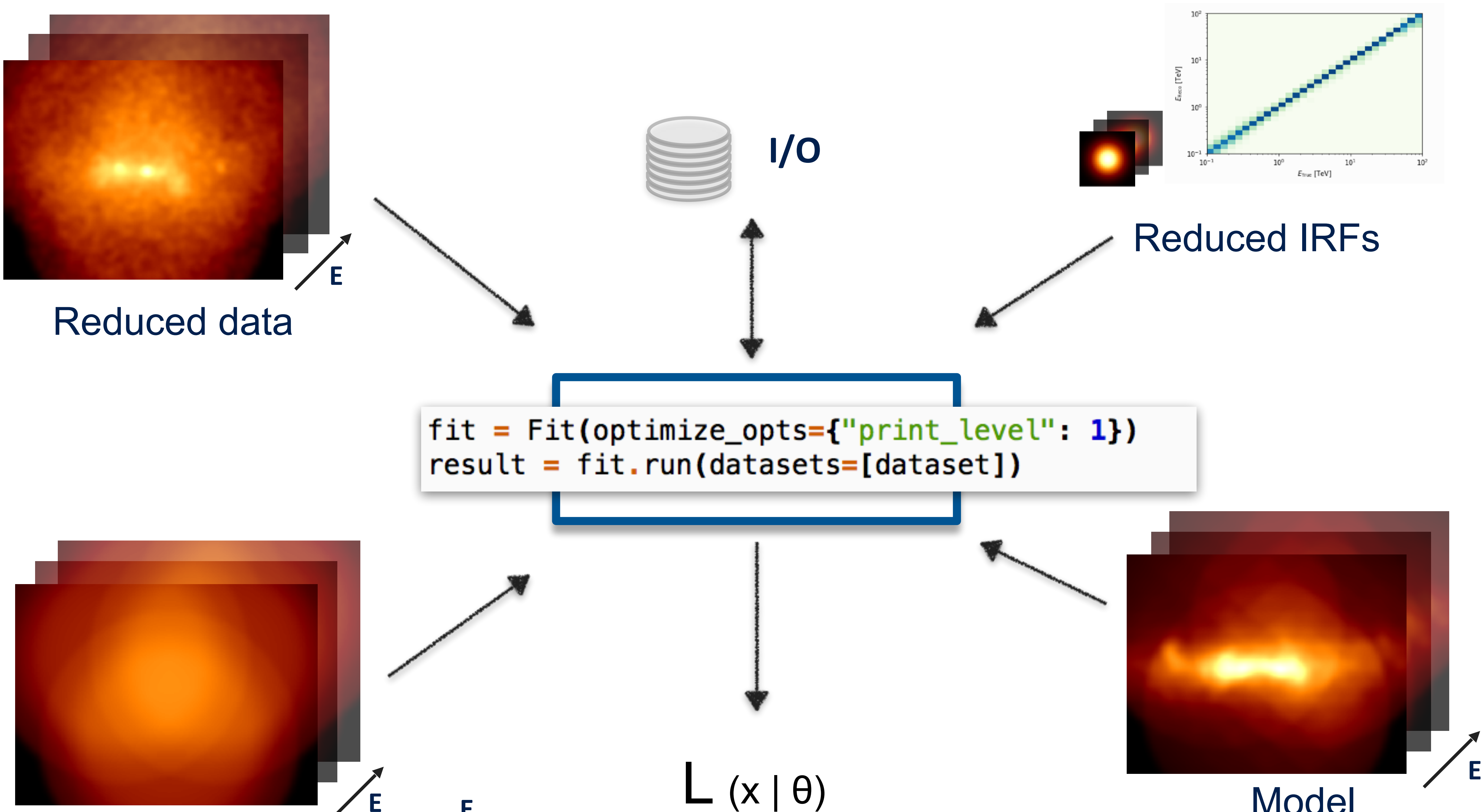
Loss Landscape image

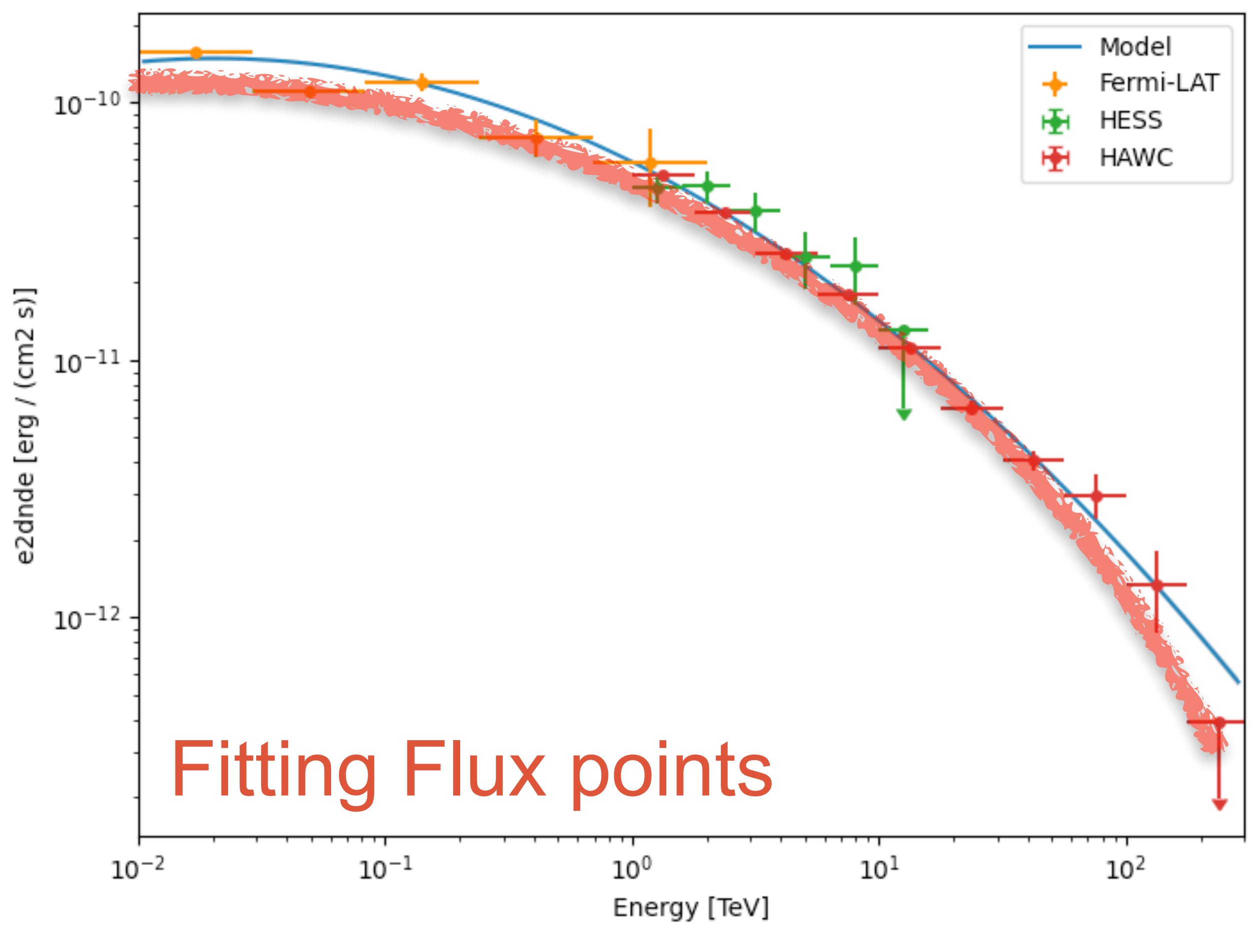
Datasets modeling and fitting

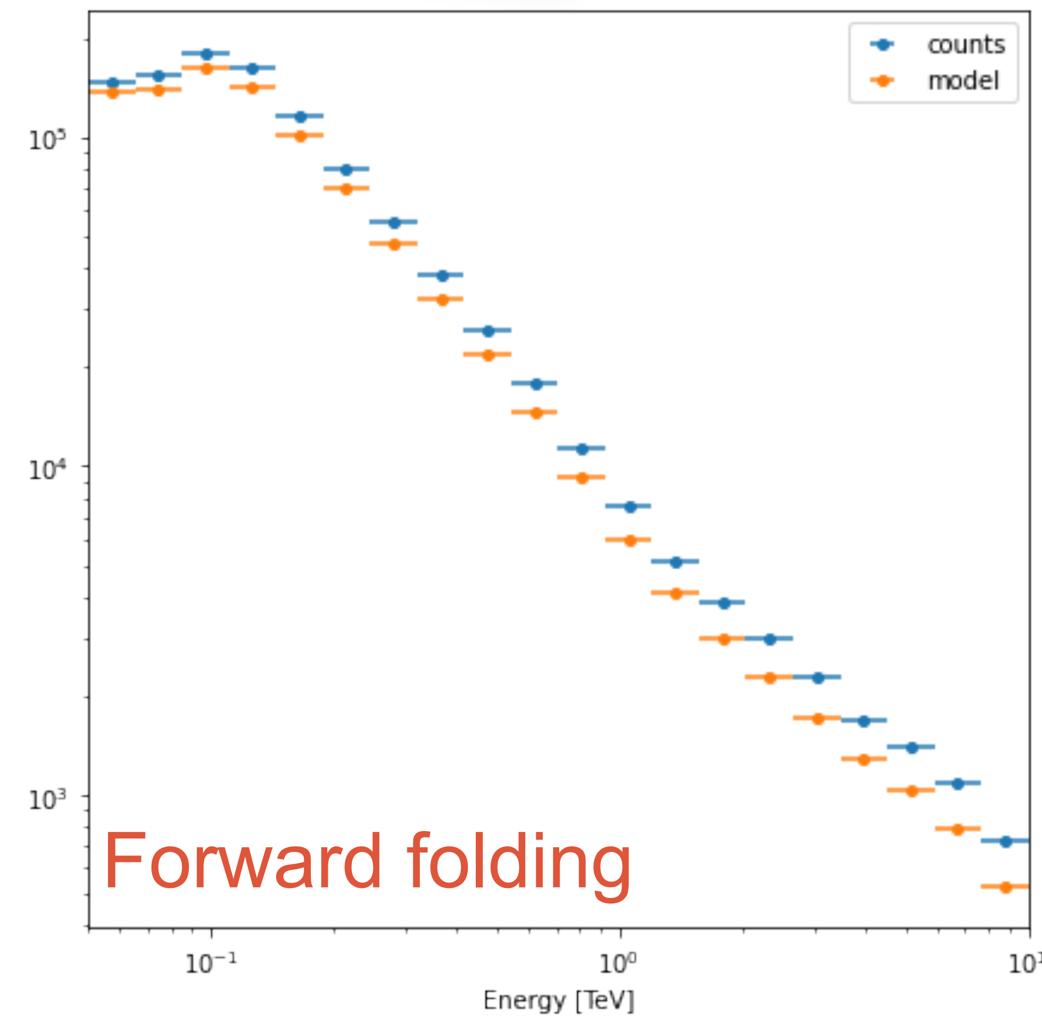
A binned analysis



see: [Dataset fitting tutorial]

What is forward folding ?





- Data are transformed to physical information
- Flux point modeling : a chi2 fit on flux points
  - Loss of statistical information
  - No handling of correlation between points

- Data are not transformed: N_{obs}
- The physical model is
  - Flux -> N_{pred} counts
- Proper statistical treatment
  - In particular for low counts

