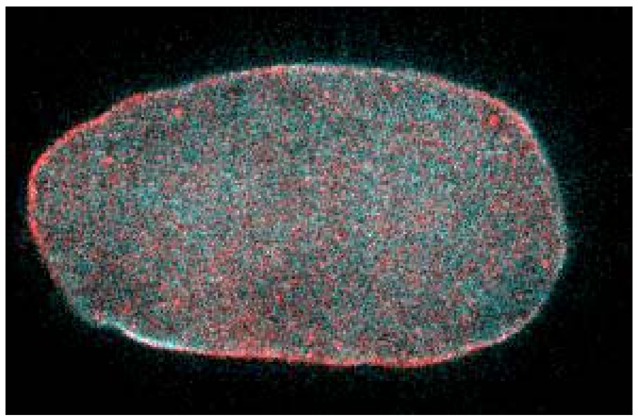
Enabling data-driven discovery in biology: Statistical learning of *interpretable* mathematical models from microscopy videos

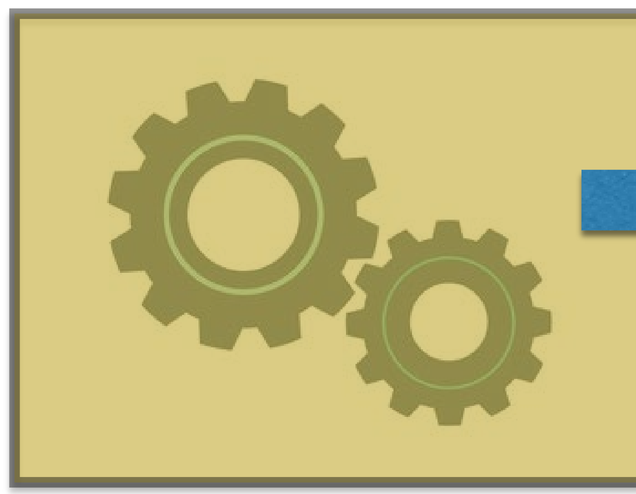
Suryanarayana Maddu, Bevan Cheeseman, Ivo F. Sbalzarini, Christian Mueller



Grill lab (MPI-CBG)

DATA

Drescher lab, MPI for terrestrial microbiology



Inference box

ODE/PDE

$$\frac{dy}{dt} = ty^2 + c$$

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + \frac{\partial^4 u}{\partial x^4} + u\frac{\partial u}{\partial x}$$

center for systems biology dresden

TECHNISCHE UNIVERSITÄT DRESDEN

FLATIRON INSTITUTE Center for Computational Mathematics

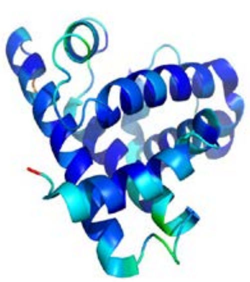
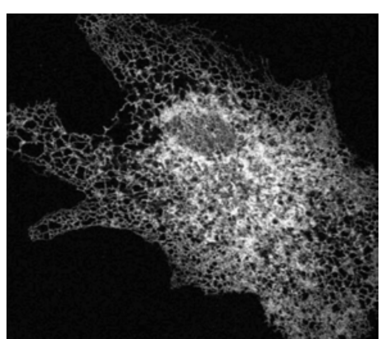
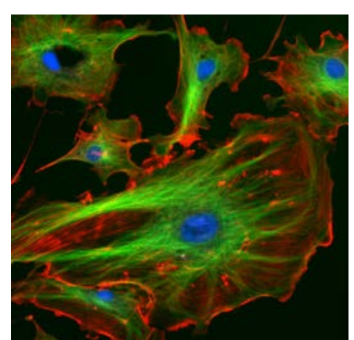
CBG Max Planck Institute of Molecular Cell Biology and Genetics

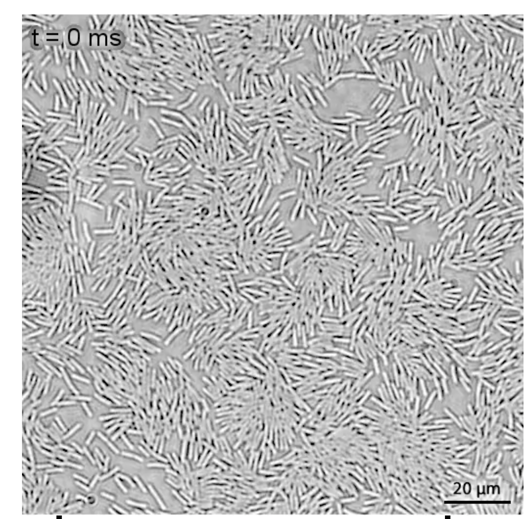
Life in a nutshell

- Highly organized
- Regulated
- Complex shapes
- Non-equilibrium
- Non-linear
- Coupled

Images: Wikipedia

Life at different scales

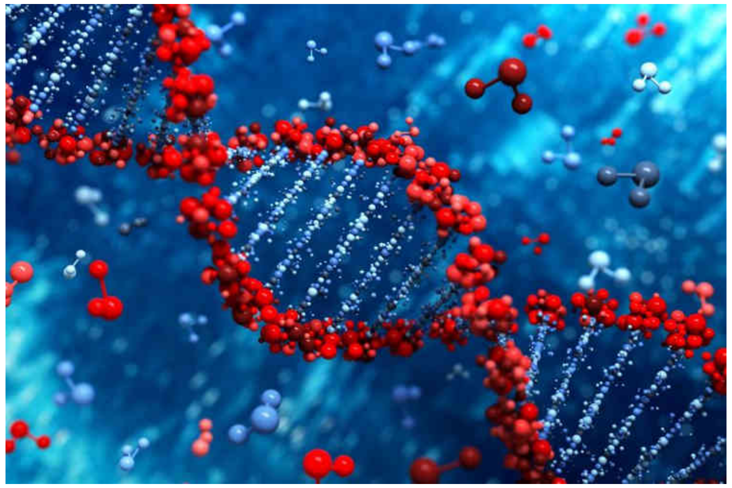




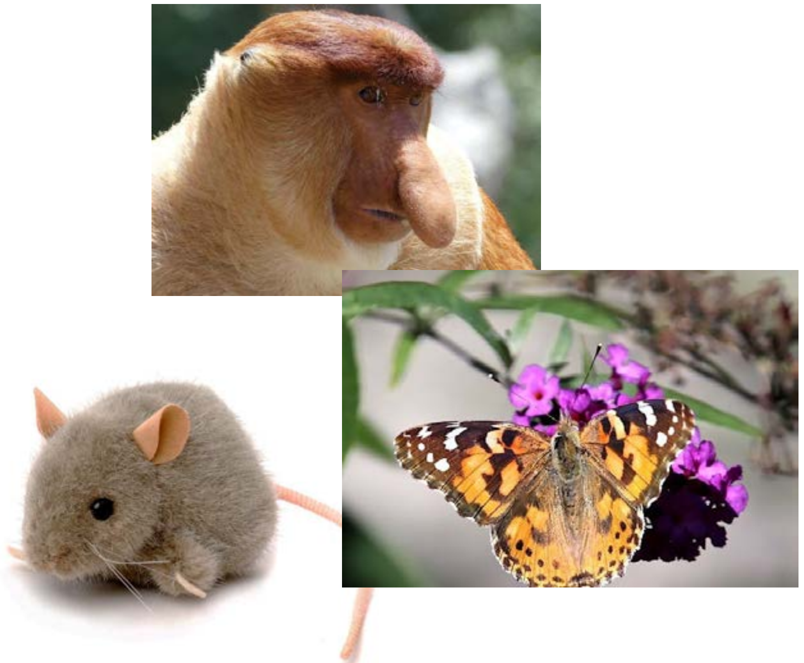


Molecules — Sub-cellular — Cells — tissue/organs — Organisms — Ecosystems

The big question(s) in biology

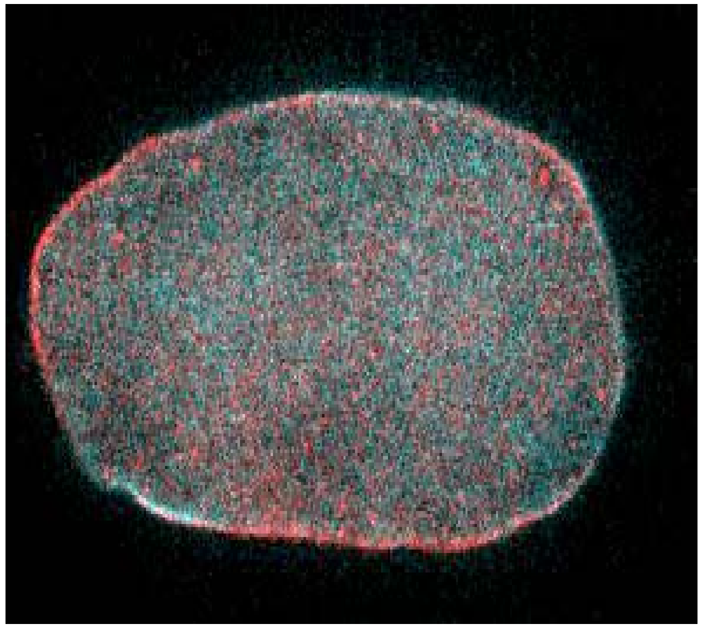
Genotype - Phenotype ?



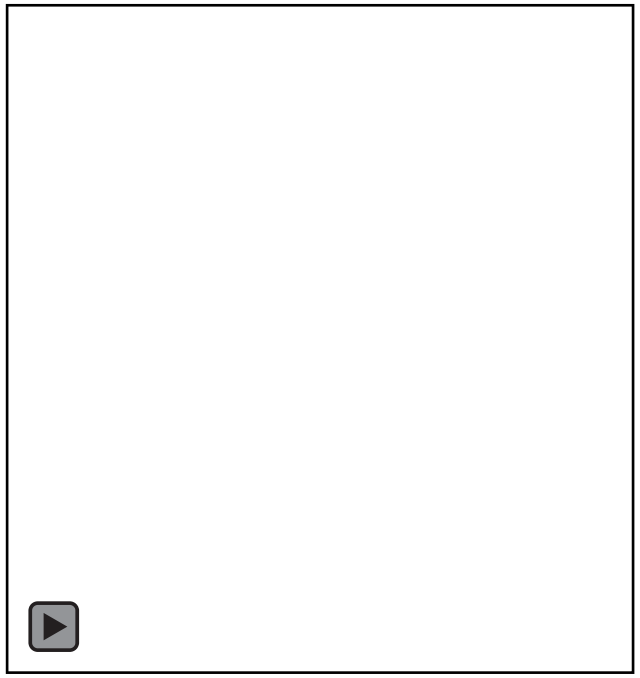


Images: Wikipedia

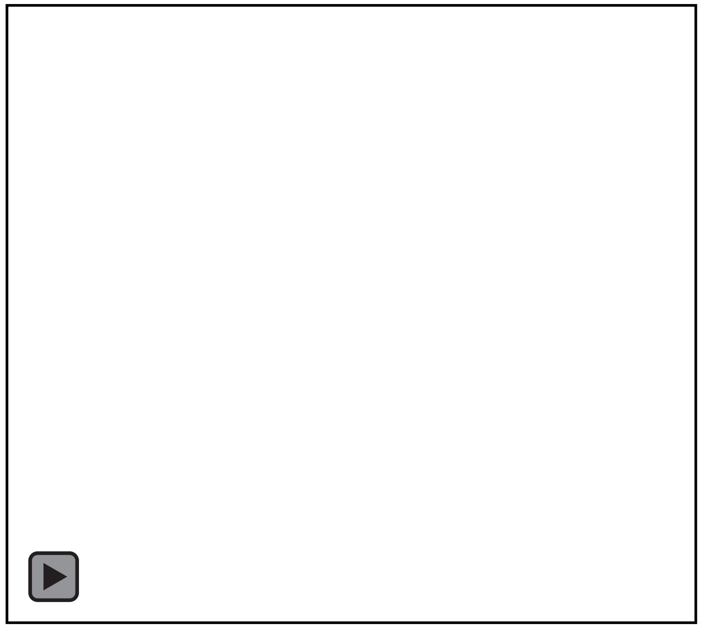
Individual interactions —> emergent dynamics ?



Cellular scale
(Grill lab,MPI-CBG)



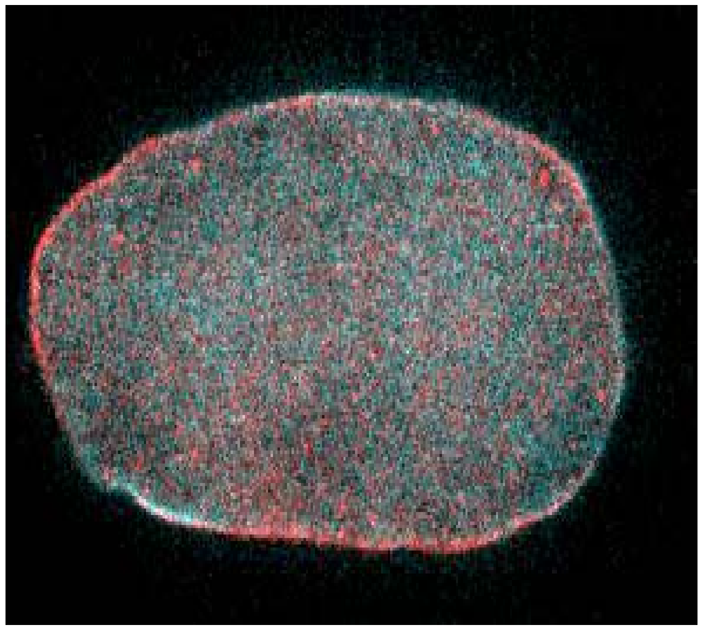
Tissue scale
(Tomancak lab, MPI-CBG)



Eco-systems scale
Drescher lab (MPI-terMic)

Mathematical models for emergent dynamics

Individual interactions —> emergent dynamics ?



