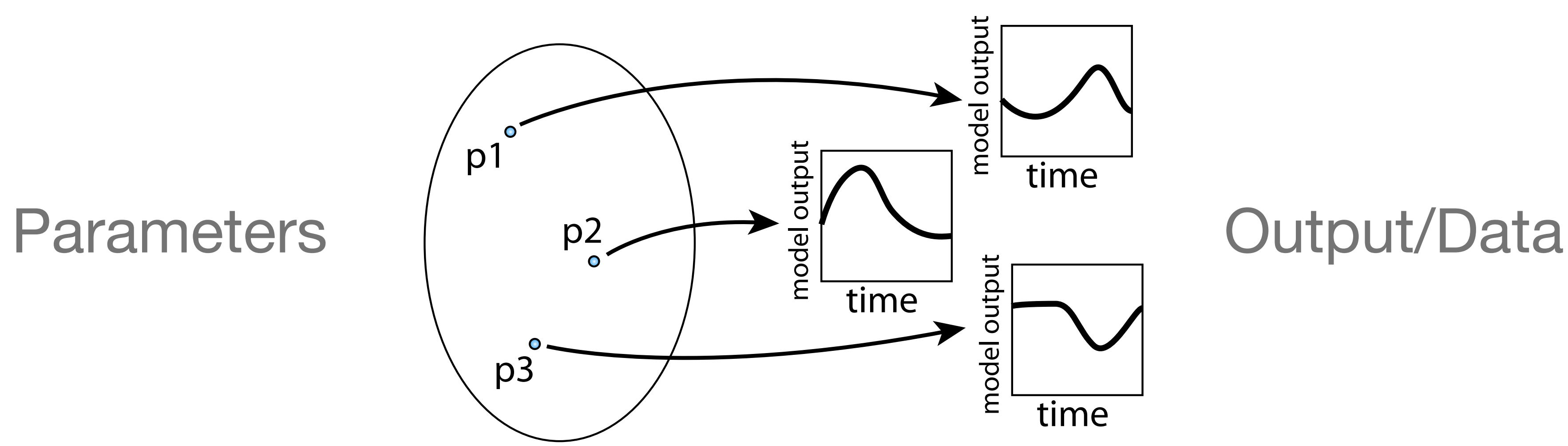
Introduction to Structural & Practical Identifiability

Marisa Eisenberg
Epid 814

Identifiability

- Identifiability — Is it possible to uniquely determine the parameters from the data?

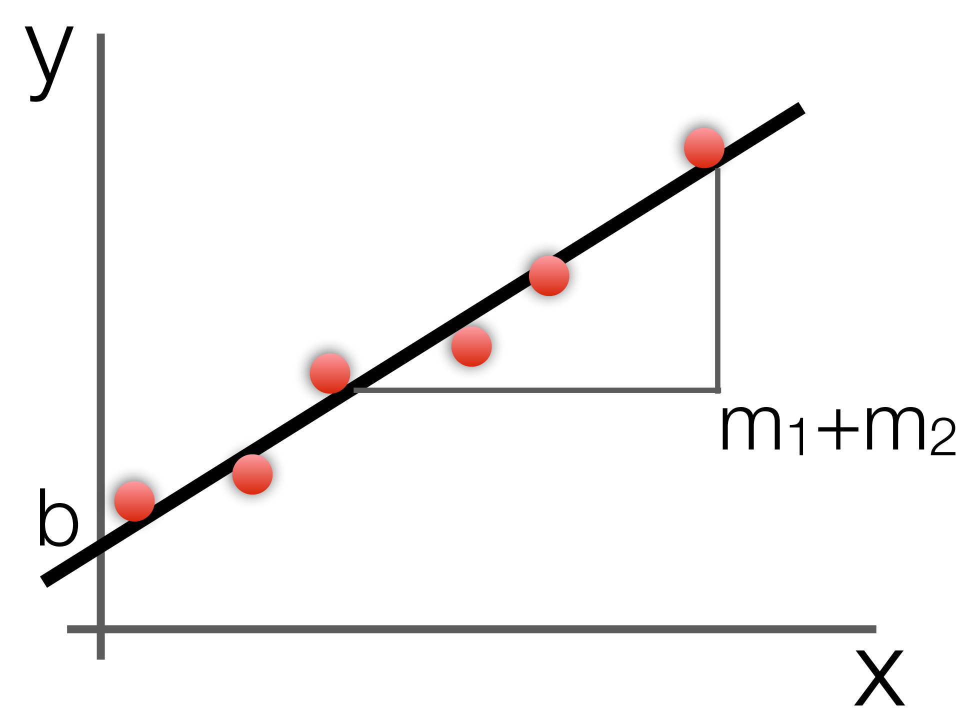


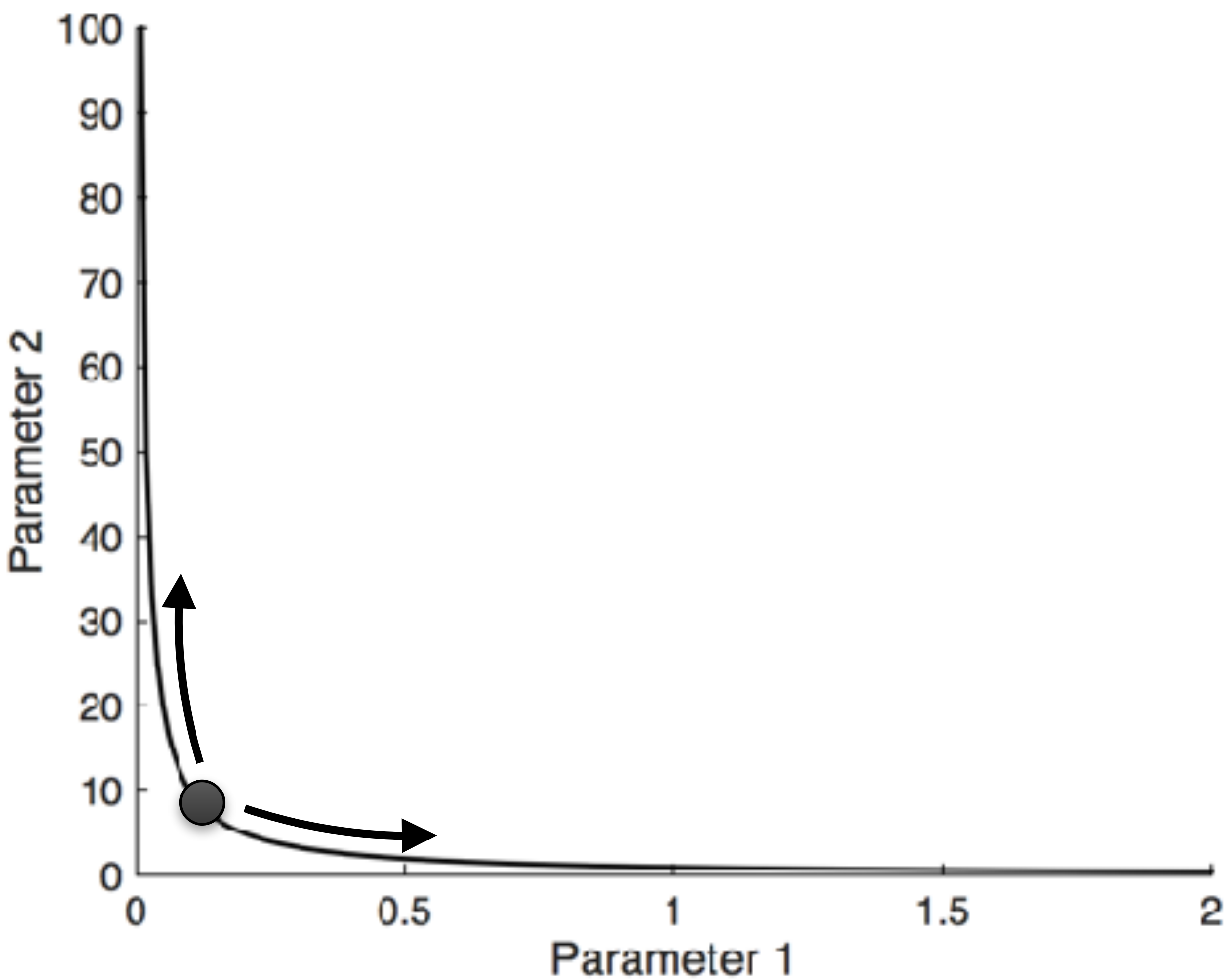
- Important problem in parameter estimation
- Many different approaches - statistics, applied math, engineering/systems theory

Ollivier 1990, Ljung & Glad 1994, Evans & Chappell 2000, Audoly et al 2003, Hengl et al. 2007, Chis et al 2011

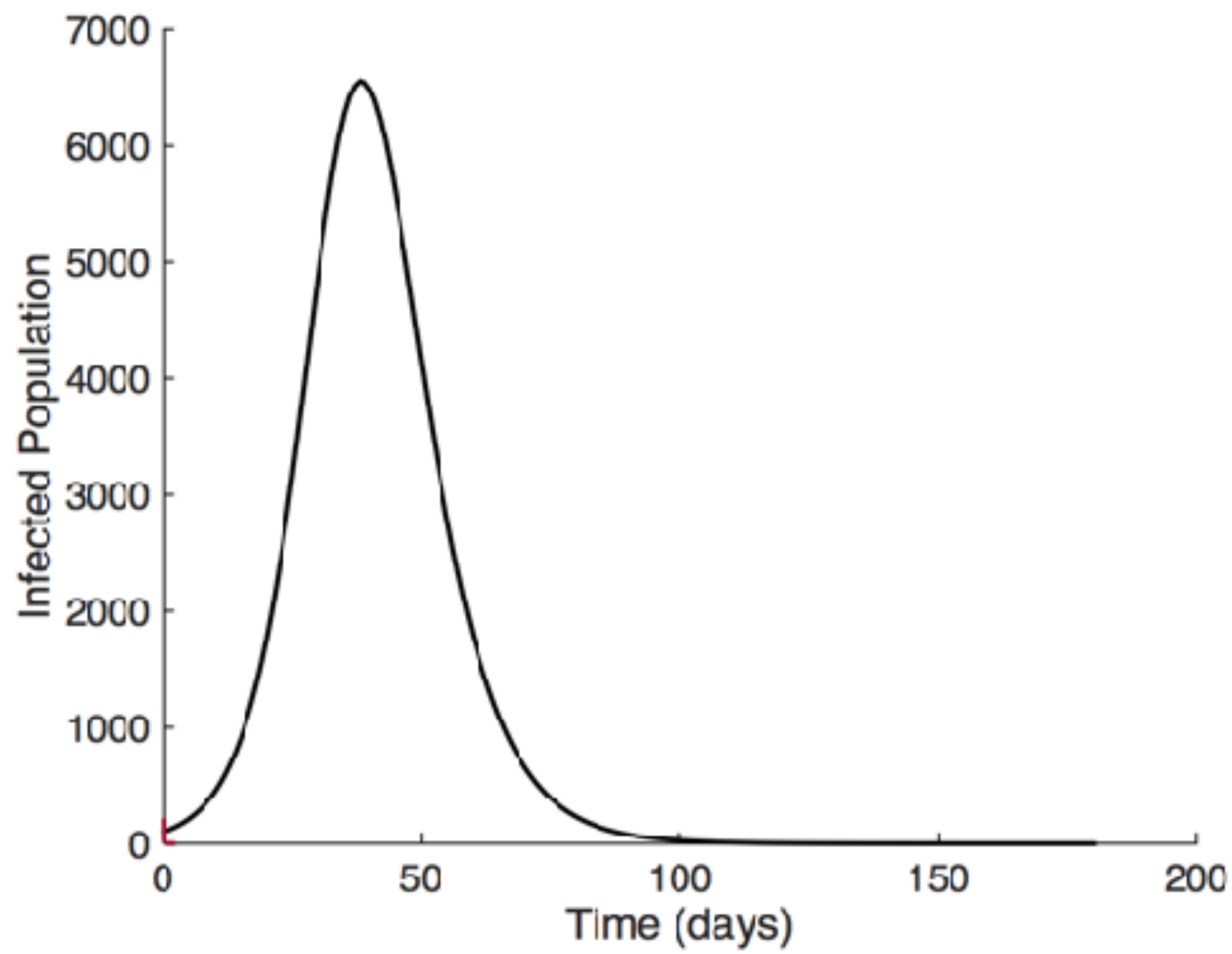
Identifiability

- Practical vs. Structural
  - Broad, sometimes overlapping categories
  - Noisy vs. perfect data
- Example: $y = (m_1 + m_2)x + b$
- Unidentifiability - can cause serious problems when estimating parameters
- Identifiable combinations

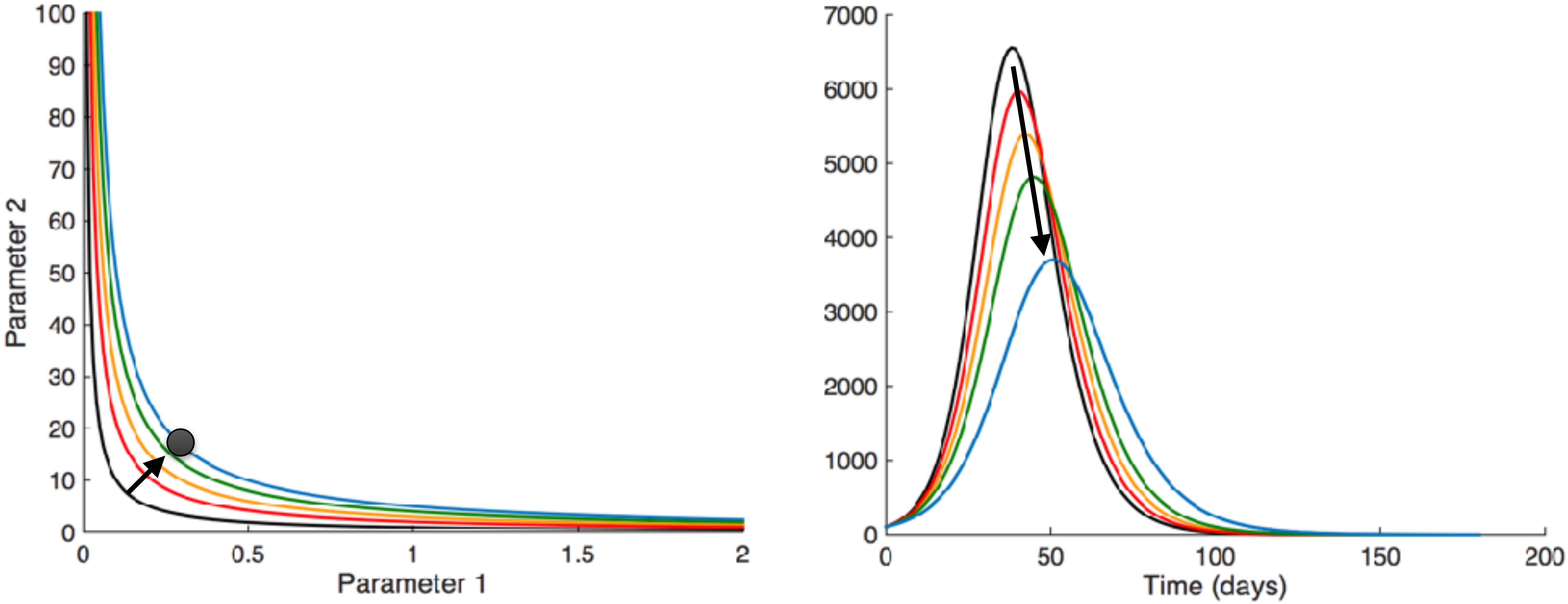




100
90
80
70
60
50
40
30
20
10
0
0
0.5
1
1.5
2
Parameter 2
Parameter 1



7000
6000
5000
4000
3000
2000
1000
0
0
50
100
150
200
Infected Population
Time (days)



Parameter 2
Parameter 1
Time (days)

Structural Identifiability

- Assumes best case scenario - data is known perfectly at all times
- Unrealistic!
- But, necessary condition for practical identifiability with real, noisy data

Structural Identifiability

- Reveals identifiable combinations and how to restructure the model so that it is identifiable

- Can give a priori information, help direct experiment design

Categories to consider

- Structural vs. practical identifiability
- Analytical vs. numerical methods
- Global vs. local results (in parameter space)

Key Concepts

- Identifiability vs. unidentifiability
  - Practical vs. structural, local vs. global
  - Can be in between, e.g. quasi-identifiable
- Identifiable Combinations
- Reparameterization
- Related questions: observability, distinguishability & model selection

Methods we'll talk about today

- Fisher information matrix - structural or practical, local, analytical or numerical method
- Profile likelihood - structural or practical, local, numerical method
- Differential Algebra Approach - structural identifiability, global, analytical method

Simple Methods

- Simulated data approach
- If you have a small system, you can even plot the likelihood surface (typically can't though—more on this with profile likelihoods)

Numerical Methods for Identifiability Analysis

Numerical Approaches to Identifiability

- Analytical approaches can be slow, sometimes have limited applicability
- Wide range of numerical approaches
  - Sensitivities/Fisher Information Matrix
  - Profile Likelihood
  - Many others (e.g. Bayesian approaches, etc.)

