

Four ways to fit an ion channel model

Michael Clerx



Gary Mirams
Nottingham



Kylie Beattie
Oxford



Chon Lok Lei
Oxford



David Gavaghan
Oxford

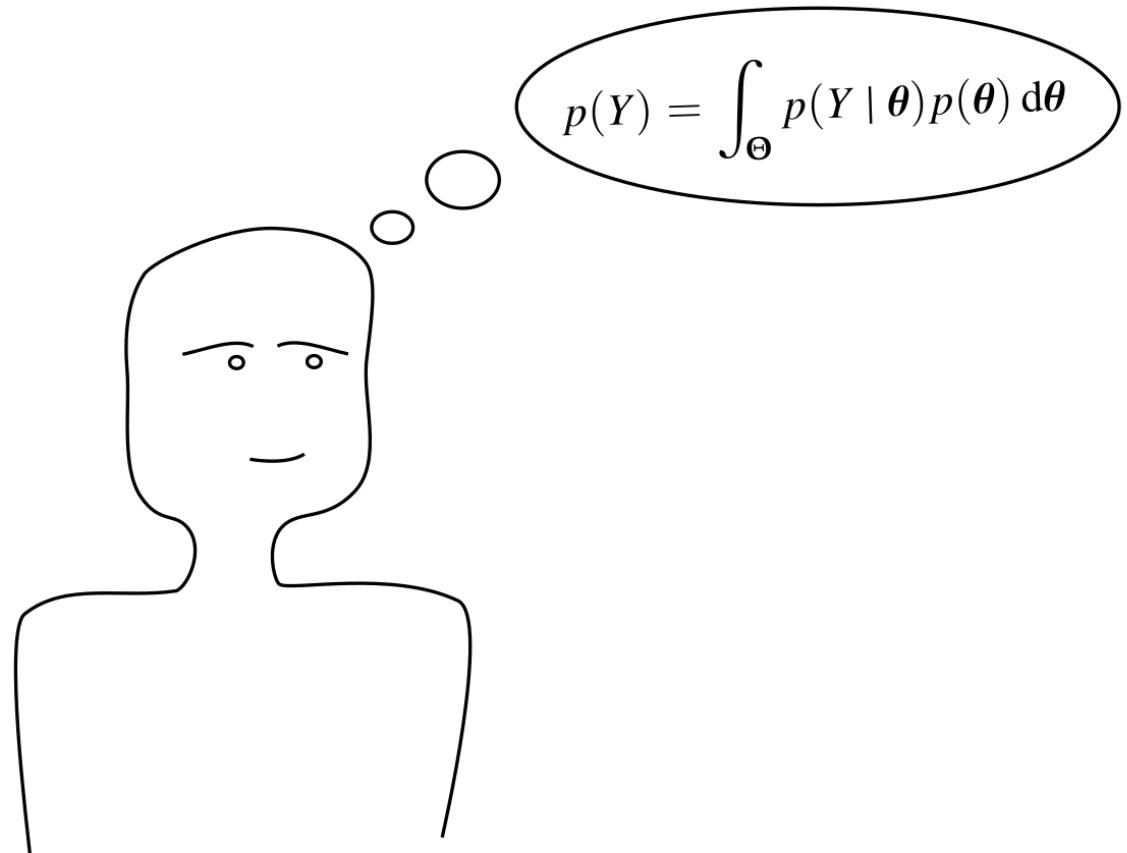
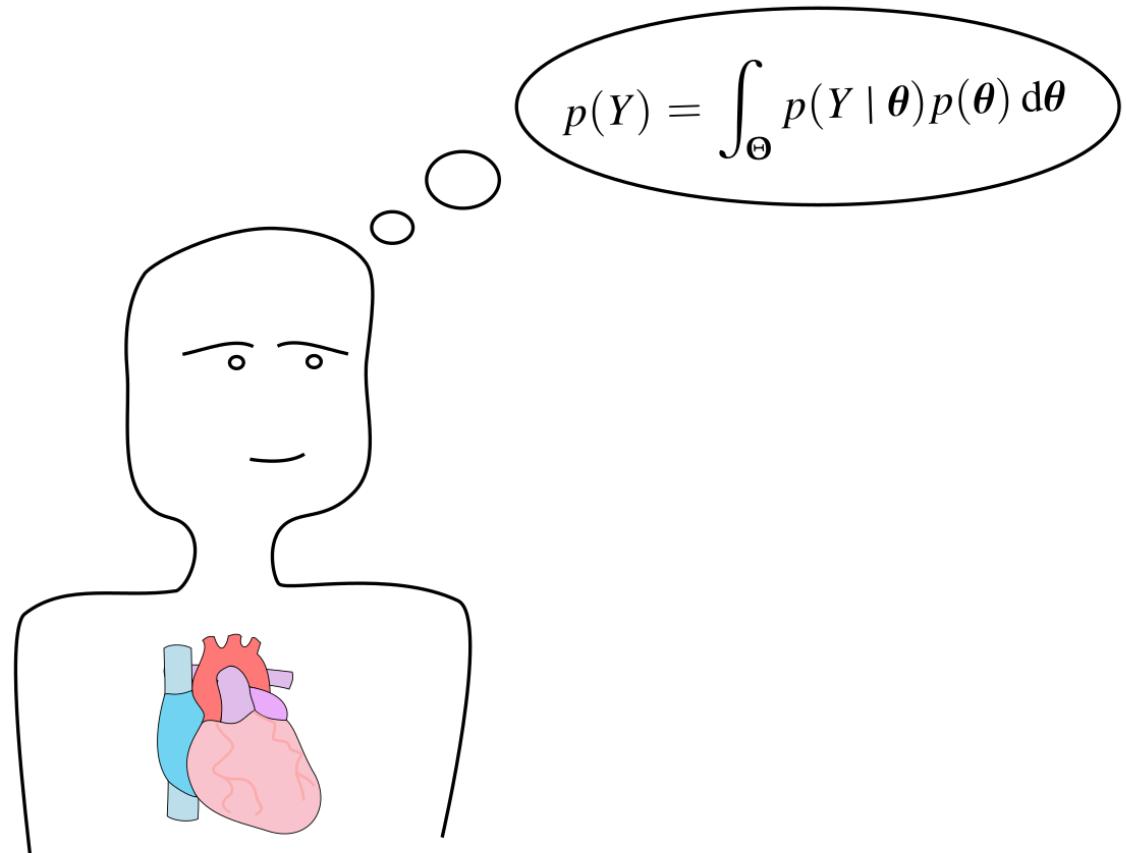
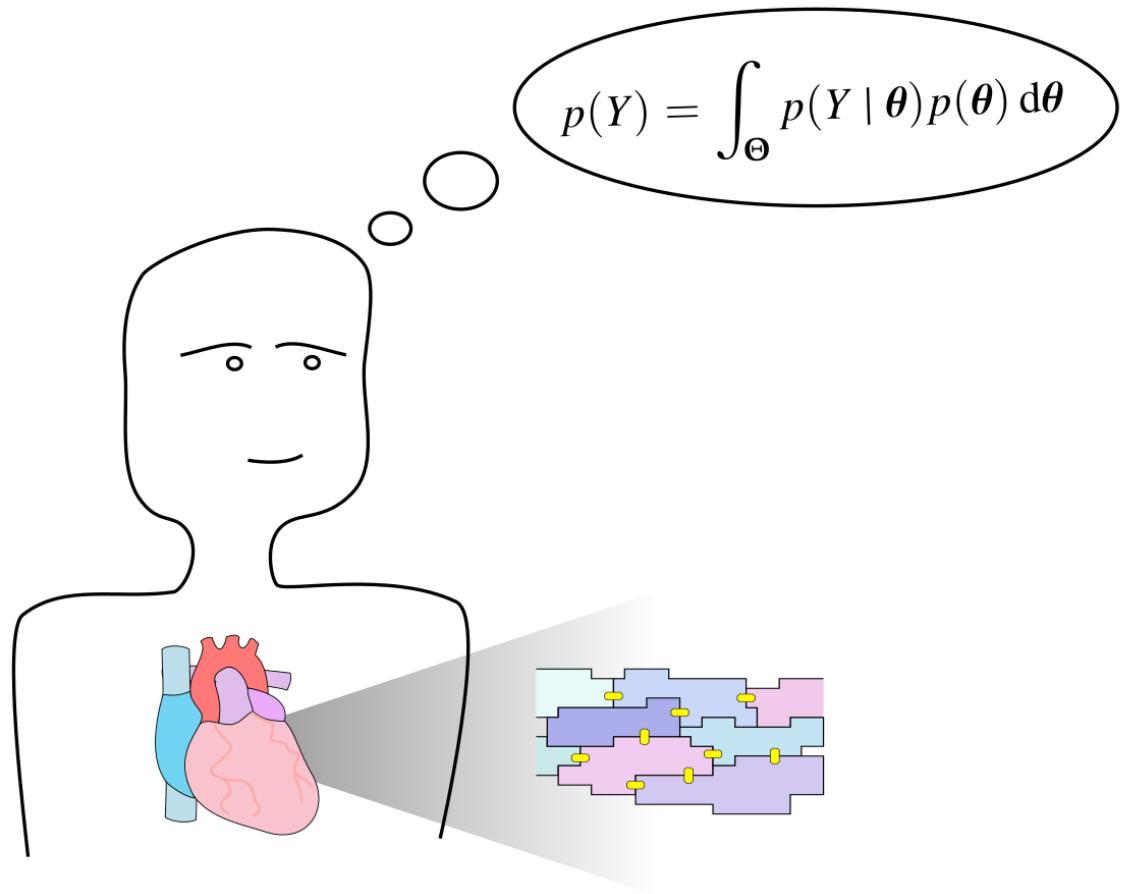
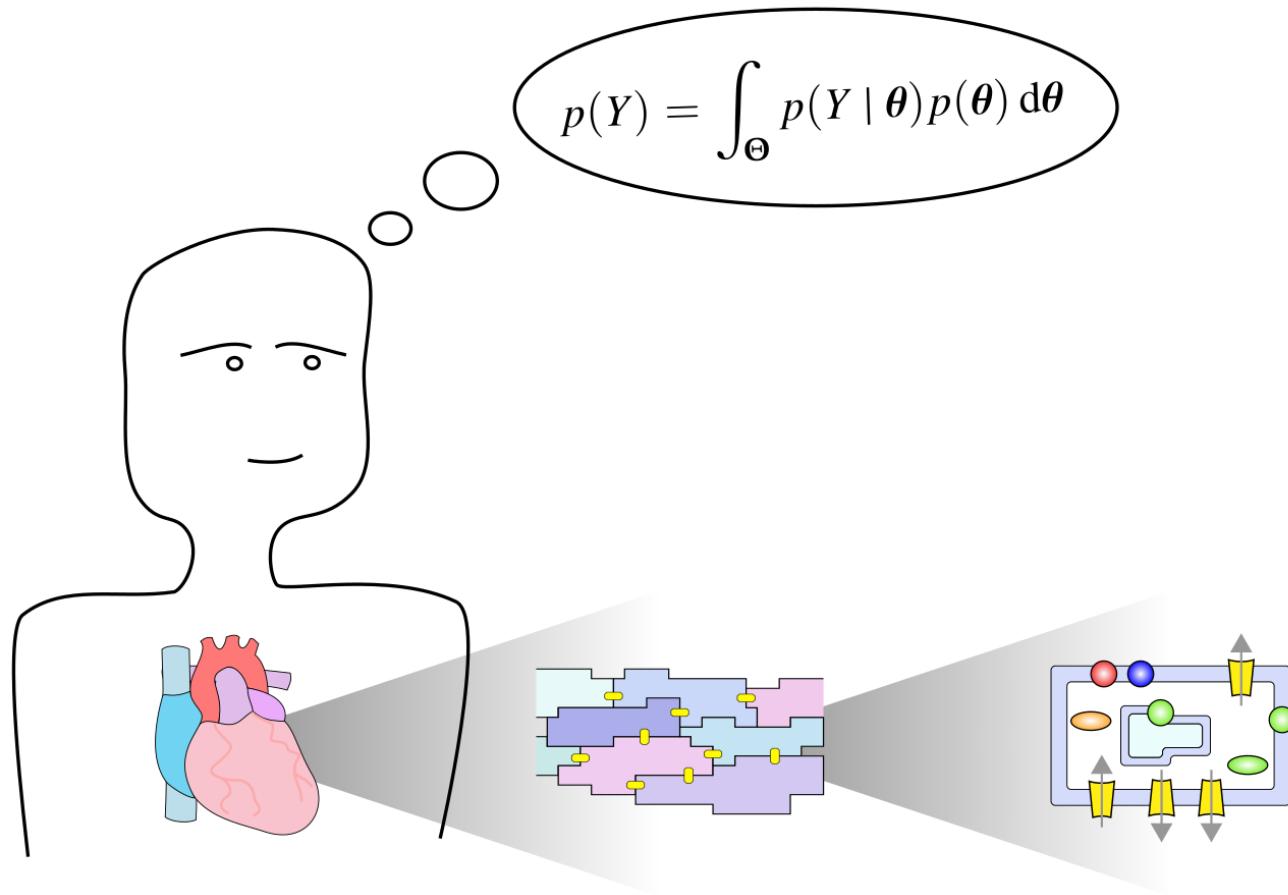
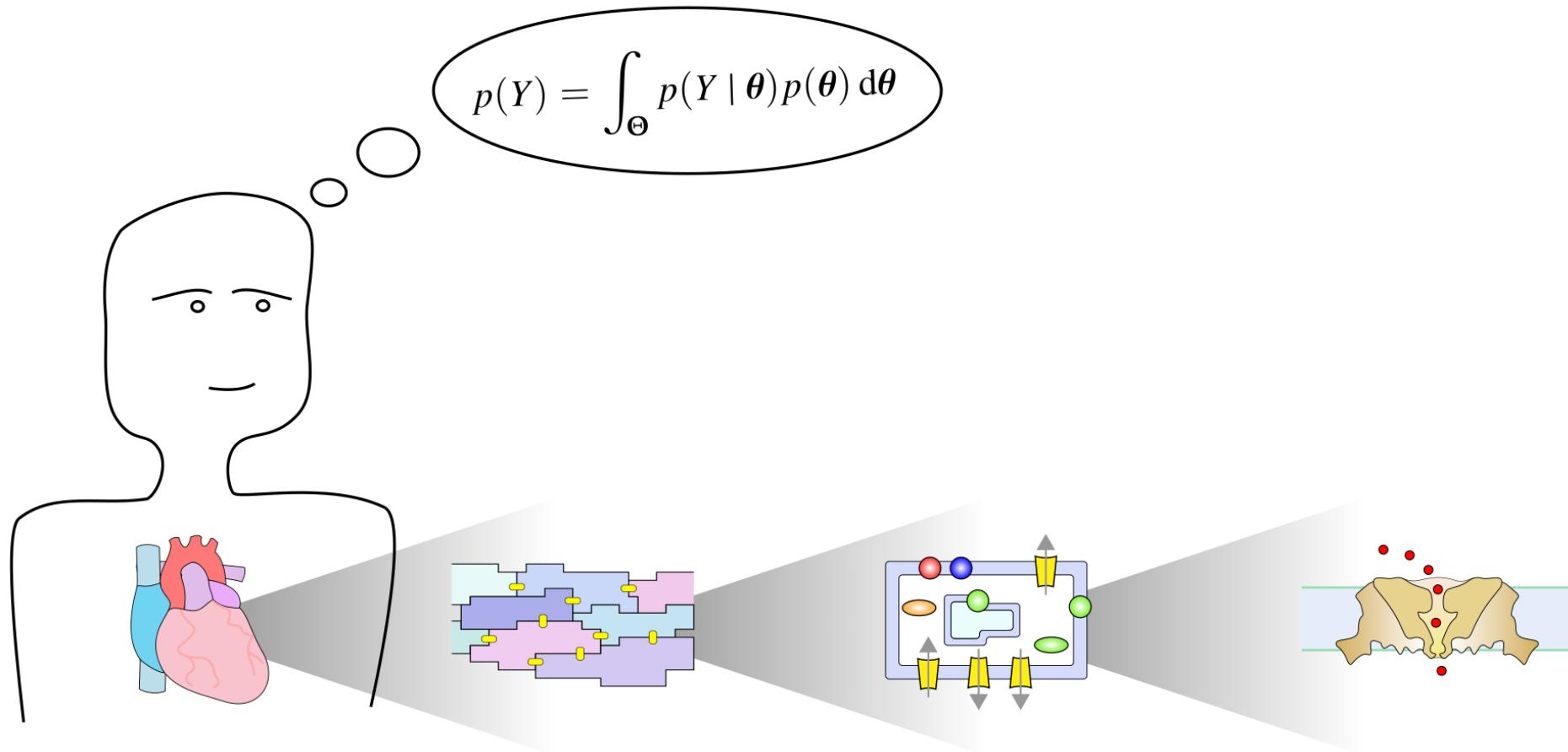


