



Information geometry for Review multiparameter models: new perspectives on the origin of simplicity

Katherine N Quinn , Michael C Abbott , Mark K Transtrum , Benjamin B Machta and James P Sethna ... and James P Sethnas * 0

PRL **104,** 060201 (2010)

PHYSICAL REVIEW LETTERS

Why are Nonlinear Fits to Data so Challenging?

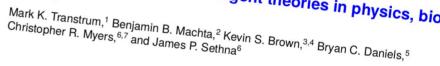
Mark K. Transtrum,* Benjamin B. Machta, † and James P. Sethna ‡

Parameter Space Compression Underlies Emergent Theories and Predictive Models

Benjamin B. Machta, 1,2 Ricky Chachra, 1 Mark K. Transtrum, 1,3 James P. Sethna 1*

THE JOURNAL OF CHEMICAL PHYSICS 143, 010901 (2015)

Perspective: Sloppiness and emergent theories in physics, biology,



PHYSICAL REVIEW E 83, 036701 (2011)

Geometry of nonlinear least squares with applications to sloppy models and optimization

Mark K. Transtrum, Benjamin B. Machta, and James P. Sethna

PRL 113, 098701 (2014)

PHYSICAL REVIEW LETTERS

Model Reduction by Manifold Boundaries

Mark K. Transtrum^{1,*} and Peng Oiu²



week ending

29 AUGUST 2014

Bridging Mechanistic and Phenomenological Models of Complex Biological Systems

Mark K. Transtrum¹*, Peng Qiu²



james sethna



mark transtrum



benjamin machta



katherine quinn

sloppyness → hierarchy of parameter importance

relevant for:

- robustness and evolvability of complex biological systems
- effectiveness of simple models
 (the unreasonable success of spherical cows)
- developing coarse-grained descriptions of complex models
- parameter identifiability

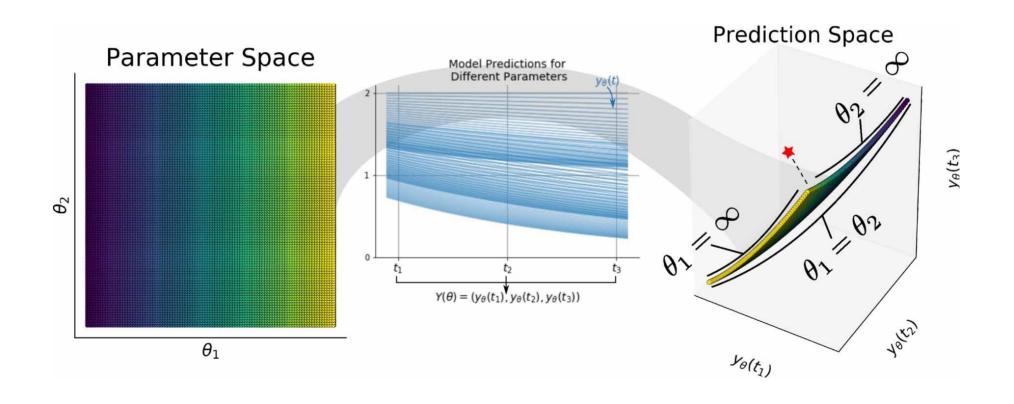
nonlinear least-squares

model:
$$x = y_{\theta}(t)$$
 $\begin{cases} \text{data: } \{\mathbf{x}, \mathbf{t}\} \equiv \{(x_1, t_1), \dots, (x_M, t_M)\} \\ \text{parameters: } \boldsymbol{\theta} \equiv (\theta_1, \theta_2, \dots, \theta_D) \end{cases}$

$$p(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{M} \left[x_i - y_{\boldsymbol{\theta}}(t_i)\right]^2\right\}$$

parameter manifold

$$y_{\theta}(t) = e^{-\theta_1 t} + e^{-\theta_2 t}, \qquad \theta_{\mu} \geqslant 0$$



deviations in parameter space

$$D_{\mathrm{KL}}(\boldsymbol{\theta}', \boldsymbol{\theta}) \equiv \sum_{x} \left[p(\mathbf{x}|\boldsymbol{\theta}) - p(\mathbf{x}|\boldsymbol{\theta}') \right] \log \left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{p(\mathbf{x}|\boldsymbol{\theta}')} \right)$$

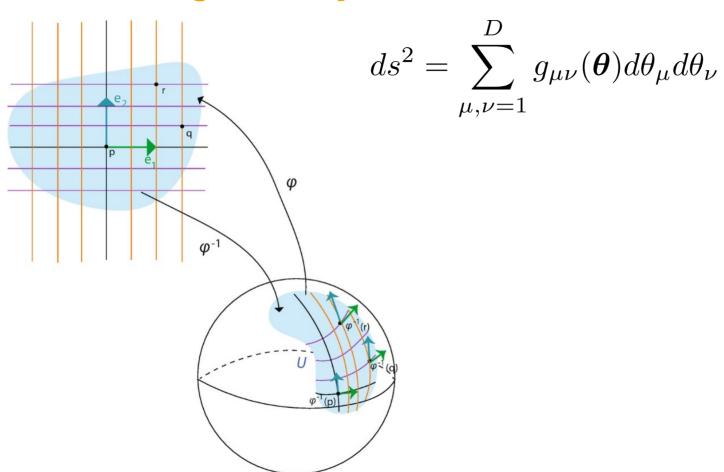
$$D_{\mathrm{KL}}(\boldsymbol{\theta}, \boldsymbol{\theta}) = 0$$
 $\frac{\partial D_{\mathrm{KL}}}{\partial \theta_{\mu}}(\boldsymbol{\theta}, \boldsymbol{\theta}) = 0$

$$D_{\mathrm{KL}}(\boldsymbol{\theta} + d\boldsymbol{\theta}, \boldsymbol{\theta}) = \sum_{\mu,\nu=1}^{D} g_{\mu\nu}(\boldsymbol{\theta}) d\theta_{\mu} d\theta_{\nu} + \cdots$$