Numerical Approaches to Identifiability

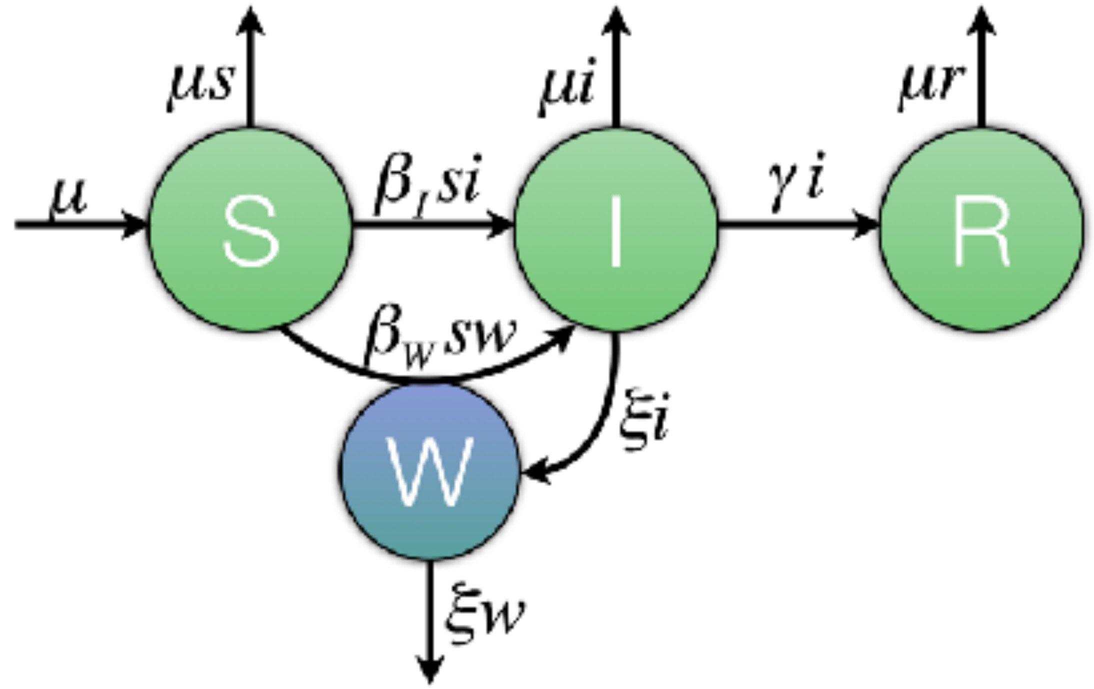
- Most can do both structural & practical identifiability
- Wide range of applicable models, often (relatively) fast
- Typically only local

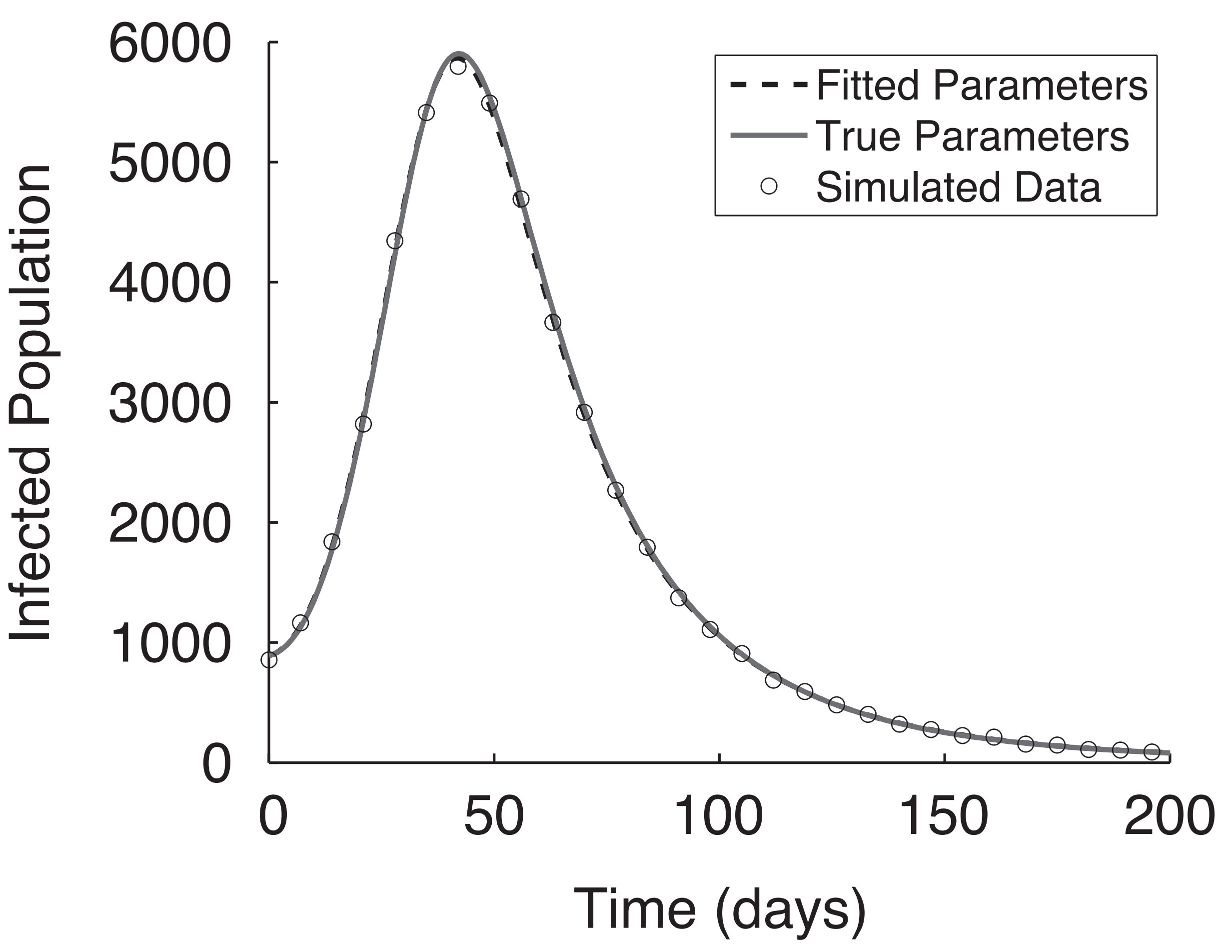
Simple Simulation Approach

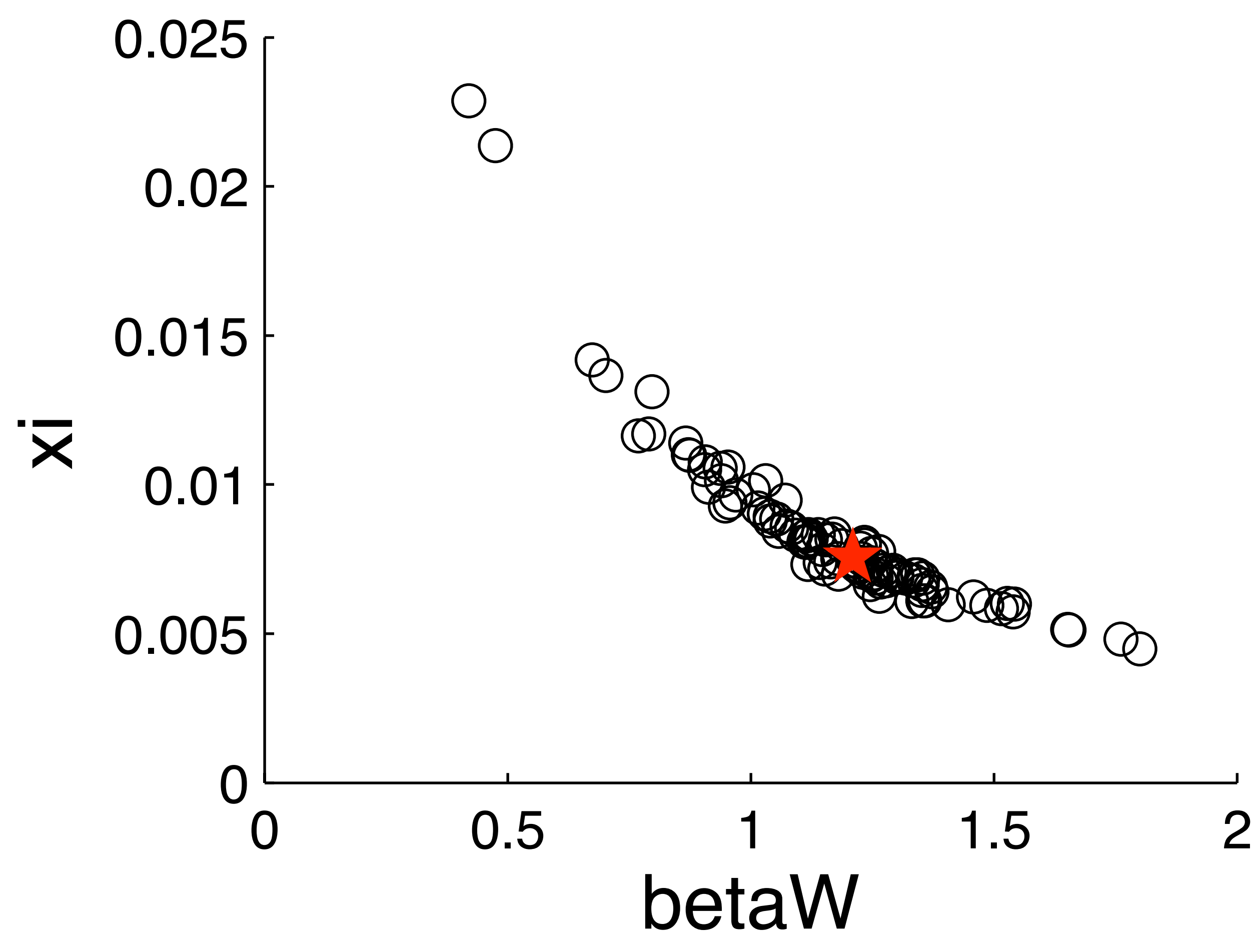
- Simulate data using a single set of ‘true’ parameter values
  - Without noise for structural identifiability
  - With noise for practical identifiability (in this case generate multiple realizations of the data)

Simple Simulation Approach

- Fit your simulated data from multiple starting points and see where your estimates land
- If they all return to the ‘true’ parameters, likely identifiable, if they do not—may be problems
- Note—unidentifiability when estimating with ‘perfect’, noise-free simulated data is most likely structural


μs
μi
μr
μ
S
$\beta_I si$
I
γi
R
$\beta_W sw$
ξi
W
ξw


Fitted Parameters
True Parameters
Simulated Data
Infected Population
Time (days)
0
1000
2000
3000
4000
5000
6000
0
50
100
150
200


xi
betaW
0
0.005
0.01
0.015
0.02
0.025
0
0.5
1
1.5
2

Parameter Sensitivities

- Output sensitivity matrix (design matrix)
- Closely related to identifiability
- Insensitive parameters
- Dependencies between columns

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

Fisher Information Matrix

- FIM - N_P x N_P matrix

$$\left[\mathcal{I}\left(\theta\right)\right]_{i,j} = \mathrm{E}\left[\left(\frac{\partial}{\partial\theta_i}\log f(X;\theta)\right)\left(\frac{\partial}{\partial\theta_j}\log f(X;\theta)\right)\middle|\theta\right]$$

- Useful in testing practical & structural ID - represents amount of information that the output **y** contains about parameters **p**

- Cramer-Rao Bound: $FIM^{-1} \leq Cov(\mathbf{p})$

- Rank(FIM) = number of identifiable parameters/ combinations

Fisher Information Matrix

- Special case when errors are normally distributed

$$F = X^T W X$$

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

W = weighting matrix

Fisher Information Matrix

- For looking at structural ID, often just use

$$F = X^T X$$

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

Identifiability & the FIM

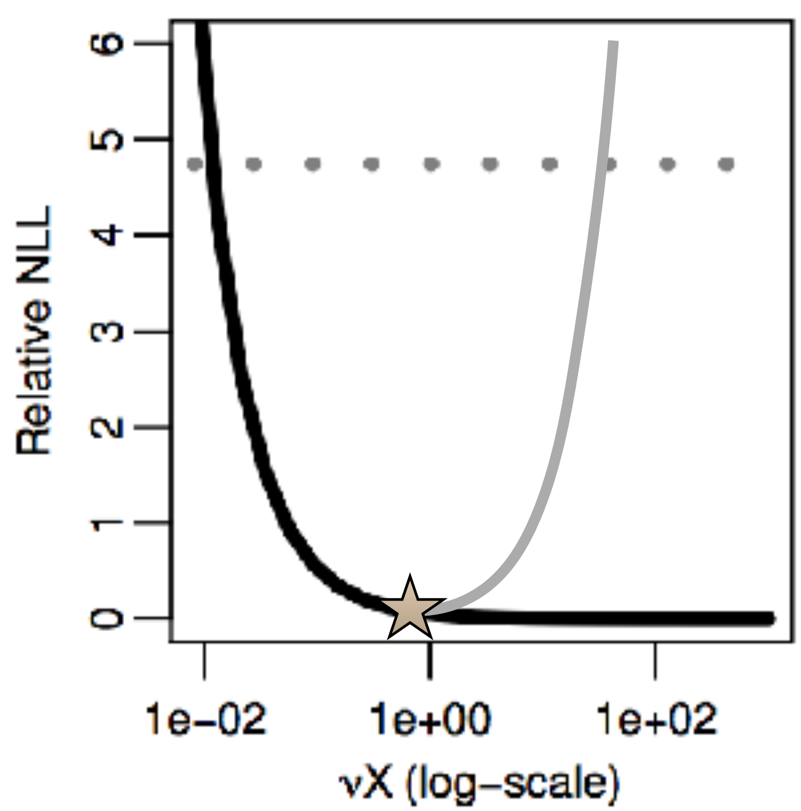
- Covariance matrix/confidence interval estimates from Cramér-Rao bound: $\text{Cov} \geq \text{FIM}^{-1}$
  - e.g. large confidence interval ⇒ probably at least practically unID
  - Often can detect structural unID as 'near-infinite' (gigantic) variances in $\text{Cov} \sim \text{FIM}^{-1}$

