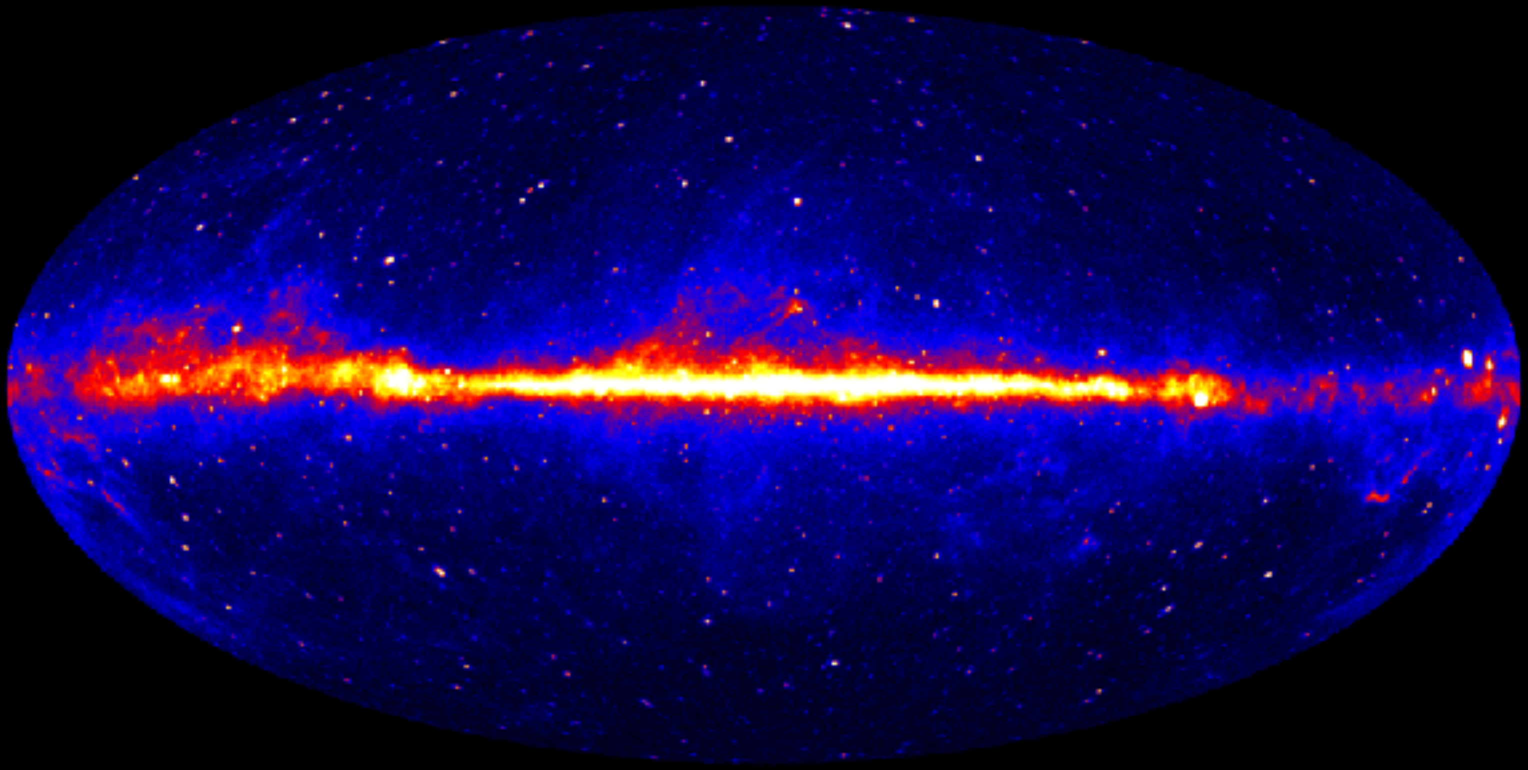
Dark Information: Forecasting with the Fisher Matrix

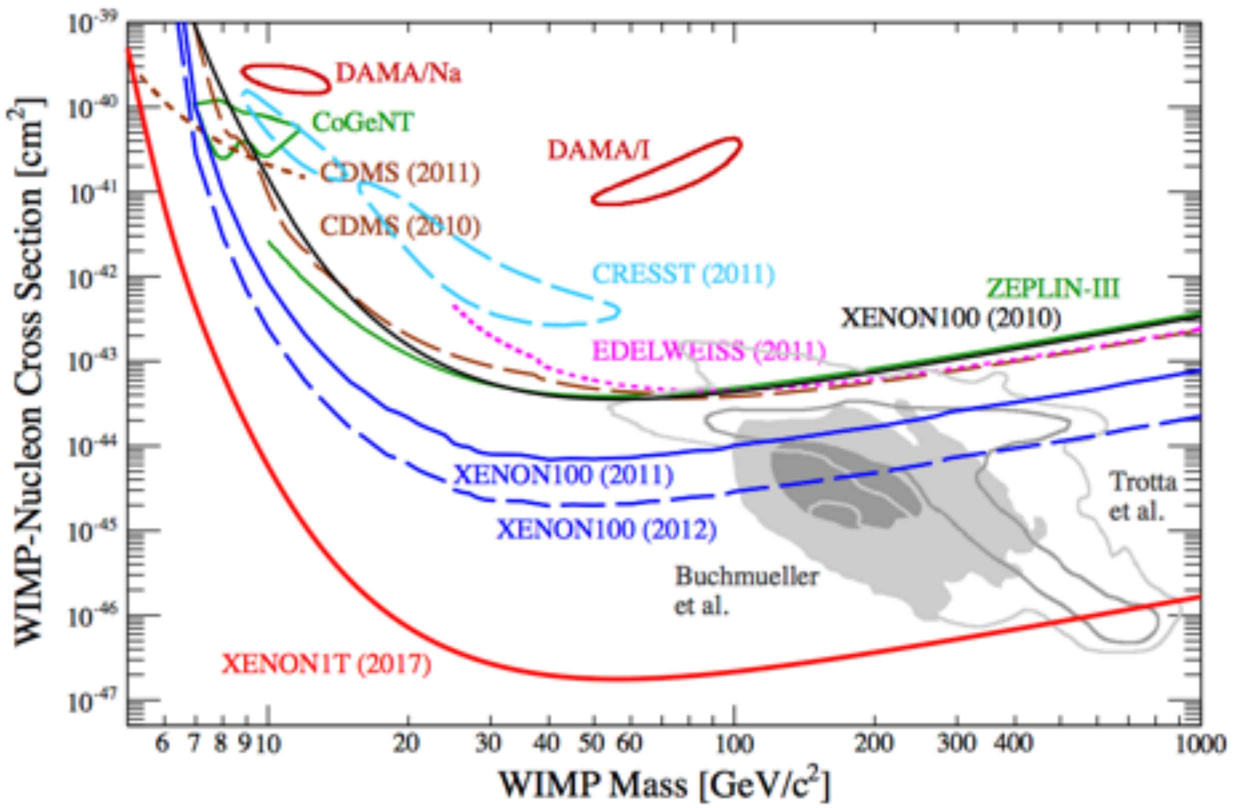


T. D. P. Edwards
and C. Weniger

https://arxiv.org/abs/1704.05458

Forecasting

- Estimating the sensitivity of existing or future experiments for the detection of astrophysical or new physics signals is a ubiquitous task, usually requiring the calculation of the expected exclusion and discovery limits