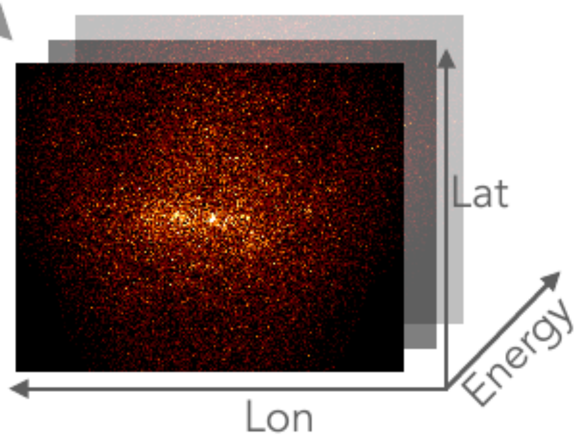
Poisson Log-Likelihood

List of gamma-like events…

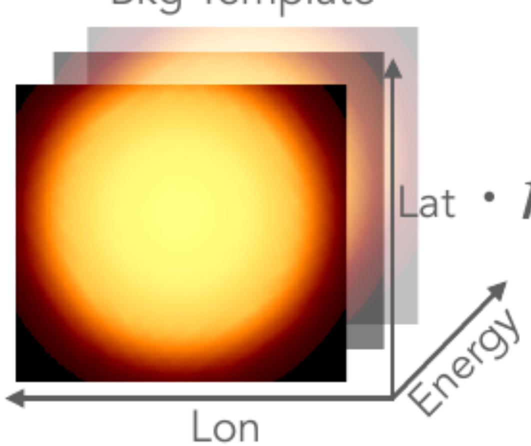
EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
int64	float64	float32	float32	float32
5407363825684	123890826.66805482	84.97964	23.89347	10.352011
5407363825695	123890826.69749284	84.54751	21.004095	4.0246882
5407363825831	123890827.23673964	85.39696	19.41868	2.2048872

…binned into…

Counts

$N_{Obs} =$ 

Bkg Template

$N_{Bkg} =$ 

$\cdot \, n_0 \cdot \left(\frac{E}{E_0}\right)^{-\Gamma}$

"Cash statistics": summed over all "bins"

$$\mathscr{C} = 2 \sum_i N^i_{Pred} - N^i_{Obs} \cdot \log N^i_{Pred}$$

i: spectral channels or 3D voxels

$$N_{Pred} = N_{Bkg} + \sum_{Src} N_{Pred,Src}$$

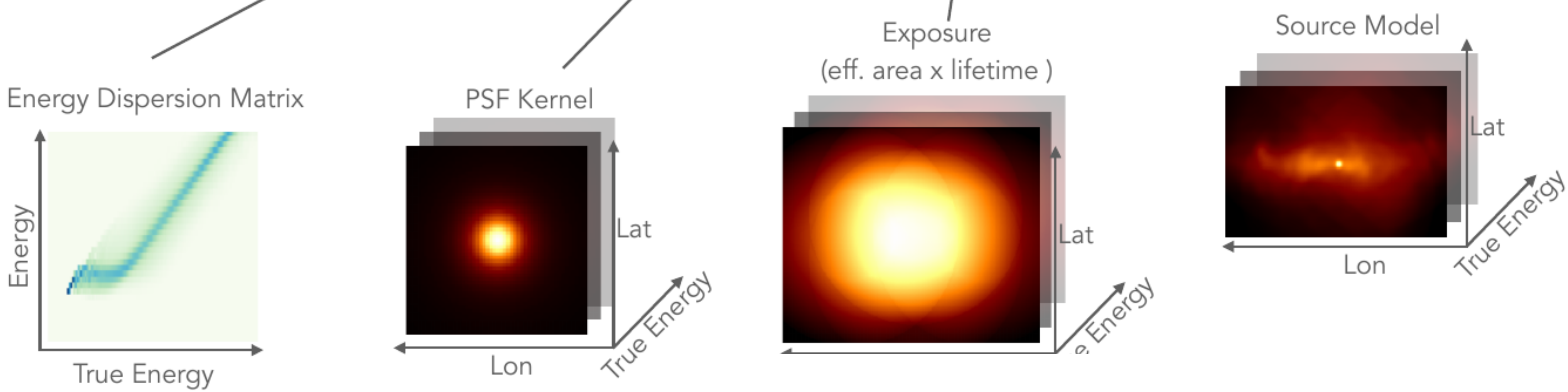
- Predicted counts are **computed per model component** ("source / object") and summed
- A **"global" background model** template with "correction parameters" is added

How is N_{pred} obtained ?

An analytical source model or template is **"forward folded" through the instrument response function (IRF)** to predict the measured number of counts…

