(scalable, flexible, principled)

Neural simulation-based supernova Ia cosmology

SICRET: 2209.06733

SIDE-real: 2403.07871

SimSIMS: 2311.15650

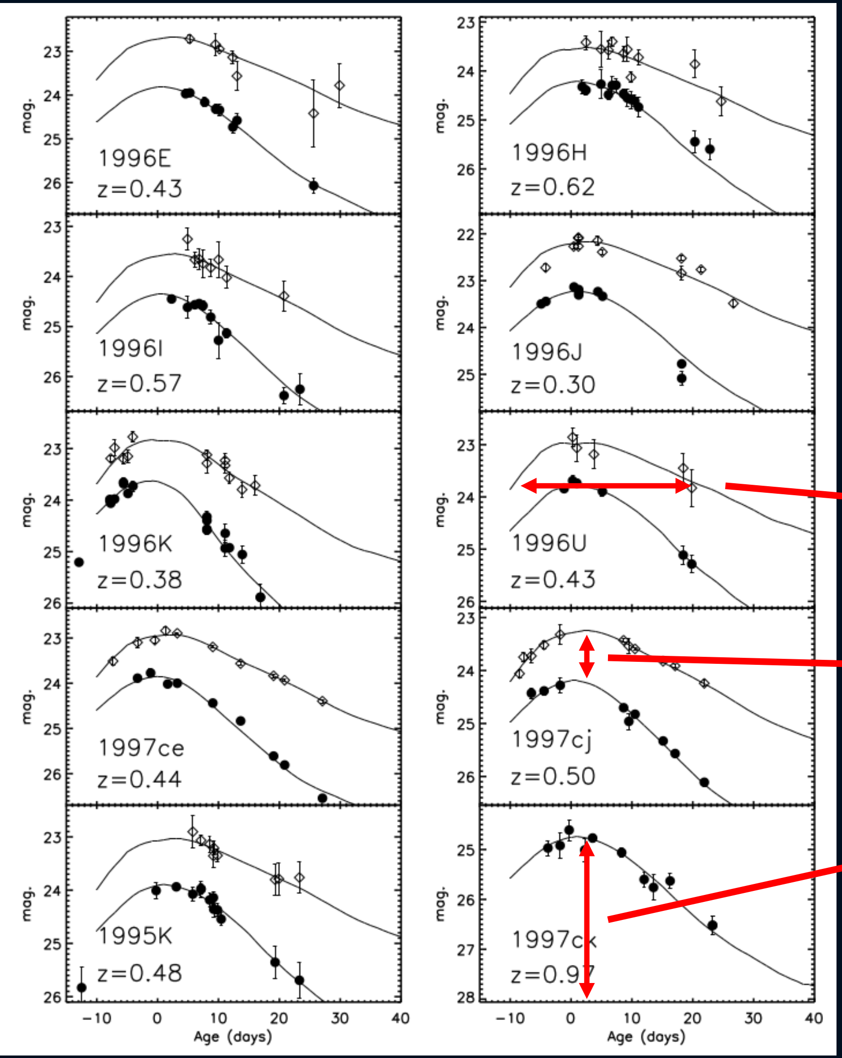
Konstantin Karchev

with Roberto Trotta, Christoph Weniger

MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

SISSA DATASCIENCE
Machine Learning for the Natural Sciences

42 SNæ Ia = $1 million (a Nobel prize)



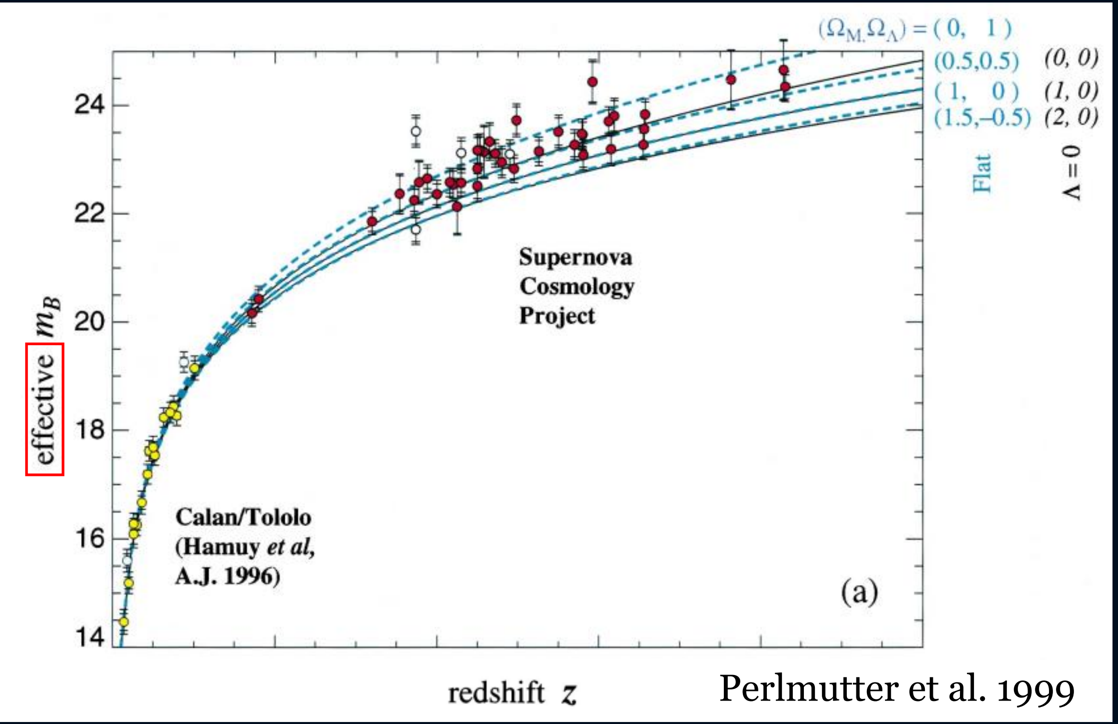
Riess et al. 1999

hand-crafted summaries

$x_1^s \pm \sigma_{x_1}^s$
“stretch”

$c^s \pm \sigma_c^s$
“colour”

