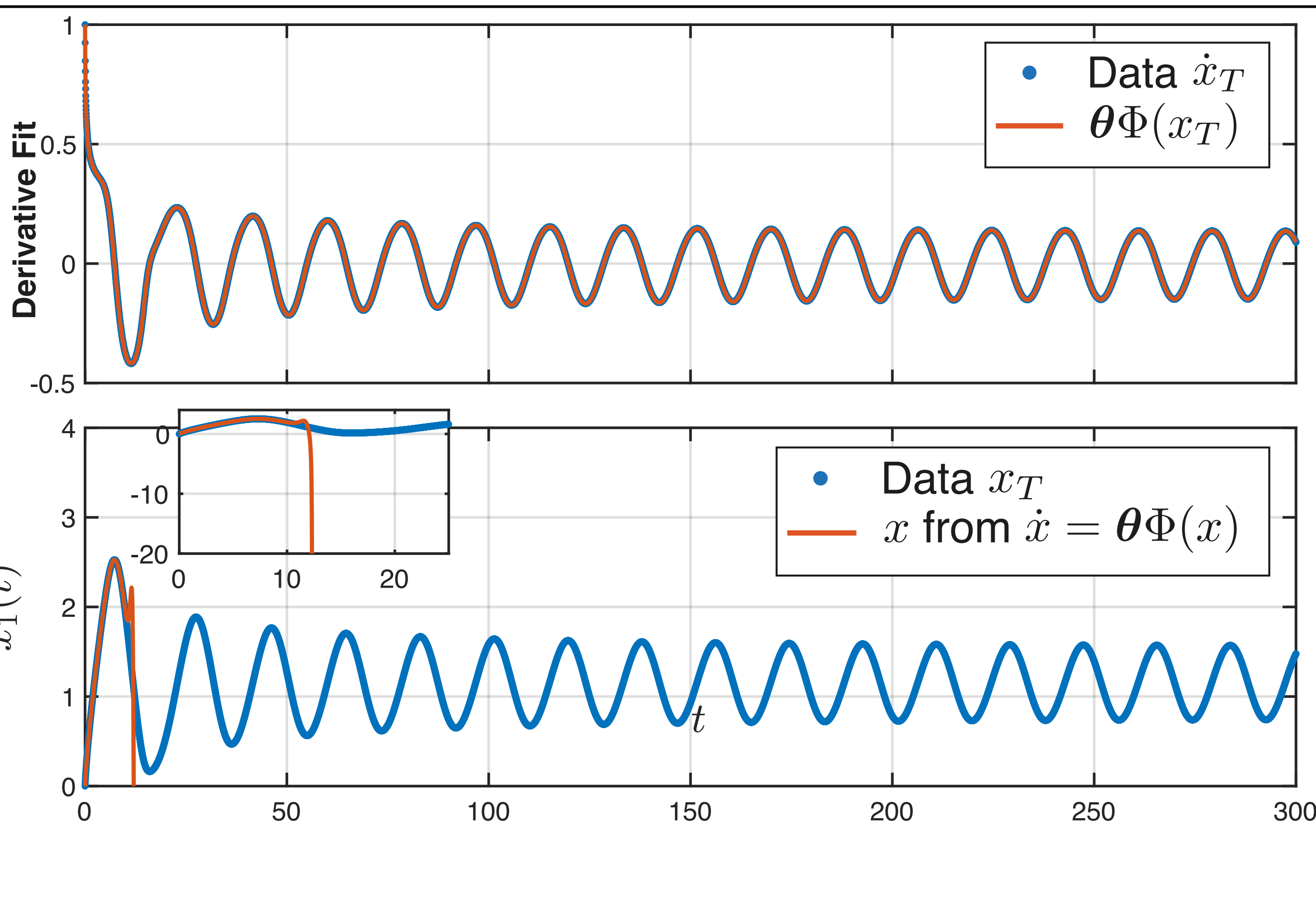
Library  
(m functions)

$$\min_{\theta} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \theta\Phi(x_T(t)) \right\|_2^2 dt$$



# Example 1: Indirect method

**Desired trajectory**  $\{x_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta\Phi(x)$

- $n = 10$  species
- Bimolecular reactions (up to quadratic)
- No sparsity (full model)

$$\theta = \begin{bmatrix} \theta_{1,1} & \cdots & \theta_{1,66} \\ \vdots & \ddots & \vdots \\ \theta_{10,1} & \cdots & \theta_{10,66} \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_{10} \\ x_1^2 \\ x_1 x_2 \\ \vdots \\ x_{10}^2 \end{bmatrix}_{66 \times 1}$$

Library  
(m functions)

$$\min_{\theta} \frac{1}{2} \int_0^T \left\| \dot{x}_T(t) - \theta\Phi(x_T(t)) \right\|_2^2 dt$$

