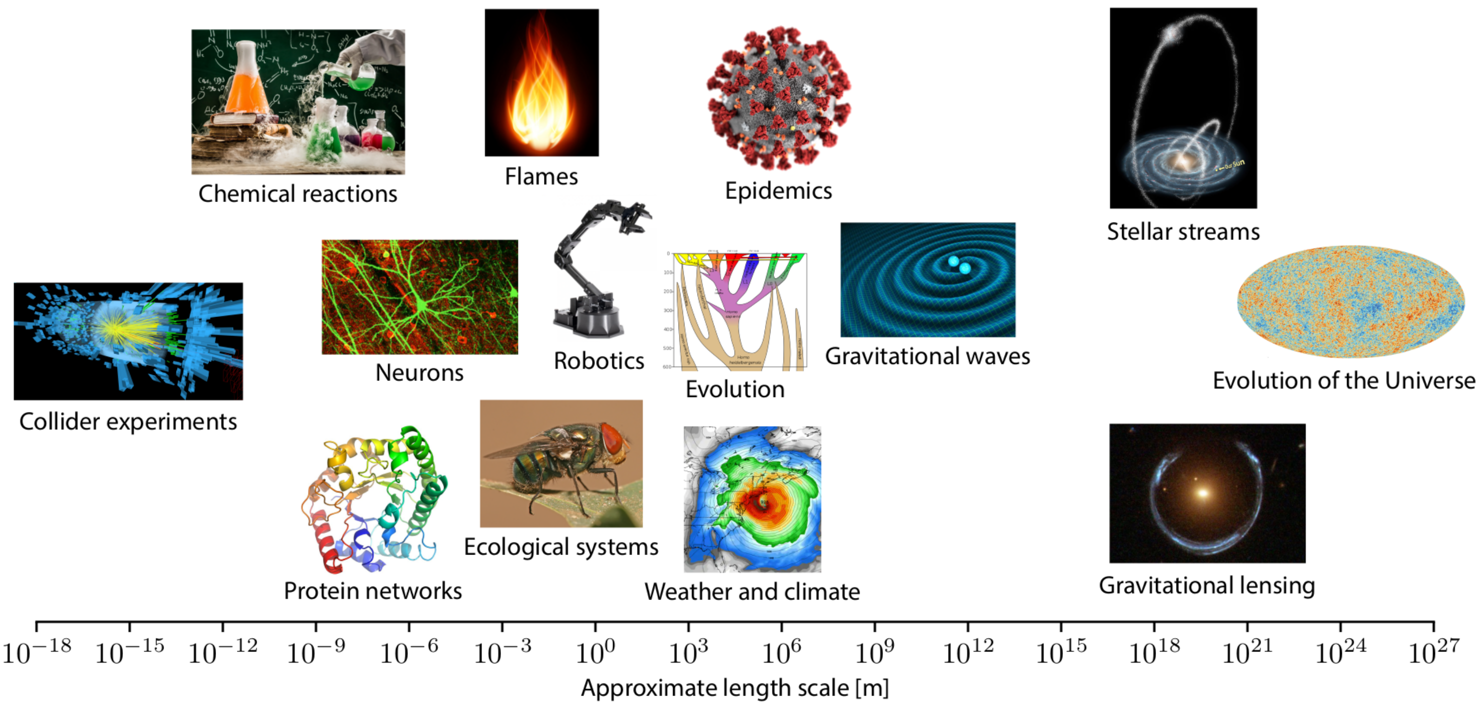
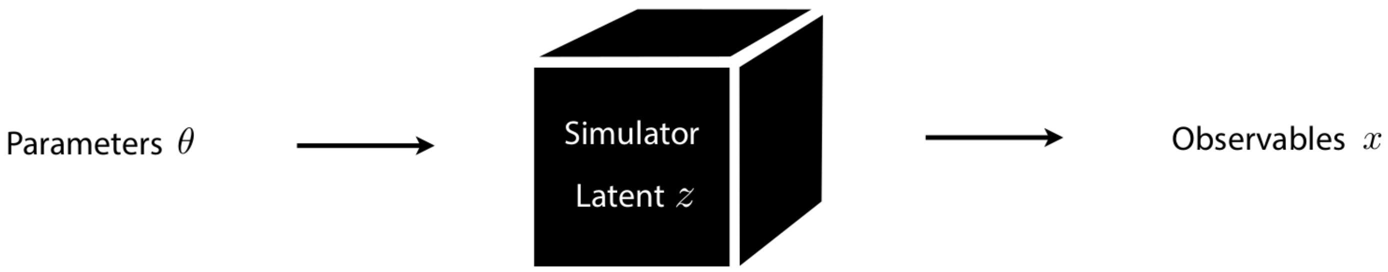
Neural Ratio Estimation For Simulation-based Inference