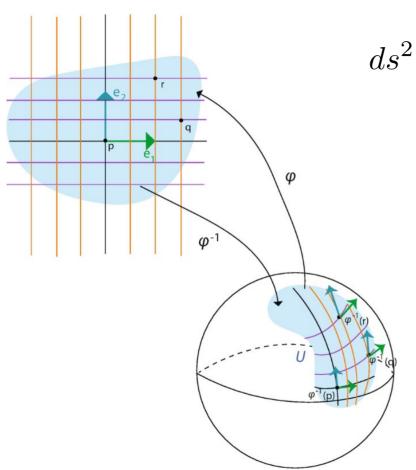
$$g_{\mu\nu}(\boldsymbol{\theta}) = \sum_{x} p(\mathbf{x}|\boldsymbol{\theta}) \frac{\partial \log p(\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_{\mu}} \frac{\partial \log p(\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_{\nu}}$$

fisher's information matrix

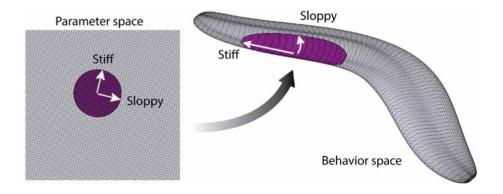
riemannian geometry



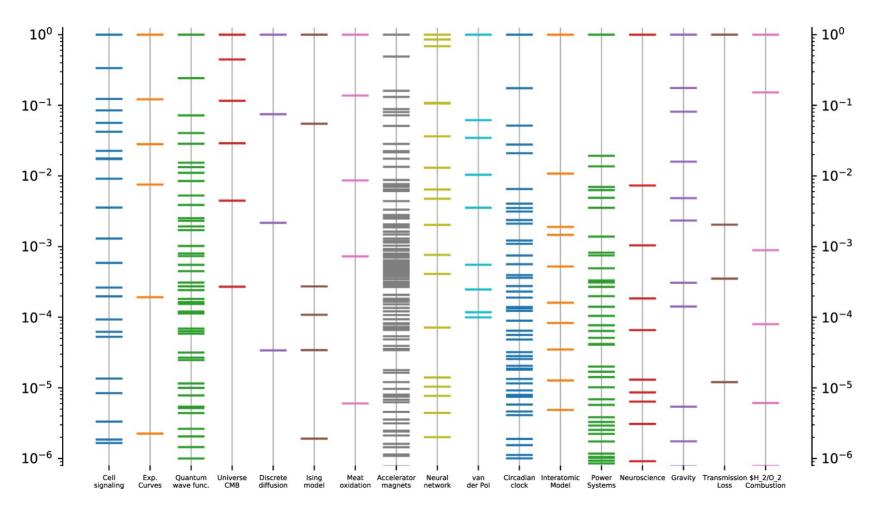
riemannian geometry



$$ds^2 = \sum_{\mu,\nu=1}^{D} g_{\mu\nu}(\boldsymbol{\theta}) d\theta_{\mu} d\theta_{\nu}$$



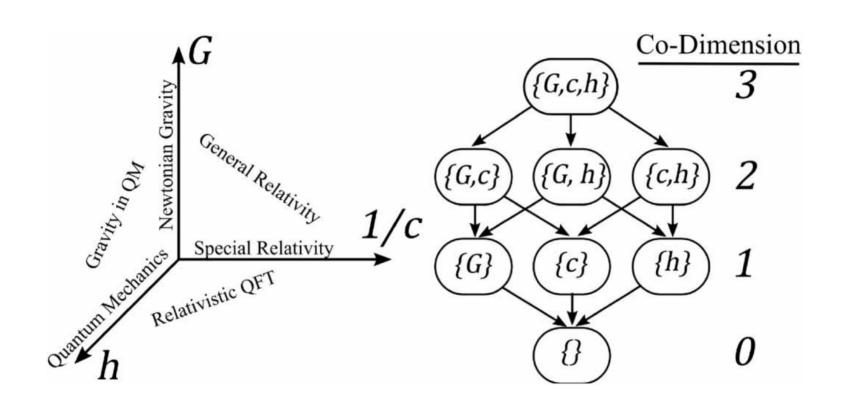
spectrum of the metric tensor



size of the manifold: hyperribbon

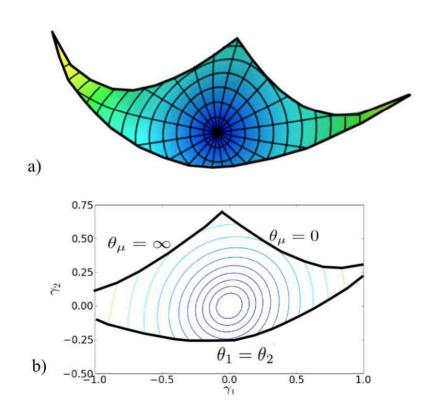
$$L = \int \sqrt{ds^2} = \int_0^1 d\tau \, \sqrt{\sum_{\mu,\nu} g_{\mu\nu}(\theta)} \frac{\partial \theta_\mu}{\partial \tau} \frac{\partial \theta_\nu}{\partial \tau} \qquad \qquad \theta(\tau) \, \, \text{geodesic}$$

