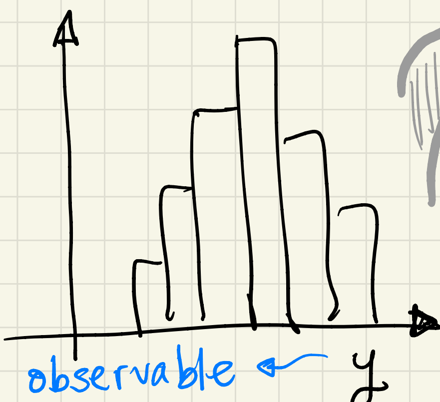
MODERN APPROACHES & TOOLS FOR STAT. INFERENCE

Bryan Zaldivar
(IFIC, Valencia)

MOTIVATION

- Increasing complexity of physics experiments

Simple experiment



Bin j : $n_j^{obs} \sim \text{Poisson}(\lambda_j)$

Maximum Likelihood Estimator

$$\lambda_j = \lambda_j^{bckg} + \lambda_j^{signal}(\theta)$$

params. of interest

- Good enough if
  - Dataset is sufficiently large
  - Binning is " fine

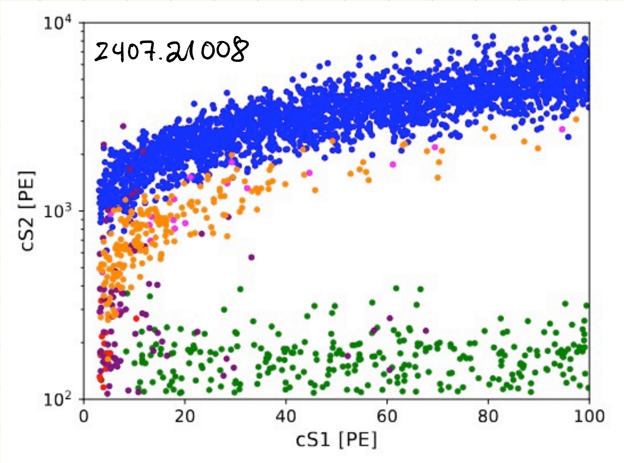
MOTIVATION

Complex experiments

MOTIVATION

Complex experiments

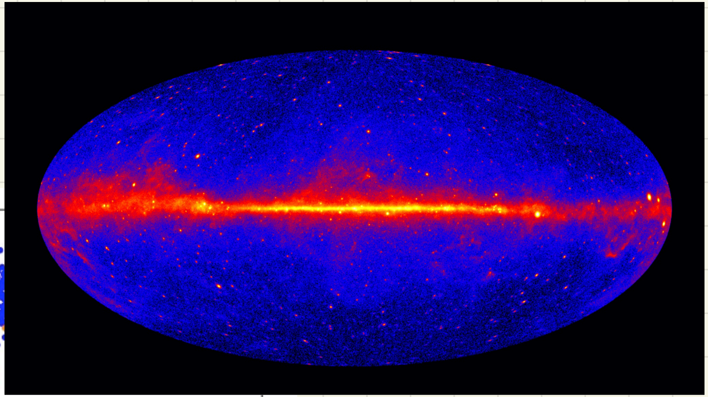
Simulation of Xenon-nT

