"Hyper-K physics workshop at DIPC"

(June 2023)

"MACHINE LEARNING IN PARTICLE PHYSICS"

Bryan Zaldivar (Valencia)

Outline

- What does machine learning bring to physics?
- Some examples of different uses of ML for Dark Matter [DISCLAIMER]
- Good practices when applying ML to physics

MOTIVATION

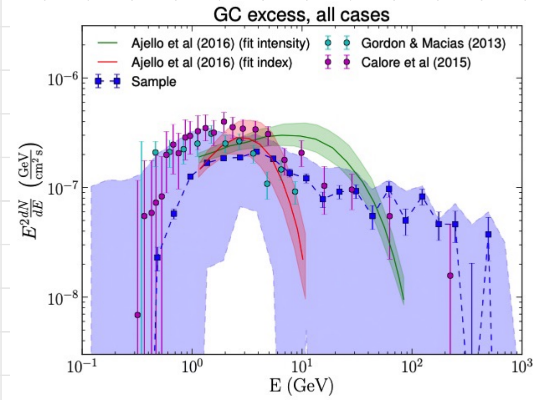
* Physical knowledge of background is limited
  (common problem in astrophysics)

MOTIVATION

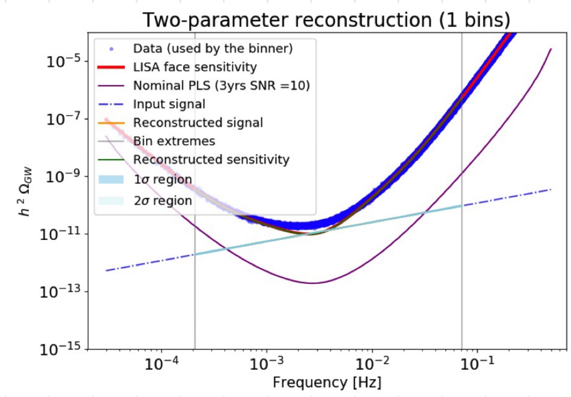
* Physical knowledge of background is limited
(common problem in astrophysics)

e.g. • γ-rays
• GW's

Fermi-LAT, 1704.03910



Caprini et al, 1906.09244

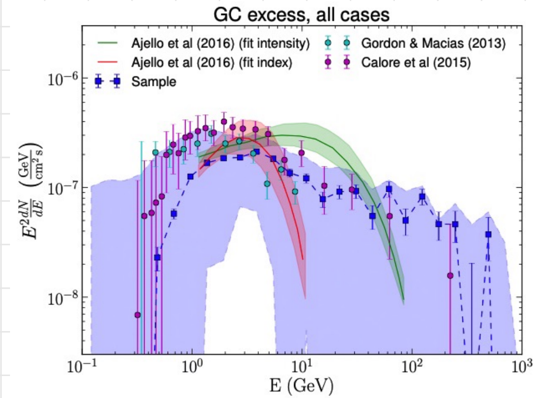


MOTIVATION

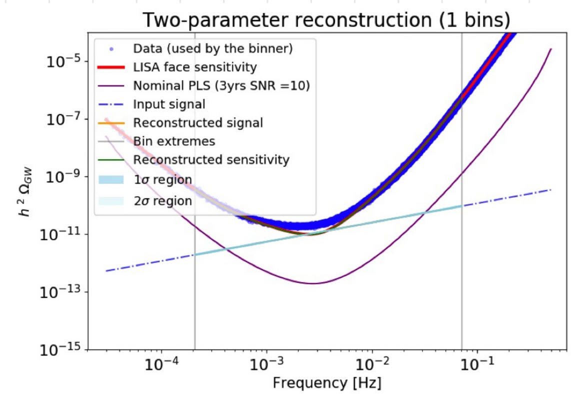
* Physical knowledge of background is limited
(common problem in astrophysics)

e.g. • γ-rays
• GW's

Fermi-LAT, 1704.03910



Caprini et al, 1906.09244



Signal + Background $\Longleftrightarrow$ Data

interest
(Physical model)

nuisance
(data-driven model)

MOTIVATION

* Physics well known, but observables very complicated to compute

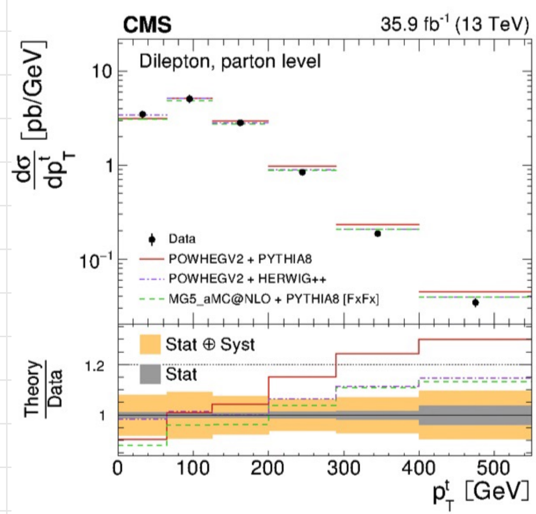
(complex topology, particle misidentification, etc)

MOTIVATION

* Physics well known, but observables very complicated to compute

(complex topology, particle misidentification, etc)

TOP-17-014-PAS

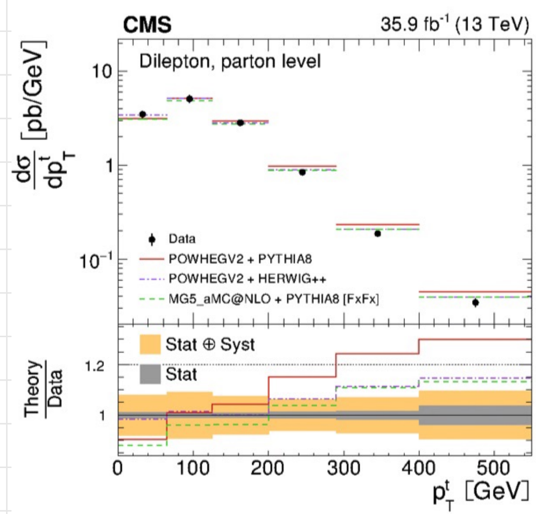


MOTIVATION

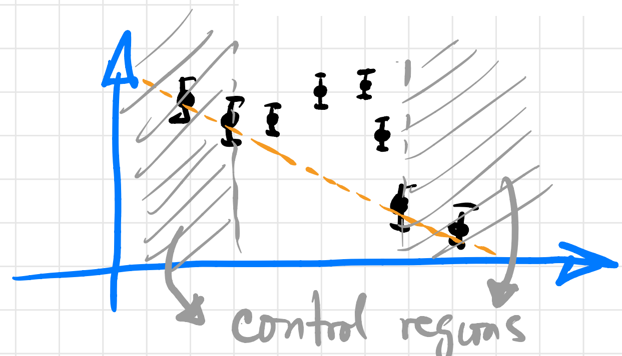
* Physics well known, but observables very complicated to compute

(complex topology, particle misidentification, etc)