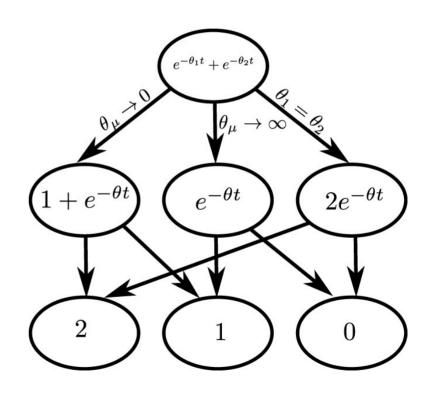
effective theories as manifold boundaries



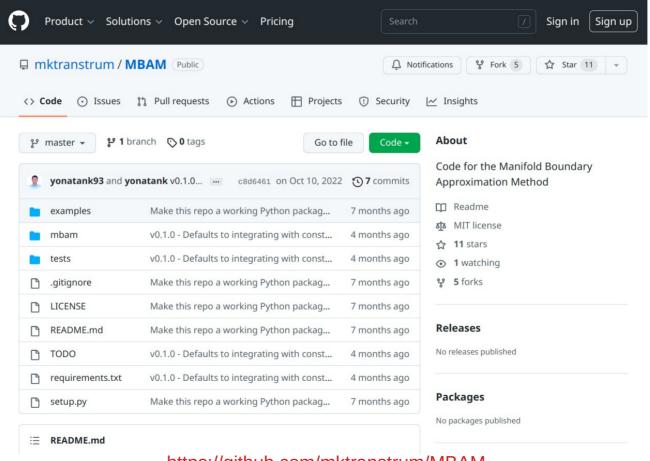
effective theories as manifold boundaries

$$y_{\theta}(t) = e^{-\theta_1 t} + e^{-\theta_2 t}, \qquad \theta_{\mu} \ge 0$$

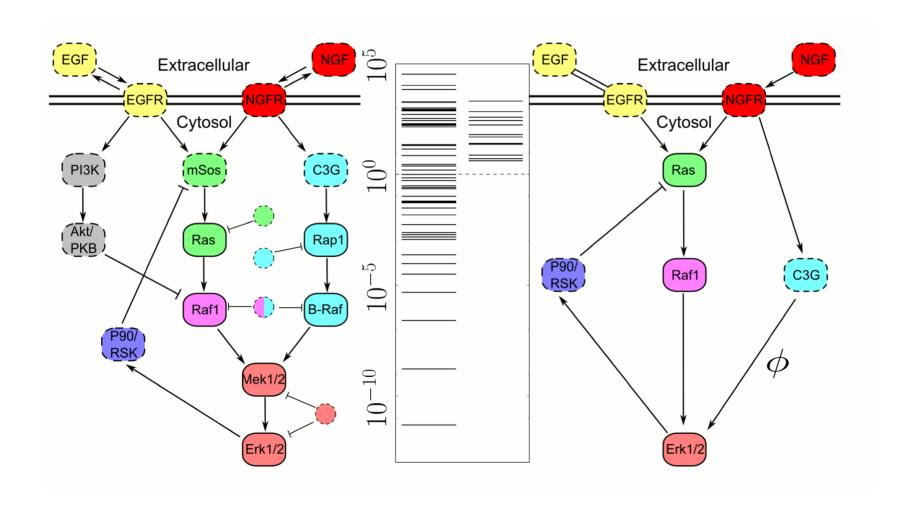




- 1. identify the sloppiest direction on the manifold
- 2. follow the shortest geodesic path to a boundary
- 3. use the limiting combination of parameters to simplify the model
- 4. fit data with the simplified model
- 5. iterate



