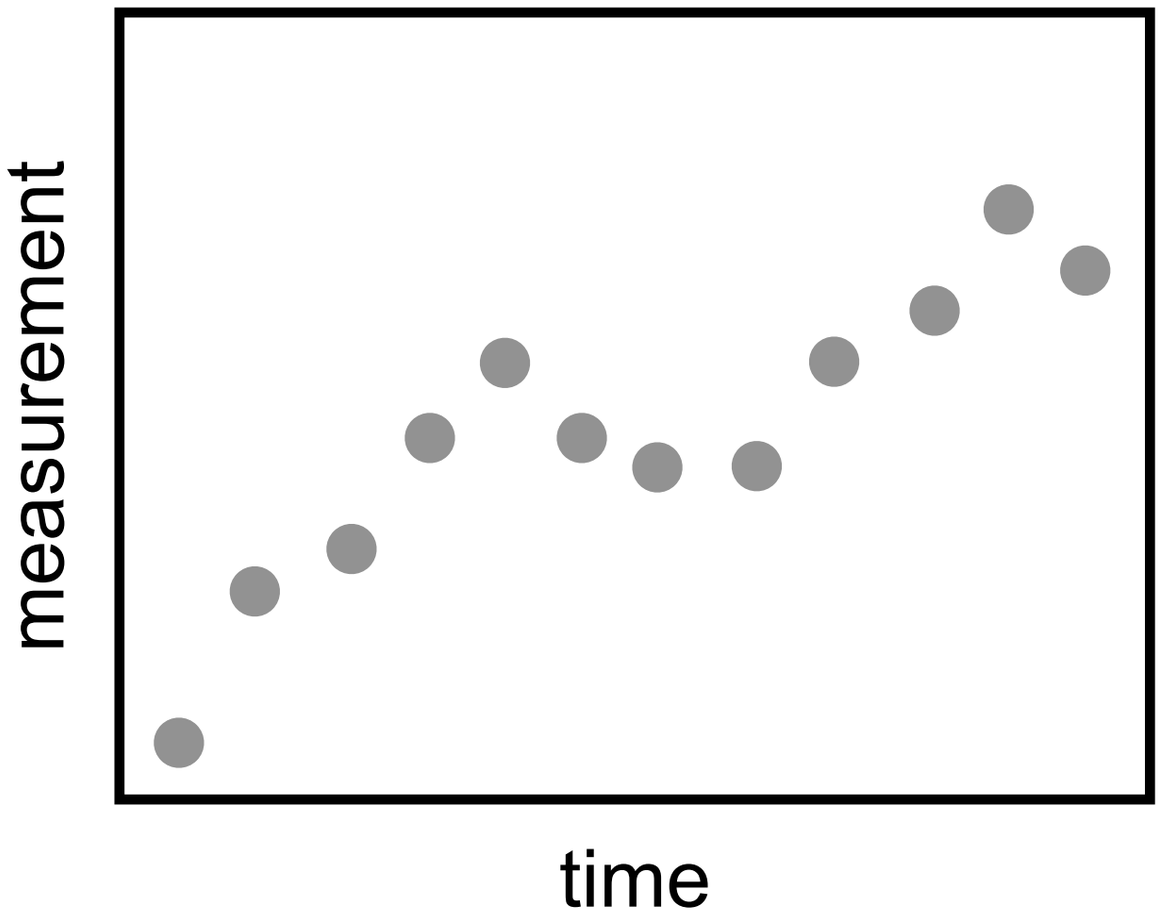
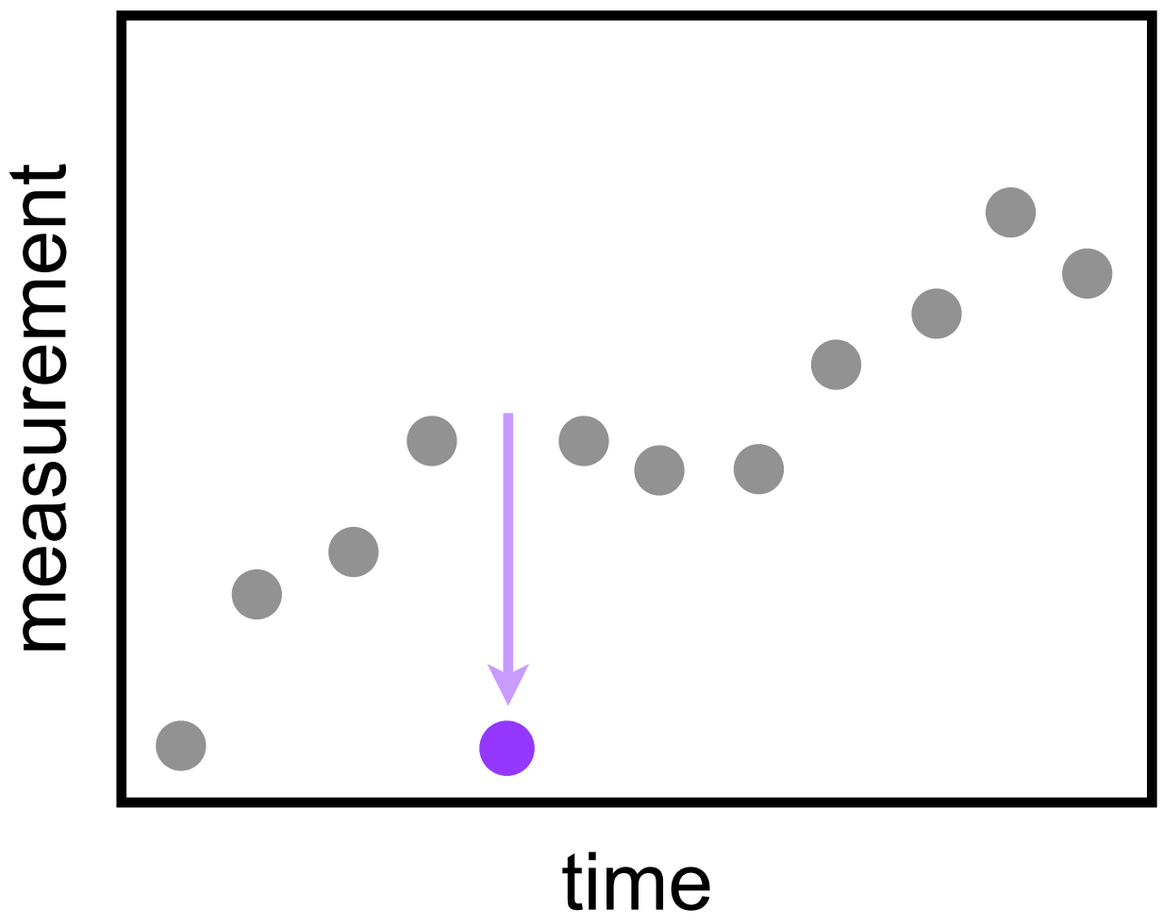
How can we deal with outliers?

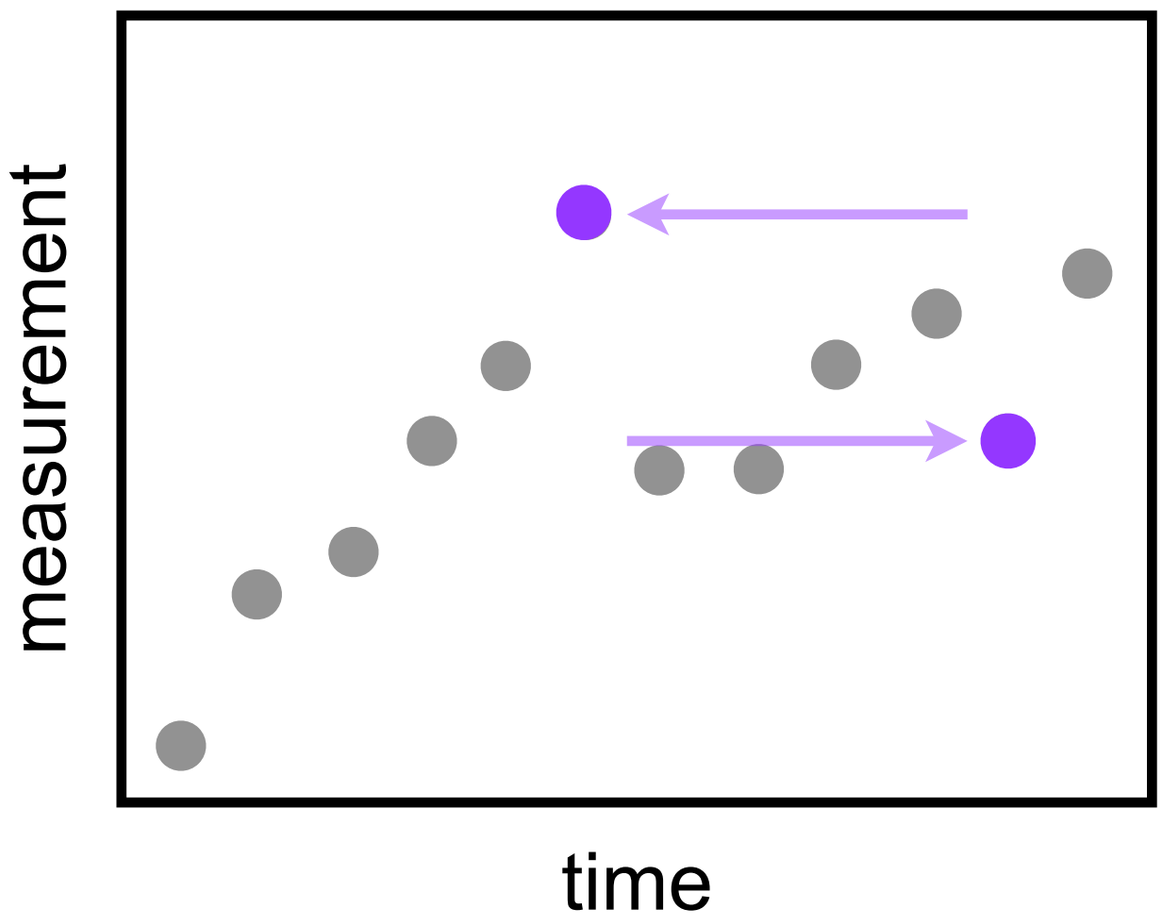
Outlier in measurement data



Outlier in measurement data

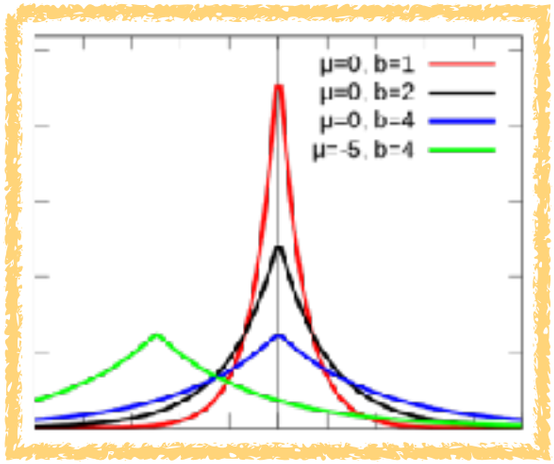


Outlier in measurement data



How to handle extreme outlier?

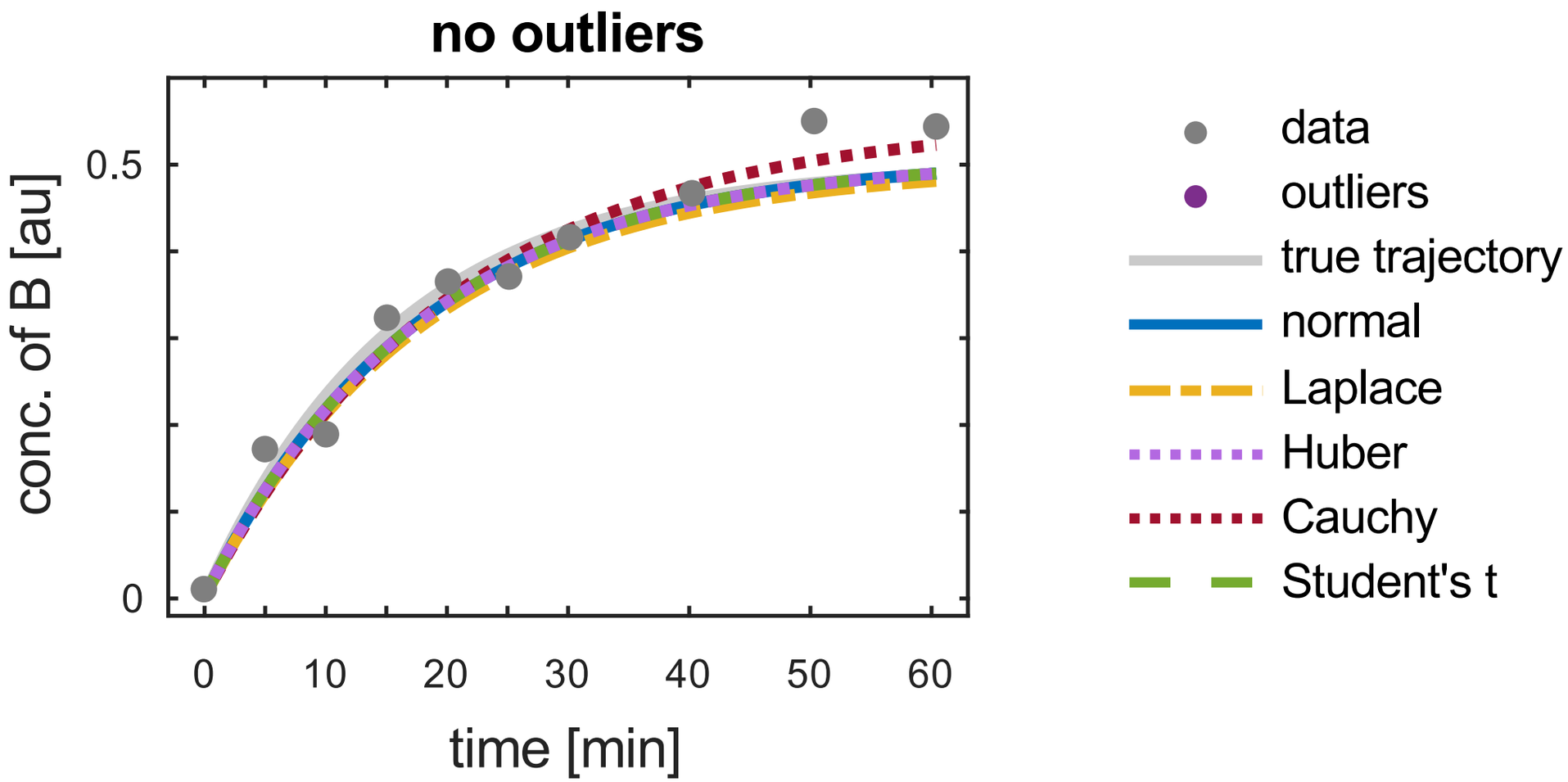
Probability distributions with "heavier" tails



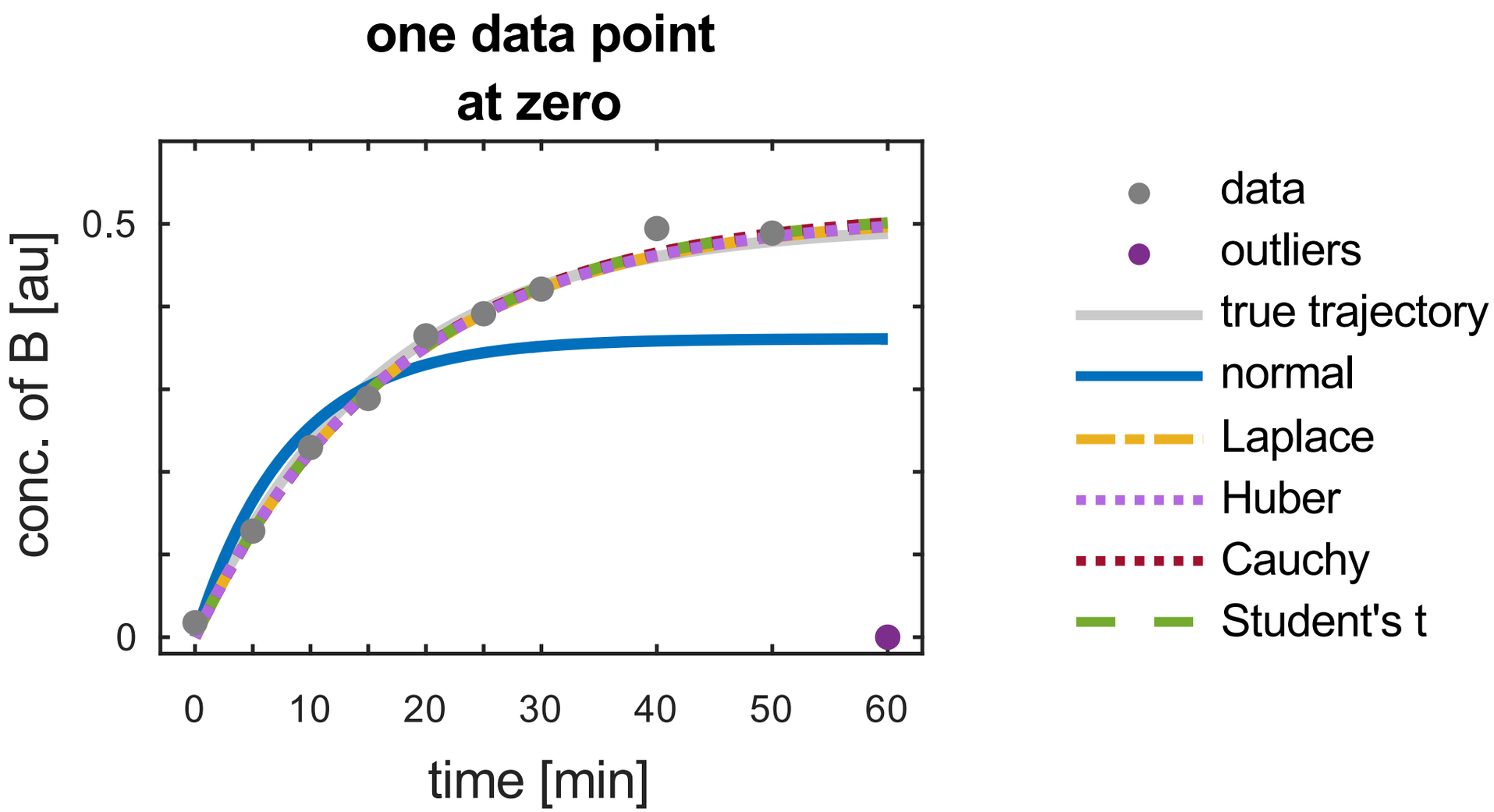
	probability density	distribution parameters φ
normal	$p(\bar{y}\|y,\sigma_n) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{1}{2}\left(\frac{\bar{y}-y}{\sigma_n}\right)^2\right)$	standard deviation $\sigma_n > 0$
Huber	$p(\bar{y}\|y,\sigma_h,\kappa) = s\cdot\begin{cases}\exp\left(-\frac{1}{2}\left(\frac{\bar{y}-y}{\sigma_h}\right)^2\right), & \left\|\frac{\bar{y}-y}{\sigma_h}\right\| \le \kappa \\ \exp\left(-\frac{1}{2}\left(2\kappa\left\|\frac{\bar{y}-y}{\sigma_h}\right\| - \kappa^2\right)\right), & \left\|\frac{\bar{y}-y}{\sigma_h}\right\| > \kappa\end{cases}$ with $s = \left(\sqrt{2\pi}\sigma_h \mathrm{erf}\left(\frac{\kappa}{\sqrt{2}}\right) + \frac{2\sigma_h}{\kappa}\exp\left(-\frac{1}{2}\kappa^2\right)\right)^{-1}$	scale $\sigma_h > 0$, tuning $\kappa > 0$
Laplace	$p(\bar{y}\|y,b) = \frac{1}{2b}\exp\left(-\frac{\|\bar{y}-y\|}{b}\right)$	scale $b > 0$
Cauchy	$p(\bar{y}\|y,\gamma) = \frac{1}{\pi\gamma}\frac{\gamma^2}{(\bar{y}-y)^2+\gamma^2}$	scale $\gamma > 0$
Student's t	$p(\bar{y}\|y,\sigma_t,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma_t}\left(1+\frac{1}{\nu}\left(\frac{\bar{y}-y}{\sigma_t}\right)^2\right)^{-\frac{\nu+1}{2}}$	scale $\sigma_t > 0$, degree $\nu > 0$

Performance using these distributions?