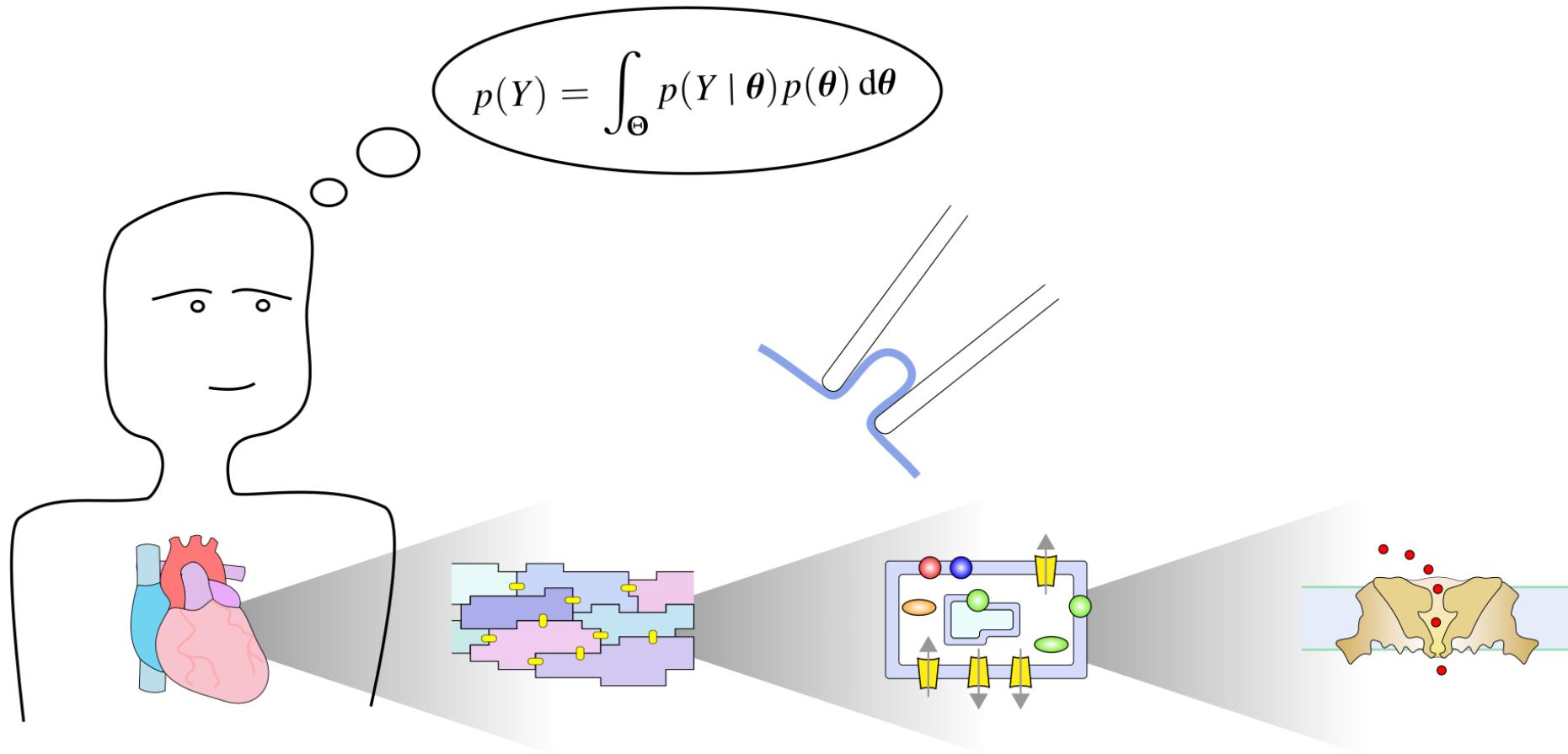
Figure 1: A lovely day

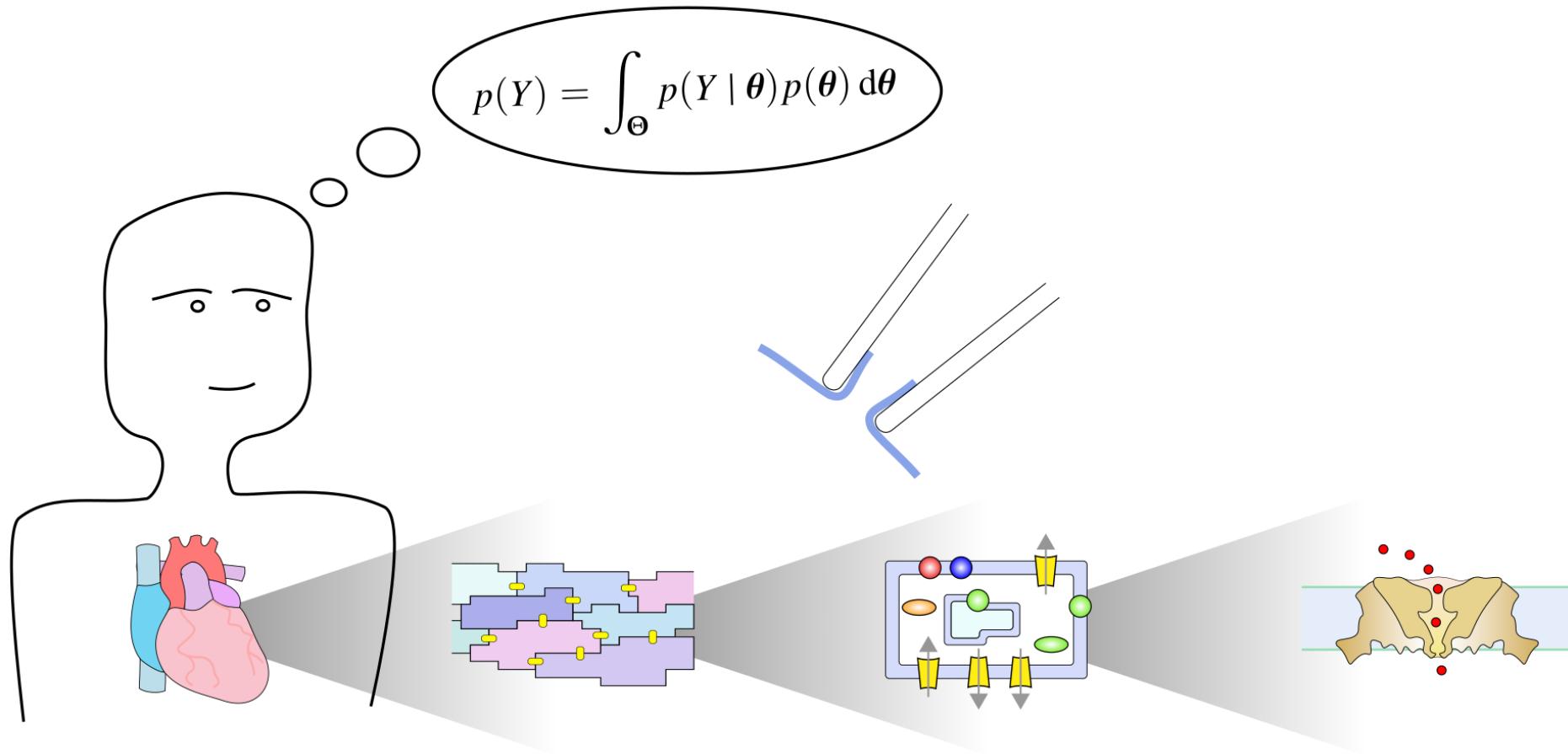


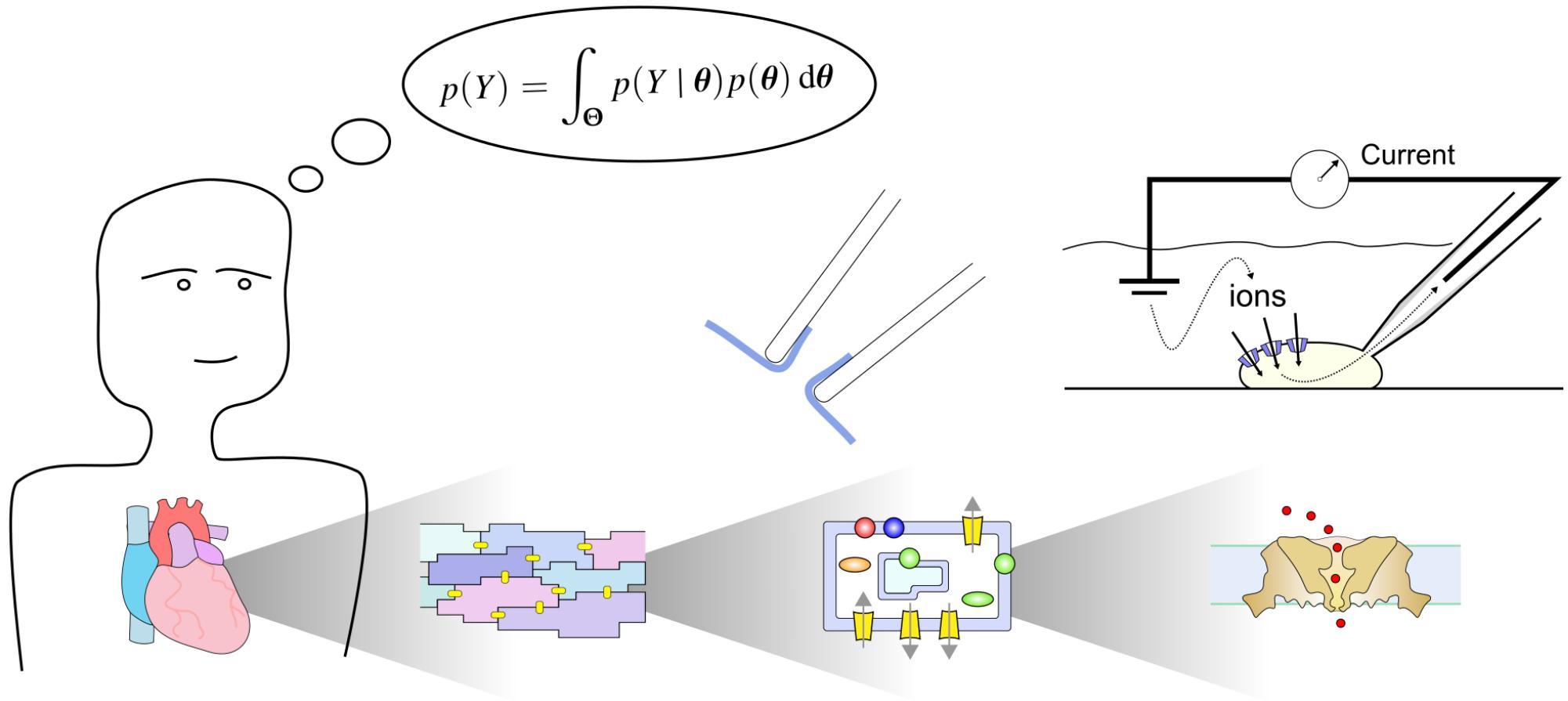


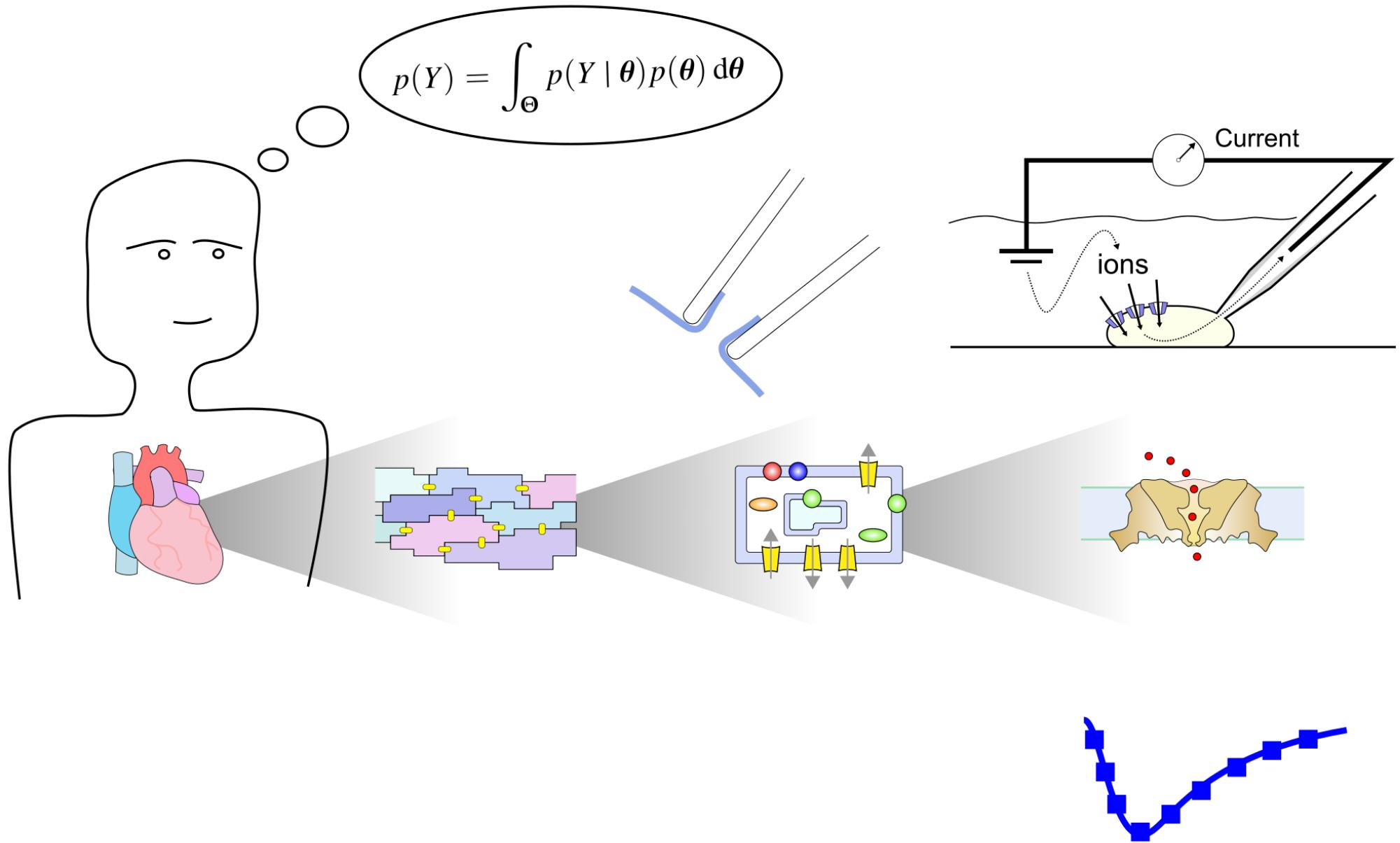


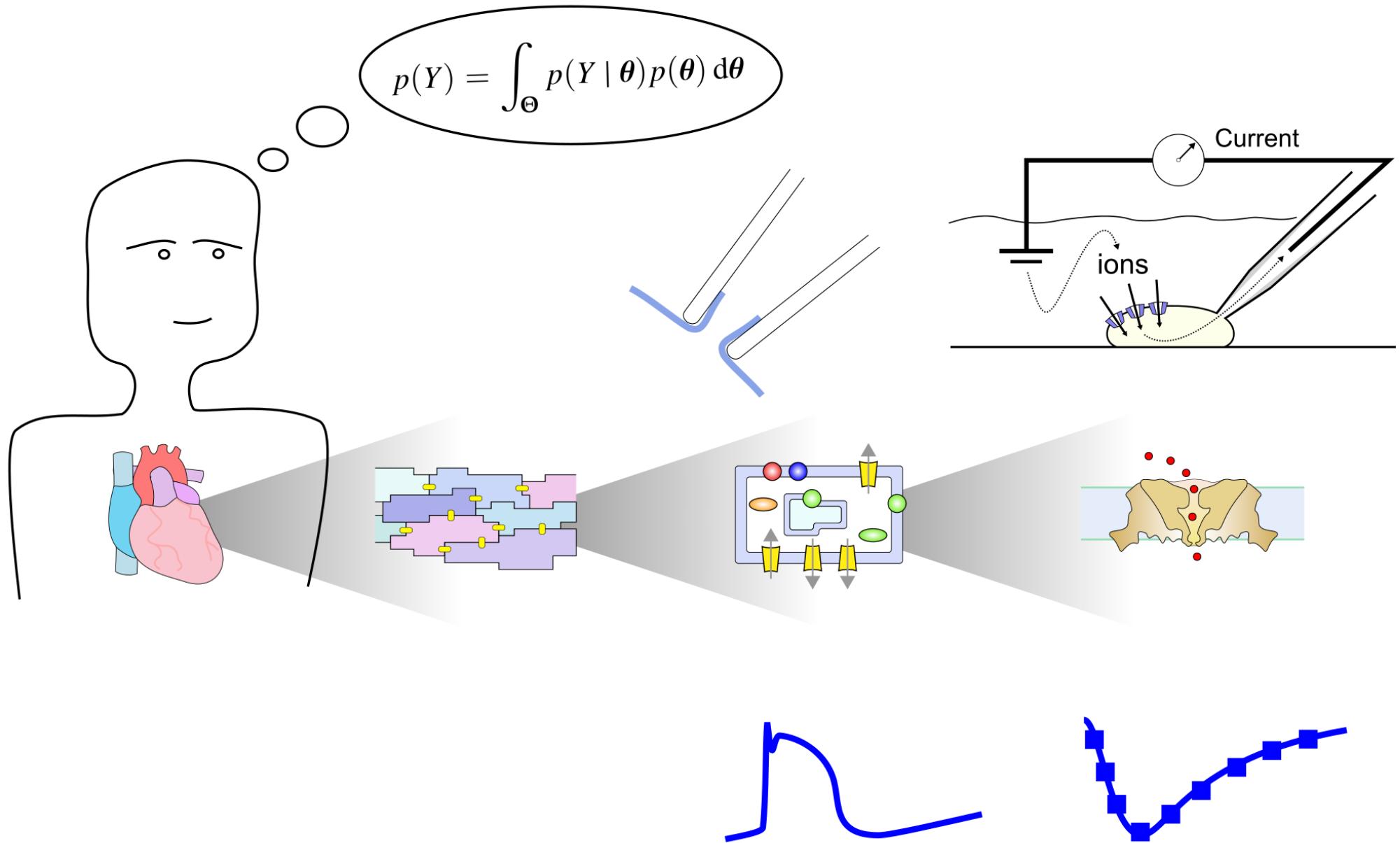


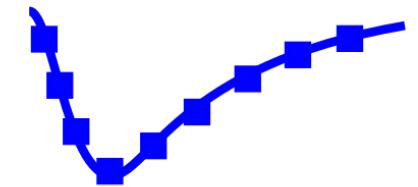
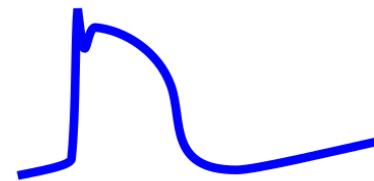
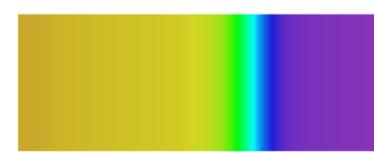
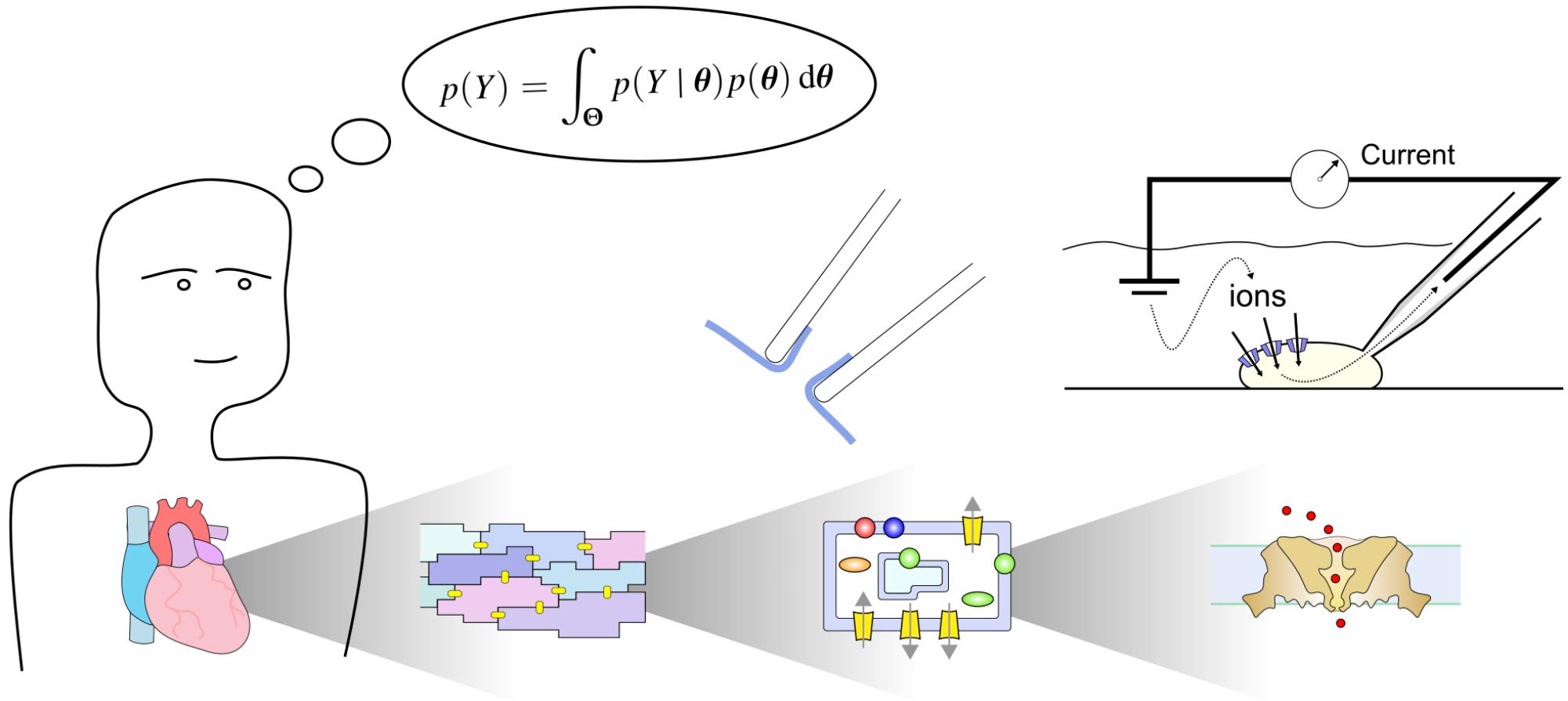