Nordstat 2021
June 23

Gilles Louppe
g.louppe@uliege.be

Simulation-based inference





Prediction:

- Well-motivated mechanistic, causal model
- Simulator can generate samples $x \sim p(x|\theta)$

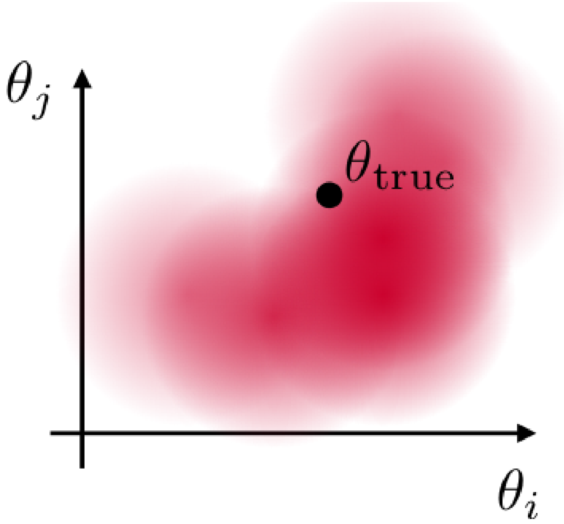
Inference:

- Interactions between low-level components lead to challenging inverse problems
- Likelihood $p(x|\theta) = \int \mathrm{d}z \; p(x, z|\theta)$ is intractable

Start with

- a simulator that lets you generate N samples $x_i \sim p(x_i|\theta_i)$ (for parameters θ_i of our choice),
- observed data $x_{\text{obs}} \sim p(x_{\text{obs}}|\theta_{\text{true}})$,
- a prior $p(\theta)$.

Then, estimate the posterior $p(\theta|x_{\text{obs}})$.



Amortizing Bayes

The Bayes rule can be rewritten as

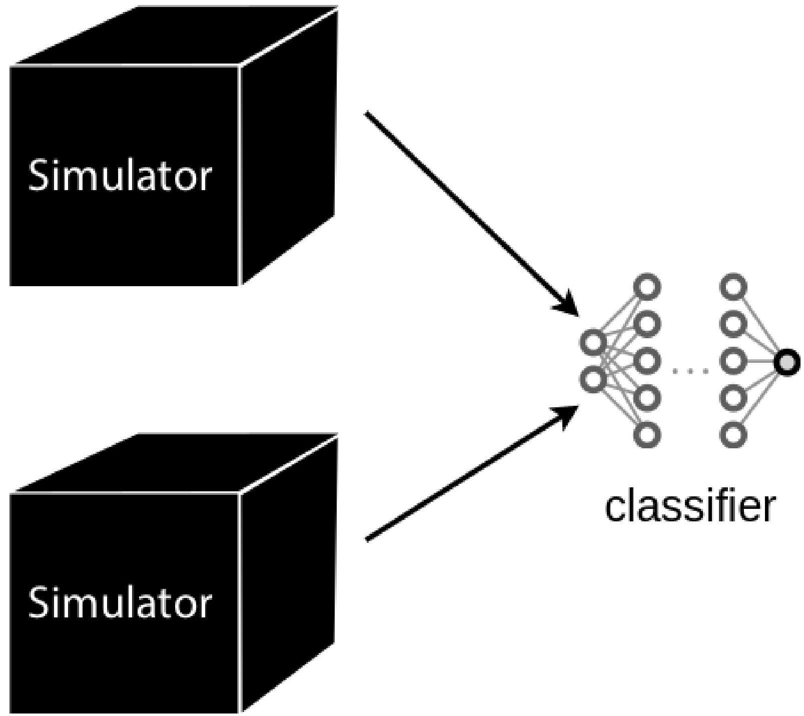
$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = r(x|\theta)p(\theta) \approx \hat{r}(x|\theta)p(\theta),$$

where $r(x|\theta) = \frac{p(x|\theta)}{p(x)}$ is the likelihood-to-evidence ratio.

The ratio can be learned with machine learning, even when neither the likelihood nor the evidence can be evaluated!