Cellular scale
(Grill lab,MPI-CBG)

Tissue scale
(Tomancak lab, MPI-CBG)

Eco-systems scale
Drescher lab (MPIterMic)

- Spatial organisation
- Temporal dynamics
- Environmental influences
- Physics of interactions
- Regulatory mechanism

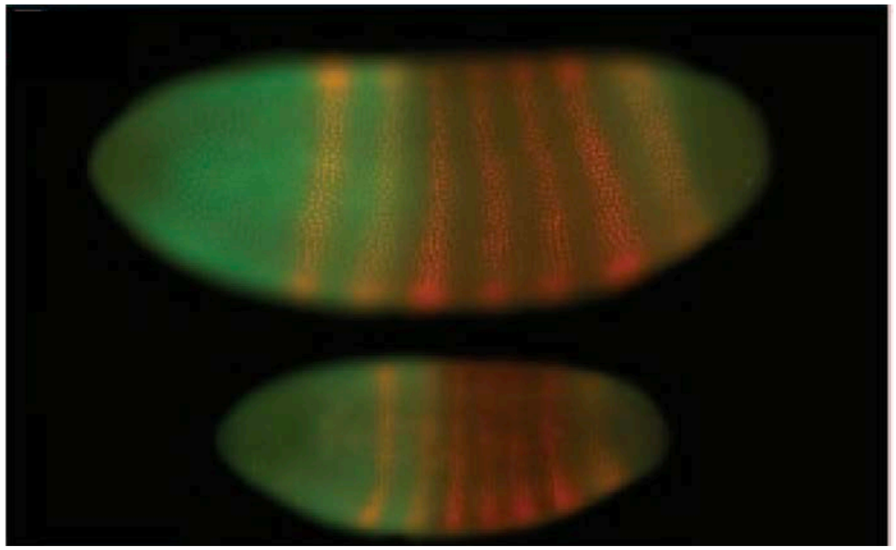
PDE

→ u, $\mathbf{x}$, t

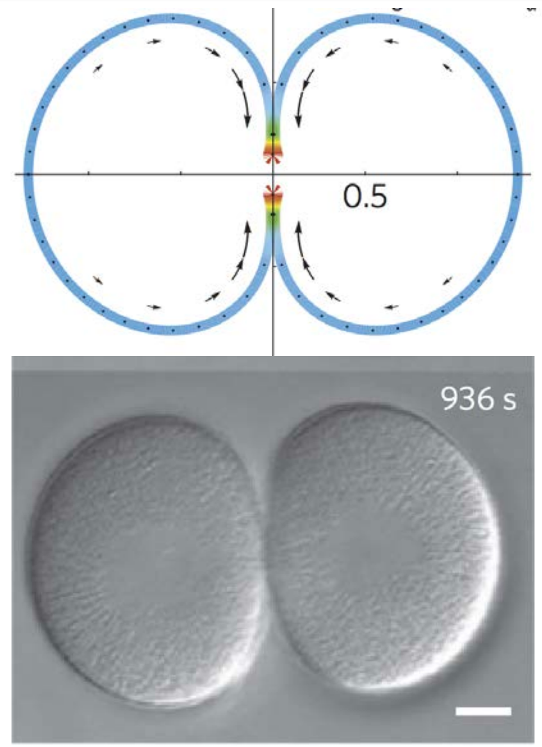
→ Initial/boundary Conditions

→ PDE components

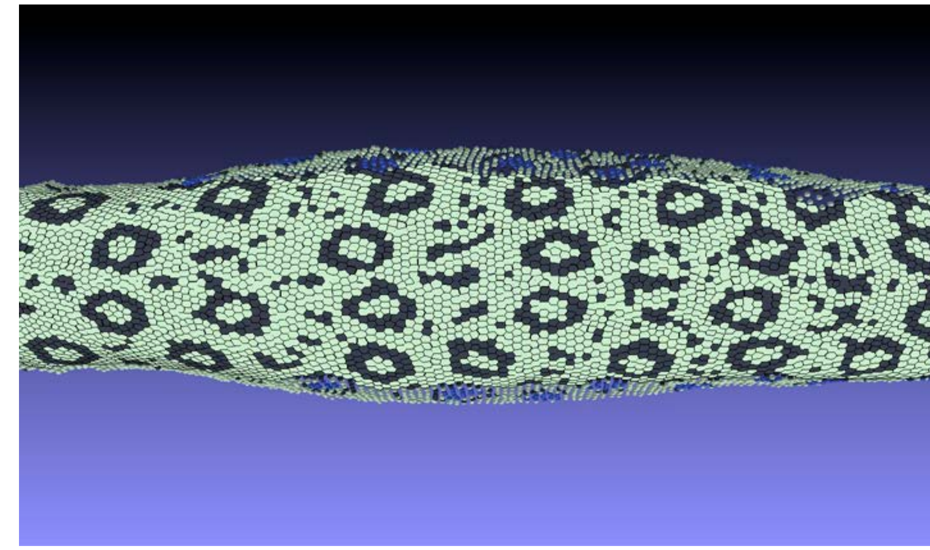
Modeling emergent dynamics : PDE success



Gregor et al, PNAS 2005

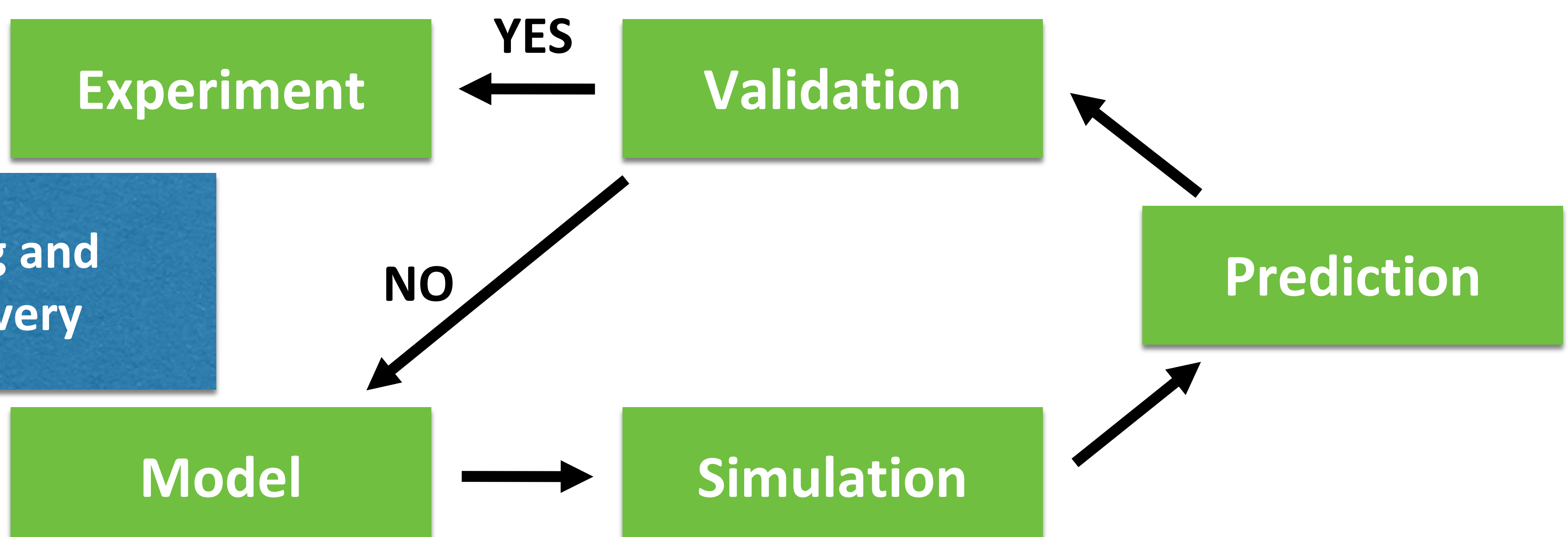


J Prost et al, Nature 2015



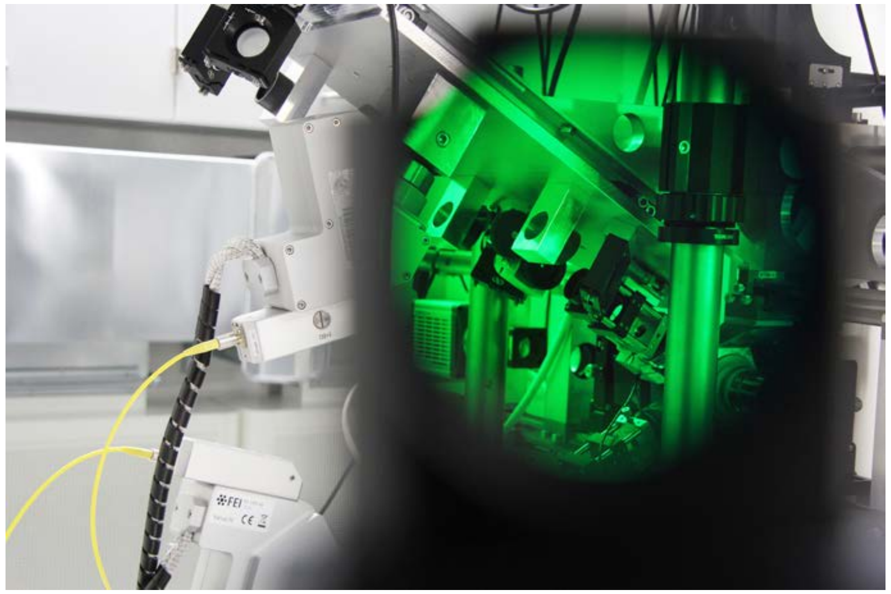
L Manukyan et al, Nature 2017

Hypothesis testing and Knowledge discovery



new paradigm

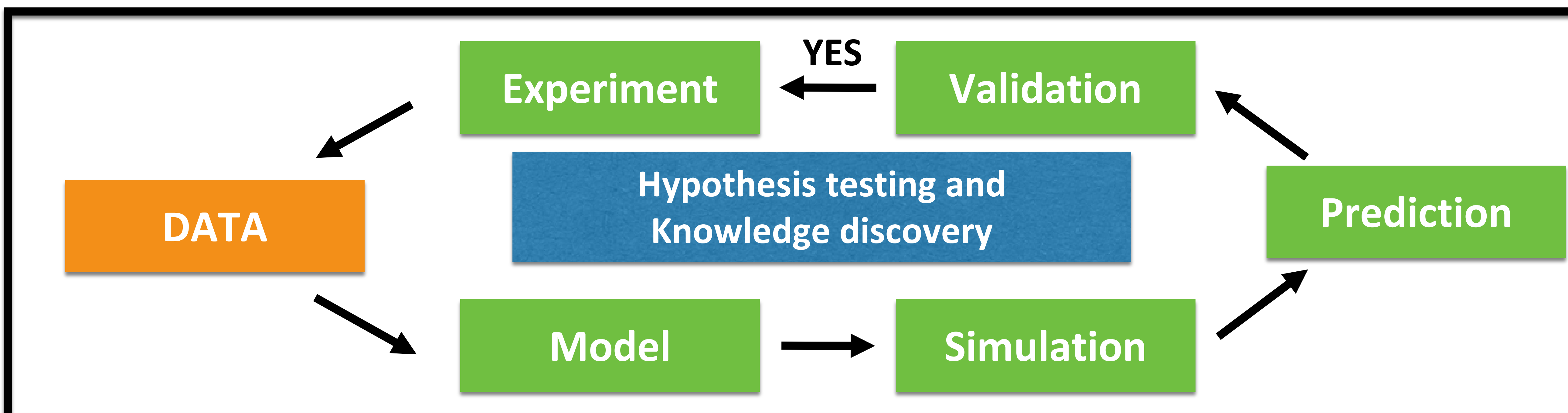
Biological microscopy data ERA



Self-driving light sheet microscope
- MPI-CBG/CSBD

BIOLOGICAL DATA-SETS

- long-term imaging of single molecules - **fluorescence microscopy**
- simultaneous measurements of multiple biosensors - **live cell microscopy**
- development of entire organisms - **lattice light-sheet microscopy, SPIM**



Microscopy data —> mathematical models PDE/ODEs

Formal definition of data-driven PDE inference

Generic non-linear, space time-dependent, parametric systems

State-variable

Parameters

$$\frac{\partial u}{\partial t} = \mathcal{N}\left(\left[u, u^2, u_{xx}, uu_x, \ldots\right], x, t, \Theta, \Sigma\right)$$