Identifiability & the FIM

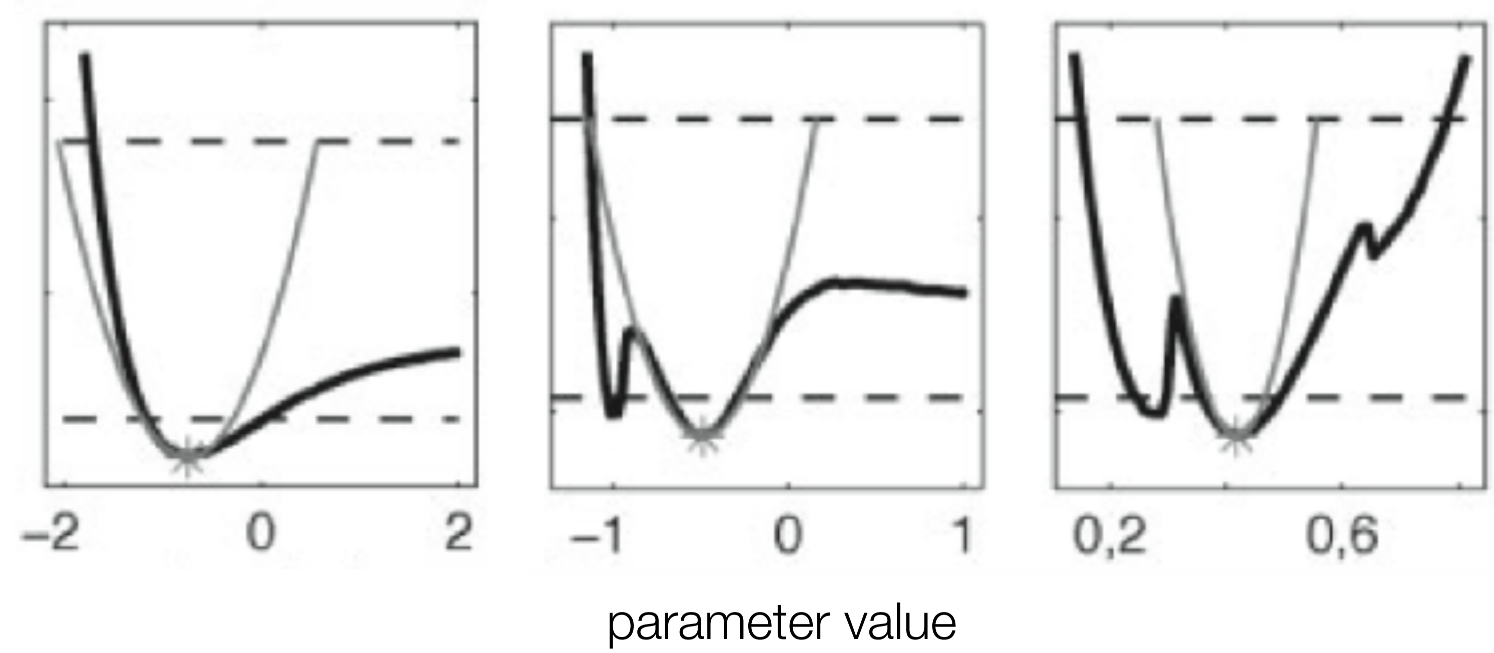
- **Rank of the FIM** is number of identifiable combinations/parameters - can do a lot by testing sub-FIMs and versions of the FIM
- Use FIM to find blocks of related parameters & how many to fix (not estimate)
- Identifiable combinations - can often see what parameters are related, but don't know form
  - Interaction of combinations

Identifiability & the FIM

- But, be careful—FIM is local & asymptotic
- Local approximation of the curvature of the likelihood



Brouwer, Meza, Eisenberg 2017



Raue et al. 2010

Profile Likelihood

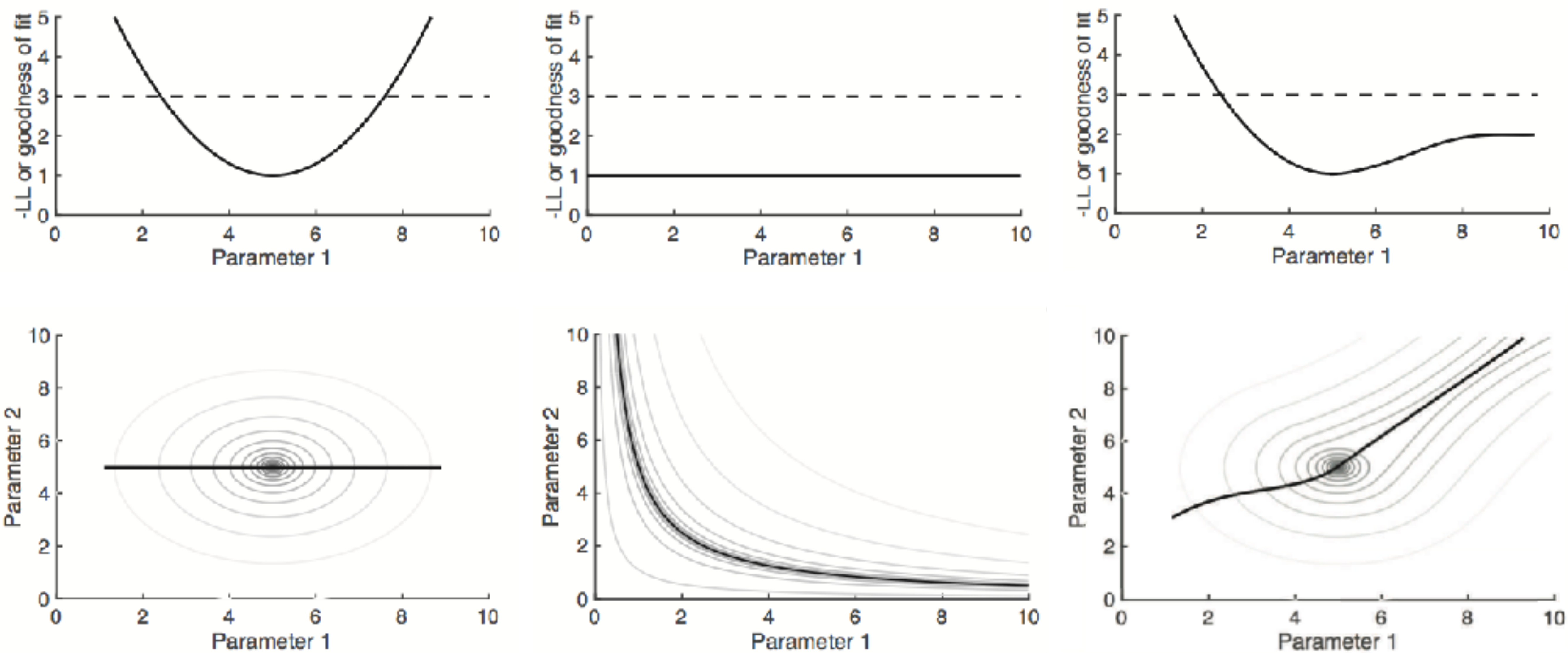
- Want to examine likelihood surface, but often high-dimensional
- Basic Idea: ‘profile’ one parameter at a time, by fixing it to a range of values & fitting the rest of the parameters
- Gives best fit at each point
- Evaluate curvature of likelihood to determine confidence bounds on parameter (and to evaluate parameter uncertainty)

Profile Likelihood

- Choose a range of values for parameter p_i
- For each value, fix p_i to that value, and fit the rest of the parameters
- Report the best likelihood/RSS/cost function value for that p_i value
- Plot the best likelihood values for each value of p_i— this is the profile likelihood

Profile Likelihoods

identifiable | structurally unidentifiable | practically unidentifiable



Profile Likelihood & ID

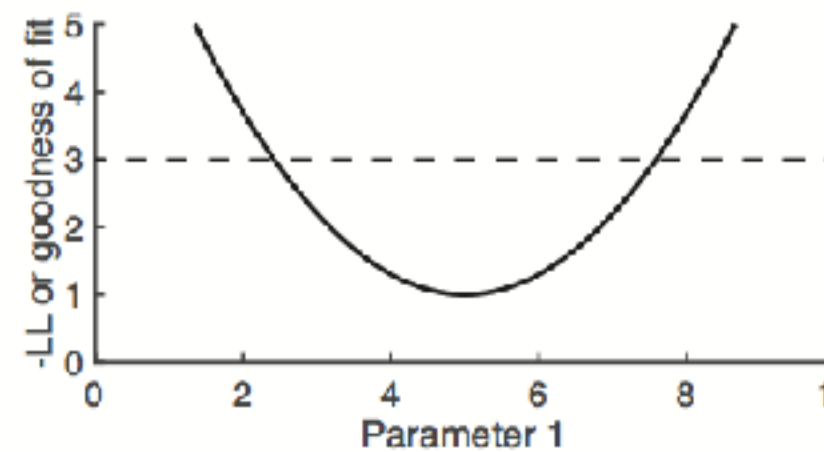
- Can generate confidence bounds based on the curvature of the profile likelihood
- Flat or nearly flat regions indicate identifiability issues
- Can generate simulated ‘perfect’ data to test structural identifiability

Profile-based Confidence Intervals

- The shape of the likelihood—more specifically, the likelihood ratio:

$$2(NLL(p) - NLL(\hat{p}))$$

  is approximately χ^2-distributed when the sample size is large

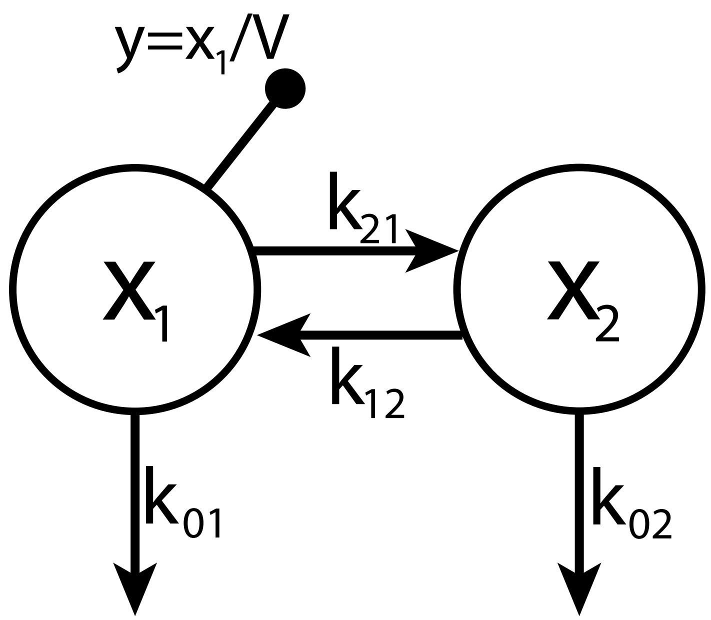


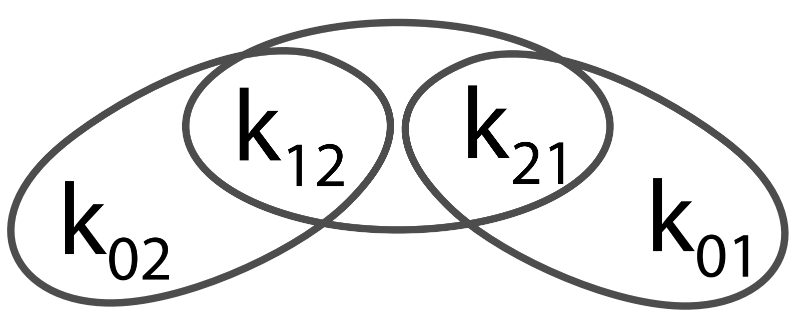
- From this, we can calculate a threshold to define a confidence interval, based on the appropriate percentile of the χ^2

Profile Likelihood

- Can also help reveal the form of identifiable combinations
  - Look at relationships between parameters when profiling
  - However, can be problematic when too many degrees of freedom
- Similar to pairwise plots with sampling-based methods (e.g. MCMC)

2-Compartment Example





$$1/V = a_1 \Rightarrow V = 1/a_1$$

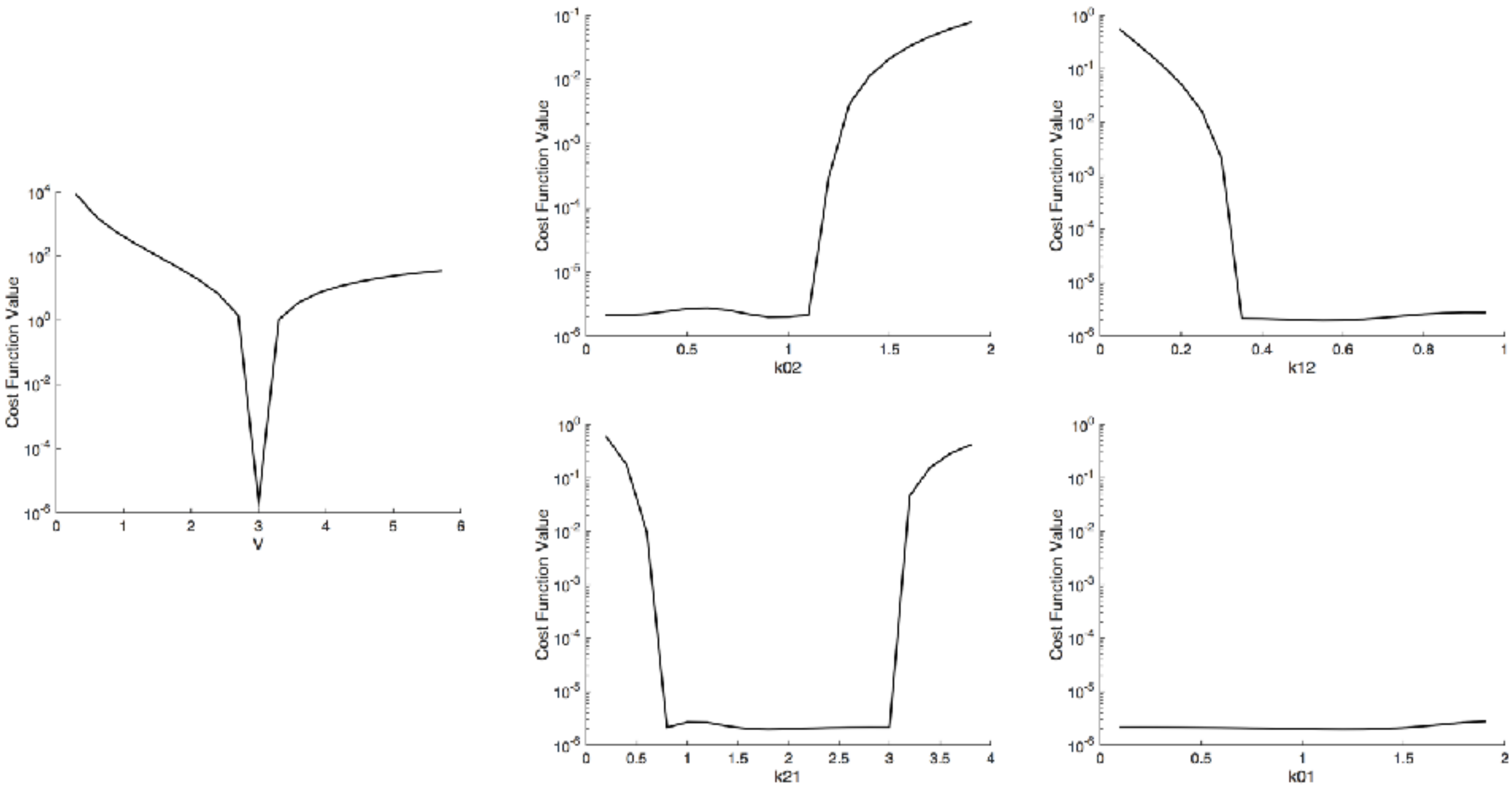
$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