$m^s \pm \sigma_m^s$
brightness

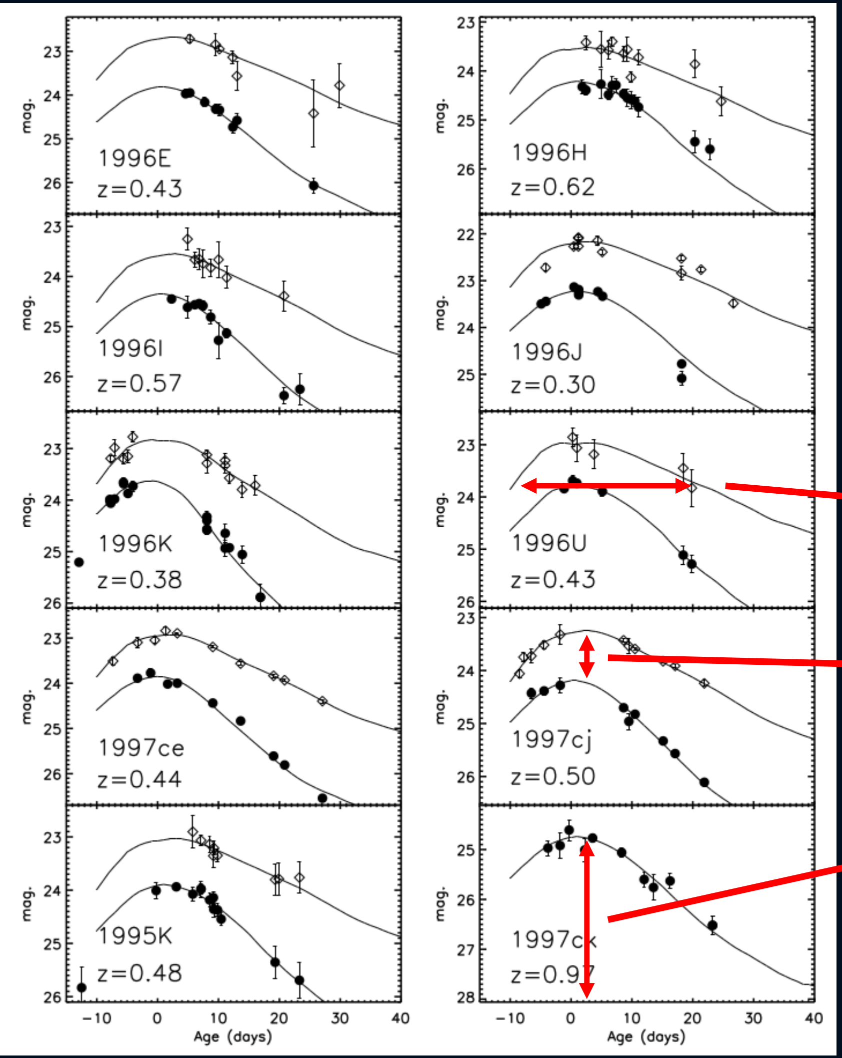


Perlmutter et al. 1999

$$m^s + \alpha x_1^s - \beta c^s = M + \mu(z^s, \mathcal{C}) + \text{"noise"}$$

$\mathcal{C} \pm \sigma_{\mathcal{C}}$
posterior

42 SNæ Ia = $1 million

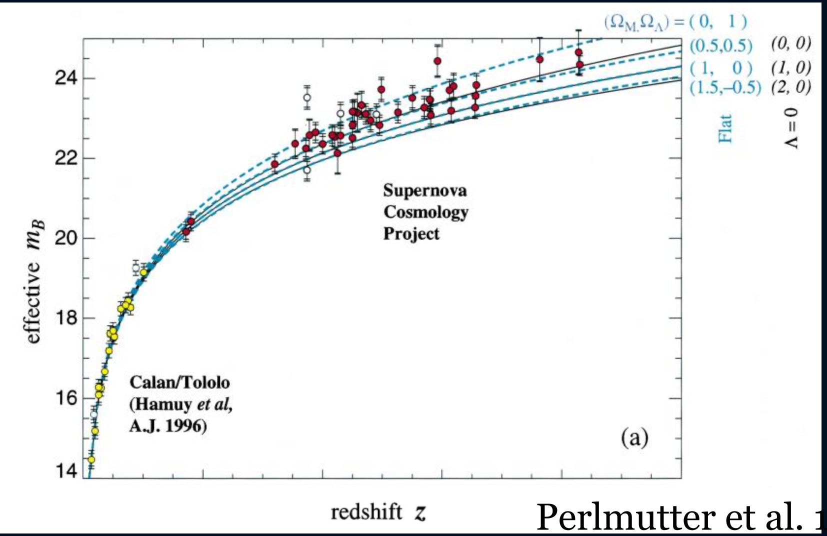


$x_1^s \pm \sigma_{x_1}^s$
“stretch”

$c^s \pm \sigma_c^s$
“colour”

$m^s \pm \sigma_m^s$
brightness

Riess et al. 1999



Perlmutter et al. 1999

$$m^s + \alpha x_1^s - \beta c^s = M + \mu(z^s, \mathcal{C}) + \text{"noise"}$$

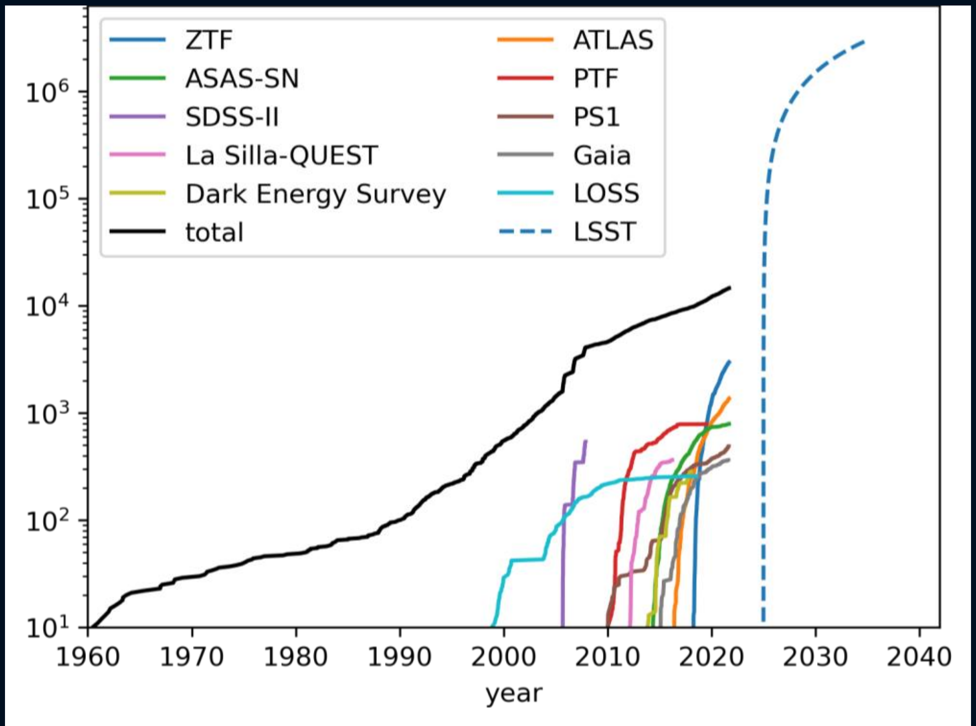
hand-crafted summary statistics:

- extract limited information
- unknown / complicated population distributions
- can lack full physical interpretation (e.g. intrinsic vs. extrinsic colour)

LSST: ~10^6 SNæ = \$1.9 billion

- Noisy light curves with irregular cadence
  - classic fits and GP take time…
- Modelling systematics
  - effect of dust on standardisation
  - correlations with host and evolution
- Photometric redshift
  - complicated uncertainty
- Selection biases
  - current "bias correction" is ad-hoc

cumulative SNae discovered



Neural simulation-based inference

- A new paradigm for marginal inference
  - ➢ model = simulator (arbitrarily* complex)

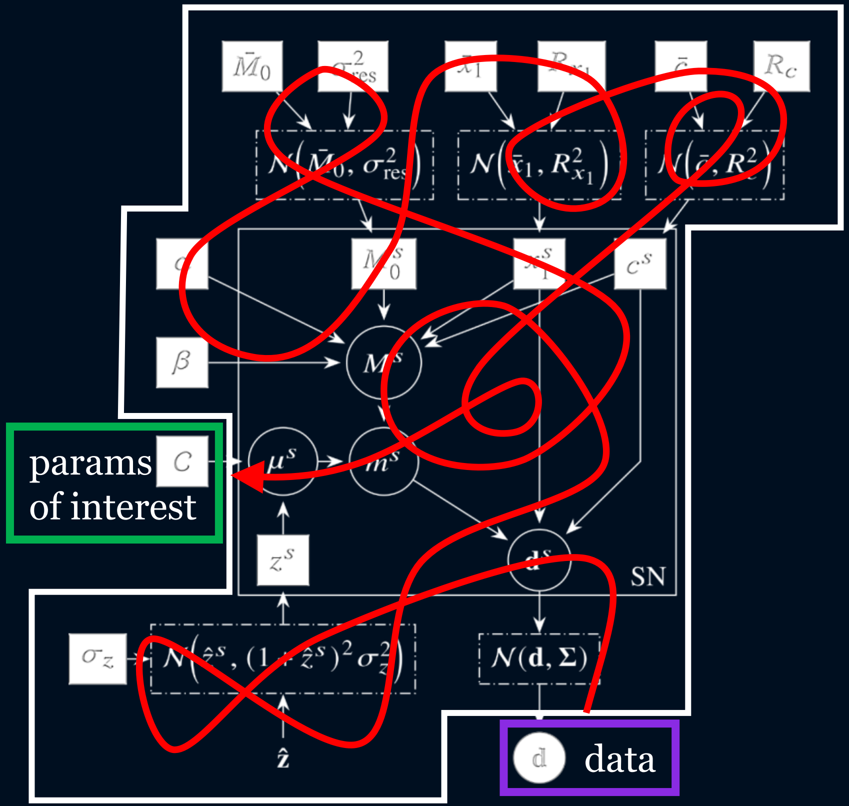
- Leverages neural networks for inference
  - neural likelihood estimation
  - neural posterior estimation
  - neural ratio estimation
  - neural model comparison

$$\boxed{p(\theta|d)} = \boxed{\frac{\boxed{p(d|\theta)}}{\boxed{p(d|M)}}}\, p(\theta)$$