Dynamics

Stochastic effects

$$\frac{\partial u}{\partial t} = \beta_0 + \beta_1 u + \beta_2 u \frac{\partial u}{\partial x} + \beta_3 u^2 + \ldots.. + \beta_k u^3 \frac{\partial^2 u}{\partial x^2} + \ldots.,$$

Data and measurement model

State-variable measurements

$$v\left(x_m, t_n\right) = \mathcal{F}\left(u_{m,n}, x_m, t_n, \Xi\right)$$

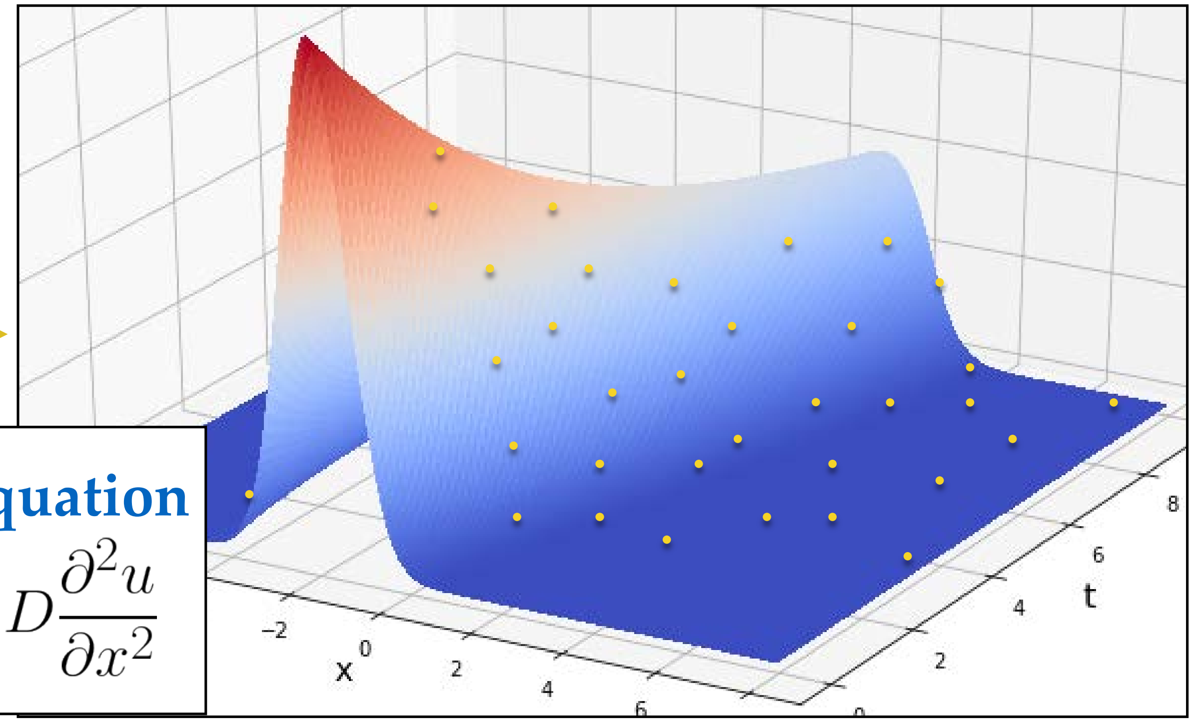
Measurement model

Measurement noise

Given data $u_{m,n}$ and $v\left(x_m, t_n\right) \approx \frac{\partial u}{\partial t}$, build $\mathcal{F}(\cdot) \approx \mathcal{N}(\cdot)$?

CASE STUDY: 1D BURGERS EQUATION inference

$u_{m,n}$



urgers Equation

$$u\frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2}$$

Representative microscopy data

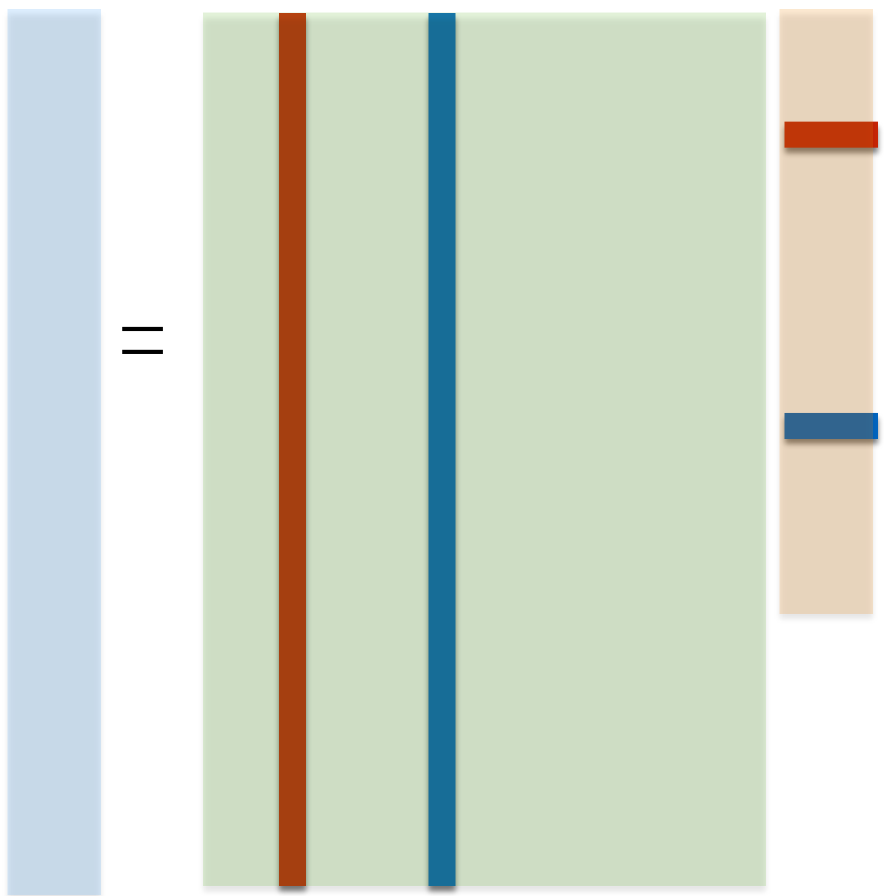
$$\begin{bmatrix} u_t(x_0,t_0) \\ u_t(x_1,t_0) \\ u_t(x_2,t_0) \\ \vdots \\ u_t(x_{n-1},t_m) \\ u_t(x_n,t_m) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & u(x_0,t_0) & u_x(x_0,t_0) & \dots & u^5u_{xxx}(x_0,t_0) \\ 1 & u(x_1,t_0) & u_x(x_1,t_0) & \dots & u^5u_{xxx}(x_1,t_0) \\ 1 & u(x_2,t_0) & u_x(x_2,t_0) & \dots & u^5u_{xxx}(x_2,t_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u(x_{n-1},t_m) & u_x(x_{n-1},t_m) & \dots & u^5u_{xxx}(x_{n-1},t_m) \\ 1 & u(x_n,t_m) & u_x(x_n,t_m) & \dots & u^5u_{xxx}(x_n,t_m) \end{bmatrix}}_{\text{Example dictionary for state variable u in one spatial dimension}} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix}$$

$U_t(N \times 1)$ $\Theta(N \times p)$ $\beta(p \times 1)$

N - #points sampled, p - dictionary size

The belief system - SPARSITY

Sparse - regression



$U_t(N \times 1)$ $\Theta(N \times p)$ $\beta(p \times 1)$

Least-squares formulation

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^p} ||U_t - \Theta\beta||_2^2$$

- Linear-least square solution (N>p)

Sparse regression formulation

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^p} ||U_t - \Theta\beta||_2^2 + \lambda||\beta||_0$$

data fitting term

Sparsity

- NP hard - heuristic algorithms exist

Rudy et al, Sci Advances (2017)

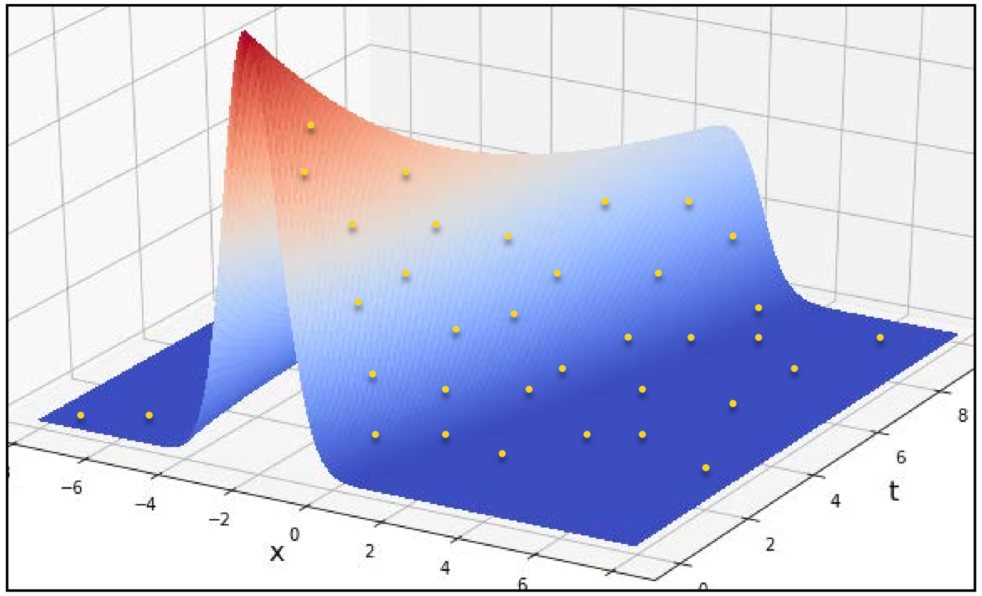
Relaxed sparse regression form

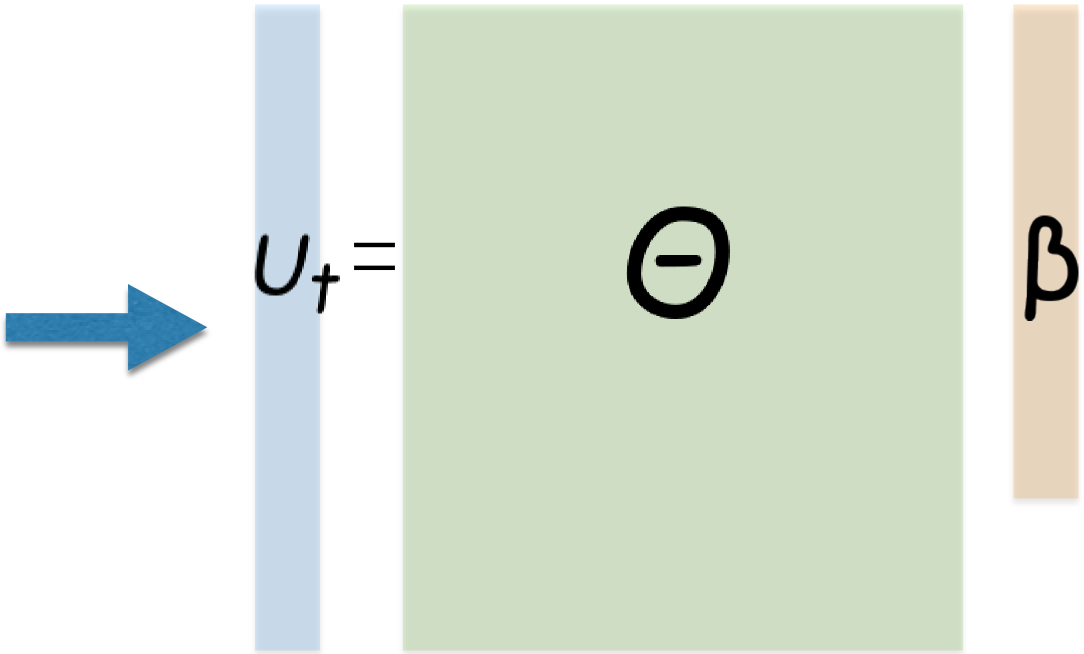
$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^p} ||U_t - \Theta\beta||_2^2 + \lambda||\beta||_{1,2}$$

- Lasso, Randmized Lasso, Elastic Net

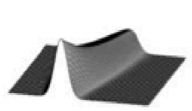
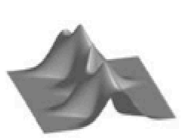
H Schaeffer et al, Royal Society (2016)

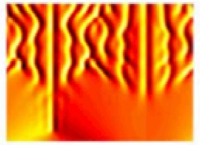
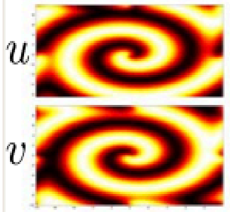
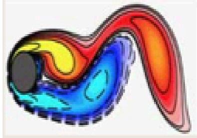
Success with sparse regression for PDE inference





$$\hat{\beta} = \arg\min_{\beta} \| U_t - \Theta\beta \|_2^2 + \lambda \|\beta\|_0$$