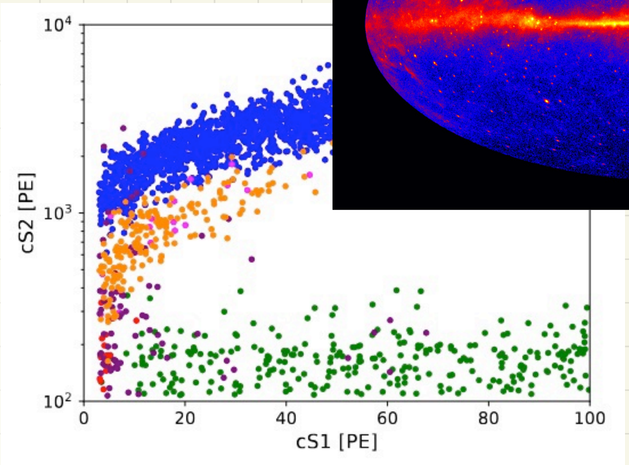


MOTIVATION

Complex experiments

Fermi-LAT integrated γ-ray sky

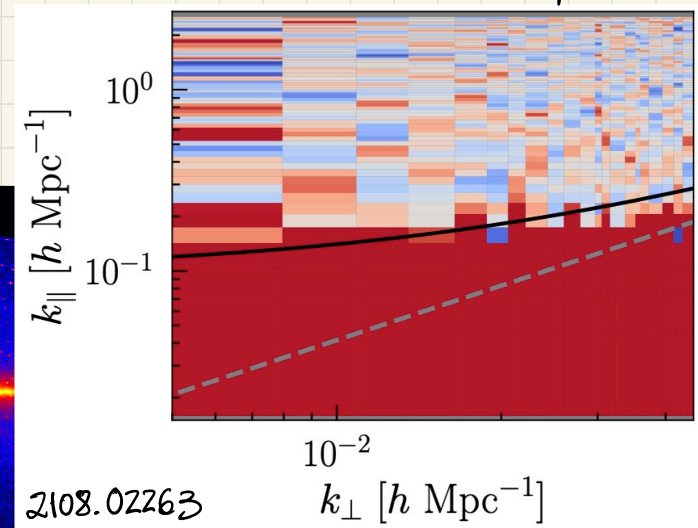




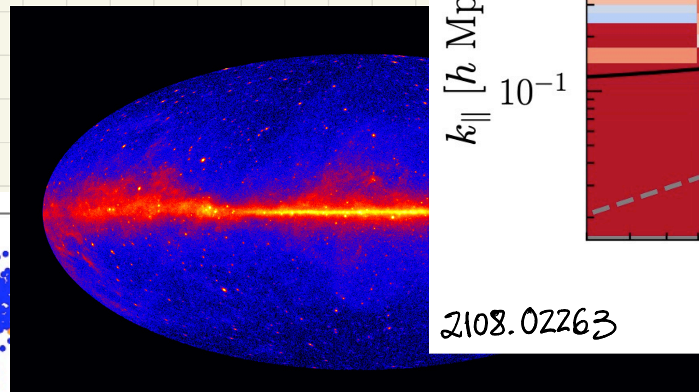
MOTIVATION

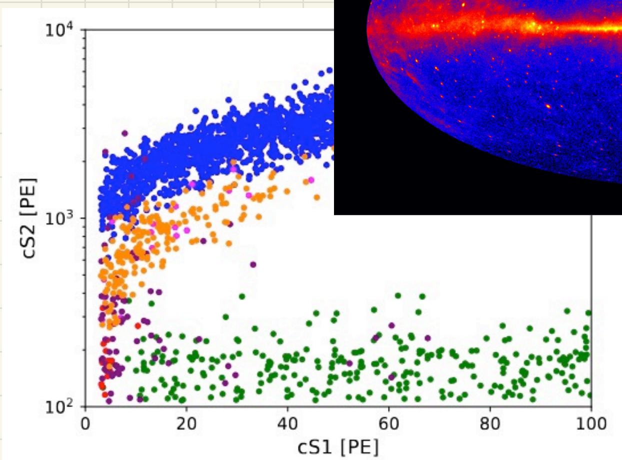
Complex experiments

HERA measured 21cm Power Spectrum



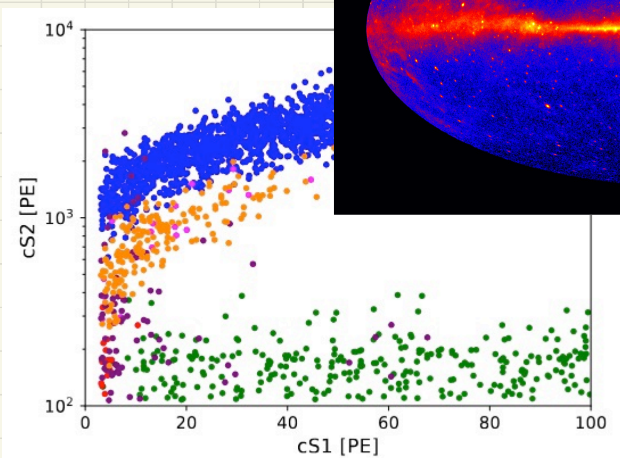
2108.02263

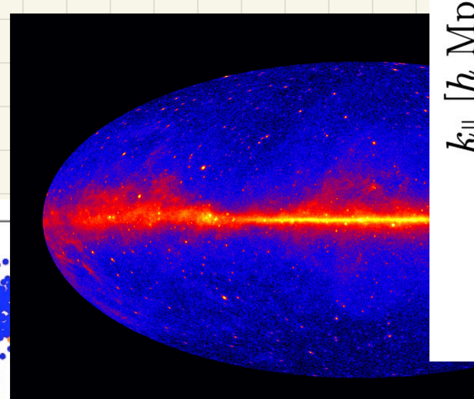


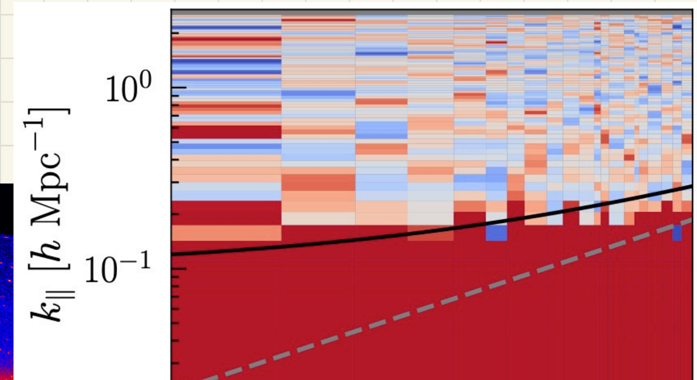


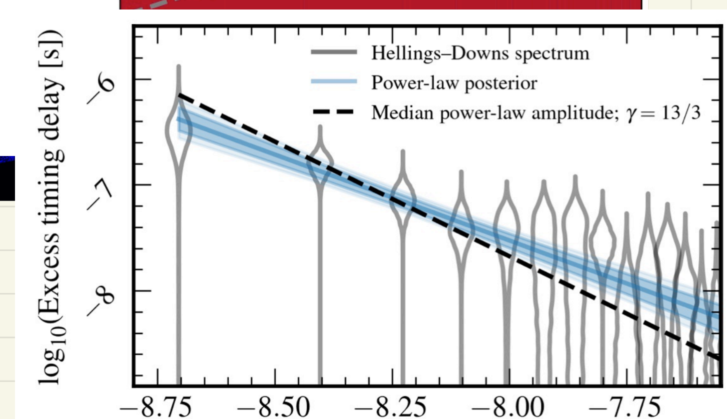
MOTIVATION

$\boxed{\text{Complex experiments}}$









2306.16213

NANOGrav PTA timing delay distrib.