  - ➢ data = NN input (arbitrarily* big and complex)

Simulation-based inference

- The old paradigm: joint likelihood-based inference



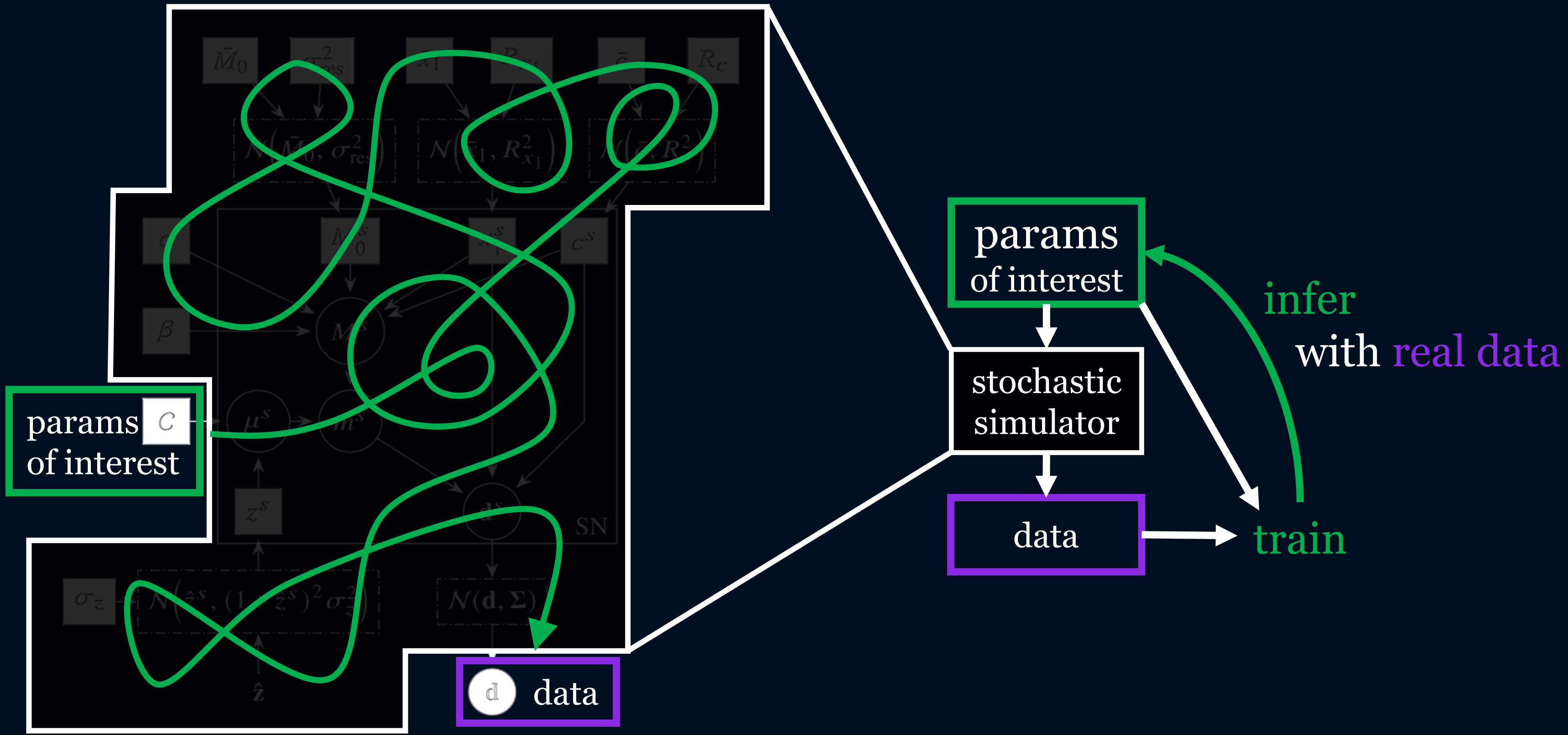
$$p(\theta|d) \propto \int p(d|\theta,\nu)\, p(\theta,\nu)\, d\nu$$

Requires

- calculating all probabilities
  - ➢ limited to simple analytic descriptions
- inferring all parameters jointly
  - ➢ unfavourable scaling, e.g. $\mathcal{O}(N^2)$