$x, \theta \sim p(x, \theta)$

$x, \theta \sim p(x)p(\theta)$



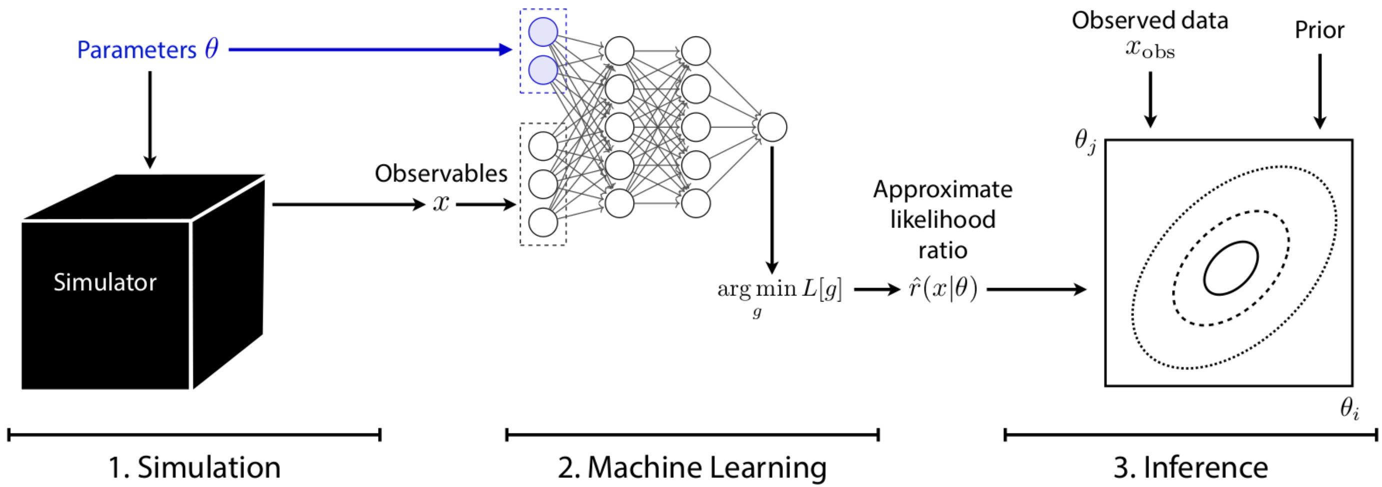
The solution d found after training approximates the optimal classifier

$$d(x,\theta) \approx d^*(x,\theta) = \frac{p(x,\theta)}{p(x,\theta) + p(x)p(\theta)}.$$

Therefore,

$$r(x|\theta) = \frac{p(x|\theta)}{p(x)} = \frac{p(x,\theta)}{p(x)p(\theta)} \approx \frac{d(x,\theta)}{1 - d(x,\theta)} = \hat{r}(x|\theta).$$