- Histogram analysis becomes largely insufficient

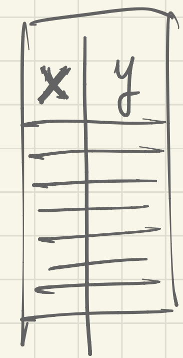
- Histogram analysis becomes largely insufficient
- Need to consider instead the distribution of the observables themselves

$$\boxed{p(y|x;\theta)}$$

x	y

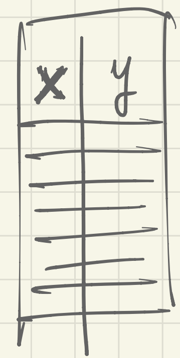
- Histogram analysis becomes largely insufficient
- Need to consider instead the distribution of the observables themselves

$$\boxed{p(y|x;\theta)}$$



How to make inference on θ?

- Histogram analysis becomes largely insufficient
- Need to consider instead the distribution of the observables themselves $\boxed{p(y|x;\theta)}$



How to make inference on θ ?

usual practice

Assume a parametric shape

e.g. $\mathcal{N}(y \mid \mu(x;\theta), \sigma)$

MLE

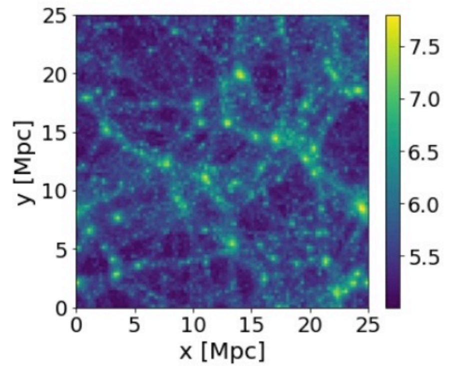
Bayesian, $p(\theta|y)$

* MCMC
* Variational Inference

However the true likelihood is intractable in typical simulations nowadays:

However the true likelihood is intractable in typical simulations nowadays:

[2206.11312]



(a)

Figure 1. Logarithmic surface density of dark matter

Latent variables: e.g. positions of the DM halos in N-body sims.

However the true likelihood is intractable in typical simulations nowadays:

[2206.11312]

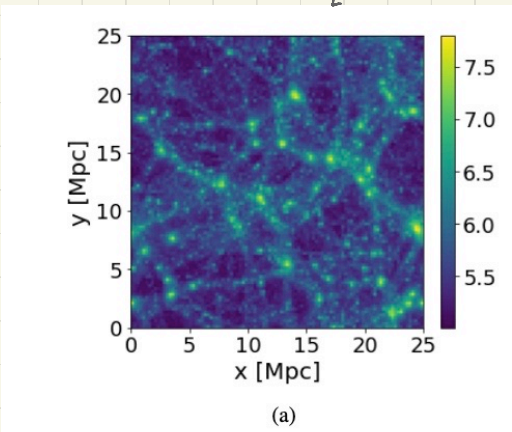
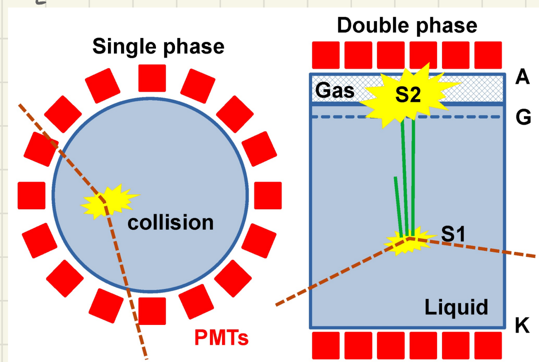


Figure 1. Logarithmic surface density of dark matter

[2310.20472]



Latent variables

e.g. positions of the DM halos in N-body sims.

Positions of the collision points in Direct Detection sims.

However the true likelihood is intractable in typical simulations nowadays:

[2206.11312]

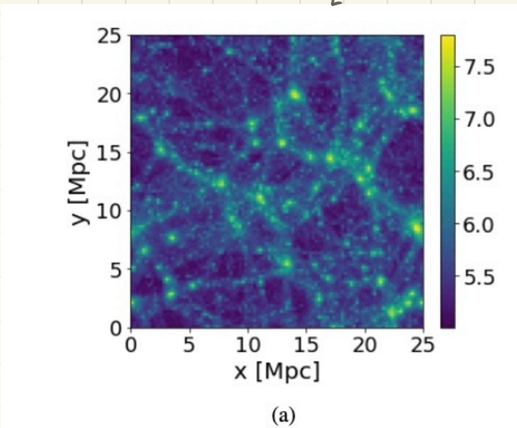
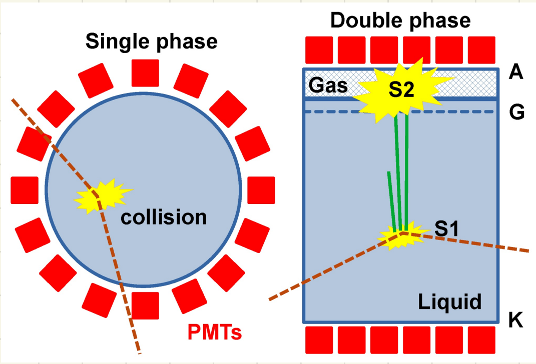
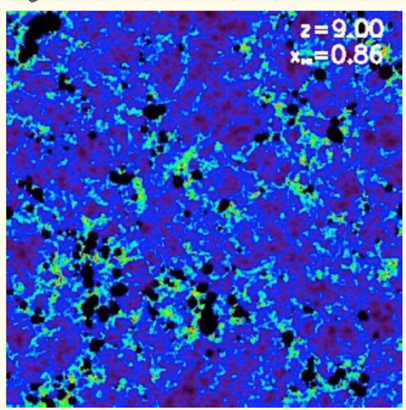


Figure 1. Logarithmic surface density of dark matter

[2310.20472]



[SKA 1210.0197]



Latent variables { e.g. positions of the DM halos in N-body sims.