TOP-17-014-PAS



Data-driven methods to infer the background from inter/extrapolations



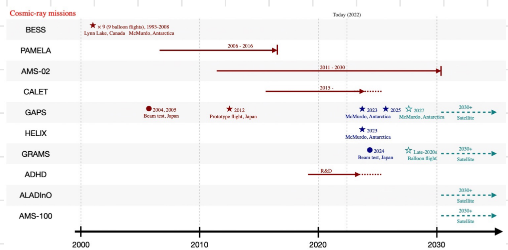
MOTIVATION

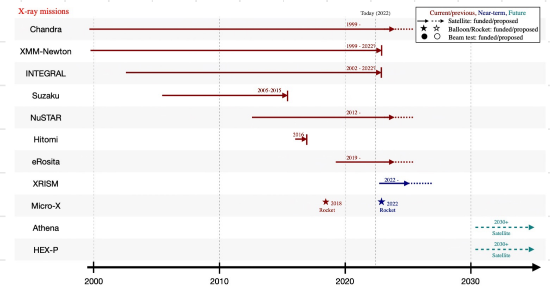
Snowmass, 2209.07426

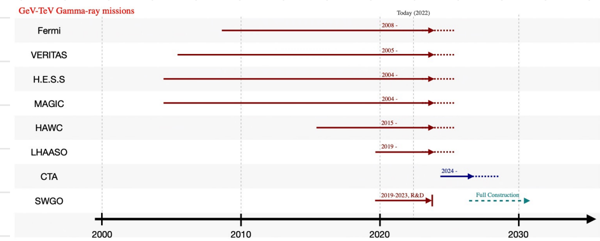
Name	Technology	Target	Active Mass	Experiment Location	Start Ops	End Ops
Currently Running or Under Construction						
LZ	TPC	LXe	7,000 kg	SURF	2021	2026
PandaX-4T	TPC	LXe	4,000 kg	CJPL	2021	2025
XENONnT	TPC	LXe	7,000 kg	LGNS	2021	2025
DEAP-3600	Scintillator	LAr	3,300 kg	SNOLAB	2016	2025
Darkside-20k	TPC	LAr	50 t	LNGS	2027	2035
DAMA/LIBRA	Scintillator	NaI	250 kg	LNGS	2003	
ANAIS-112	Scintillator	NaI	112 kg	Canfranc	2017	2022
SABRE PoP	Scintillator	NaI	5 kg	LNGS	2021	2022
COSINE-200	Scintillator	NaI	200 kg	YangYang	2022	2025
CDEX-10	Ionization (77K)	Ge	10 kg	CJPL	2016	
EDELWEISS III (High Field)	Cryo Ionization / HV	Ge	33 g	LSM	2019	
SuperCDMS CUTE	Cryo Ionization / HV	Ge/Si	5 kg/1 kg	SNOLAB	2020	2022
SuperCDMS SNOLAB	Cryo Ionization / HV	Ge/Si	11 kg/3 kg	SNOLAB	2023	2028
CRESST-III (HW Tests)	Bolometer Scintillation	CaWO4		LNGS	2020	
PICO-40	Bubble Chamber	C3F8	35 kg	SNOLAB	2020	
NEWS-G	Gas Drift	CH4		SNOLAB	2020	2025
DAMIC-M prototype	CCD Skipper	Si	18 g	LSM	2022	2023
DAMIC-M	CCD Skipper	Si	1 kg	LSM	2024	2025
SENSEI	CCD Skipper	Si	2 g	Fermilab	2019	2020
SENSEI	CCD Skipper	Si	100 g	SNOLAB	2021	2023

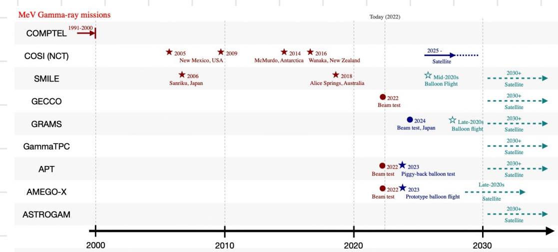
Name	Technology	Target	Active Mass	Experiment Location	Start Ops	End Ops
Planned						
SABRE (North)	Scintillator	NaI	50 kg	LNGS	2022	2027
SABRE (South)	Scintillator	NaI	50 kg	SUPL	2022	2027
COSINE-200 South Pole	Scintillator	NaI	200 kg	South Pole	2023	
COSINUS	Bolometer Scintillator	NaI		LNGS	2023	
Darwin / XLZD (US LXe G3)	TPC	LXe	50,000 kg	undetermined	2028	2033
ARGO	TPC or Scintillator	LAr	300 t	SNOLAB	2030	2035
CDEX-100 / 1T	Ionization (77K)	Ge	100-1000 kg	CJPL	202X	
PICO-500	Bubble Chamber	C3F8	430 kg	SNOLAB	2021	
Concept or R&D						
Oscura	CCD Skipper	Si	10 kg Si	SNOLAB	2025	2028
SBC	Bubble Chamber	LAr	1 t	SNOLAB	2028	
SNOWBALL	Supercooled Liquid H2O					
DarkSide-LowMass	TPC	LAr	1.5 t			
ALETHEIA	TPC	He		China Inst. At. Energy		
TESSERACT	Cryo TES	LHe, SiO_2, Al_2O_3, GaAs		undetermined	2026	
CYGNO	Gas Directional	He + CF_4	0.5 - 1 kg	LNGS	2024	
CYGNUS	Gas Directional	He + SF_6/CF_4		Multiple sites		
Windchime	Accelerometer array			Multiple sites		
MAGNETO-χ	Cryogenic MMC	Diamond, Sapphire, etc.				

Experiment	Final state	Threshold/sensitivity	Field of view	Location
Current experiments				
Fermi	Photons	10 MeV – 10^3 GeV	Wide	Space
HESS	Photons	30 GeV - 100 TeV	Targeted	Namibia
VERITAS	Photons	85 GeV - > 30 TeV	Targeted	USA
MAGIC	Photons	30 GeV - 100 TeV	Targeted	Spain
HAWC	Photons	300 GeV - >100 TeV	Wide	Mexico
LHAASO (partial)	Photons	10 TeV - 10 PeV	Wide	China
KASCADE	Photons	100 TeV - 10 PeV	Wide	Germany
KASCADE-Grande	Photons	10 - 100 PeV	Wide	Italy
Pierre Auger Observatory	Photons	1 EeV - 1 ZeV	Wide	Argentina
Telescope Array	Photons	1 - 100 EeV	Wide	USA
IceCube	Neutrinos	100 TeV - 100 EeV	Wide	Antarctica
ANITA	Neutrinos	EeV - ZeV	Wide	Antarctica
Pierre Auger Observatory	Photons & Neutrinos	1 EeV - 1 ZeV	Wide	Argentina
Future experiments				
CTA	Photons	20 GeV - 300 TeV	Targeted	Chile & Spain
SWGO	Photons	100 GeV - 1 PeV	Wide	South America
IceCube-Gen2	Neutrinos	10 TeV - 100 EeV	Wide	Antarctica
LHAASO (full)	Photons	100 GeV - 10 PeV	Wide	China
KM3NeT	Neutrinos	100 GeV - 10 PeV	Wide	Mediterranean Sea
AugerPrime	Photons & Neutrinos	1 EeV - 1 ZeV	Wide	Argentina
POEMMA	Neutrinos	20 PeV - 100 EeV	Wide	Space









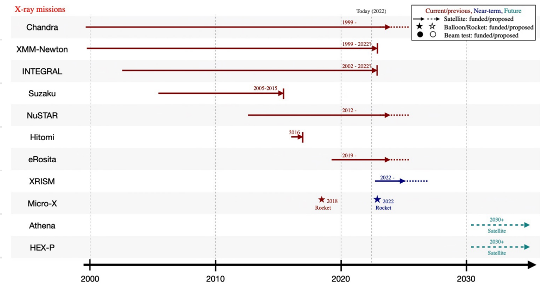
MOTIVATION

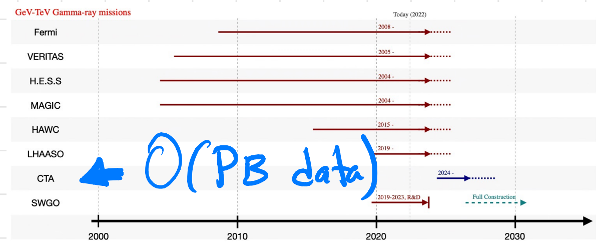
Snowmass, 2209.07426

Name	Technology	Target	Active Mass	Experiment Location	Start Ops	End Ops
Currently Running or Under Construction						
LZ	TPC	LXe	7,000 kg	SURF	2021	2026
PandaX-4T	TPC	LXe	4,000 kg	CJPL	2021	2025
XENONnT	TPC	LXe	7,000 kg	LGNS	2021	2025
DEAP-3600	Scintillator	LAr	3,300 kg	SNOLAB	2016	2025
Darkside-20k	TPC	LAr	50 t	LNGS	2027	2035
DAMA/LIBRA	Scintillator	NaI	250 kg	LNGS	2003	
ANAIS-112	Scintillator	NaI	112 kg	Canfranc	2017	2022
SABRE PoP	Scintillator	NaI	5 kg	LNGS	2021	2022
COSINE-200	Scintillator	NaI	200 kg	YangYang	2022	2025
CDEX-10	Ionization (77K)	Ge	10 kg	CJPL	2016	
EDELWEISS III (High Field)	Cryo Ionization / HV	Ge	33 g	LSM	2019	
SuperCDMS CUTE	Cryo Ionization / HV	Ge/Si	5 kg/1 kg	SNOLAB	2020	2022
SuperCDMS SNOLAB	Cryo Ionization / HV	Ge/Si	11 kg/3 kg	SNOLAB	2023	2028
CRESST-III (HW Tests)	Bolometer Scintillation	CaWO4		LNGS	2020	
PICO-40	Bubble Chamber	C3F8	35 kg	SNOLAB	2020	
NEWS-G	Gas Drift	CH4		SNOLAB	2020	2025
DAMIC-M prototype	CCD Skipper	Si	18 g	LSM	2022	2023
DAMIC-M	CCD Skipper	Si	1 kg	LSM	2024	2025
SENSEI	CCD Skipper	Si	2 g	Fermilab	2019	2020
SENSEI	CCD Skipper	Si	100 g	SNOLAB	2021	2023