Inference

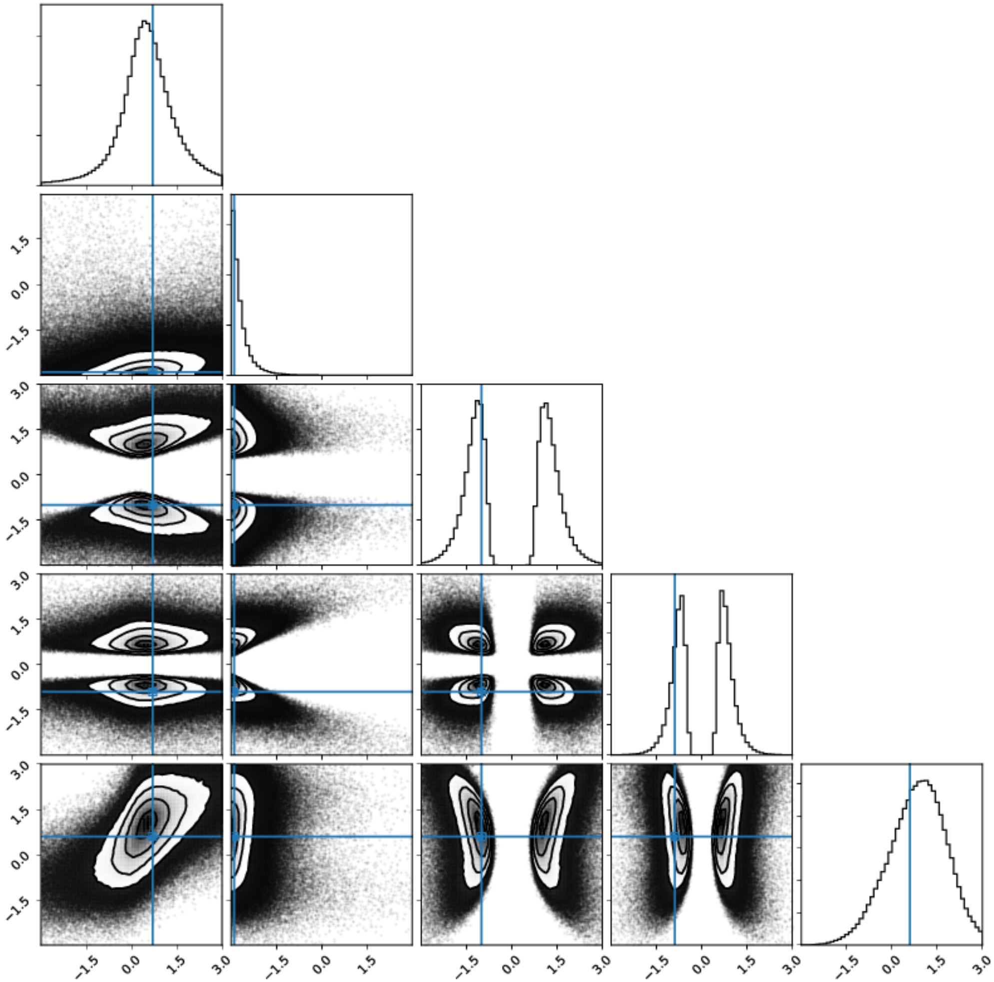


Run simulator and save data

Train NN classifier, interpret as likelihood ratio estimator

Amortized: cheap to repeat for new data

Diagnostics



How to assess that the approximate posterior is not wrong?

Coverage

- For every $x, \theta \sim p(x, \theta)$ in a validation set, compute the $1 - \alpha$ credible interval based on $\hat{p}(\theta|x) = \hat{r}(x|\theta)p(\theta)$.
- The fraction of samples for which θ is contained within the interval corresponds to the empirical coverage probability.
- If the empirical coverage is larger that the nominal coverage probability $1 - \alpha$, then the ratio estimator $\hat{r}$ passes the diagnostic.

Convergence towards the nominal value θ^*

If the approximation $\hat{r}$ is correct, then the posterior

$$\hat{p}(\theta|\mathcal{X}) = \frac{p(\theta)p(\mathcal{X}|\theta)}{p(\mathcal{X})} = p(\theta)\left[\int p(\theta')\prod_{x\in\mathcal{X}}\frac{p(x|\theta')}{p(x|\theta)}d\theta'\right]^{-1}$$
$$\approx p(\theta)\left[\int p(\theta')\prod_{x\in\mathcal{X}}\frac{\hat{r}(x|\theta')}{\hat{r}(x|\theta)}d\theta'\right]^{-1}$$

should concentrate around θ^* as the number of observations

$$\mathcal{X} = \{x_1, ..., x_n\},$$

for $x_i \sim p(x|\theta^*)$, increases.