# Example 2: Glycolytic Oscillator

$$\begin{cases} \dot{x}_1 = J_0 - \frac{k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_2 = k_2 x_2 x_5 - N k_2 x_2 - k_6 x_2 x_5 + \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_3 = N k_2 x_2 - A k_3 x_3 - k_2 x_2 x_5 + k_3 x_3 x_6 \\ \dot{x}_4 = \kappa x_7 - \kappa x_4 + A k_3 x_3 - k_3 x_3 x_6 - k_4 x_4 x_5 \\ \dot{x}_5 = N k_2 x_2 - k_2 x_2 x_5 - k_4 x_4 x_5 - k_6 x_2 x_5 \\ \dot{x}_6 = 2 A k_3 x_3 - k_5 x_6 - 2 k_3 x_3 x_6 - \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_7 = \kappa \psi x_4 - x_7 (k + \kappa \psi) . \end{cases}$$

- 7-species model
- Mass-action and Hill-type nonlinearities
- SINDy (indirect method) was shown to produce good derivative fits, but fail to capture trajectories beyond initial phase.

# Example 2: Glycolytic Oscillator

$$\begin{cases} \dot{x}_1 = J_0 - \frac{k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_2 = k_2 x_2 x_5 - N k_2 x_2 - k_6 x_2 x_5 + \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_3 = N k_2 x_2 - A k_3 x_3 - k_2 x_2 x_5 + k_3 x_3 x_6 \\ \dot{x}_4 = \kappa x_7 - \kappa x_4 + A k_3 x_3 - k_3 x_3 x_6 - k_4 x_4 x_5 \\ \dot{x}_5 = N k_2 x_2 - k_2 x_2 x_5 - k_4 x_4 x_5 - k_6 x_2 x_5 \\ \dot{x}_6 = 2 A k_3 x_3 - k_5 x_6 - 2 k_3 x_3 x_6 - \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_7 = \kappa \psi x_4 - x_7 (k + \kappa \psi) . \end{cases}$$

- 7-species model
- Mass-action and Hill-type nonlinearities
- SINDy (indirect method) was shown to produce good derivative fits, but fail to capture trajectories beyond initial phase.

Brunton et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." *PNAS* (2016).



# Example 2: Glycolytic Oscillator

$$\begin{cases} \dot{x}_1 = J_0 - \frac{k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_2 = k_2 x_2 x_5 - N k_2 x_2 - k_6 x_2 x_5 + \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_3 = N k_2 x_2 - A k_3 x_3 - k_2 x_2 x_5 + k_3 x_3 x_6 \\ \dot{x}_4 = \kappa x_7 - \kappa x_4 + A k_3 x_3 - k_3 x_3 x_6 - k_4 x_4 x_5 \\ \dot{x}_5 = N k_2 x_2 - k_2 x_2 x_5 - k_4 x_4 x_5 - k_6 x_2 x_5 \\ \dot{x}_6 = 2 A k_3 x_3 - k_5 x_6 - 2 k_3 x_3 x_6 - \frac{2 k_1 x_1 x_6}{\left(\frac{x_6}{K_1}\right)^q + 1} \\ \dot{x}_7 = \kappa \psi x_4 - x_7 (k + \kappa \psi) . \end{cases}$$

- 7-species model
- Mass-action and Hill-type nonlinearities
- SINDy (indirect method) was shown to produce good derivative fits, but fail to capture trajectories beyond initial phase.

Brunton et al. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." *PNAS* (2016).