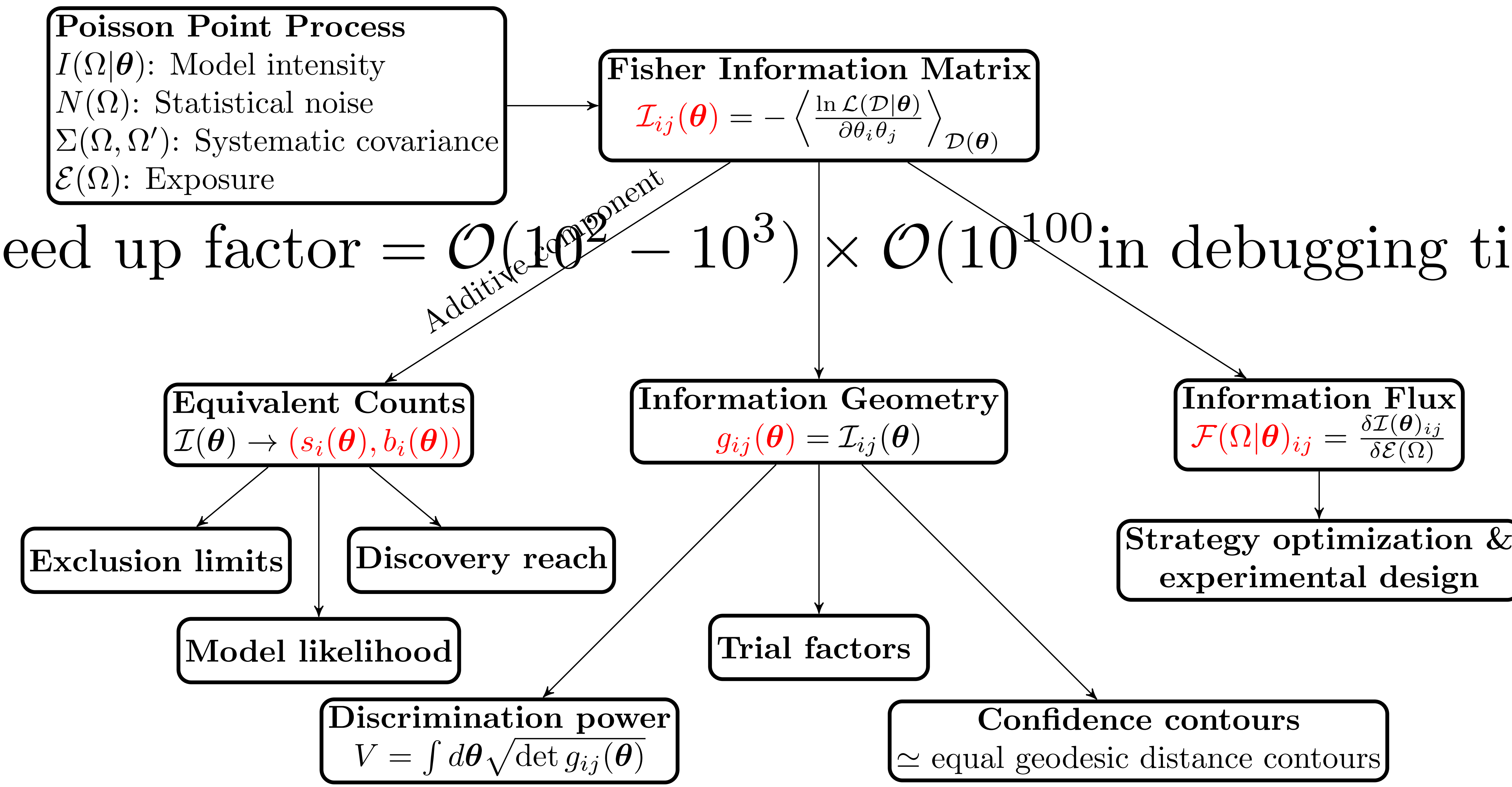


https://arxiv.org/pdf/1206.6288

- What is the maximum information that can be in principle extracted from a given observation?
- Information gain here corresponds to the reduction of the uncertainty associated to the model parameters of interest

Equivalent Counts Method

Speed up factor $= \mathcal{O}(10^2 - 10^3) \times \mathcal{O}(10^{100}$ in debugging time)



Poisson Likelihoods and Model definitions

Poisson Likelihood is used when considering counting experiments such as the gamma-ray satellite Fermi

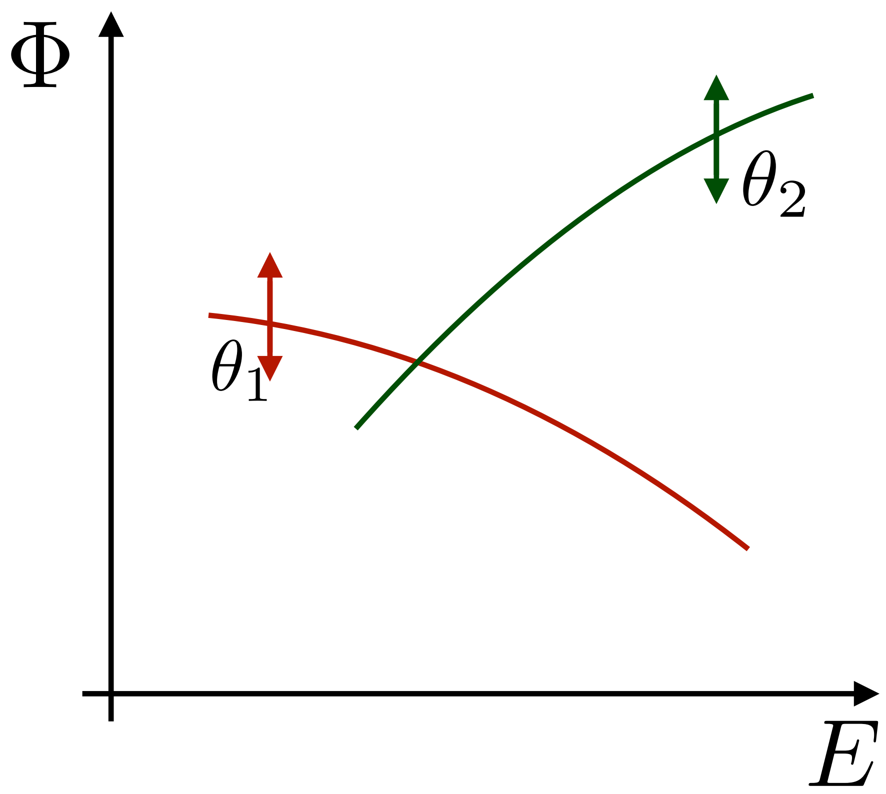
$$\Phi(E|\theta) = \sum_{i=1}^{n} \theta_i \Psi_i(E)$$