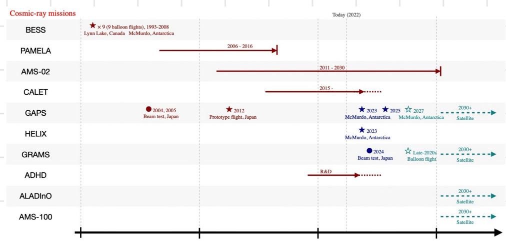
Name	Technology	Target	Active Mass	Experiment Location	Start Ops	End Ops
Planned						
SABRE (North)	Scintillator	NaI	50 kg	LNGS	2022	2027
SABRE (South)	Scintillator	NaI	50 kg	SUPL	2022	2027
COSINE-200 South Pole	Scintillator	NaI	200 kg	South Pole	2023	
COSINUS	Bolometer Scintillator	NaI		LNGS	2023	
Darwin / XLZD (US LXe G3)	TPC	LXe	50,000 kg	undetermined	2028	2033
ARGO	TPC or Scintillator	LAr	300 t	SNOLAB	2030	2035
CDEX-100 / 1T	Ionization (77K)	Ge	100-1000 kg	CJPL	202X	
PICO-500	Bubble Chamber	C3F8	430 kg	SNOLAB	2021	
Concept or R&D						
Oscura	CCD Skipper	Si	10 kg Si	SNOLAB	2025	2028
SBC	Bubble Chamber	LAr	1 t	SNOLAB	2028	
SNOWBALL	Supercooled Liquid H2O					
DarkSide-LowMass	TPC	LAr	1.5 t			
ALETHEIA	TPC	He		China Inst. At. Energy		
TESSERACT	Cryo TES	LHe, SiO_2, Al_2O_3, GaAs		undetermined	2026	
CYGNO	Gas Directional	He + CF_4	0.5 - 1 kg	LNGS	2024	
CYGNUS	Gas Directional	He + SF_6/CF_4		Multiple sites		
Windchime	Accelerometer array			Multiple sites		
MAGNETO-χ	Cryogenic MMC	Diamond, Sapphire, etc.				

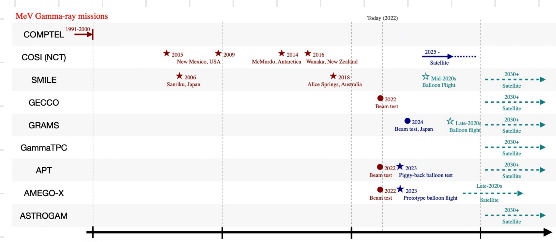
Experiment	Final state	Threshold/sensitivity	Field of view	Location
Current experiments				
Fermi	Photons	10 MeV – 10^3 GeV	Wide	Space
HESS	Photons	30 GeV - 100 TeV	Targeted	Namibia
VERITAS	Photons	85 GeV - > 30 TeV	Targeted	USA
MAGIC	Photons	30 GeV - 100 TeV	Targeted	Spain
HAWC	Photons	300 GeV - >100 TeV	Wide	Mexico
LHAASO (partial)	Photons	10 TeV - 10 PeV	Wide	China
KASCADE	Photons	100 TeV - 10 PeV	Wide	Germany
KASCADE-Grande	Photons	10 - 100 PeV	Wide	Italy
Pierre Auger Observatory	Photons	1 EeV - 1 ZeV	Wide	Argentina
Telescope Array	Photons	1 - 100 EeV	Wide	USA
IceCube	Neutrinos	100 TeV - 100 EeV	Wide	Antarctica
ANITA	Neutrinos	EeV - ZeV	Wide	Antarctica
Pierre Auger Observatory	Photons & Neutrinos	1 EeV - 1 ZeV	Wide	Argentina
Future experiments				
CTA	Photons	20 GeV - 300 TeV	Targeted	Chile & Spain
SWGO	Photons	100 GeV - 1 PeV	Wide	South America
IceCube-Gen2	Neutrinos	10 TeV - 100 EeV	Wide	Antarctica
LHAASO (full)	Photons	100 GeV - 10 PeV	Wide	China
KM3NeT	Neutrinos	100 GeV - 10 PeV	Wide	Mediterranean Sea
AugerPrime	Photons & Neutrinos	1 EeV - 1 ZeV	Wide	Argentina
POEMMA	Neutrinos	20 PeV - 100 EeV	Wide	Space





O(PB data)





O(40) experiments operative in near future!

MOTIVATION

* Statistical bottleneck

MOTIVATION

* Statistical bottleneck

- More complex datasets

$\Downarrow$

More complex physical modeling

$\Downarrow$

More complex simulator

MOTIVATION

* Statistical bottleneck

- More complex datasets

↓

More complex physical modeling

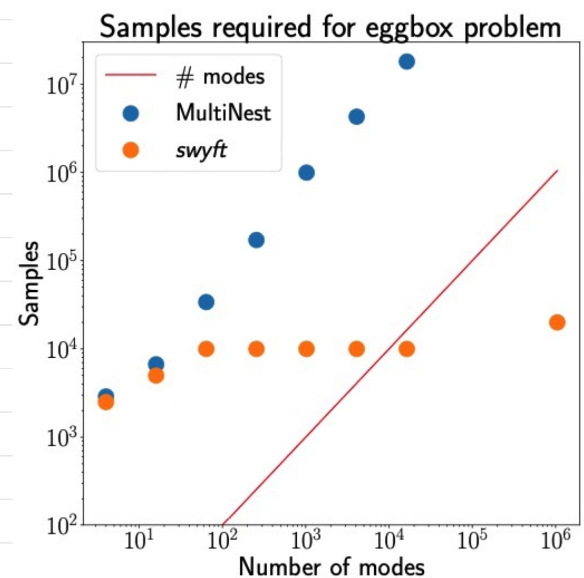
↓

More complex simulators

↓

Better statistical treatment!

Miller et al, 2011.13951



- better scaling
- more descriptive
- Higher statistical power

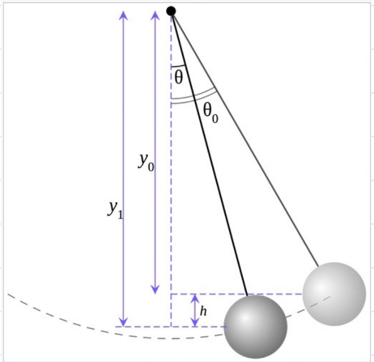
MOTIVATION

* Hints about the underlying physics

MOTIVATION

* Hints about the underlying physics

- The intrinsic dimension of → # physical variables
  - Single pendulum : 2
  - Lava lamp : ?



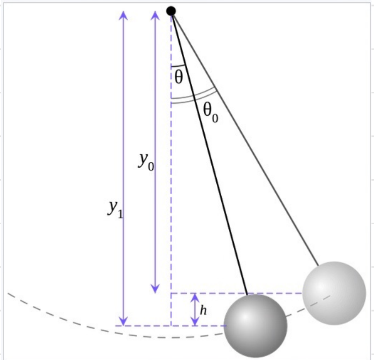
MOTIVATION

* Hints about the underlying physics

- The intrinsic dimension of

physical variables

- Single pendulum : 2
- Lava lamp : ?