https://github.com/mktranstrum/MBAM



```
\frac{d}{dt}[EGF] = kruEGF \cdot [bEGFR] - krbEGF \cdot [EGF] \cdot [fGFR]
       \frac{d}{dt}[NGF] = kruNGF \cdot [bNGFR] - krbNGF \cdot [NGF] \cdot [fNGFR]
   \frac{d}{dt}[\text{fEGFR}] = \text{kruEGF} \cdot [\text{bEGFR}] - \text{krbEGF} \cdot [\text{EGF}] \cdot [\text{fEGFR}]
   \frac{d}{dt} [bEGFR] = krbEGF · [EGF] · [fEGFR] - kruEGF · [bEGFR]
   \frac{d}{dt}[fNGFR] = kruNGF \cdot [bNGFR] - krbNGF \cdot [NGF] \cdot [fNGFR]
   \frac{d}{dt} \left[ \text{bNGFR} \right] \ = \ \text{krbNGF} \cdot \left[ \text{NGF} \right] \cdot \left[ \text{fNGFR} \right] - \text{kruNGF} \cdot \left[ \text{bNGFR} \right]
       \frac{d}{dt} [SosA] = \frac{\text{kEGF} \cdot [\text{bEGFR}] \cdot [SosI]}{([SosI] + \text{KmEGF})}
                                 kdSos \cdot [P90RskA] \cdot [SosA]
                                        ([SosA] + KmdSos)
\frac{d}{dt} [P90RskA] = \frac{\text{kpP90Rsk} \cdot [\text{ErkA}] \cdot [P90RskI]}{([P90RskI] + \text{KmpP90Rsk})}
      \frac{d}{dt} [RasA] = \frac{kSos \cdot [SosA] \cdot [RasI]}{([RasI] + KmSos)}
                                 kRasGap \cdot [RasGapA] \cdot [RasA]
                                        ([RasA] + KmRasGap)
     \frac{d}{dt} \left[ \text{Raf1A} \right] = \frac{\text{kRasToRaf1} \cdot \left[ \text{RasA} \right] \cdot \left[ \text{Raf1I} \right]}{\left( \left[ \text{Raf1I} \right] + \text{KmRasToRaf1} \right)}
                                  kdRaf1 \cdot [Raf1PPtase] \cdot [Raf1A]
                                          ([Raf1A] + KmdRaf1)
                                  kdRaf1ByAkt \cdot [AktA] \cdot [Raf1A]
                                      ([Raf1A] + KmRaf1ByAkt)
```

```
\frac{d}{dt} [BRafA] = \frac{kRap1ToBRaf \cdot [Rap1A] \cdot [BRafI]}{([BRafI] + KmRap1ToBRaf)}
                     -\frac{\text{kdBRaf} \cdot [\text{Raf1PPtase}] \cdot [\text{BRafA}]}{\text{(BRafA)}}
                                          ([BRafA] + KmdBRaf)
     \frac{d}{dt} [\text{MekA}] = \frac{\text{kpRaf1} \cdot [\text{Raf1A}] \cdot [\text{MekI}]}{([\text{MekI}] + \text{KmpRaf1})}
                               +\frac{\mathrm{kpBRaf}\cdot[\mathrm{BRafA}]\cdot[\mathrm{MekI}]}{([\mathrm{MekI}]+\mathrm{KmpBRaf})}
                                kdMek \cdot [PP2AA] \cdot [MekA]
                                       ([MekA] + KmdMek)
      ([ErkI] + KmpMekCytoplasmic)
                                   kdErk \cdot [PP2AA] \cdot [ErkA]
                                       ([ErkA] + KmdErk)
    \frac{d}{dt} [PI3KA] = \frac{kPI3K \cdot [bEGFR] \cdot [PI3KI]}{([PI3KI] + KmPI3K)}
                               +\frac{\text{kPI3KRas} \cdot [\text{RasA}] \cdot [\text{PI3KI}]}{([\text{PI3KI}] + \text{KmPI3KRas})}
      \frac{d}{dt} [AktA] = \frac{kAkt \cdot [PI3KA] \cdot [AktI]}{([AktI] + KmAkt)}
     \frac{d}{dt} [C3GA] = \frac{kC3GNGF \cdot [bNGFR] \cdot [C3GI]}{([C3GI] + KmC3GNGF)}
    \frac{d}{dt} [\text{Rap1A}] = \frac{\text{kC3G} \cdot [\text{C3GA}] \cdot [\text{Rap1I}]}{([\text{Rap1I}] + \text{KmC3G})}
                                   kRapGap \cdot [RapGapA] \cdot [Rap1A]
                                          ([Rap1A] + KmRapGap)
```

15 independent diff. eqs.48 parameters

uncertainties up to 500 %!

$$\frac{d}{dt}[\text{NGF}] = -\theta_1[\text{NGF}][\text{fNGFR}]$$

$$\frac{d}{dt}[\text{bNGFR}] = \theta_1[\text{NGF}][\text{fNGFR}]$$

$$[\text{bEGFR}] = \begin{cases} 1 & \text{EGF Present} \\ 0 & \text{Otherwise} \end{cases}$$

$$\frac{d}{dt}[\text{RasA}] = \theta_2[\text{bEGFR}] + \theta_3[\text{bNGFR}] - [\widetilde{\text{P90}}][\text{RasA}]$$

$$\frac{d}{dt}[\widetilde{\text{Raf1A}}] = \theta_4[\text{RasA}] - \theta_5 \frac{[\widetilde{\text{Raf1A}}]}{[\widetilde{\text{Raf1A}}] + \theta_6}$$

$$\frac{d}{dt}[\text{C3GA}] = \theta_7[\text{bNGFR}][\text{C3GI}]$$

$$[\text{Rap1A}] = \theta_8[\text{C3GA}]$$

$$[\text{MekA}] = [\widetilde{\text{Raf1A}}][\text{MekI}] + \theta_9[\text{Rap1A}]$$

$$\frac{d}{dt}[\text{ErkA}] = -\theta_{10}[\text{ErkA}] + \theta_{11}[\text{MekA}][\text{ErkI}]$$

$$\frac{d}{dt}[\widetilde{\text{P90}}] = \theta_{12}[\text{ErkA}].$$

7 independent diff. eqs.

12 "effective" parameters

```
\theta_1 = \text{krbNGF}
 \theta_2 = \frac{[SosI](kEGF)(KmRasGap)(kSos)}{[RasGapA](KmEGF)(kRasGap)}
 \theta_3 = \frac{[SosI](kNGF)(KmRasGap)(kSos)}{[RasGapA](KmNGF)(kRasGap)}
 \theta_4 = \frac{(kRasToRaf1)(kpRaf1)(KmdMek)}{[PP2AA](kdMek)(KmpRaf1)}
 \theta_5 = \frac{[\text{Raf1PPtase}](\text{kdRaf1})(\text{kpRaf1})(\text{KmdMek})}{[\text{PP2AA}](\text{kdMek})(\text{KmpRaf1})}
 \theta_6 = \frac{(\text{KmdRaf1})(\text{kpRaf1})(\text{KmdMek})}{[\text{PP2AA}](\text{kdMek})(\text{KmpRaf1})}
  \theta_7 = kC3GNGF/KmC3GNGF
 \theta_8 = \frac{(\text{KmRapGap})(\text{kC3G})}{[\text{RapGapA}](\text{kRapGaP})}
 \theta_9 = \frac{[BRafI](kRap1ToBRaf)(KmdBRaf)(kpBRaf)(KmdMek)}{[PP2AA][Raf1PPtase](kdBRaf)(KmRap1ToBRaf)(kdMek)}
\theta_{10} = [PP2AA](kdErk)/KmdErk
\theta_{11} = \text{kpMekCytoplasmic/KmpMekCytoplasmic}
\theta_{12} = \frac{[P90/RSKI](kpP90Rsk)(kdSos)}{(KmpP90Rsk)(KmdSos)}.
```