PDE		Form
	KdV	$u_t + 6uu_x + u_{xxx} = 0$
	Burgers	$u_t + uu_x - \epsilon u_{xx} = 0$
	Schrödinger	$iu_t + \frac{1}{2}u_{xx} - \frac{x^2}{2}u = 0$

	KS	$u_t + uu_x + u_{xx} + u_{xxxx} = 0$
u, v 	Reaction Diffusion	$u_t = 0.1\nabla^2 u + \lambda(A)u - \omega(A)v$ $v_t = 0.1\nabla^2 v + \omega(A)u + \lambda(A)v$ $A^2 = u^2 + v^2, \omega = -\beta A^2, \lambda = 1 - A^2$
	Navier-Stokes	$\omega_t + (\mathbf{u} \cdot \nabla)\omega = \frac{1}{Re}\nabla^2\omega$

Rudy et al, Sci Advances (2017)

Barriers to cross for application to real experimental data

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^p} \|U_t - \Theta\beta\|_2^2 + \lambda\|\beta\|_0$$

data fitting term

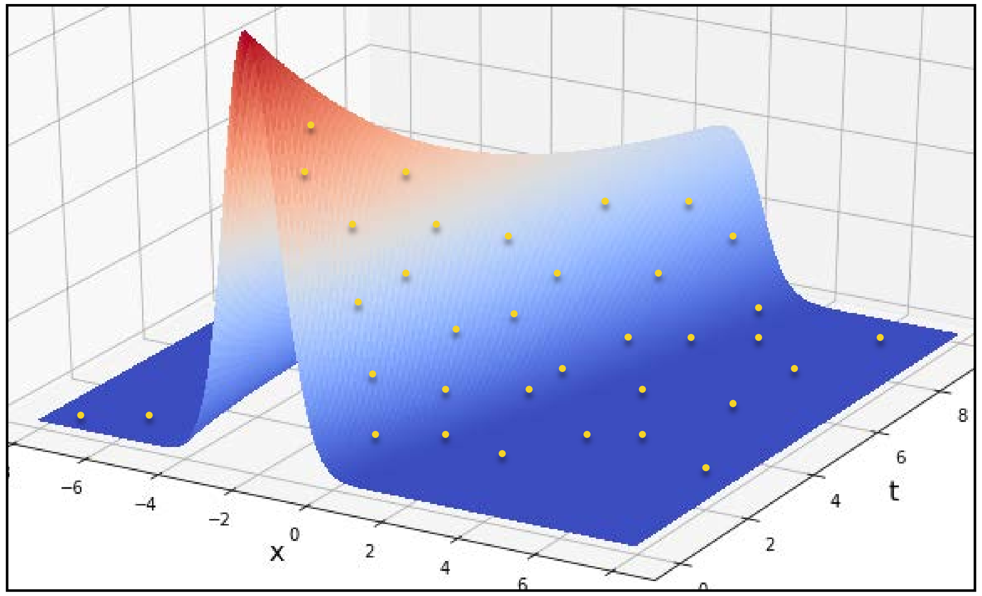
Sparsity

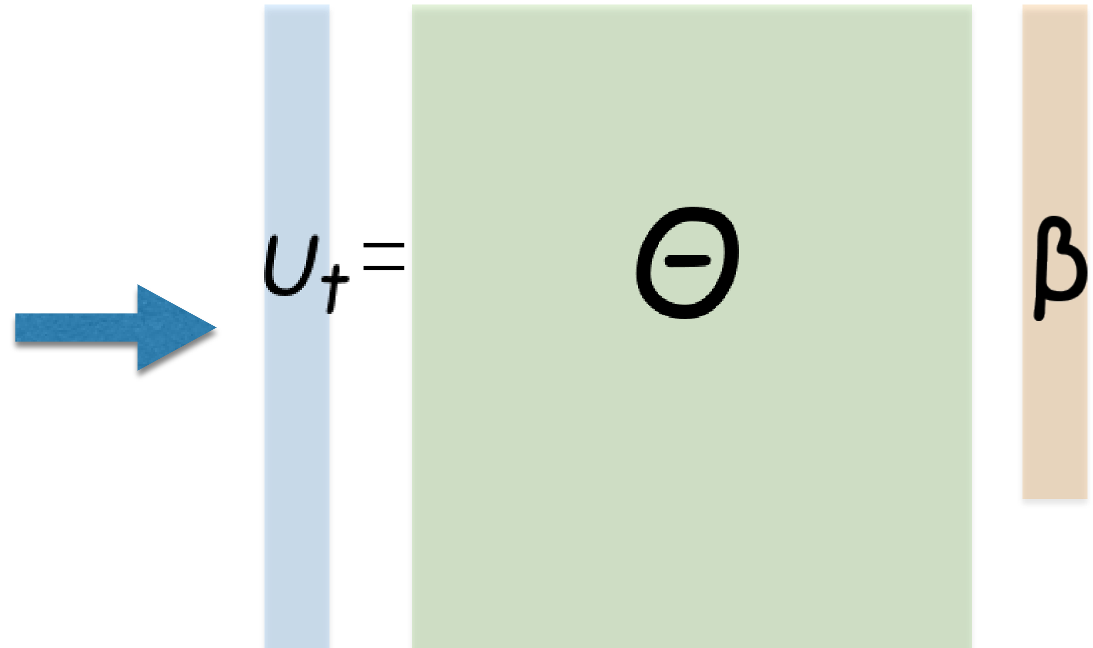
- Parametric dependency of the algorithm with change in design, What is the right **complexity parameter** lambda ?
- Does the algorithm always give the right support ? Is this the best way to solve the hard L_0 problem ?
- How can we compare between algorithms ? Can we guarantee consistency for varying design's ?

- Derivative computation from noisy data
- Noise performance (1%)

parameter dependence

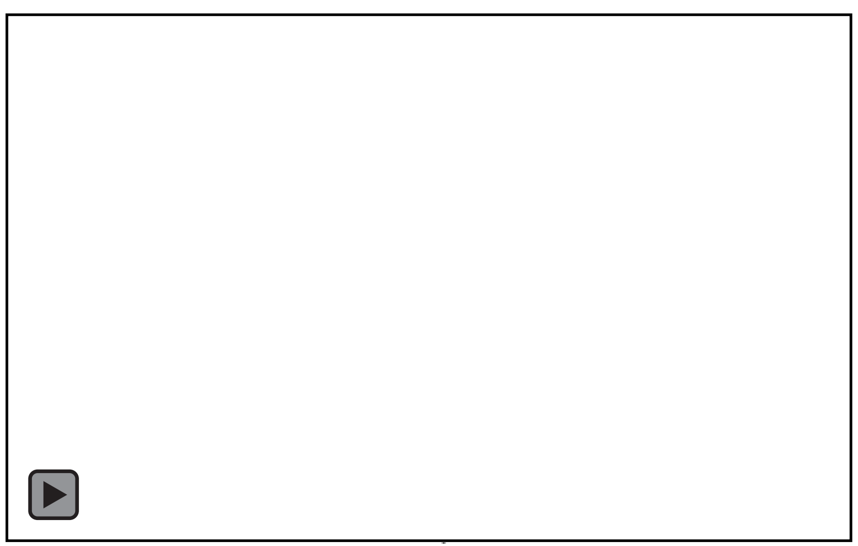
Regularization paths for 1D BURGERS EQUATION inference





$$\hat{\beta} = \arg\min_{\beta} \|U_t - \Theta\beta\|_2^2 + \lambda\|\beta\|_1$$

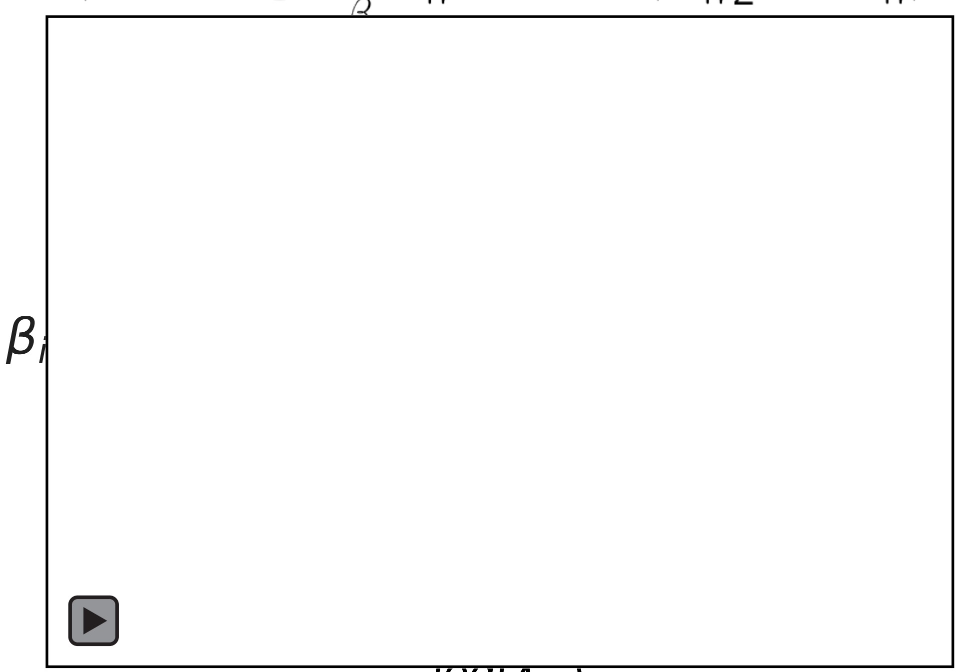
β_i



$-log(\lambda^*)$

$$\hat{\beta} = \arg\min_{\beta} \|U_t - \Theta\beta\|_2^2 + \lambda\|\beta\|_0$$