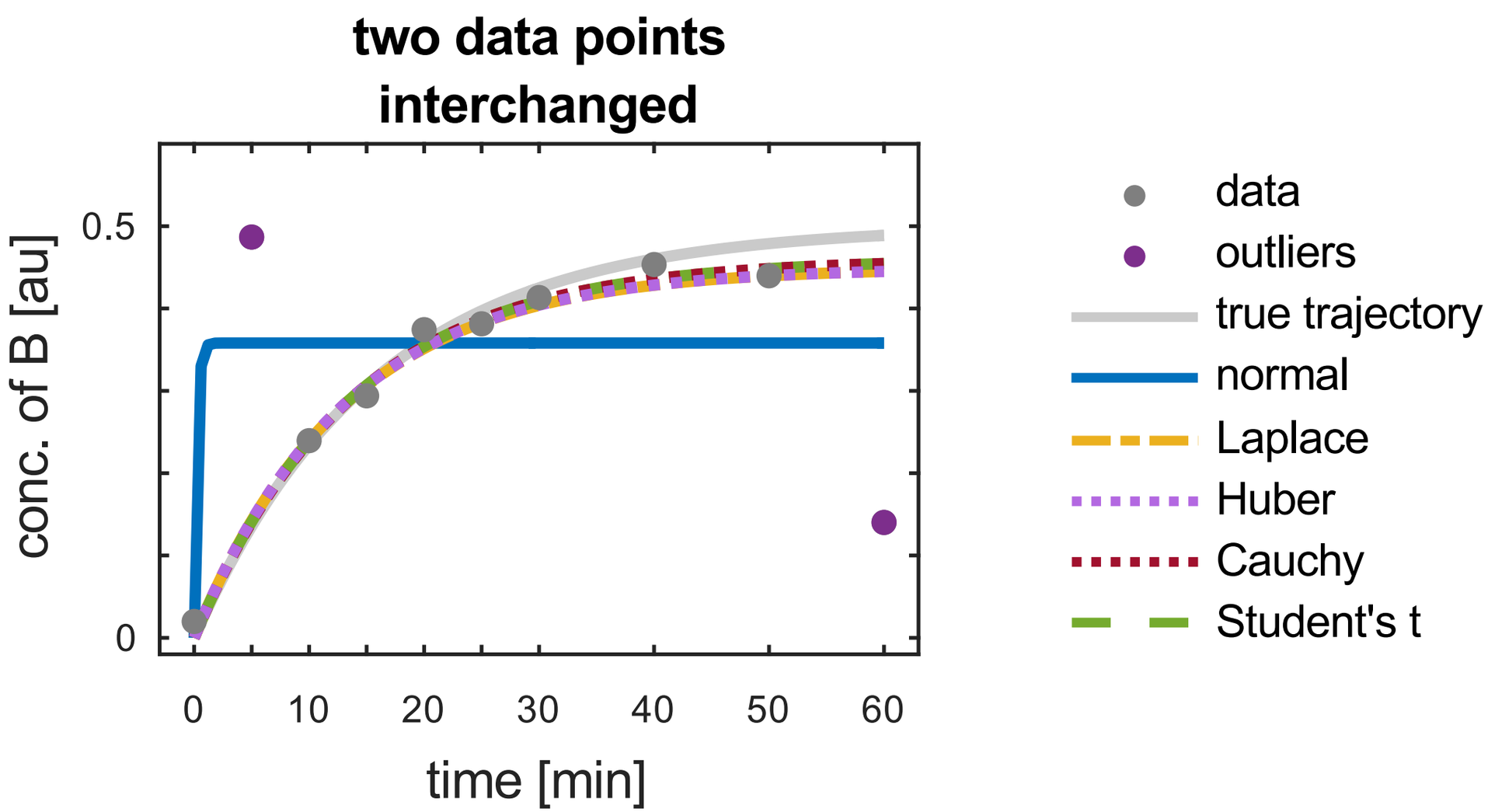
Improved robustness to outliers



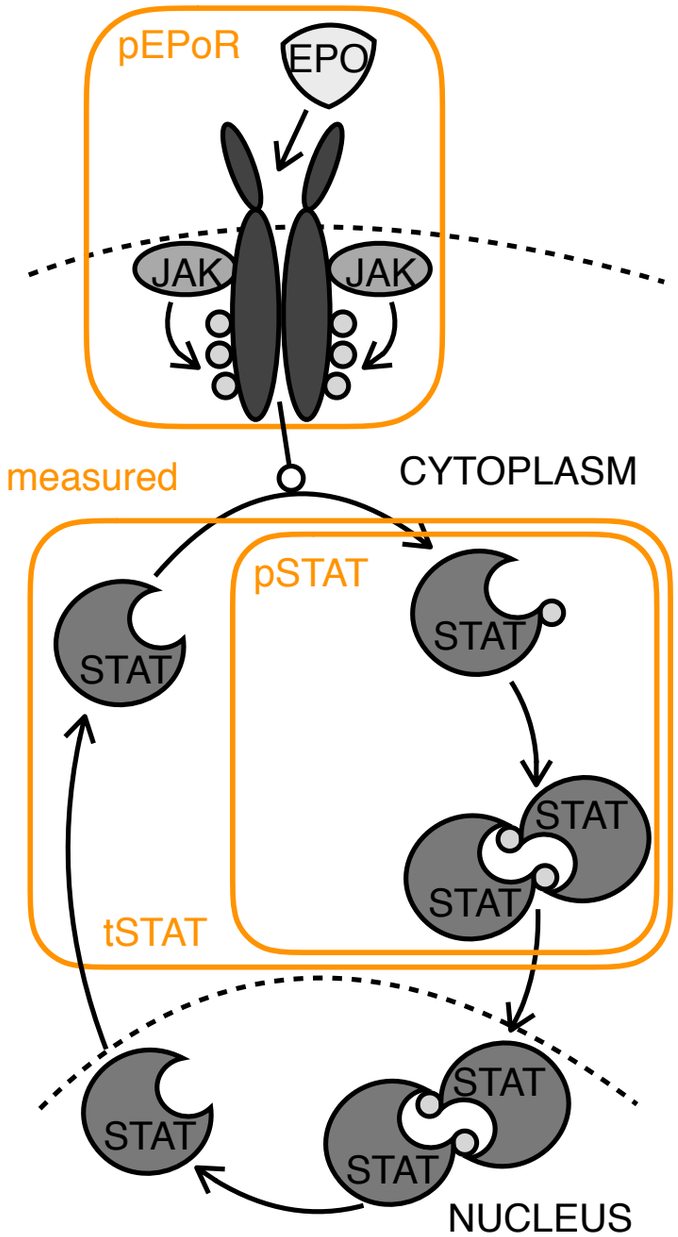
Improved robustness to outliers



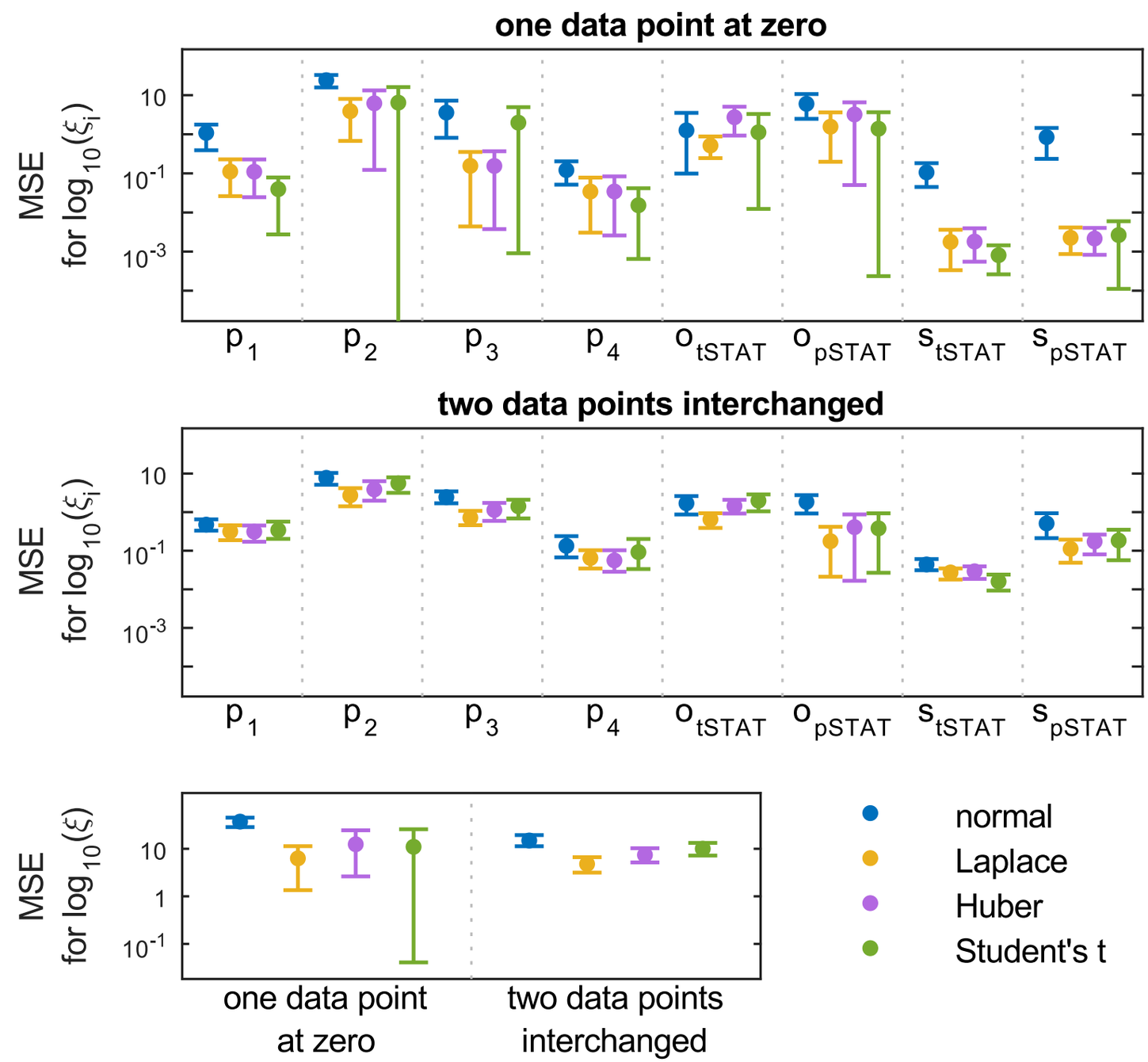
Improved robustness to outliers



Evaluation

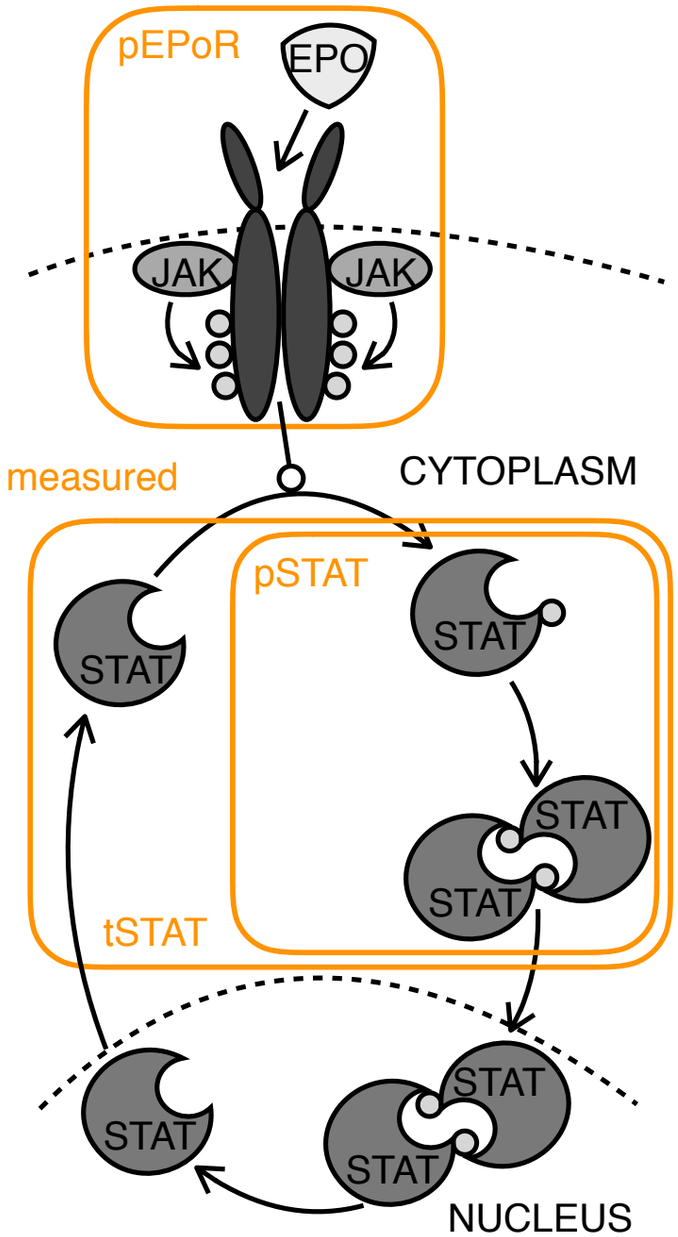


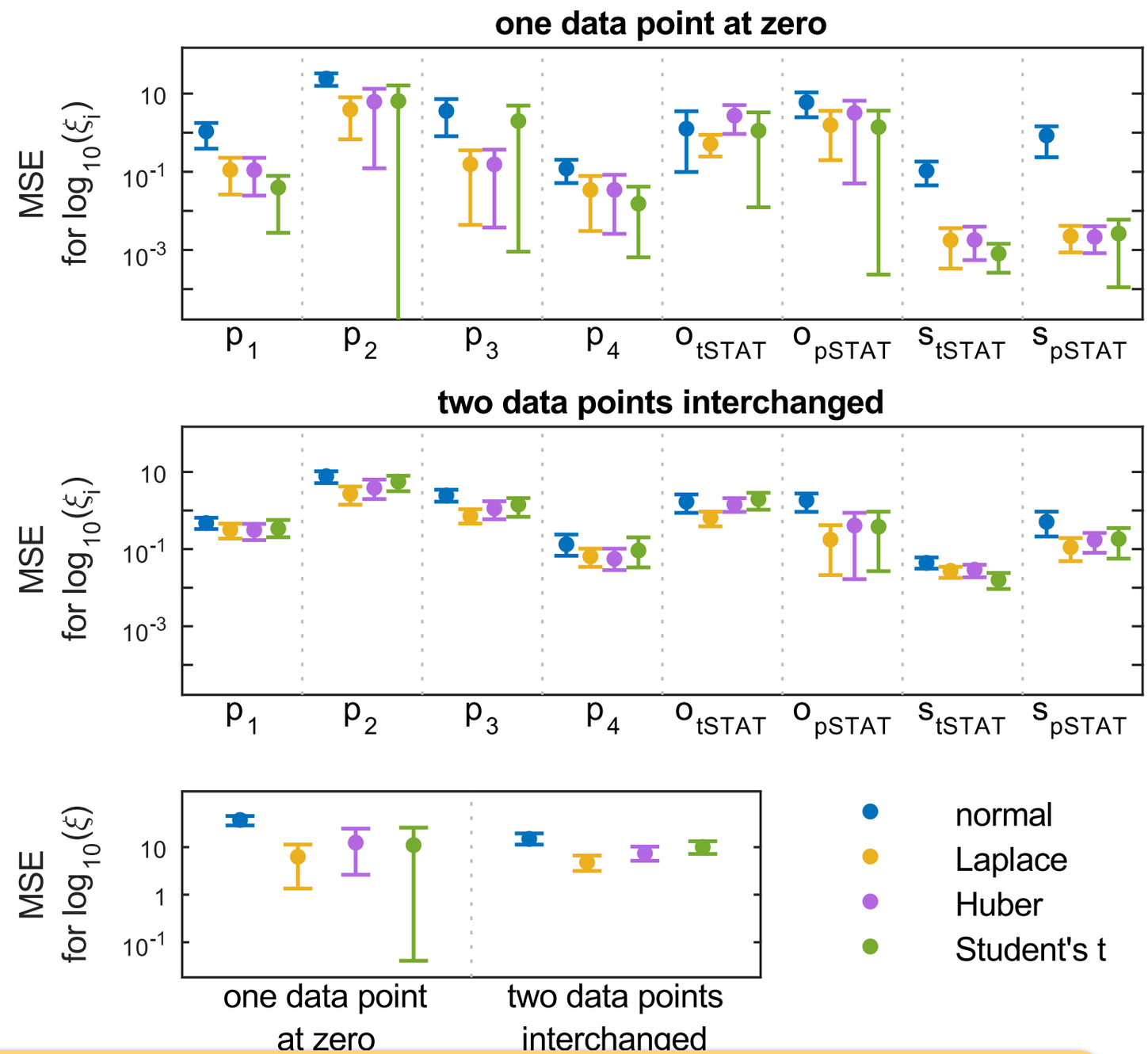
Raue et al., *Bioinformatics*, 2009



C. Maier, C. Loos, and J. Hasenauer. Robust parameter estimation for dynamical systems from outlier-corrupted data. Bioinformatics, 33(5):718–725, 2017.