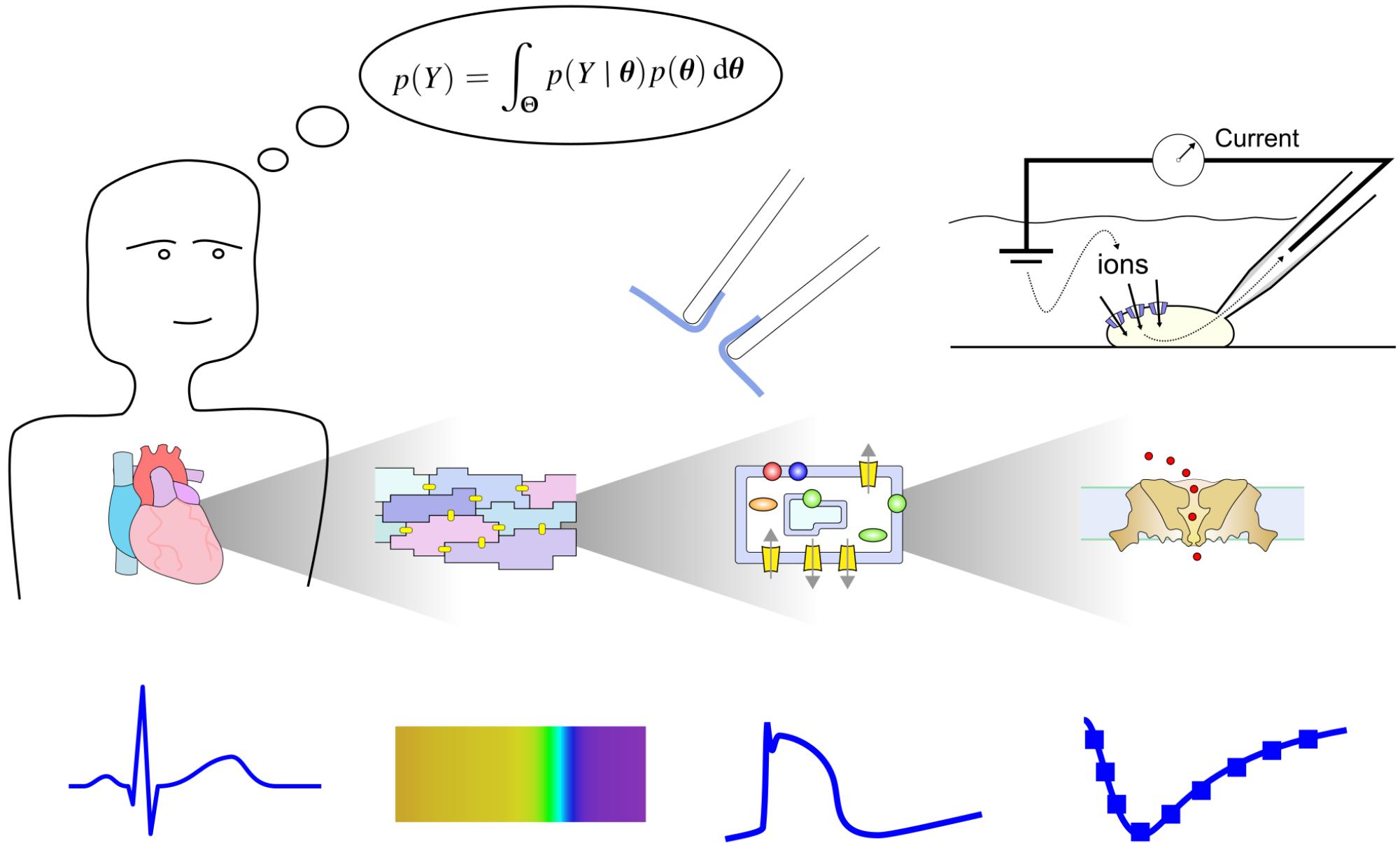


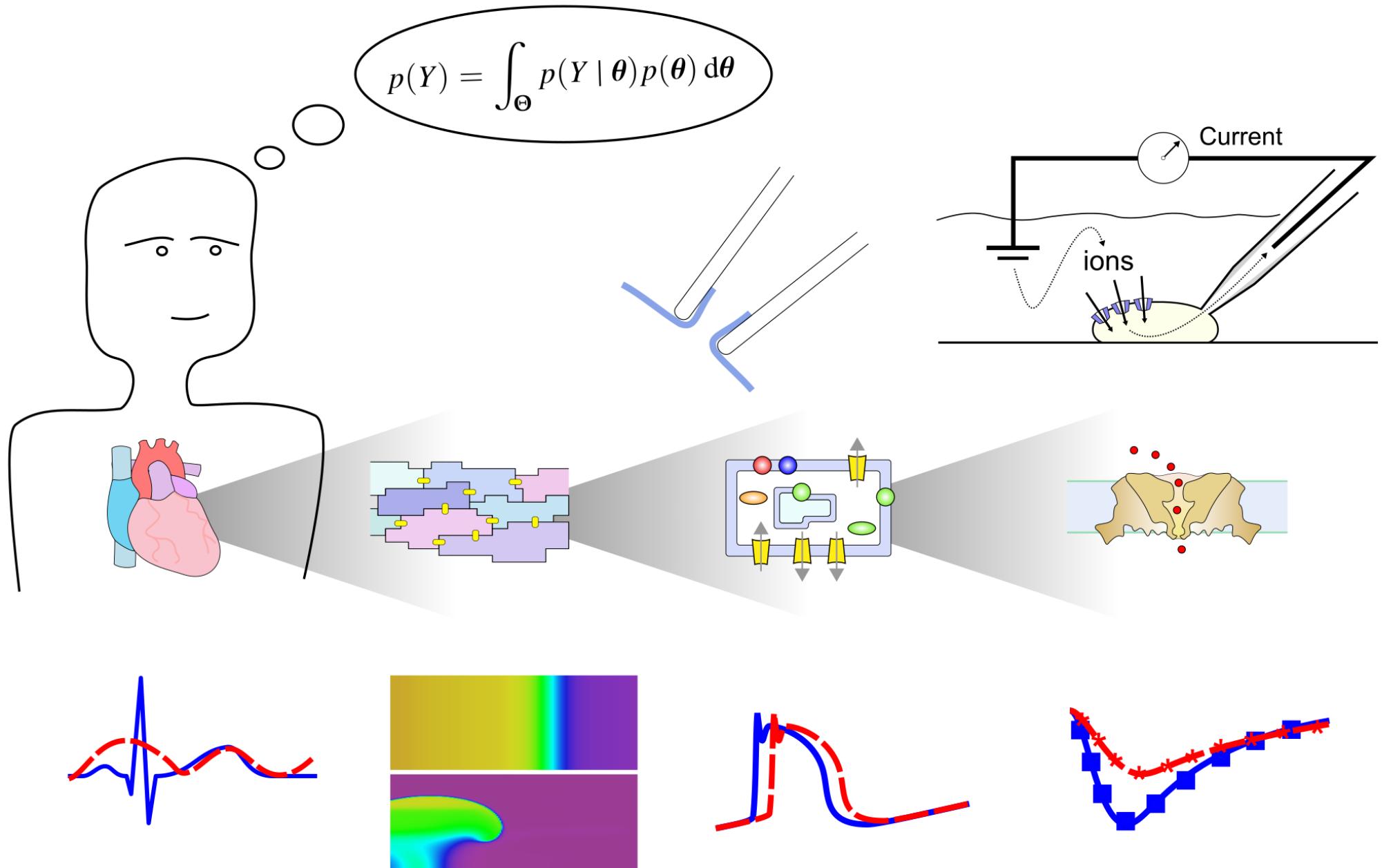






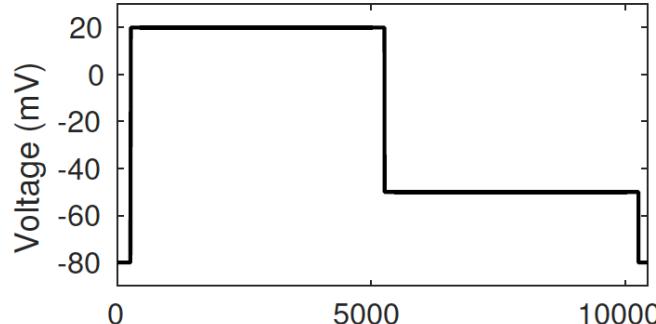




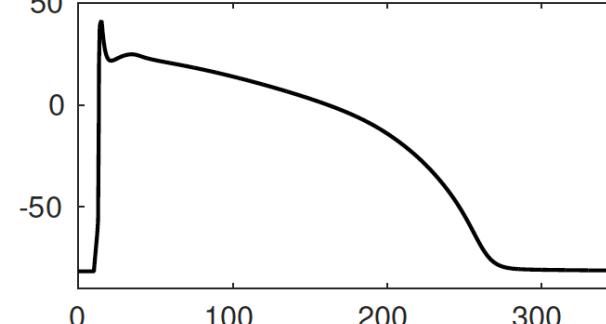


Our example will be IKr well studied, but published models disagree...

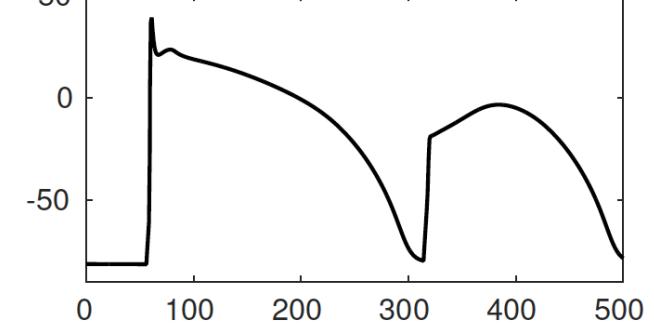
i)



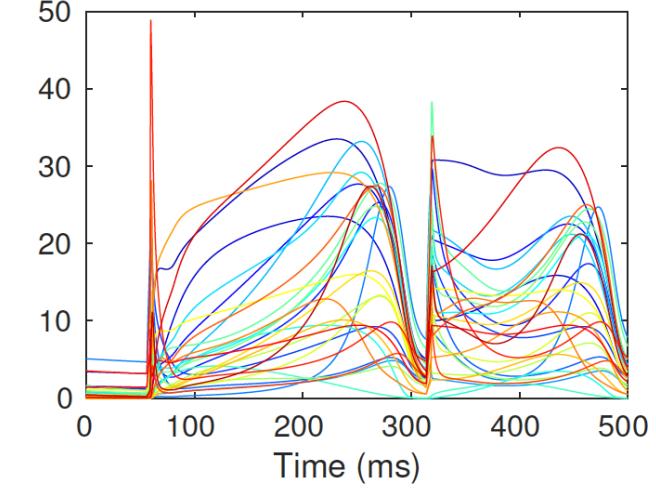
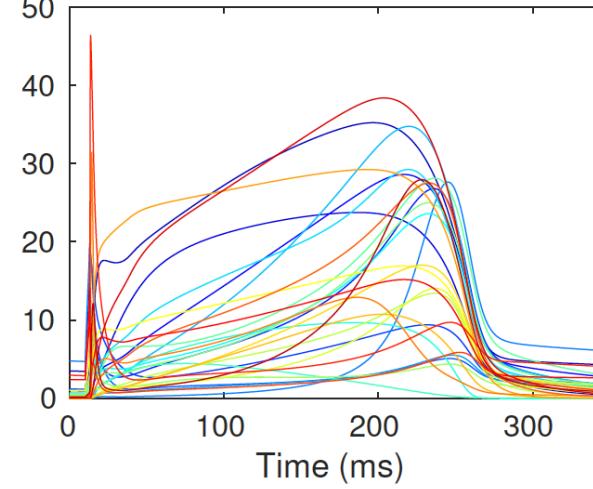
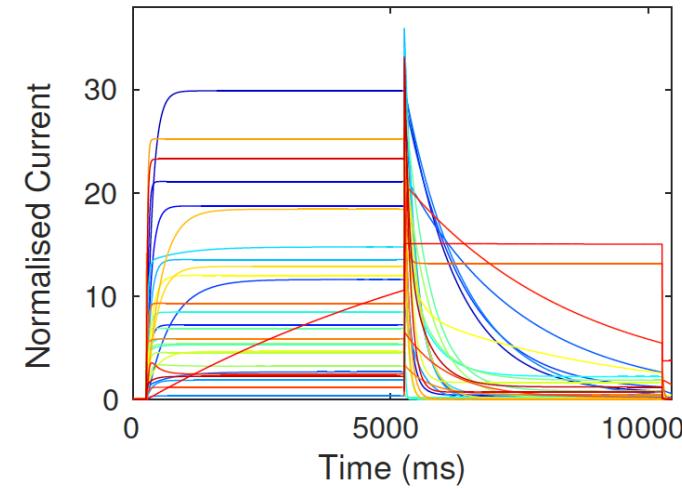
ii)



iii)

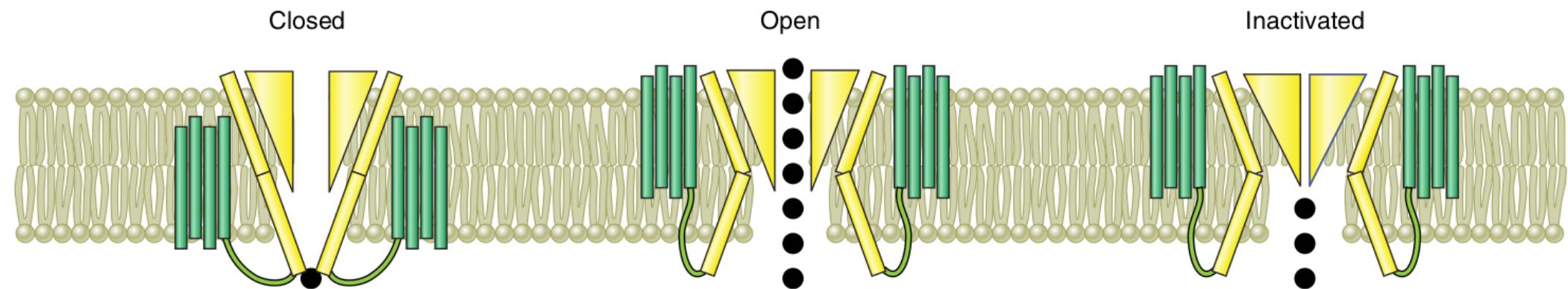


Normalised Current



Beattie et al. 2018; Sinusoidal voltage protocols for rapid characterisation of ion channel kinetics

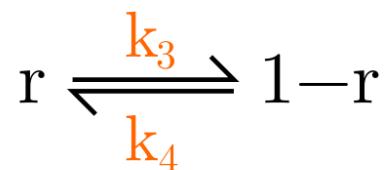
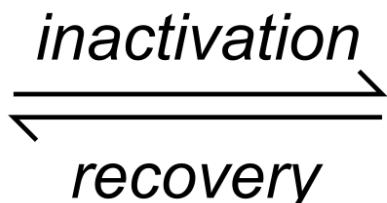
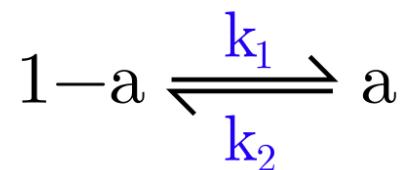
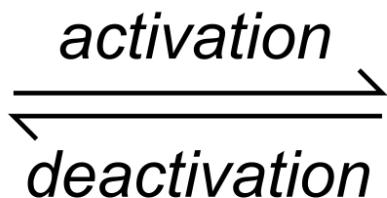
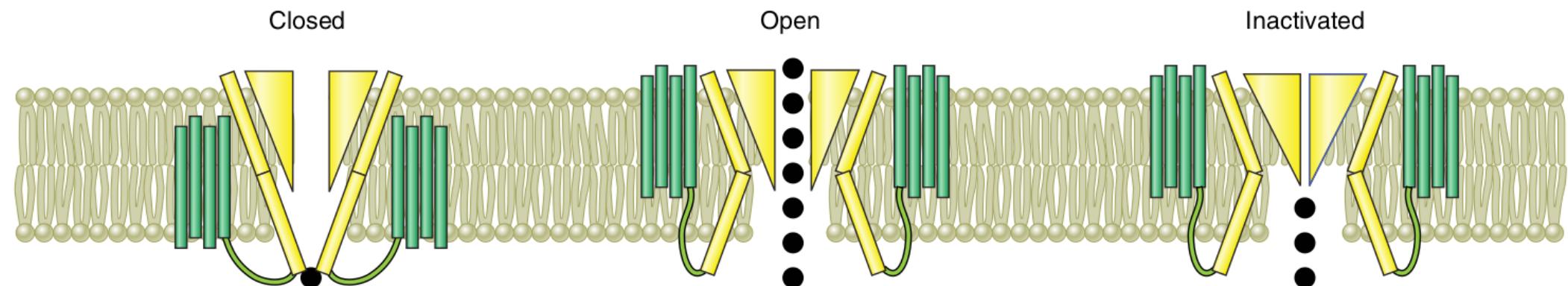
I_{Kr} channels “*activate*” and “*inactivate*”



Vandenberg et al. 2012; hERG K channels: structure, function, and clinical significance

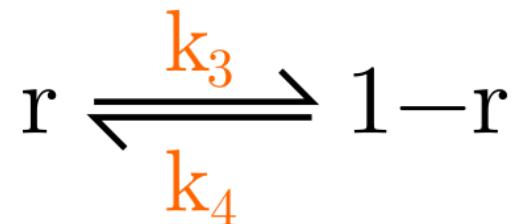
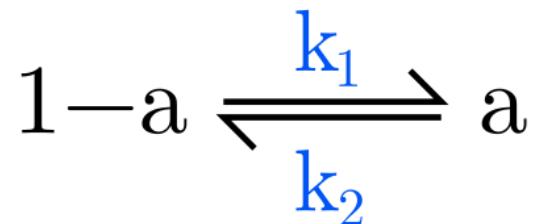
I_{Kr} channels “activate” and “inactivate”

We can model this as two independent processes



$$I(V,t) = g \times a \times r \times (V - E)$$

I_{Kr} channels “activate” and “inactivate”