Positions of the collision points in Direct Detection sims.

Initial conditions of density and velocity fields in 21cm brightness temp. sims.

However the true likelihood is intractable in typical simulations nowadays:

However the true likelihood is intractable in typical simulations nowadays:

Latent variables η $\rightsquigarrow$ As many as $\mathcal{O}(10^5)$ commonly

Likelihood $\quad p(y|x;\theta) = \int d\eta \; p(y|x;\eta,\theta)\, p(\eta|\theta)$

Intractable integral in general

However the true likelihood is intractable in typical simulations nowadays:

Latent variables η $\rightsquigarrow$ As many as $\mathcal{O}(10^5)$ commonly

Likelihood $p(y|x;\theta) = \int d\eta \; p(y|x;\eta,\theta)\, p(\eta|\theta)$

Intractable integral in general

* Likelihood-based approaches condemned to fail unless Central Limit Th. or other hint applies

However the true likelihood is intractable in typical simulations nowadays:

Latent variables η $\Rightarrow$ As many as $\mathcal{O}(10^5)$ commonly

Likelihood $\quad p(y|x;\theta) = \int d\eta \; p(y|x;\eta,\theta)\, p(\eta|\theta)$

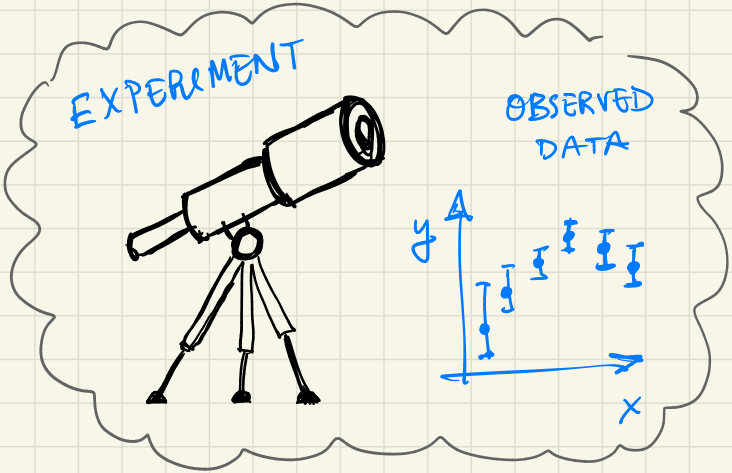
Intractable integral in general

* Likelihood-based approaches condemned to fail unless Central Limit Th. or other hint applies

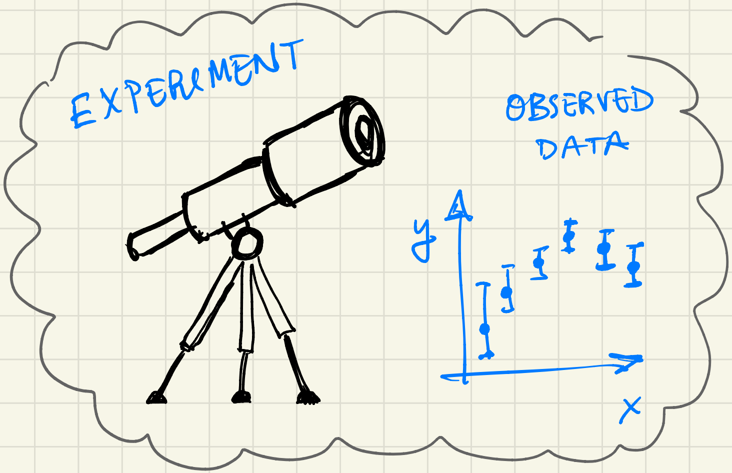
$\Rightarrow$ Embrace "Likelihood-free" approaches

a.k.a. Simulation-based Inference (SBI)