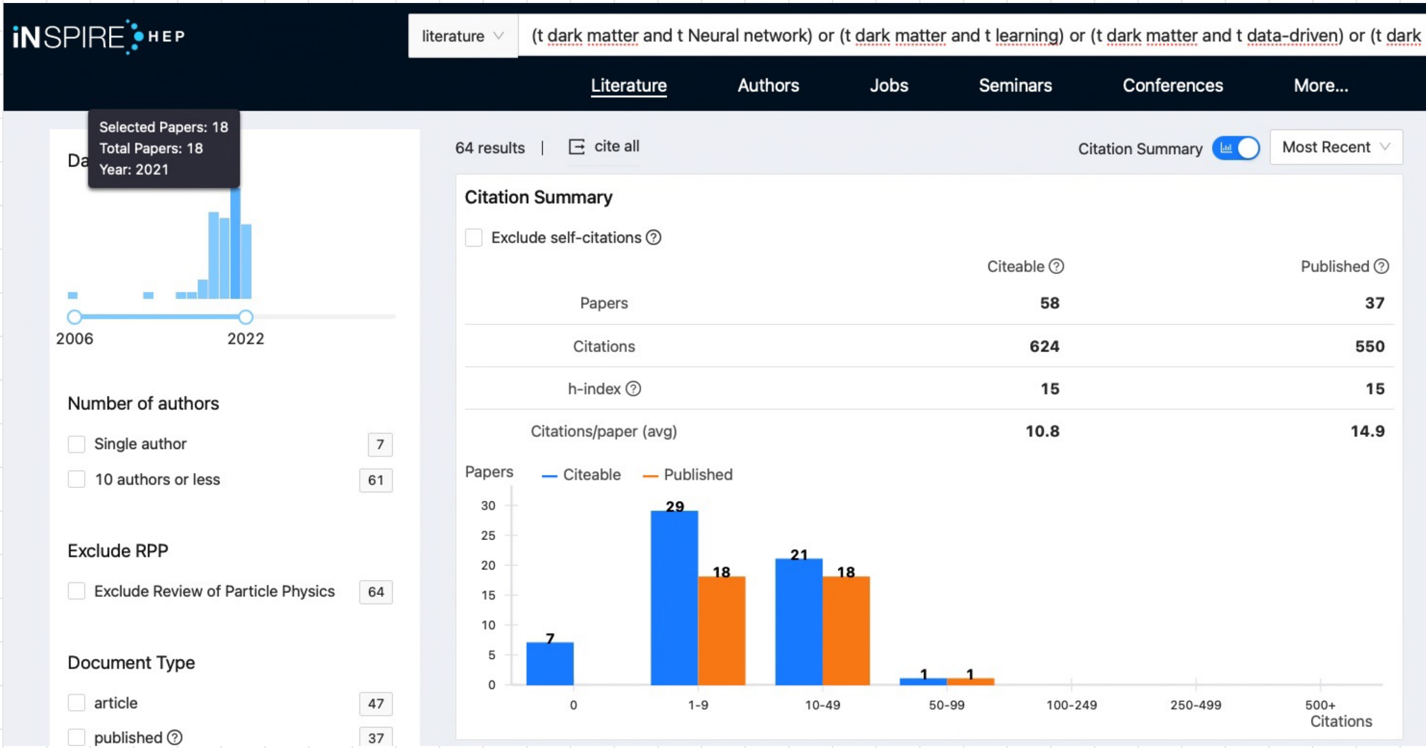
Statistical model (ML)

ID = 7 − 8

(see 2112.10755)

DARK MATTER WITH MACHINE LEARNING

[likely same trend in other areas]



Two main ML paradigms
used in physics

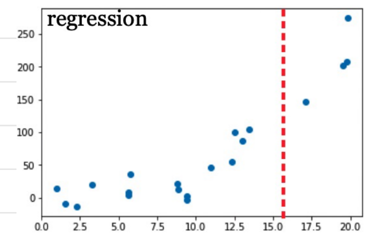
SUPERVISED LEARNING

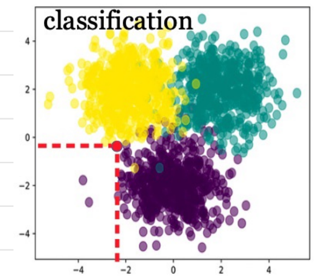
Tasks

1) Estimate underlying distribution

$$y(\vec{x}) \sim p\left(f(\vec{x};\vec{\theta}), \sigma\right)$$

$$\text{data} = \{\vec{x}_i, y_i\}$$





SUPERVISED LEARNING

Tasks

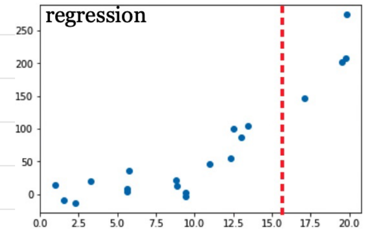
$\text{data} = \{\vec{x}_i, y_i\}$

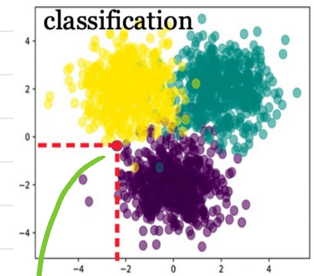
1) Estimate underlying distribution

$$y(\vec{x}) \sim p\left(f(\vec{x};\vec{\theta}), \sigma\right)$$

- Assume shape of p,
  estimate $f(\vec{x},\vec{\theta})$ with a ML model

[useful for estimating nuisance contributions,
(c.f. Motivation #1)]





2) Prediction for new $\vec{x}$

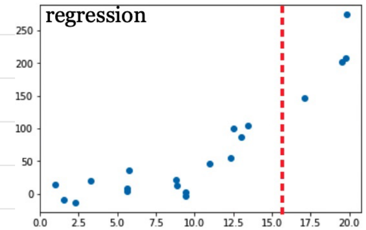
SUPERVISED LEARNING

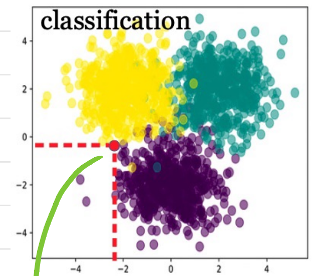
Tasks

data = $\{\vec{x}_i, y_i\}$

1) Estimate underlying distribution

$$y(\vec{x}) \sim p\left(f(\vec{x};\vec{\theta}), \sigma\right)$$





- Assume shape of p,
  estimate $f(\vec{x},\vec{\theta})$ with a ML model

  [useful for estimating nuisance contributions,
  (c.f. Motivation #1)]

2) Prediction for new $\vec{x}$

- $f(\vec{x},\vec{\theta})$ given by a physics model (infer physical parameters $\vec{\theta}$)
  Estimate shape of p with a ML procedure

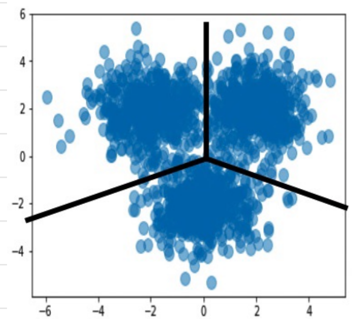
UNSUPERVISED LEARNING

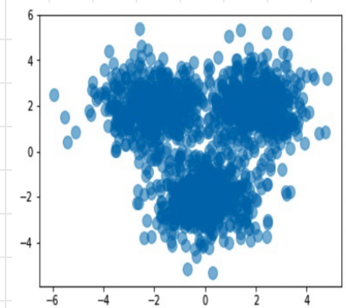
data $= \{ \vec{x}_i \}$

Tasks

1) Clusterize the data

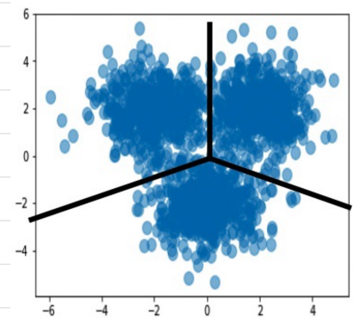
(very useful e.g. in collider physics)

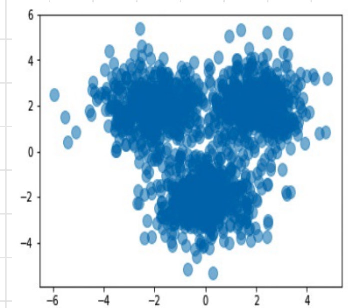




UNSUPERVISED LEARNING

data $= \{\vec{x}_i\}$

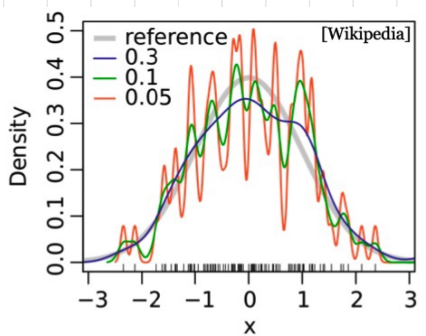




Tasks

1) Clusterize the data
(very useful e.g. in collider physics)

2) Probability density estimation
(astrophysics, cosmology, everywhere...)