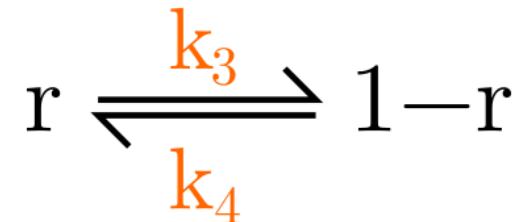
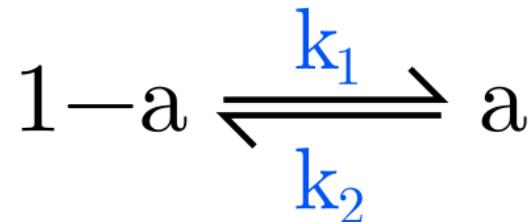
$$\frac{da}{dt} = k_1 (1-a) - k_2 a$$

$$\frac{dr}{dt} = k_4 (1-r) - k_3 r$$

$$k_i = a_i \exp(b_i V)$$

$$I(V,t) = g \times a \times r \times (V - E)$$

I_{Kr} channels “activate” and “inactivate”



$$\frac{da}{dt} = k_1 (1-a) - k_2 a$$

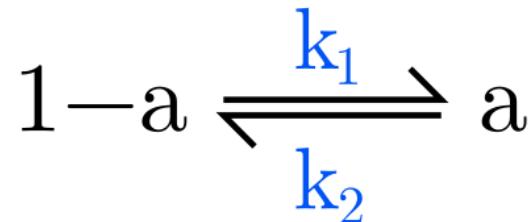
$$\frac{dr}{dt} = k_4 (1-r) - k_3 r$$

$$k_i = a_i \exp(b_i V)$$

$$I(V,t) = g \times a \times r \times (V - E)$$

↑ ↑ ↑
Output States Input

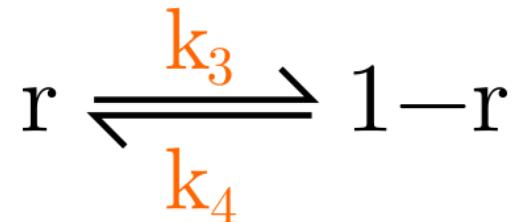
I_{Kr} channels “activate” and “inactivate”



$$\frac{da}{dt} = \frac{a_\infty - a}{\tau_a}$$

$$\tau_a = 1/(k_1 + k_2)$$

$$a_\infty = k_1 / (k_1 + k_2)$$



$$\frac{dr}{dt} = \frac{r_\infty - r}{\tau_r}$$

$$\tau_r = 1/(k_4 + k_3)$$

$$r_\infty = k_4 / (k_4 + k_3)$$

$$k_i = a_i \exp(b_i V)$$

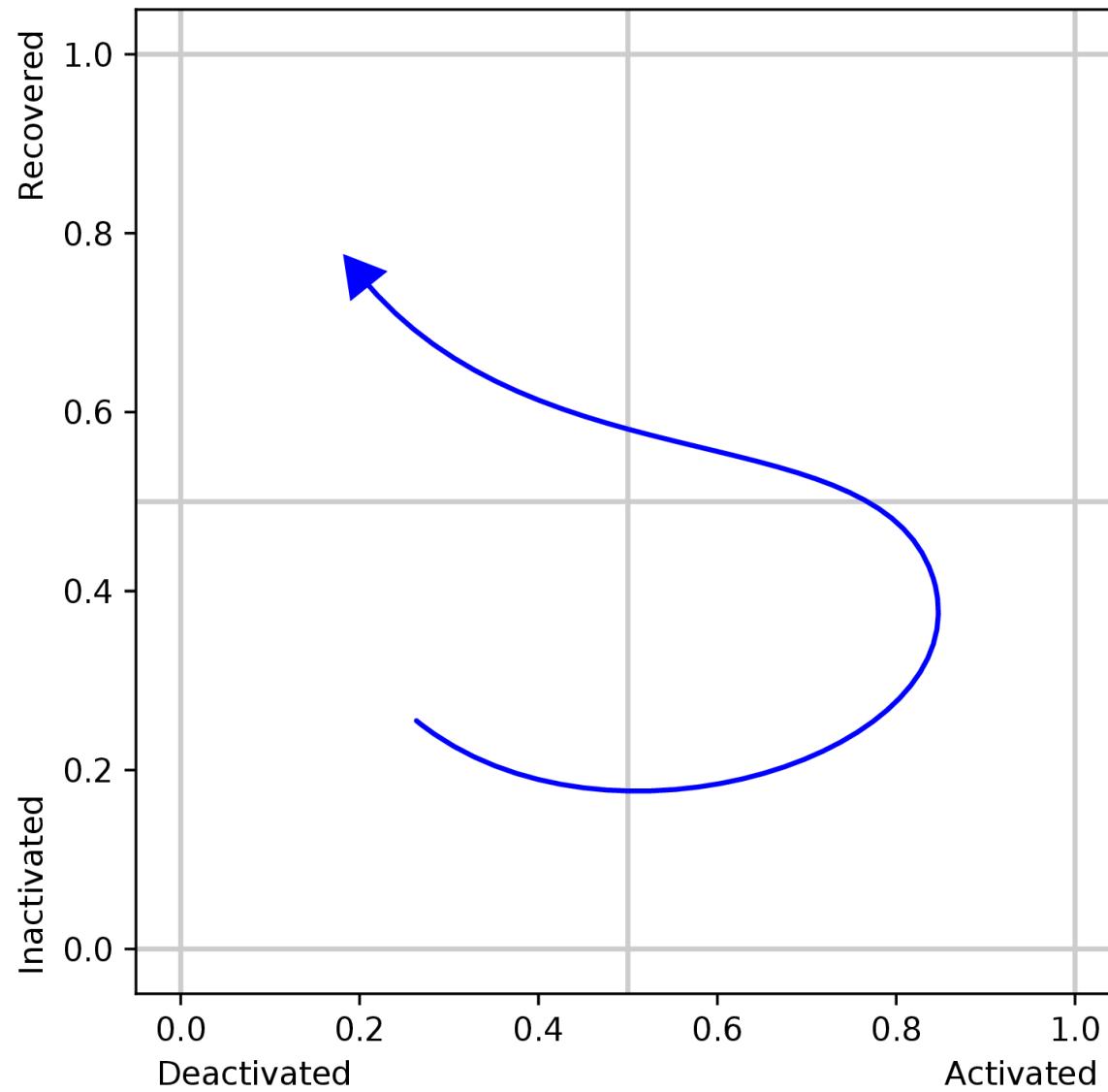
$$I(V,t) = g \times a \times r \times (V - E)$$

How do we parametrise this model?

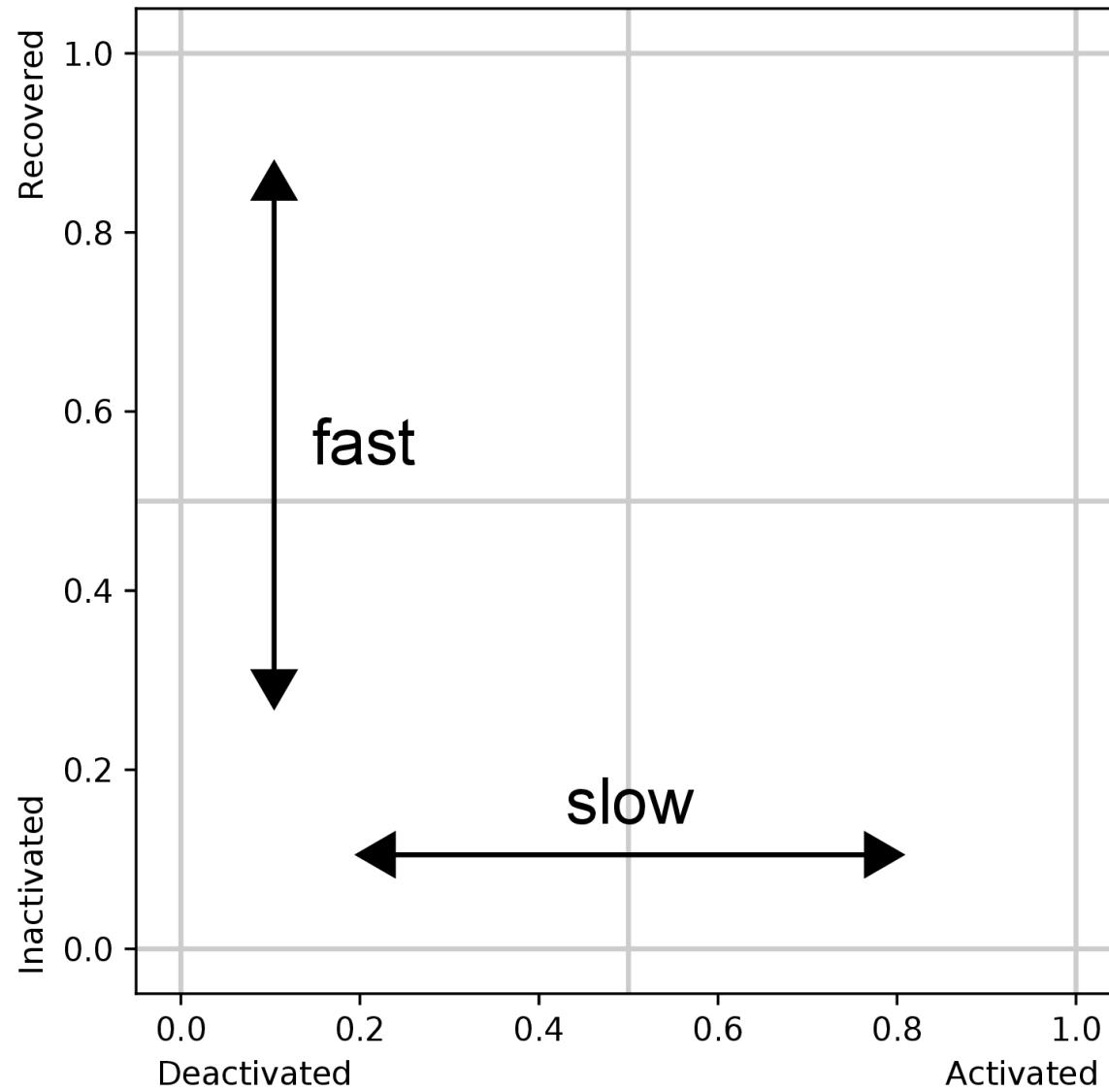
The first method (~1950) is very “involved”:

it requires knowledge of the system
and assumes that the two time constants
have very different scales

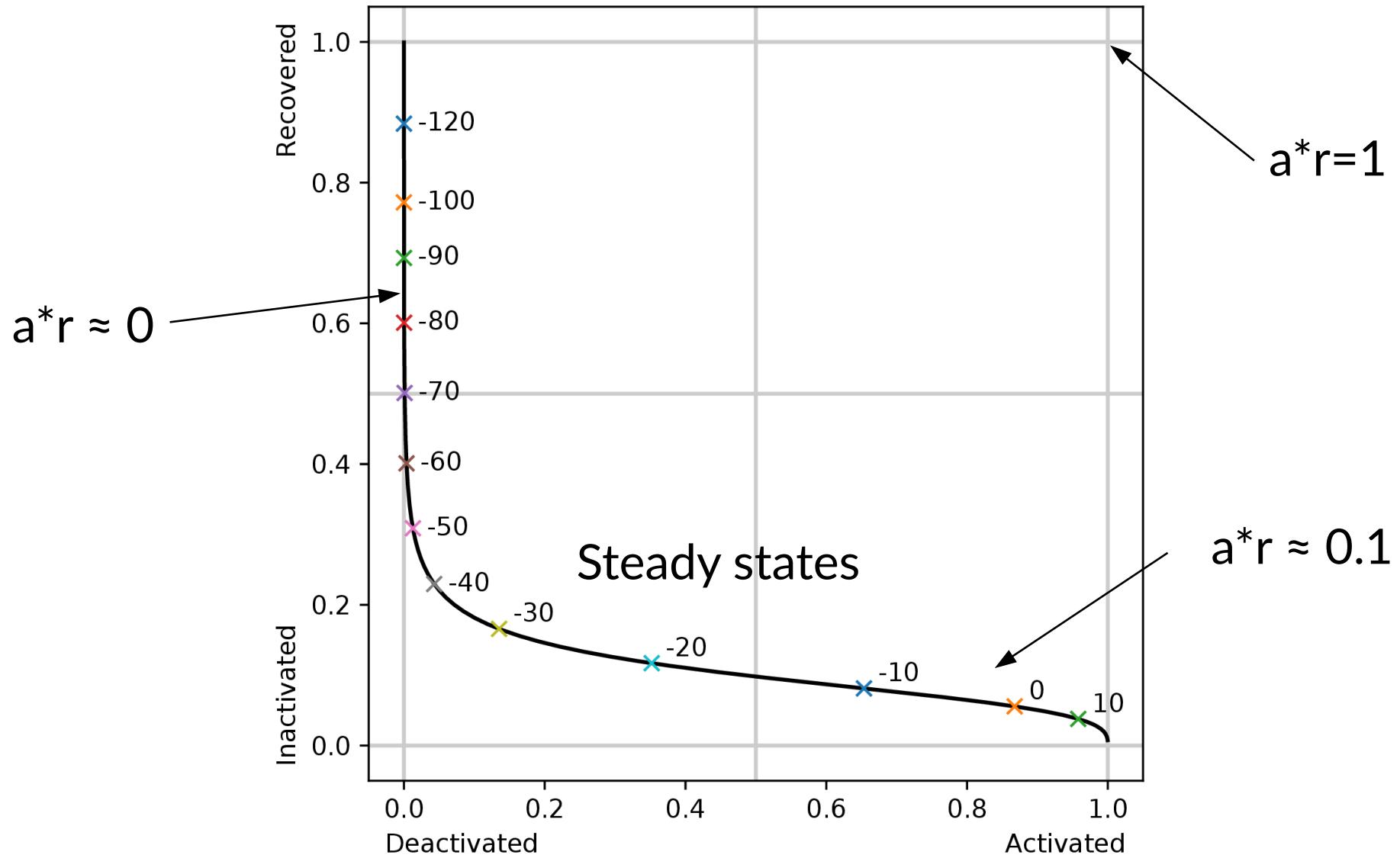
We can plot the state on a phase plane



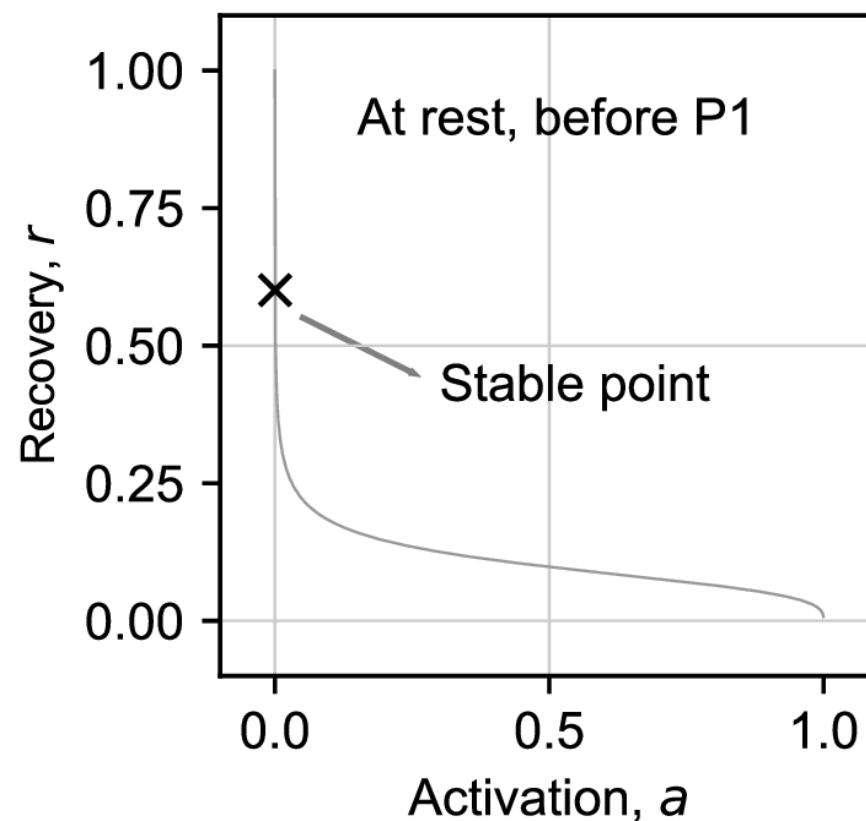
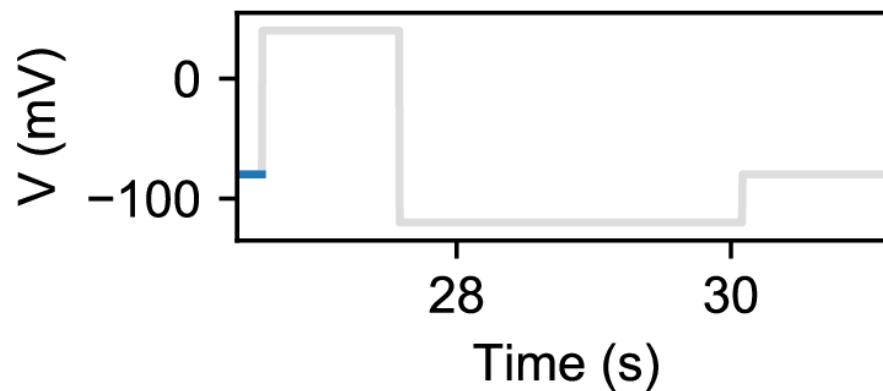
Activation = slow, Inactivation = fast



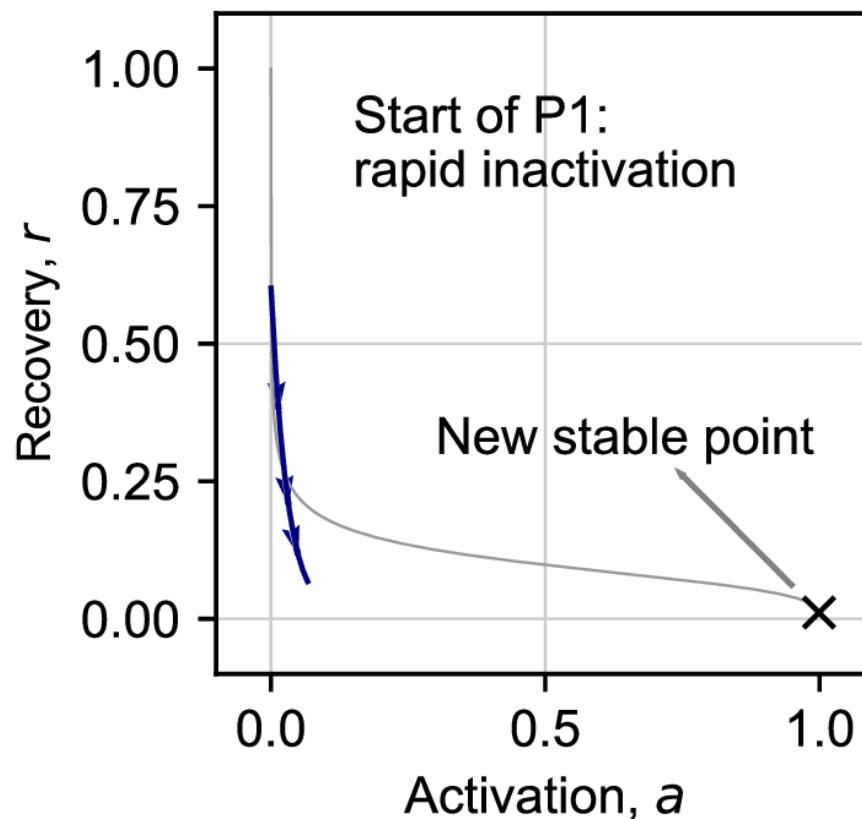
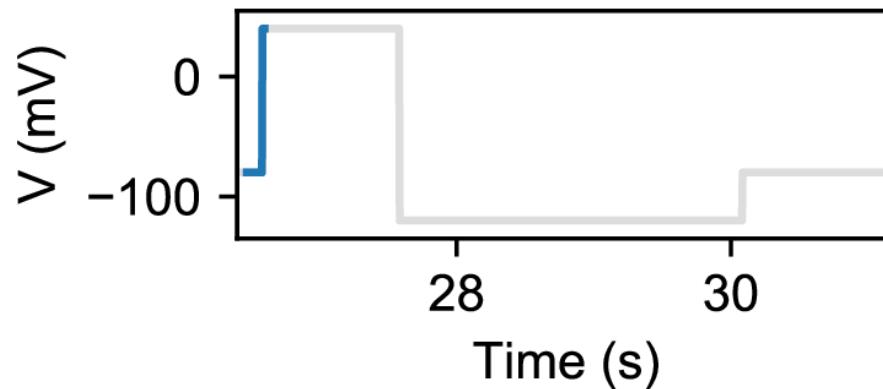
Steady states yield little output