# Example 2: Glycolytic Oscillator

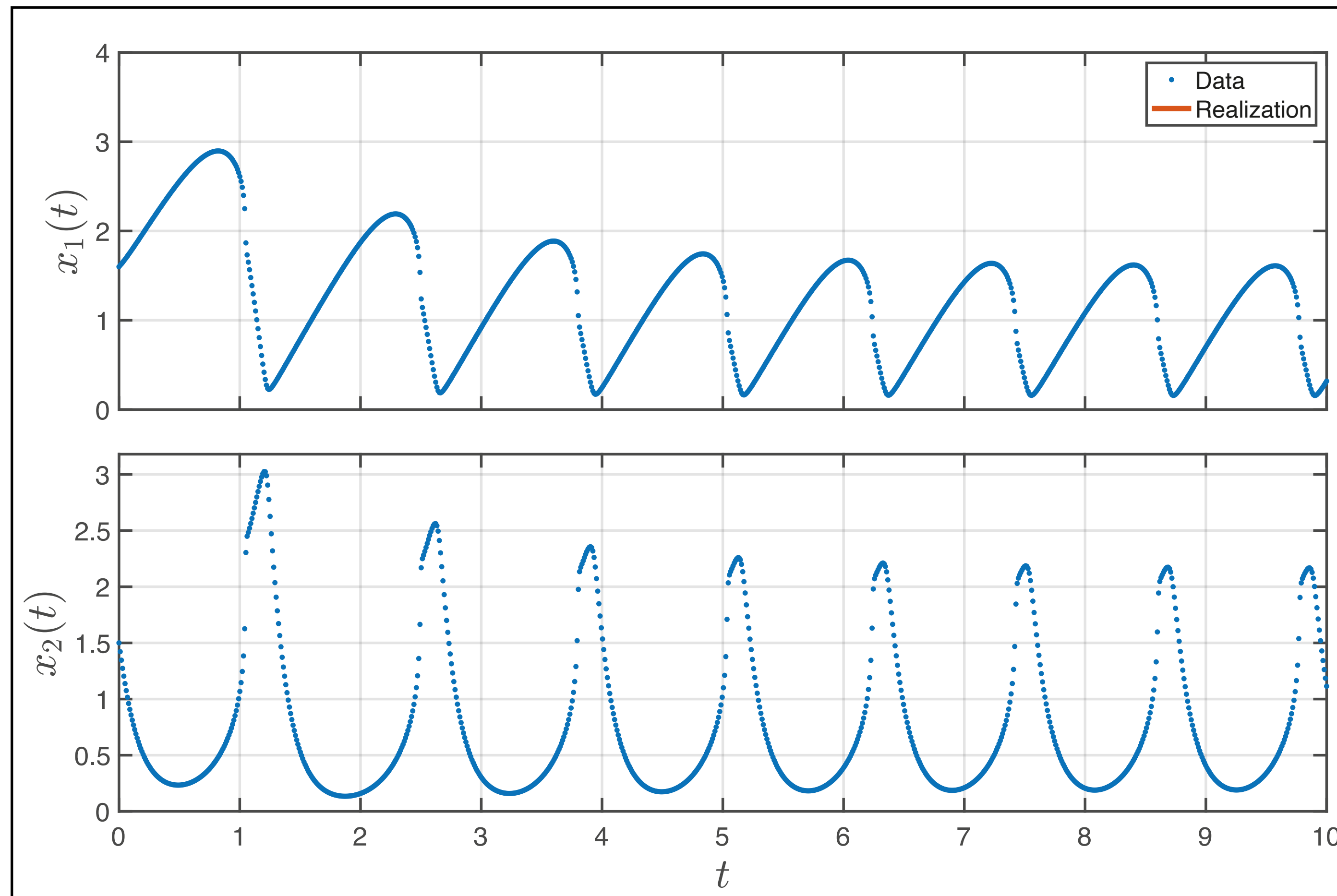
# Example 2: Glycolytic Oscillator



# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$

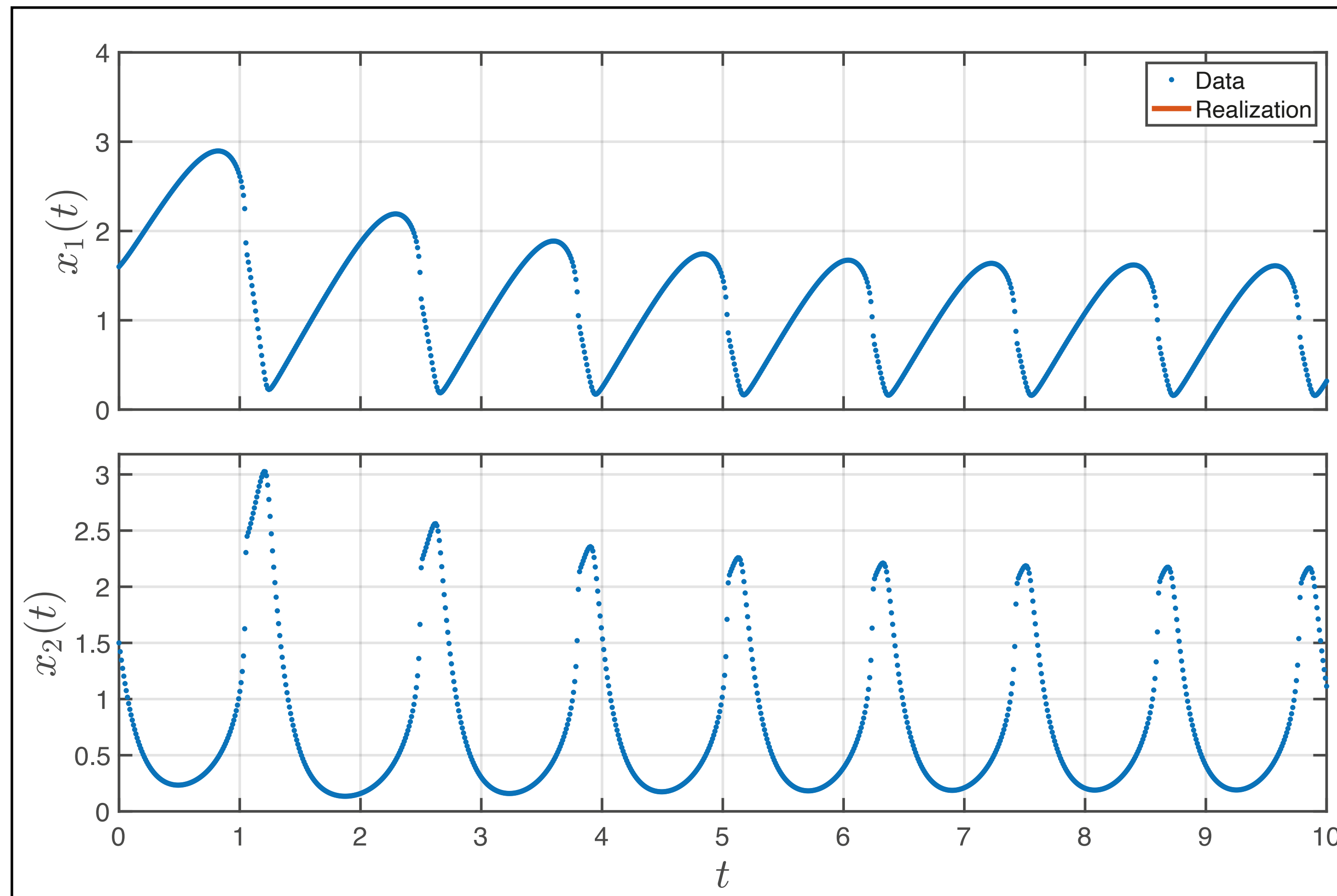