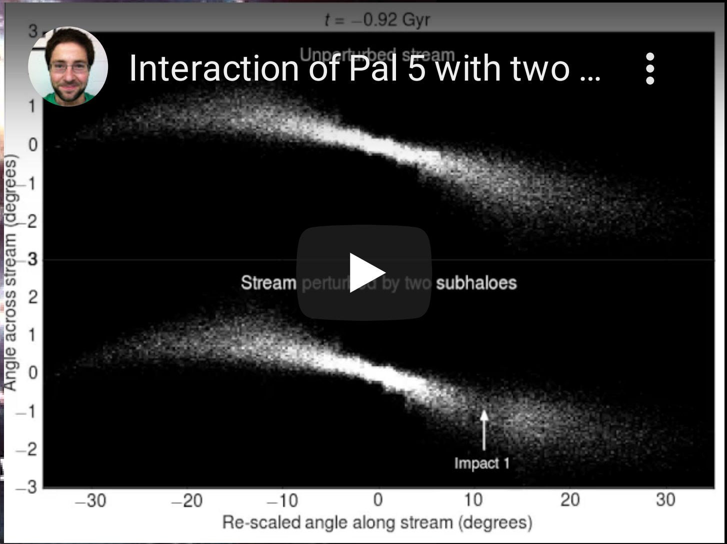
ROC AUC score

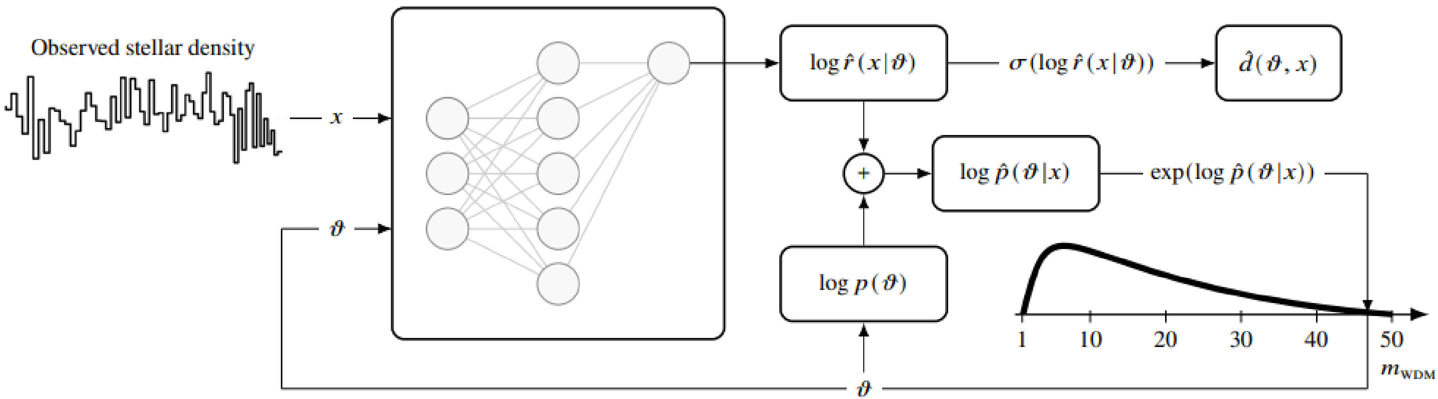
The ratio estimator $\hat{r}\,(x|\theta)$ is only exact when samples x from the reweighted marginal model $p(x)\,\hat{r}\,(x|\theta)$ cannot be distinguished from samples x from a specific likelihood $p(x|\theta)$.

Therefore, the predictive ROC AUC performance of a classifier should be close to 0.5 if the ratio is correct.

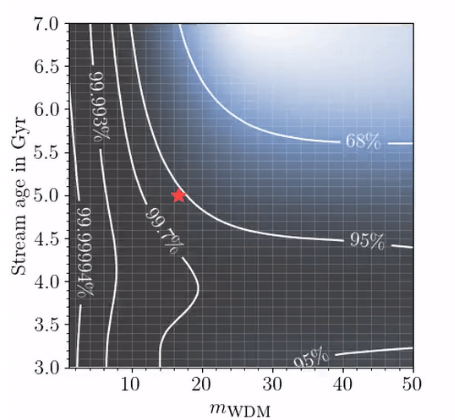
Constraining dark matter with stellar streams

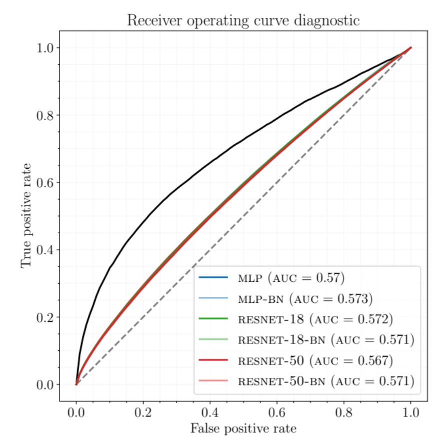
.]