Evaluation





Parameter estimation using Laplace and Huber distribution outperforms alternative methods.

C. M

Bioinformatics, 33(5):718–725, 2017.

Literature

Generic methods

- P.J. Huber et al. Robust estimation of a location parameter. The Annals of Mathematical Statistics, 35:73-101, 1964.
- K.L. Lange et al. Robust statistical modeling using the t distribution. Journal of American Statistical Association, 84, 881–896, 1989.
- D. Peel, and G. J. McLachlan. Robust mixture modelling using the t distribution. Statistics and Computing, 10, 339–348, 2000.
- A. Tarantola. Inverse Problem Theory and Methods for Model
- Parameter Estimation. SIAM, Philadelphia. 2005.

Application in systems biology

- C. Maier, C. Loos, and J. Hasenauer. Robust parameter estimation for dynamical systems from outlier-corrupted data. Bioinformatics, 33(5):718–725, 2017.

How well are the parameters determined?

The “vehicle problem” car freaks face

Task: Given only a few information, find the right vehicle.

Vehicle 1
Lamborghini
Aventador lp700-4

Vehicle 2
the tractor of my uncle

Vehicle 3
mini cooper

Dataset 1

0 - 100 km/h: 2.7 sec

Dataset 2

4.2 l / 100 km

Dataset 3

4 wheels

Problem: Data are often not sufficient to determine the unknowns uniquely.
⇒ identifiability problems and uncertainties

Resolving a crime

Task: Find the murderer.

Data:
- dead body
- fingerprints
- foodprints
- DNA
- ...

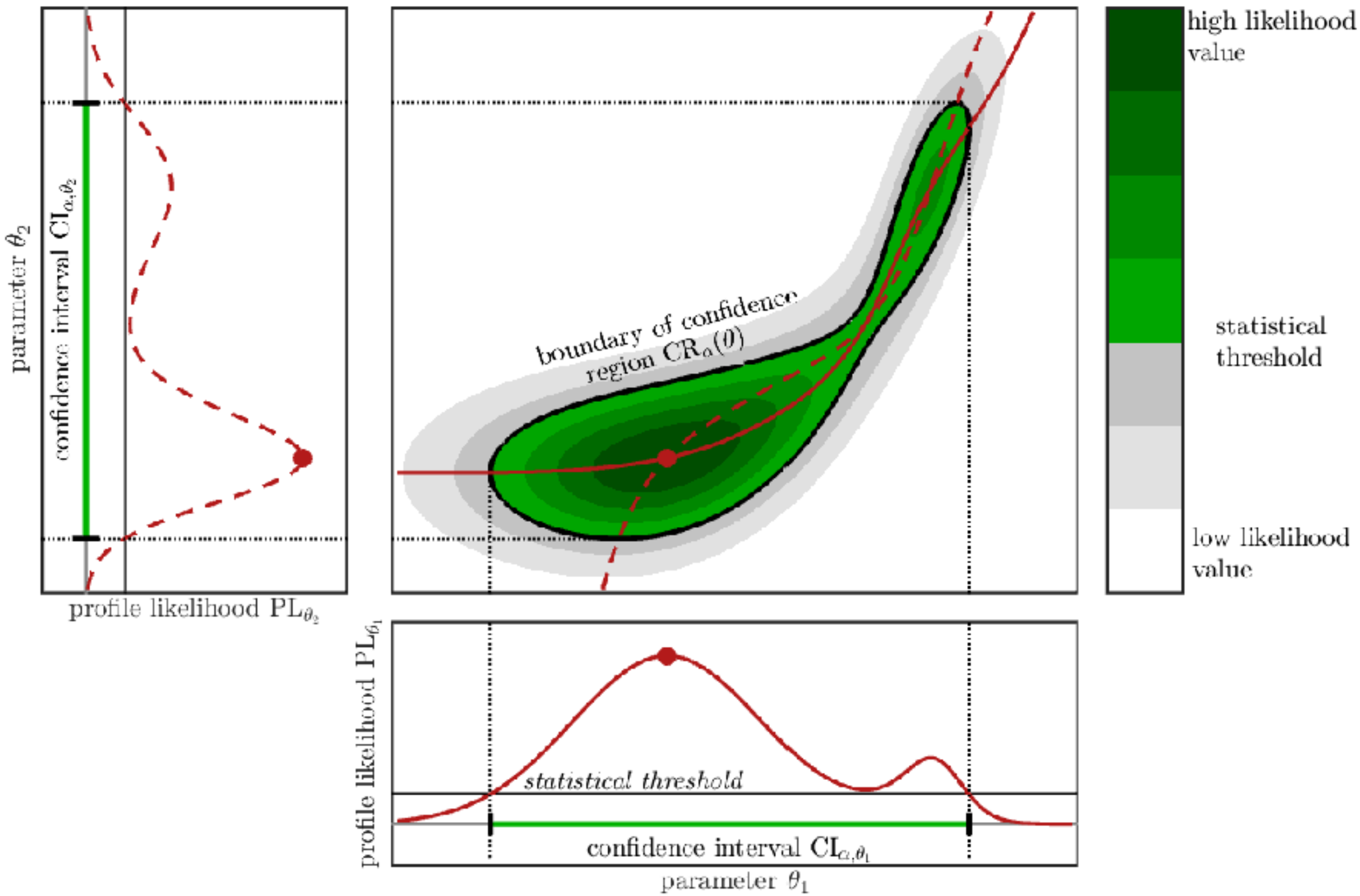
Prior information:
- regular suspects
- ...

Estimator
(clever detectives)

Uncertainty / Identifiability analysis
(judge)

Problem of general importance
(8 million viewers)!

Confidence interval and confidence region



Practical identifiability

Practical identifiability
A model property $g(\theta)$ is called practically identifiable if its confidence interval $\mathrm{CI}_{\alpha,g(\theta)}$ is bounded; otherwise it is called practically non-identifiable.

In practices also there definitions are used:

- A parameter θ_i is practically identifiable if $\mathrm{CI}_{\alpha,\theta_i}$ is in the interior of Θ.
- A parameter θ_i is practically identifiable if $\mathrm{CI}_{\alpha,\theta_i}$ does not contain zero.

Remark 1: The practical identifiability depends on the confidence level α.

Remark 2: practical identifiability $\underset{\not\Leftarrow}{\Rightarrow}$ structural identifiability

How well are the parameters determined?
- Profile likelihood calculation

Model-based data integration: Uncertainty analysis

Profile likelihoods

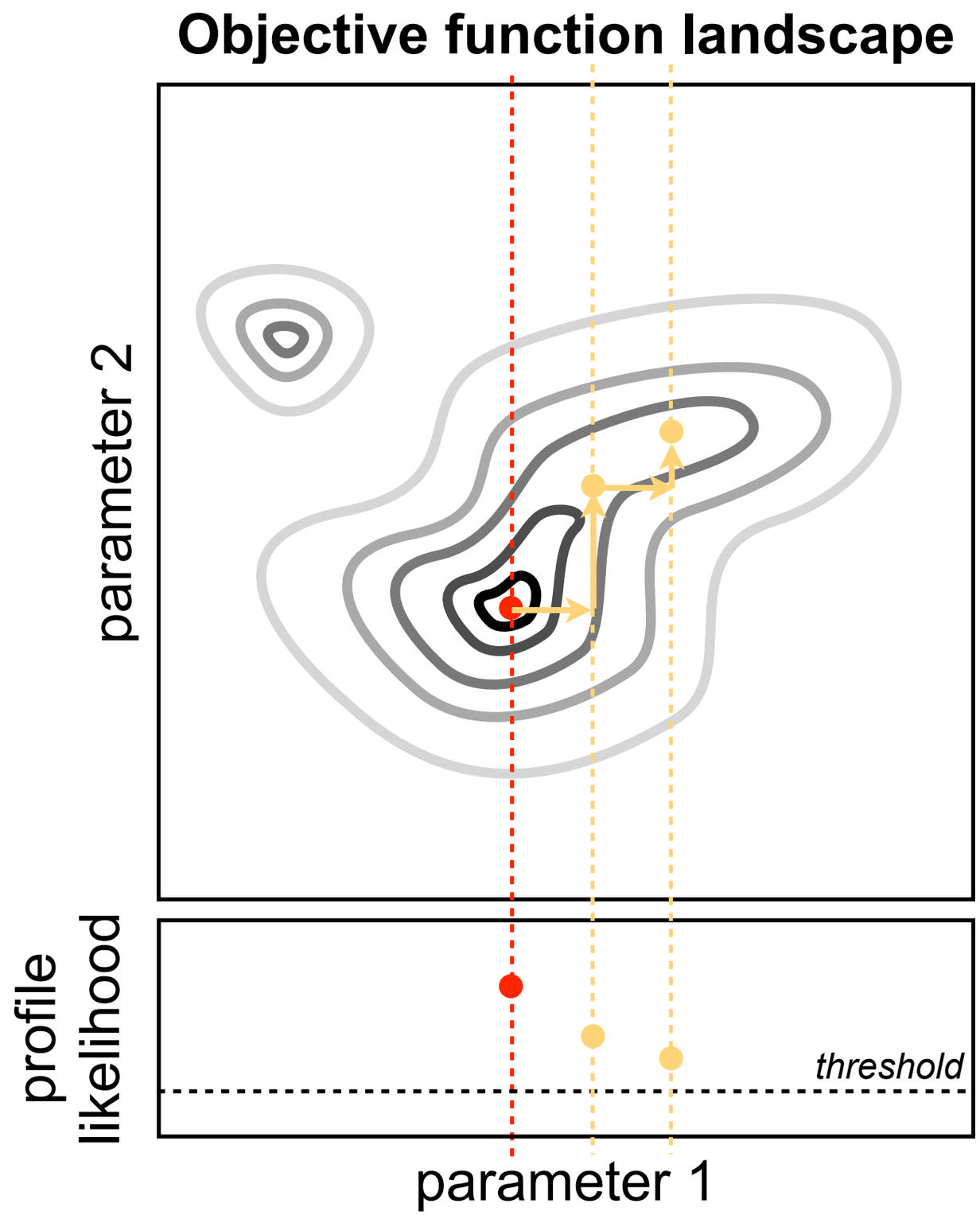
are computed by repeated local optimisation

$$\mathrm{PL}_{\theta_i}(c) = \max_{\substack{\theta_j, j \neq i \\ \theta_i = c}} p(\mathcal{D}|\theta)$$

Confidence intervals

are derived from the likelihood ratio (test)

$$\frac{\mathrm{PL}_{\theta_i}(c)}{p(\mathcal{D}|\theta^{\mathrm{ml}})}$$



Model-based data integration: Uncertainty analysis

Profile likelihoods

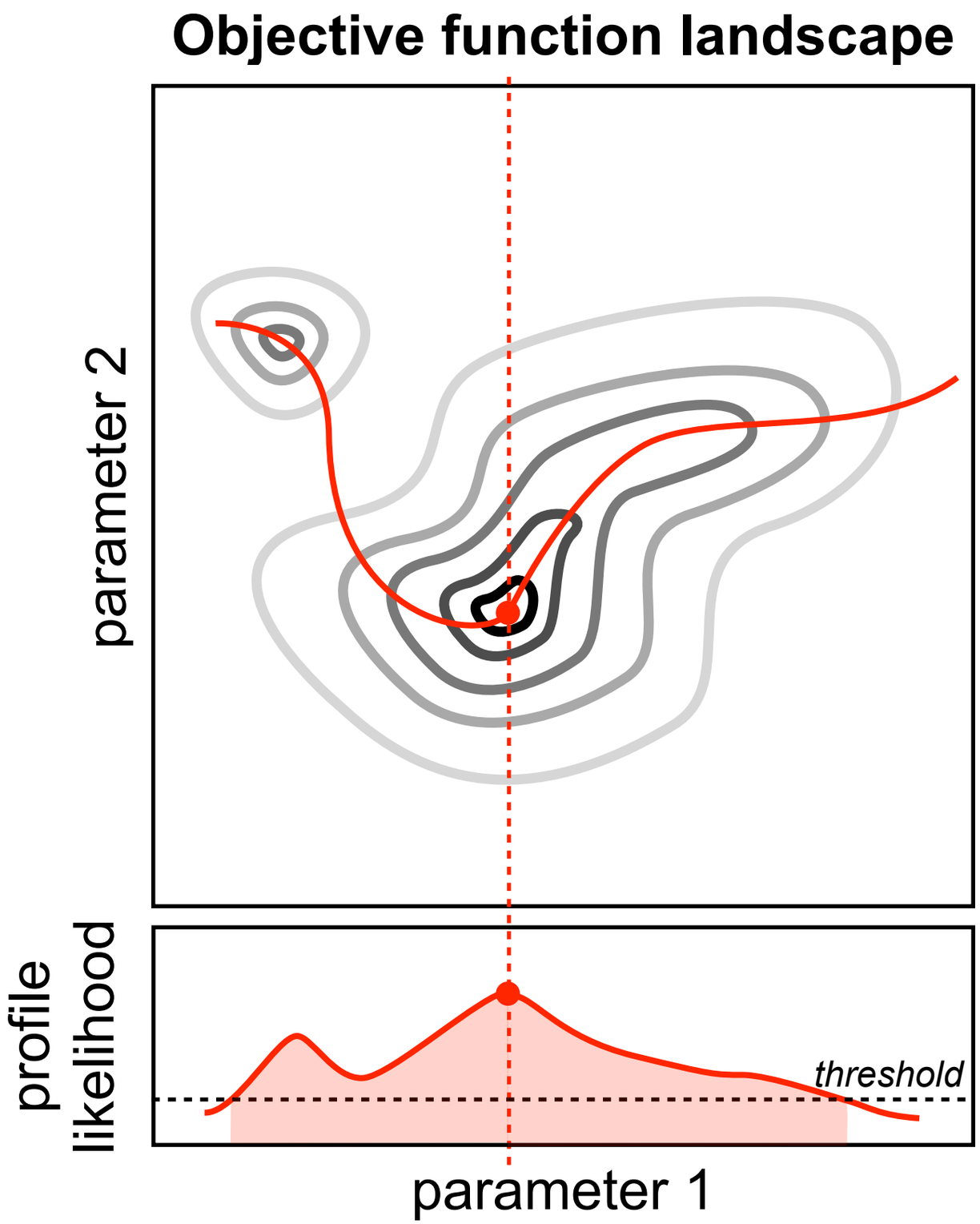
are computed by repeated local optimisation

$$\mathrm{PL}_{\theta_i}(c) = \max_{\substack{\theta_j, j \neq i \\ \theta_i = c}} p(\mathcal{D}|\theta)$$

Confidence intervals

are derived from the likelihood ratio (test)

$$\frac{\mathrm{PL}_{\theta_i}(c)}{p(\mathcal{D}|\theta^{\mathrm{ml}})}$$



Model-based data integration: Uncertainty analysis

Profile likelihoods

are computed by repeated local optimisation

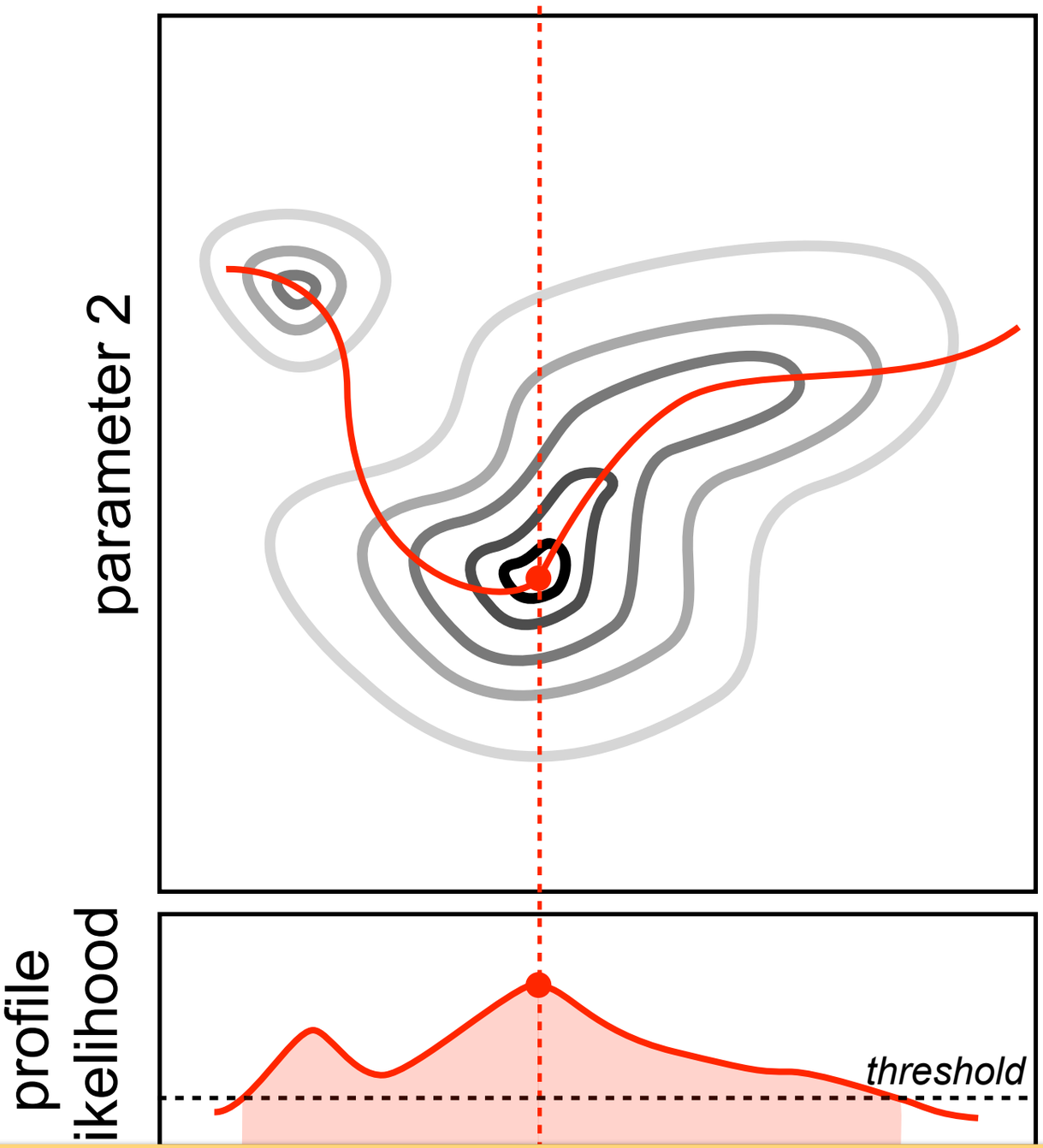
$$\mathrm{PL}_{\theta_i}(c) = \max_{\substack{\theta_j, j \neq i \\ \theta_i = c}} p(\mathcal{D}|\theta)$$

Confidence intervals

are derived from the likelihood ratio (test)

$$\frac{\mathrm{PL}_{\theta_i}(c)}{p(\mathcal{D}|\theta^{\mathrm{ml}})}$$

Objective function landscape



Efficient optimisation-based uncertainty analysis approach.