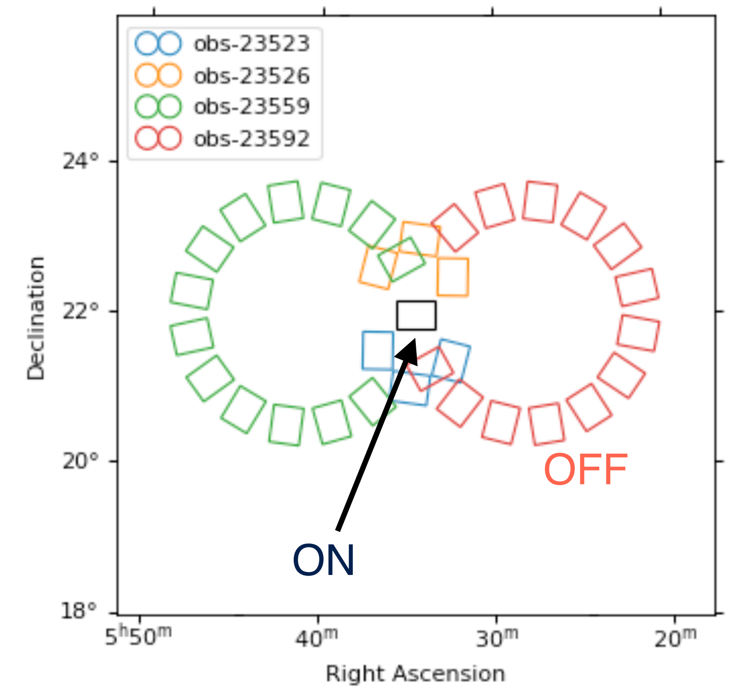
$$N_{Pred,Src} = \mathrm{EDISP}_{Src}(\mathrm{PSF}_{Src}(\mathscr{E}_{Src} \cdot f_{Src}))$$

$$f_{Src} = f_{Spectral}(E) \cdot f_{Spatial}(E,l,b) \cdot f_{Temporal}(t)$$

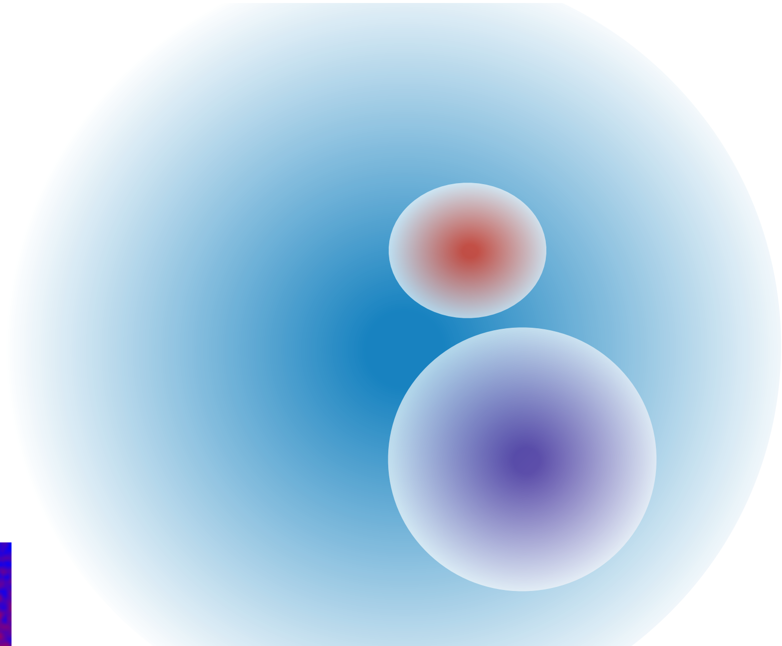


1D or 3D analysis ?

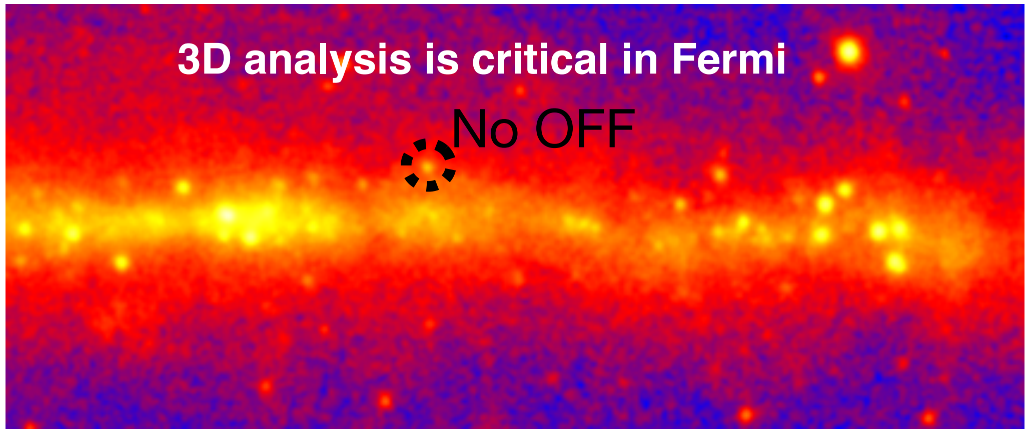
- Classical spectral analysis : ON/OFF background subtraction
  - Pros: background from the data (same observing conditions), no spatial assumption needed, less IRFs systematics
  - Cons: cannot disentangle overlapping components, limited number of OFF in complex regions