**Model:**  $\dot{x} = \theta \Phi(x)$





# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



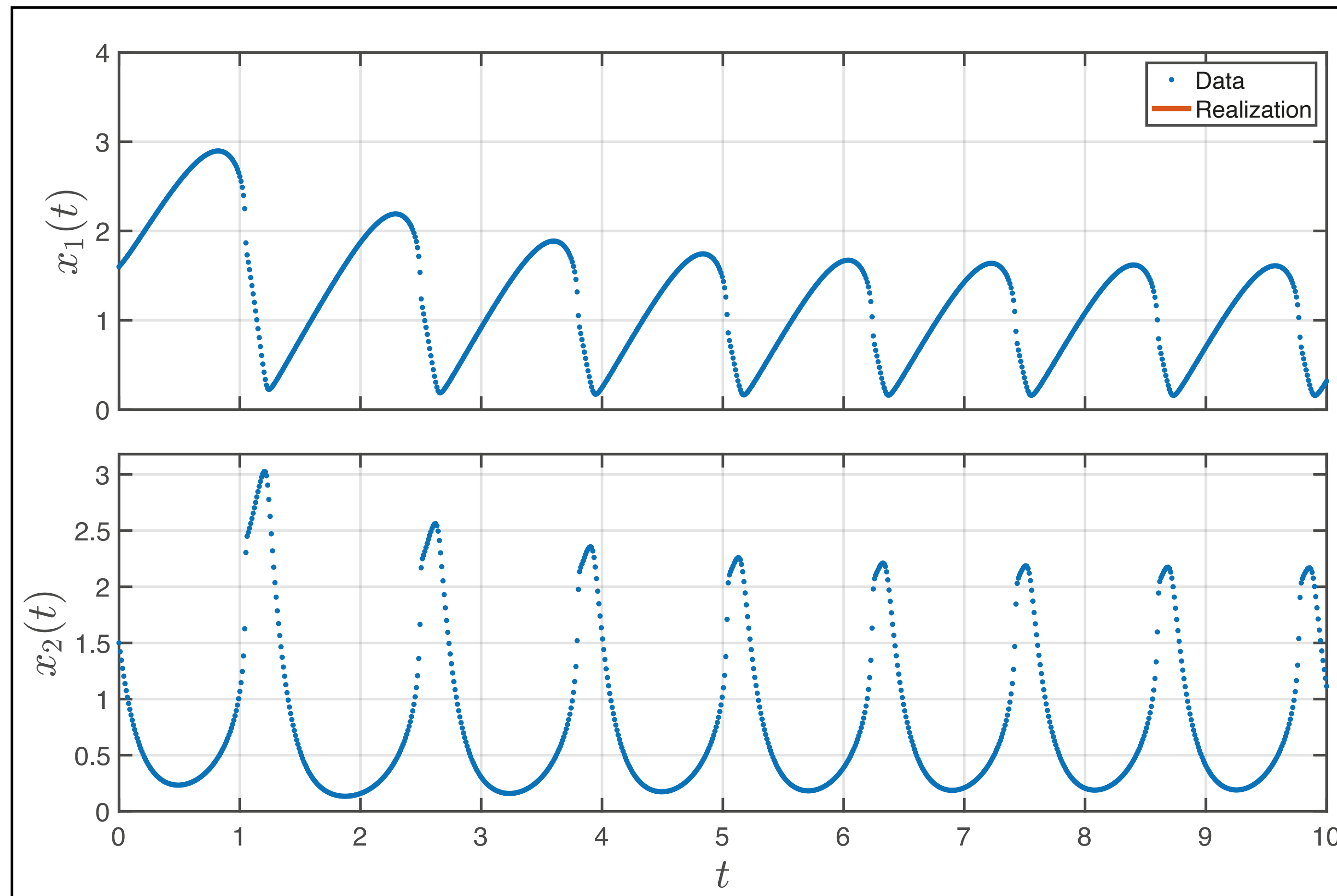
**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 3$  species (2 observed + 1 hidden)