$$\frac{d}{dt}[NGF] = -\theta_1[NGF][fNGFR]$$

$$\frac{d}{dt}[bNGFR] = \theta_1[NGF][fNGFR]$$

$$[bEGFR] = \begin{cases} \frac{1}{0} \stackrel{EGF}{Otherwise} \\ \frac{d}{dt}[RasA] = \theta_2[bEGFR] + \theta_3[bNGFR] - [\widetilde{P90}][RasA] \end{cases}$$

$$\frac{d}{dt}[\widetilde{Raf1A}] = \theta_4[RasA] - \theta_5 \frac{[\widetilde{Raf1A}]}{[\widetilde{Raf1A}] + \theta_6}$$

$$\frac{d}{dt}[C3GA] = \theta_7[bNGFR][C3GI]$$

$$[Rap1A] = \theta_8[C3GA]$$

$$[MekA] = [\widetilde{Raf1A}][MekI] + \theta_9[Rap1A]$$

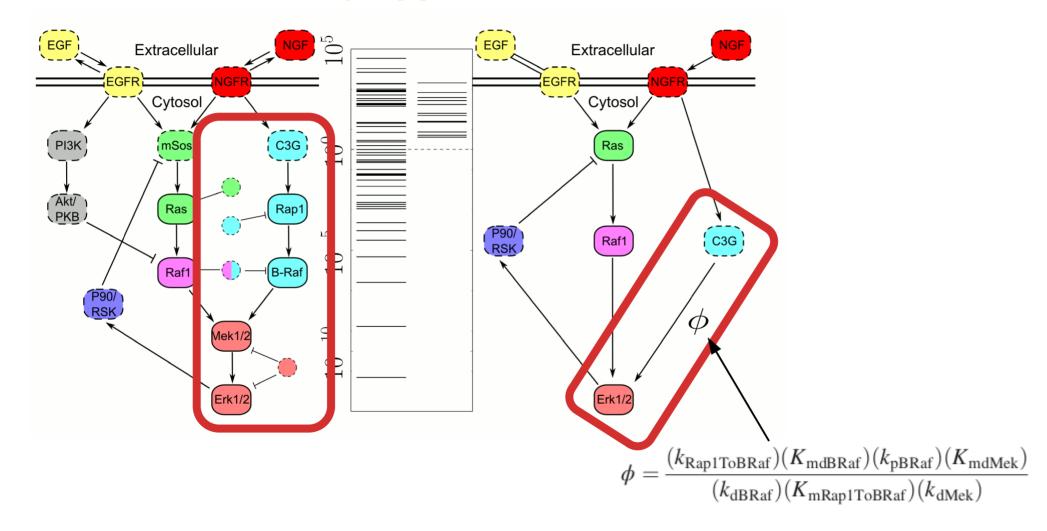
$$\frac{d}{dt}[ErkA] = -\theta_{10}[ErkA] + \theta_{11}[MekA][ErkI]$$

$$\frac{d}{dt}[\widetilde{P90}] = \theta_{12}[ErkA].$$

parameter	Reduced Model Value	Uncertainty
$ heta_1$	2.37×10^{-3}	1.1×10^{-3}
$ heta_2$	9.34×10^{-2}	2.7×10^{-2}
θ_3	7.57×10^{-1}	3.9×10^{-1}
$ heta_4$	9.88×10^{-1}	5.4×10^{-1}
$ heta_5$	3.40×10^{-1}	2.1×10^{-1}
θ_6	2.70×10^{-1}	1.9×10^{-1}
$ heta_7$	1.11×10^{0}	7.8×10^{-1}
θ_8	1.30×10^{-1}	3.6×10^{-2}
$ heta_9$	1.75×10^{0}	5.5×10^{-1}
θ_{10}	2.56×10^{-1}	1.0×10^{-1}
θ_{11}	4.51×10^{0}	1.6×10^{0}
θ_{12}	8.21×10^{-1}	3.5×10^{-1}

7 independent diff. eqs.

12 "effective" parameters

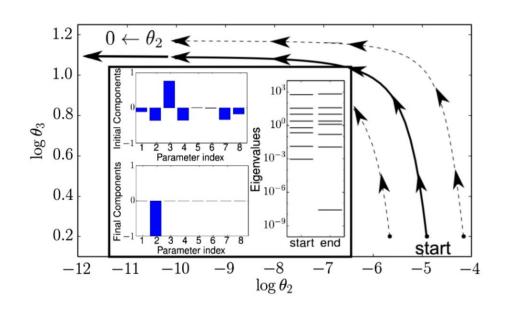


example 1: sum of exponentials

$$y(t, \mathbf{A}, \lambda) = A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} + A_3 e^{-\lambda_3 t} + A_4 e^{-\lambda_4 t}, \qquad \boldsymbol{\theta} = (A_{\mu}, \lambda_{\mu}) \geqslant 0$$

example 1: sum of exponentials

$$y(t, \mathbf{A}, \lambda) = A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} + A_3 e^{-\lambda_3 t} + A_4 e^{-\lambda_4 t}, \qquad \boldsymbol{\theta} = (A_{\mu}, \lambda_{\mu}) \geqslant 0$$