Model-based data integration: Uncertainty analysis

Profile likelihoods

are computed by repeated local optimisation

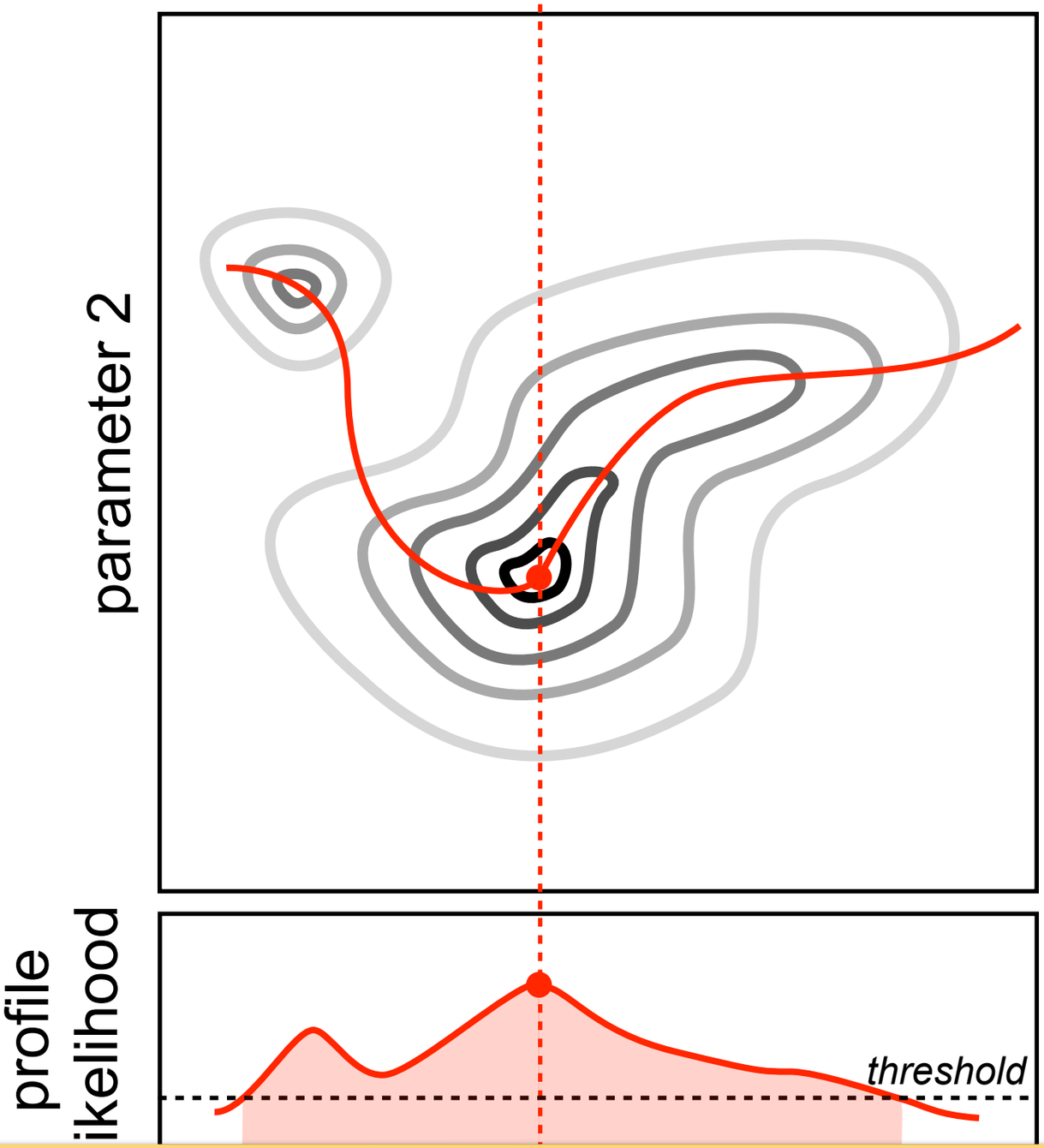
$$\mathrm{PL}_{\theta_i}(c) = \max_{\substack{\theta_j, j \neq i \\ \theta_i = c}} p(\mathcal{D}|\theta)$$

Confidence intervals

are derived from the likelihood ratio (test)

$$\frac{\mathrm{PL}_{\theta_i}(c)}{p(\mathcal{D}|\theta^{\mathrm{ml}})}$$

Objective function landscape

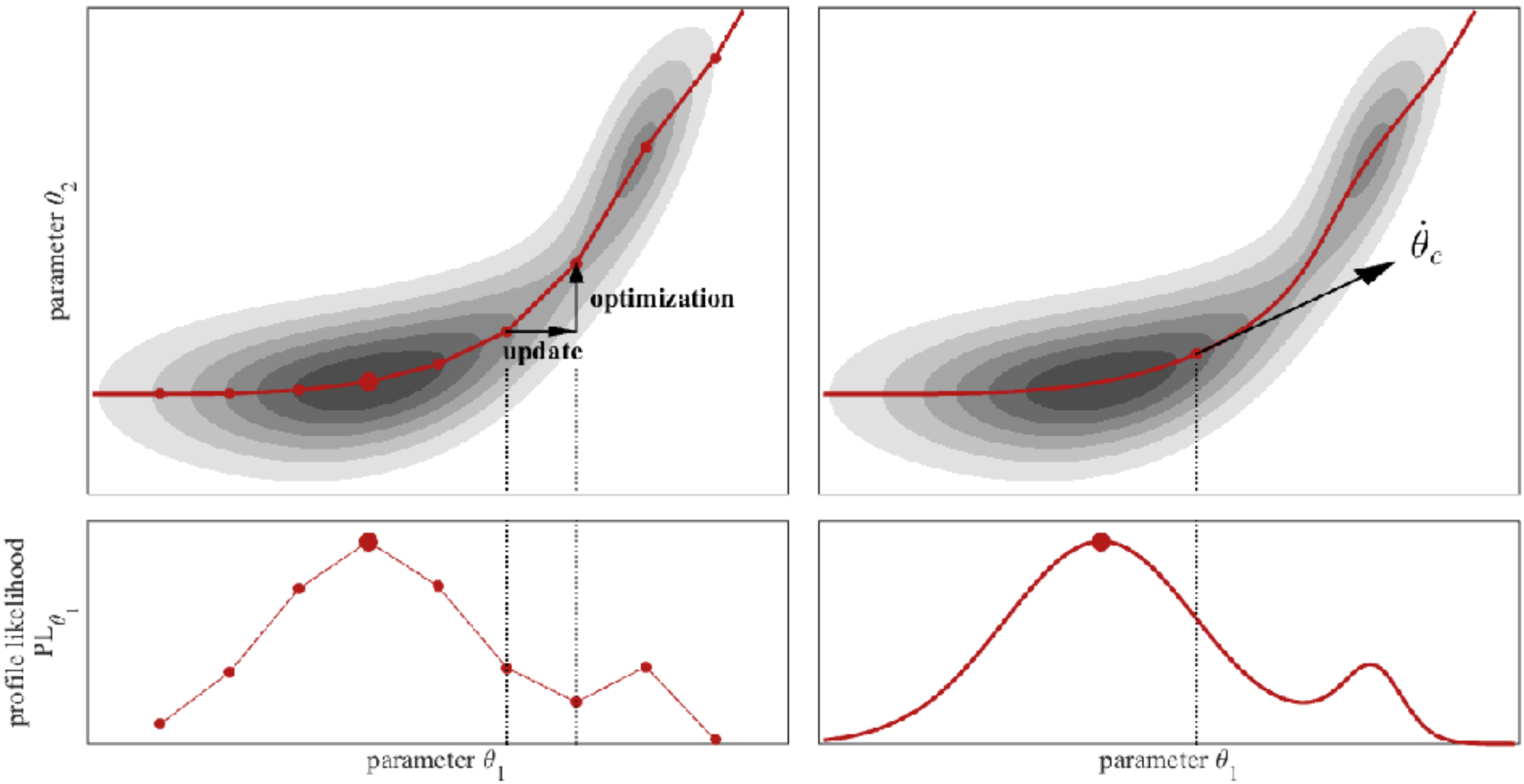


All model properties have to be analysed separately.

Methods for profile calculation

optimisation-based profile calculation

integration-based profile calculation



How well are the parameters determined?
- Sampling approaches

Sampling methods

Bayes' theorem

likelihood

prior

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} \propto p(\mathcal{D}|\theta)p(\theta)$$

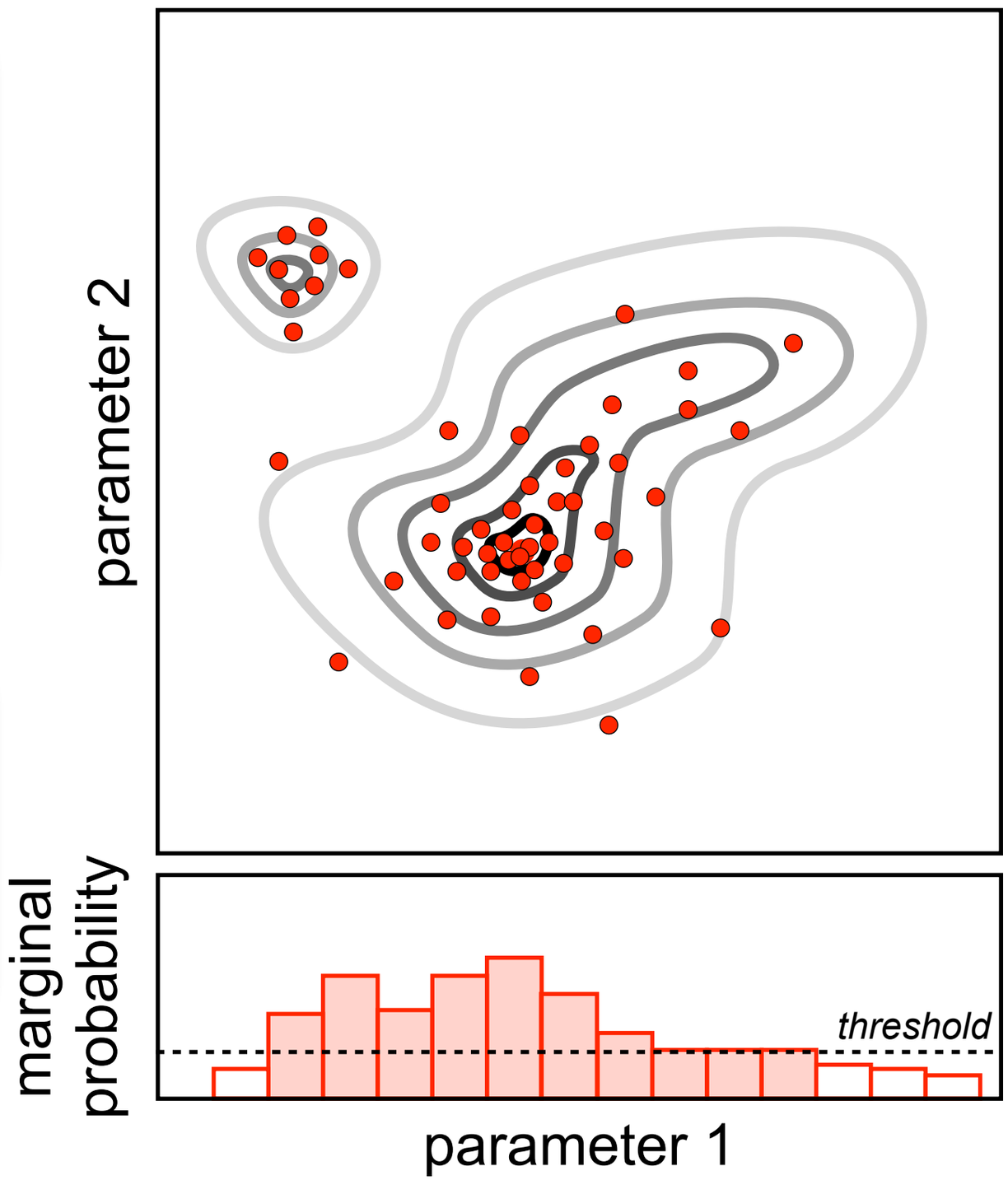
posterior

marginal

Sampling of (unnormalised) posterior distribution using

using Markov chain Monte-Carlo sampling.

Objective function landscape



Sampling methods

Bayes' theorem

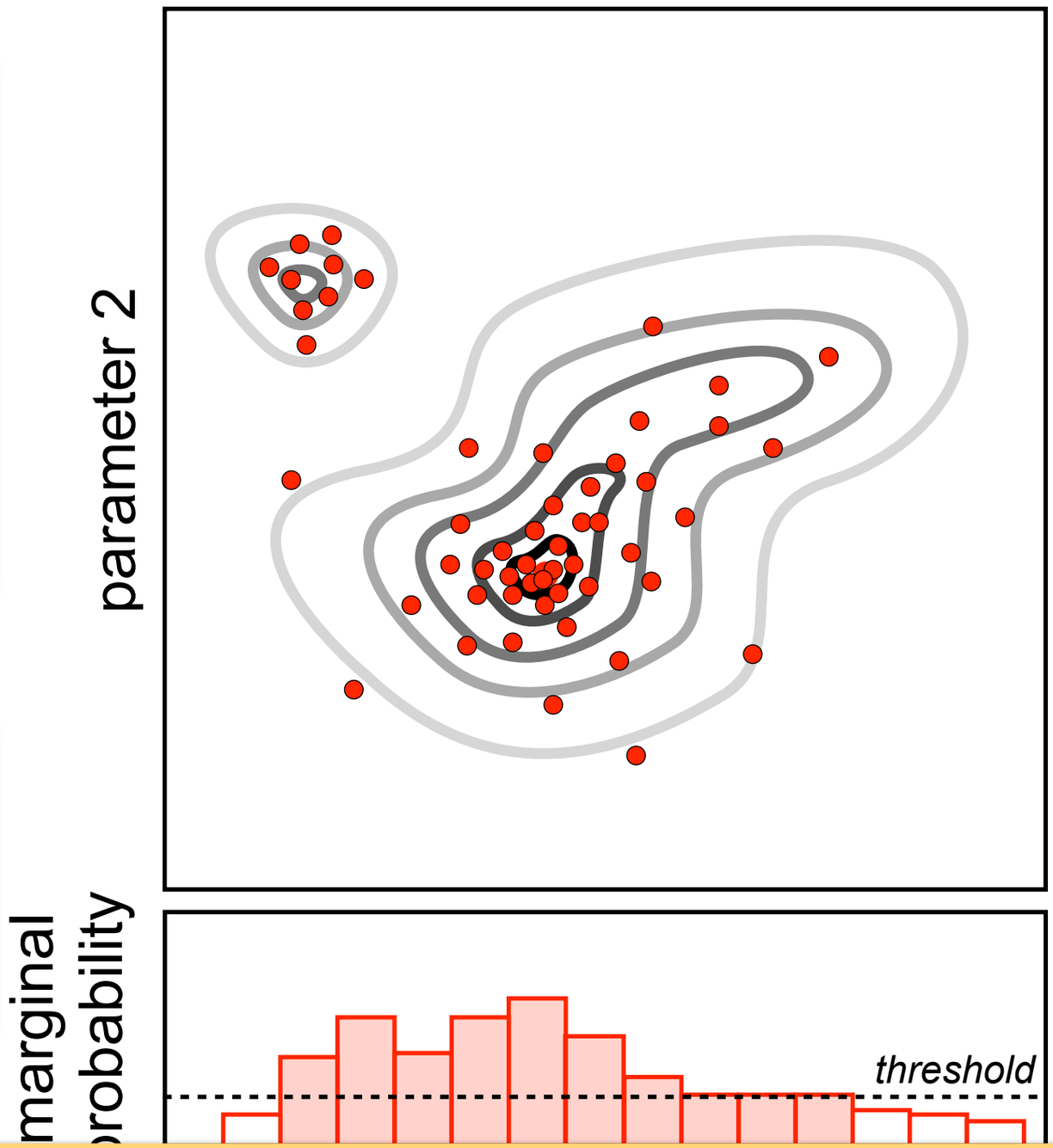
$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} \propto p(\mathcal{D}|\theta)p(\theta)$$

likelihood *prior* *posterior* *marginal*

Sampling of (unnormalised) posterior distribution using

using Markov chain Monte-Carlo sampling.

Objective function landscape



Sampling-based uncertainty analysis approach.