β_i

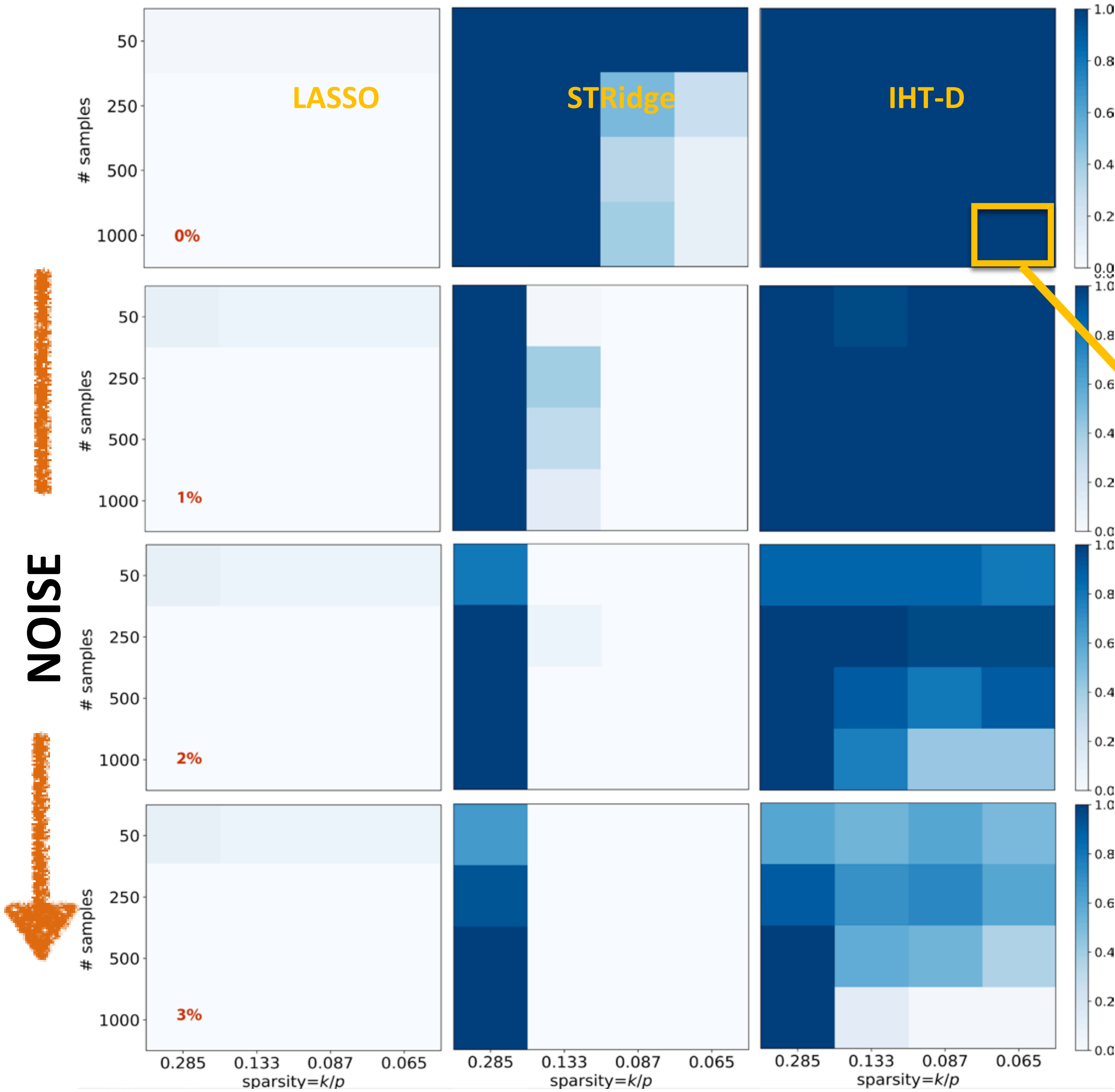


$-log(\lambda^*)$

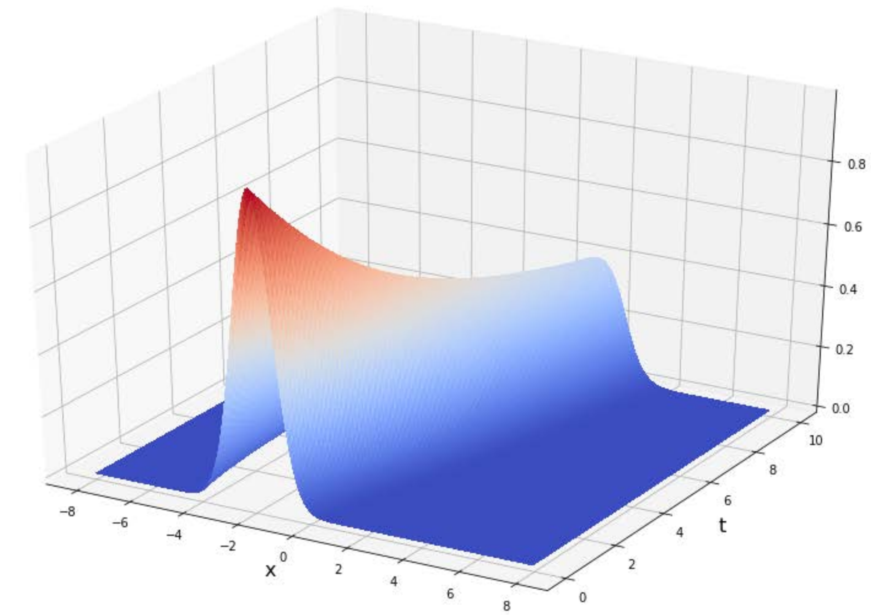
But which algorithm has the best recoverability?

Best sparsity-promoter

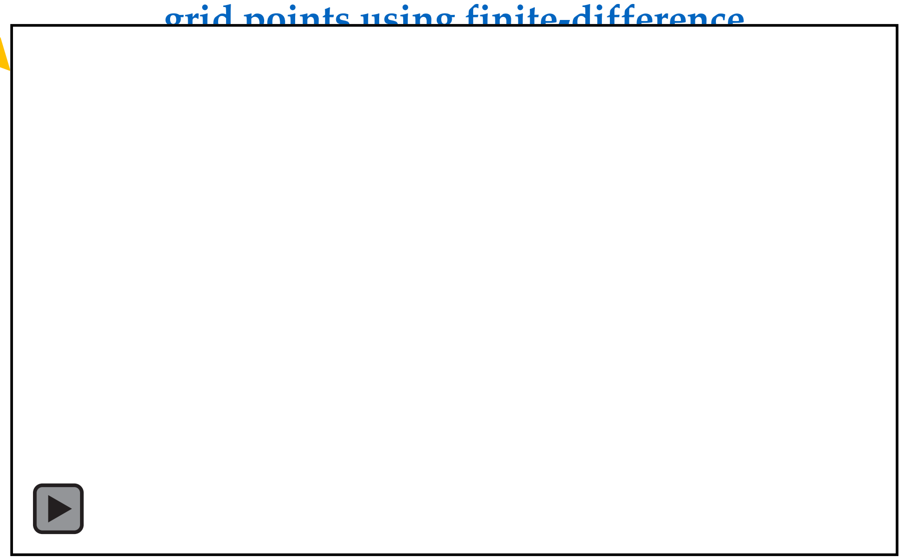
Which sparsity promoter ?



Success probability through oracle lambda for noise = 0-3%



Numerical simulation of 1D Burgers on 128 space grid points using finite-difference



k = # true components
p = dictionary size
STRidge parameter = 1e-5
Polynomial differentiation

IHT-D shows better performance in comparison to STRidge and LASSO

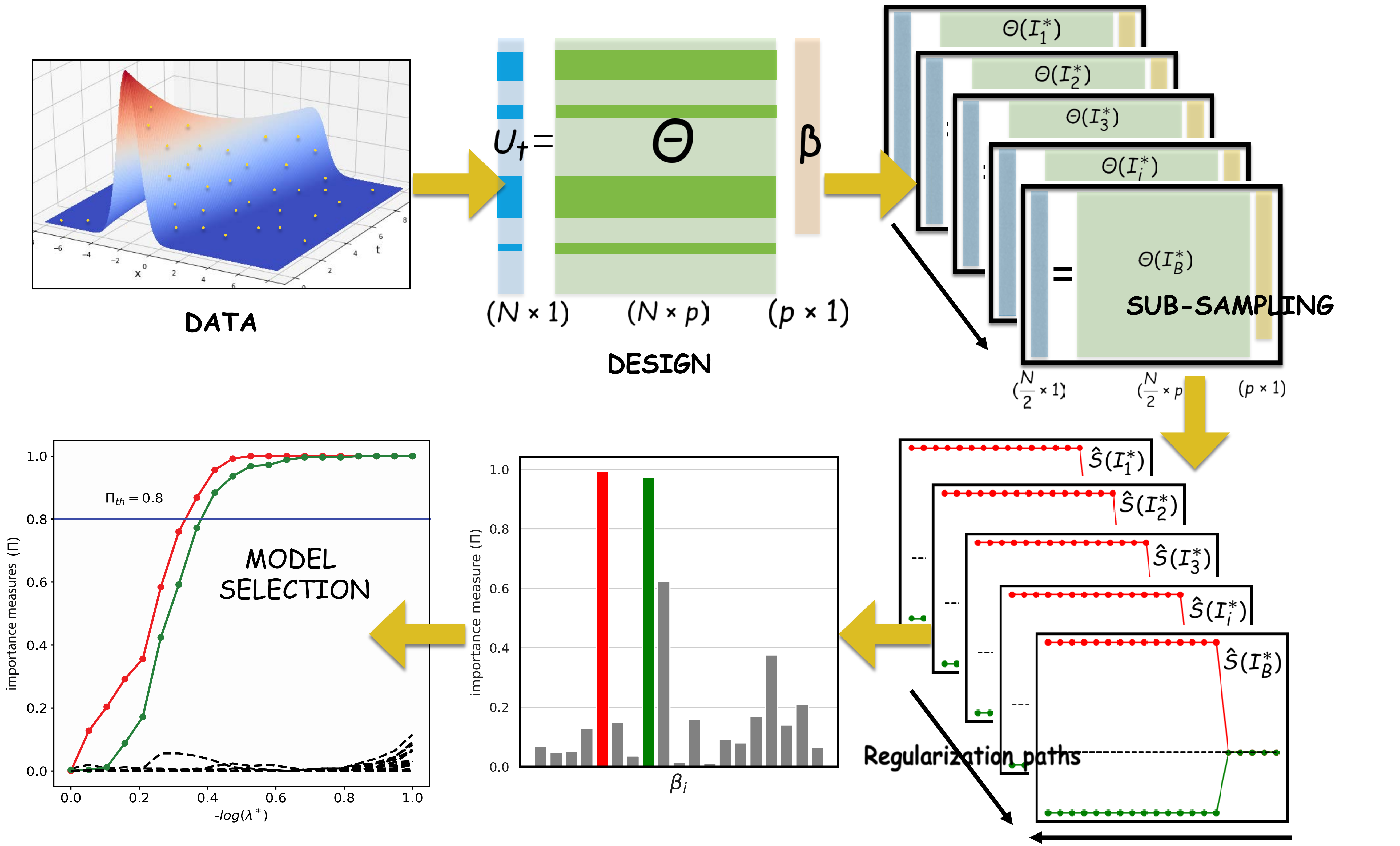
center for systems biology dresden

TECHNISCHE UNIVERSITÄT DRESDEN

FLATIRON INSTITUTE Center for Computational Mathematics

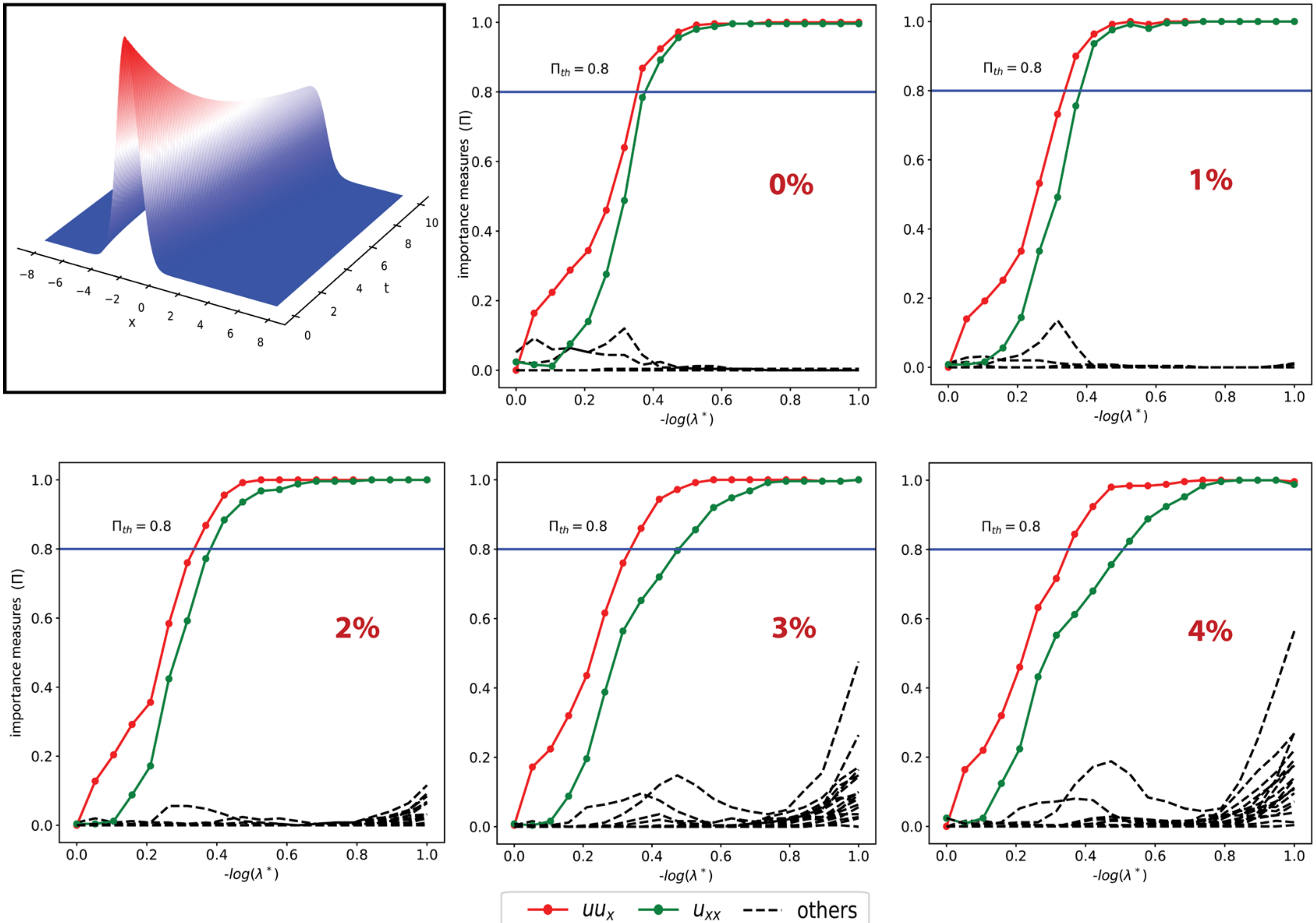
CBG Max Planck Institute of Molecular Cell Biology and Genetics parameter dependence