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 3$  species (2 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)

$$\theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,10} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,10} \\ \theta_{3,1} & \theta_{3,2} & \cdots & \theta_{3,10} \end{bmatrix}$$

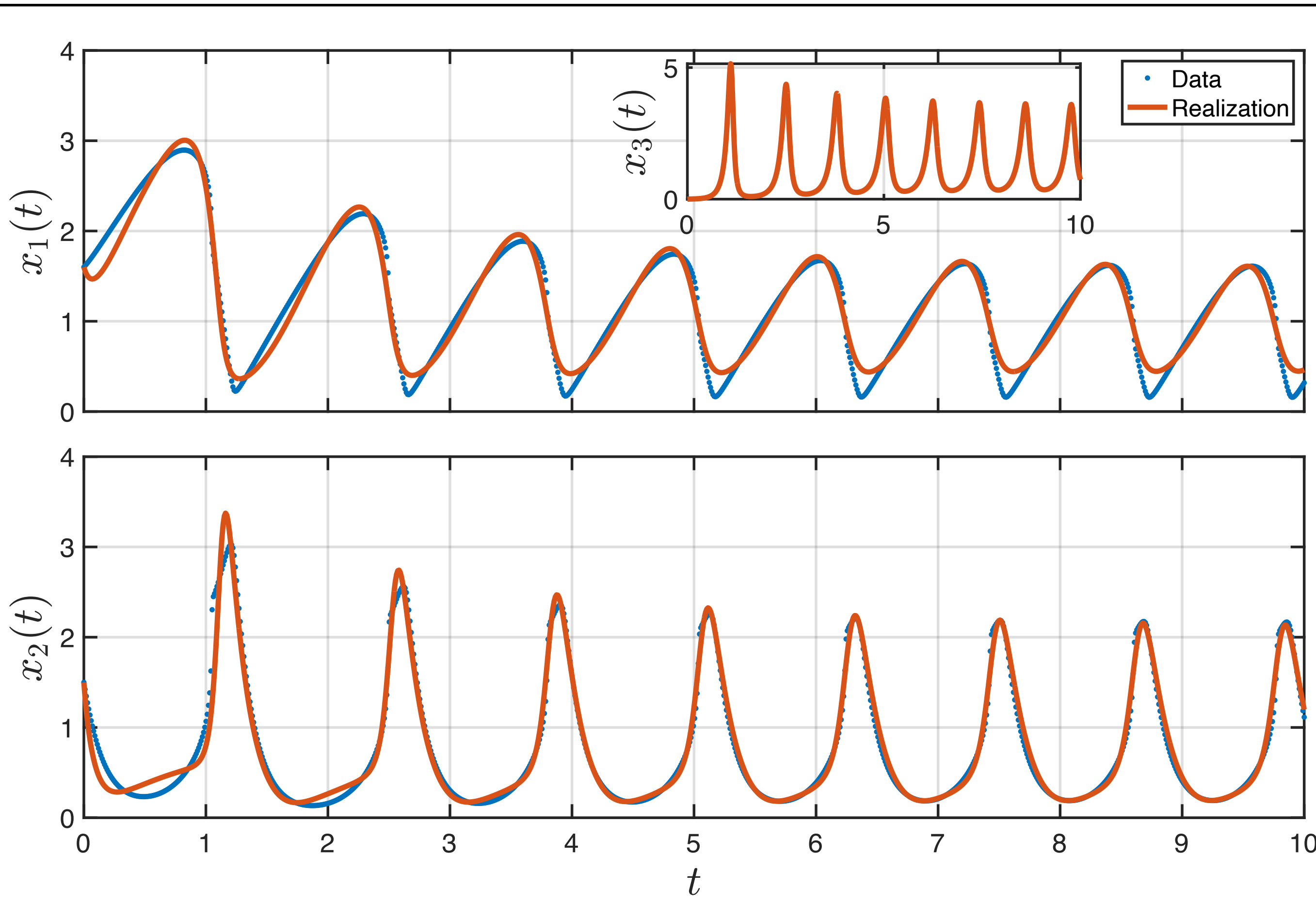
$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1^2 \\ x_1 x_2 \\ x_1 x_3 \\ x_2^2 \\ x_2 x_3 \\ x_3^2 \end{bmatrix}$$