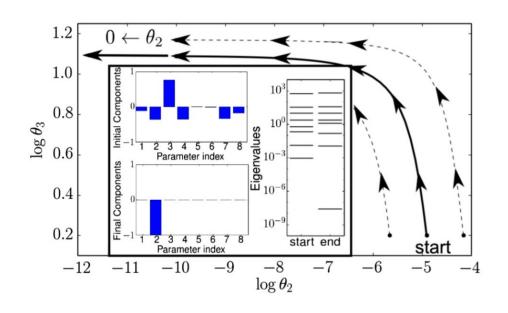


$$\ddot{\theta}^{\mu} + \Gamma^{\mu}_{\alpha,\beta} \dot{\theta}^{\alpha} \dot{\theta}^{\beta} = 0$$

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(\frac{\partial g_{\nu\beta}}{\partial \theta^{\alpha}} + \frac{\partial g_{\nu\alpha}}{\partial \theta^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial \theta^{\nu}} \right)$$

example 1: sum of exponentials

$$y(t, \mathbf{A}, \lambda) = A_1 e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} + A_3 e^{-\lambda_3 t} + A_4 e^{-\lambda_4 t}, \qquad \boldsymbol{\theta} = (A_{\mu}, \lambda_{\mu}) \geqslant 0$$



$$\ddot{\theta}^{\mu} + \Gamma^{\mu}_{\alpha,\beta} \dot{\theta}^{\alpha} \dot{\theta}^{\beta} = 0$$

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(\frac{\partial g_{\nu\beta}}{\partial \theta^{\alpha}} + \frac{\partial g_{\nu\alpha}}{\partial \theta^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial \theta^{\nu}} \right)$$

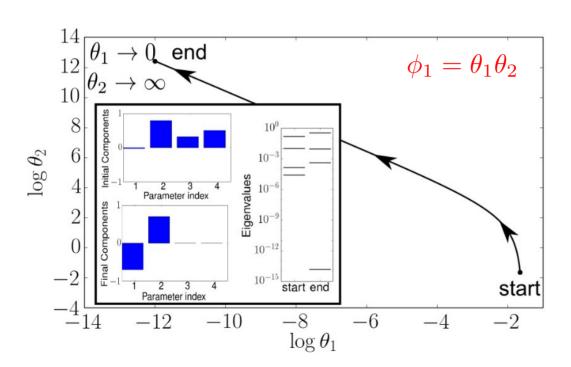
$$\tilde{y}(t, \tilde{\mathbf{A}}, \tilde{\boldsymbol{\lambda}}) = \tilde{A}_2 + \tilde{A}_1 e^{-\tilde{\lambda}_2 t} + \tilde{A}_3 e^{-\tilde{\lambda}_3 t} + \tilde{A}_4 e^{-\tilde{\lambda}_4 t}$$

example 2: kinetics of enzymatic reaction

$$y(u, \boldsymbol{\theta}) = \frac{\theta_1(u^2 + \theta_2 u)}{u^2 + \theta_3 u + \theta_4}$$

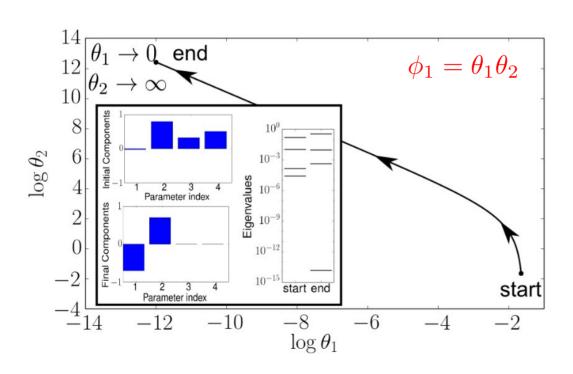
example 2: kinetics of enzymatic reaction

$$y(u, \boldsymbol{\theta}) = \frac{\theta_1(u^2 + \theta_2 u)}{u^2 + \theta_3 u + \theta_4}$$



example 2: kinetics of enzymatic reaction

$$y(u, \boldsymbol{\theta}) = \frac{\theta_1(u^2 + \theta_2 u)}{u^2 + \theta_3 u + \theta_4}$$



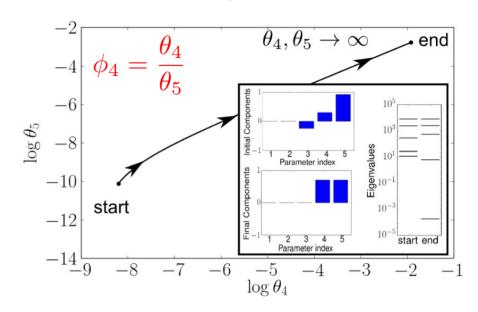
$$y(u,\phi) = \frac{\phi_1 u}{u^2 + \phi_2 u + \phi_3}$$

example 3: thermal isomerisation of α -pinene

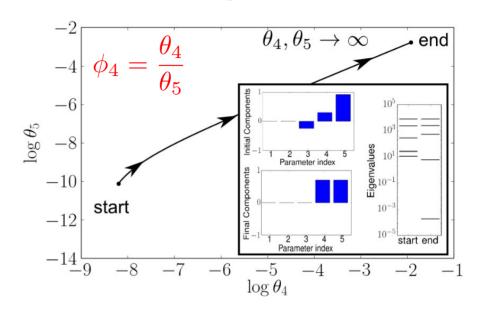
$$\dot{y}_1 = -(\theta_1 + \theta_2)y_1
\dot{y}_2 = \theta_1 y_1
\dot{y}_3 = \theta_2 y_1 - (\theta_3 + \theta_4)y_3 + \theta_5 y_5
\dot{y}_4 = \theta_3 y_3
\dot{y}_5 = \theta_4 y_3 - \theta_5 y_5$$