Stability-based model selection for 1D BURGERS EQUATION



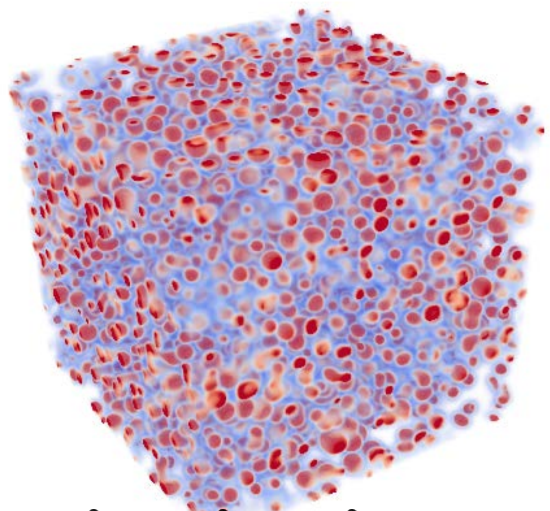
Robustness/ Consistency

Stability-based model selection for 1D BURGERS EQUATION



Design N=200, p=20

Stability-based model selection for 3D Gray-Scott reaction diffusion system

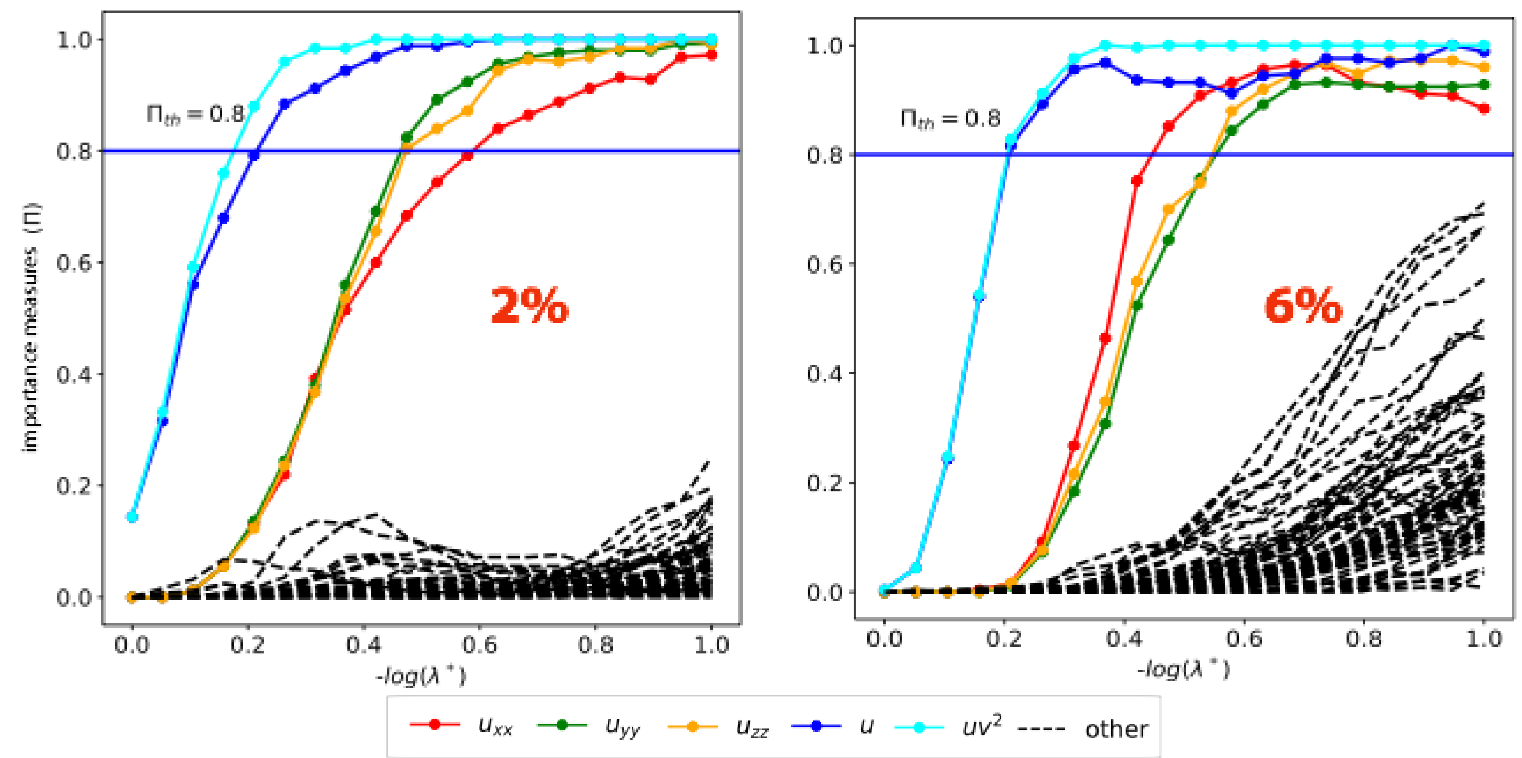


$$u_t = D_u \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) - uv^2 + f(1-u),$$

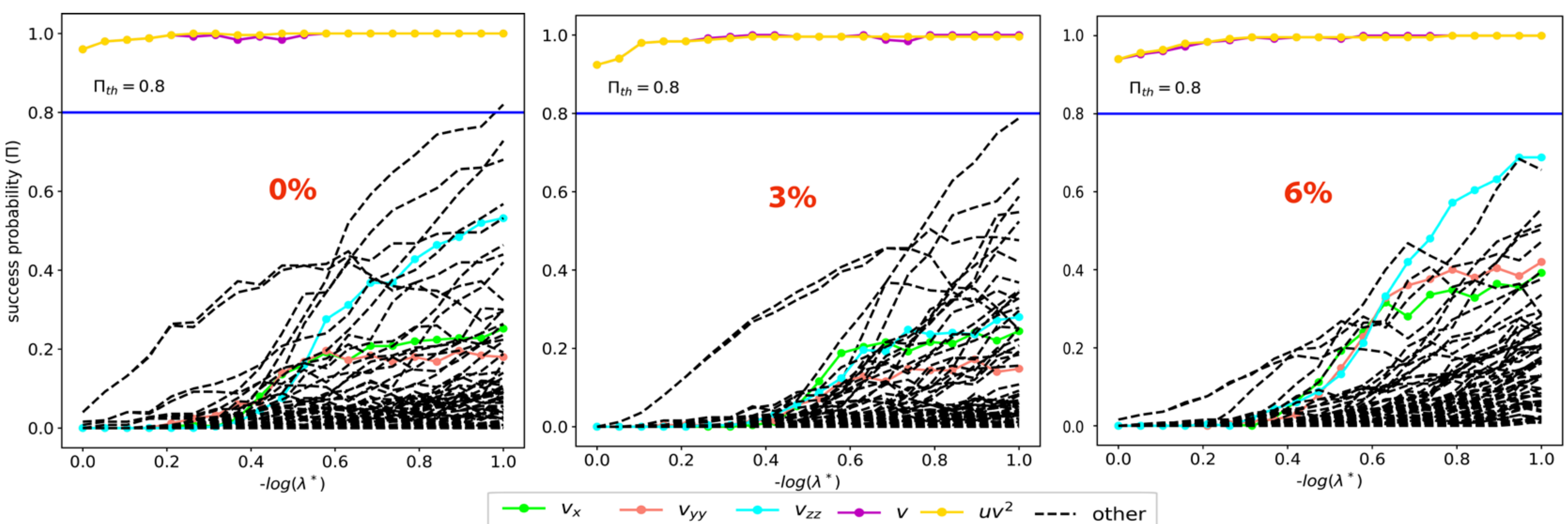
$$v_t = D_v \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) + uv^2 - (f+k)v.$$