Library  
(m functions)

# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 3$  species (2 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )

$$\theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,10} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,10} \\ \theta_{3,1} & \theta_{3,2} & \cdots & \theta_{3,10} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

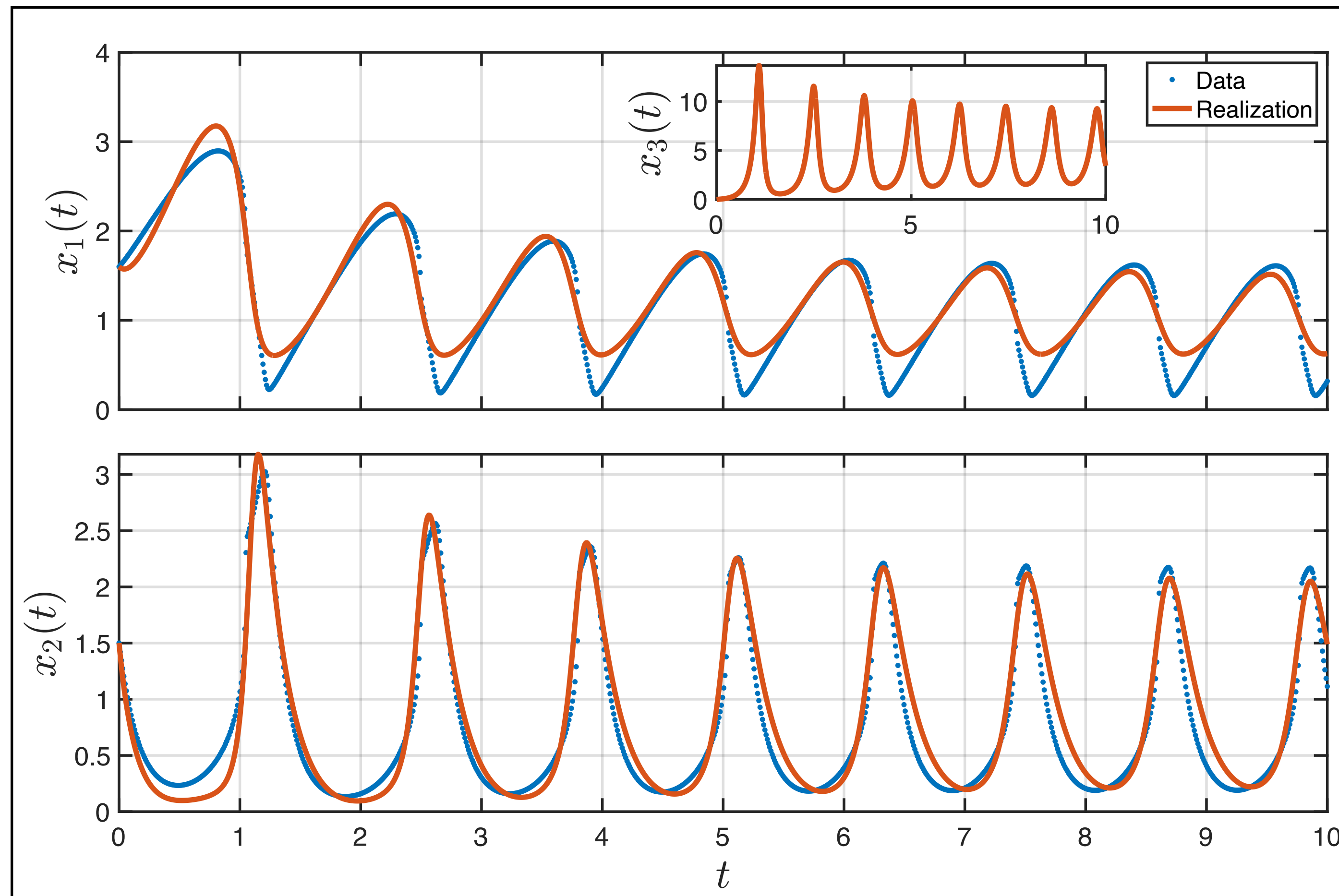
$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1^2 \\ x_1 x_2 \\ x_1 x_3 \\ x_2^2 \\ x_2 x_3 \\ x_3^2 \end{bmatrix}$$