$$\ln \mathcal{L}_{\text{pois}}(\mathcal{D}|\theta) = \int dE \;\; [\mathcal{C}(E) \ln \Phi(E|\theta) - \Phi(E|\theta)]$$

Fixed shape parameter

Normalisation, parameter we wish to set a limit on

$$\mathcal{I}_{ij}^{\text{pois}}(\theta) = \int dE \, \frac{\partial \Phi(E|\theta)}{\partial \theta_i} \frac{1}{\Phi(E|\theta)} \frac{\partial \Phi(E|\theta)}{\partial \theta_j}$$



GRAPPA
GRavitation AstroParticle Physics Amsterdam

What is the Fisher Information Matrix

$$\mathcal{I}_{ij}(\theta) \equiv \left\langle \left(\frac{\partial \ln \mathcal{L}(\mathcal{D}|\theta)}{\partial \theta_i} \right) \left(\frac{\partial \ln \mathcal{L}(\mathcal{D}|\theta)}{\partial \theta_j} \right) \right\rangle_{\mathcal{D}(\theta)} = - \left\langle \frac{\partial^2 \ln \mathcal{L}(\mathcal{D}|\theta)}{\partial \theta_i \partial \theta_j} \right\rangle_{\mathcal{D}(\theta)}$$

- **The Fisher Information is a description of the curvature of the likelihood**
- **Curvature of the likelihood surface gives us a description of the variance**
- **The Cramér-Rao bound is based on the Fisher information matrix, which quantifies how 'sharply peaked' the likelihood function describing the observational data is around its maximum value**

$$\mathrm{cov}\left[\hat{\theta}_i, \hat{\theta}_j\right] \equiv \langle (\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j) \rangle_{\mathcal{D}(\theta)} \geq (\mathcal{I}(\hat{\theta})^{-1})_{ij}$$

- **Bound is 'asymptotically efficient' when the bound is saturated in the large sample limit**

Equivalent Counts Method

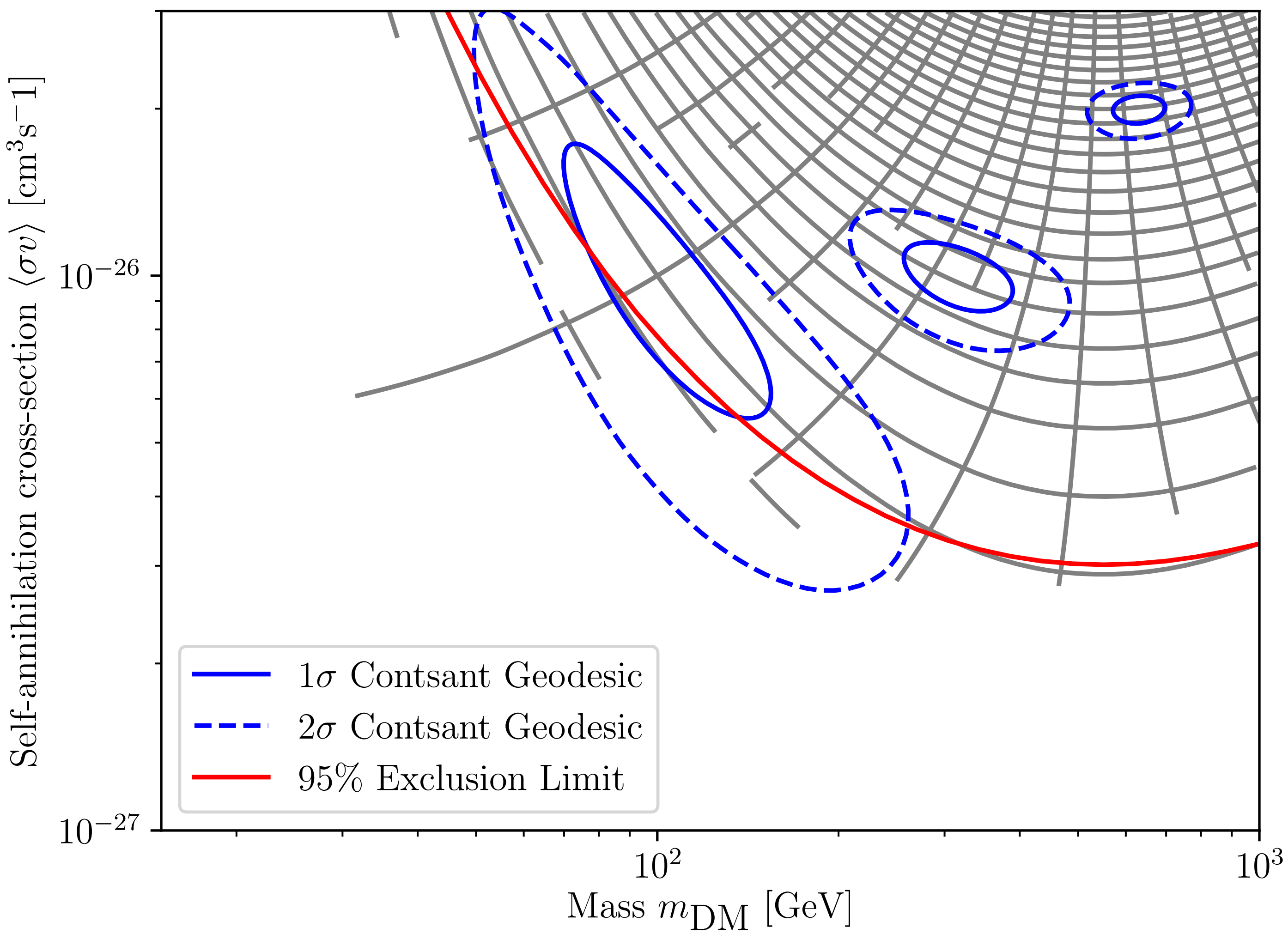
Logic:

- **Signal to Noise of events in a single bin example tells us about the significance of the signal**
- **Extend same technique to multi-bin case**
- **Not all signal events statistically contribute if they are drowned out by large backgrounds**
- **Convenient to define significant signal and background events using the FIM**

$$s_i(\theta) = \frac{\theta_i^2}{\sigma_i^2(\theta) - \sigma_i^2(\theta_0)}$$

$$b_i(\theta) = \frac{\theta_i^2 \sigma_i^2(\theta_0)}{(\sigma_i^2(\theta) - \sigma_i^2(\theta_0))^2}$$

Indirect DM Search - CTA Projections



- Previous analysis was **extremely slow**. Systematic limits easily accounted for in the FIM
- Streamlines roughly depicts the model density of the parameter space
- Starting to tap into the field of **Information Geometry** for useful parameter space visualisations and observation strategies

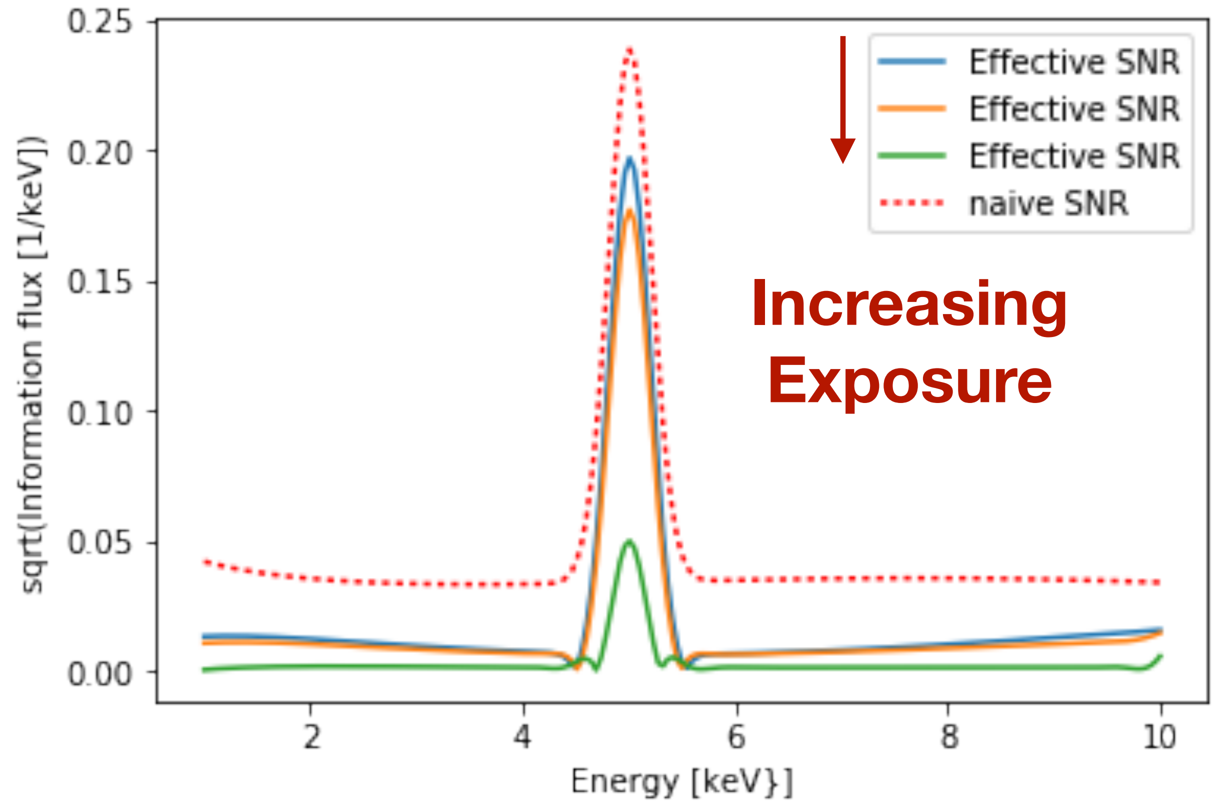
- Treat the FIM as a **metric** and use this to plot constant geodesic ellipses. The **size and shape of the ellipses show how far** in the parameter space we need to travel to rule out parameter space

Strategy Optimisation

- It is possible to include an additional term to the likelihood that describes background correlated systematics. In addition we can look how the information is distributed over you binned variable, we call this object the Effective *Fisher information flux*

$$\frac{d\widetilde{\mathcal{F}}_{11}}{dE}(E_k) = \sum_{ij} \frac{I_1}{I_2}(E_i) D^{-1}_{ik} \frac{1}{\Delta E_k^2 \mathcal{E}(E_k)^2 I_2(E_k)} D^{-1}_{kj} \frac{I_1}{I_2}(E_j)$$

- The diagonal part of the Fisher information flux corresponds to the square of the SNR of component i, and the non-diagonal parts provide information about the degeneracy of the components pairs (i, j)



Python Package now public

Index

Classes

- EffectiveCounts
  - __init__
  - discoveryreach
  - effectivecounts
  - totalcounts
  - upperlimit
- Funkfish
  - __init__
  - EffectiveCounts
  - Swordfish
  - TensorField
  - iminuit
- Swordfish
  - __init__
  - effectivefishermatrix
  - effectiveinfoflux
  - fishermatrix
  - infoflux
  - lnL
  - mu
  - profile_lnL
  - variance

Sub-modules

- swordfish.metricplot

swordfish module

`swordfish` is a Python tool to study the information yield of counting experiments.

NOTE: The package is stable, but still in beta phase. Use for production only if you know what you're doing.

Motivation

With **swordfish** you can quickly and accurately forecast experimental sensitivities without all the fuss with time-intensive Monte Carlos, mock data generation and likelihood maximization.

With `swordfish` you can

- Calculate the expected upper limit or discovery reach of an instrument.
- Derive expected confidence contours for parameter reconstruction.
- Visualize confidence contours as well as the underlying information metric field.
- Calculate the *information flux*, an effective signal-to-noise ratio that accounts for background systematics and component degeneracies.

A large range of experiments in particle physics and astronomy are statistically described by a Poisson point process. The **swordfish** module implements at its core a rather general version of a Poisson point process, and provides easy access to its information geometrical properties. Based on this information, a number of common and less common tasks can be performed.