$$k = 0.053,\ f = 0.014,\ D_u = 2e^{-5},\ D_v = 1e^{-5}$$

periodic boundary conditions

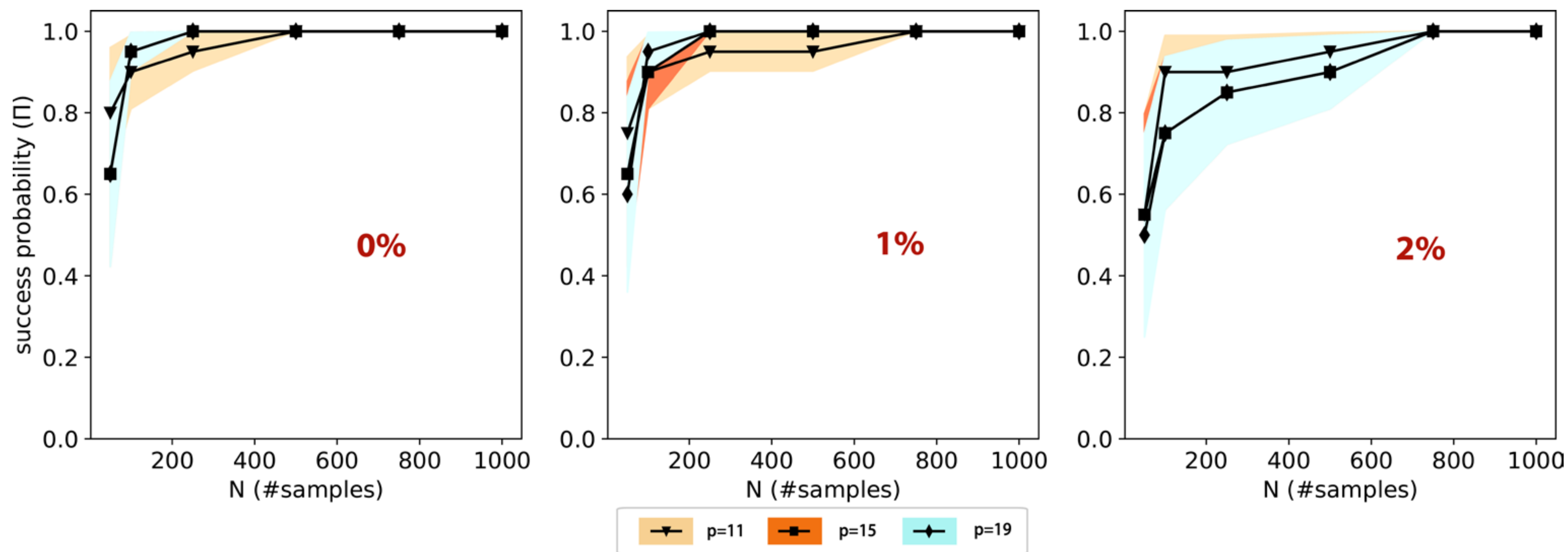


Design N=400, p=69

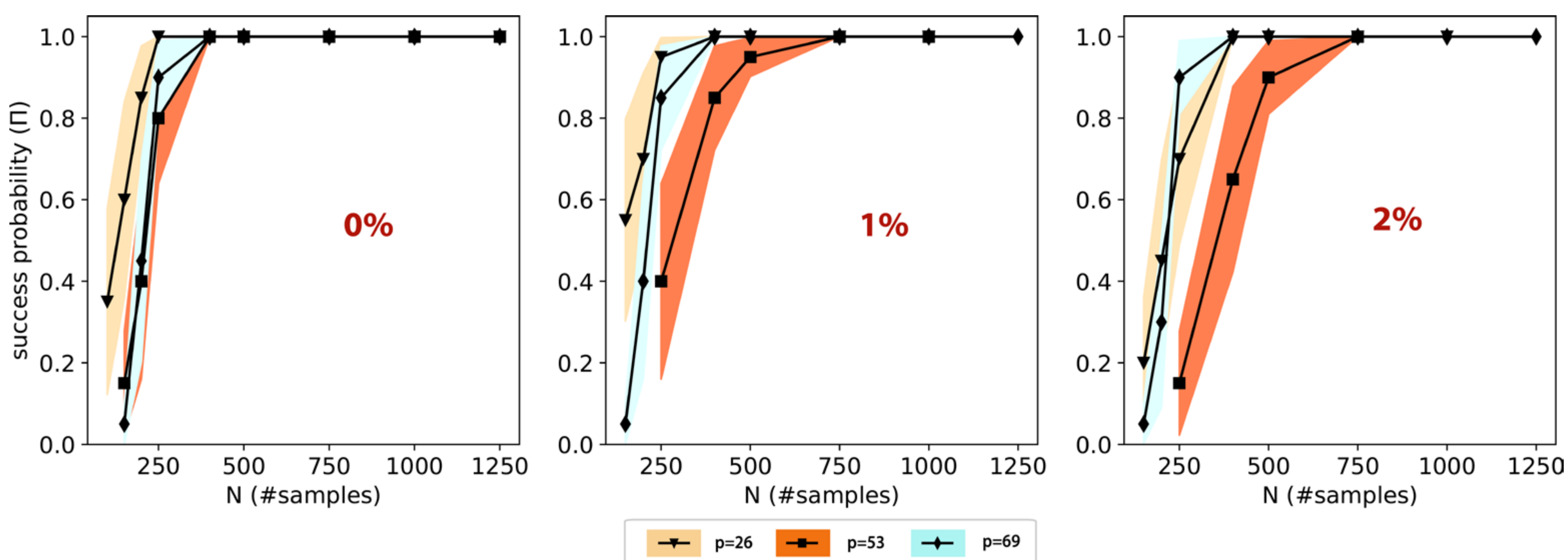


Achievability results

1D Burgers equation

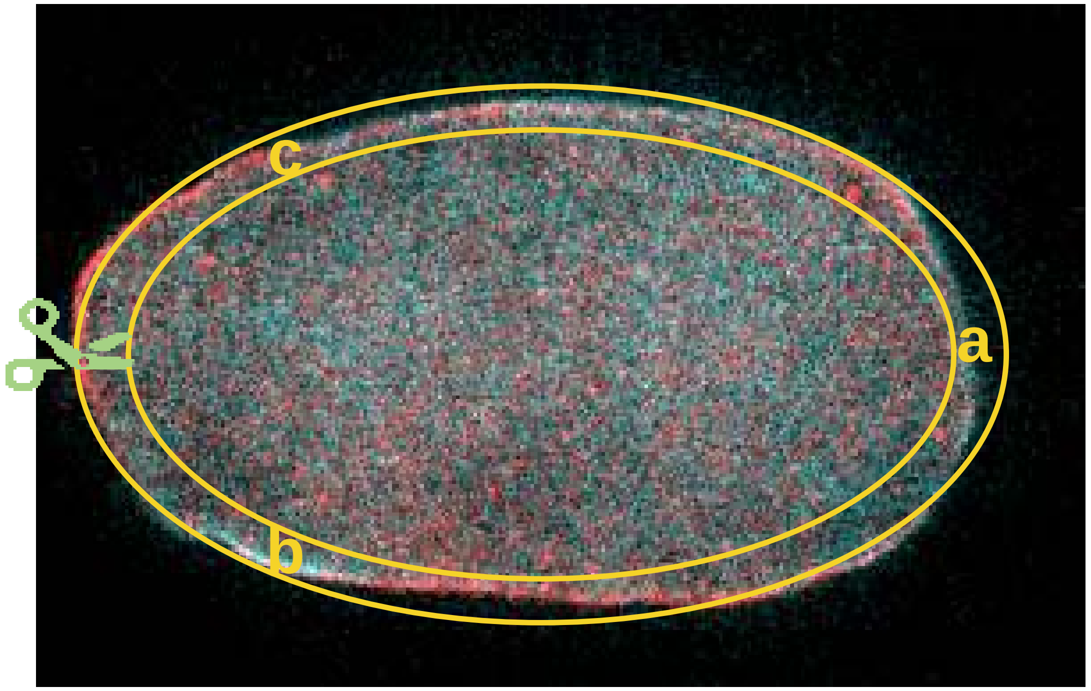


3D Gray-scott equation

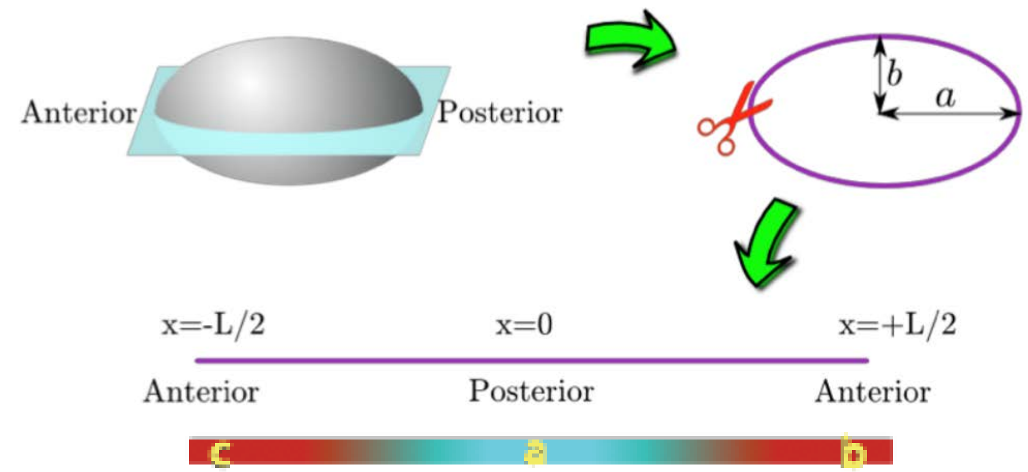


Experimental data

PAR Proteins patterning in the *Caenorhabditis elegans* zygote



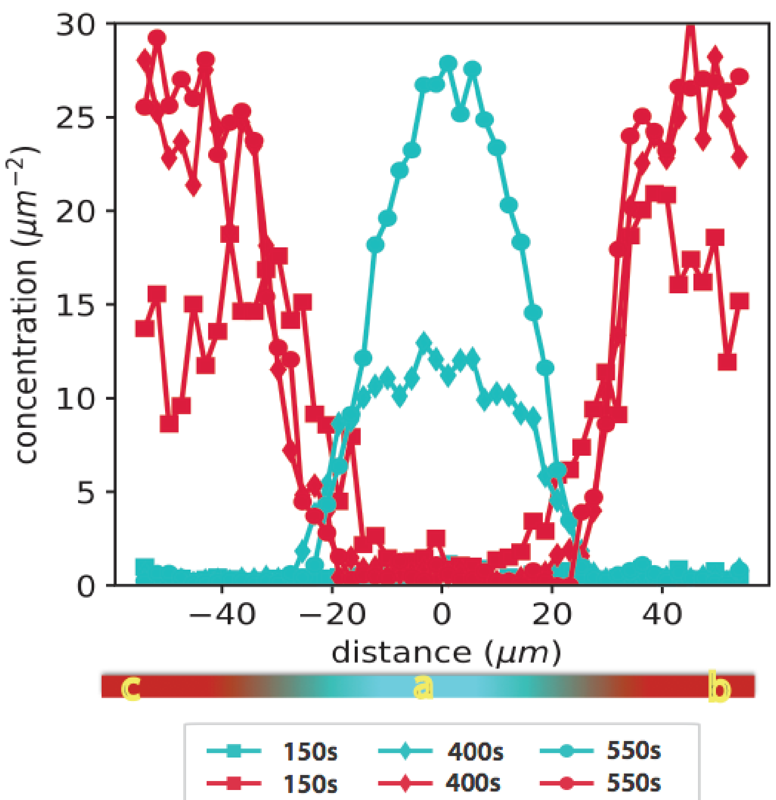
Grill lab, MPI-CBG

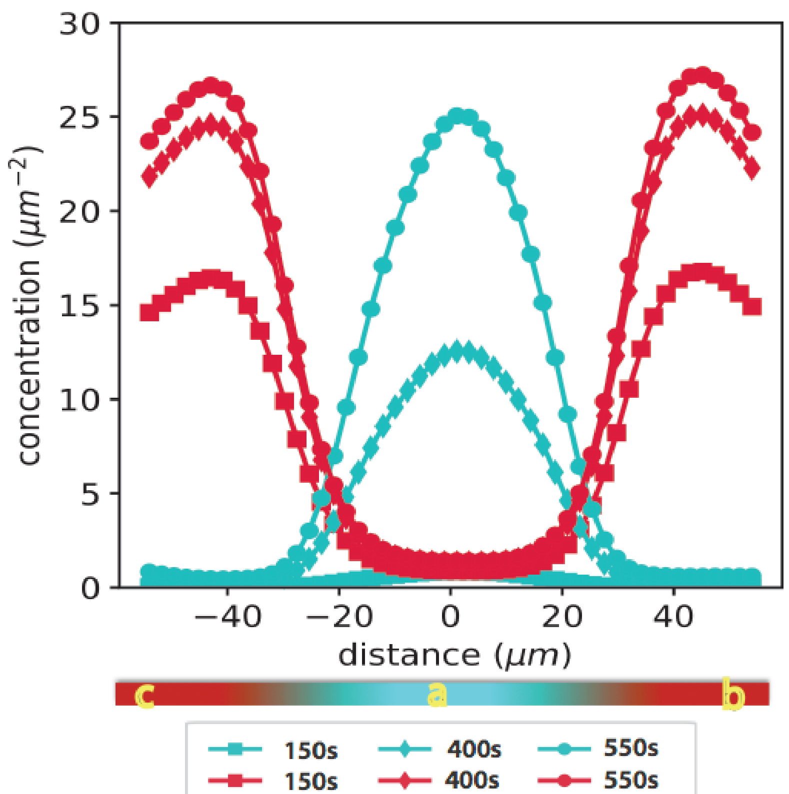


$$\partial_t A = D_A \partial_x^2 A - \partial_x(vA) + R_A \qquad R_A = k_{on,A} A_{cyto} - k_{off,A} A - k_{AP} P^{\alpha} A$$

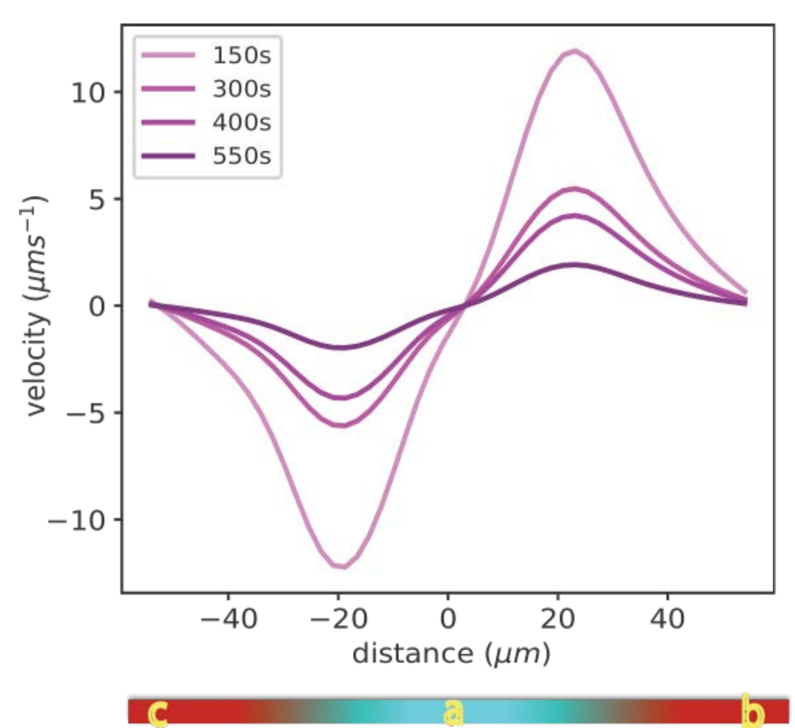
$$\partial_t P = D_P \partial_x^2 P - \partial_x(vP) + R_P \qquad R_P = k_{on,P} P_{cyto} - k_{off,P} P - k_{PA} A^{\beta} P$$

Protein Concentration



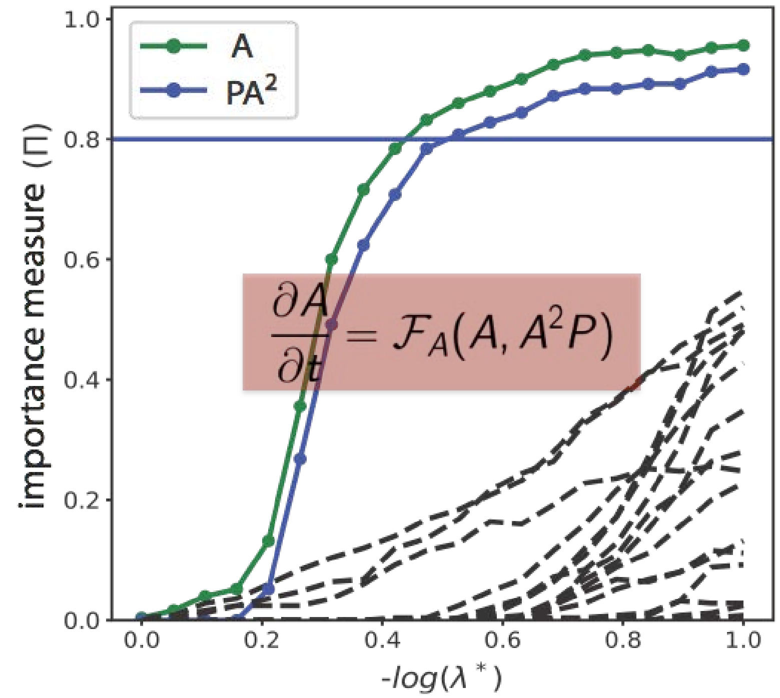


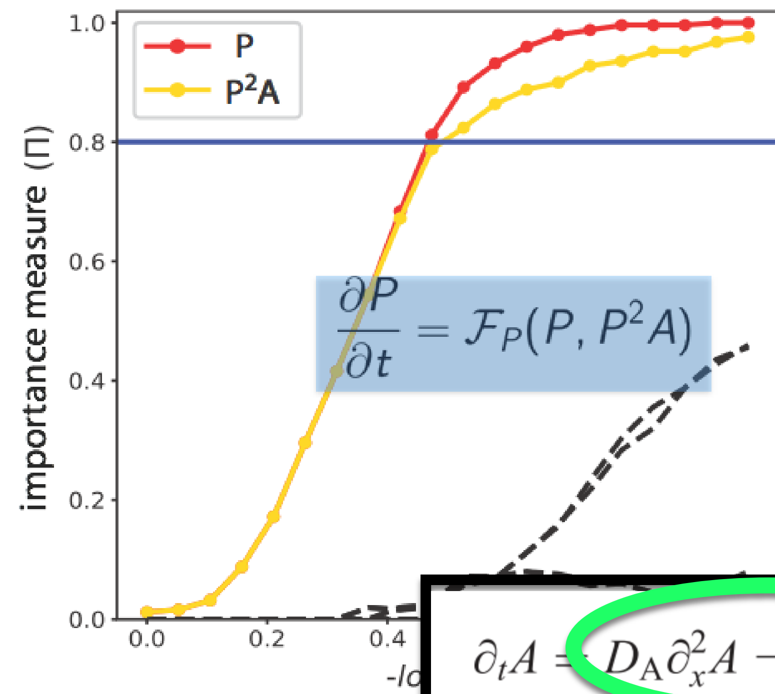
cortical flow field

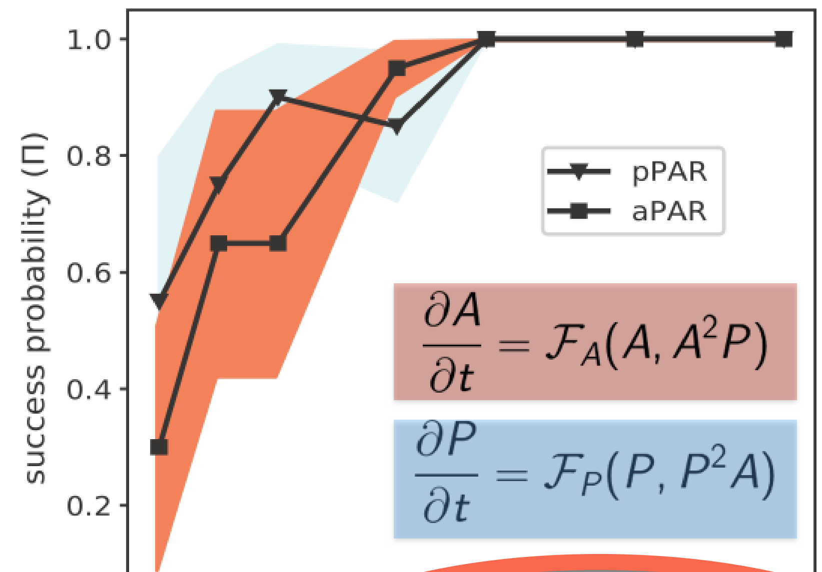


Goehring et al, Science (2011)