Library  
(m functions)



# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

- $n = 3$  species (2 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )

$$\theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,10} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,10} \\ \theta_{3,1} & \theta_{3,2} & \cdots & \theta_{3,10} \end{bmatrix}$$

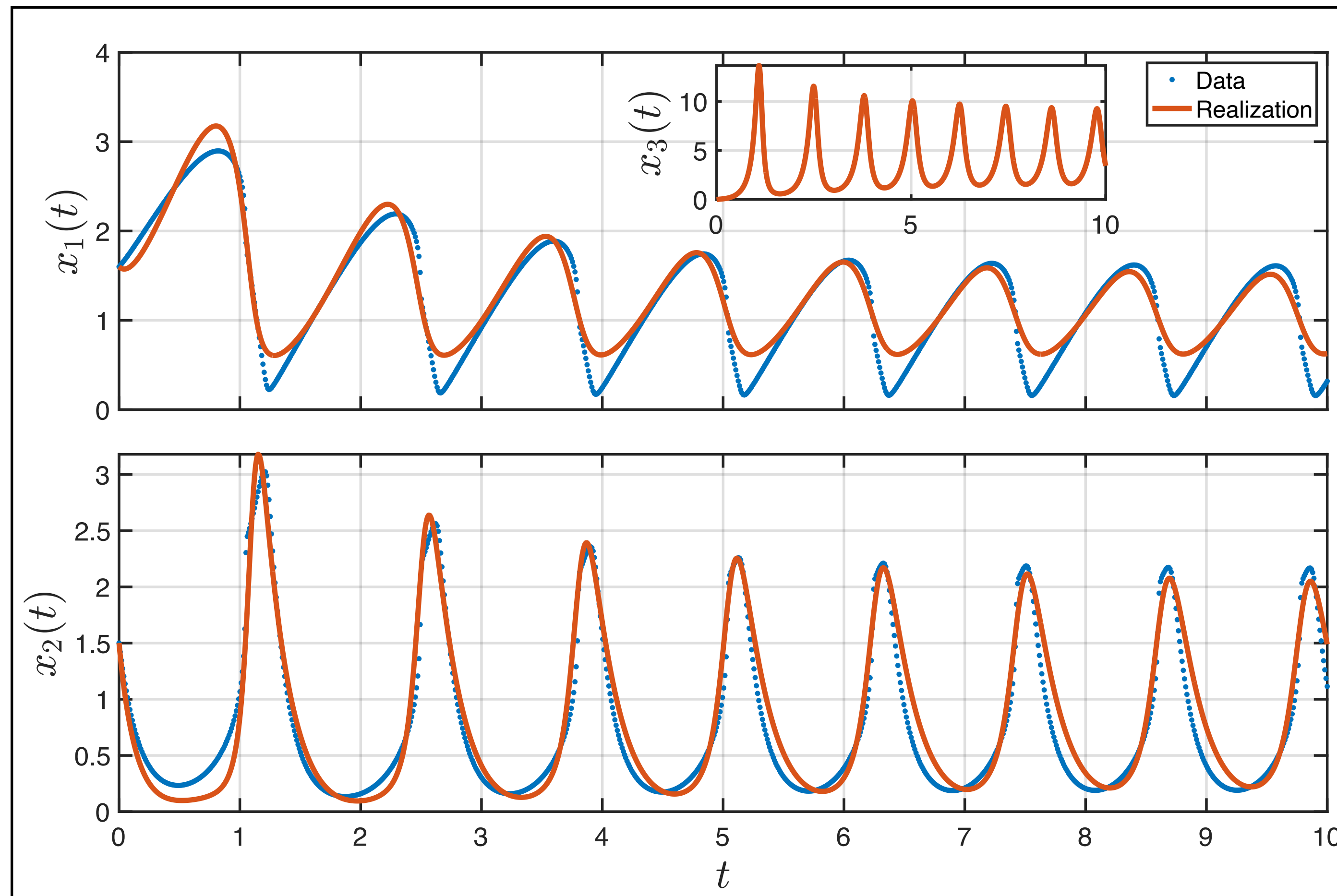
$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1^2 \\ x_1 x_2 \\ x_1 x_3 \\ x_2^2 \\ x_2 x_3 \\ x_3^2 \end{bmatrix}$$

Library  
(m functions)