- 3D analysis (x, y, E, like Fermi-LAT) :
  - Pros: Suited for complex regions (can disentangle overlapping sources), morphological and spectral analysis in one step, sensitivity gain
  - Cons: Assumption on the background, spatial model for complex source



Cannot define clear
ON & OFF regions
Cross- contamination



Note on the loss function

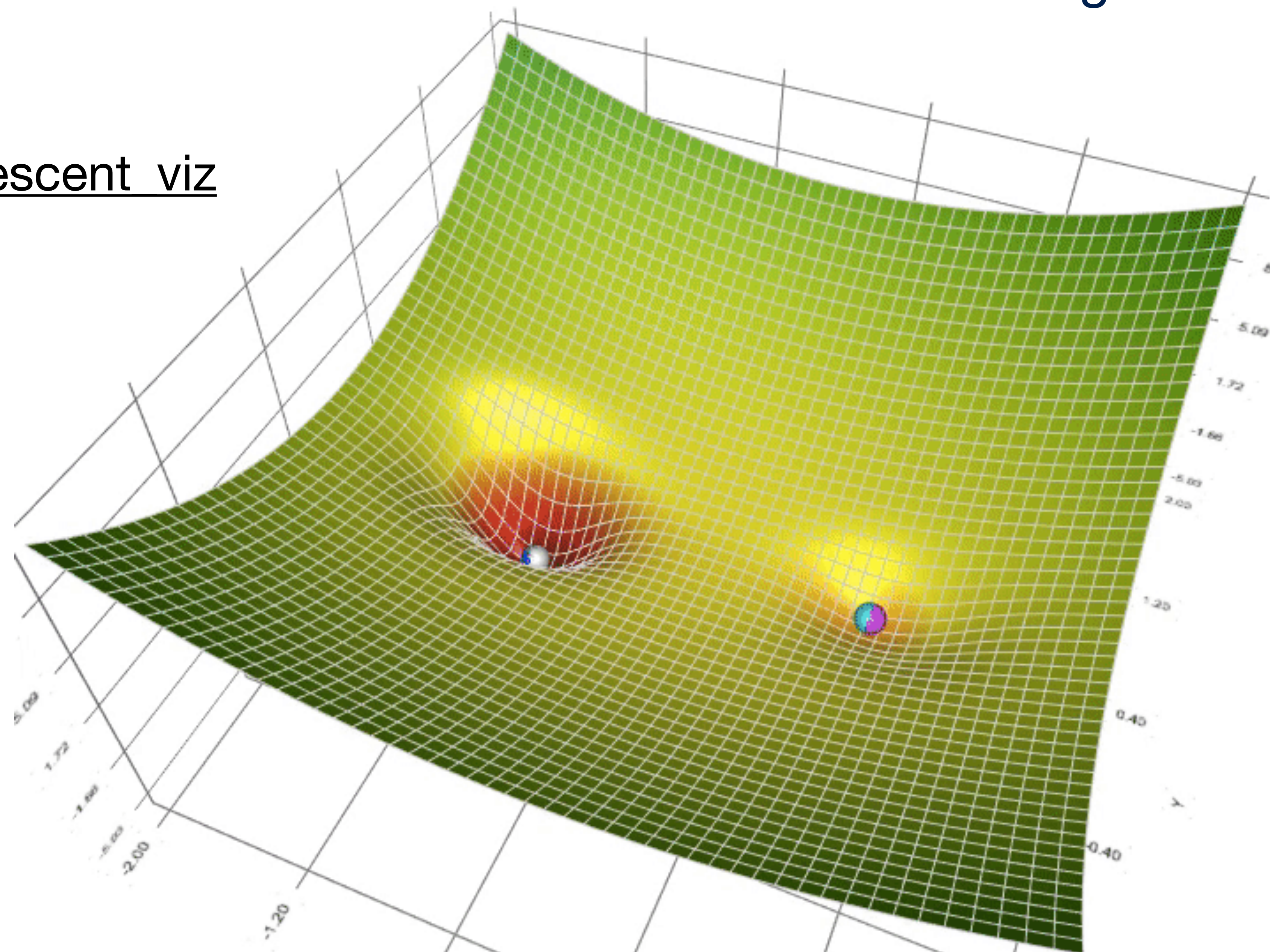
- Model parameter estimation is performed through maximum likelihood technique:
  - Cash statistics is used for counts data with a known background
    - The 3D analysis with a model background in the IRF
  - Wstat statistics is used for counts data with a measured background
    - Typically the 1D analysis where the bkg is estimated from the OFF regions
    - Or a 3D analysis with ON/OFF estimation