Sampling methods

Bayes' theorem

likelihood

prior

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} \propto p(\mathcal{D}|\theta)p(\theta)$$

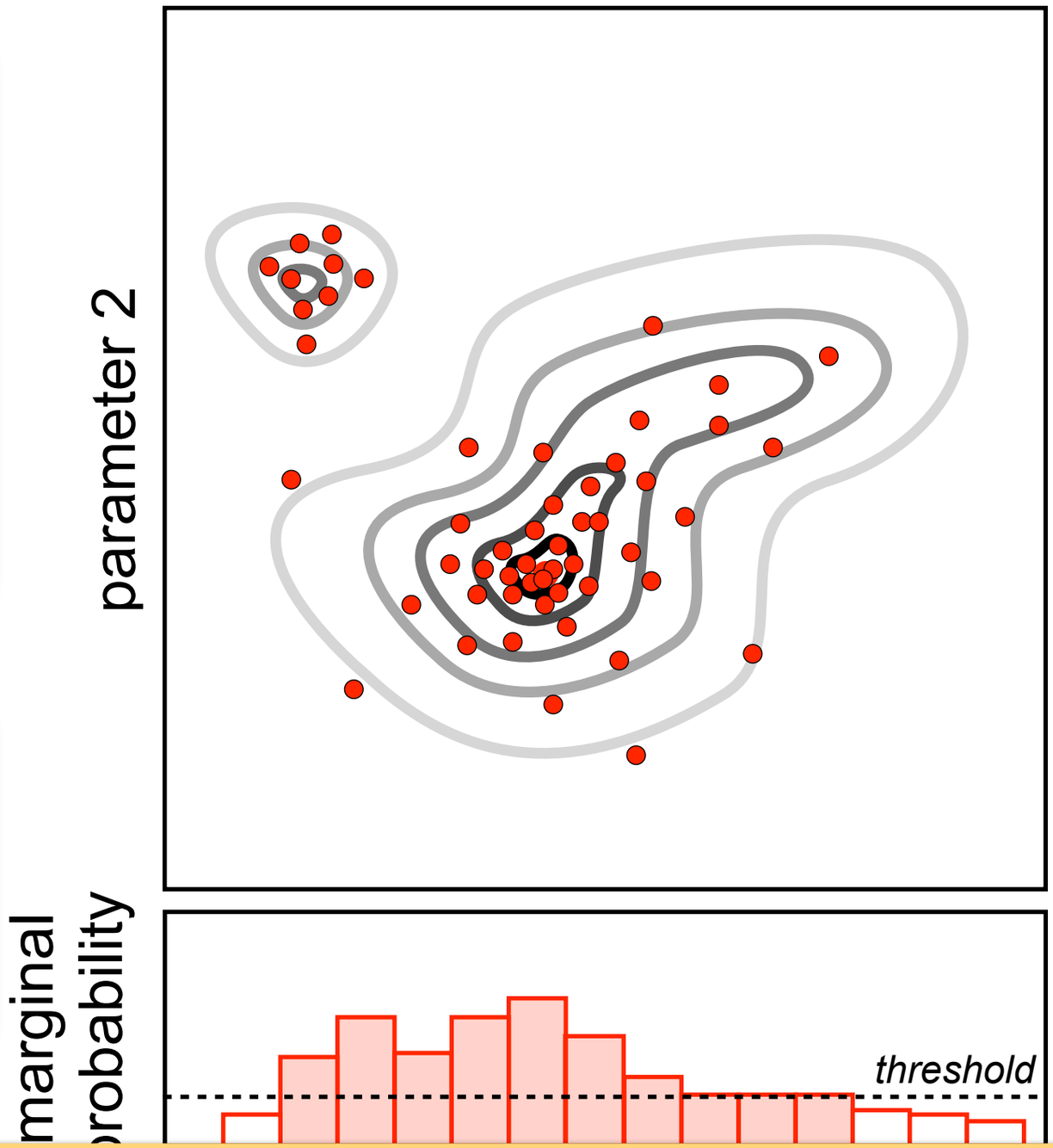
posterior

marginal

Sampling of (unnormalised) posterior distribution using

using Markov chain Monte-Carlo sampling.

Objective function landscape



Holistic approach but computationally demanding.

Single-chain sampling methods

- Metropolis-Hastings
- adaptive Metropolis
- Riemann manifold Monte Carlo
- delayed rejection Metropolis
- Hamiltonian Monte Carlo

Algorithm

start at random $\theta^{(0)}$
repeat
    generate sample, $\theta' \sim q(\theta|\mu^{(k)},\Sigma^{(k)})$

    accept it with probability $r = \frac{p(\theta'|D)}{p(\theta^{(k)}|D)} \frac{q(\theta^{(k)}|\mu^{(k+1)},\Sigma^{(k+1)})}{q(\theta'|\mu^{(k)},\Sigma^{(k)})}$

    if rand < r
        set $\theta^{(k+1)} = \theta'$
    else
        set $\theta^{(k+1)} = \theta^{(k)}$
    end
until number of desired samples reached

Multi-chain sampling methods

- parallel hierarchical sampling
- parallel tempering

Parallel hierarchical sampling

Idea: Multiple chains with different proposals for better mixing

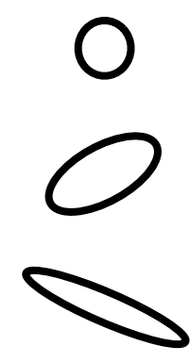
Mother chain: $\theta^{(0,0)}, \theta^{(0,1)}, \theta^{(0,2)}, \theta^{(0,3)}, \theta^{(0,4)}, \theta^{(0,5)} \ldots$

Auxiliary chains: $\theta^{(1,0)}, \theta^{(1,1)}, \theta^{(1,2)}, \theta^{(1,3)}, \theta^{(1,4)}, \theta^{(1,5)} \ldots$

$\theta^{(2,0)}, \theta^{(2,1)}, \theta^{(2,2)}, \theta^{(2,3)}, \theta^{(2,4)}, \theta^{(2,5)} \ldots$

$\theta^{(3,0)}, \theta^{(3,1)}, \theta^{(3,2)}, \theta^{(3,3)}, \theta^{(3,4)}, \theta^{(3,5)} \ldots$

$\ldots$



Note: For the “in-chain-sampling” any single-chain sample can be used.

Multi-chain sampling methods

- parallel hierarchical sampling
- parallel tempering

Parallel tempering

Idea: Sample simpler (tempered) distributions

$$p_t(\theta|D) = c\ p(\theta|D)^t\ p(\theta)$$

and use them to bridge different parts of the posterior.

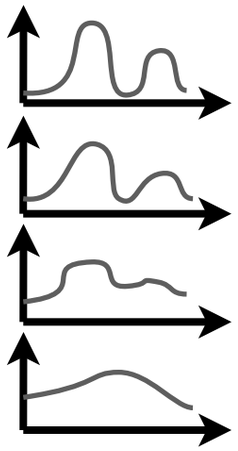
t = 1: $\theta^{(0,0)}$, $\theta^{(0,1)}$, $\theta^{(0,2)}$, $\theta^{(0,3)}$, $\theta^{(0,4)}$, $\theta^{(0,5)}$ ...

t = 0.5: $\theta^{(1,0)}$, $\theta^{(1,1)}$, $\theta^{(1,2)}$, $\theta^{(1,3)}$, $\theta^{(1,4)}$, $\theta^{(1,5)}$ ...

t = 0.25: $\theta^{(2,0)}$, $\theta^{(2,1)}$, $\theta^{(2,2)}$, $\theta^{(2,3)}$, $\theta^{(2,4)}$, $\theta^{(2,5)}$ ...

t = 0.1: $\theta^{(3,0)}$, $\theta^{(3,1)}$, $\theta^{(3,2)}$, $\theta^{(3,3)}$, $\theta^{(3,4)}$, $\theta^{(3,5)}$ ...

...



Note: For the “in-chain-sampling” any single-chain sample can be used.

Multi-chain sampling methods

- parallel hierarchical sampling
- parallel tempering

Parallel tempering

Idea: Sample simpler (tempered) distributions

$$p_t(\theta|D) = c\ p(\theta|D)^t\ p(\theta)$$

and use them to bridge different parts of the posterior.

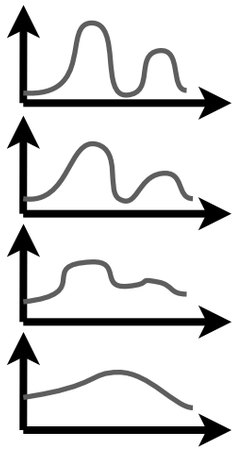
t = 1: $\theta^{(0,0)}$, $\theta^{(0,1)}$, $\theta^{(0,2)}$, $\theta^{(0,3)}$, $\theta^{(0,4)}$, $\theta^{(0,5)}$ ...

t = 0.5: $\theta^{(1,0)}$, $\theta^{(1,1)}$, $\theta^{(1,2)}$, $\theta^{(1,3)}$, $\theta^{(1,4)}$, $\theta^{(1,5)}$ ...

t = 0.25: $\theta^{(2,0)}$, $\theta^{(2,1)}$, $\theta^{(2,2)}$, $\theta^{(2,3)}$, $\theta^{(2,4)}$, $\theta^{(2,5)}$ ...

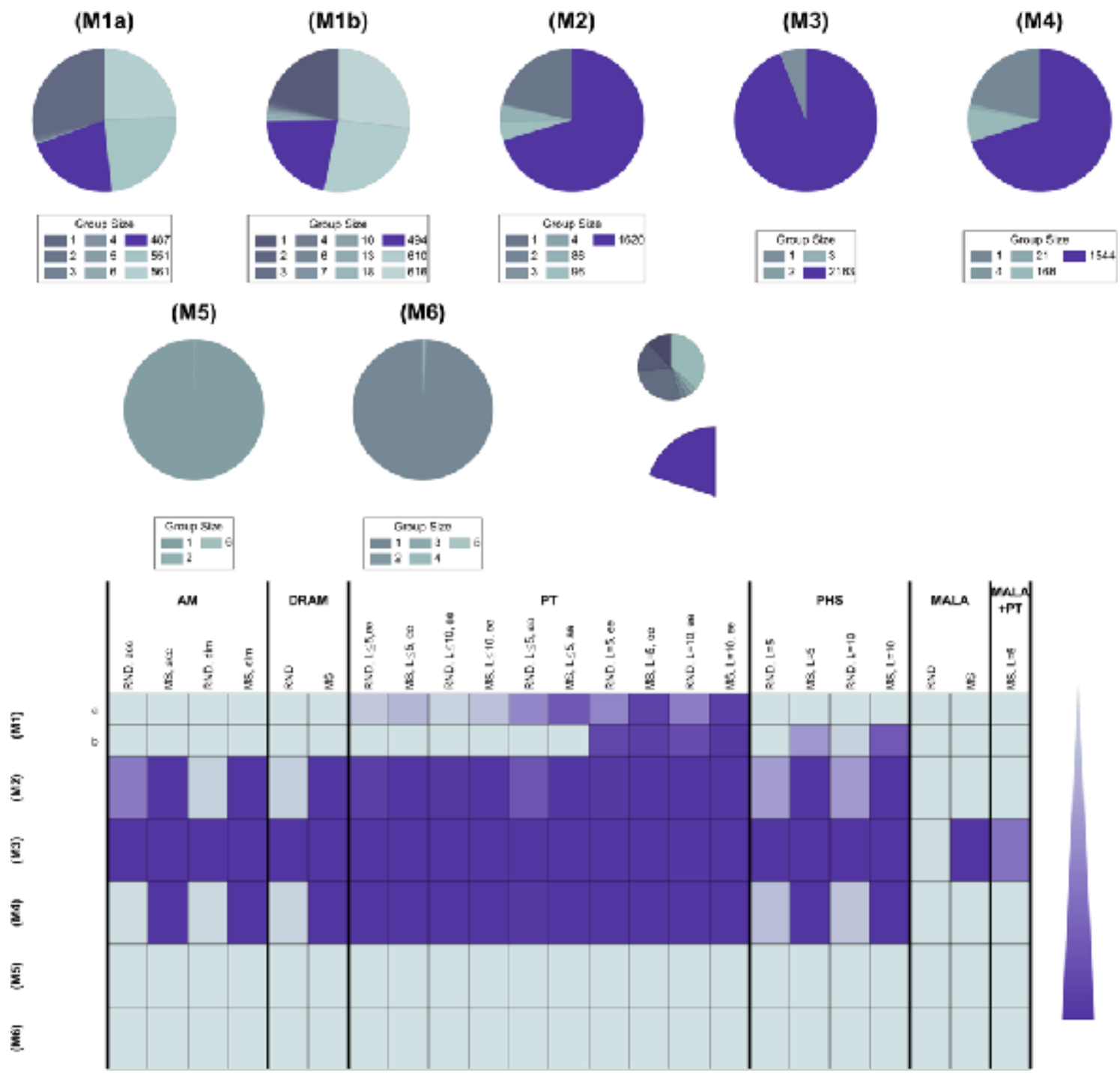
t = 0.1: $\theta^{(3,0)}$, $\theta^{(3,1)}$, $\theta^{(3,2)}$, $\theta^{(3,3)}$, $\theta^{(3,4)}$, $\theta^{(3,5)}$ ...

...



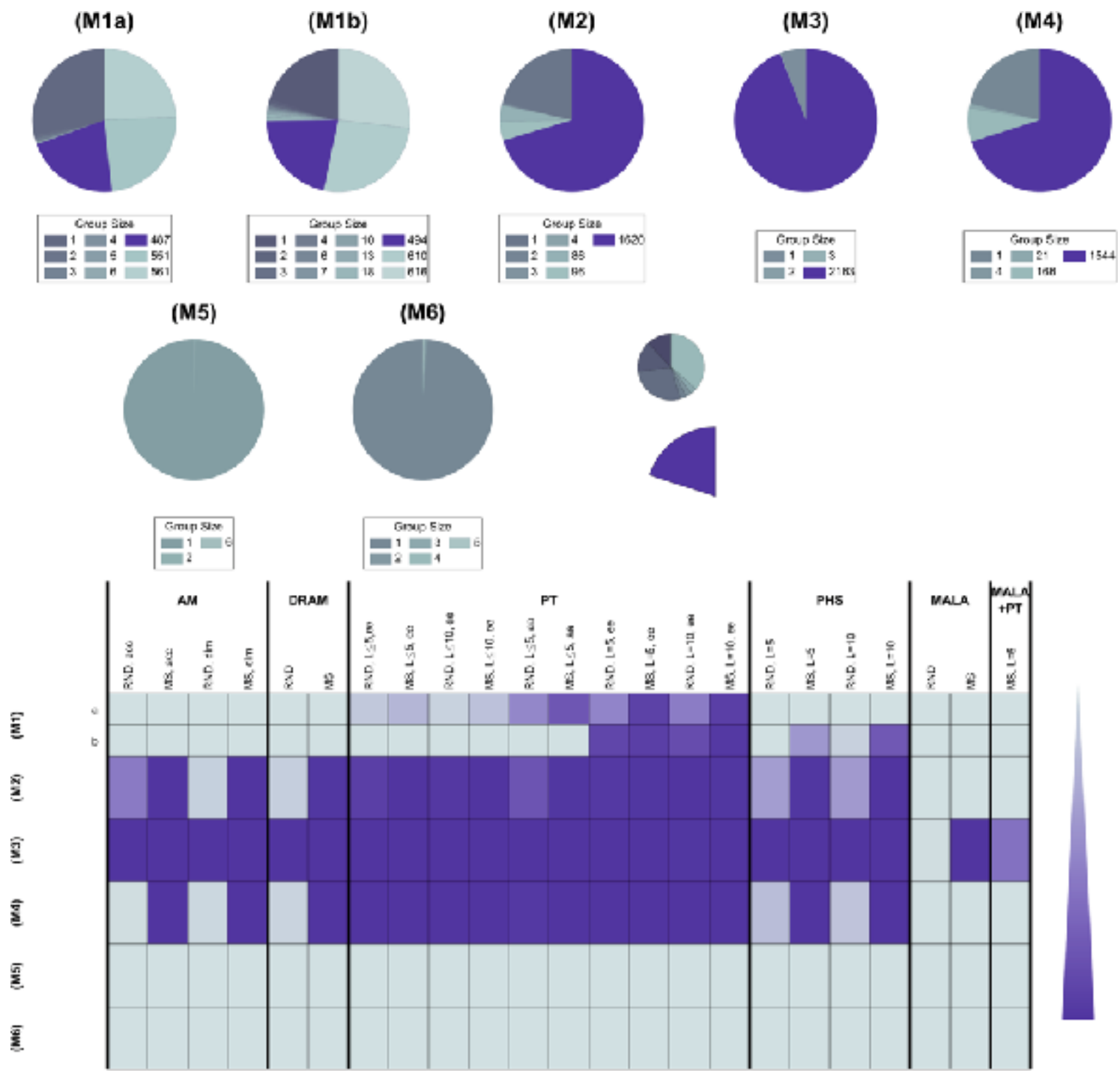
Multi-chain approaches can easily exploit parallel architectures.

Benchmarking of methods



Ballnus et al. Comprehensive benchmarking of Markov chain Monte Carlo methods for dynamical systems. BMC Systems Biology, 11(63), 2017.

Benchmarking of methods



Check the constancy and reproducibility of the results!

Literature

Profile likelihood

- A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics, 25(25):1923-1929, 2009.
- H. Hass, C. Kreutz, J. Timmer, and D. Kaschek. Fast integration-based prediction bands for ordinary differential equation models. Bioinformatics, 32(8):1204-1210, 2016.
- P. Stapor, F. Fröhlich, and J. Hasenauer. Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis. Bioinformatics, 34(13):i151-i159, 2018.

Sampling

- H. Haario, M. Laine, A. Mira, and E. Saksman. DRAM: Efficient adaptive MCMC. Statistics and Computing, 16(4):339-354, 2006.
- T.-R. Xu et al. Inferring signaling pathway topologies from multiple perturbation measurements of specific biochemical species. Science Signalling, 3(113):ra20, 2010.
- M. Girolami, and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods Journal of the Royal Statistical Society B, 73(2):123-214, 2011.
- S. Hug, A. Raue, J. Hasenauer, J. Bachmann, U. Klingmüller, J. Timmer, and F. J. Theis. High-dimensional Bayesian parameter estimation: Case study for a model of JAK2/STAT5 signaling. Mathematical Biosciences, 246(2):293-304, 2013.
- Ballnus et al. Comprehensive benchmarking of Markov chain Monte Carlo methods for dynamical systems. BMC Systems Biology, 11(63), 2017.

Literature

Profile likelihood

- A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics, 25(25):1923-1929, 2009.
- H. Hass, C. Kreutz, J. Timmer, and D. Kaschek. Fast integration-based prediction bands for ordinary differential equation models. Bioinformatics, 32(8):1204-1210, 2016.
- P. Stapor, F. Fröhlich, and J. Hasenauer. Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis. Bioinformatics, 34(13):i151-i159, 2018.

Sampling

- H. Haario, M. Laine, A. Mira, and E. Saksman. DRAM: Efficient adaptive MCMC. Statistics and Computing, 16(4):339-354, 2006.
- T.-R. Xu et al. Inferring signaling pathway topologies from multiple perturbation measurements of specific biochemical species. Science Signalling, 3(113):ra20, 2010.
- M. Girolami, and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods Journal of the Royal Statistical Society B, 73(2):123-214, 2011.
- S. Hug, A. Raue, J. Hasenauer, J. Bachmann, U. Klingmüller, J. Timmer, and F. J. Theis. High-dimensional Bayesian parameter estimation: Case study for a model of JAK2/STAT5 signaling. Mathematical Biosciences, 246(2):293-304, 2013.