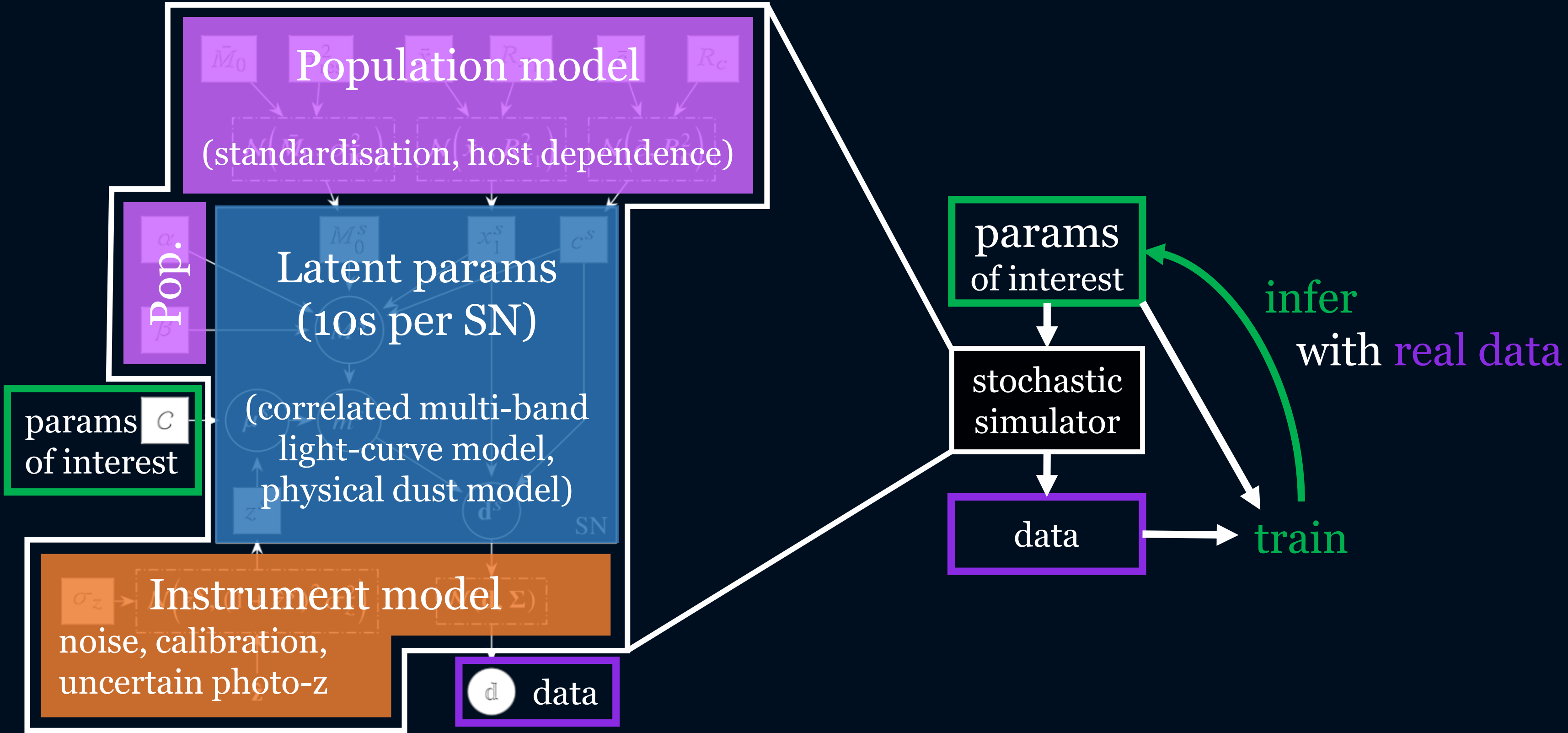
Simulation-based inference

- Model = simulator
  - can be arbitrarily complex



Simulation-based inference

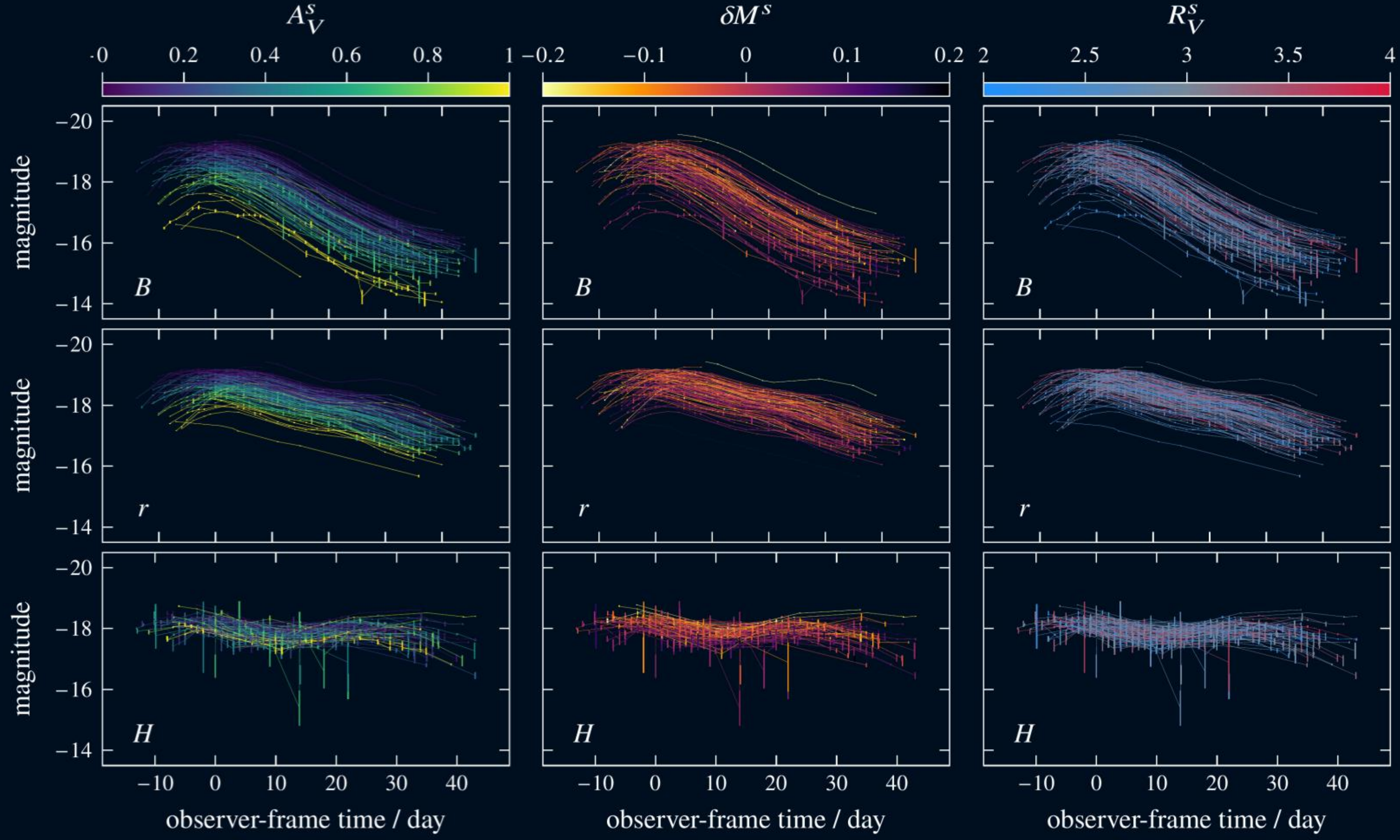
- Model = simulator
  - often more interpretable / grounded / “physical”



SICRET (2023):
cosmology with
SALT parameters.

SIDE-real (2024):
dust inference with
the BayeSN light
curve model

SLiCsim: realistic uncertain light curves



visible
(more variability)

infrared
(less variability)

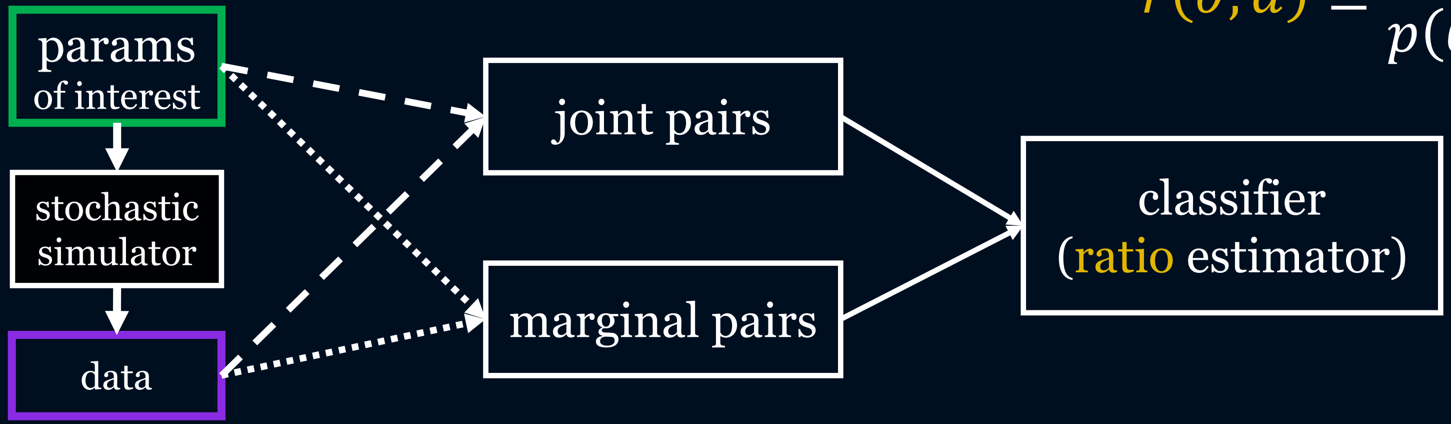
SIDE-real (2024):
dust inference with
the BayeSN light
curve model

Neural simulation-based inference

- Neural ratio estimation

$$-\text{BCELoss} = \mathbb{E}_{p(d|\theta)p(\theta)}\left[\ln \frac{\hat{r}_{\text{NN}}}{1+\hat{r}_{\text{NN}}}\right] + \mathbb{E}_{p(d)p(\theta)}\left[\ln \frac{1}{1+\hat{r}_{\text{NN}}}\right]$$

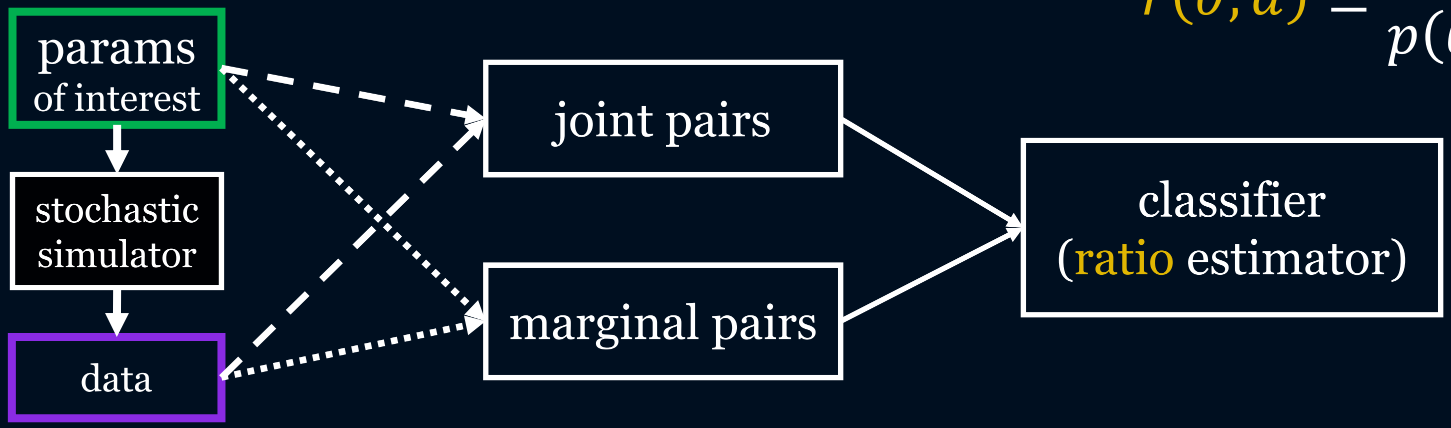
$$r(\theta, d) \equiv \frac{p(\theta, d)}{p(\theta)p(d)}$$