SBI in two slides

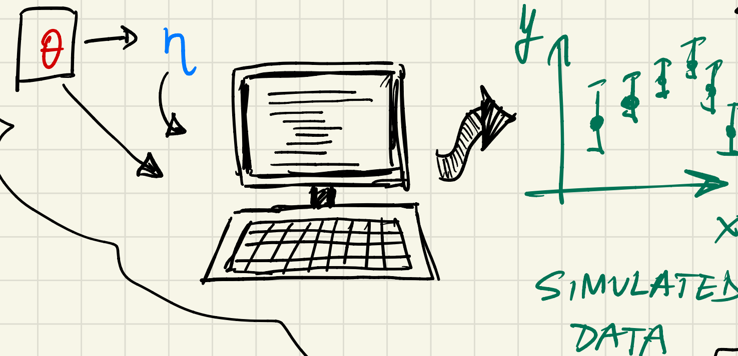


SBI in two slides

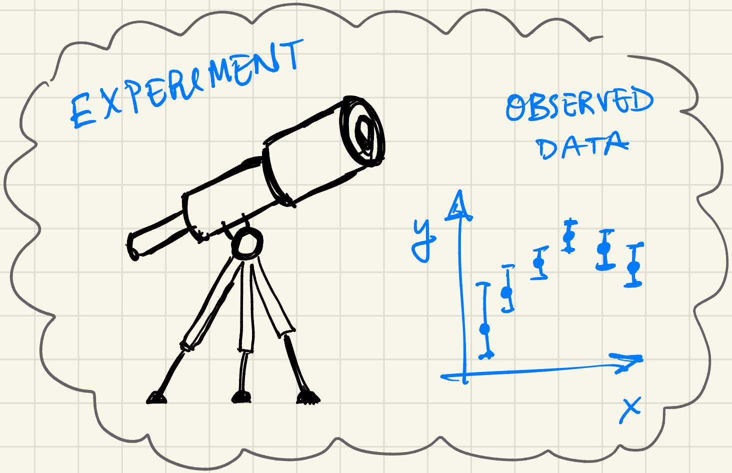


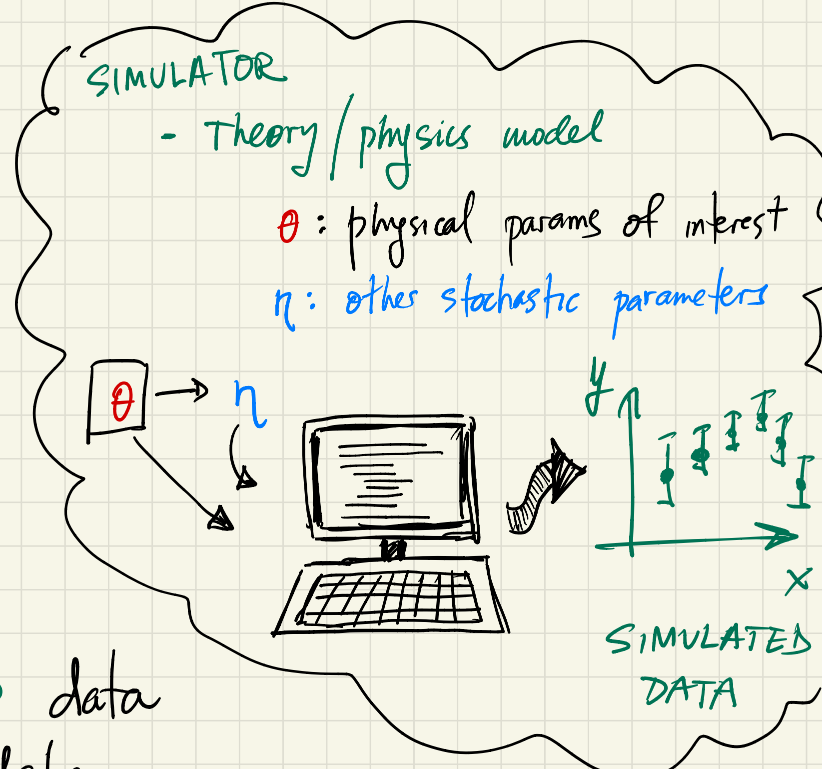
SIMULATOR

- Theory / physics model
  - θ : physical params of interest
  - η : other stochastic parameters



SBI in two slides





Aim: Solving the inverse problem

How? Selecting sets of SIMULATED data compatible with OBSERVED data

* Decades old idea (see "Approximate Bayesian Computation")

ABC ~ 80's

with computational disadvantages

* Decades old idea (see "Approximate Bayesian Computation")

ABC ~ 80's

with computational disadvantages

* Meteoric revival in the last few years thanks to
  - Old good statistical theorems
  - Nowadays computational power and modern optimisation & **ML** algorithms

* Decades old idea (see "Approximate Bayesian Computation")

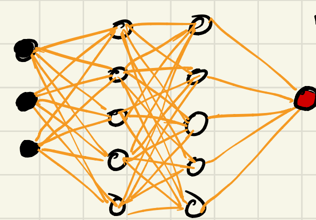
ABC ~ 80's

with computational disadvantages

* Meteoric revival in the last few years thanks to
  - Old good statistical theorems
  - Nowadays computational power and modern optimisation & **ML** algorithms

$p(\theta \,|\, \text{OBS. DATA})$

or

$p(\text{OBS DATA} \,|\, \theta)$ $\approx f\Big($  $\Big)$

Inputs ●

Output ●

} depend on the SBI flavor

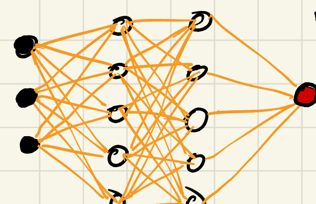
* Decades old idea (see "Approximate Bayesian Computation")

ABC ~ 80's

with computational disadvantages

* Meteoric revival in the last few years thanks to

- Old good statistical theorems
- Nowadays computational power and modern optimisation & **ML** algorithms

$p(\theta \mid \text{OBS. DATA})$ or $p(\text{OBS DATA} \mid \theta)$ $\approx f\Big($  $\Big)$

$f: \dfrac{p(\text{DATA} \mid \theta)}{p(\text{DATA})} \Rightarrow$ Neural Ratio Estimation

Inputs ●
Output ●

} depend on the SBI flavor

* Decades old idea (see "Approximate Bayesian Computation")

ABC ~ 80's

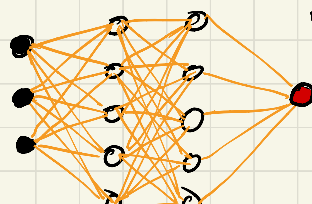
with computational disadvantages

* Meteoric revival in the last few years thanks to
  - Old good statistical theorems
  - Nowadays computational power and modern optimisation & ML algorithms

$p(\theta \mid \text{OBS. DATA})$

or

$p(\text{OBS DATA} \mid \theta)$

$\approx f\Big($  $\Big)$

$f : \dfrac{p(\text{DATA} \mid \theta)}{p(\text{DATA})}$ ⟹ Neural Ratio Estimation

$f : q_w(\theta)$ ⟹ Neural Posterior Estimation

variational distrib.