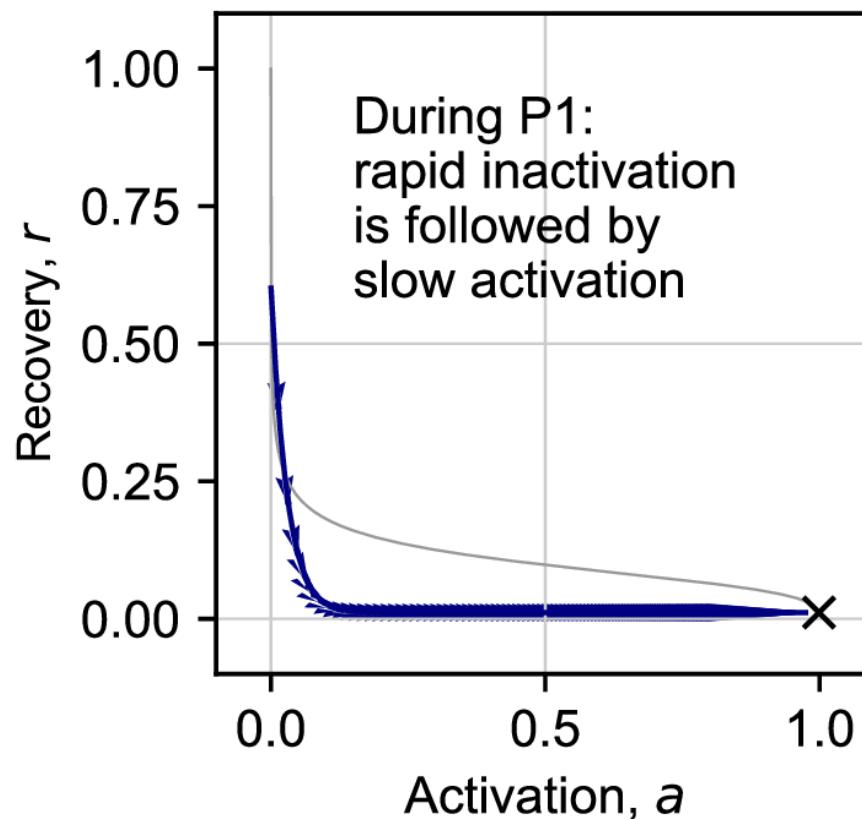
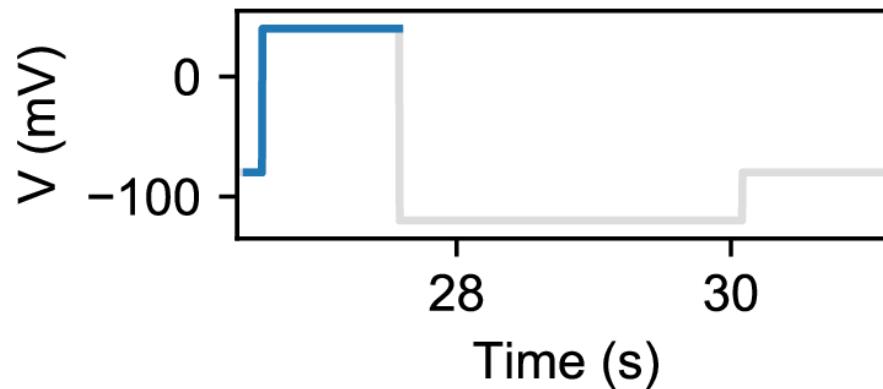
An example step sequence



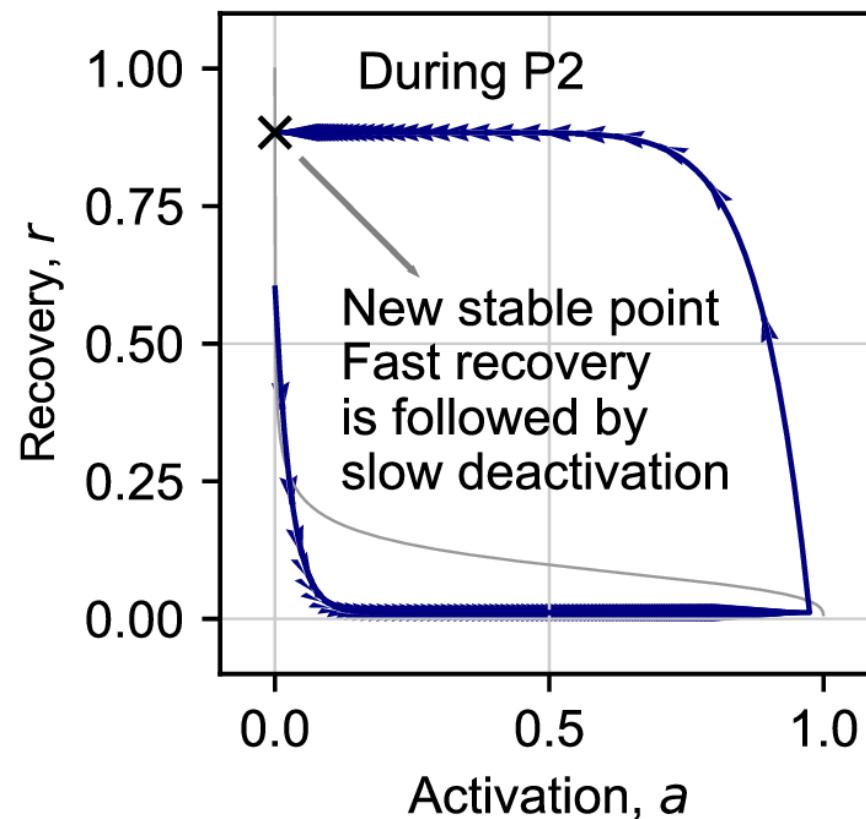
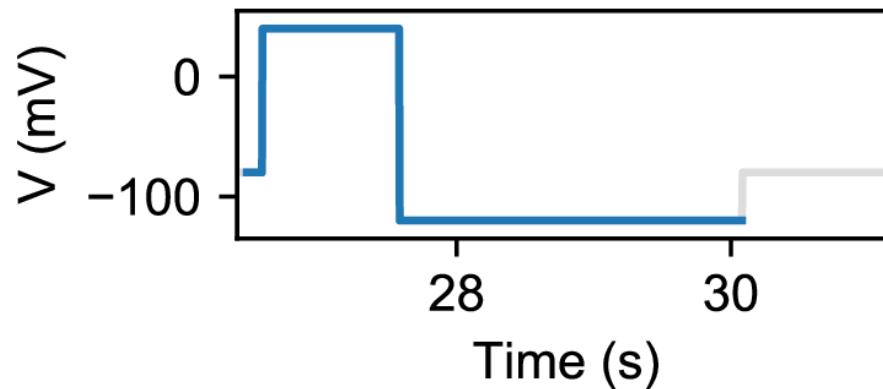
An example step sequence



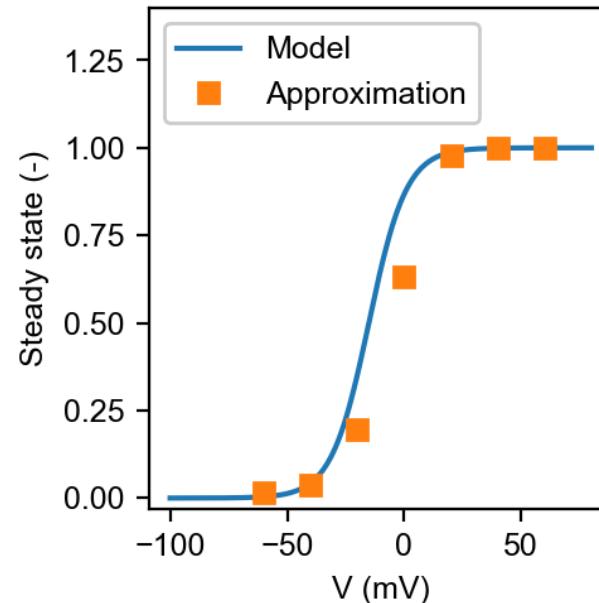
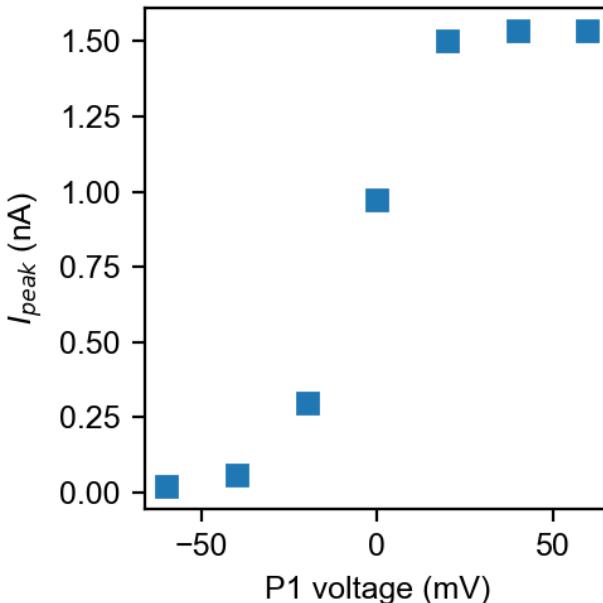
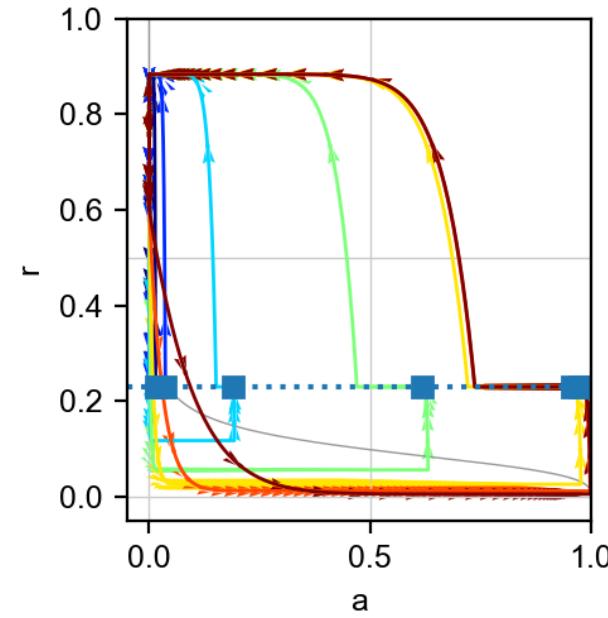
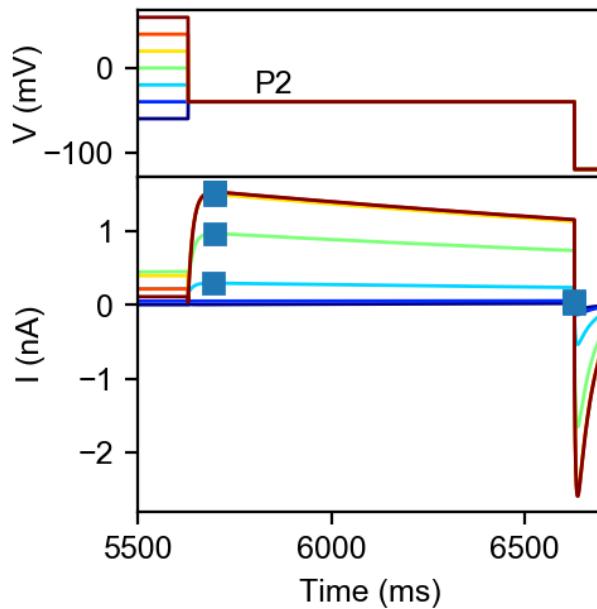
An example step sequence



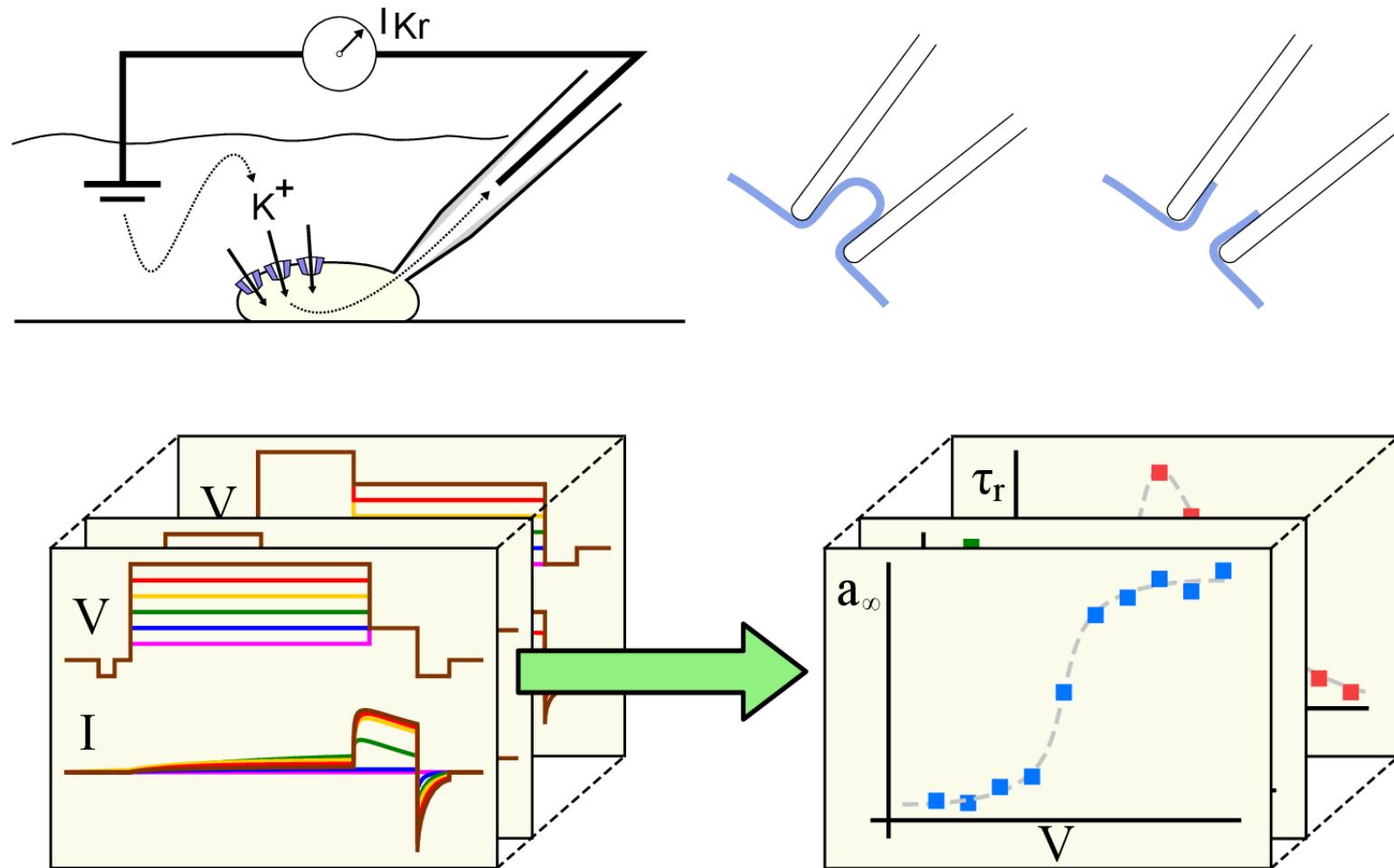
An example step sequence



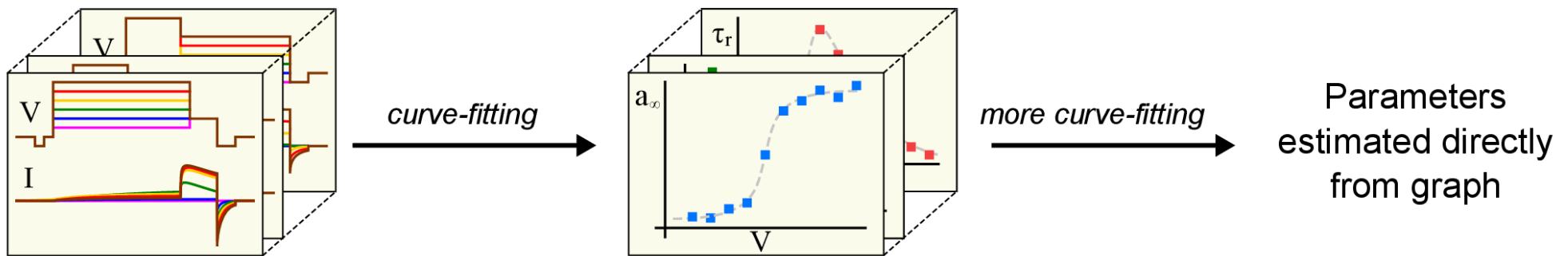
Example: Steady-states of activation



Similar protocols yield r_∞ , τ_a , τ_r



“Method 1”



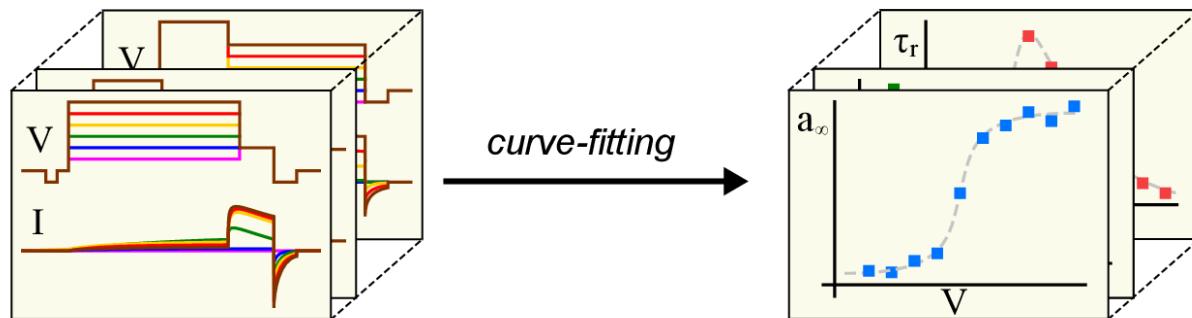
One more curve fitting step
gives us our 9 parameters