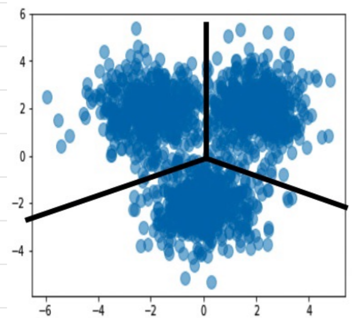


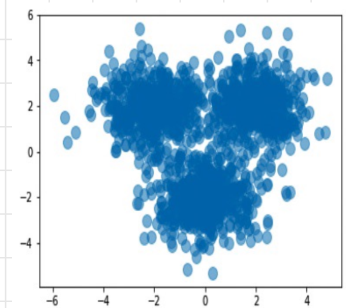
UNSUPERVISED LEARNING

data = $\{\vec{x}_i\}$

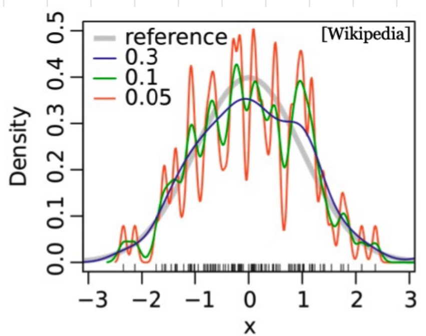
Tasks

1) Clusterize the data
(very useful e.g. in collider physics)

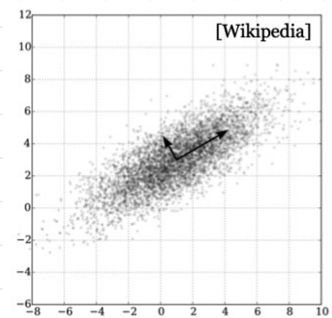




2) Probability density estimation
(astrophysics, cosmology, everywhere...)



3) Dimensionality reduction
(everywhere)



Some examples of applications

[by no means exhaustive!!]

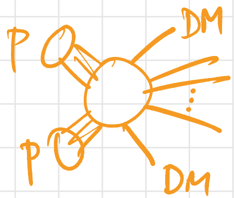
I DM searches at the LHC (Banerjee et al, 1705.02327)

(I) DM searches at the LHC (Banerjee et al, 1705.02327)

- What is the motivating issue?

Final state topology may be quite complex

$\Rightarrow$ potentially large phase space!



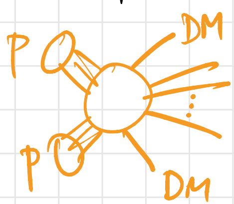
traditional analyses (e.g. cut-based) may give poor performance

(I) DM searches at the LHC (Banerjee et al, 1705.02327)

• What is the motivating issue?

Final state topology may be quite complex

$\Rightarrow$ potentially large phase space!



traditional analyses (e.g. cut-based) may give poor performance

• How to address it?

Frame it into a ML classification problem

(optimize signal-to-bckg ratio)

ML classifier is (formally) an approx to the

$$TS = \frac{p(x\,|\,H_0)}{p(x\,|\,H_1)}$$

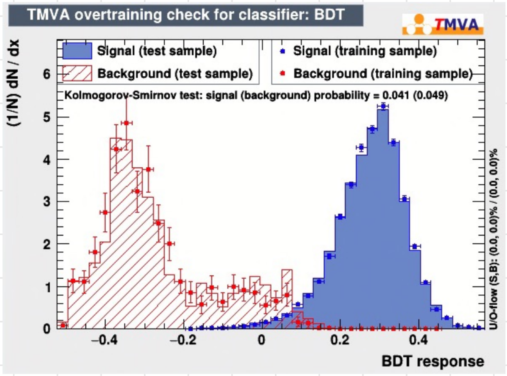
highest statistical power

$\hookrightarrow$ Neural network, Boosted Decision Trees, ...

I) DM searches at the LHC

(Banerjee et al, 1705.02327)

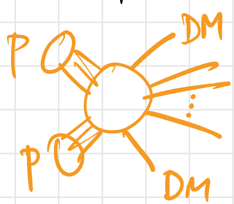
↳ simple case study with a monojet signal



- What is the motivating issue?

Final state topology may be quite complex

$\Rightarrow$ potentially large phase space!



traditional analyses (e.g. cut-based) may give poor performance

- How to address it?

Frame it into a ML classification problem

(optimize signal-to-bckg ratio)

ML classifier is (formally) an approx to the

$$TS = \frac{p(x \mid H_0)}{p(x \mid H_1)}$$

} highest statistical power

↳ Neural network, Boosted Decision Trees, ...

II Direct Detection searches (Herrero-Garcia et al 2110.12248)

II Direct Detection searches (Herrero-Garcia et al 2110.12248)

Motivation

A priori different DM candidates will produce different patterns in detectors

Seems complicated to consider all possible types of signals

II Direct Detection searches (Herrero-Garcia et al 2110.12248)

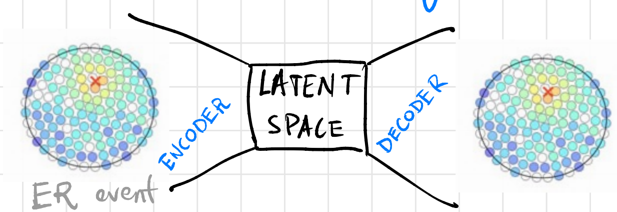
Motivation

A priori different DM candidates will produce different patterns in detectors

Seems complicated to consider all possible types of signals

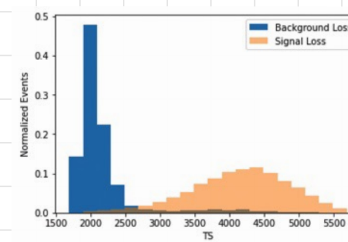
Solution: "Anomaly awareness" (unsupervised) [2007.14462]

1) Train a NN to learn the background



2) Feed it with signal events

The resulting distribution should be distinguishable from the bckg one



(II) Direct Detection searches (Herrero-Garcia et al 2110.12248)

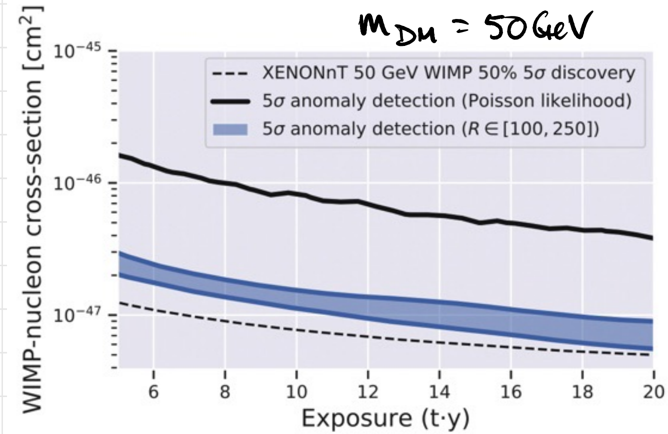
Motivation

A priori different DM candidates will produce different patterns in detectors

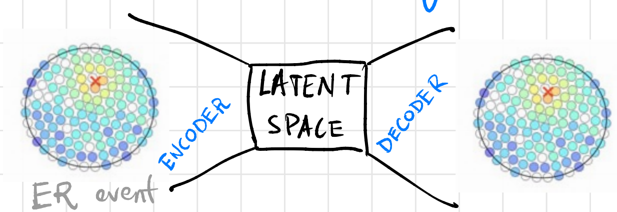
Seems complicated to consider all possible types of signals

Solution: "Anomaly awareness" (unsupervised) [2007.14462]