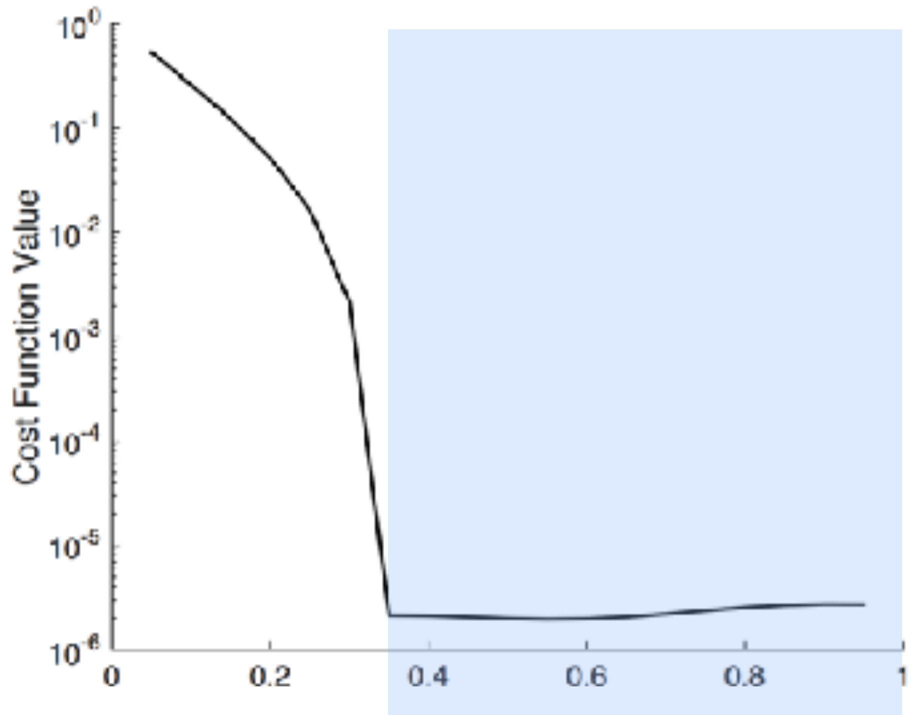
Eisenberg & Hayashi, Math Biosciences 2014

Profile Likelihoods

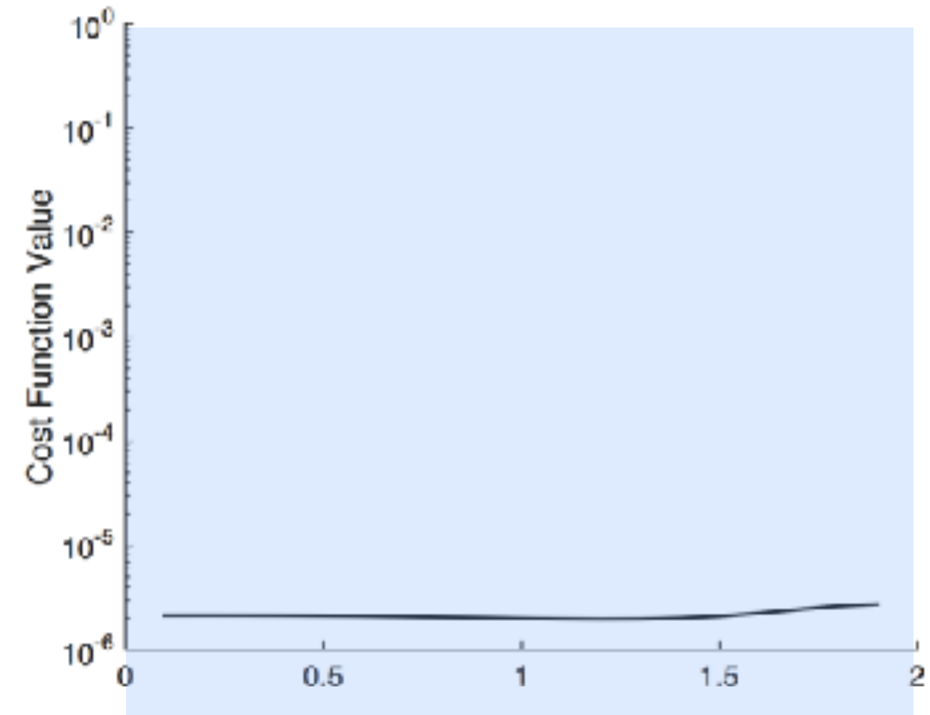


Eisenberg & Hayashi, Math Biosciences 2014

Parameter Relationships



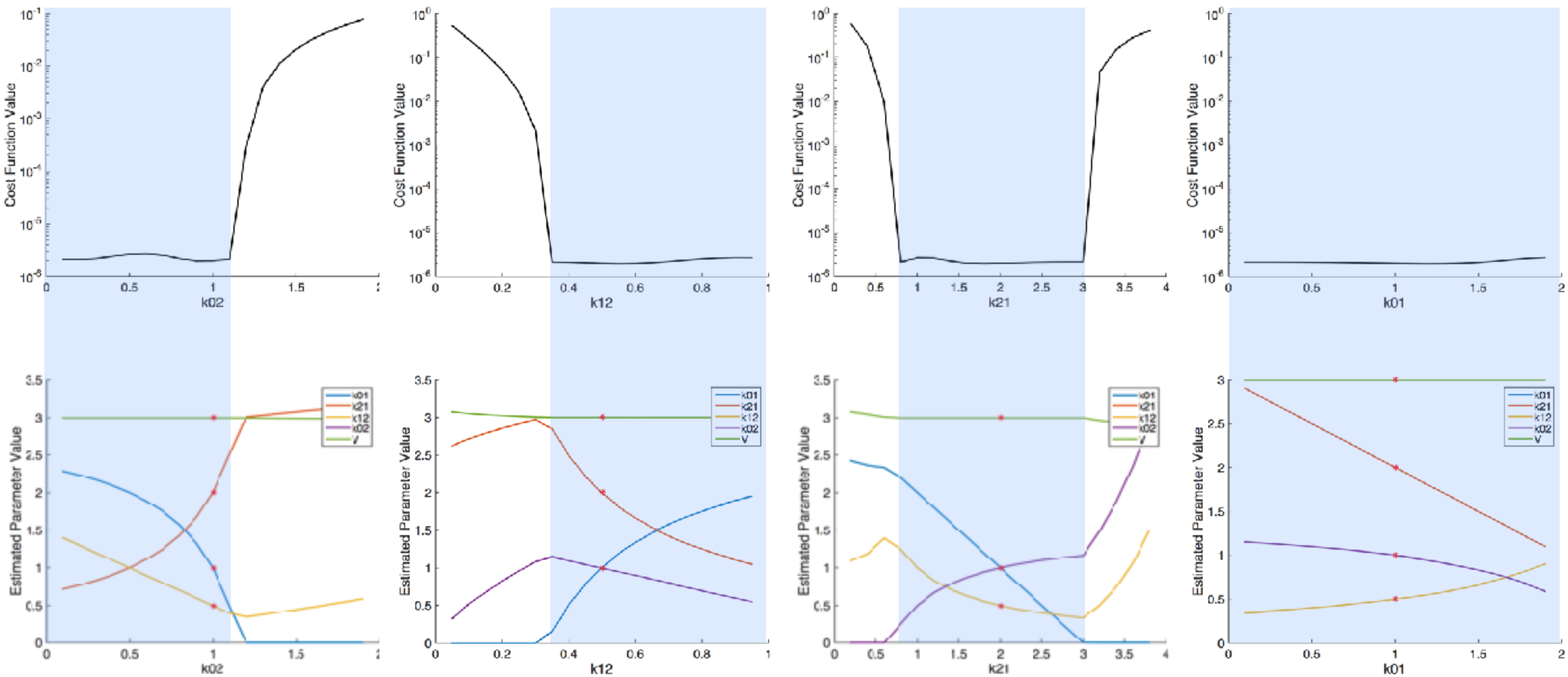
k12



k01

Eisenberg & Hayashi, Math Biosciences 2014

Parameter Relationships



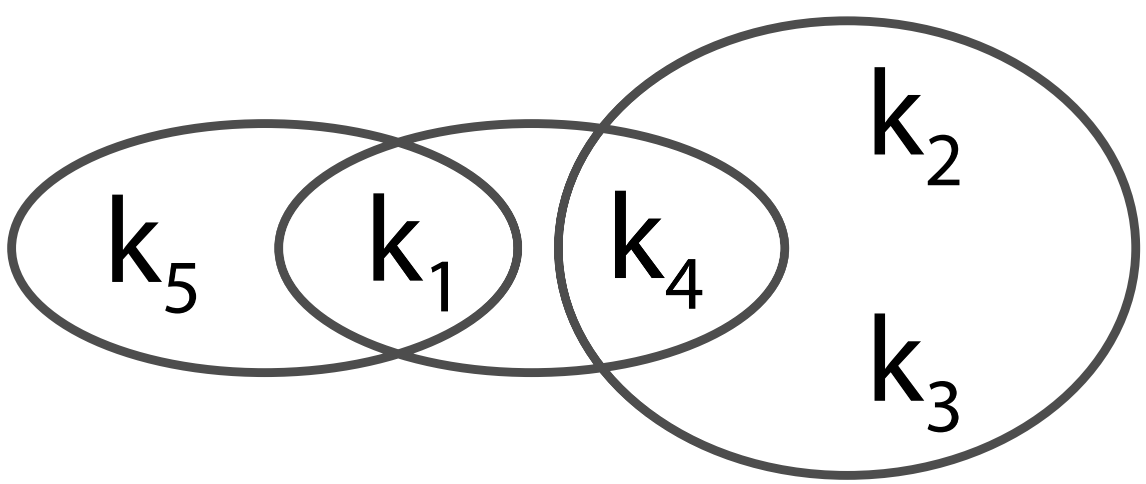
Eisenberg & Hayashi, Math Biosciences 2014

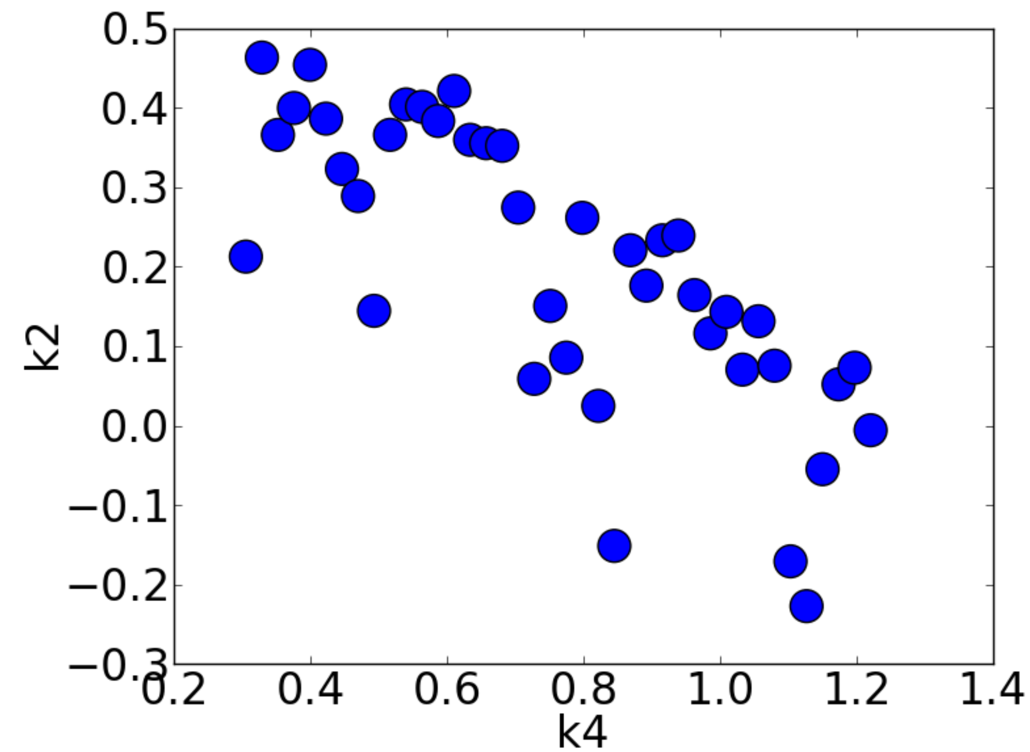
Some potential issues

$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$

$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$

$$y = x_1 / V$$

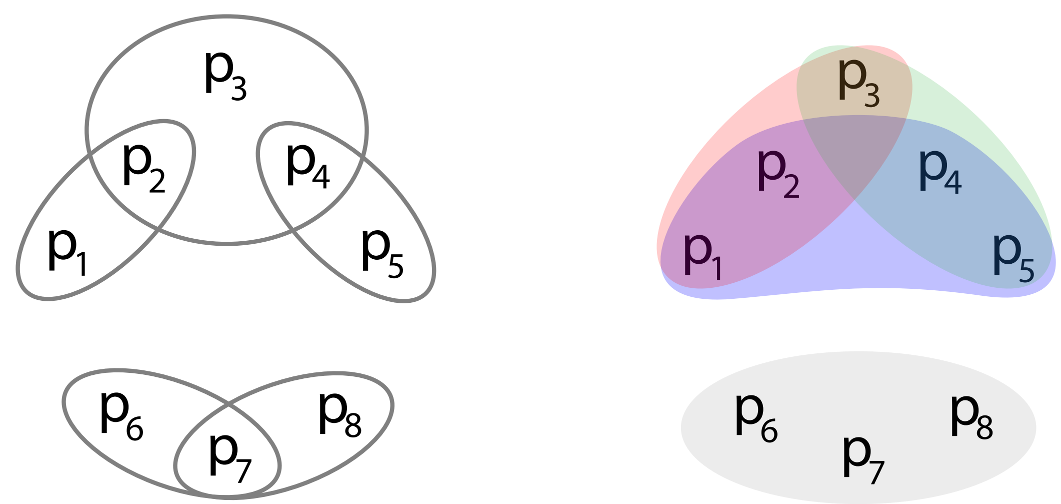




Eisenberg & Hayashi, Math Biosciences 2014

FIM Subset Approach

- Basic idea - evaluate the rank of the FIM for subsets of parameters to elucidate the structure of the identifiable combinations