Neural simulation-based inference

- Neural ratio estimation
  - simplicity of NN architecture (binary classifier)
  - flexibility with prior
  - allows Bayesian *and* frequentist inference

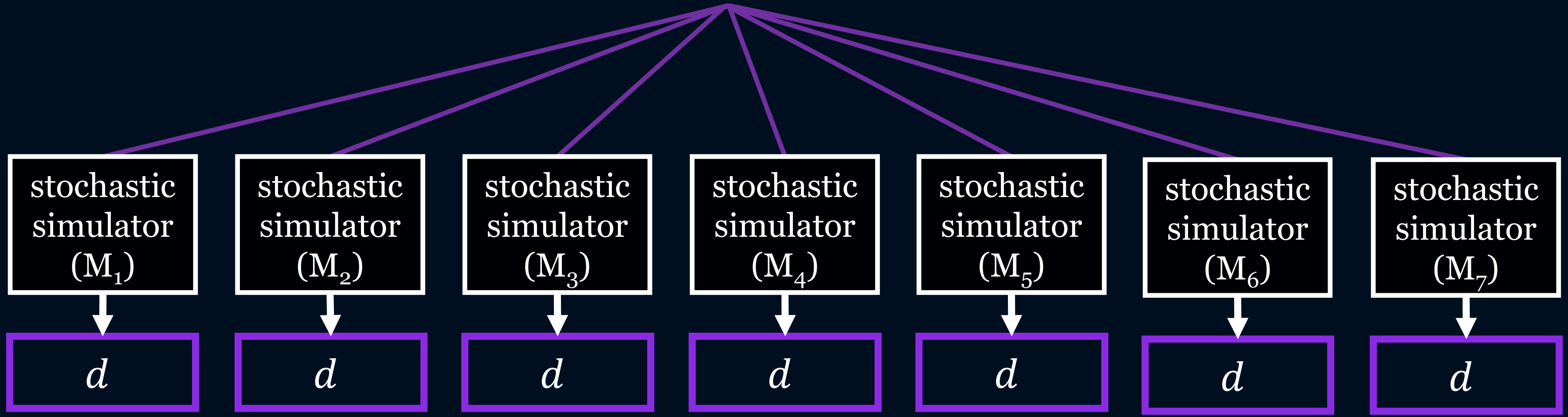
$$r(\theta, d) \equiv \frac{p(\theta, d)}{p(\theta)p(d)} = \frac{p(\theta|d)}{p(\theta)} = \frac{p(d|\theta)}{p(d)}$$



Beyond SB *inference*

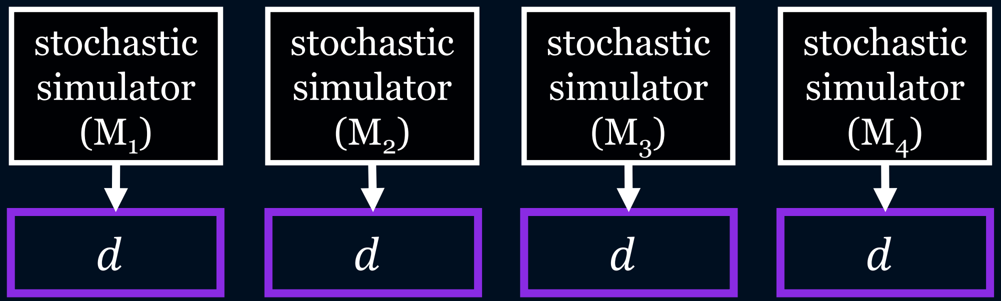
- Neural ratio estimation
- Neural model comparison

$$-\mathrm{NLLLoss} = \mathbb{E}_{\mathrm{p}(d|M)p(M)}[\ln \hat{p}_{\mathrm{NN}}(M|d)]$$



Beyond SB *inference*

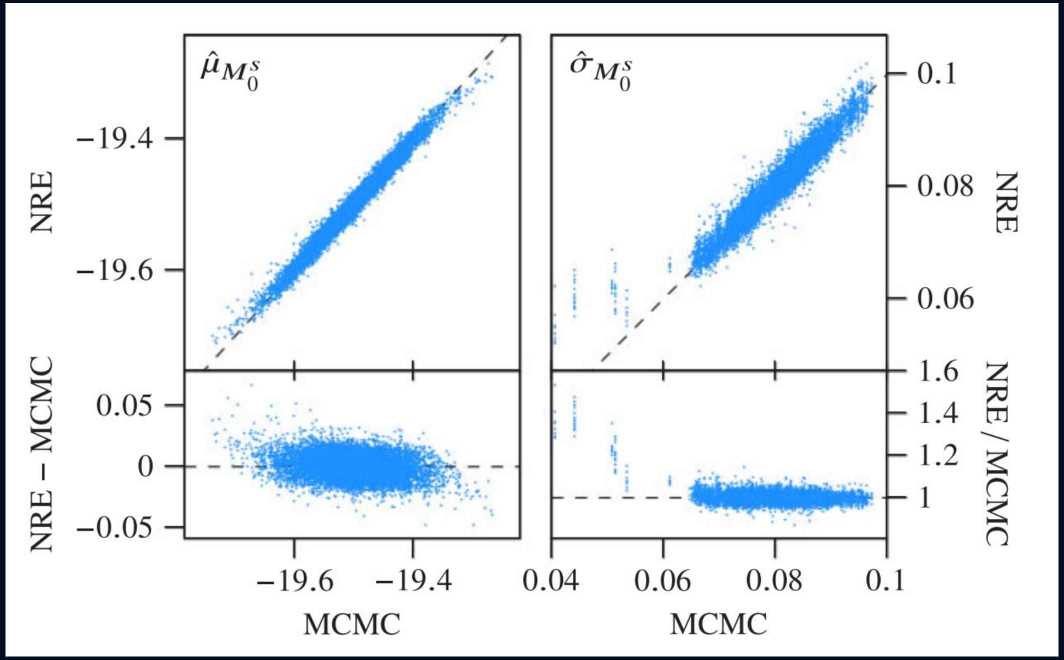
- Neural ratio estimation
- Neural model comparison
  - n-class classifier with trivial cross-entropy loss
  - gives direct access to Bayes factor / posterior over models

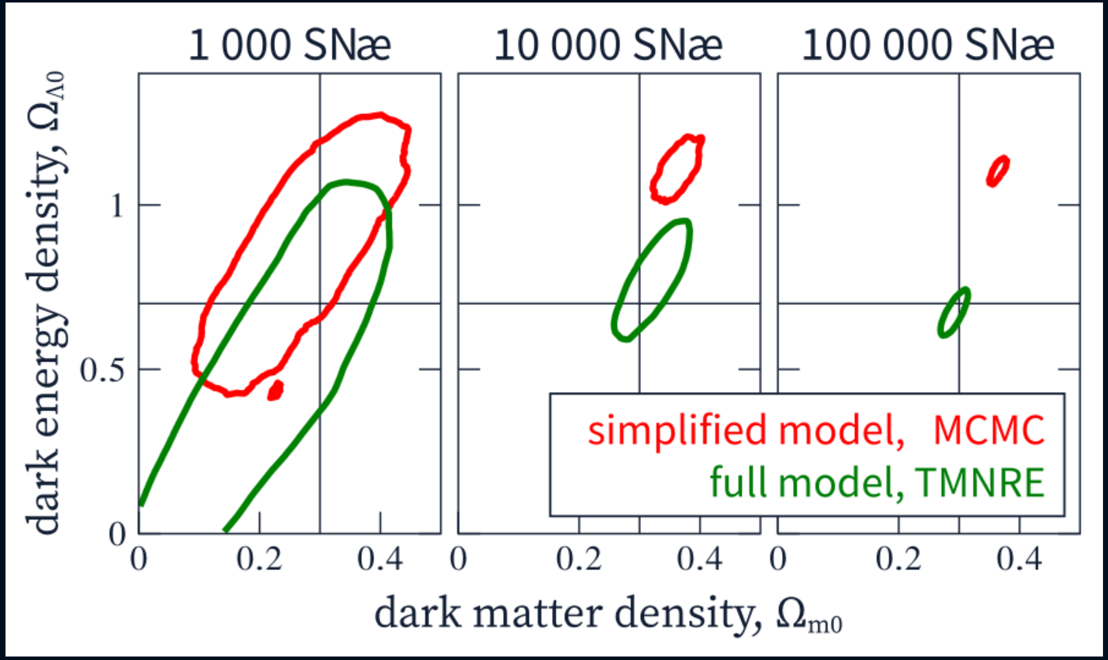


$$-\mathrm{NLLLoss} = \mathbb{E}_{\mathrm{p}(d|M)p(M)}[\ln \hat{p}_{\mathrm{NN}}(M|d)]$$