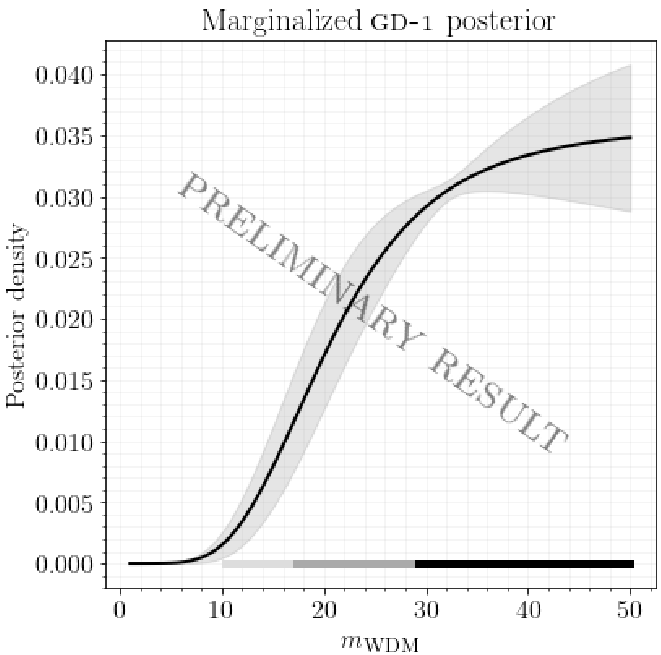


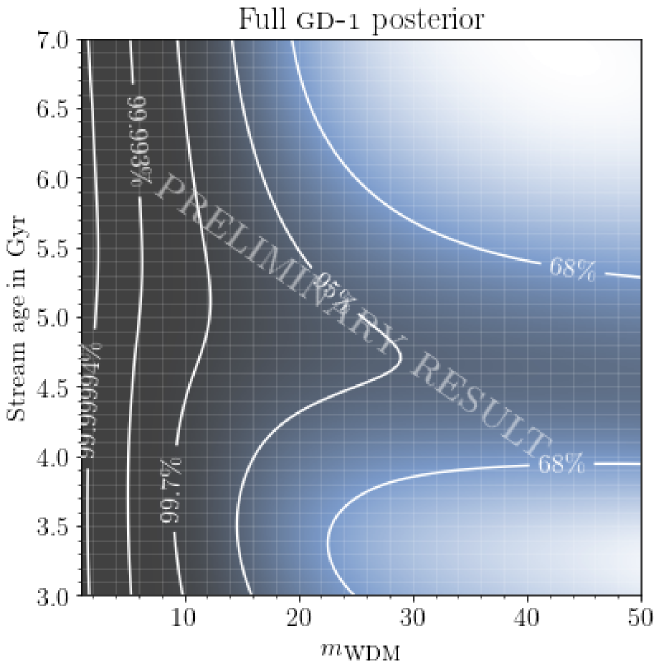
Architecture	68% CR	95% CR
$\hat{r}(x\|\vartheta)$ with $\vartheta \triangleq (m_{\rm WDM})$		
MLP	0.685 $_{\pm 0.004}$	0.954 $_{\pm 0.002}$
MLP-BN	0.687 $_{\pm 0.006}$	0.951 $_{\pm 0.002}$
RESNET-18	0.667 $_{\pm 0.004}$	0.943 $_{\pm 0.002}$
RESNET-18-BN	0.672 $_{\pm 0.004}$	0.945 $_{\pm 0.001}$
RESNET-50	0.671 $_{\pm 0.005}$	0.947 $_{\pm 0.003}$
RESNET-50-BN	0.678 $_{\pm 0.004}$	0.949 $_{\pm 0.004}$
$\hat{r}(x\|\vartheta)$ with $\vartheta \triangleq (m_{\rm WDM}, t_{\rm age})$		
MLP	0.685 $_{\pm 0.005}$	0.953 $_{\pm 0.002}$
MLP-BN	0.685 $_{\pm 0.004}$	0.952 $_{\pm 0.003}$
RESNET-18	0.666 $_{\pm 0.005}$	0.945 $_{\pm 0.002}$
RESNET-18-BN	0.671 $_{\pm 0.003}$	0.945 $_{\pm 0.003}$
RESNET-50	0.674 $_{\pm 0.006}$	0.944 $_{\pm 0.002}$
RESNET-50-BN	0.677 $_{\pm 0.004}$	0.947 $_{\pm 0.003}$





Coverage **Convergence to θ^*** **ROC AUC score**





Preliminary results for GD-1 suggest a preference for CDM over WDM.

In summary

- Much of modern science is based on simulators making precise predictions, but in which inference is challenging.
- Machine learning enables powerful inference methods.
- They work in problems from the smallest to the largest scales.
- Further advances in machine learning will translate into scientific progress.

Thanks!








References

- Hermans, J., Banik, N., Weniger, C., Bertone, G., & Louppe, G. (2020). Towards constraining warm dark matter with stellar streams through neural simulation-based inference. arXiv preprint arXiv:2011.14923.
- Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference. Proceedings of the National Academy of Sciences, 117(48), 30055-30062.
- Hermans, J., Begy, V., & Louppe, G. (2019). Likelihood-free MCMC with Approximate Likelihood Ratios. arXiv preprint arXiv:1903.04057.
- Cranmer, K., Pavez, J., & Louppe, G. (2015). Approximating likelihood ratios with calibrated discriminative classifiers. arXiv preprint arXiv:1506.02169.

The end.