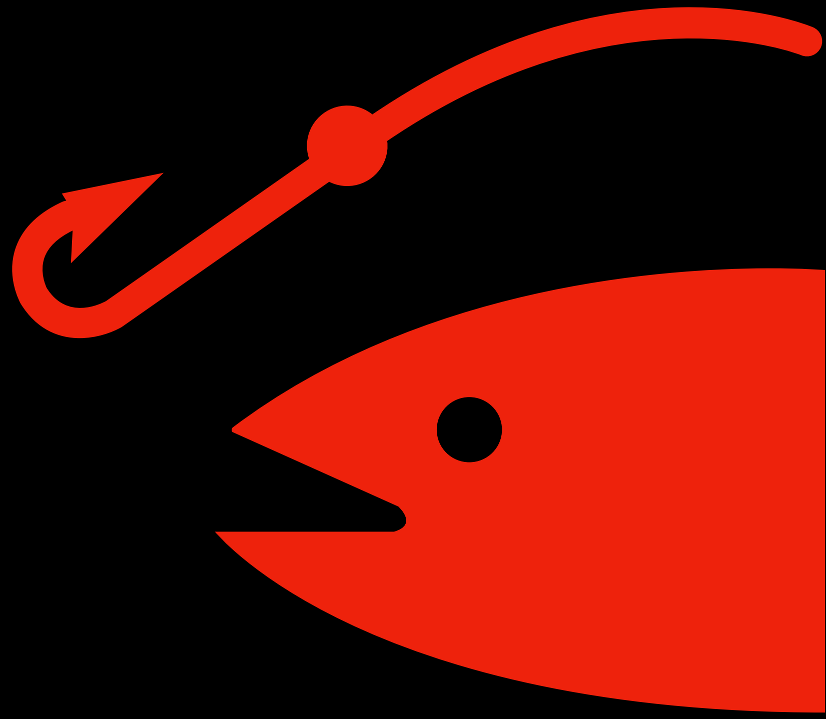
Get started

Most of the functionality of `swordfish` is demonstrated in two jupyter notebooks.

- Linear models and effective counts method
- Non-linear models and confidence contours

WITH DOCUMENTATION

https://github.com/cweniger/swordfish



Thanks for listening!

https://arxiv.org/abs/1704.05458

What do people normally do…

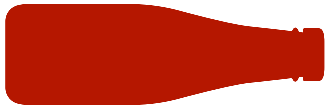
- There are many methods both for calculating your expected sensitivity and expected signal. Most prominent is the use of Monte Carlo Simulations

Start with basic Physics

Create Model

Sample from Model using Monte Carlo methods

Calculate limits off mock data and repeat simulations to check validity



Bottle neck of the analysis. Often takes a long time with little flexibility

Limit and Reach

Exclusion Limits

$$s_1(\theta^U) = Z(\alpha) \cdot \sqrt{s_1(\theta^U) + b_1(\theta^U)}$$

where $\theta^U = (\theta_1^U, \theta_2, \ldots, \theta_n)^T$

Discovery Reach

$$\left(s_1(\theta^D) + b_1(\theta^D)\right) \ln\left(\frac{s_1(\theta^D) + b_1(\theta^D)}{b_1(\theta^D)}\right) - s_1(\theta^D) = \frac{Z(\alpha)^2}{2}$$

where $\theta^D = (\theta_1^D, \theta_2, \ldots, \theta_n)^T$

- Solve for θ_1^{U} to calculate the exclusion limit
- α is the statistical significance
- Z(α) is connected to the desired confidence limit inverse of the standard normal cumulative distribution e.g. for CL=95%, Z(0.05)=1.64
- Solve for θ_1^{D} to calculate the discovery reach

Useful properties of the Formalism

Properties:

- **Multiplicativity of Poisson Likelihoods**
  —> Additive Fisher Information

$$\mathcal{I}_{\mathrm{Tot}} = \mathcal{I}_1 + \mathcal{I}_2 + \ldots$$

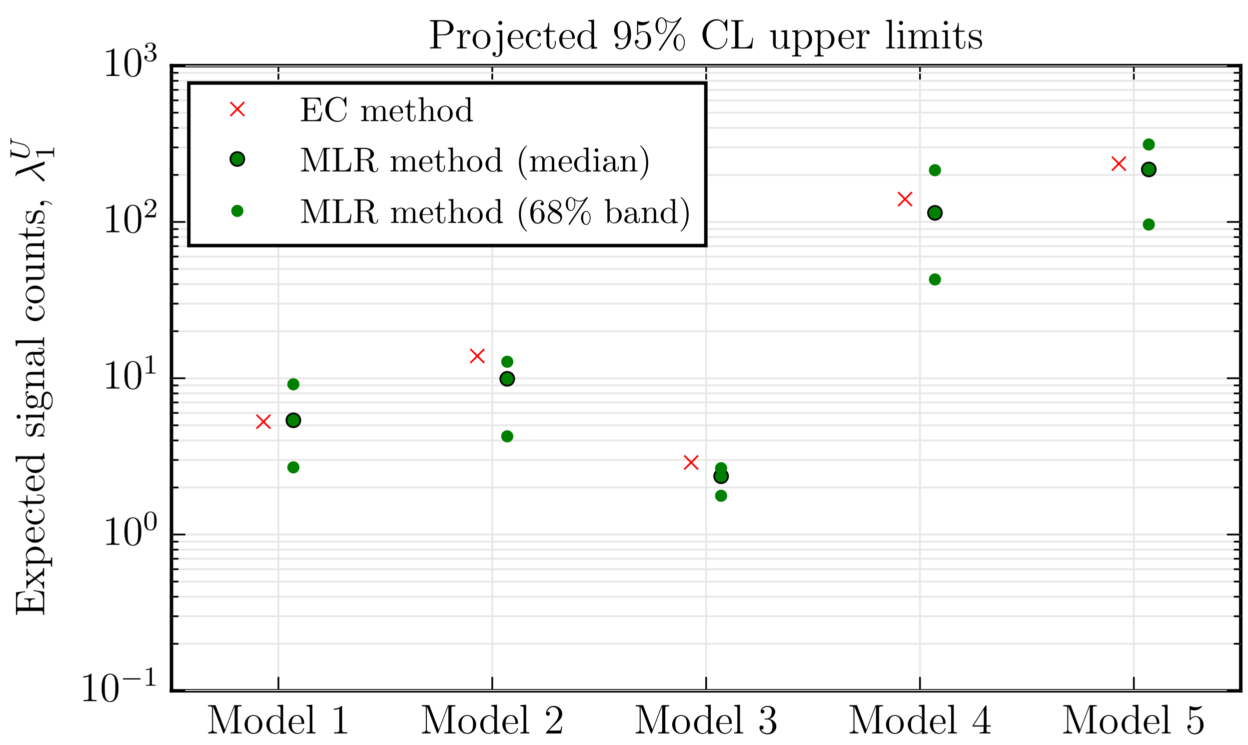
- **Easily *reduced***

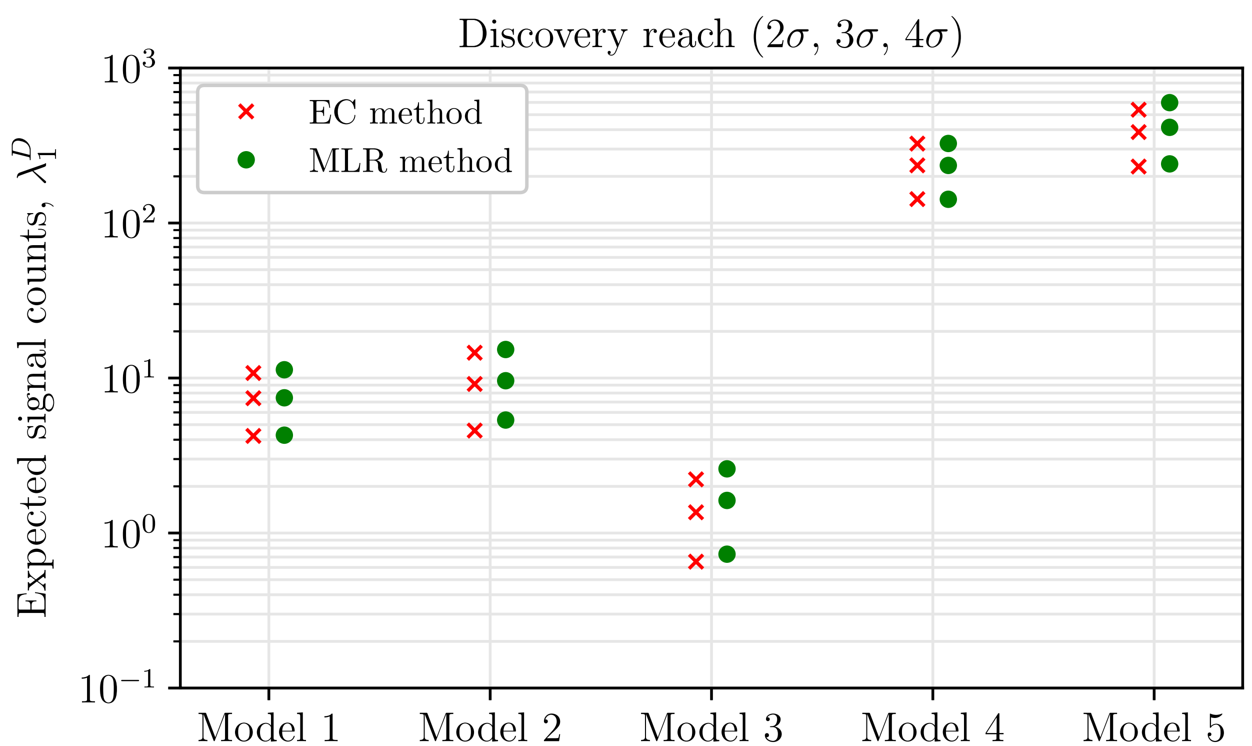
$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_A & \mathcal{I}_C^T \\ \mathcal{I}_C & \mathcal{I}_B \end{pmatrix} \qquad \widetilde{\mathcal{I}}_A = \mathcal{I}_A - \mathcal{I}_C^T \, \mathcal{I}_B^{-1} \, \mathcal{I}_C$$

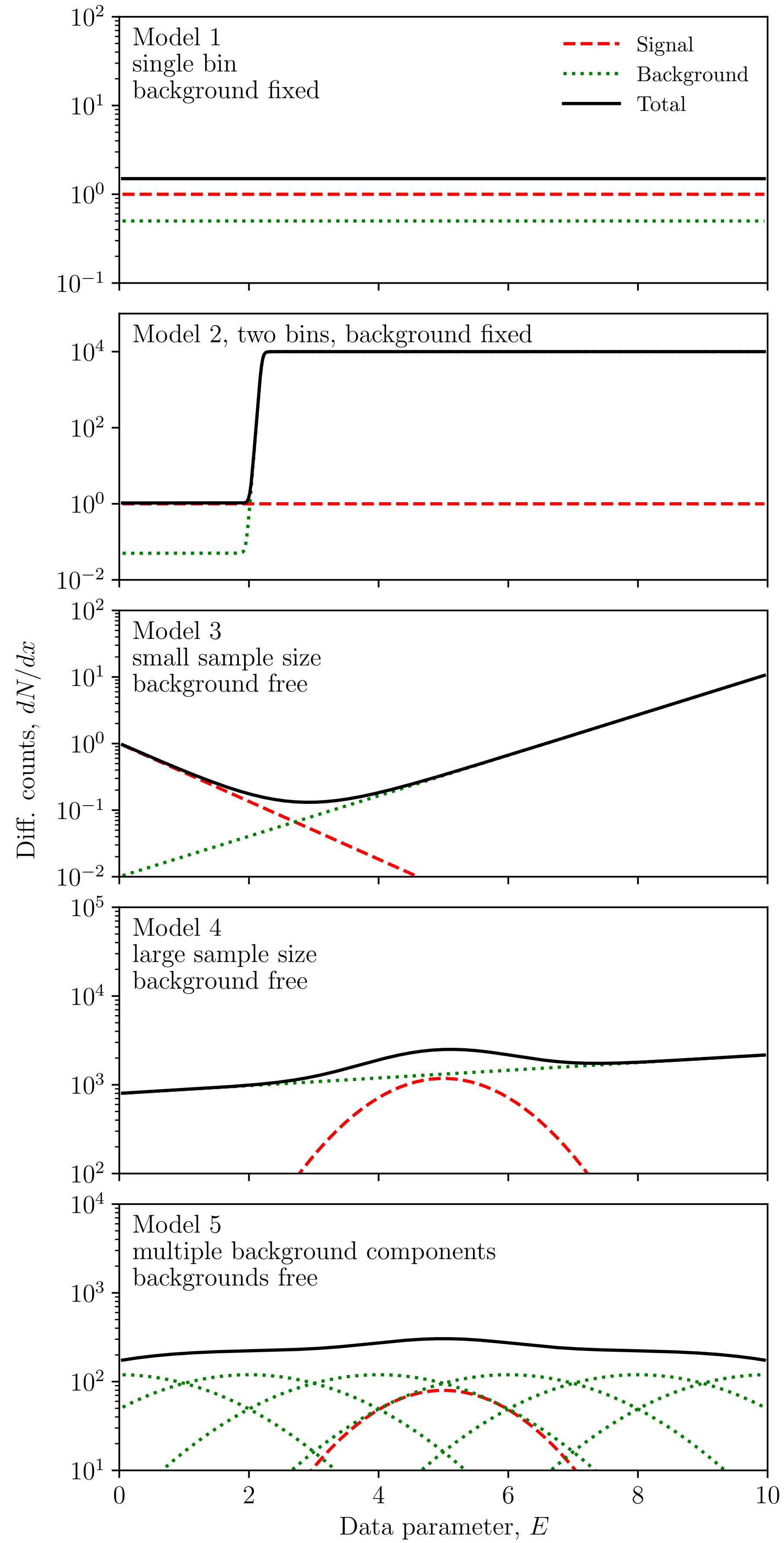
- **Simple to map to standard Dark matter searches - Direct, Indirect, Collider etc.**