- Can then combine this with profile likelihood approach by Raue et al. to determine the form of the combinations

Eisenberg & Hayashi, Math Biosciences 2014

FIM Subset Approach

- Use the FIM rank to select subsets of parameters which are *nearly full rank* (i.e. which become full rank if any single parameter is fixed)

- Use these subsets when likelihood profiling to determine all parameter relationships

- Polynomial interpolation to recover identifiable combinations

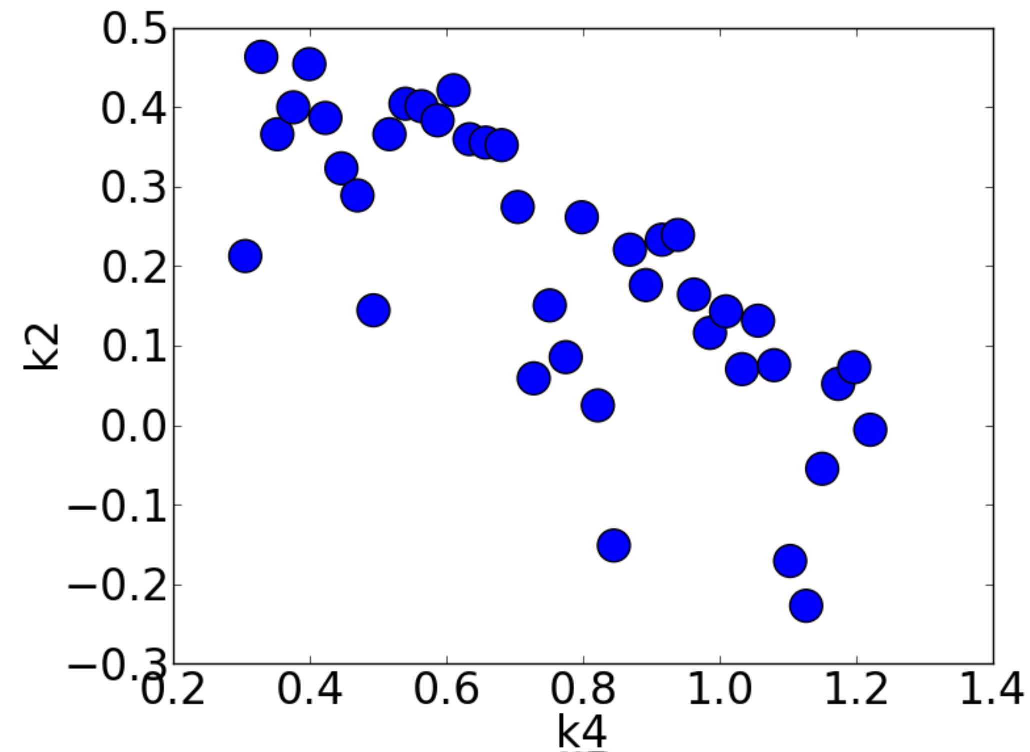
Eisenberg & Hayashi, Math Biosciences 2014

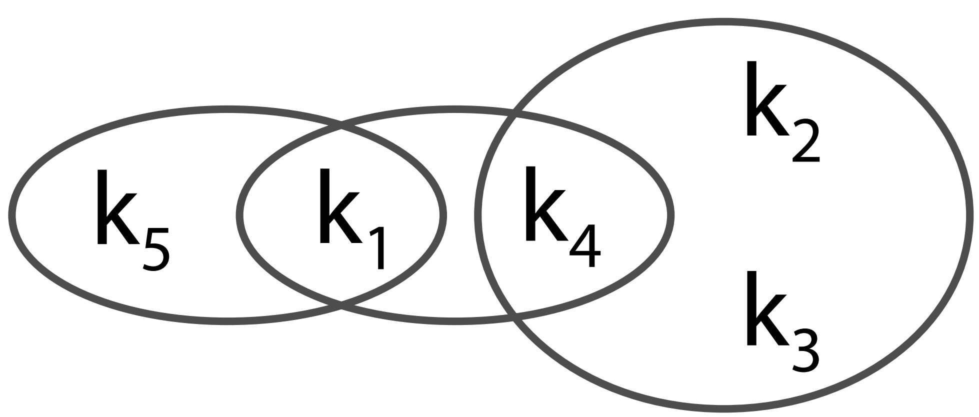
Example Model

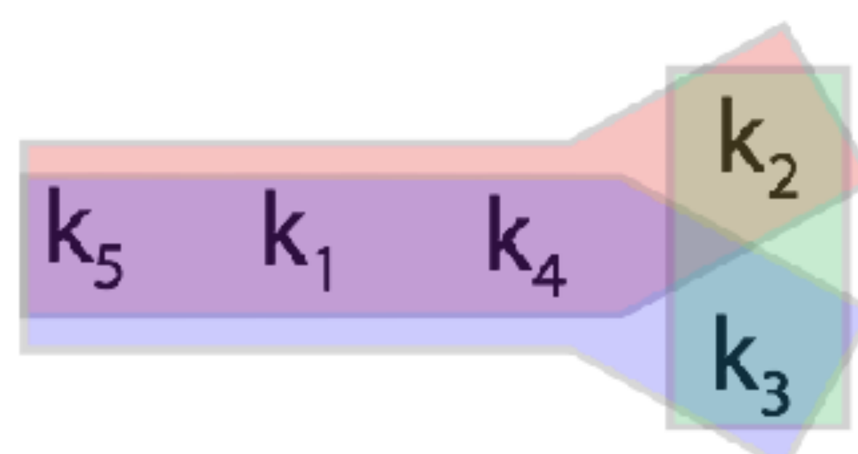
$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$

$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$

$$y = x_1 / V$$



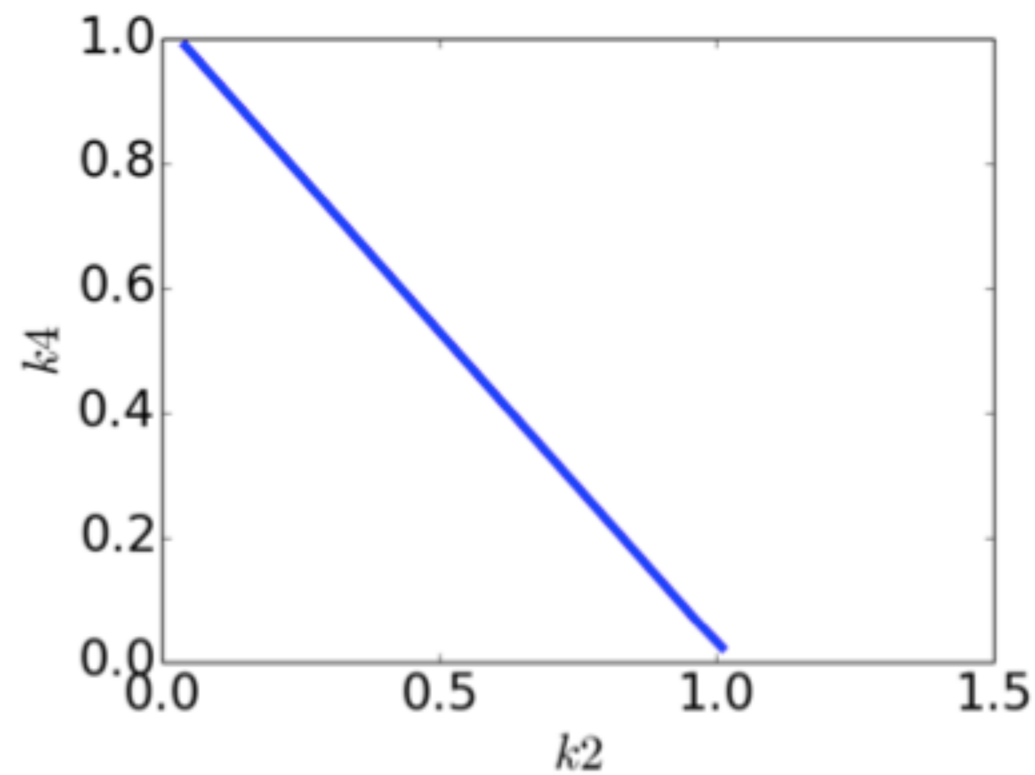


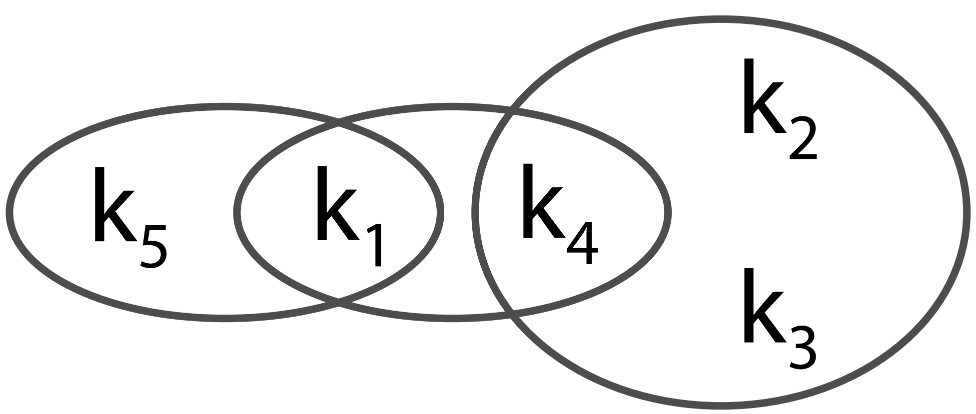


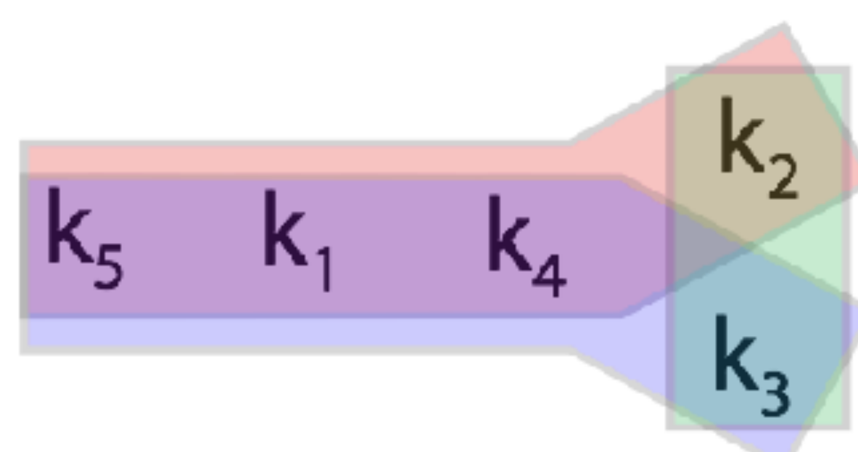
Eisenberg & Hayashi 2014, *in review*

Example Model

$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$
$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$
$$y = x_1 / V$$







Eisenberg & Hayashi 2014, *in review*

Analytical Methods for Structural Identifiability

Methods for Structural Identifiability

- **Laplace transform** - linear models only
- **Taylor series approach** - more broad application, but only local info & may not terminate
- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem
- **Differential algebra approach** - rational function ODE models, global info

Bellman 1970, Cobelli & DiStefano 1980, Evans & Chappell 2000, Ollivier 1990, Ljung & Glad 1994, Audoly et al 2003

Methods for Structural Identifiability

- **Laplace transform** - linear models only
- **Taylor series approach** - more broad application, but only local info & may not terminate
- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem
- **Differential algebra approach** - rational function ODE models, global info

Bellman 1970, Cobelli & DiStefano 1980, Evans & Chappell 2000, Ollivier 1990, Ljung & Glad 1994, Audoly et al 2003

Differential Algebra Approach

- Basic idea: use substitution & differentiation to eliminate all variables except for observed output (y)
- Clear (divide by) the coefficient for highest derivative term(s)
- This is called the **input-output equation(s)**
- Contains all structural identifiability info for the model

Differential Algebra Approach

- Use the coefficients to solve for identifiability of the model
- If unidentifiable, determine identifiable combinations
- Find identifiable reparameterization of the model?
- Easier to see with an example—

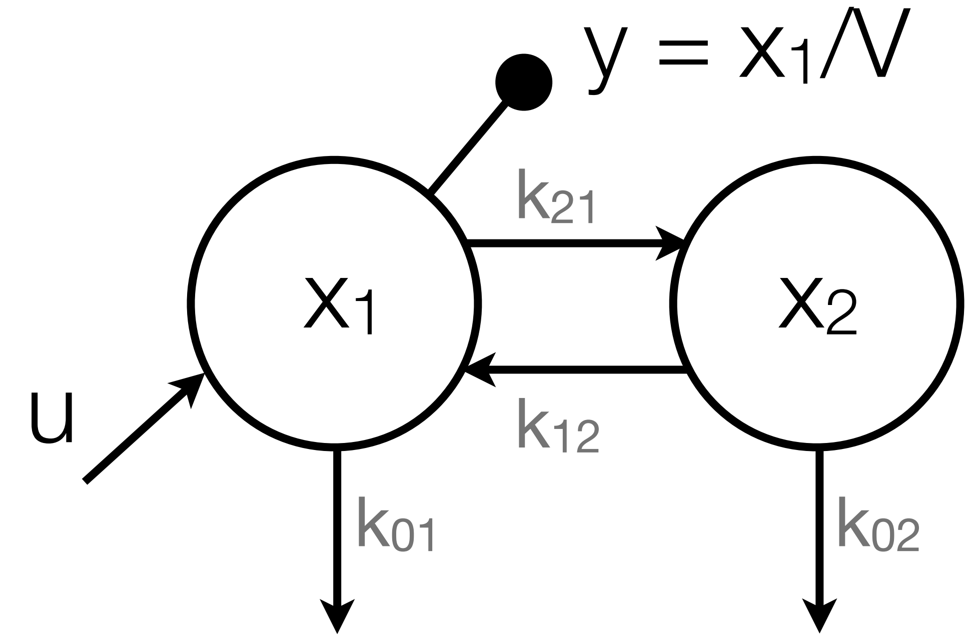
2-Compartment Example

- Linear 2-Comp Model

$$\dot{x}_1 = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)x_1$$

$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$

$$y = x_1 / V$$



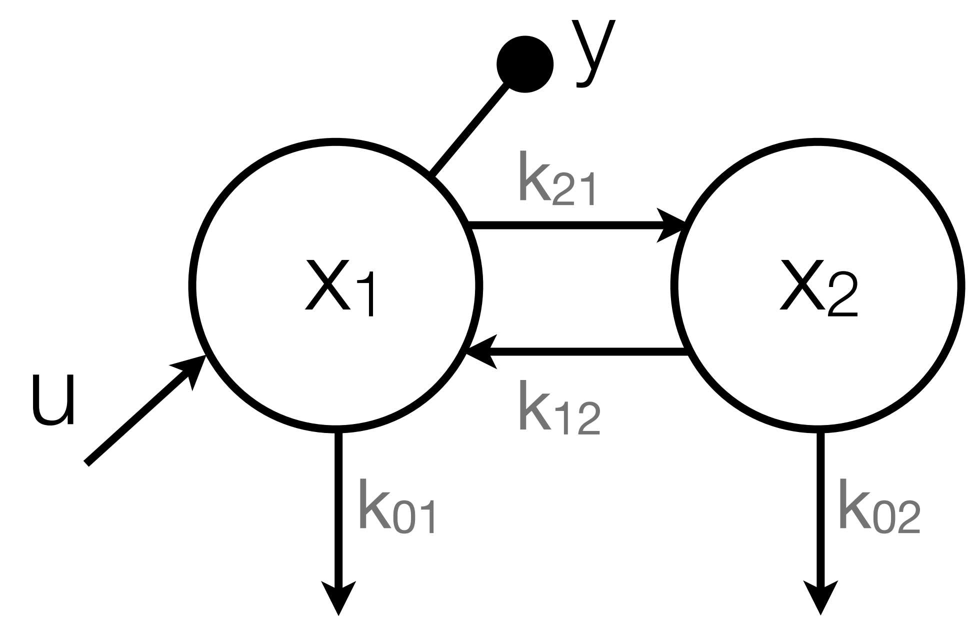
- state variables (x)
- measurements (y)
- known input (u) (e.g. IV injection)