Method 1

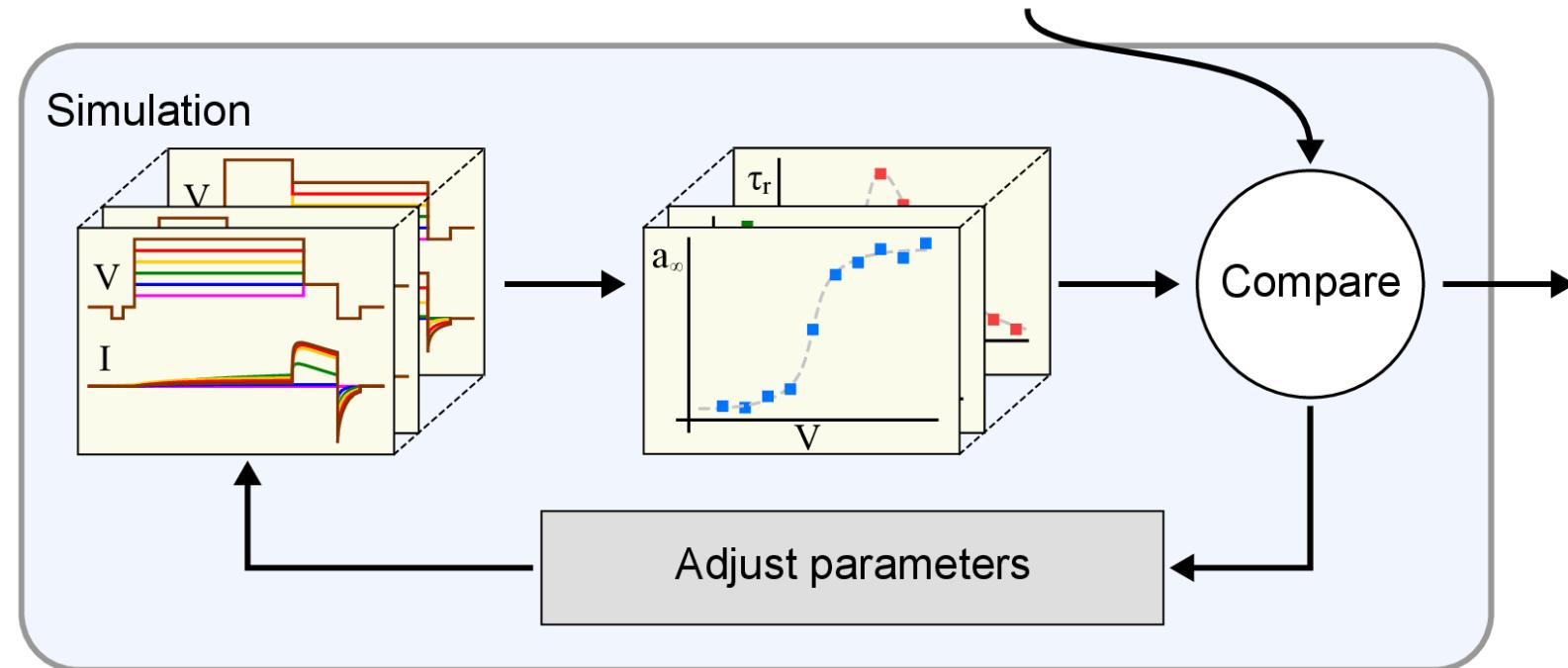
- Downsides:
 - Limited accuracy
 - Model-dependent
 - Multiple experiments needed
- Nevertheless, these protocols became established

“Method 2”

Experiment



Simulation



Parameters
estimated through
simulation and
post-processing

Method 2

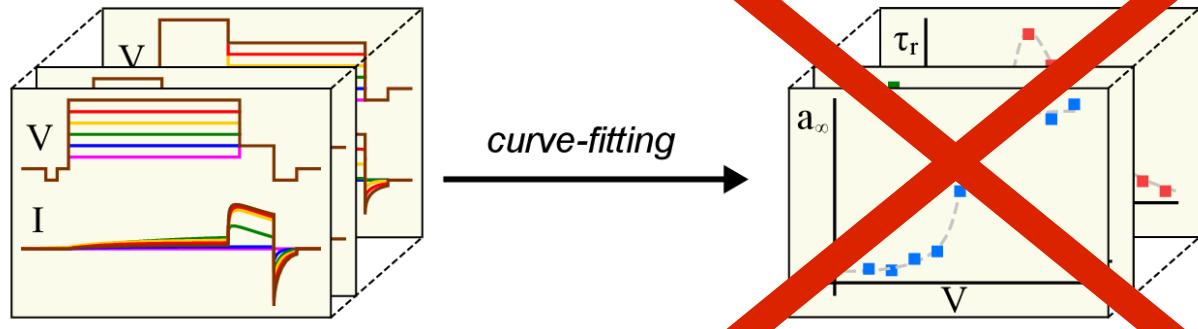
- Method 2 was invented by modellers
 - Removes some inaccuracies*
 - Allows different models to be fit to same data**
- Crucially, allows modellers to do what they want using digitised data from the literature

* if a good fit can be made

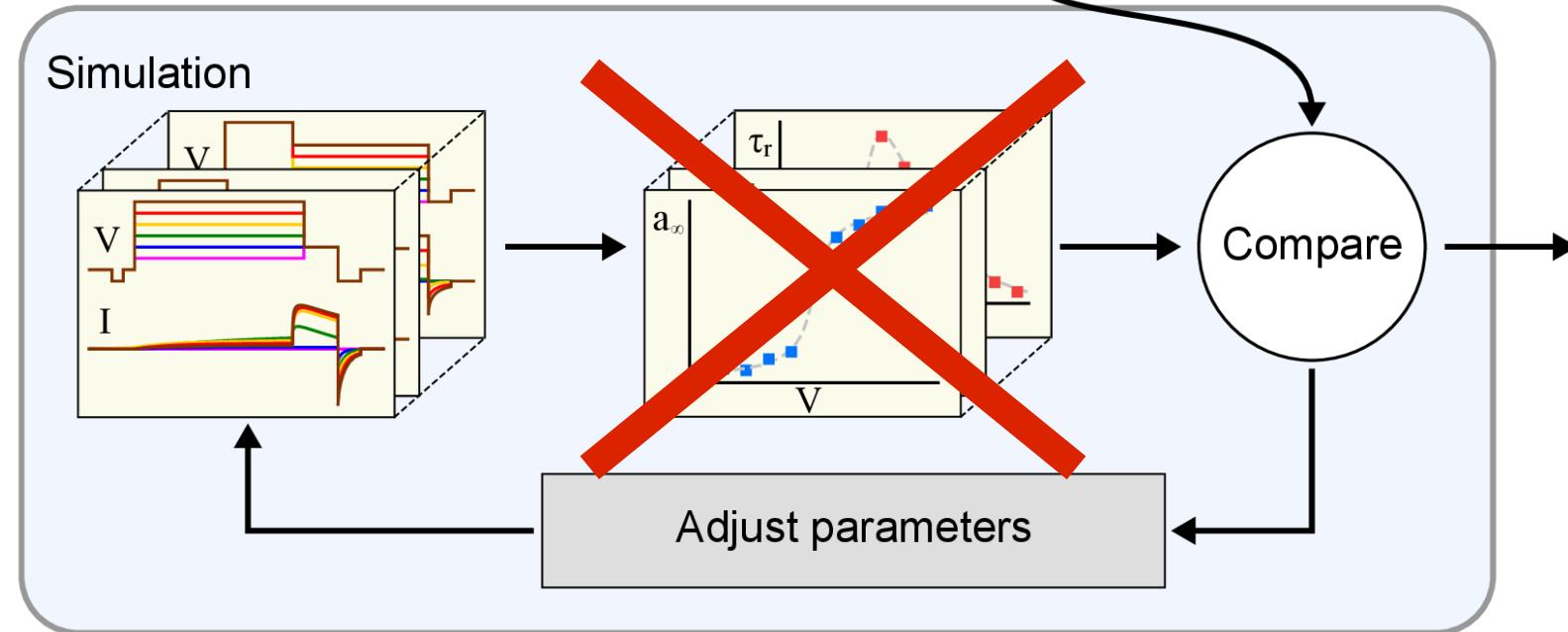
** if the data contains enough information

Method 2

Experiment

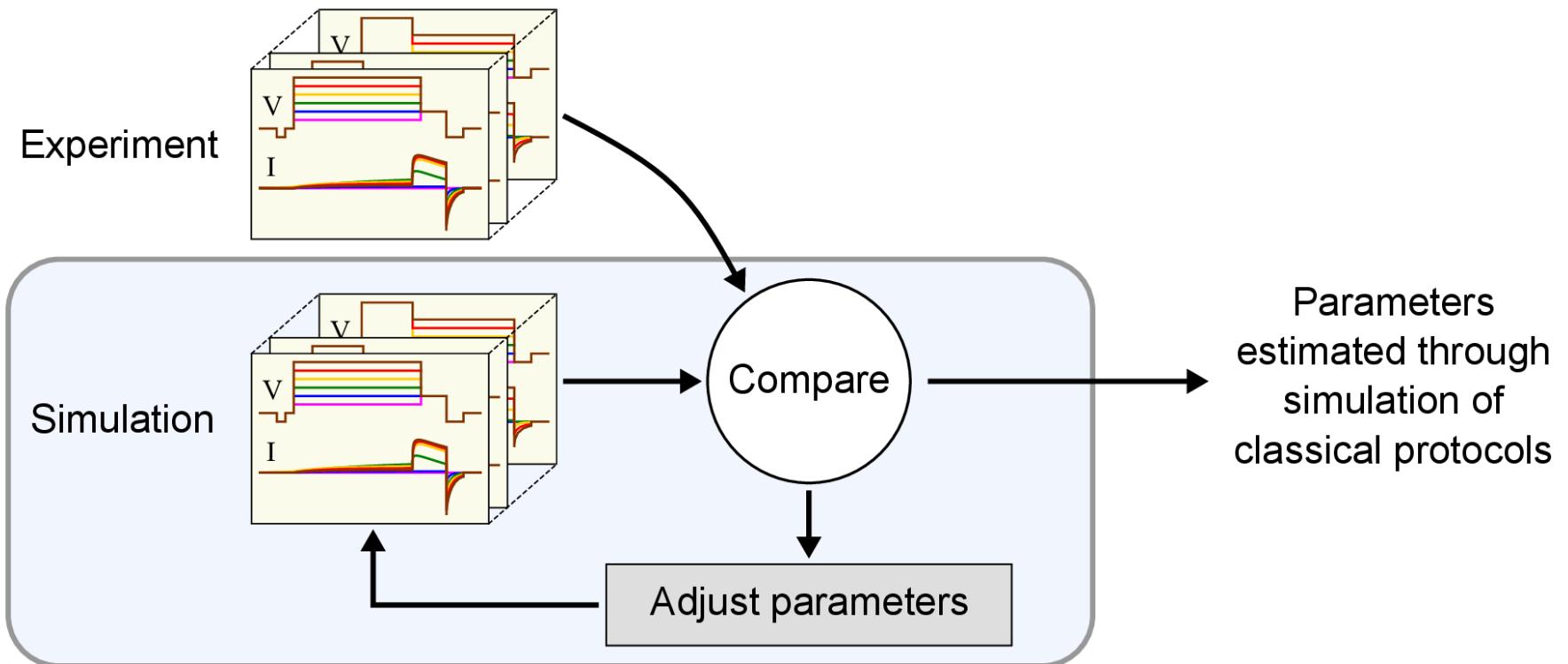


Simulation

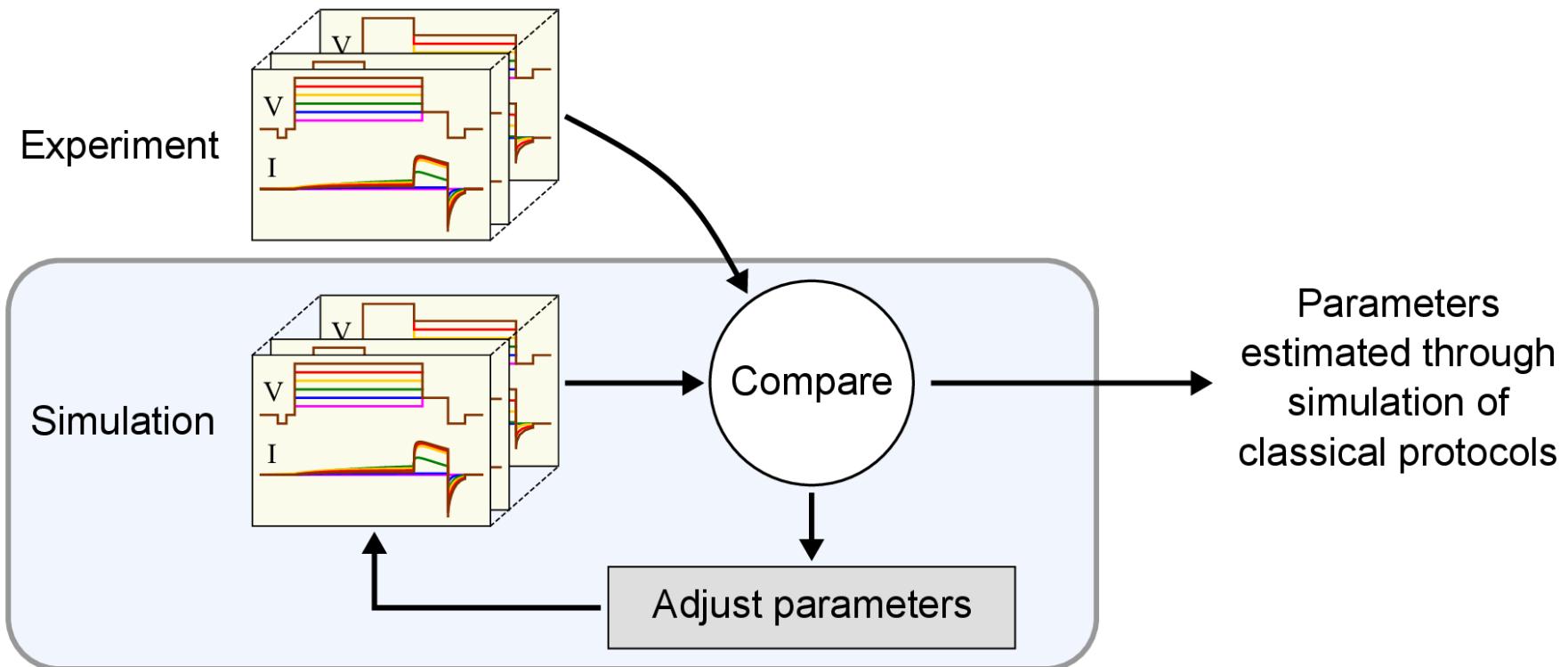


Parameters
estimated through
simulation and
post-processing

“Method 3”



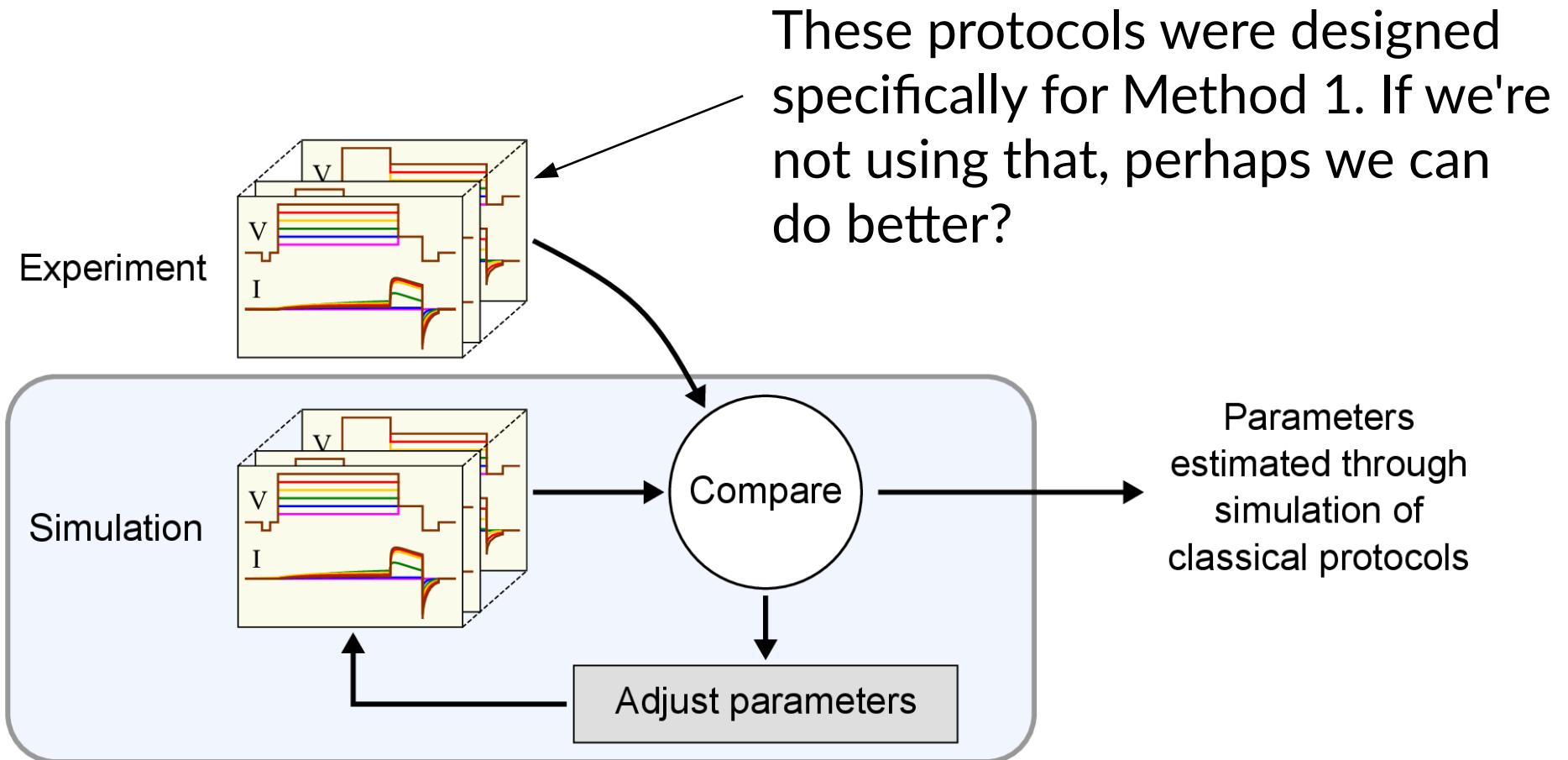
Method 3



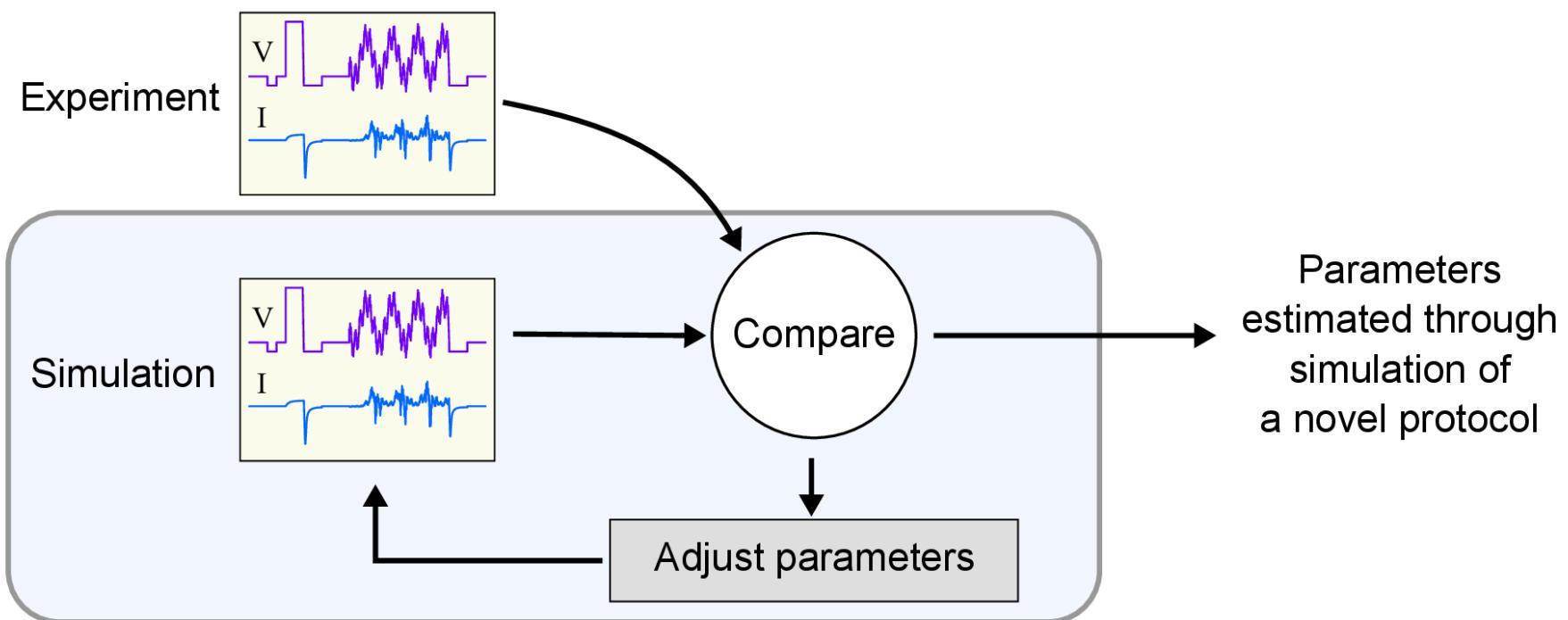
30 years!