Scalability of uncertainty analysis methods is still very limited and limiting!

What tools are available?

Software tools

COmplex PAthway SImulator (COPASI)

- Simple handling of model, experimental data and experimental conditions
- Parameter optimisation

Data2Dynamics (D2D)

- Simple handling of model, experimental data and experimental conditions
- Highly automised
- Multi-experiment fitting for ODEs
- Parameter optimisation and uncertainty analysis

Advanced MATLAB Interface for CVODE and IDAS (AMICI)

- 1st and 2nd order forward sensitivity analysis
- 1st order adjoints sensitivity analysis
- Support for state and output events

Parameter EStimation TOolbox (PESTO & pyPESTO)

- Parameter optimisation
- Uncertainty analysis
- Flexible interface

MEtaheuristics for systems biology and bIoinformatics Global Opt. (MEIGO)

- Flexible toolbox for global optimisation
- Support of integer variables
- Deterministic methods

Software tools

COmplex PAthway SImulator (COPASI)

- Simple handling of model, experimental data and experimental conditions
- Parameter optimisation

Data2Dynamics (D2D)

- Simple handling of model, experimental data and experimental conditions
- Highly automised
- Multi-experiment fitting for ODEs
- Parameter optimisation and uncertainty analysis

Advanced MATLAB Interface for CVODE and IDAS (AMICI)

- 1st and 2nd order forward sensitivity analysis
- 1st order adjoints sensitivity analysis
- Support for state and output events

Parameter EStimation TOolbox (PESTO & pyPESTO)

- Parameter optimisation
- Uncertainty analysis
- Flexible interface

MEtaheuristics for systems biology and bIoinformatics Global Opt. (MEIGO)

If you miss a feature, contact the developers.

They are often motivated and helpful.

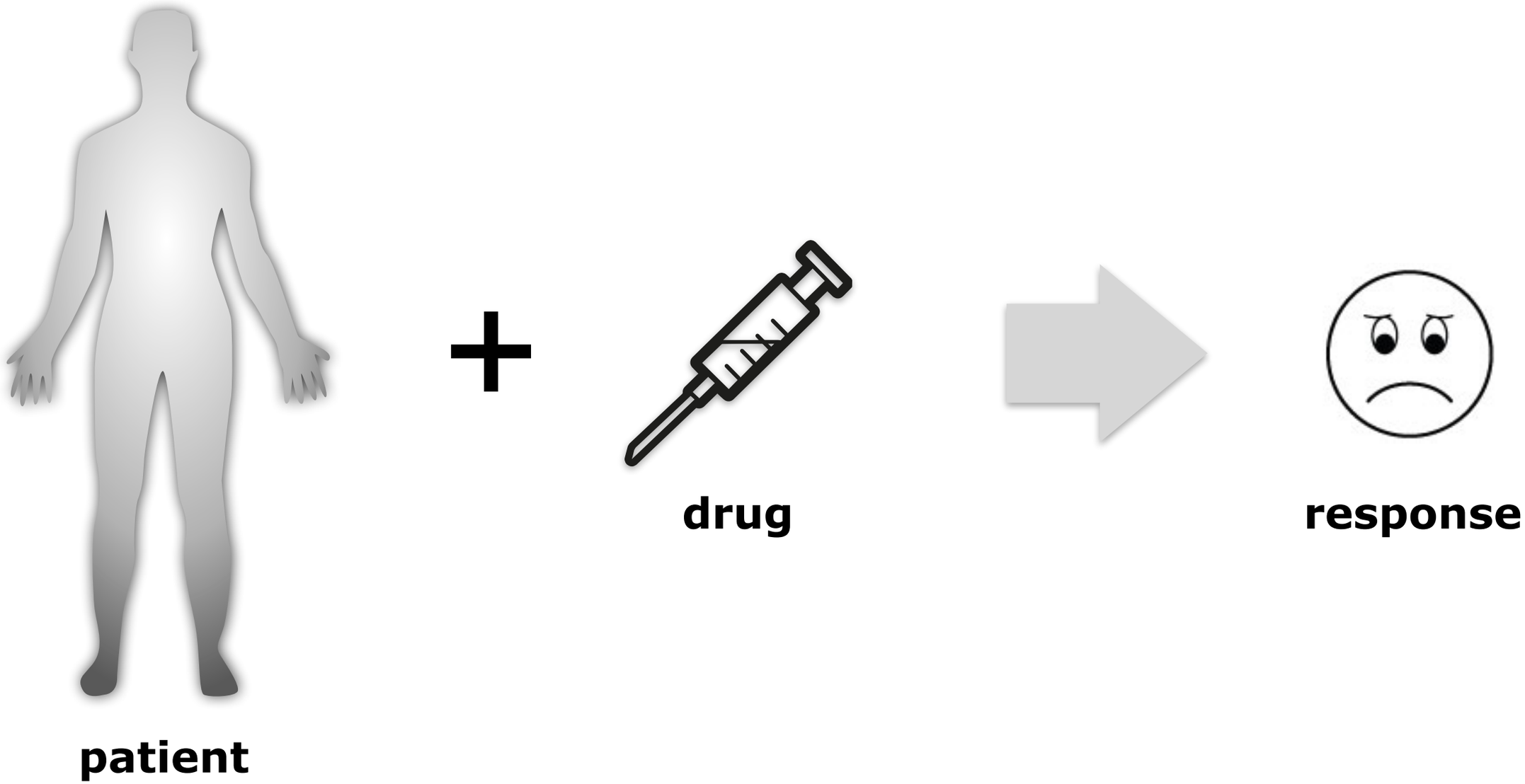
Literature

Software

- S. Hoops et al., COPASI - a COmplex PAthway SImulator. Bioinformatics, 22(24): 3067-3074, 2006.
- E. Balsa-Canto, and J. R. Banga. AMIGO, a toolbox for advanced model identification in systems biology using global optimization. Bioinformatics, 27(16):2311-2313, 2011
- J. A. Egea et al. MEIGO: An open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. BMC Bioinf., 15(136), 2014
- A. Raue et al. Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics, (21):3558-3560, 2015.
- P. Stapor et al. PESTO: Parameter EStimation TOolbox. Bioinformatics, btx676, 2017.
- F. Fröhlich et al. Parameter estimation for dynamical systems with discrete events and logical operations. Bioinformatics, 33(7):1049-1056, 2017.

A large-scale model of cancer signalling

Drug response prediction



Selection of the right drug for the right patient?

Modelling of biochemical reaction networks

Prior knowledge

KEGG STRING ChEMBL REACTOME

Experimental data

CCLE Genomics England NAKO Nationale Kohorte 23andMe TCGA

data integration

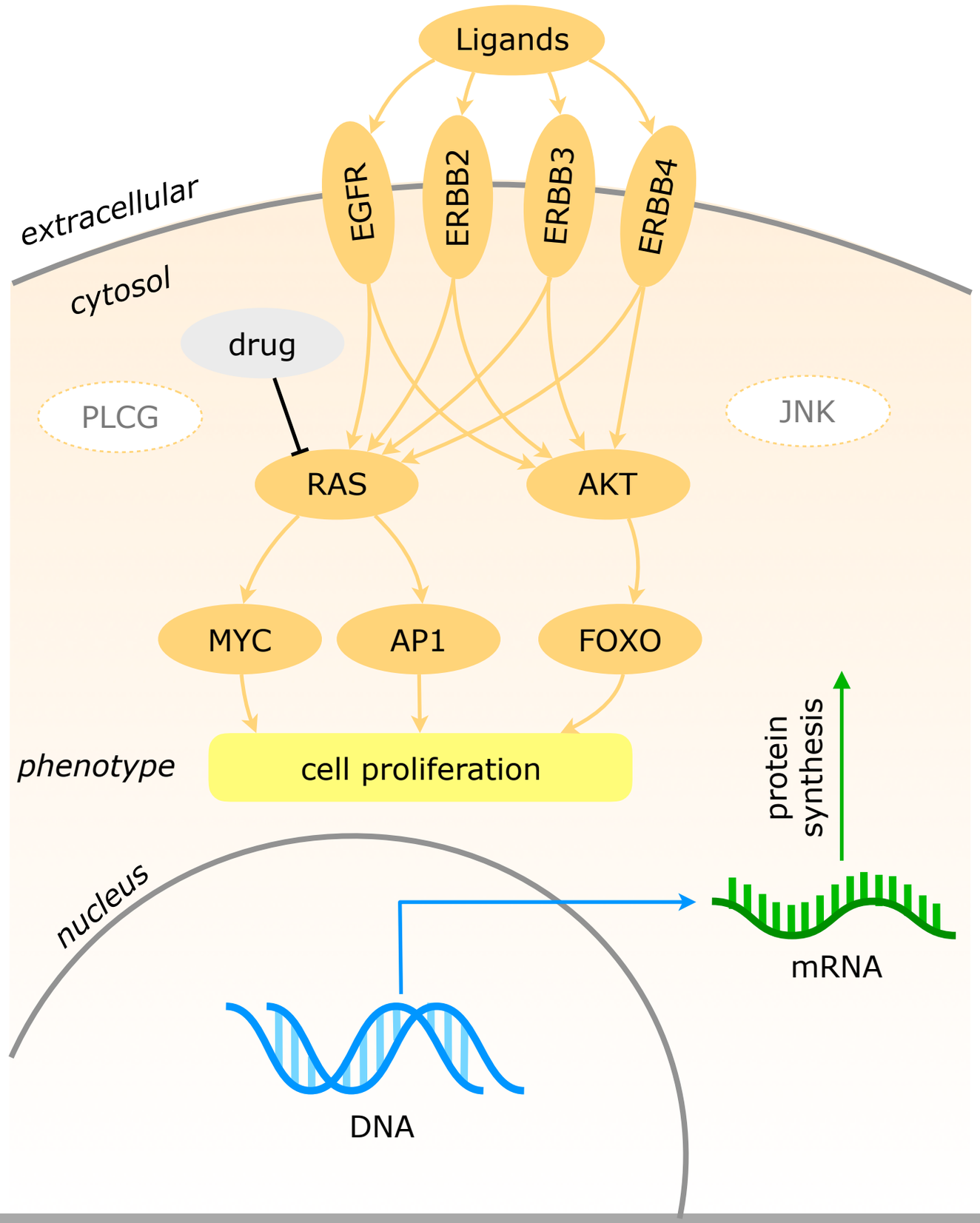
Predictive mechanistic model

$$\dot{x} = Sv(x, \theta)$$

$$y = h(x, \theta)$$

Structure of the model?

Mathematical model



Protein-protein interaction network

Mechanistic description:

KRAS:GDP ⇄ KRAS:GTP

RAF + KRAS:GTP → pRAF + KRAS:GTP

KRAS*:GDP ⇄ KRAS*:GTP

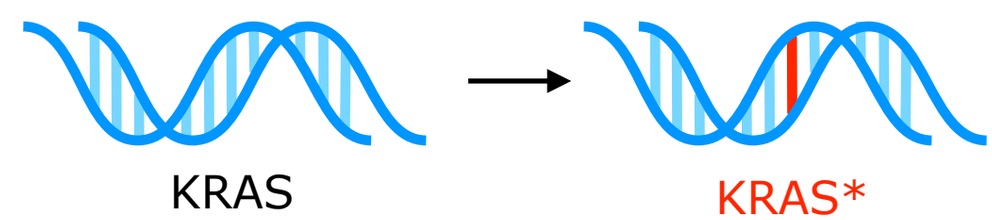
RAF + KRAS*:GTP → pRAF + KRAS*:GTP

Gene expression profile

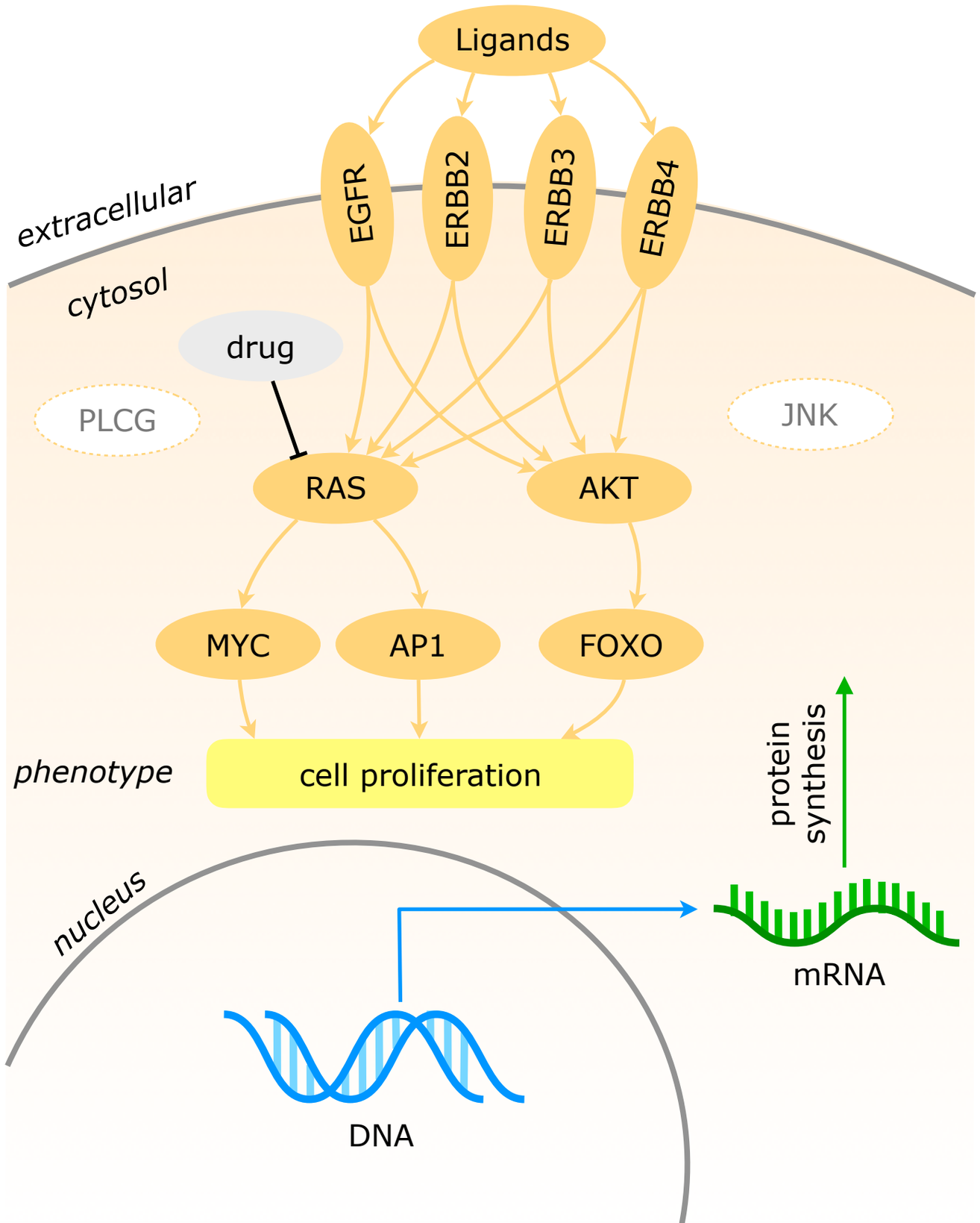
Mechanistic description of protein synthesis:

Genetic profile

Mutation patterns:



Mathematical model



Protein-protein interaction network

$$\dot{x} = S\, v(x, \theta, g, e, d)$$

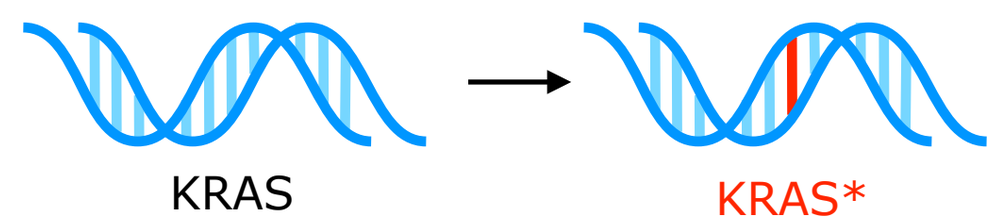
$$y = h(x, \theta)$$

Gene expression profile

Mechanistic description of protein synthesis:

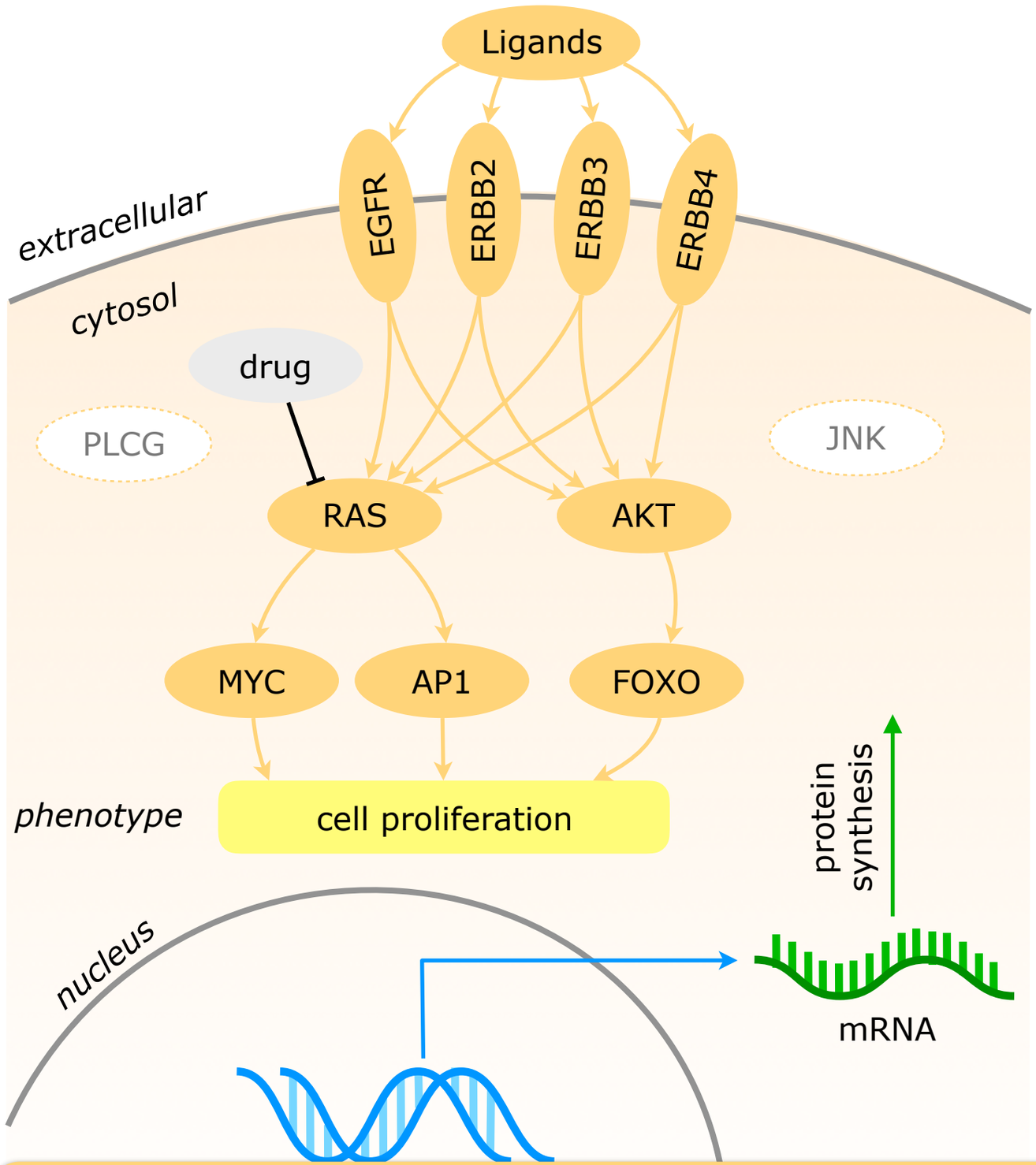
Genetic profile

Mutation patterns:

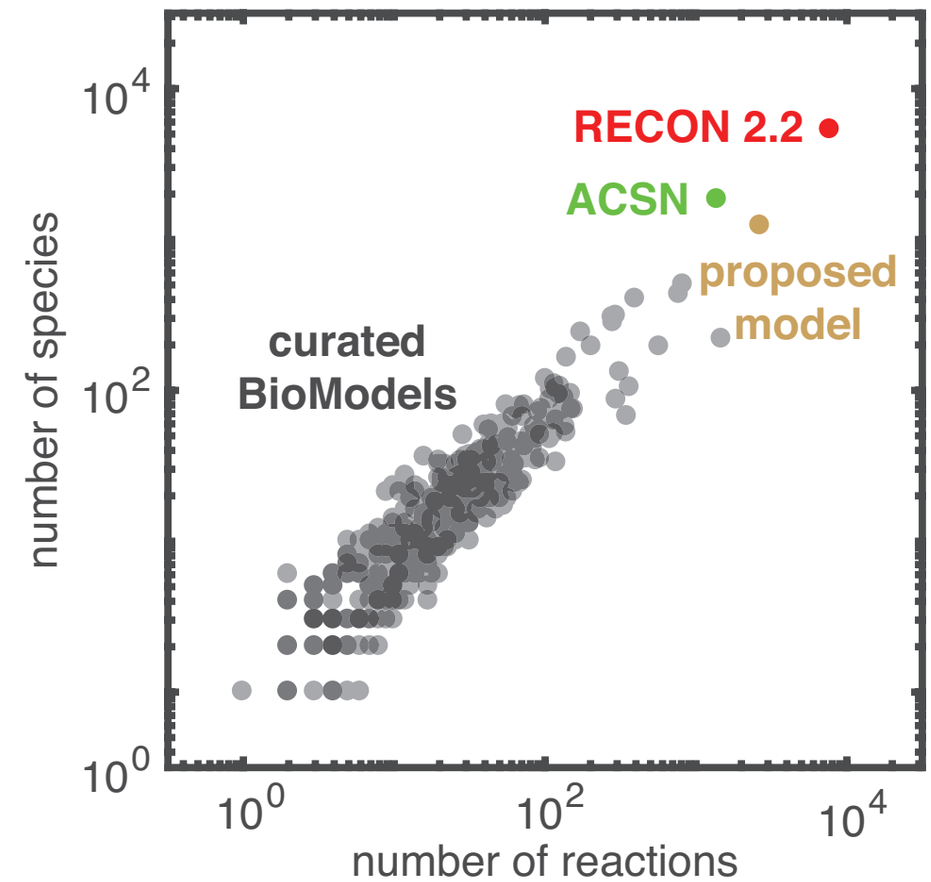


with several employees of the AlacrisTheranostics GmbH

Mathematical model



Model size



Dataset size

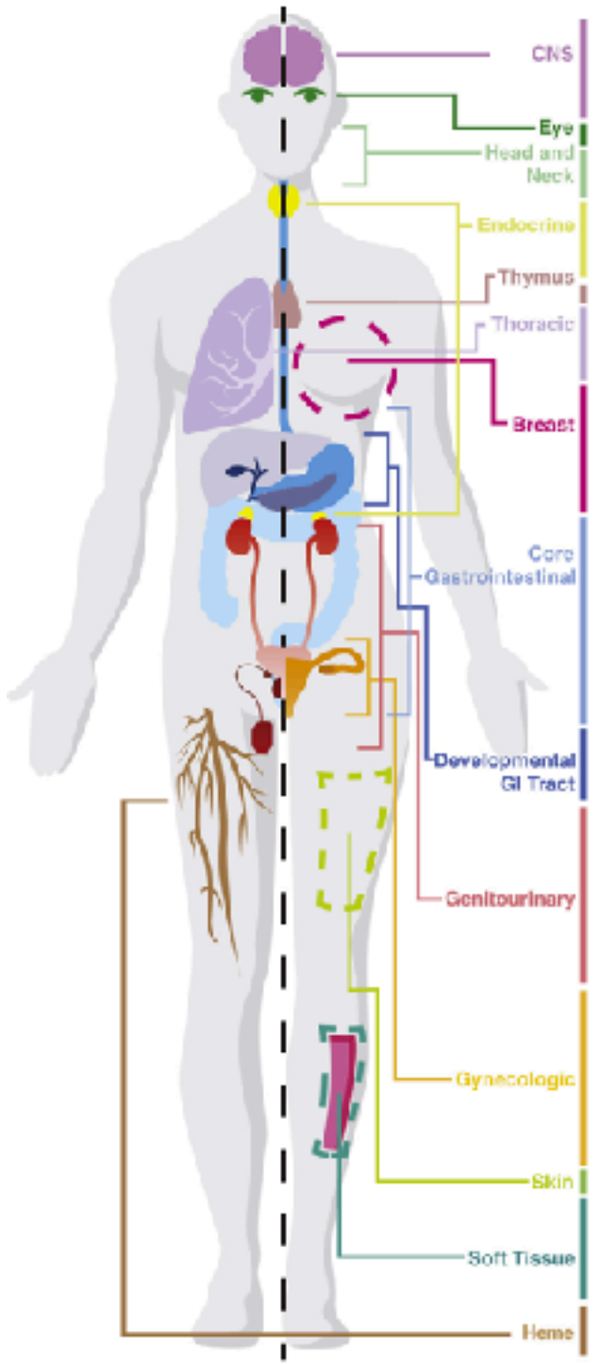
Cell lines: 120

Drugs: 7

Drug concentrations: 8

Do we really need such a “big” model?

Signaling pathways frequently altered in cancer

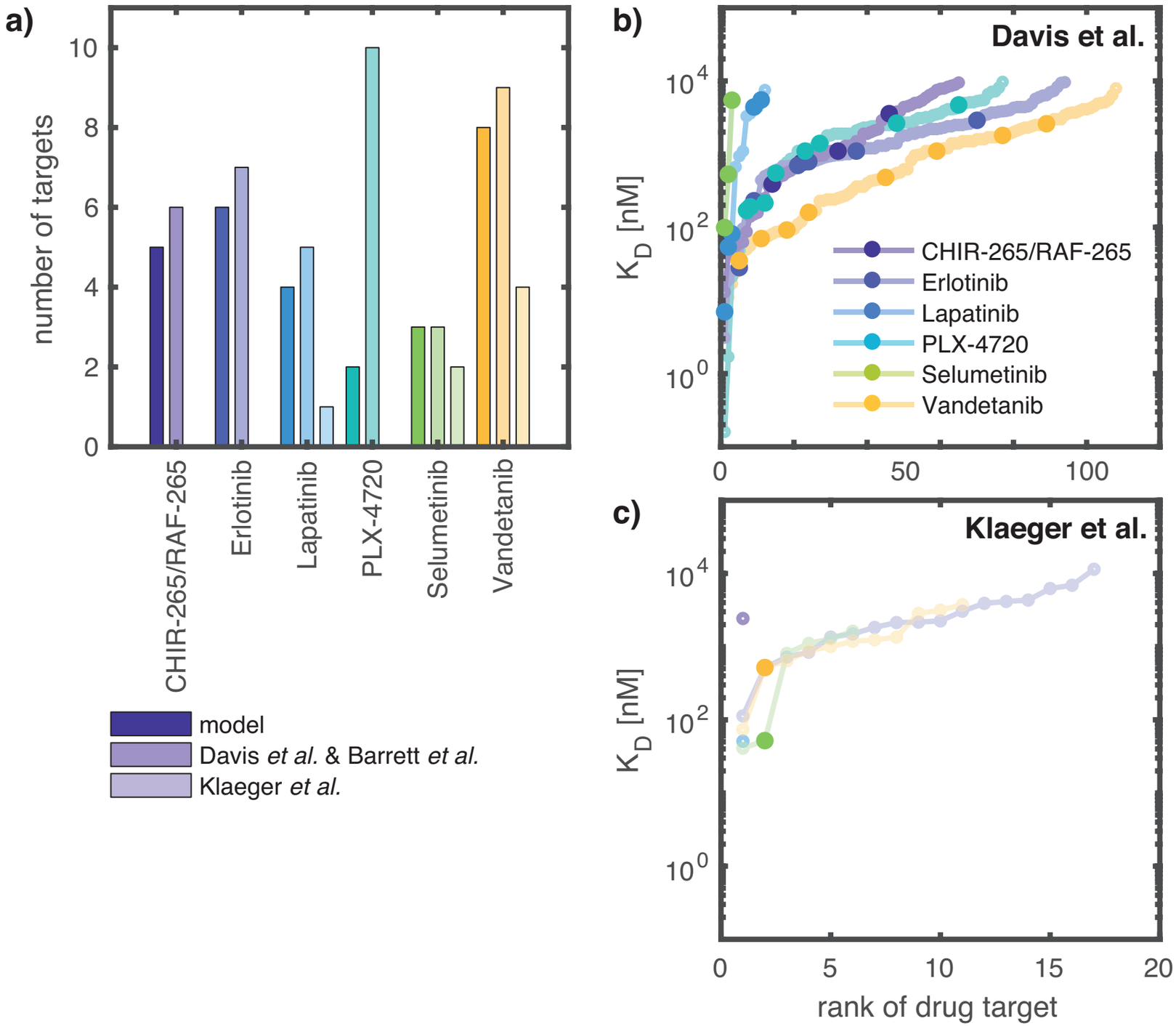


	Alteration frequencies										CIN	FGA	TMB
GBM	77	86	57	48	18	8	6	10	2		119	0.37	4.3
LGG IDHwt	82	64	47	29	27	5	1	5			58	0.25	4.3
LGG IDHmut-codel	9	45	22	5	26	66	3	99	60		34	0.21	0.8
LGG IDHmut	19	28	15	92	24	8	10	92	21	1	60	0.17	0.8
UVM	6	6	4	2	10	1	2	10		1	65	0.28	0.4
HNSC HPV+	26	32	60	11	25	8	4	6	11	1	83	0.32	3.3
HNSC HPV−	45	86	39	82	36	16	20	42	13	13	108	0.43	4.8
THCA	84	14	4	1	4	13	2	1	2		16	0.03	0.4
ACC	22	30	16	28	11	41	7	5	1		146	0.76	1.6
PCPG	32	15	6	6	11	10	1	4	1		81	0.33	0.3
THYM	14	9	4	7	6	7	1	7	3	2	26	0.09	0.8
LUAD	74	56	38	61	21	19	23	23	10	15	118	0.48	8.2
MESO	9	54	13	21	9	6	7	40	2		98	0.41	0.8
LUSC	54	79	68	86	31	18	12	28	11	25	158	0.61	7.7
BRCA LumA	28	31	62	25	14	15	12	5	4	1	101	0.34	2.0
BRCA LumB	44	48	48	49	25	31	26	15	10	2	211	0.60	2.0
BRCA Her2-enriched	82	40	60	78	18	17	29	10	8	1	230	0.53	4.3
BRCA Basal	46	51	53	91	38	11	39	14	8	4	248	0.67	2.7
BRCA Normal	36	36	33	31	3	6	19	3			53	0.16	1.5
STES Squamous	50	89	53	96	38	13	22	21	13	23	189	0.59	3.1
STES CIN	63	74	33	76	21	26	21	16	23	2	222	0.58	3.4
STES EBV	50	100	80	13	83	67	7	10	17		52	0.22	4.1
STES GS	31	39	18	24	31	20	12	4	20	2	66	0.10	2.1
STES MSI-POLE	71	64	64	49	79	70	19	54	57	2	85	0.19	37.1
CRC MSI-POLE	99	74	68	49	74	95	52	64	55	1	45	0.09	66.9
CRC GS	88	45	53	19	29	90	21	10	38	5	55	0.23	2.9
CRC CIN	66	36	32	84	23	91	17	8	22	1	115	0.54	2.9
LIHC	22	69	25	37	26	43	19	12	7	7	121	0.45	2.9
CHOL	56	53	17	19	8	17	19	17	3	6	100	0.58	1.8
PAAD	78	70	19	59	14	12	14	7	41		62	0.26	3.5
KIRC	14	14	17	8	8	7	8	5	3	3	49	0.25	1.6
KIRP	17	12	8	4	12	9	6	11	1	6	49	0.35	2.2
KICH	5	23	15	32	3	3	2	3	5		77	0.60	0.9
BLCA	64	81	46	62	42	20	18	28	9	9	150	0.50	6.8
PRAD	15	28	32	21	13	35	11	5	6	1	92	0.16	1.3
TGCT sem	63	8	11	6	6		2				70	0.54	0.4
TGCT non-sem	20	7	5	5	16	2		10	2		99	0.67	0.4
OV	58	48	49	96	28	10	40	21	5	5	318	0.79	2.4
UCEC CN high	61	43	86	90	32	18	31	13	5	5	296	0.67	1.9
UCEC CN low	37	9	95	10	14	54	10	7	1	5	42	0.15	2.1
UCEC MSI-POLE	71	31	98	42	64	70	30	55	31	19	30	0.08	71.2
UCS	61	70	79	91	54	18	27	18	4	4	247	0.71	3.5
CESC Adeno	63	21	56	19	30	14	18	14	21	5	95	0.38	3.8
CESC Squamous	32	19	59	12	35	12	5	33	11	10	101	0.44	5.2
SKCM	94	77	33	28	27	23	10	25	7	1	131	0.53	22.1
SARC DDLPS	43	83	20	85	17	15	7	9	7		450	0.36	1.1
SARC LMS	31	55	33	71	14	11	4	4	1	4	177	0.69	1.8
SARC MFS/UPS	48	74	32	68	34	20	8	21	6	4	328	0.66	3.0
SARC other	25	30	15	5	5	5		10			104	0.33	1.2
DLBC	24	76	8	19	70	70	14	35	14		90	0.29	3.5
LAML	49	17	3	9	18	11	2	3	1	1	28	0.05	1.1
	46	45	33	29	23	15	11	10	7	1			
	RTK/RAS	Cell cycle	PI3K	p53	Notch	Wnt	Myc	Hippo	TGFβ	Nrf2			

Sanchez-Vega et al., Oncogenic signaling pathways in The Cancer Genome Atlas, Cell, 173:321–337, 2018

with Daniel Weindl

Polypharmacology of anti-cancer drugs



Parameter estimation for models and datasets of this size?