2-Compartment Example

$$\dot{x}_1 = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)x_1$$

$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$

$$y = x_1 / V$$

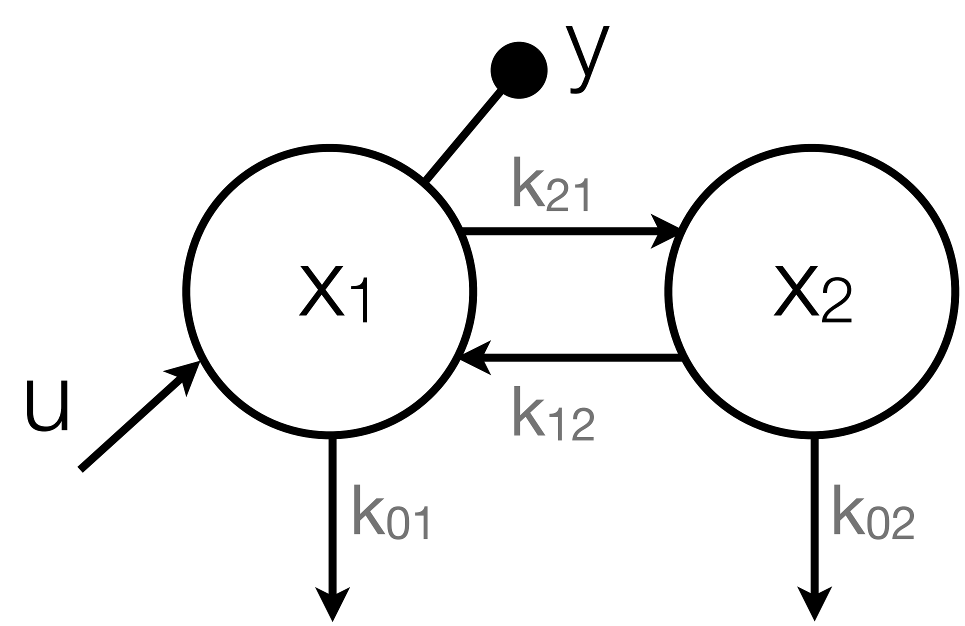


2-Compartment Example

$$y = x_1 / V$$

$$\dot{x}_1 = u + k_{12} x_2 - \left(k_{01} + k_{21}\right) x_1$$

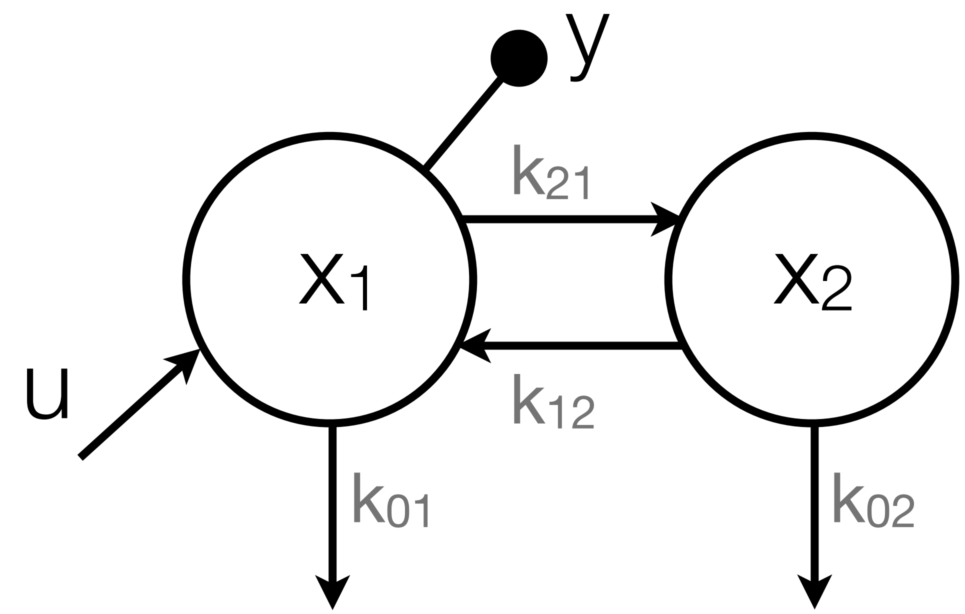
$$\dot{x}_2 = k_{21} x_1 - \left(k_{02} + k_{12}\right) x_2$$



2-Compartment Example

$$\dot{y}V = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)yV$$

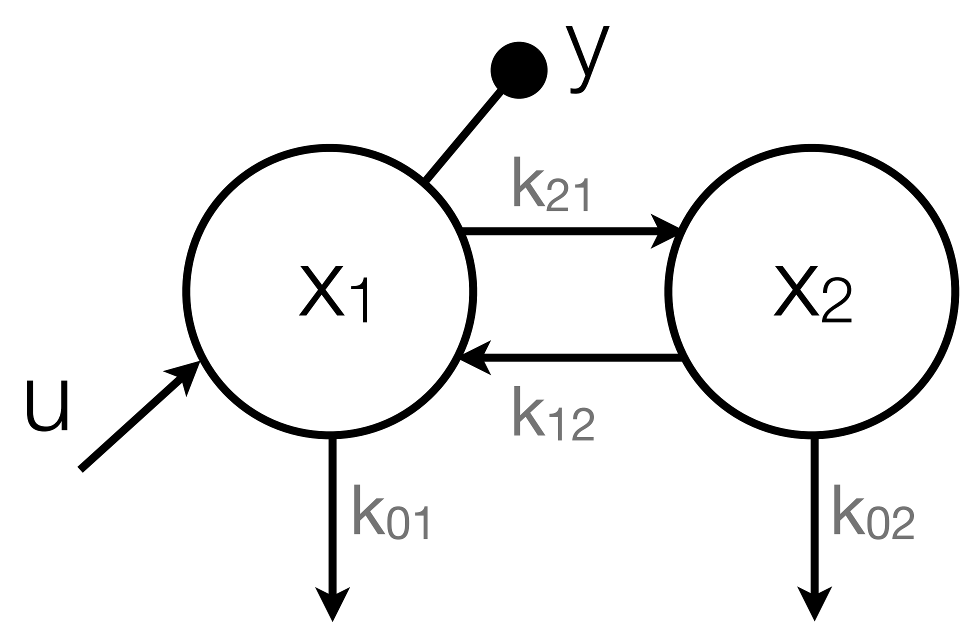
$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$



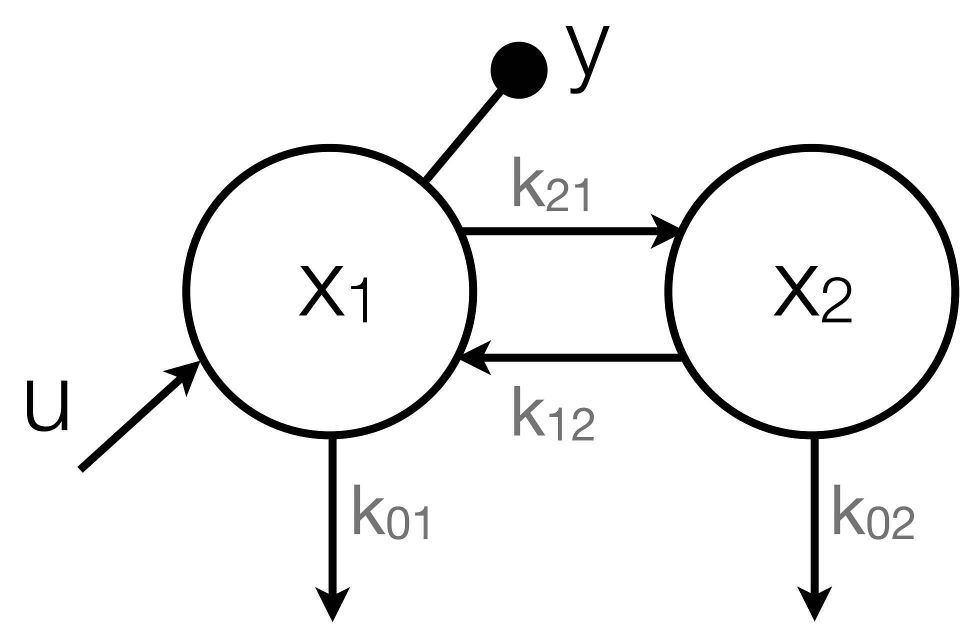
2-Compartment Example

$$\dot{x}_2 = k_{21} x_1 - \left(k_{02} + k_{12}\right) x_2$$

$$\dot{y}V = u + k_{12} x_2 - \left(k_{01} + k_{21}\right) yV$$



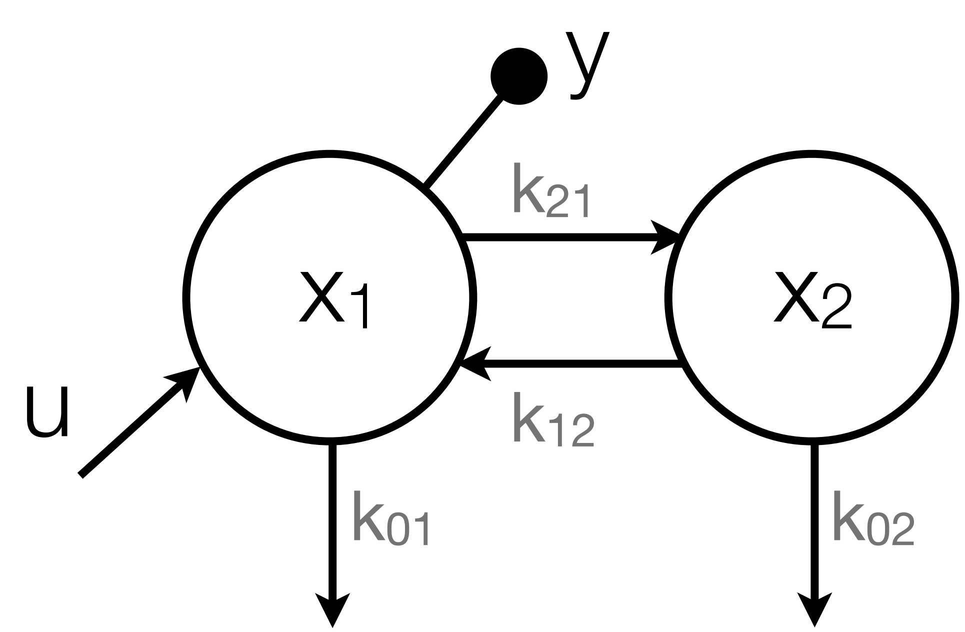
2-Compartment Example



$$\ddot{y} + \left(k_{01} + k_{21} + k_{12} + k_{02}\right)\dot{y} -$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)y - u\left(k_{12} + k_{02}\right)/V - \dot{u}/V = 0$$

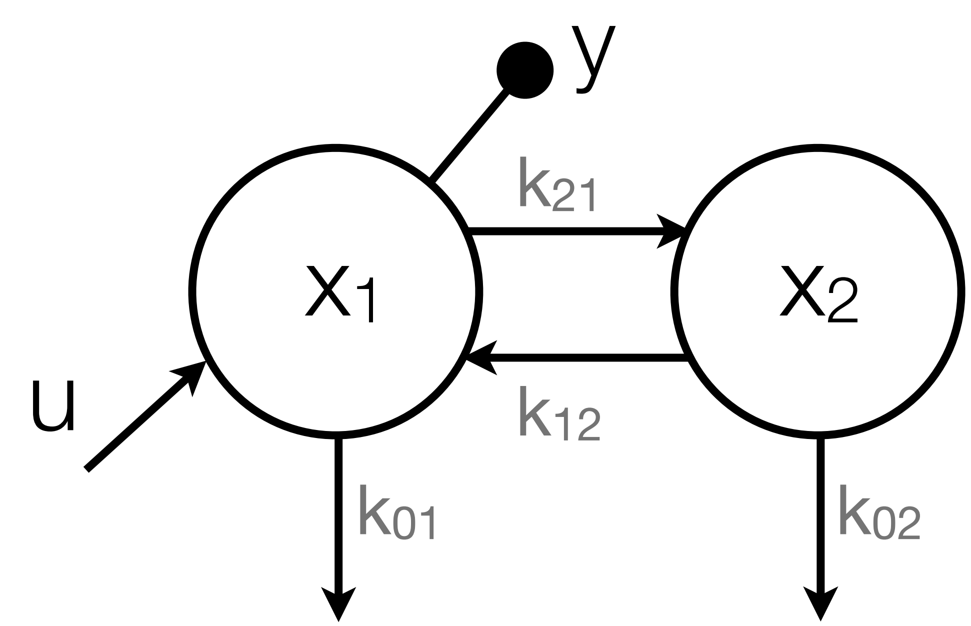
2-Compartment Example



$$\ddot{y} + \left(k_{01} + k_{21} + k_{12} + k_{02}\right)\dot{y} -$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)y - u\left(k_{12} + k_{02}\right)/V - \dot{u}/V = 0$$