$m_{DM} = 50\,\mathrm{GeV}$

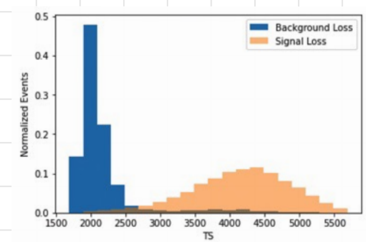


1) Train a NN to learn the background



2) Feed it with signal events

The resulting distribution should be distinguishable from the bckg one



III Indirect Detection Searches

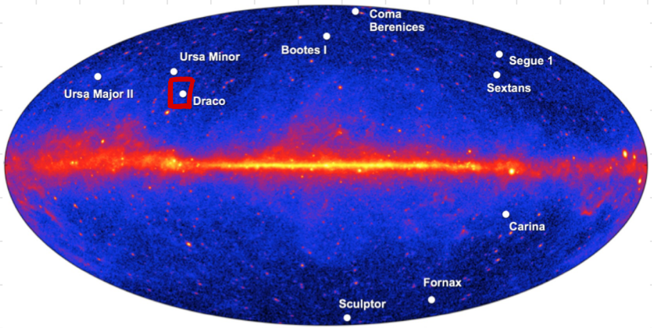
(modeling the background)

(Calore, Serpico, Zaldivar, 1803.05508)

III Indirect Detection Searches
(modeling the background)

(Calore, Serpico, Zaldivar, 1803.05508)

Fermi γ-rays at dSphs



Motivation

1) Typical bckg estimation has limitations
   * Templates are biased (limited knowledge of physics)
   * Local fitting of templates (no global consistency)
   * No clear way of accounting for prediction uncertainties

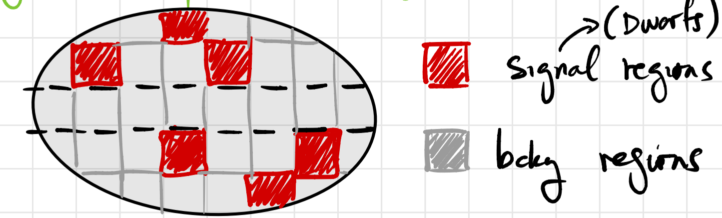
(Calore, Serpico, Zaldivar, 1803.05508) cont...

Aiming for a global, unbiased estimation of the background probability distribution

(Calore, Serpico, Zaldivar, 1803.05508) cont...

Aiming for a global, unbiased estimation of the background probability distribution

Idea:



(Calore, Serpico, Zaldivar, 1803.05508) cont...

Aiming for a global, unbiased estimation of the background probability distribution

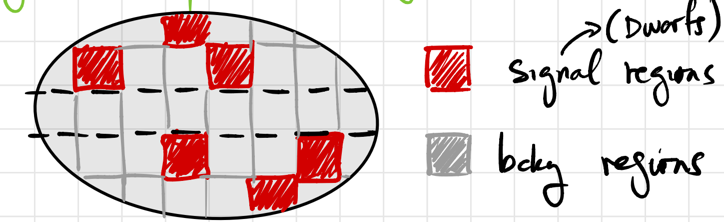
$$\text{Data} = \{\vec{x}_i, y_i\}_{i=1}^{N}$$

$N \to$ all pixels of the sky

$\vec{x}_i$: spatial coordinates

y_i: counts

Idea:



$$p(\vec{x}, y) = \frac{1}{N} \sum_i k_\alpha(\vec{x}-\vec{x}_i) \cdot g_\beta(y-y_i)$$

↳ unbiased estimator of the true PDF

$k_\alpha(\cdot)$, $g_\beta(\cdot)$: kernel functions

(Calore, Serpico, Zaldivar, 1803.05508) cont...

Aiming for a global, unbiased estimator of the background probability distribution

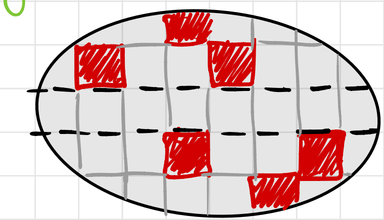
$$\text{Data} = \{\vec{x}_i , y_i\}_{i=1}^{N}$$

$N \to$ all pixels of the sky

spatial coordinates ($\vec{x}_i$)

counts (y_i)

Idea:



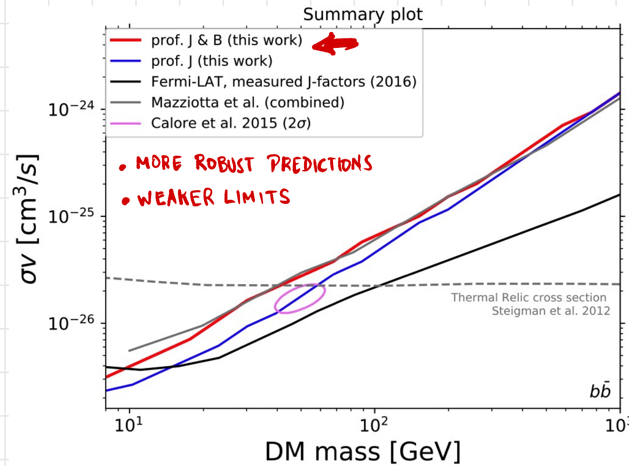
signal regions $\to$ (Dwarfs)

bckg regions

$$p(\vec{x}, y) = \frac{1}{N} \sum_i k_\alpha(\vec{x}-\vec{x}_i) \cdot g_\beta (y - y_i)$$

$\to$ unbiased estimator of the true PDF

$k_\alpha(\cdot), g_\beta(\cdot)$: kernel functions



IV DM searches from Fermi's unidentified PLS's

(Gammaldi et al, 2207.09307)

IV DM searches from Fermi's unidentified PLS's

(Gammaldi et al, 2207.09307)

- 4FGL catalog ≈ 5000 sources $\Rightarrow$ 1/3 are unID's — Is DM shining there?

  (pulsars, quasars, blazars, ...)

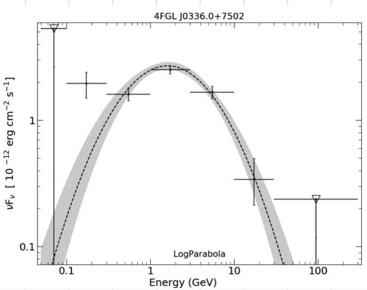
IV DM searches from Fermi's unidentified PLS's

(Gammaldi et al, 2207.09307)

- 4FGL catalog ≈ 5000 sources → 1/3 are unID's → Is DM shining there?

(pulsars, quasars, blazars, ...)

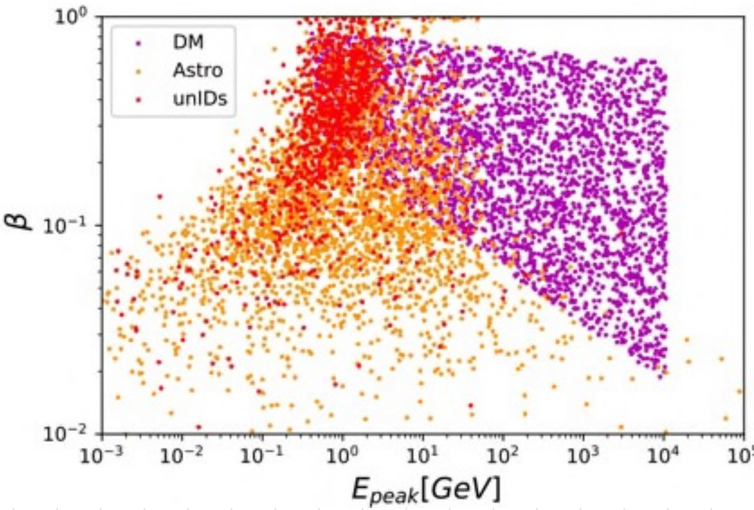
Astrophysical sources vs. Dark Matter sources



✓ ✓ Energy spectrum

✓ X Exp. uncertainties

✓ X Detection significance



Idea: DM inherits statistical properties of the unID's

(Gammaldi et al, 2207.09307) cont...

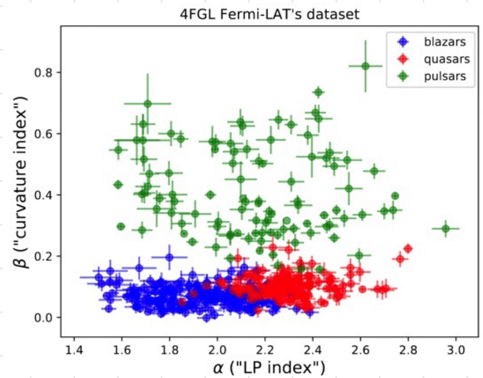
A standard classification problem?

Not so: input data have uncertainties

(Gammaldi et al, 2207.09307) cont...