Stability-based PDE inference for PAR system experimental data



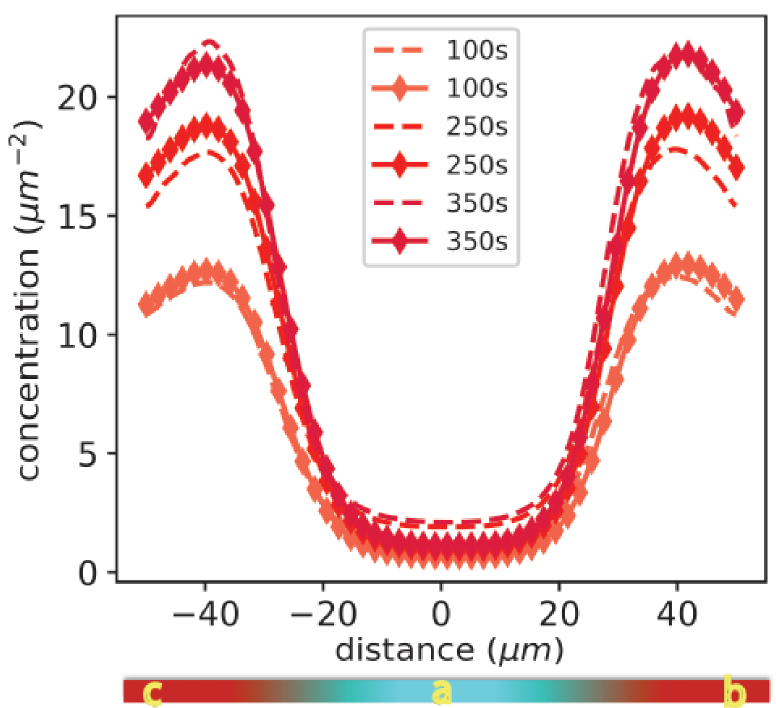


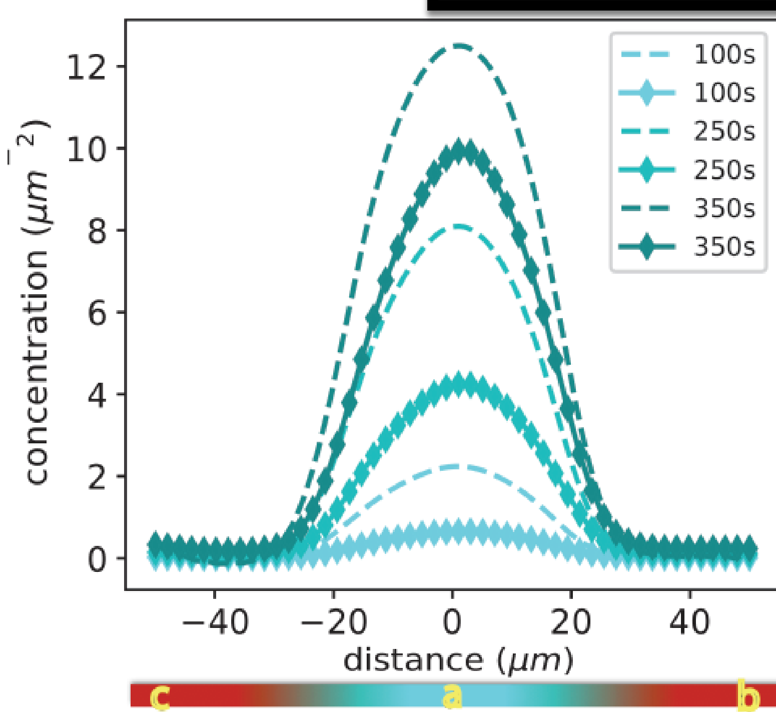


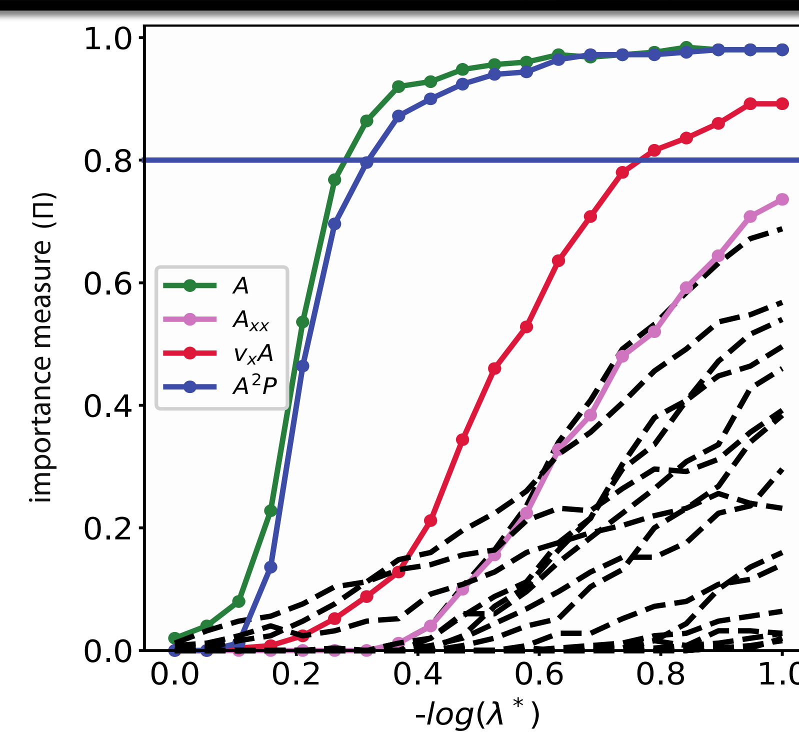
$$\partial_t A = D_A \partial_x^2 A - \partial_x(vA) + R_A \qquad R_A = k_{on,A} A_{cyto} - k_{off,A} A - k_{AP} P^{\alpha} A$$

$$\partial_t P = D_P \partial_x^2 P - \partial_x(vP) + R_P \qquad R_P = k_{on,P} P_{cyto} - k_{off,P} P - k_{PA} A^{\beta} P$$

Design N=500, p=20







Next steps

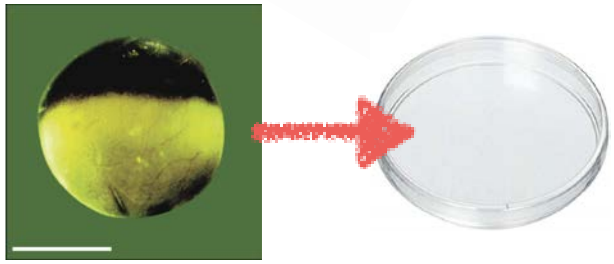
What kind of problems can be addressed with such methods ?

red flour beetle
(Tribolium castaneum)

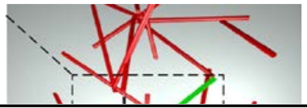
Tissue flows in Tribolium embryo
(Tomancak lab, MPI-CBG)

Can active gel mechanical models be Inferred from the intensity and flow data ?

Xenopus laevis



J Foster, eLife(2015)

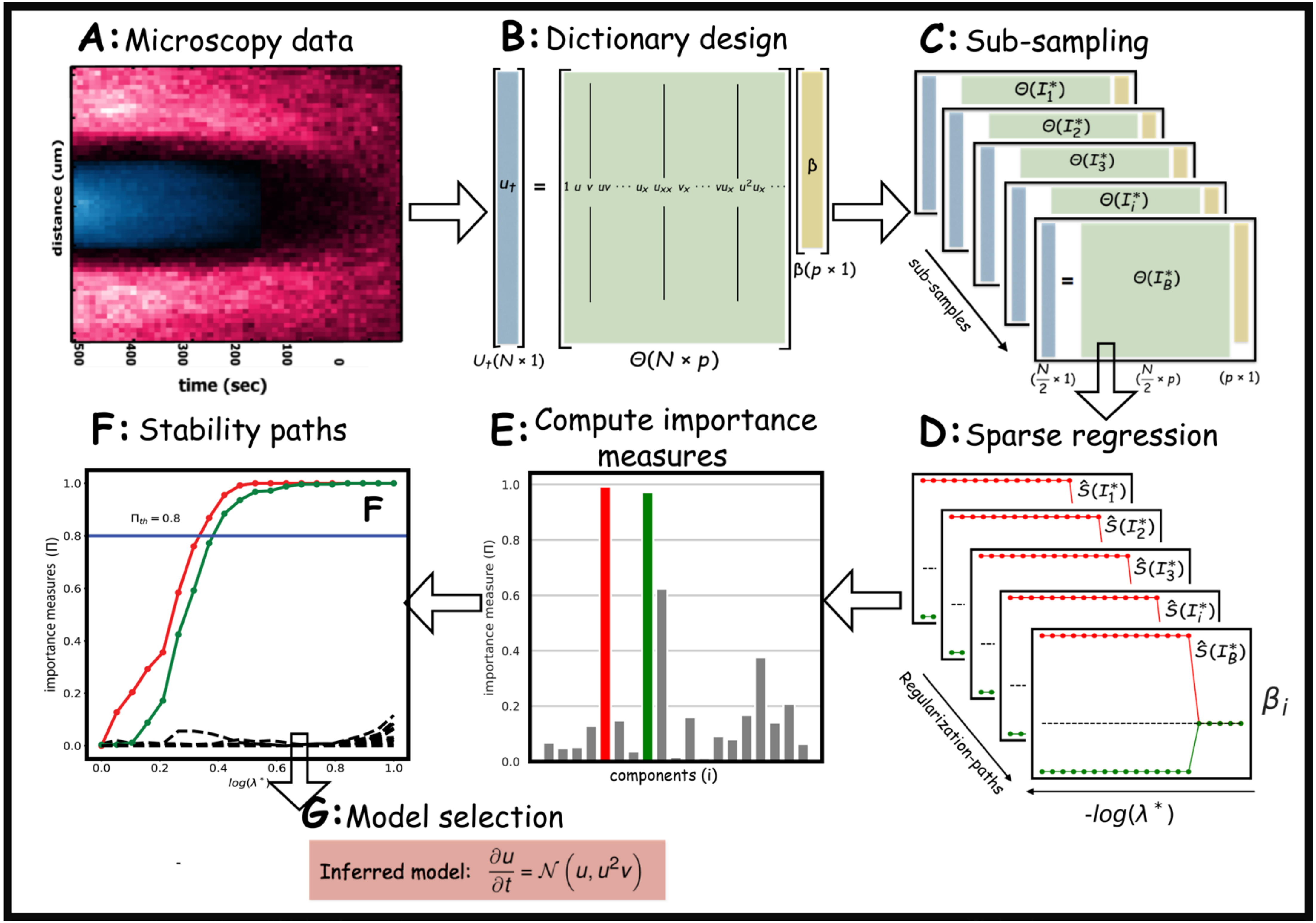


Sebastian Fürthauer (Flatiron Institute,NY)

Can we bridge the gap using data-driven coarse-grained modeling ?

center for systems biology dresden | TECHNISCHE UNIVERSITÄT DRESDEN | FLATIRON INSTITUTE Center for Computational Mathematics | CBG Max Planck Institute of Molecular Cell Biology and Genetics

Stability-based PDE inference from limited noisy spatial-temporal data
data



Acknowledgements

joint work with

Dr. Christian Mueller
Flatiron Institute, New York

Prof Ivo F. Sbalzarini,
CSBD, MPI-CBG, TU Dresden

Dr. Bevan Cheeseman,
University of Cambridge

supported by

MAX-PLANCK-GESELLSCHAFT

FLATIRON INSTITUTE
Center for Computational Mathematics

TECHNISCHE UNIVERSITÄT DRESDEN

MOSAIC Group
https://mosaic.mpi-cbg.de/

THANK YOU