Advantages of neural SBI

- scales to future-sized data sets
- avoids biases due to simplifications

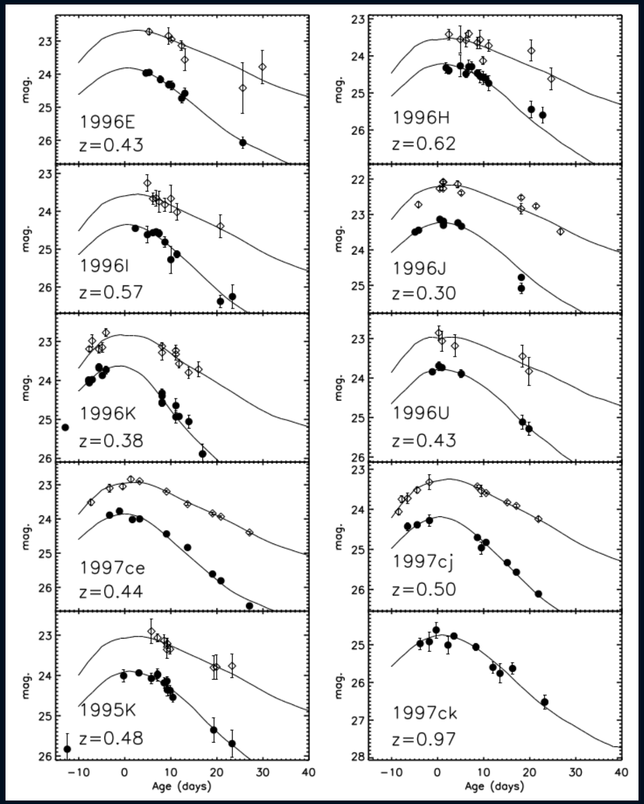




from SICRET (2023)

Advantages of neural SBI

- data can be arbitrarily* complex



$\left\{ \begin{array}{c} \boldsymbol{d}^1 \\ \boldsymbol{d}^2 \\ \ldots \end{array} \right\}$ NN optimised object 1 summary

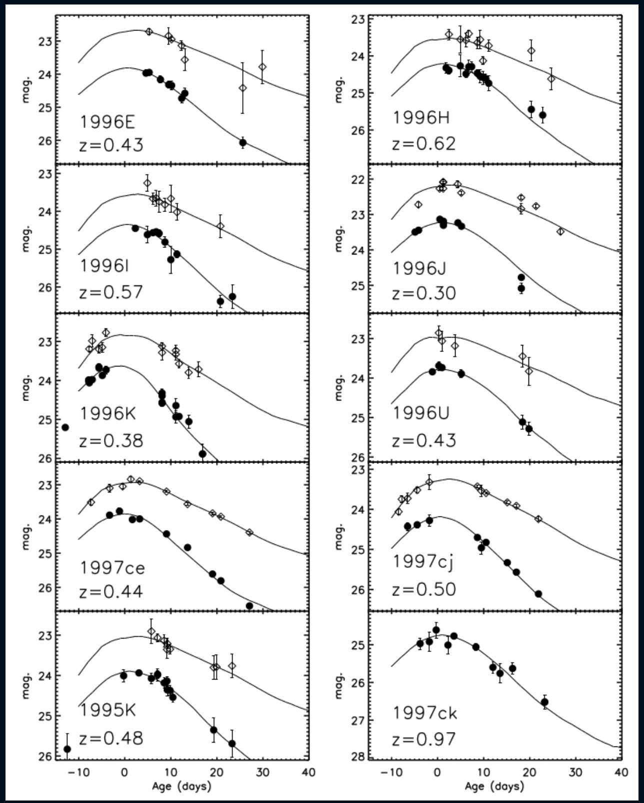
$\left\{ \begin{array}{c} \boldsymbol{d}^{25} \\ \boldsymbol{d}^{26} \\ \ldots \end{array} \right\}$ NN optimised object 2 summary

$\left\{ \begin{array}{c} \boldsymbol{d}^{42} \\ \boldsymbol{d}^{43} \\ \ldots \end{array} \right\}$ NN optimised object 3 summary

NN optimised full data set summary tailored for the inference task

Advantages of neural SBI

➢ data can be arbitrarily* complex: wealth of NN architectures



$\left\{ \begin{array}{c} \boldsymbol{d}^1 \\ \boldsymbol{d}^2 \\ \dots \end{array} \right\}$ NN optimised object 1 summary

$\left\{ \begin{array}{c} \boldsymbol{d}^{25} \\ \boldsymbol{d}^{26} \\ \dots \end{array} \right\}$ NN optimised object 2 summary

$\left\{ \begin{array}{c} \boldsymbol{d}^{42} \\ \boldsymbol{d}^{43} \\ \dots \end{array} \right\}$ NN optimised object 3 summary

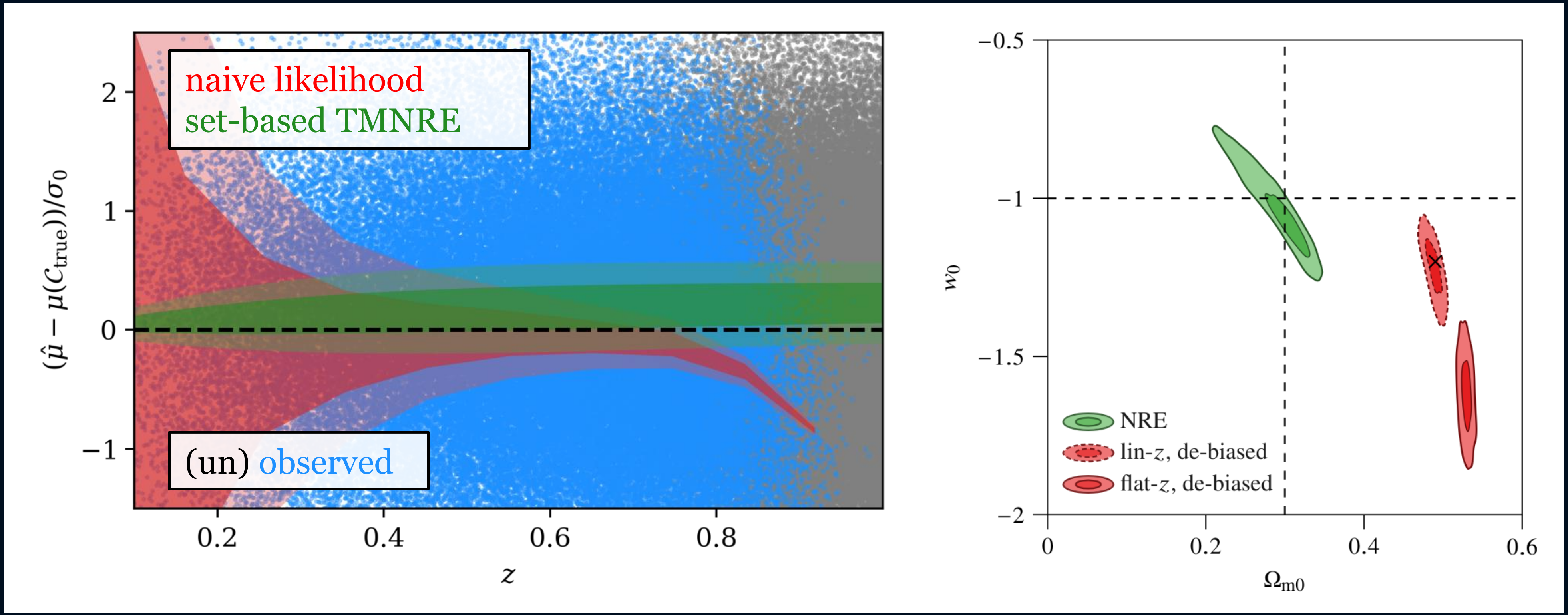
$$g_{NN}\left[\theta, \sum_i f_{\mathrm{NN}}\left(d_i, \theta\right)\right]$$

Conditioned deep set (Zaheer et al. 2017),

Transformers?