$$\underbrace{\frac{dN}{d\Omega}}_{\Phi(\Omega)} = \underbrace{\langle \sigma v \rangle}_{\theta_1^U} \underbrace{\frac{1}{2\pi} \frac{1}{m_{\mathrm{DM}}^2} \int dE \frac{dN}{dE} \int ds \rho^2(r(s,l,b))}_{\Psi_s}$$

Comparison with Monte Carlo

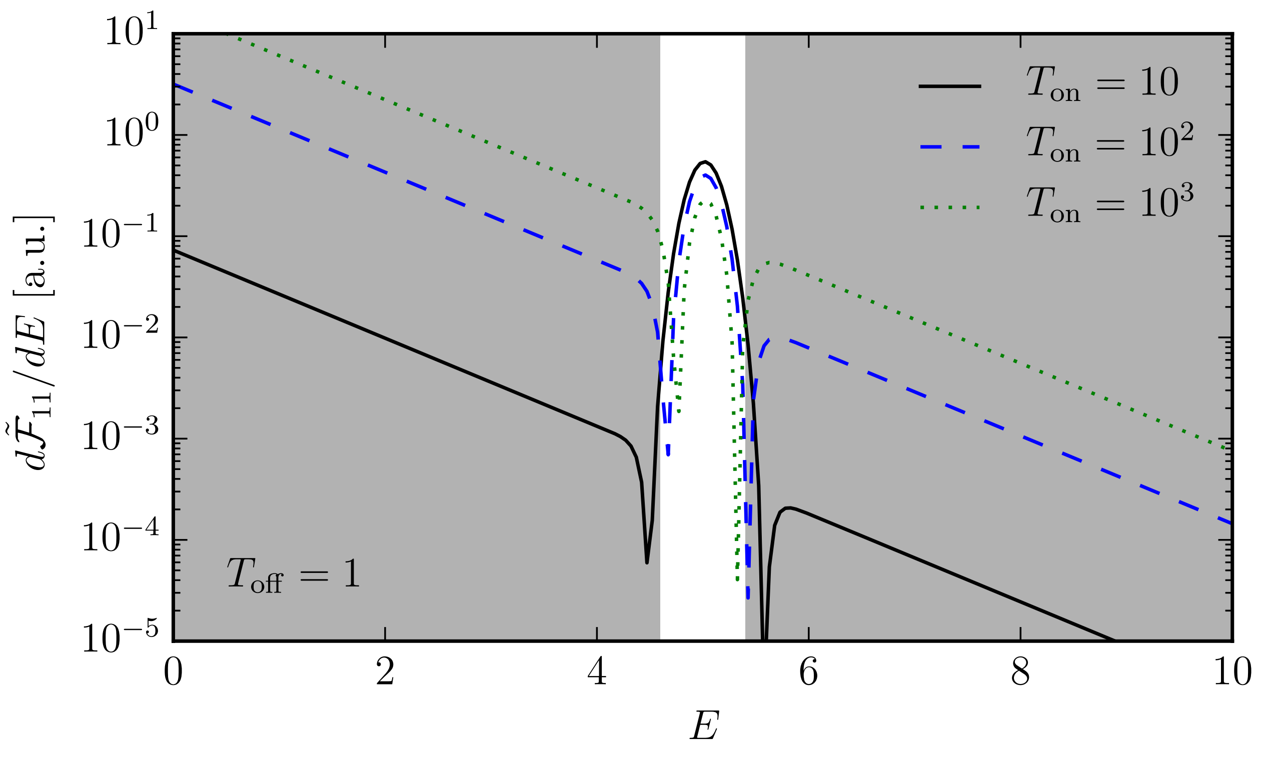


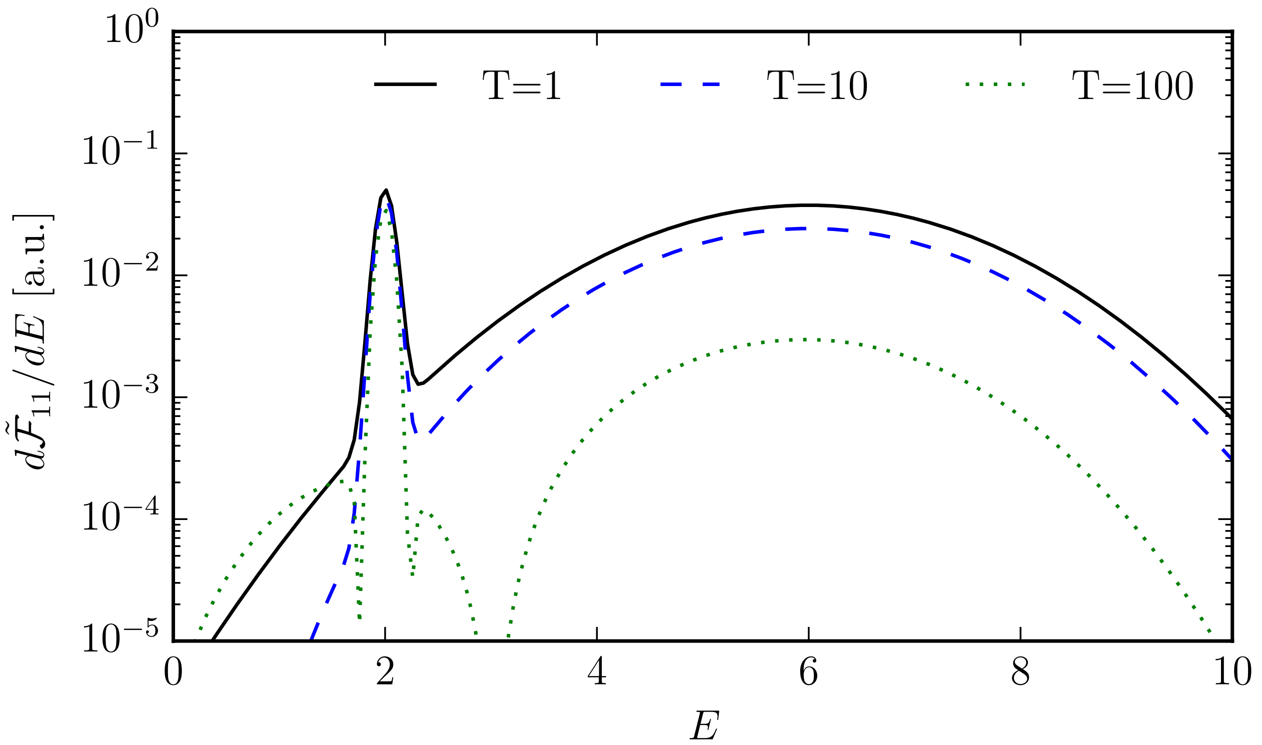




Non-locality and Saturation

- **The resulting Fisher Information Flux is non-local in space due to the appearance of the full FIM occurring in the *Effective Information Flux***
- **With non-locality, we mean that the information flux at E depends in general on the past observation history of E'/=E**
- **The differential information flux accounts for the fact that the broad feature becomes increasingly degenerate with the flat background at late times but the sharp peak remains useful**





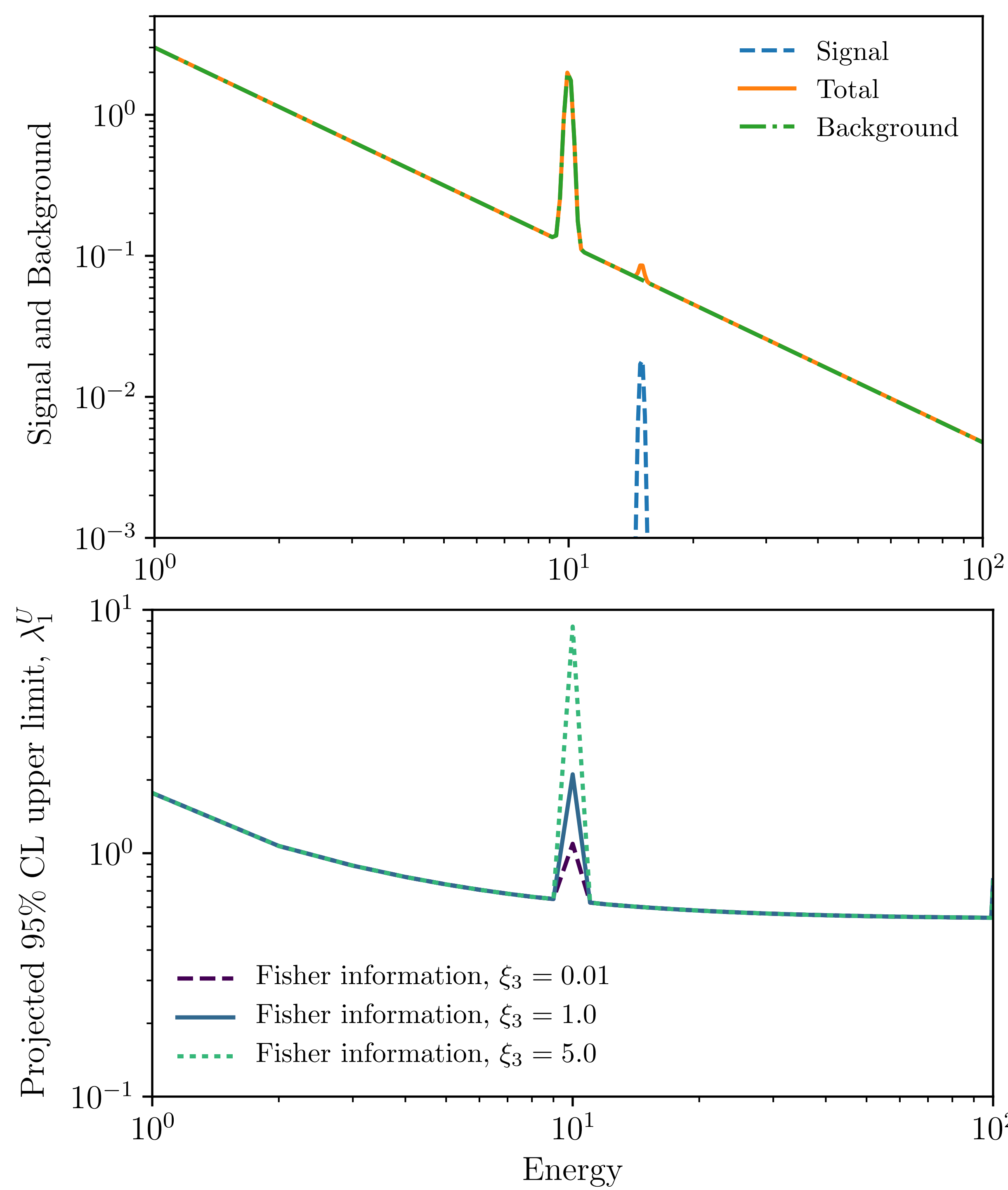
Systematics

- In many cases of practical importance, additional information about nuisance parameters is available, which must be included in the sensitivity projections to obtain realistic results

$$\mathcal{L}(\mathcal{D}|\theta) = \mathcal{L}(\mathcal{D}|\theta)_{\text{pois}} \times \prod_i \mathcal{N}(\theta_i^A|\mu = \theta_i, \sigma^2 = \xi_i^2)$$

$$\mathcal{I}_{ij} = \mathcal{I}_{ij}^{\text{pois}} + \mathcal{I}_{ij}^{\text{syst}}$$

- In general, systematic uncertainties in the background will be correlated as a function of energy
- Can encode the complicated covariances by inserting a Gaussian random field with zero mean



$$\widetilde{\mathcal{I}}_{11} = \sum_{ij} \frac{\Psi_1}{\Psi_2}(E_i) D_{ij}^{-1} \frac{\Psi_1}{\Psi_2}(E_j) \;, \quad D_{ij} \equiv \frac{\delta_{ij}}{\Delta E_i \Psi_2(E_i)} + \Sigma_\delta(E_i, E_j)$$

Covariance function