A standard classification problem? Not so: input data have uncertainties } surprisingly not addressed before in the ML literature!

[Villacampa et al, 2001.10523]



Gaussian Process model

$\vec{x} \sim N(\vec{\mu}, \vec{\sigma})$

latent variables to be inferred

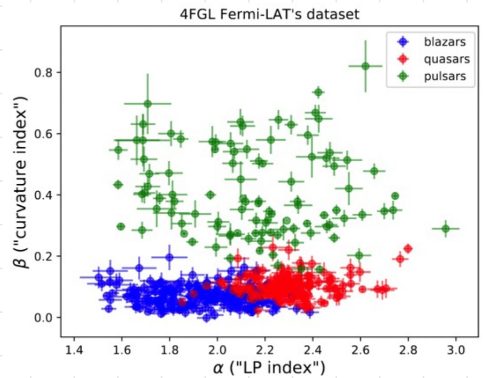
(Gammaldi et al, 2207.09307) cont...

A standard classification problem?

Not so: input data have uncertainties } surprisingly not addressed before in the ML literature!

- Another way ⇒ uncertainties as separate input variables

[Villacampa et al, 2001.10523]



* Train a NN on Astro-vs-DM

* NN output is probability for DM class

* Make predictions for unID's

Gaussian Process model

$\vec{x} \sim N(\vec{\mu}, \vec{\sigma})$

latent variables to be inferred

(Gammaldi et al, 2207.09307) cont...

A standard classification problem?

Not so: input data have uncertainties } surprisingly not addressed before in the ML literature!

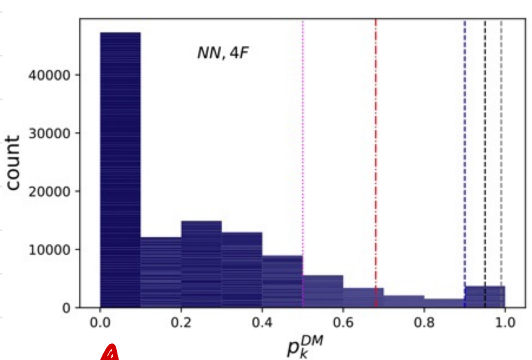
- Another way ⇒ uncertainties as separate input variables

[Villacampa et al, 2001.10523]

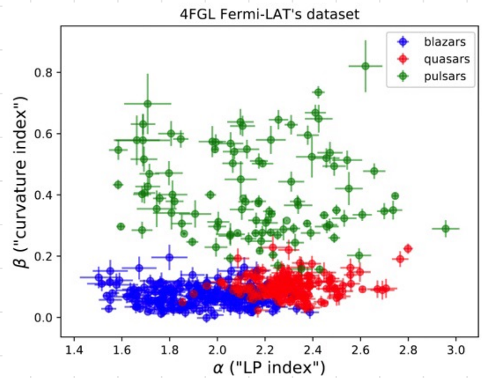
* Train a NN on Astro-vs-DM

* NN output is probability for DM class

* Make predictions for unID's



No statistically significant evidence for DM



Gaussian Process model

$\vec{x} \sim N(\vec{\mu}, \vec{\sigma})$

latent variables to be inferred

V Warm dark matter mass from strong lensing

(Anau Montel et al, 2205.09126)

V Warm dark matter mass from strong lensing

(Anau Montel et al, 2205.09126)

What is the motivating issue?

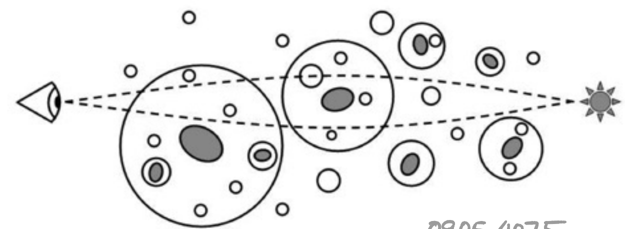
Characterizing collective subhalo properties from gravitational lensing is very difficult!

(V) Warm dark matter mass from strong lensing

(Anau Montel et al, 2205.09126)

What is the motivating issue?

Characterizing collective subhalo properties from gravitational lensing is very difficult!



0905.4075

nuisance
- Size and position of the main lens
- Brightness distribution and position of the source
- Position, mass, distance, etc of all the subhalos

→ Warm dark matter mass ← parameter of interest

Typical MCMC algorithms don't scale well with dimensionality of the param. space

(Anau Montel et al, 2205.09126) cont...

- getting samples from $p(\vec{\theta}, \vec{\alpha} \mid \text{data})$ is typically intractable

(Anau Montel et al, 2205.09126) cont...

- getting samples from $p(\vec{\theta}, \vec{\alpha} \,|\, \text{data})$ is typically intractable

Alternative

- $p(\vec{\theta} \,|\, \text{data}) = p(\text{data} \,|\, \vec{\theta})\, p(\vec{\theta}) / p(\text{data})$

$$p(\text{data} \,|\, \vec{\theta}) = \int d\vec{\alpha}\; p(\text{data} \,|\, \vec{\alpha}, \vec{\theta})\; p(\vec{\alpha} \,|\, \vec{\theta})$$ $\Rightarrow$ also intractable!

(Anau Montel et al, 2205.09126) cont...

- getting samples from $p(\vec{\theta}, \vec{\alpha} \mid \text{data})$ is typically intractable

Alternative

- $p(\vec{\theta} \mid \text{data}) = p(\text{data} \mid \vec{\theta})\, p(\vec{\theta}) / p(\text{data})$

$$p(\text{data} \mid \vec{\theta}) = \int d\vec{\alpha}\; p(\text{data} \mid \vec{\alpha}, \vec{\theta})\, p(\vec{\alpha} \mid \vec{\theta}) \Rightarrow \text{also intractable!}$$

Solution

Estimate $p(\vec{\theta} \mid \text{data})$ directly with a neural network

$\Downarrow$

no need to evaluate the likelihood

(Anau Montel et al, 2205.09126) cont...

- getting samples from $p(\vec{\theta}, \vec{\alpha} \mid \text{data})$ is typically intractable

Alternative

- $p(\vec{\theta} \mid \text{data}) = p(\text{data} \mid \vec{\theta})\, p(\vec{\theta}) / p(\text{data})$

$$p(\text{data} \mid \vec{\theta}) = \int d\vec{\alpha}\; p(\text{data} \mid \vec{\alpha}, \vec{\theta})\, p(\vec{\alpha} \mid \vec{\theta})$$ $\Rightarrow$ also intractable!

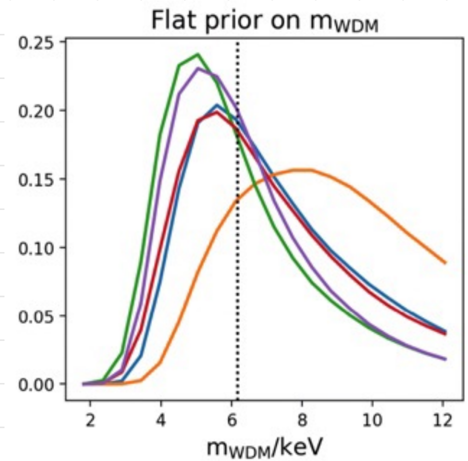
Solution

Estimate $p(\vec{\theta} \mid \text{data})$ directly with a neural network

$\Downarrow$

no need to evaluate the likelihood

$\Bigg\}$ Inferring the warm DM mass $\Rightarrow$



Important warnings to beware of

i) Occam's razor

i) Occam's razor

Facts:

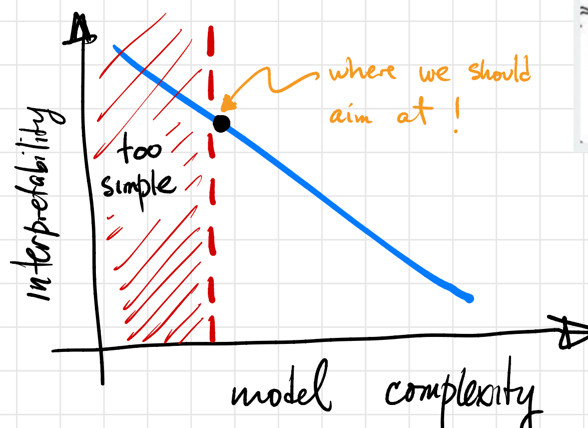
- Very large models (DNNs) have built-in overfitting control
- Data scientists typically adopt a "one-hammer-for-all" strategy (motivated by industry problems)
- Many physics applications adopting the same strategy

i) Occam's razor

Facts : - Very large models have built-in overfitting control (DNNs)