*Balser et al. 1990, Biophys J; Beaumont 1993, Mathematical Biosciences; Willms et al. 1999, J of Computational Neurosciences; and Lee et al. 2006, J of Theoretical Biology.

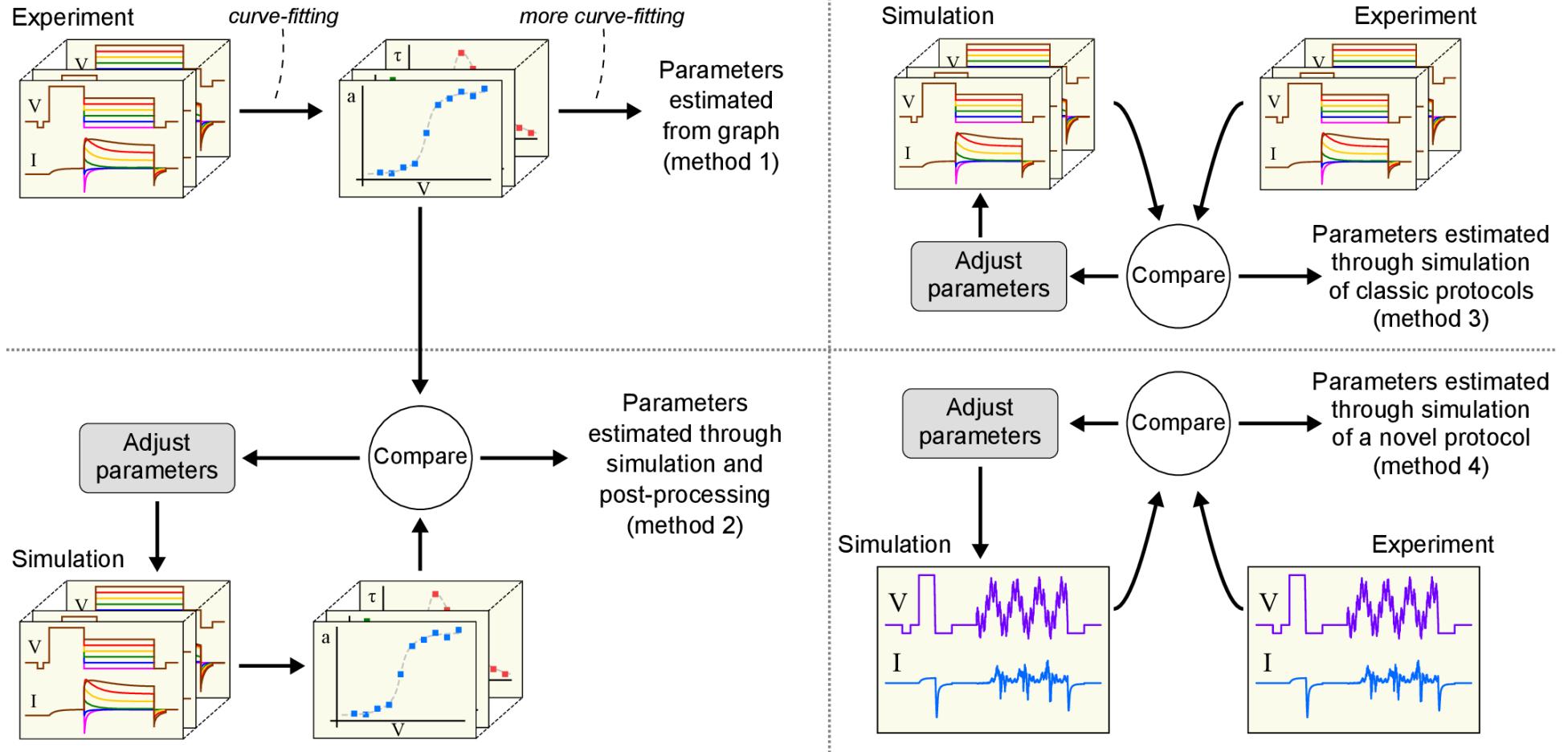
Method 3



“Method 4”



Four ways



Comparing the methods

- We measured 9 cells with old and new protocols

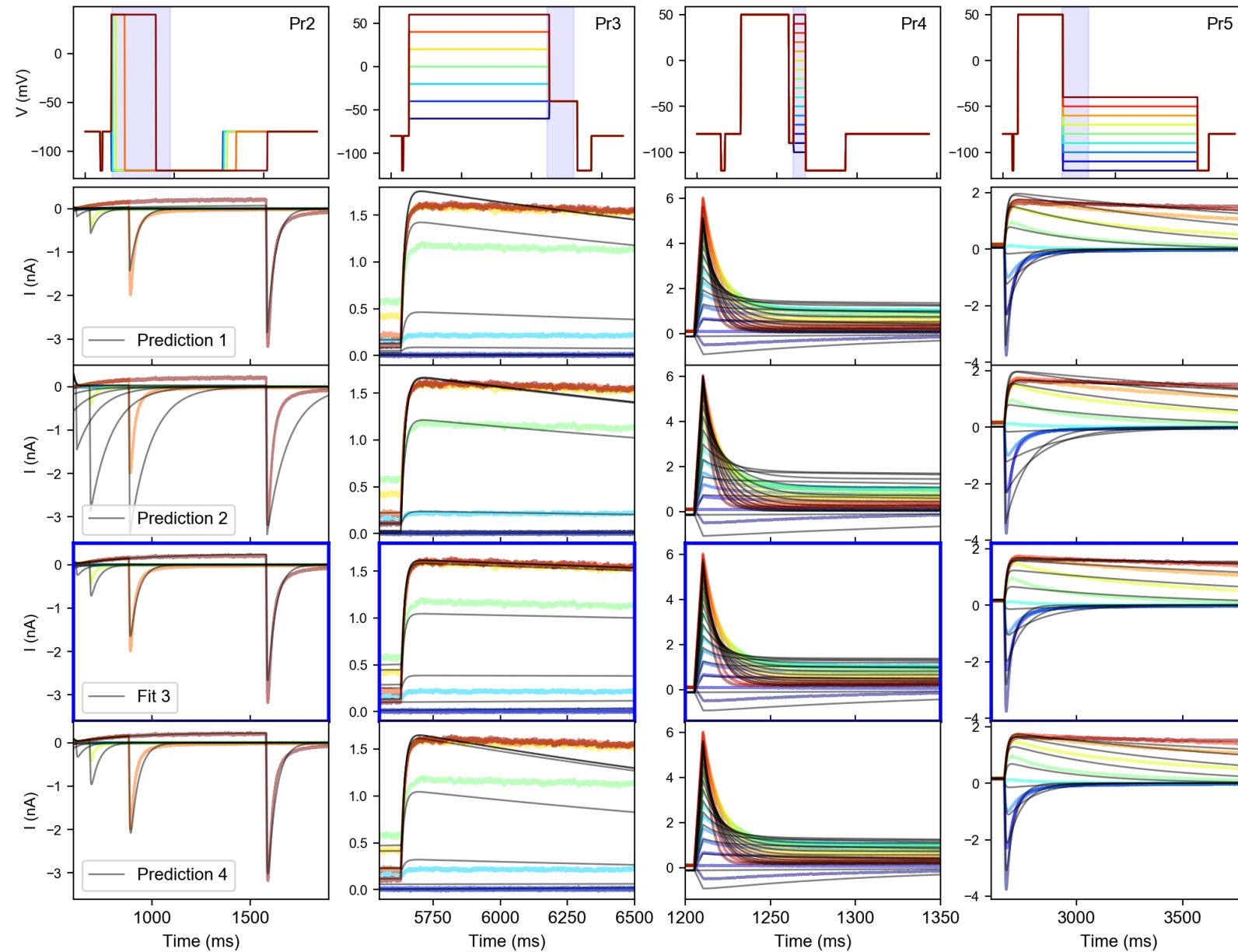
Comparing the methods

- We measured 9 cells with old and new protocols
- We fit each cell using each method
 - For methods 2, 3, 4 we used **CMA-ES**,
and a **least-squares criterion**

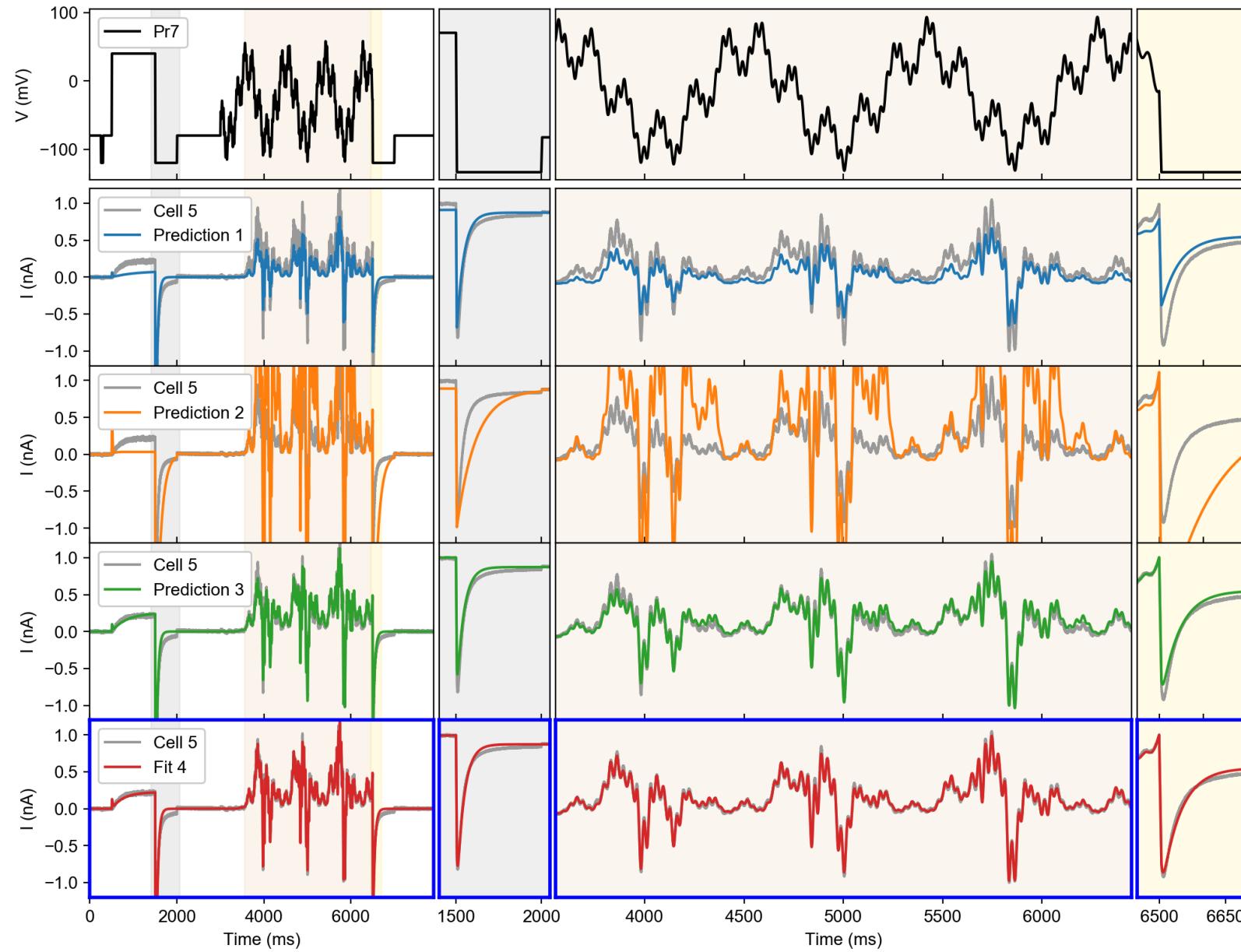
Comparing the methods

- We measured 9 cells with old and new protocols
- We fit each cell using each method
 - For methods 2, 3, 4 we used **CMA-ES**,
and a **least-squares criterion**
- We made predictions with each model
 - Cross-validation: response to the old and new protocols
 - Independent validation: response to an unused protocol

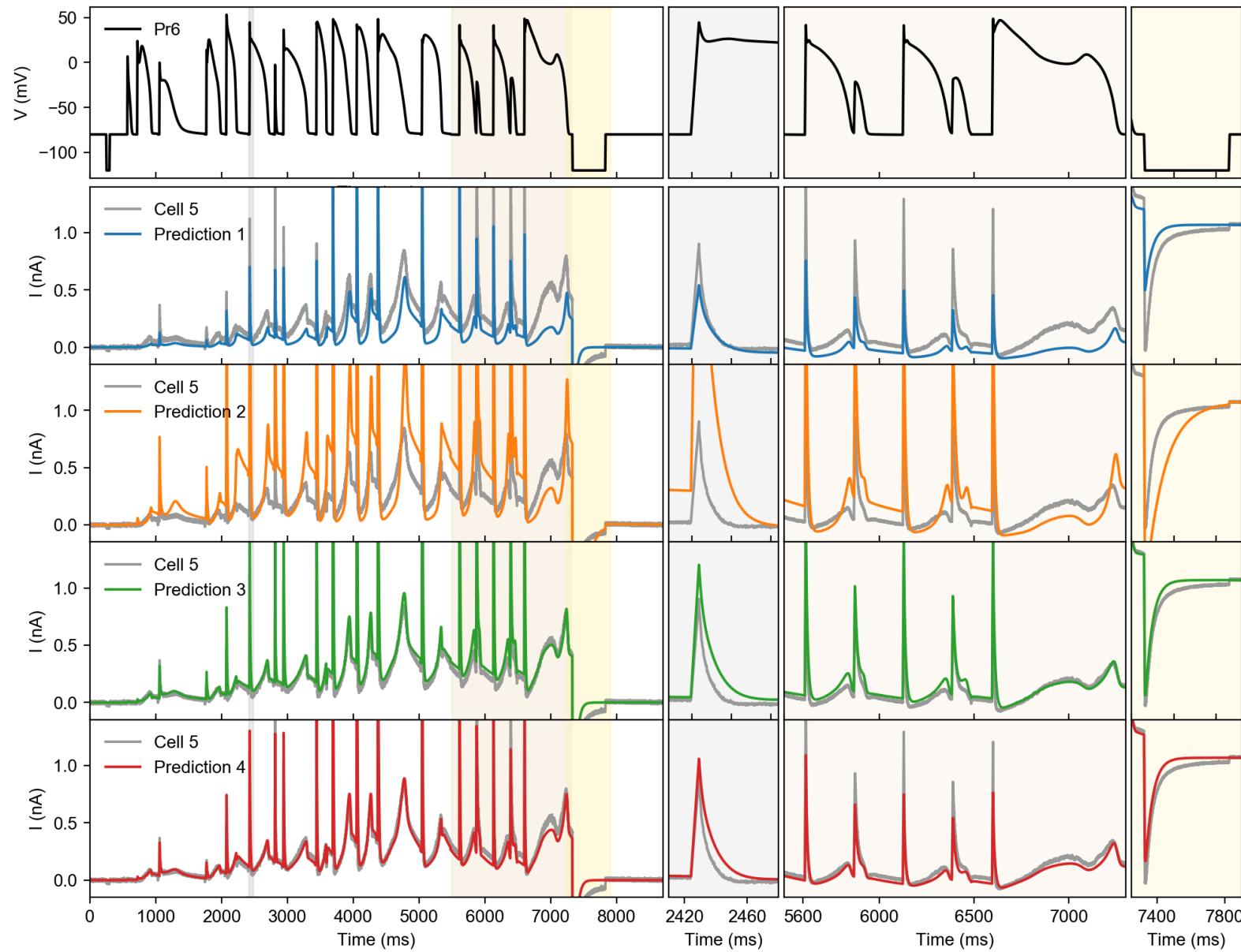
Predicted conventional protocol responses