Unique advantages of neural SBI

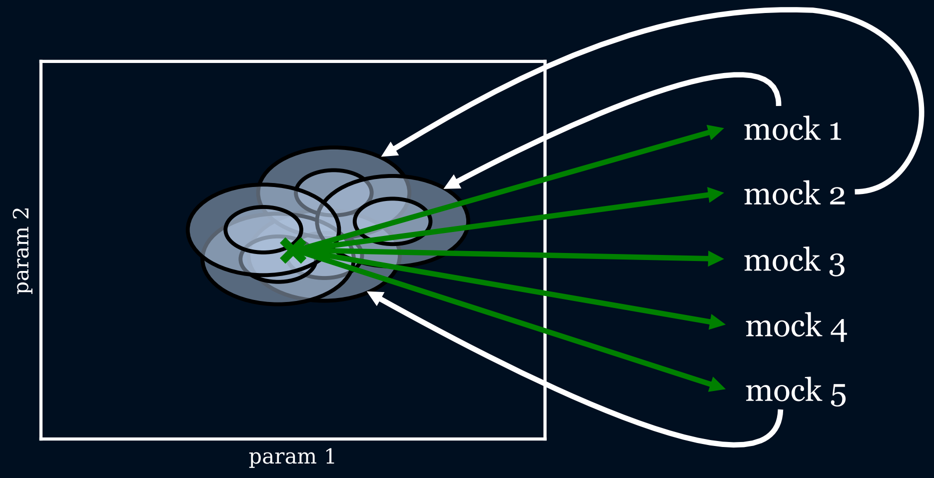
➢ intractable probabilities and varying-size data sets: 🙌 selection effects 🙌

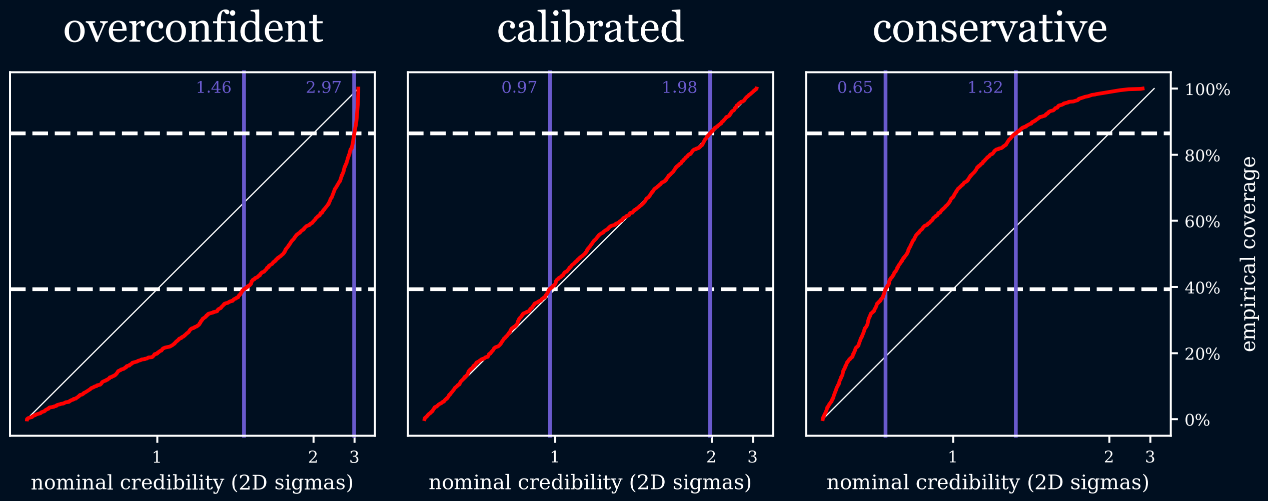


from RESSET: Ratio Estimation for Supernova Selection Effects (very soon™)

Unique advantages of neural SBI

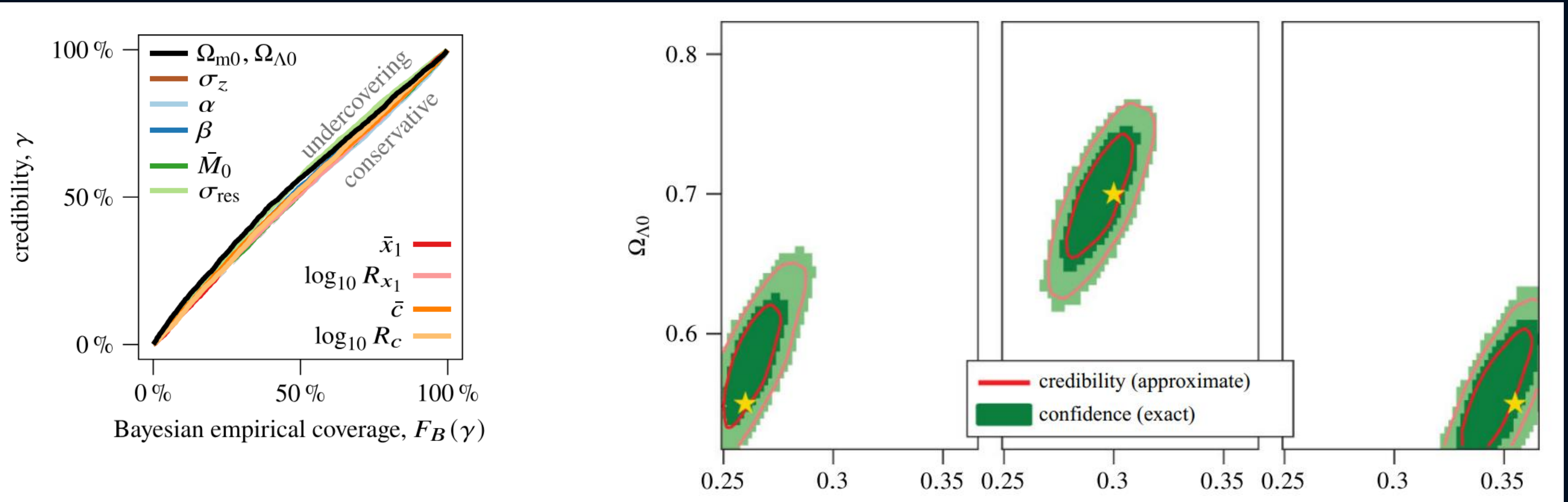
- Amortised inference
  - can be quickly validated / calibrated on simulations