example 3: thermal isomerisation of α -pinene

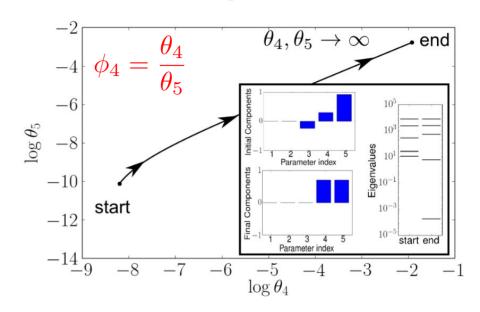
$$\dot{y}_1 = -(\theta_1 + \theta_2)y_1
\dot{y}_2 = \theta_1 y_1
\dot{y}_3 = \theta_2 y_1 - (\theta_3 + \theta_4)y_3 + \theta_5 y_5
\dot{y}_4 = \theta_3 y_3
\dot{y}_5 = \theta_4 y_3 - \theta_5 y_5$$



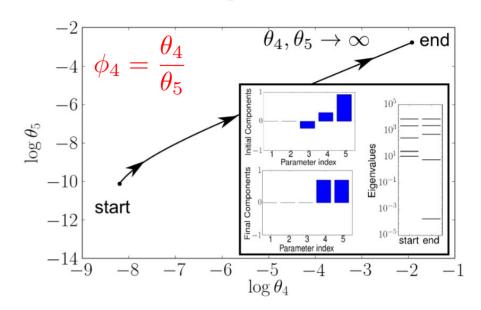
$$\dot{y}_1 = -(\theta_1 + \theta_2)y_1
\dot{y}_2 = \theta_1 y_1
\dot{y}_3 = \theta_2 y_1 - (\theta_3 + \theta_4)y_3 + \theta_5 y_5
\dot{y}_4 = \theta_3 y_3
\dot{y}_5 = \theta_4 y_3 - \theta_5 y_5$$



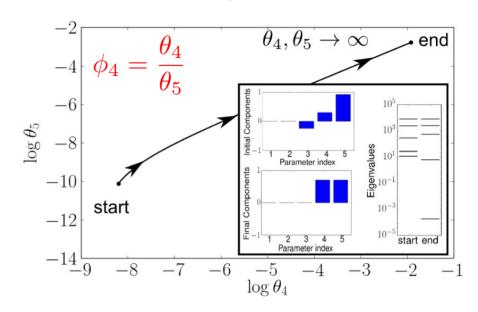
$$\dot{y}_1 = -(\theta_1 + \theta_2)y_1
\dot{y}_2 = \theta_1 y_1
\dot{y}_3 = \theta_2 y_1 - (\theta_3 + \theta_4)y_3 + \theta_5 y_5
\dot{y}_4 = \theta_3 y_3
0 = \phi_4 y_3 - y_5$$



$$\dot{y}_{1} = -(\theta_{1} + \theta_{2})y_{1}
\dot{y}_{2} = \theta_{1}y_{1}
\dot{y}_{3} = \theta_{2}y_{1} - (\theta_{3} + \theta_{4})y_{3} + \theta_{5}y_{5}
\dot{y}_{4} = \theta_{3}y_{3}
\dot{y}_{5} = \theta_{4}y_{3} - \theta_{5}y_{5}
y_{5} = \phi_{4}y_{3}$$



$$\dot{y}_1 = -(\phi_1 + \phi_2)y_1
\dot{y}_2 = \phi_1 y_1
\dot{y}_4 = \phi_3 y_3
\dot{y}_3 + \dot{y}_5 = \phi_2 y_1 - \phi_3 y_3
y_5 = \phi_4 y_3$$