Inputs ●
Output ●

depend on the SBI flavor

* Input: Parameters $\vec{\theta}$, priors $p(\vec{\theta})$,
a simulator $g: \vec{\theta} \longmapsto \{y(\vec{x})\}$

$\vec{x}$	y

* **Input:** Parameters $\vec{\theta}$, priors $p(\vec{\theta})$,

  a simulator $g: \vec{\theta} \longmapsto \{y(\vec{x})\}$

$\vec{x}$	y

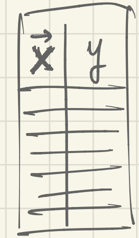
* **Algorithm:**

  ① Sample the prior N times

  $$\{\vec{\theta}_i\}, \quad i=1,..,N$$

* **Input:** Parameters $\vec{\theta}$, priors $p(\vec{\theta})$,

a simulator $g: \vec{\theta} \longmapsto \{y(\vec{x})\}$



* **Algorithm:**

1. Sample the prior N times

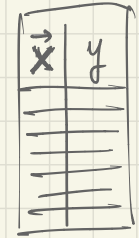
$$\{\vec{\theta}_i\}, \quad i=1,..,N$$

2. Run the simulator for each $\vec{\theta}_i$

$\vec{\theta}_i \to$  $\to \{y_i(\vec{x})\}$

$$D = \{\vec{\theta}_i \, ; \, y_i(x)\}$$

* Input: Parameters $\vec{\theta}$, priors $p(\vec{\theta})$,
  a simulator $g: \vec{\theta} \longmapsto \{y(\vec{x})\}$



* Algorithm:

① Sample the prior N times

$$\{\vec{\theta}_i\}, \quad i=1,..,N$$

② Run the simulator for each $\vec{\theta}_i$

$\vec{\theta}_i \to$  $\to \{y_i(\vec{x})\}$

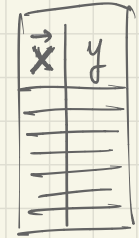
$$D = \{\vec{\theta}_i \,;\, y_i(x)\}$$

③* Train/fit our posterior approx.

$$f(D)$$

$\Downarrow$ Note:
No explicit dependence on latent variables

* Input: Parameters $\vec{\theta}$, priors $p(\vec{\theta})$,
a simulator $g: \vec{\theta} \longmapsto \{y(\vec{x})\}$



* Algorithm:

① Sample the prior N times

$\{\vec{\theta}_i\}$, $i = 1, .., N$

② Run the simulator for each $\vec{\theta}_i$

$\vec{\theta}_i \to$  $\to \{y_i(\vec{x})\}$

$D = \{\vec{\theta}_i \, ; \, y_i(x)\}$

③ * Train/fit our posterior approx.

$f(D)$

$\Downarrow$ Note:

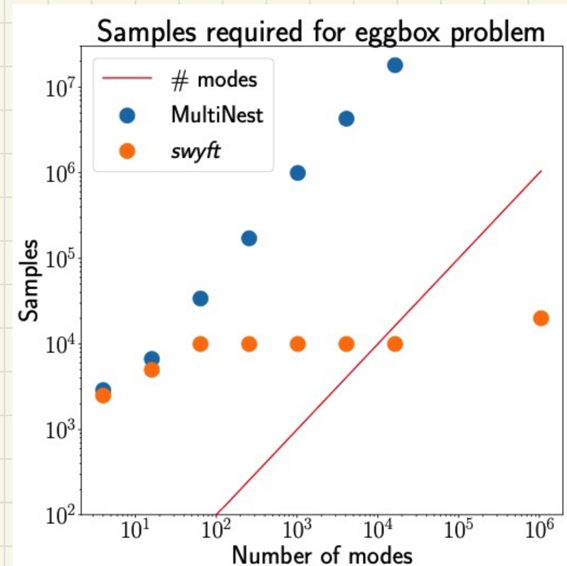
No explicit dependence on latent variables

* Pre-trained (a.k.a. "Amortised") strategy $\Rightarrow$ Inference on a new (obs) dataset takes $\approx$ no time.

SBI vs. MCMC

	SBI	MCMC
Likelihood Free	✓	✗
Exact Inference ⚠	✗	✓
Amortization	✓	✗
Latent-Variable untracking	✓	✗
Scaling with dimensionality	✓	✗

Miller et al, 2011.13951



Shallow Look at Bibliography

Approximating Likelihood Ratios with Calibrated Discriminative Classifiers #7
Kyle Cranmer (New York U.), Juan Pavez (Santa Maria U., Valparaiso), Gilles Louppe (New York U.) (Jun 6, 2015)
e-Print: 1506.02169 [stat.AP]

Mining gold from implicit models to improve likelihood-free inference #3
Johann Brehmer (New York U.), Gilles Louppe (Liege U.), Juan Pavez (Santa Maria U., Valparaiso), Kyle Cranmer (New York U.) (May 30, 2018)
Published in: *Proc.Nat.Acad.Sci.* 117 (2020) 10, 5242-5249 • e-Print: 1805.12244 [stat.ML]

The frontier of simulation-based inference #5
Kyle Cranmer (New York U., CCPP), Johann Brehmer (New York U., CCPP), Gilles Louppe (Liege U.) (Nov 4, 2019)
Published in: *Proc.Nat.Acad.Sci.* 117 (2020) 48, 30055-30062 • e-Print: 1911.01429 [stat.ML]

Towards Reliable Simulation-Based Inference with Balanced Neural Ratio Estimation
Arnaud Delaunoy, Joeri Hermans, François Rozet, Antoine Wehenkel, Gilles Louppe.
NeurIPS 2022. arXiv:2208.13624 [PDF]

A Trust Crisis In Simulation-Based Inference? Your Posterior Approximations Can Be Unfaithful
Joeri Hermans, Arnaud Delaunoy, François Rozet, Antoine Wehenkel, and Gilles Louppe.
TMLR. arXiv:2110.06581 [PDF]