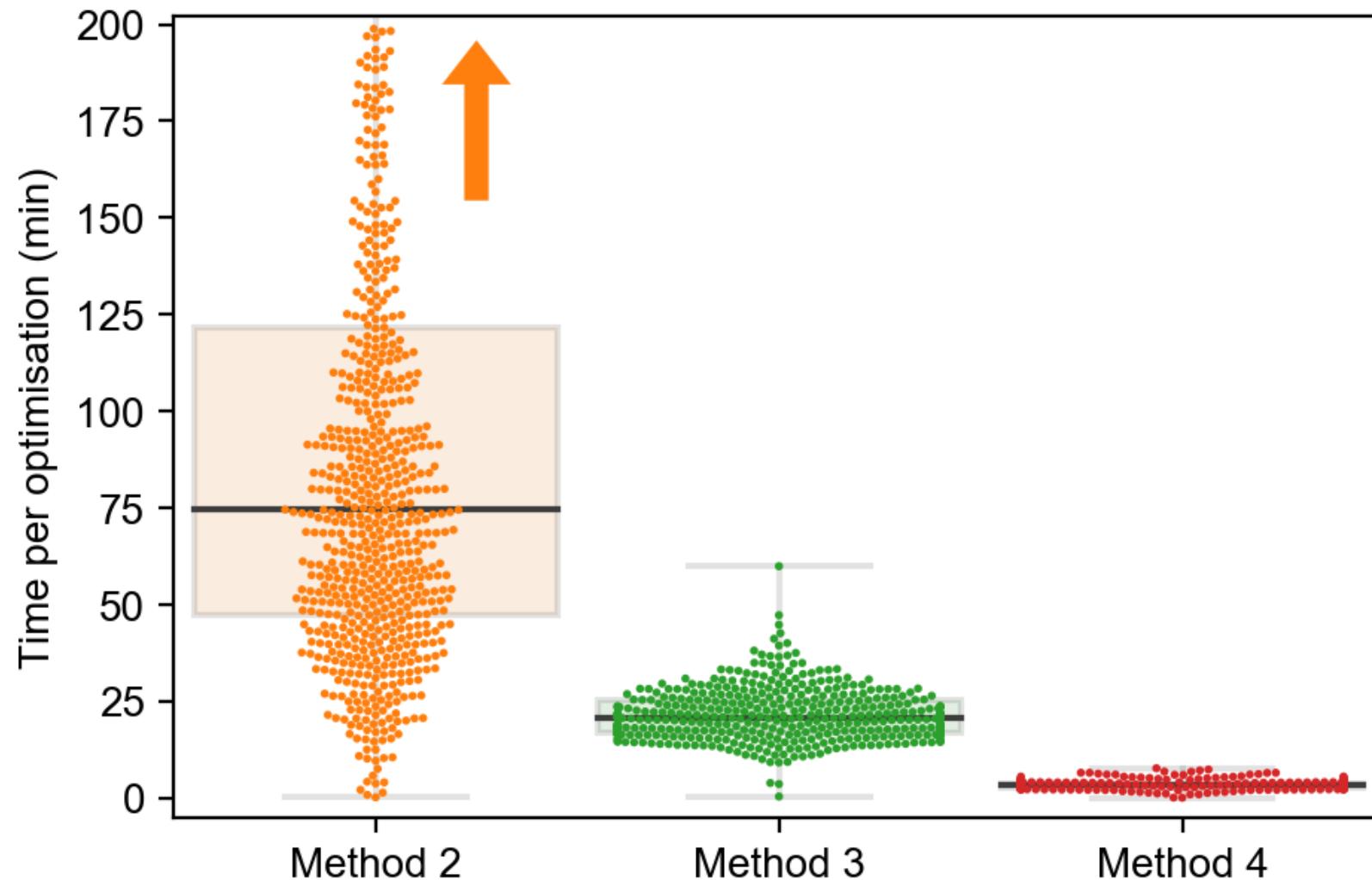
Predicted novel protocol response



What really matters: Response to AP

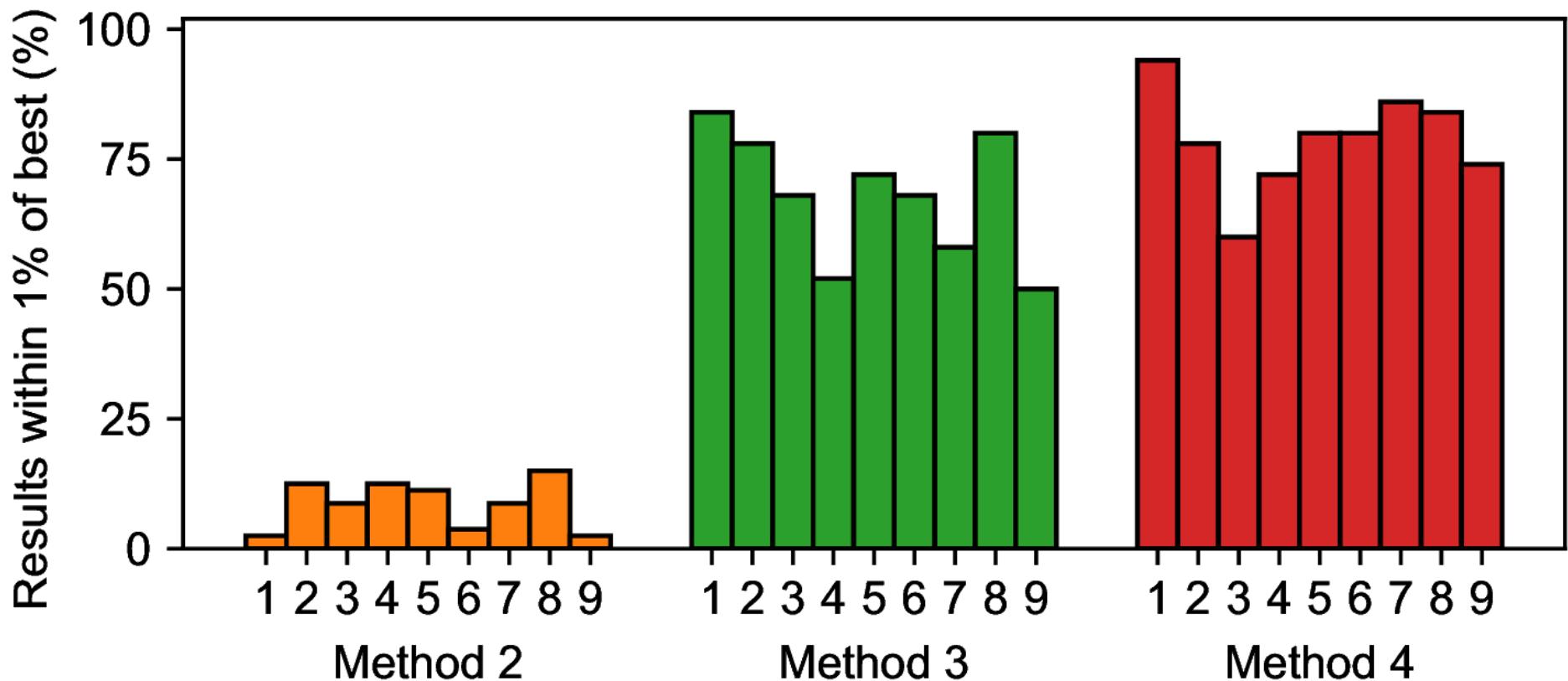


Analysis duration



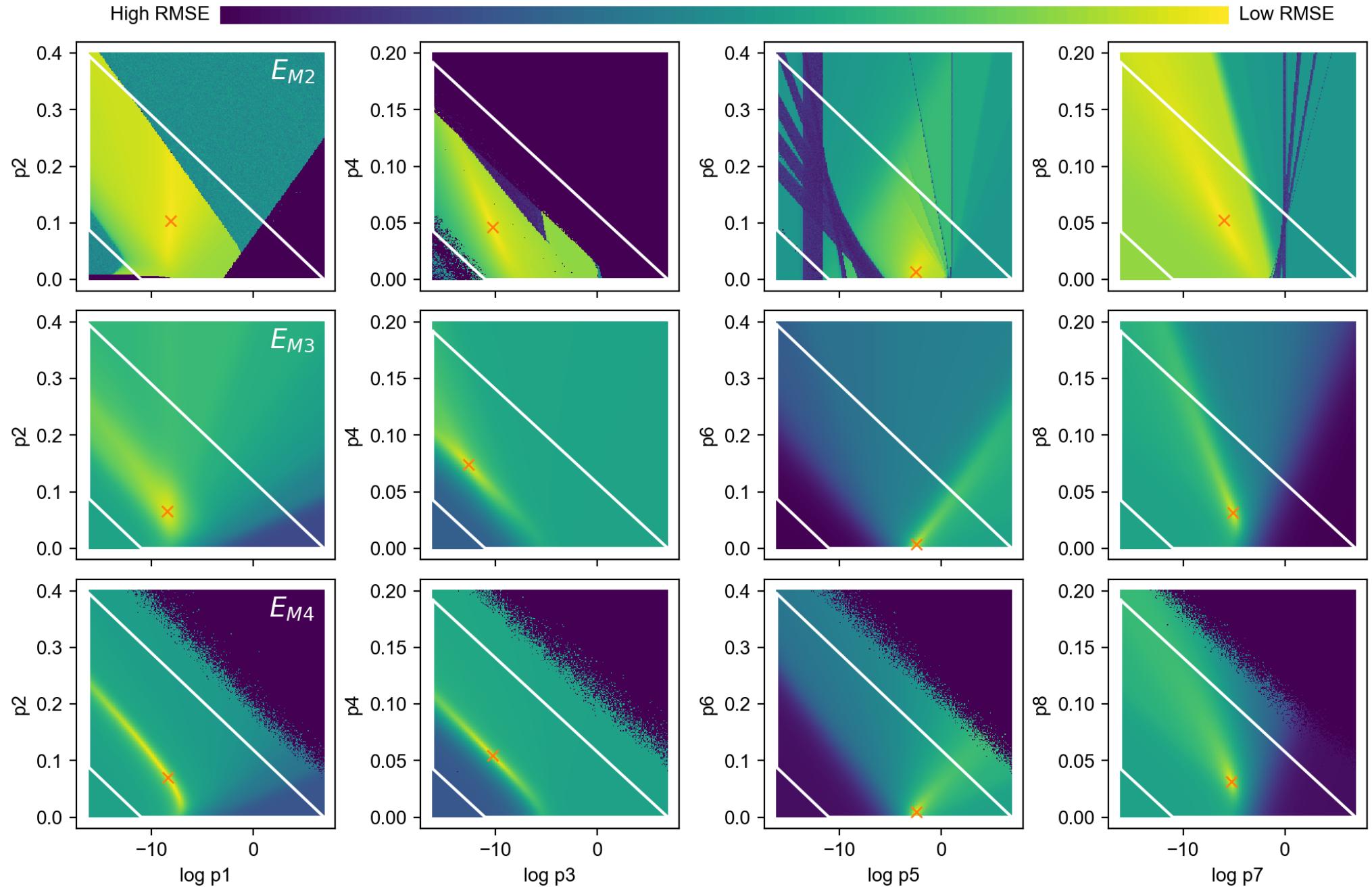
Robustness

If we run the method repeatedly,
how often do we get the best result?



Why is Method 2 so bad?

Surfaces for methods 2, 3, and 4



The post-processing method
that is **part of the error function**
in Method 2 contains lots of **branching**
and sometimes simply **fails**

This creates **artificial problems**
for your optimiser
(which no method can fix)

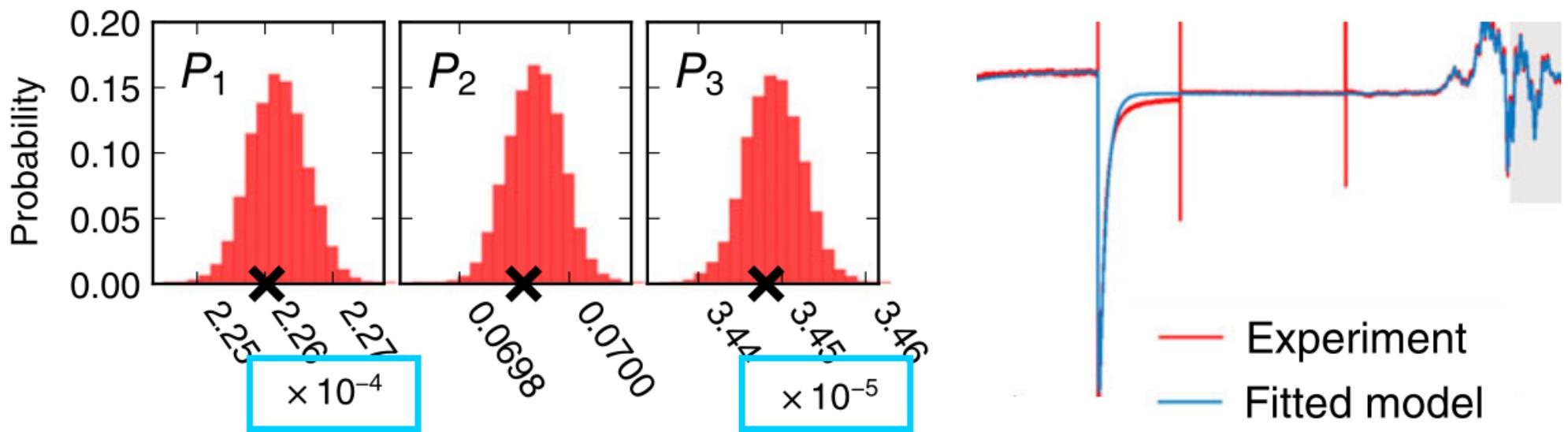
Take-home messages

- Avoid **non-smooth algorithms** in scores/likelihoods.
- Understanding the origins of the data can solve problems that algorithms cannot.
- Experimental design is a crucial inference technique.

Open questions

We would like to do U.Q.

- First attempt:
 - $I_{\text{measured}} = I_{\text{real}} + N(0, \sigma)$
- This gives us **very narrow posteriors**
 - (even though we're very uncertain)



We would like to do U.Q.

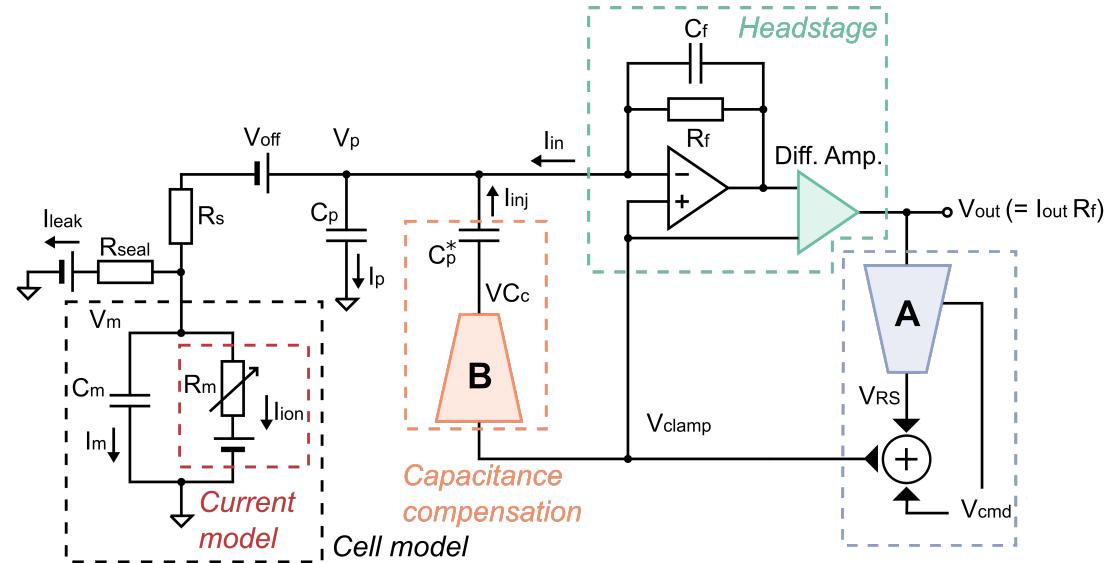
- First attempt:
 - $I_{\text{measured}} = I_{\text{real}} + N(0, \sigma)$
- This gives us **very narrow posteriors**
 - (even though we're very uncertain)
- We would like a method that says:

“These parameters are likely, given the data”
- But we have one that says:
 - *“This is the likelihood that your (obviously correlated) residuals were the result of drawing 10^6 points from a normal distribution”*

Experimental artefacts?

- Second attempt:

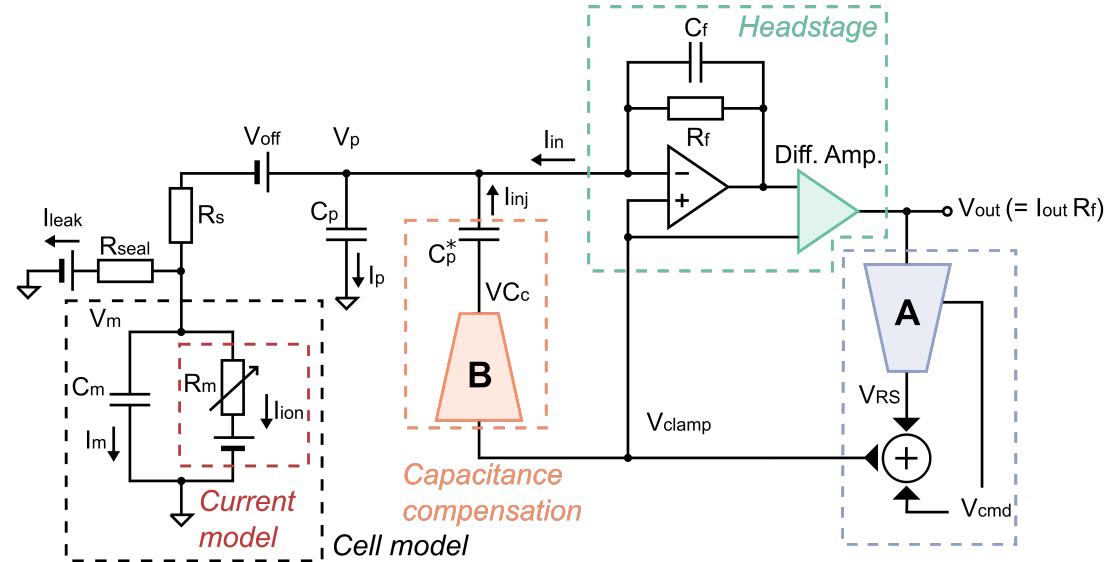
- $V_{\text{applied}} = f(V_{\text{intended}})$
- $I_{\text{measured}} = g(I_{\text{real}}) + N(0, \sigma)$



Experimental artefacts?

- Second attempt:

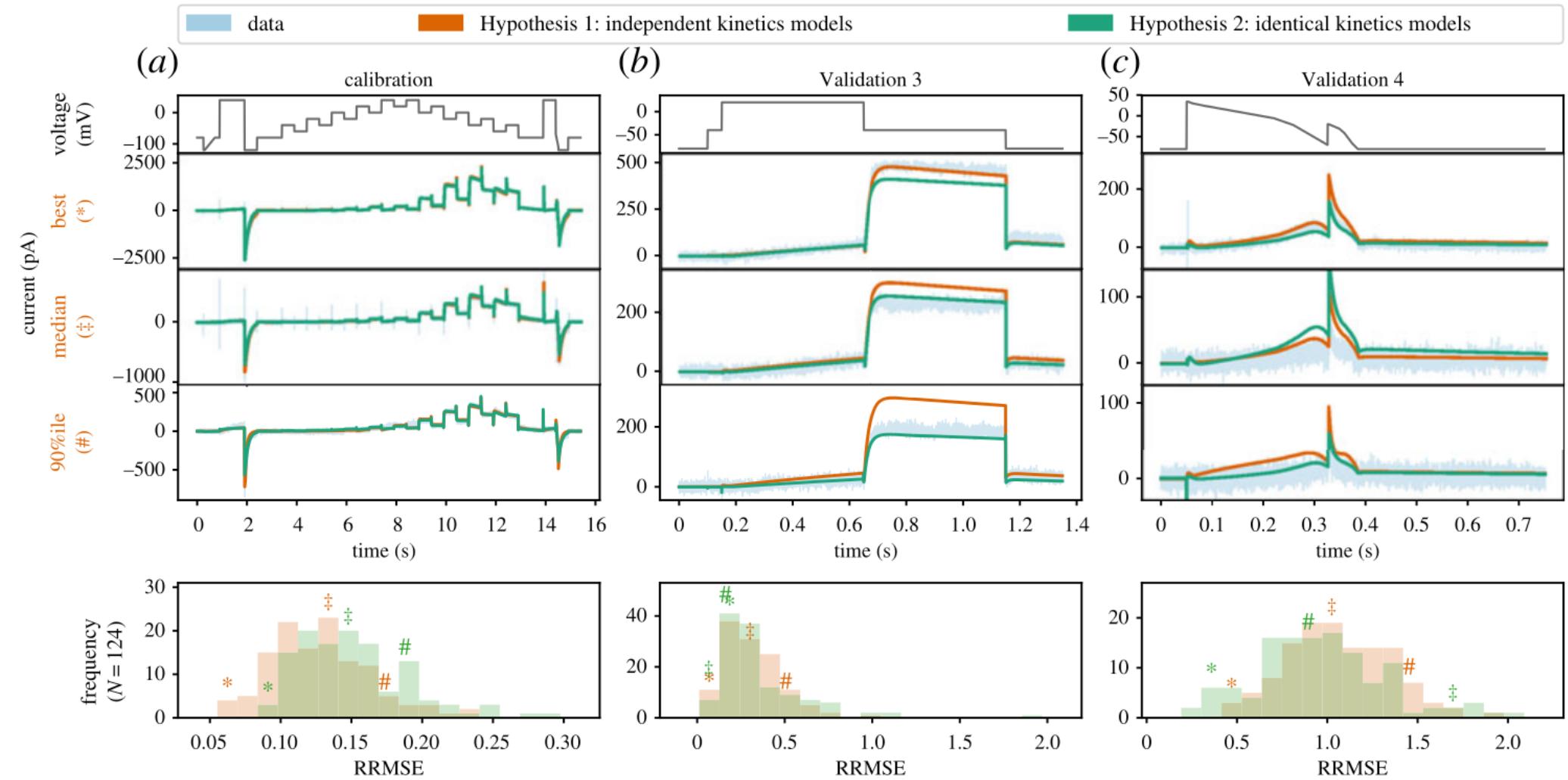
- $V_{\text{applied}} = f(V_{\text{intended}})$
- $I_{\text{measured}} = g(I_{\text{real}}) + N(0, \sigma)$



- Measured and fit 124 cells

- H1: Independent cells, no artefacts
(124 cell models, 0 error models, 124 fits)
- H2: Identical cells, independent artefacts
(1 cell model, 124 error models, one single huge fit)

H2: Worse fits, better predictions



But still discrepancy!

Are there UQ techniques
to somehow explore this discrepancy?

Or do we simply need a better model?
(or score / likelihood)

Can we use statistical techniques
to suggest better models/scores/likelihoods ?

Thank you!