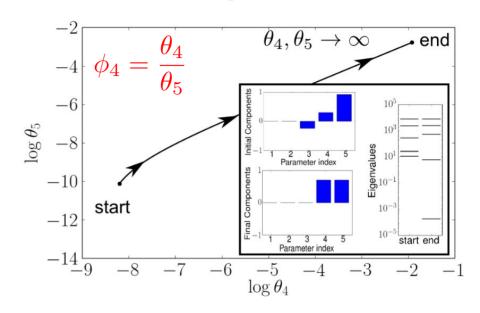


$$\dot{y}_1 = -(\phi_1 + \phi_2)y_1$$

$$\dot{y}_2 = \phi_1 y_1$$

$$\dot{y}_4 = \phi_3 y_3$$

$$\dot{y}_3 = \frac{\phi_2}{1 + \phi_4} y_1 - \frac{\phi_3}{1 + \phi_4} y_3$$



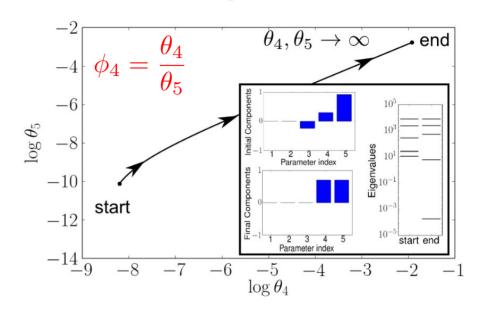
$$\dot{y}_1 = -(\phi_1 + \phi_2)y_1$$

$$\dot{y}_2 = \phi_1 y_1$$

$$\dot{y}_3 = \frac{\phi_2}{1 + \phi_4} y_1 - \frac{\phi_3}{1 + \phi_4} y_3$$

$$\dot{y}_4 = \phi_3 y_3$$

$$y_5 = \phi_4 y_3$$



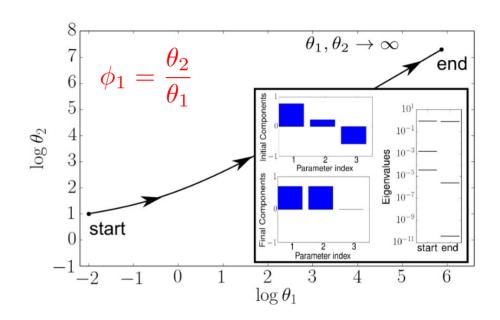
$$[\dot{E}] = -\theta_1[E][S] + (\theta_2 + \theta_3)[ES]$$
$$[\dot{S}] = -\theta_1[E][S] + \theta_2[ES]$$
$$[\dot{E}S] = \theta_1[E][S] - (\theta_2 + \theta_3)[ES]$$
$$[\dot{P}] = \theta_3[ES]$$

$$[\dot{S}] = -\theta_1[E][S] + \theta_2[ES]$$

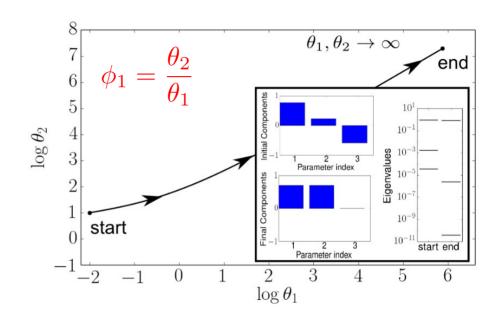
 $[\dot{P}] = \theta_3[ES]$
 $E_0 = [E] + [ES]$
 $S_0 = [S] + [ES] + [P]$

$$[\dot{S}] = -\theta_1[E][S] + \theta_2[ES]$$

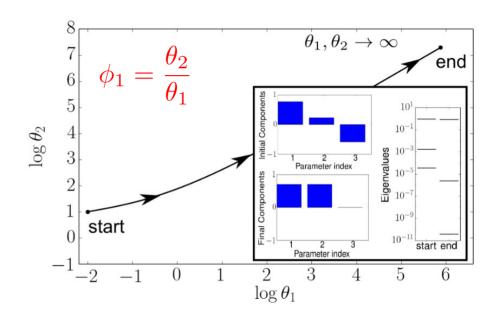
 $[\dot{P}] = \theta_3[ES]$
 $E_0 = [E] + [ES]$
 $S_0 = [S] + [ES] + [P]$



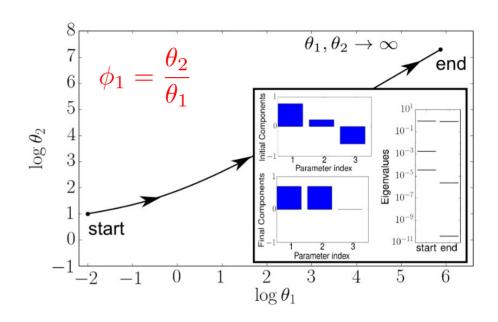
$$[\dot{S}] = -\theta_1[E][S] + \theta_2[ES]$$
$$[\dot{P}] = \theta_3[ES]$$
$$E_0 = [E] + [ES]$$
$$S_0 = [S] + [ES] + [P]$$



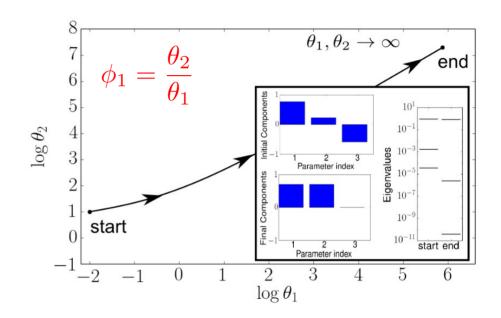
$$0 = -[E][S] + \phi_1[ES]$$
$$[\dot{P}] = \theta_3[ES]$$
$$E_0 = [E] + [ES]$$
$$S_0 = [S] + [ES] + [P]$$



$$0 = -[E][S] + \phi_1[ES]$$
$$[\dot{P}] = \theta_3[ES]$$
$$E_0 = [E] + [ES]$$
$$S_0 = [S] + [ES] + [P]$$



$$[P] = \theta_3[ES]$$
 $[E] = \frac{\phi_1 E_0}{\phi_1 + [S]}$
 $[ES] = \frac{E_0[S]}{\phi_1 + [S]}$
 $S_0 = [S] + [ES] + [P]$

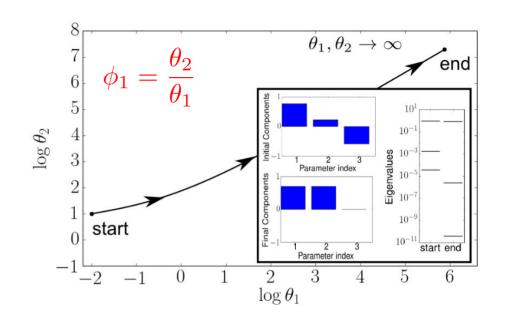


$$[\dot{P}] = \theta_3 E_0 \frac{[S]}{\phi_1 + [S]}$$

$$[E] = \frac{\phi_1 E_0}{\phi_1 + [S]}$$

$$ES] = \frac{E_0[S]}{\phi_1 + [S]}$$

$$S_0 = [S] + [ES] + [P]$$



conclusions

- complex models with many parameters neither increase our understanding nor provide the best description of a system
- spherical cows work **not** because nature is simple, **but** because most complexity is irrelevant
- mbam provides a systematic method to achieve simpler descriptions of complex models