Low-Budget Simulation-Based Inference with Bayesian Neural Networks
Arnaud Delaunoy, Maxence de la Brassinne Bonardeaux, Siddharth Mishra-Sharma, Gilles Louppe.
Pre-print. arXiv:2408.15136 [PDF]

Some of the development papers

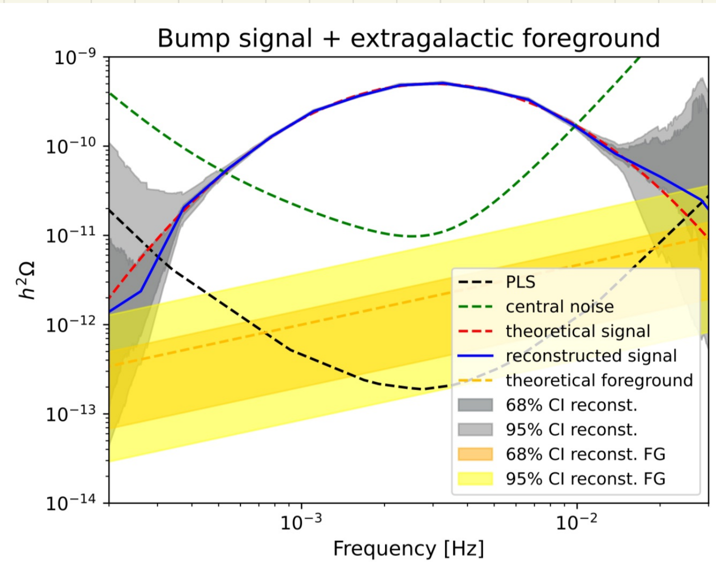
[open research line!]

Shallow Look at Bibliography

Fast likelihood-free reconstruction of gravitational wave backgrounds

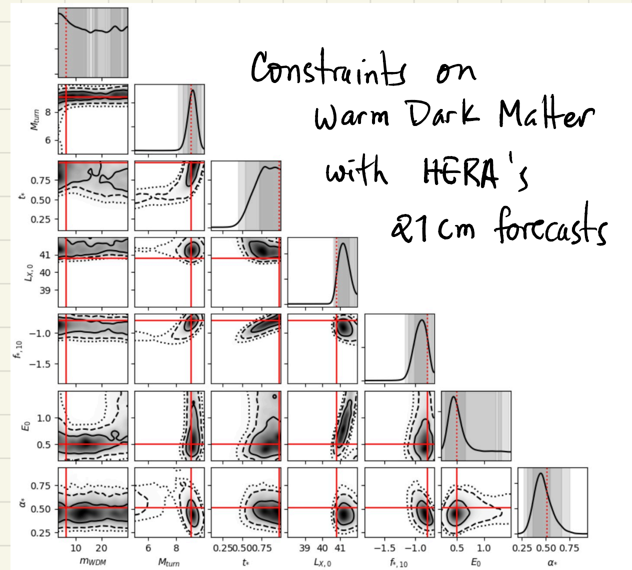
Androniki Dimitriou (Valencia U., IFIC), Daniel G. Figueroa (Valencia U., IFIC), Bryan Zaldivar IFIC) (Sep 15, 2023)

Published in: *JCAP* 09 (2024) 032 • e-Print: 2309.08430 [astro-ph.IM]



GW Backgrounds

(Decant, Dimitriou, Lopez-Honorez, Zaldivar)
TO APPEAR



Constraints on Warm Dark Matter with HERA's 21 cm forecasts

21 cm Cosmology

Shallow Look at Bibliography

Bayesian technique to combine independently-trained Machine-Learning models applied to direct dark matter detection

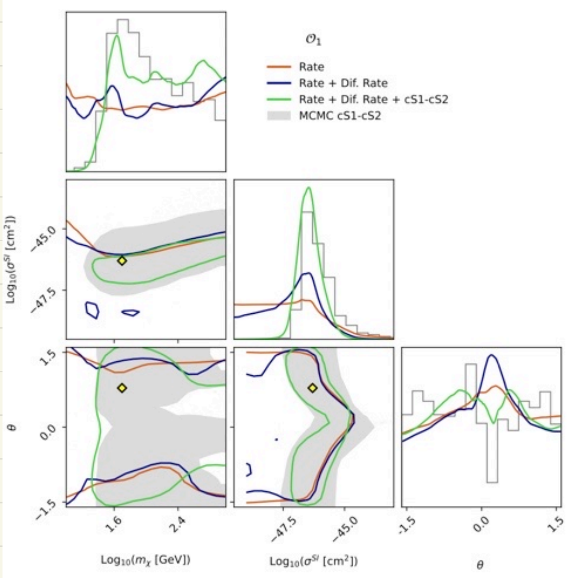
David Cerdeno (Madrid, IFT), Martin de los Rios (Madrid, IFT), Andres D. Perez (Madrid, IFT) (Jul 30, 2024)

e-Print: 2407.21008 [hep-ph]

Applying simulation-based inference to spectral and spatial information from the Galactic Center gamma-ray excess