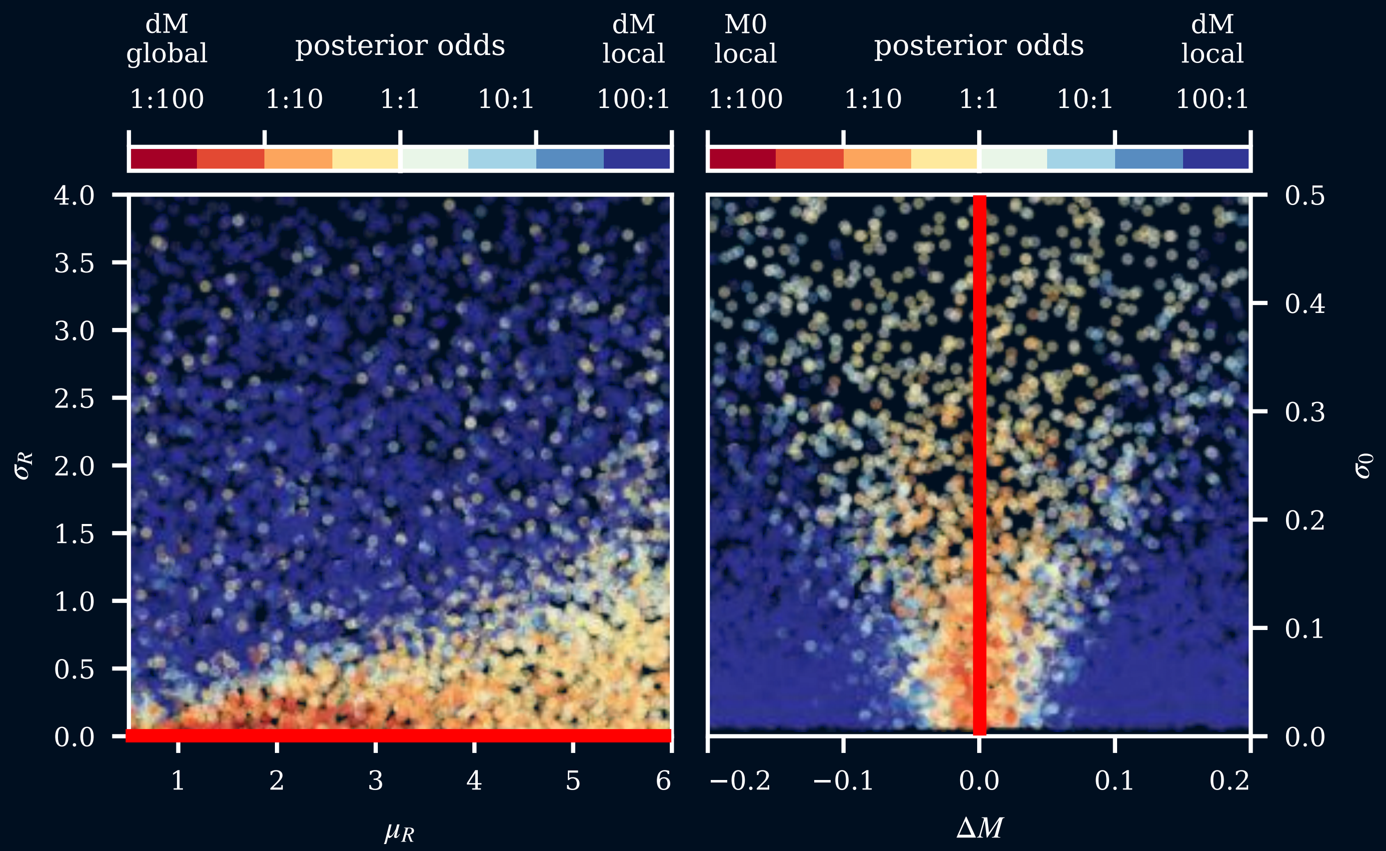
Unique advantages of neural SBI

- Amortised inference
  - can be quickly validated / calibrated on simulations
  - exact (frequentist) confidence regions



from SICRET (2023)

Amortised model comparison: Visualising Occam's razor



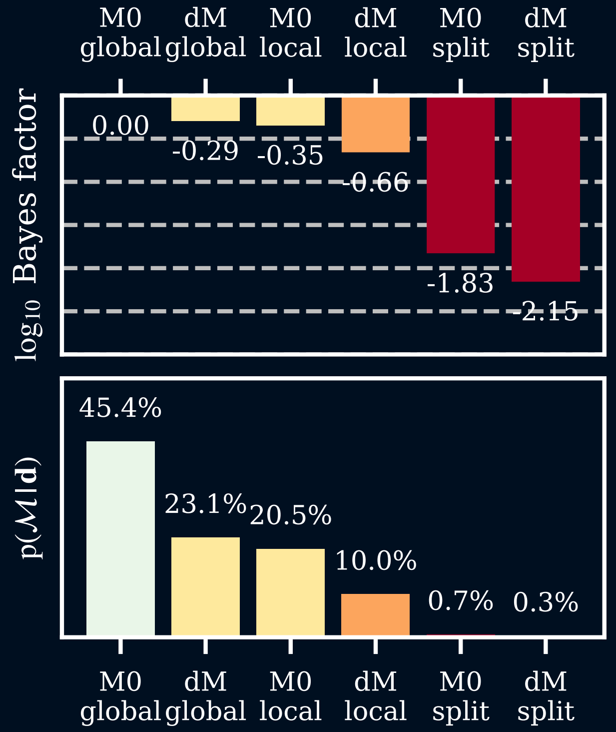
scatter of R_V^s vs. global R_V

scatter of M^s vs. magnitude step

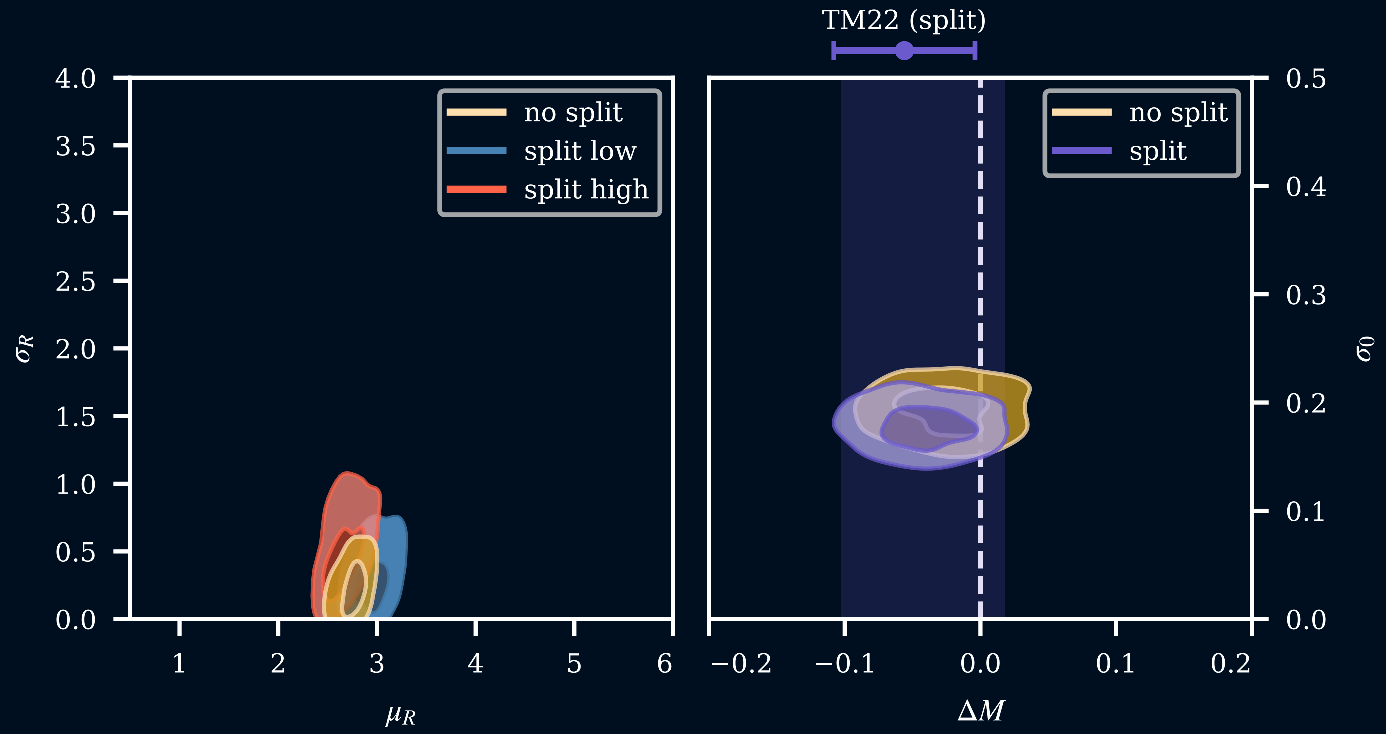
SimSIMS (2023): Model comparison for mass-magnitude step and dust laws

First results on real light curves!

neural model comparison



neural parameter inference (using NRE, reused simulations)



SNæ in the 2020s: Challenges and solutions

- Noisy light curves with irregular cadence -- deep sets
  - classic fits and GP take time… -- instant high-dim inference
- Modelling systematics -- model selection
  - effect of dust on standardisation -- flexible simulator
  - correlations with host and evolution -- stay tuned
- Photometric redshift -- stay tuned
  - complicated uncertainty, contamination -- SBI can handle it
- Selection biases -- set-based SBI (very soon)
  - current “bias correction” is ad-hoc -- fully principled SBI

Thank you for your attention!

"An illustration of a supernova explosion with swirling cosmic gases in the background, inspired by the surrealist paintings of Salvador Dali"

image by DALL·E, prompt by ChatGPT