2-Compartment Example

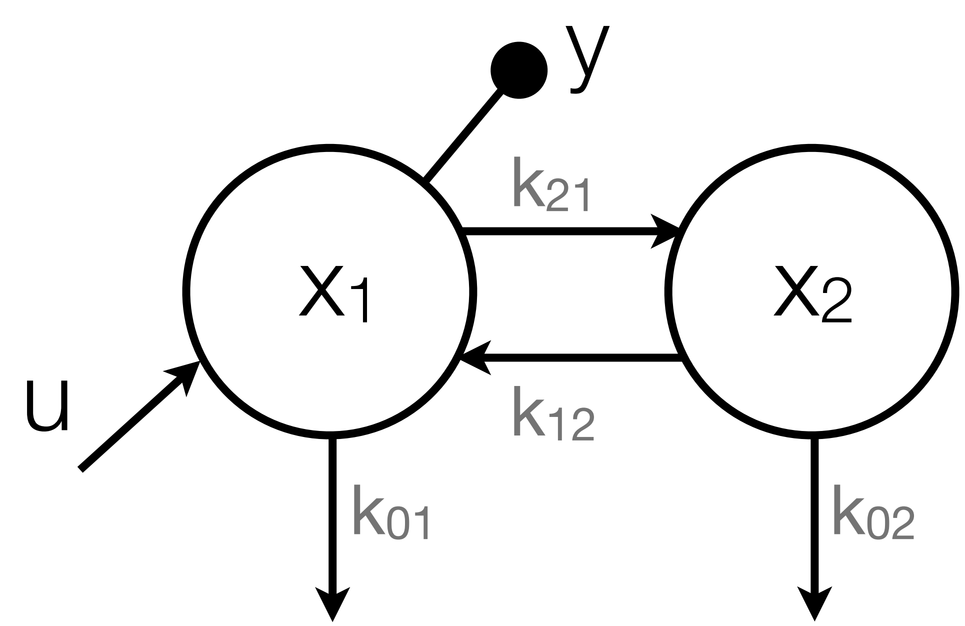


$$\ddot{y} + \left(k_{01} + k_{21} + k_{12} + k_{02}\right)\dot{y} -$$
$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)y - u\left(k_{12} + k_{02}\right)/V - \dot{u}/V = 0$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) \qquad \left(k_{12} + k_{02}\right)/V$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) \qquad 1/V$$

2-Compartment Example



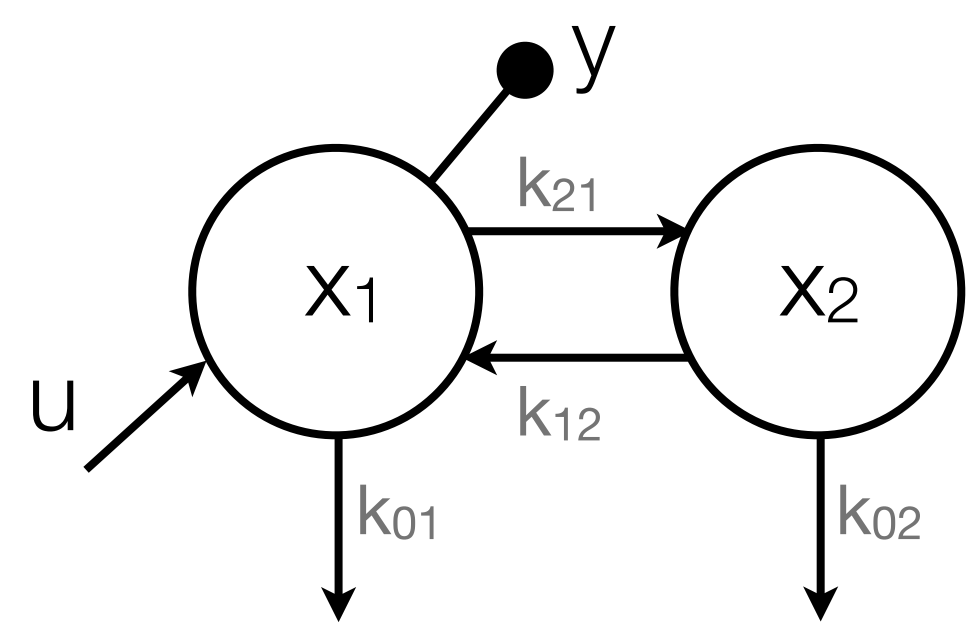
$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right)$$

$$\left(k_{12} + k_{02}\right) / V$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)$$

$$1 / V$$

2-Compartment Example



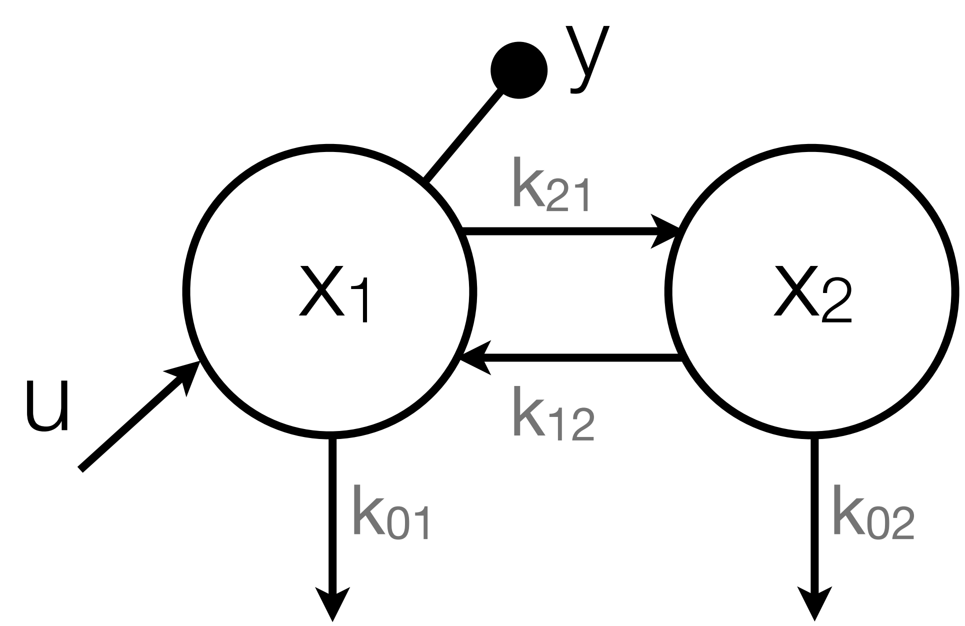
$$1 / V$$

$$\left(k_{12} + k_{02}\right) / V$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right)$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)$$

2-Compartment Example



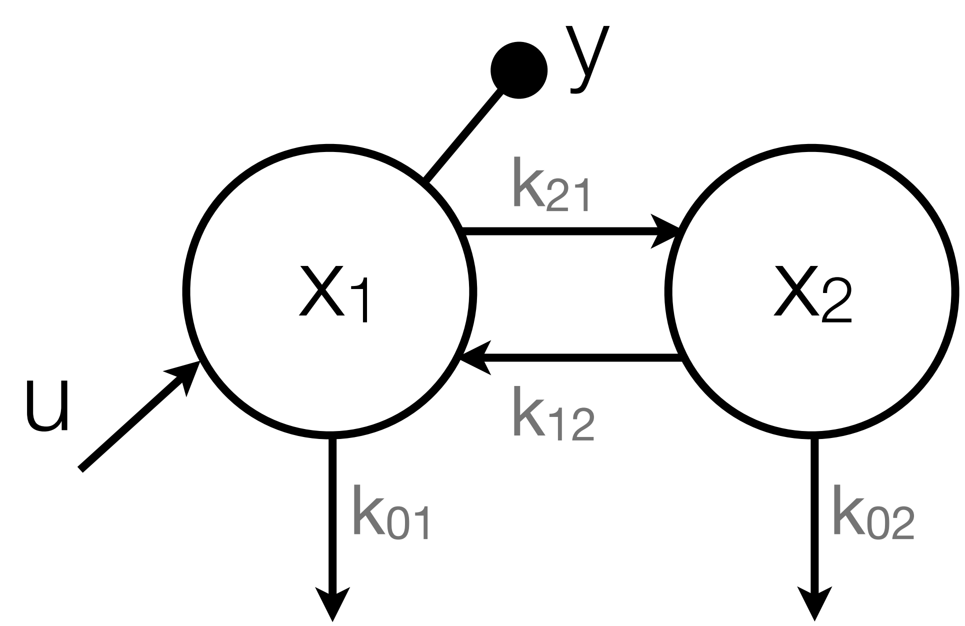
$$1 / V = a_1$$

$$\left(k_{12} + k_{02}\right) / V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

2-Compartment Example



$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

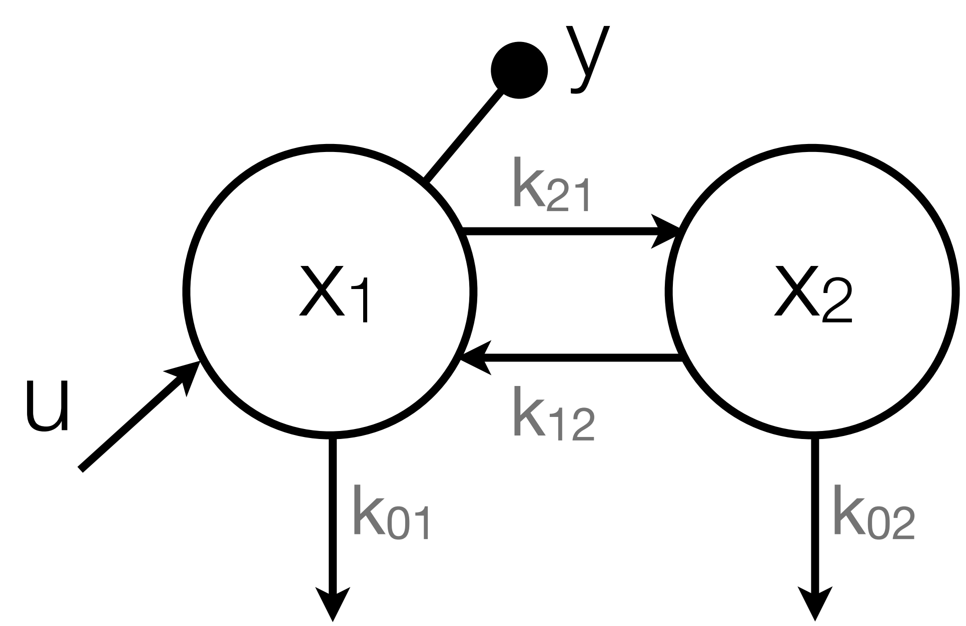
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

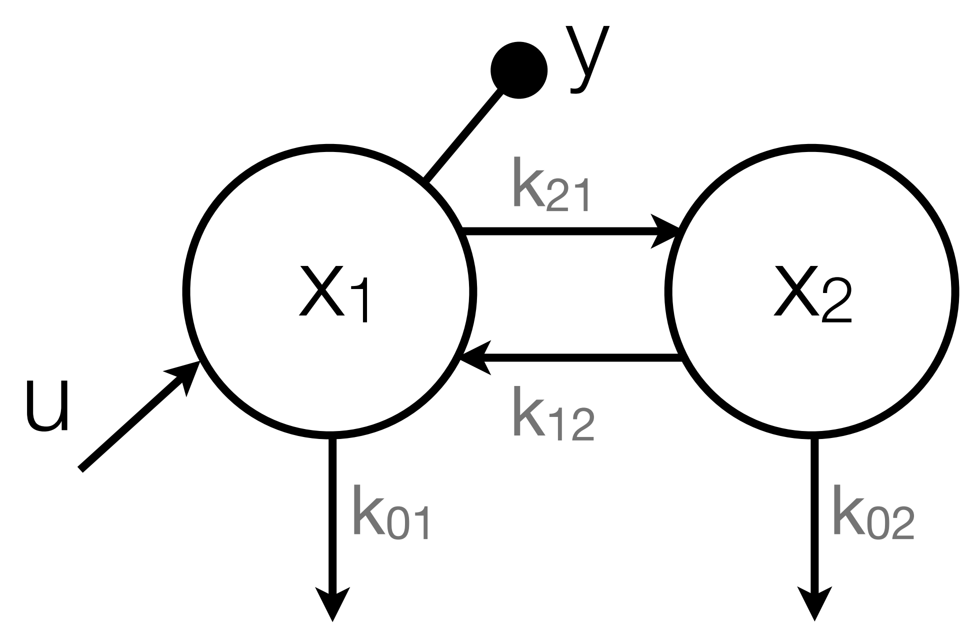
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

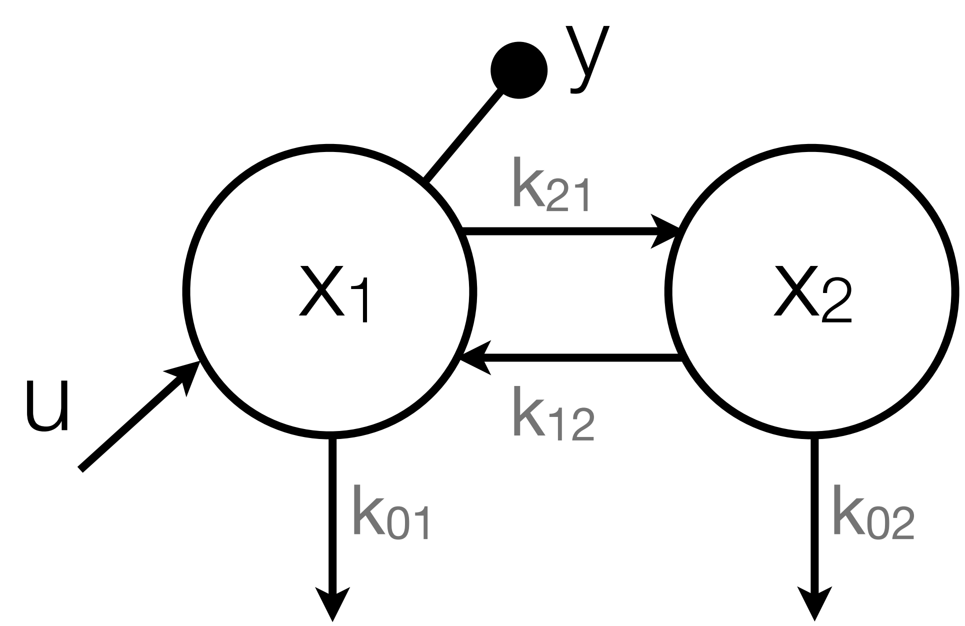
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

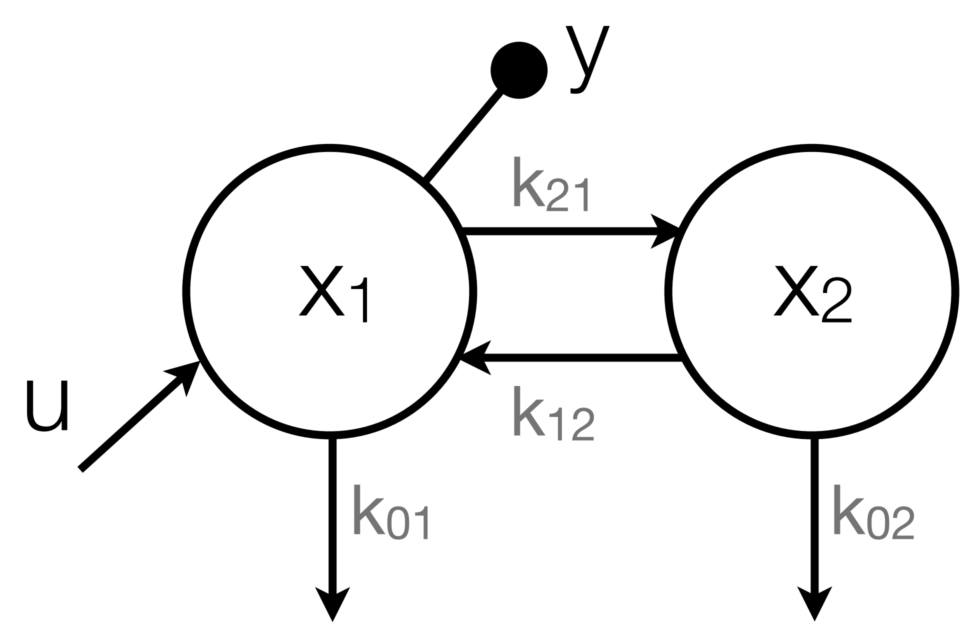
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