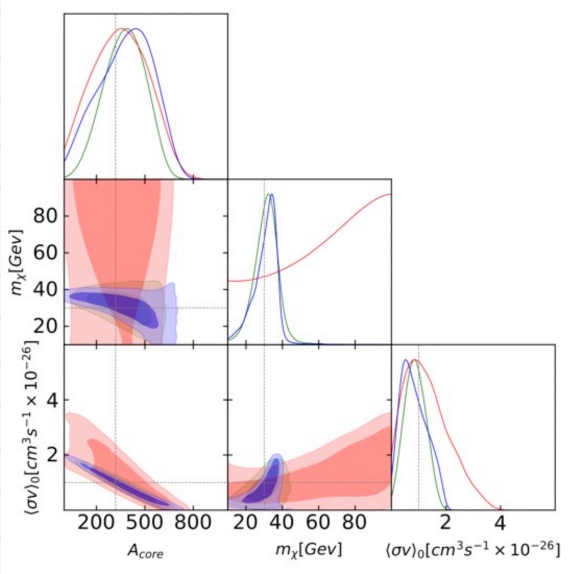
Katharena Christy (Hawaii U.), Eric J. Baxter (Inst. Astron., Honolulu), Jason Kumar (Hawaii U.) (Feb 6, 2024)

Published in: *JCAP* 07 (2024) 066 • e-Print: 2402.04549 [astro-ph.HE]

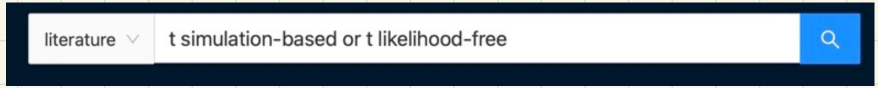


Direct Detection

Indirect Detection



Shallow Look at Bibliography



See :

some other authors
- Christoph Weniger
- Kyle Cranmer
- Benjamin Nachman
- Francesco Villaescusa-Navarro

Collaborations
- ATLAS
- DES
- COIN
- COSINUS
- LSST

PUBLIC PACKAGES

Python

Main SBI Techniques:

- Neural Posterior Estimation
- Neural Ratio Estimation
- Neural Likelihood Estimation

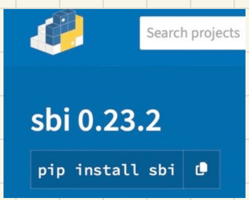
PUBLIC PACKAGES

Python

Main SBI Techniques:

- Neural Posterior Estimation
- Neural Ratio Estimation
- Neural Likelihood Estimation

Packages:



2007.09114

(Tejero-Cantero et al)

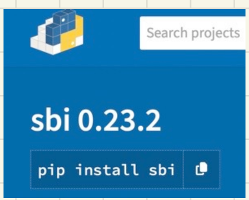
PUBLIC PACKAGES

Python

Main SBI Techniques:

- Neural Posterior Estimation
- Neural Ratio Estimation
- Neural Likelihood Estimation

Packages:



2007.09114
(Tejero-Cantero et al)

[tailored to collider physics]

MadMiner: ML based inference for particle physics

1907.10621
(Kyle Cranmer et al)

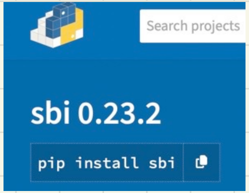
PUBLIC PACKAGES

Python

Main SBI Techniques:

- Neural Posterior Estimation
- Neural Ratio Estimation
- Neural Likelihood Estimation

Packages:



2007.09114
(Tejero-Cantero et al)

[tailored to collider Physics]

MadMiner: ML based inference for particle physics

1907.10621
(Kyle Cranmer et al)

SWYFT

2011.13951
(Weniger et al)

CONNECTING DARK MATTER CODES WITH SBI

Lagrangian

$$\mathcal{L}(\underbrace{m, y, \ldots}_{\vec{\theta}})$$

CONNECTING DARK MATTER CODES WITH SBI

Lagrangian

DM code

$$\mathcal{L}(\underbrace{m, y, \ldots}_{\vec{\theta}}) \Rightarrow \boxed{\text{MicrOmegas}}$$

- σ_{DD}, $\langle \sigma v \rangle$, ..., Ω_{DM}
- DM annihilation spectrum
- Distrib. of Recoil Energy
- DM momentum distribution

CONNECTING DARK MATTER CODES WITH SBI

Lagrangian

$\mathcal{L}(\underbrace{m, y, \ldots}_{\vec{\theta}})$

DM code

MicrOmegas

- σ_{DD}, $\langle \sigma v \rangle$, ..., Ω_{DM}
- DM annihilation Spectrum
- Distrib. of Recoil Energy
- DM momentum distribution

(as of today)

RECASTING

Experimental constraints

Collider (LHC, FCC, ...)

DD (Xenon, CDMS, ...)

DI (Fermi-LAT, AMS, Bullet C.)

Cosmo (Planck, HERA, LISA, ...)

CONNECTING DARK MATTER CODES WITH SBI

Lagrangian

$$\mathcal{L}(\underbrace{m, y, \dots}_{\vec{\theta}})$$

DM code

$\boxed{\text{MicrOmegas}}$

- σ_{DD}, $\langle \sigma v \rangle$, ..., Ω_{DM}
- DM annihilation Spectrum
- Distrib. of Recoil Energy
- DM momentum distribution

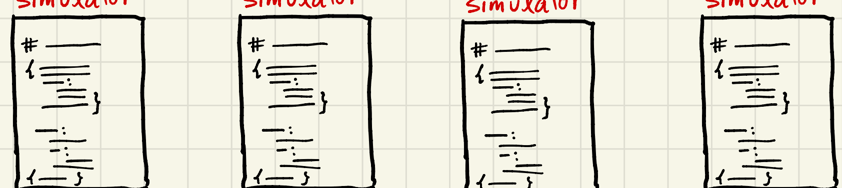
(as of today)
RECASTING

future
SBI

Experimental constraints

Collider (LHC, FCC, ...)
DD (Xenon, CDMS, ...)
DI (Fermi-LAT, AMS, Bullet C.)
Cosmo (Planck, HERA, LISA, ...)

DD simulator | DI simulator | Collider simulator | GWB simulator



Questions?