Parameterisation of large-scale pathway model

Standard optimisation methods

#cross validations (5)
x #local optimisations(10)
x #iterations(100)
x #parameters(~4000)
x #conditions(~5500)
= **~10^{11} ODE solves**

x ~2min = **~200k years**

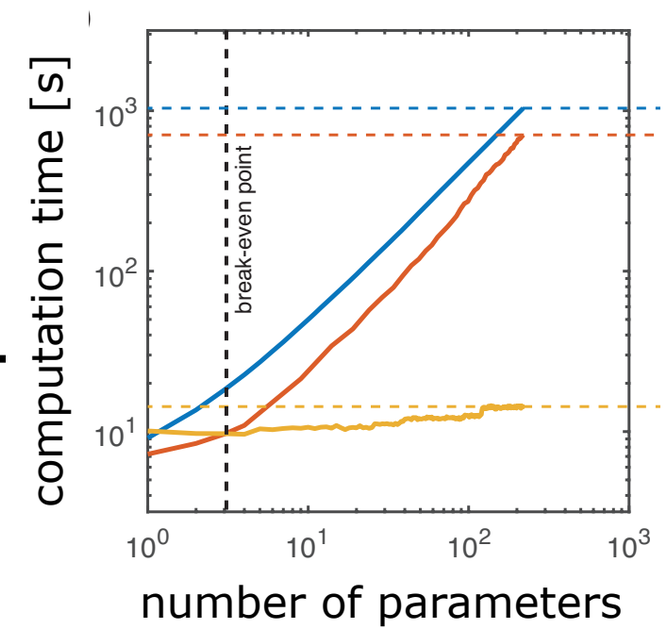
Tailored optimisation methods

#cross validations (**1**)
x #local optimisations(**1**)
x #iterations(100)
x #parameters(**~3**)
x #conditions(**~20**)
= **~$2x10^5$ ODE solves**

x ~2s = **~6 hours**

Acceleration method

parallelise (~1000x)
exploit sparsity (~60x)
adjoint sensitivities (~1500x)
hierarchical optimisation (~10x)
total speed-up: ~10^9



Parameterisation of large-scale pathway model

Standard optimisation methods

#cross validations (5)
x #local optimisations(10)
x #iterations(100)
x #parameters(~4000)
x #conditions(~5500)
= **~10^{11} ODE solves**

x ~2min = **~200k years**

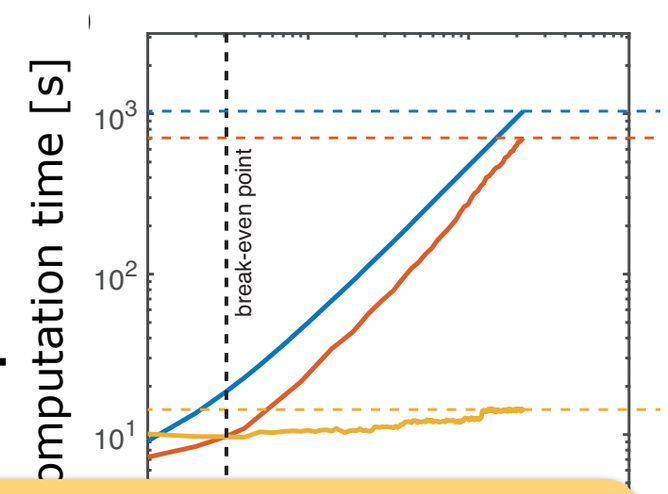
Tailored optimisation methods

#cross validations (**1**)
x #local optimisations(**1**)
x #iterations(100)
x #parameters(**~3**)
x #conditions(**~20**)
= **~$2x10^5$ ODE solves**

x ~2s = **~6 hours**

Acceleration method

parallelise (~1000x)
exploit sparsity (~60x)
adjoint sensitivities (~1500x)
hierarchical optimisation (~10x)

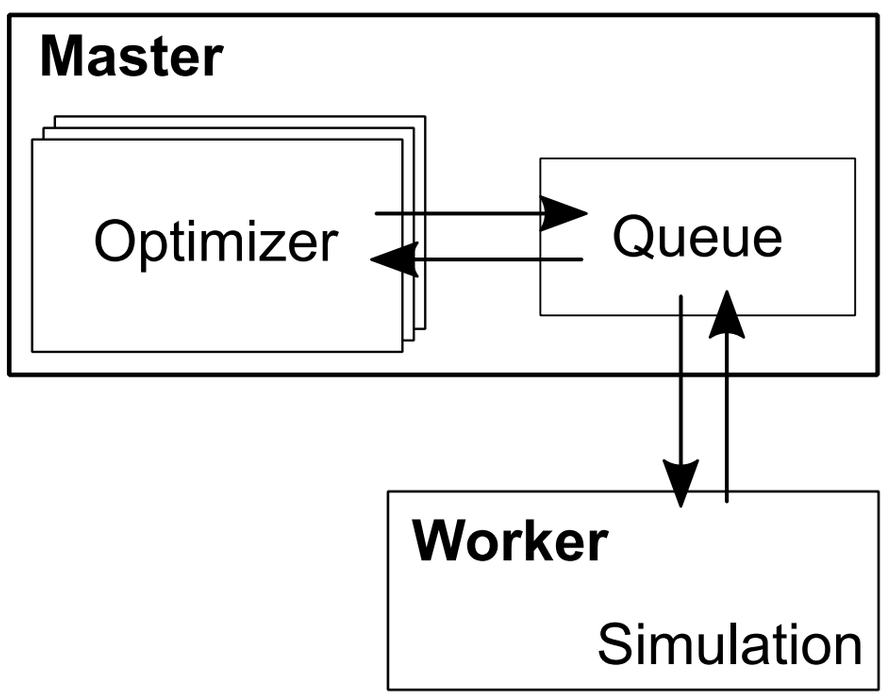


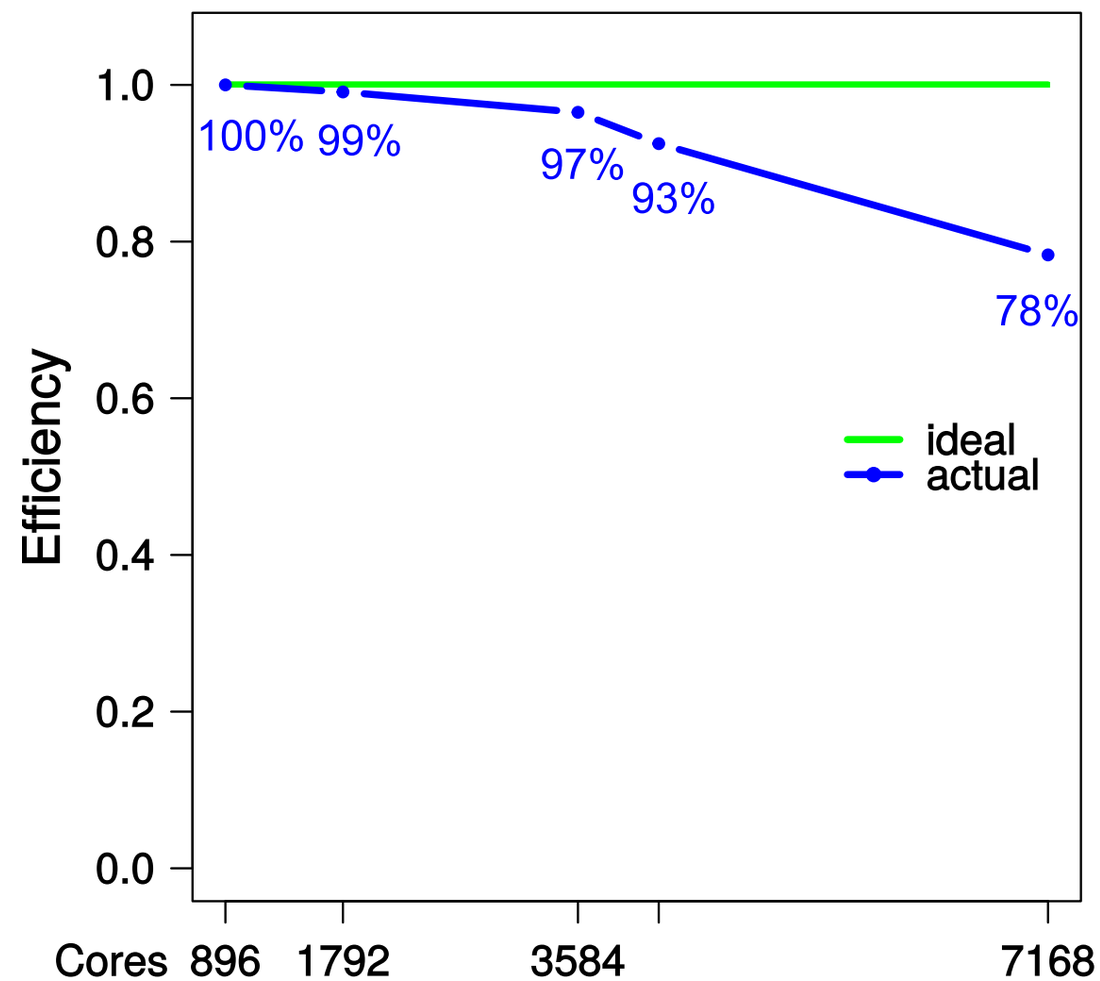
Scalable framework for data-driven modelling.

High-performance computing

with *Daniel Weindl*
Leonard Schmiester
and many others

SuperMUC (c) LRZ

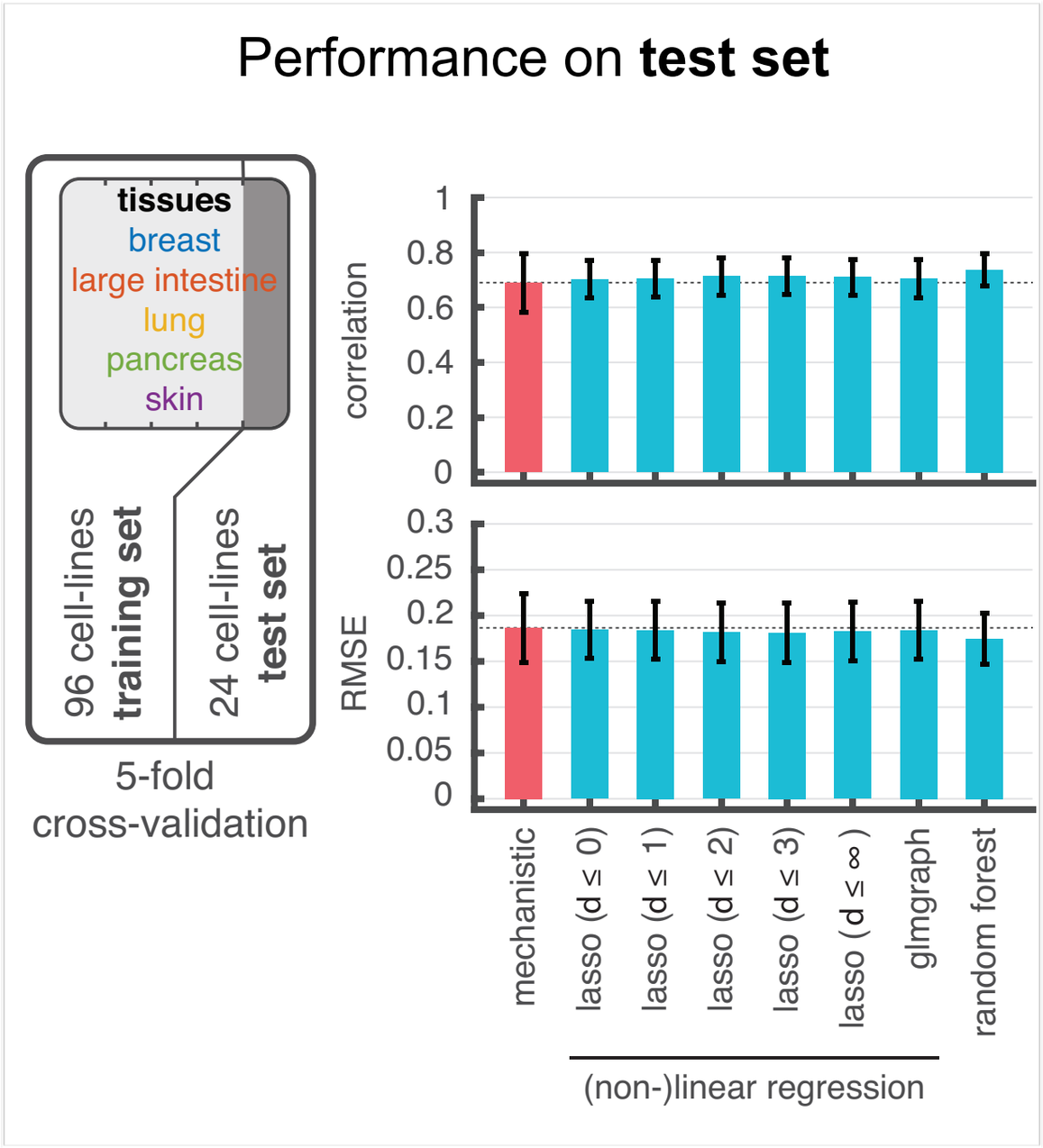


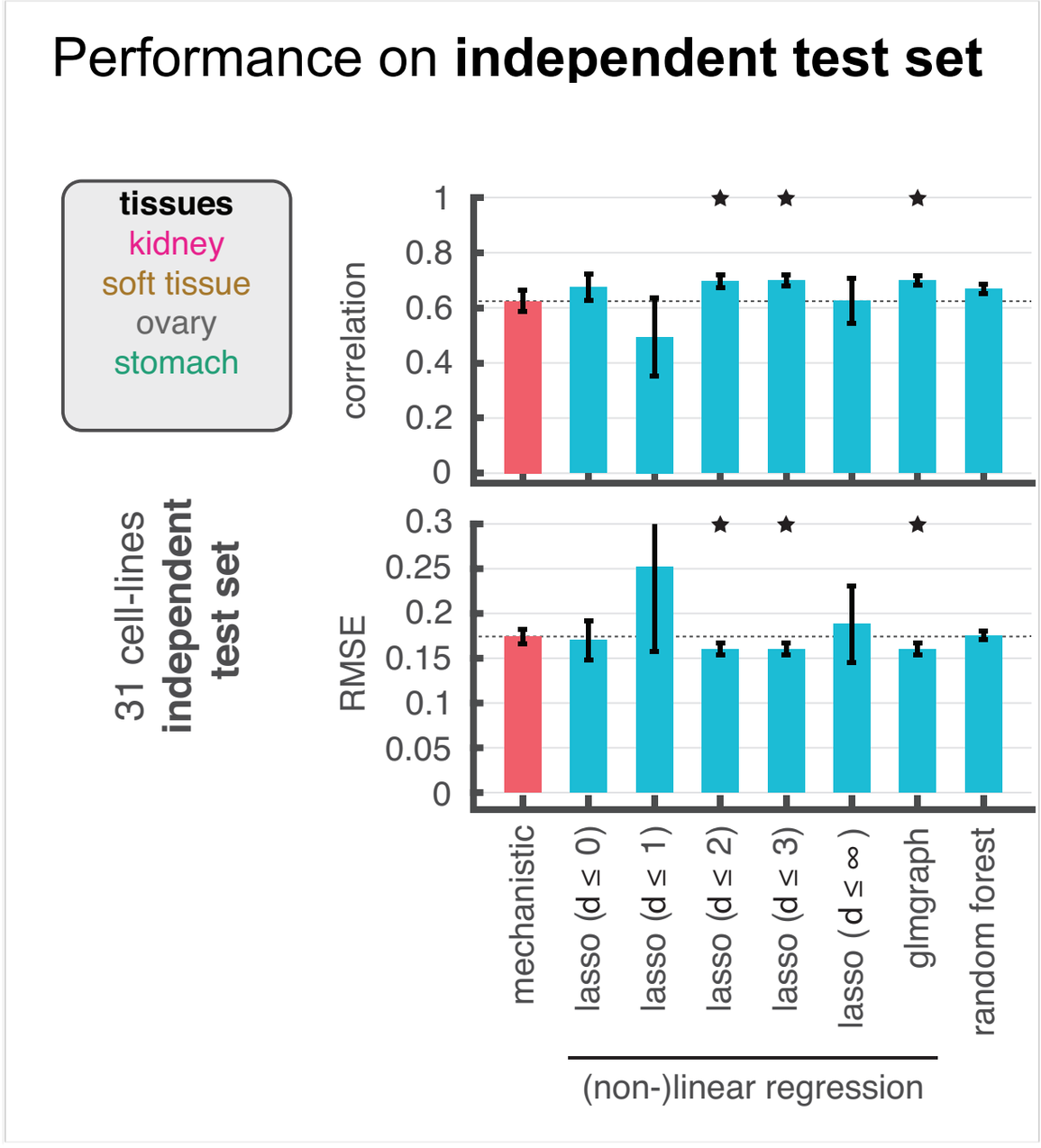


→ 10 local optimizations
in 9h @ 3584 cores
(instead of 21 months)

with Fabian Fröhlich
and many others

Prediction accuracy

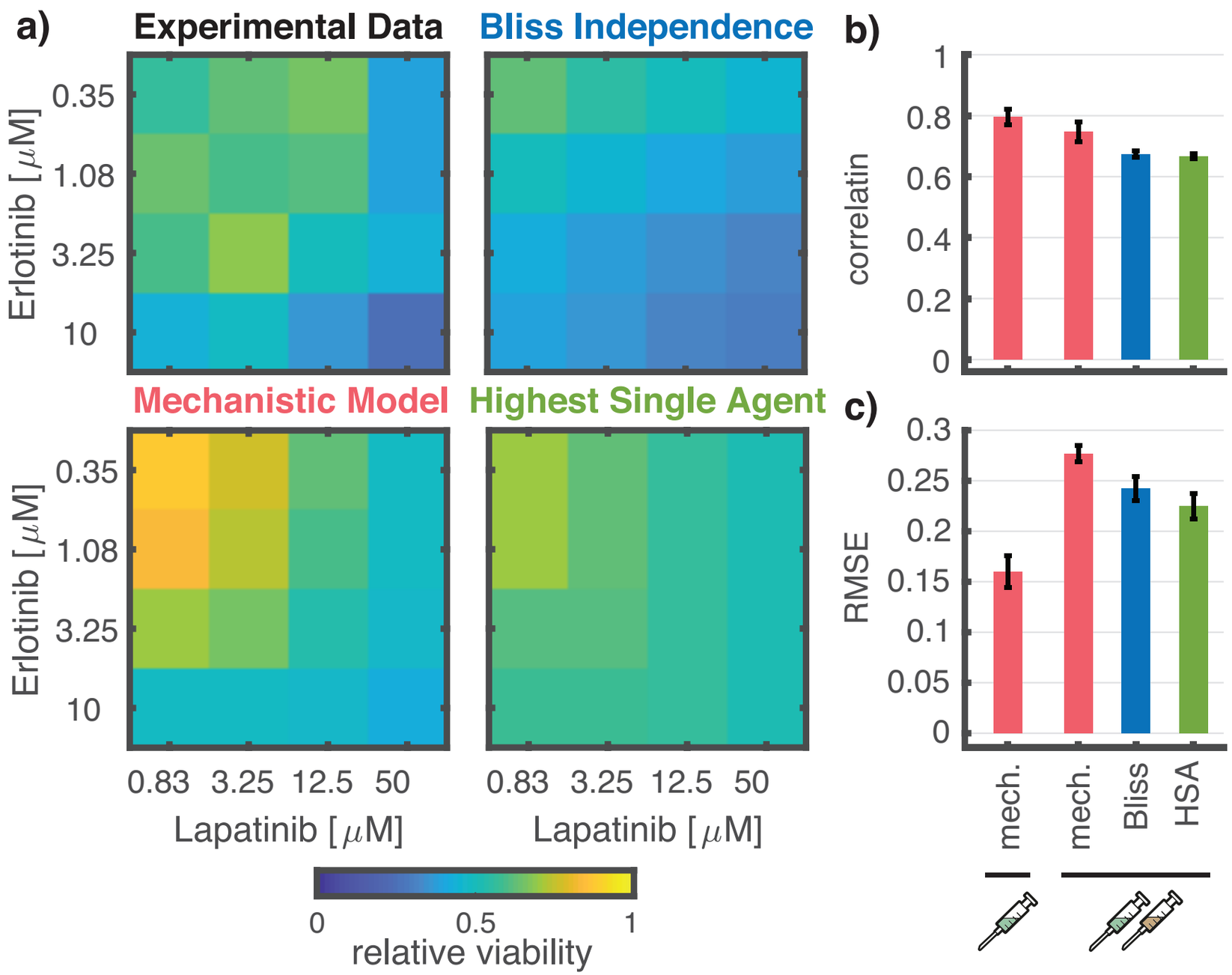




Mechanistic model reaches similar performance as well established statistical models.

with Fabian Fröhlich and many others

Prediction of combination therapies



Mechanistic approach outperforms statistical methods.

Questions or comments?

Thanks for your attention!