# Example 2: Glycolytic Oscillator

**Desired trajectory**  $\{y_T(t)\}_{0 \leq t \leq T}$



**Model:**  $\dot{x} = \theta \Phi(x)$

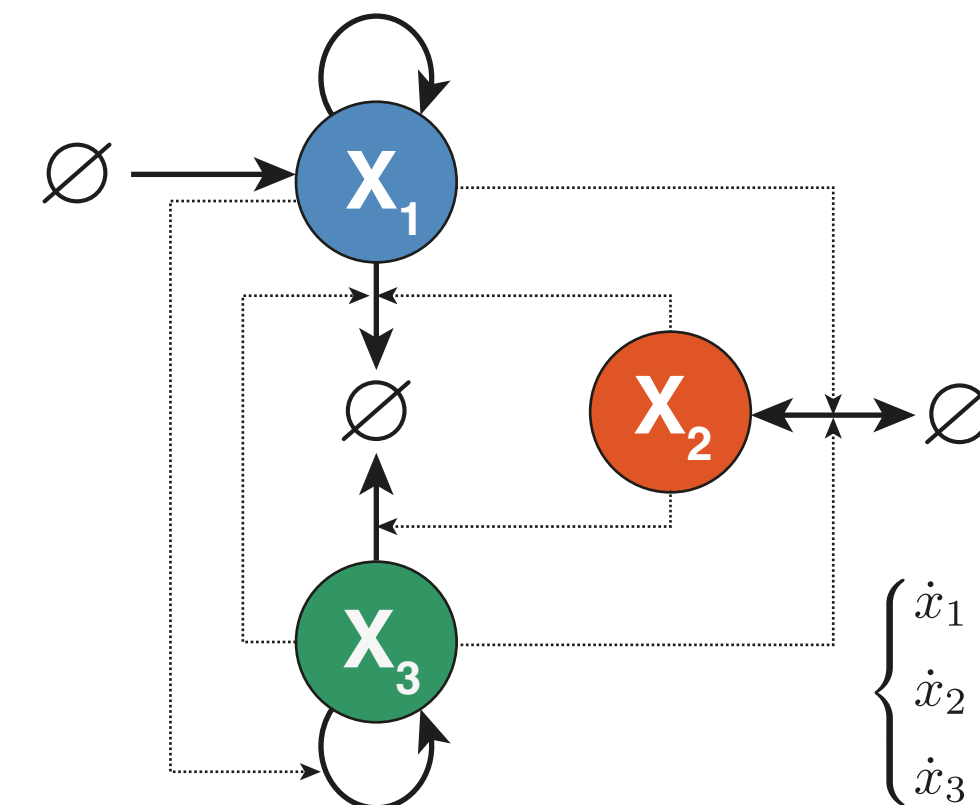
- $n = 3$  species (2 observed + 1 hidden)
- Bimolecular reactions (up to quadratic)
- First: dense fit ( $\lambda = 0$ )
- Then: sparse fit (larger  $\lambda$ )

$$\theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,10} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,10} \\ \theta_{3,1} & \theta_{3,2} & \cdots & \theta_{3,10} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\Phi(x) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1^2 \\ x_1 x_2 \\ x_1 x_3 \\ x_2^2 \\ x_2 x_3 \\ x_3^2 \end{bmatrix}$$

Library  
(m functions)



$$\begin{cases} \dot{x}_1 = \mu_0 + k_1 x_1 - (\gamma_{12} x_2 + \gamma_{13} x_3) x_1 \\ \dot{x}_2 = \eta_{21} x_1^2 + \eta_{23} x_3^2 - (\gamma_{20} + \gamma_{21} x_1 + \gamma_{22} x_2 + \gamma_{23} x_3) x_2 \\ \dot{x}_3 = k_3 x_1 + \eta_3 x_1^2 + x_3 (\alpha_1 x_1 + \alpha_3 x_3) - \gamma_{32} x_2 x_3 \end{cases}$$



# Take Home Message

$$\min_{\theta} \left\| \dot{x}_T - \theta \Phi(x_T) \right\|_2^2$$



<https://maurice-filo.github.io/>

**C.T.S.B**  
CONTROL THEORY & SYSTEMS BIOLOGY





# Take Home Message

Left minus right is not always right!

$$\dot{x} = \theta \Phi(x) \qquad \min_{\theta} \left\| \dot{x}_T - \theta \Phi(x_T) \right\|_2^2$$



[maurice.filo@bsse.ethz.ch](mailto:maurice.filo@bsse.ethz.ch)  
[mustafa.khammash@bsse.ethz.ch](mailto:mustafa.khammash@bsse.ethz.ch)

<https://maurice-filo.github.io/>

**C.T.S.B**  
CONTROL THEORY & SYSTEMS BIOLOGY