- Data scientists typically adopt a "one-hammer-for-all" strategy (motivated by industry problems)

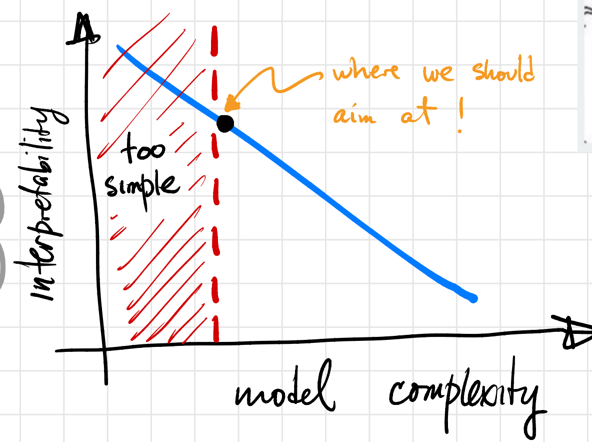
- Many physics applications adopting the same strategy



i) Occam's razor

Facts :

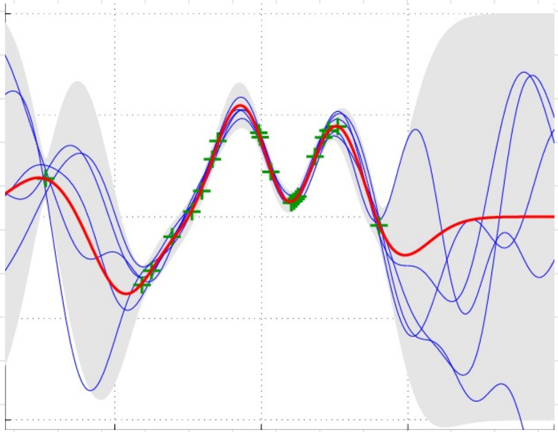
- Very large models (DNNs) have built-in overfitting control
- Data scientists typically adopt a "one-hammer-for-all" strategy (motivated by industry problems)
- Many physics applications adopting the same strategy



Remember that polymials are also -as neural nets- Universal approximators !!

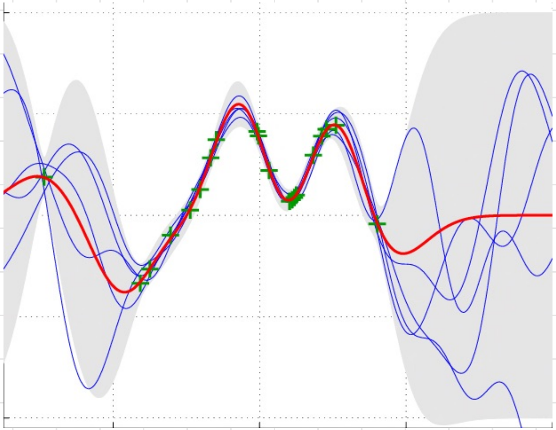
ii) Uncertainties related to the modeling itself

(a.k.a. "epistemic" uncertainties)



ii) Uncertainties related to the modeling itself

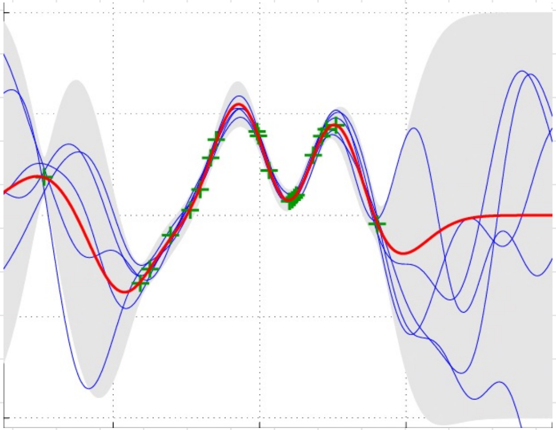
(a.k.a. "epistemic" uncertainties)



- Epistemic uncertainties could be dominant in regions with few/no data points (worse for more expressive models)

ii) Uncertainties related to the modeling itself

(a.k.a. "epistemic" uncertainties)



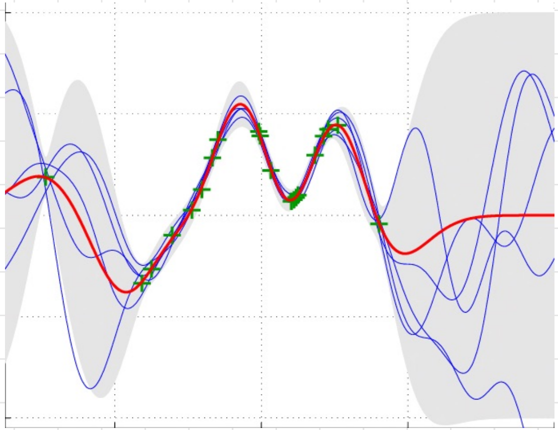
- Epistemic uncertainties could be dominant in regions with few/no data points (worse for more expressive models)

- The larger the model, the more challenging is the uncertainty estimate (large body of dedicated literature)

ii) Uncertainties related to the modeling itself

(a.k.a. "epistemic" uncertainties)

- Epistemic uncertainties could be dominant in regions with few/no data points (worse for more expressive models)



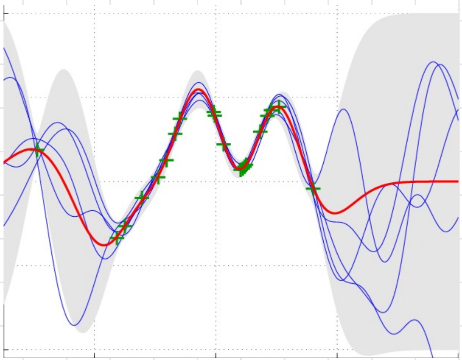
- The larger the model, the more challenging is the uncertainty estimate (large body of dedicated literature)

- Cats don't have error bars (maybe) but in science, uncertainty estimation is mandatory!

iii) ML predictions vs. physical plausibility

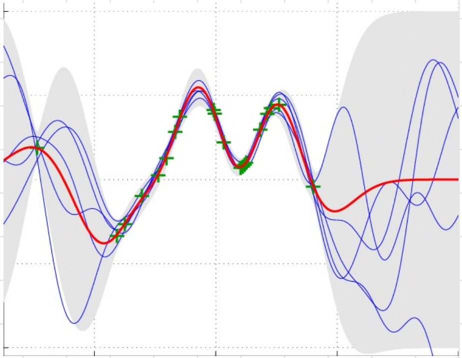
How many of these possible solutions are physically meaningful ?

[a priori no physics knowledge]



iii) ML predictions vs. physical plausibility

How many of these possible solutions are physically meaningful ?



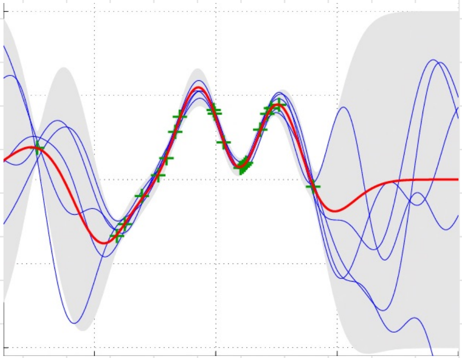
[a priori no physics knowledge]

But... "Physics-informed ML"

↳ Enforce solutions to satisfy physical laws

iii) ML predictions vs. physical plausibility

How many of these possible solutions are physically meaninful ?



[a priori no physics knowledge]

But... "Physics-informed ML"

↳ Enforce solutions to satisfy physical laws

Data

$\{\vec{x}_i, y_i\}$

Statistical (ML) model

$P(y \mid f_{ML}(\vec{x}, \vec{\theta}))$

Loss function   Partial diff. eq.

$$-\ln p(y \mid f_{ML}) + \mathcal{F}(f_{ML}, \partial_{\vec{x}} f_{ML})$$

Take-home messages

* ML is proving to be an essential tool in modern physics (unfortunately embedded in a gigantic hype)

* As physicists, we have the responsability to make use of such tools in a scientific meaningful way

* ML for science still in its infancy
  ⇒ A lot to do in Physics–ML symbiosis