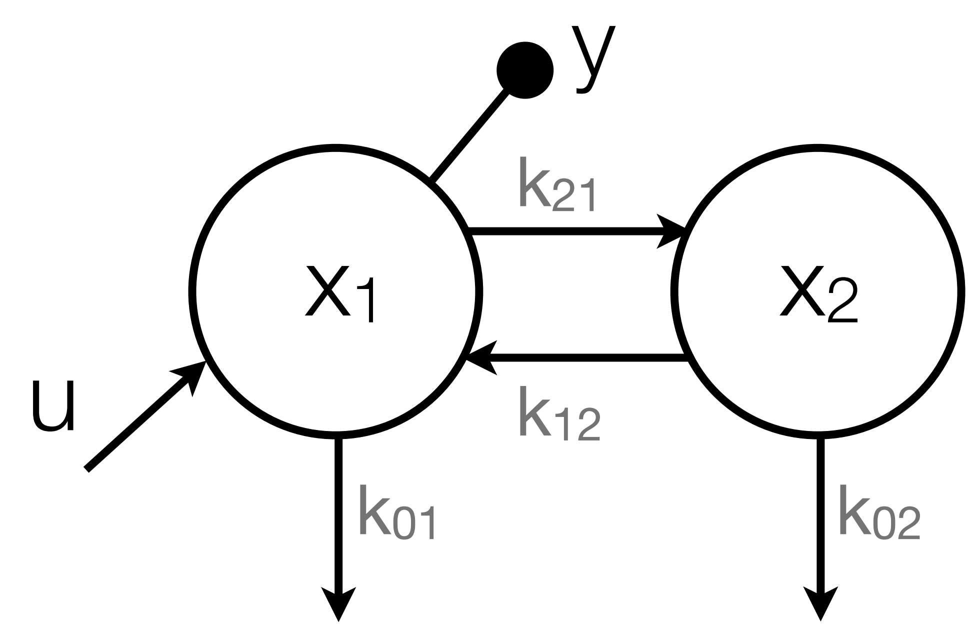
Unidentifiable

2-Compartment Example

$$\dot{x}_1 = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)x_1$$

$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$

$$y = x_1 / V$$

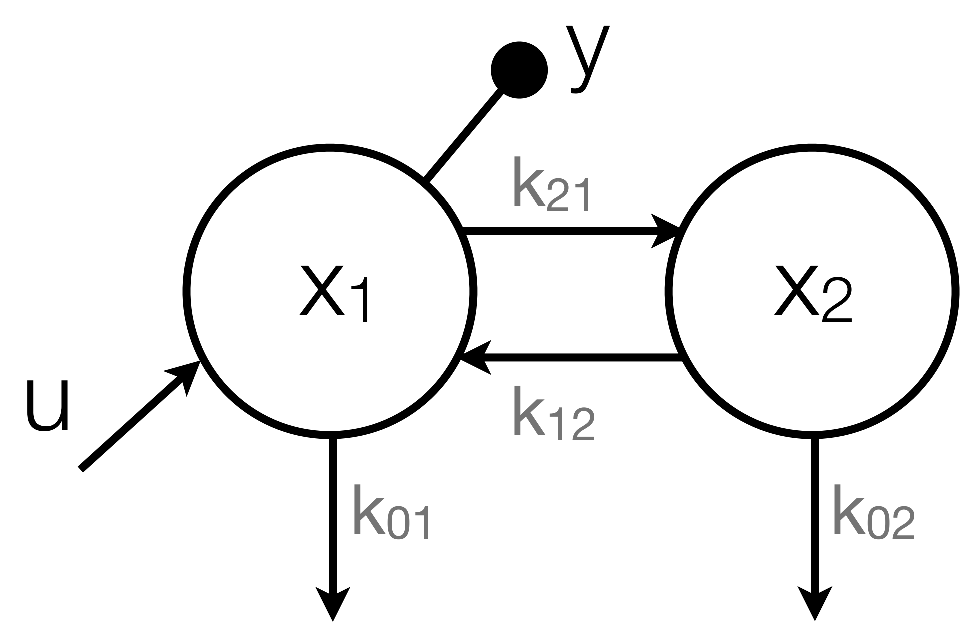


2-Compartment Example

$$\dot{x}_1 = u + \underline{k_{12}}x_2 - \left(\underline{k_{01} + k_{21}}\right)x_1$$

$$\dot{x}_2 = \underline{k_{21}}x_1 - \left(\underline{k_{02} + k_{12}}\right)x_2$$

$$y = x_1 / \underline{V}$$



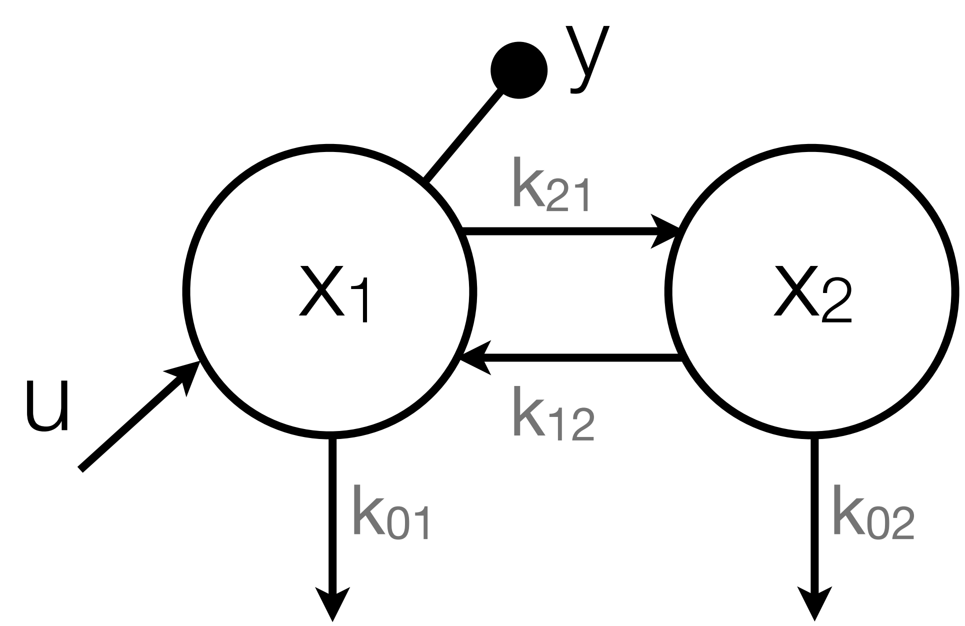
Let $\underline{x}_2 = k_{12}x_2$

2-Compartment Example

$$\dot{x}_1 = u + \underline{k_{12}}x_2 - \left(\underline{k_{01} + k_{21}}\right)x_1$$

$$\dot{x}_2 = \underline{k_{21}}x_1 - \left(\underline{k_{02} + k_{12}}\right)x_2$$

$$y = x_1 / \underline{V}$$



Let $\underline{x}_2 = k_{12}x_2$

$$\dot{x}_1 = u + \underline{x}_2 - \left(\underline{k_{01} + k_{21}}\right)x_1$$

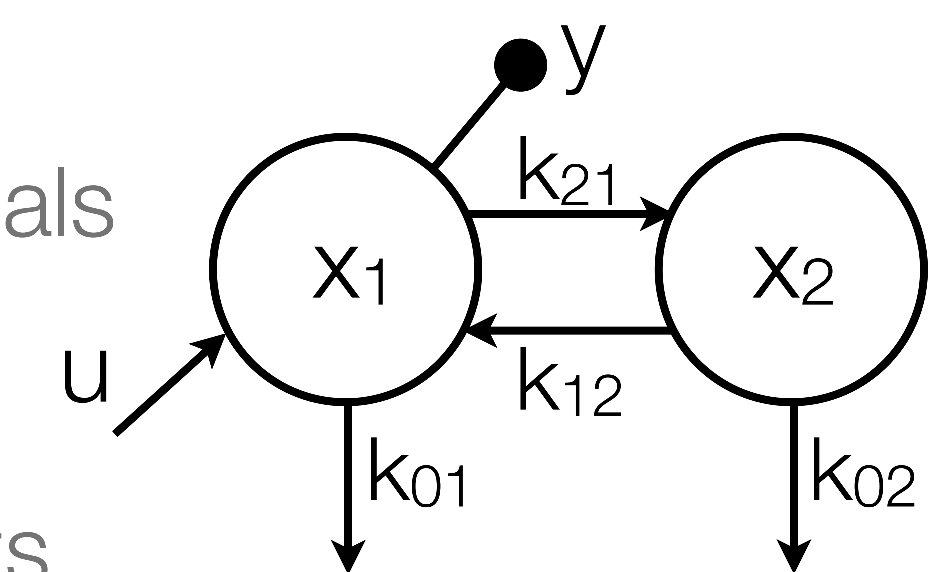
$$\dot{\underline{x}}_2 = \underline{k_{12}k_{21}}x_1 - \left(\underline{k_{02} + k_{12}}\right)\underline{x}_2$$

$$y = x_1 / \underline{V}$$

Or add information about one of the parameters

Differential Algebra Approach

- View model & measurement equations as differential polynomials



- Reduce the equations using Gröbner bases, characteristic sets, etc. to eliminate unmeasured variables (x)

- Yields **input-output equation(s)** only in terms of known variables (y, u)

- Use coefficients to test model identifiability

Ollivier 1990, Ljung & Glad 1994, Audoly et al 2003, etc.

Differential Algebra Approach

- From the coefficients, can often determine:
  - Simpler forms for identifiable combinations
  - Identifiable reparameterizations for model
- Not always easy by eye—use Gröbner bases & other methods to simplify
- Note about scaling as a useful first step (cf. nondimensionalization)

Differential Algebra Approach

- Convenient as a way to prove identifiability results for relatively broad classes of models
  - Linear compartmental models & graph structure (with Nikki Meshkat & Seth Sullivant)
  - SIR-type models (with Tony Nance)
  - Hodgkin-Huxley-type models (with Olivia Walch)

Conclusions

- Many related questions and potential issues when connecting models to data: observability, distinguishability & model selection, reparameterization & model/parameter reduction, and more

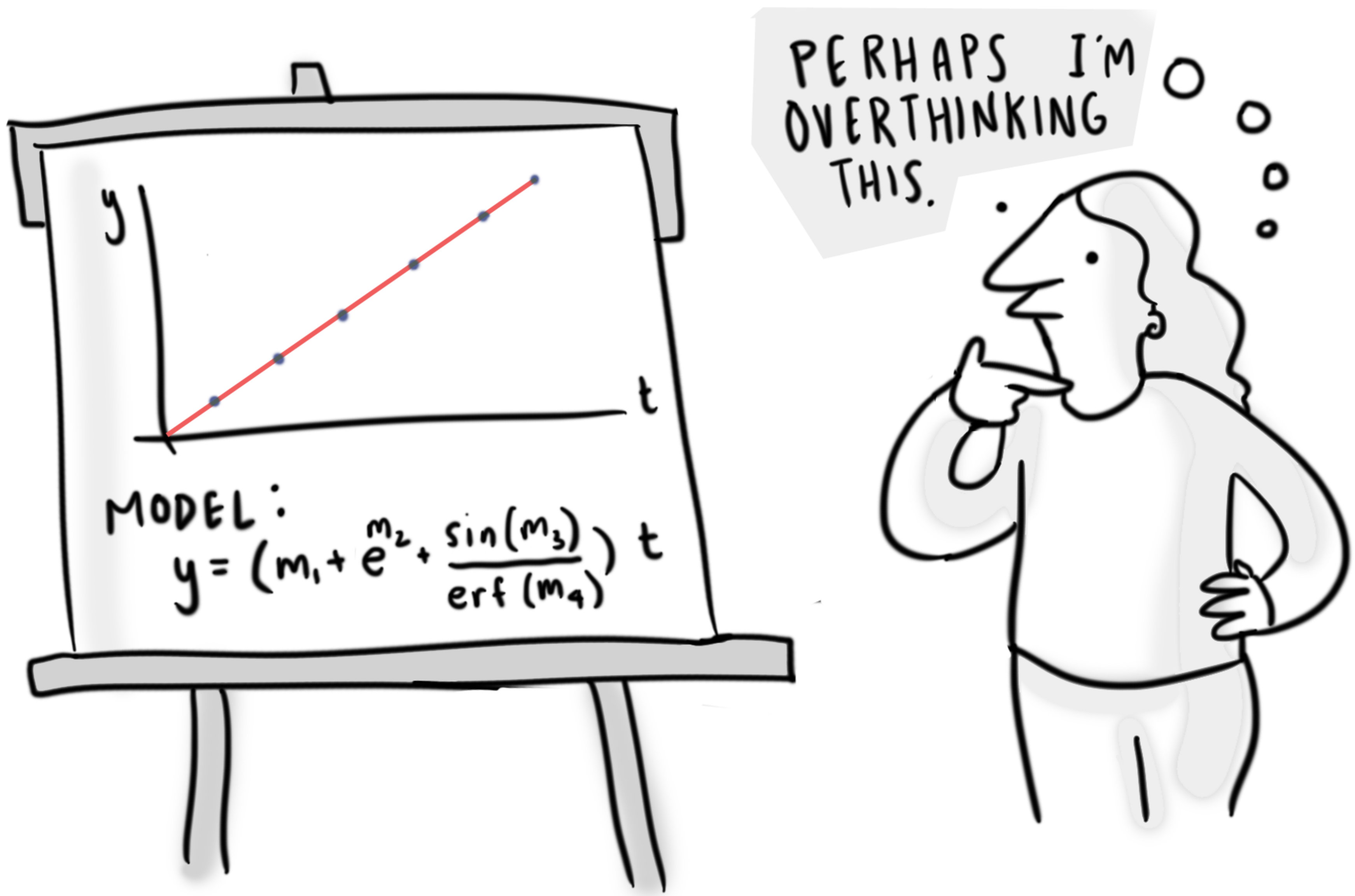
- Many other methods! (eigenvalues of FIM, sloppy models, active subspaces, Bayesian methods, & more)

- Depending on amount of data, model complexity, model type, and more, different approaches may work in different circumstances

Conclusions

- Identifiability—an important question to address when estimating model parameters
- Common problem in math bio (identifiability-robustness tradeoff)
- Many approaches, both numerical and analytical

Questions?



comic by Olivia Walch (UM):
http://imogenquest.net