What is a fit ?

- Need a loss function + minimizer :
  - Gradient Descent (e.g. Scipy minimize, iMinuit, sherpa fit, etc)
  - Markov chain Monte Carlo
  - Nested Sampling methods

github.com/lilipads/gradient_descent_viz

Never take a best-fit for granted

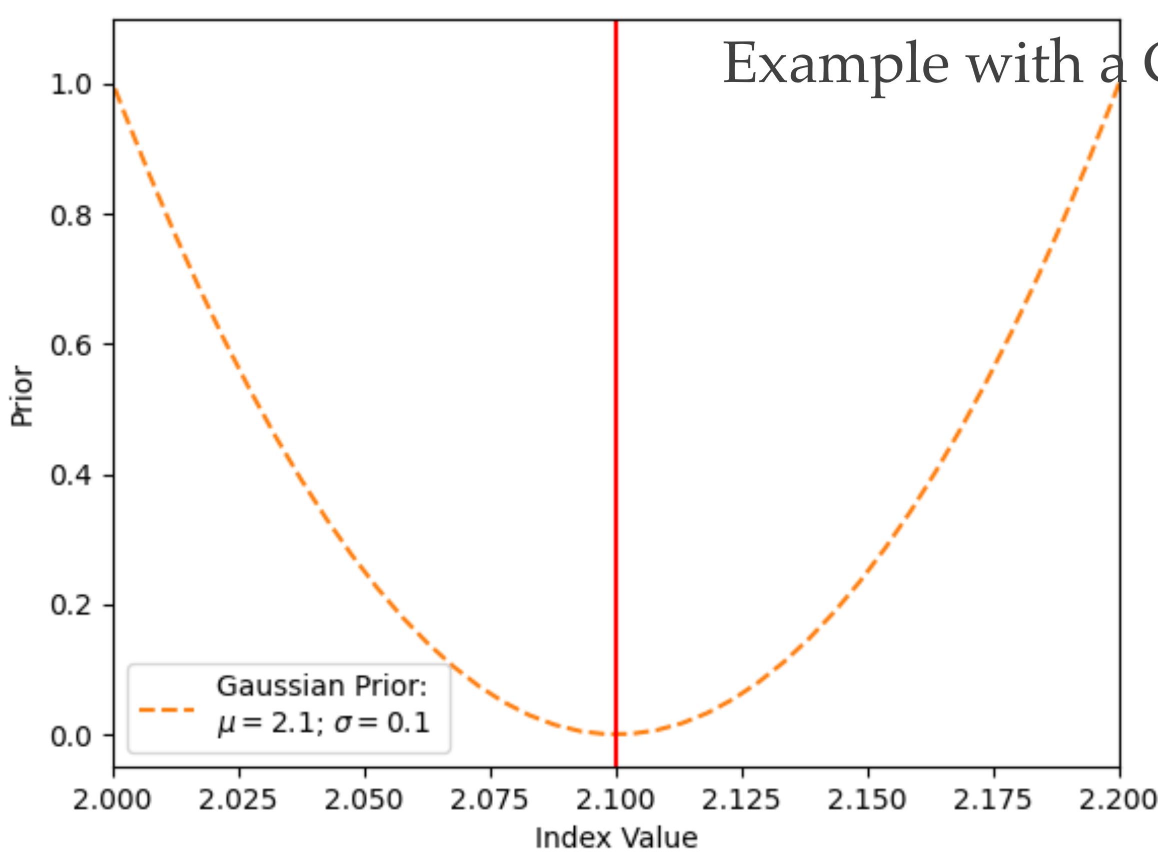


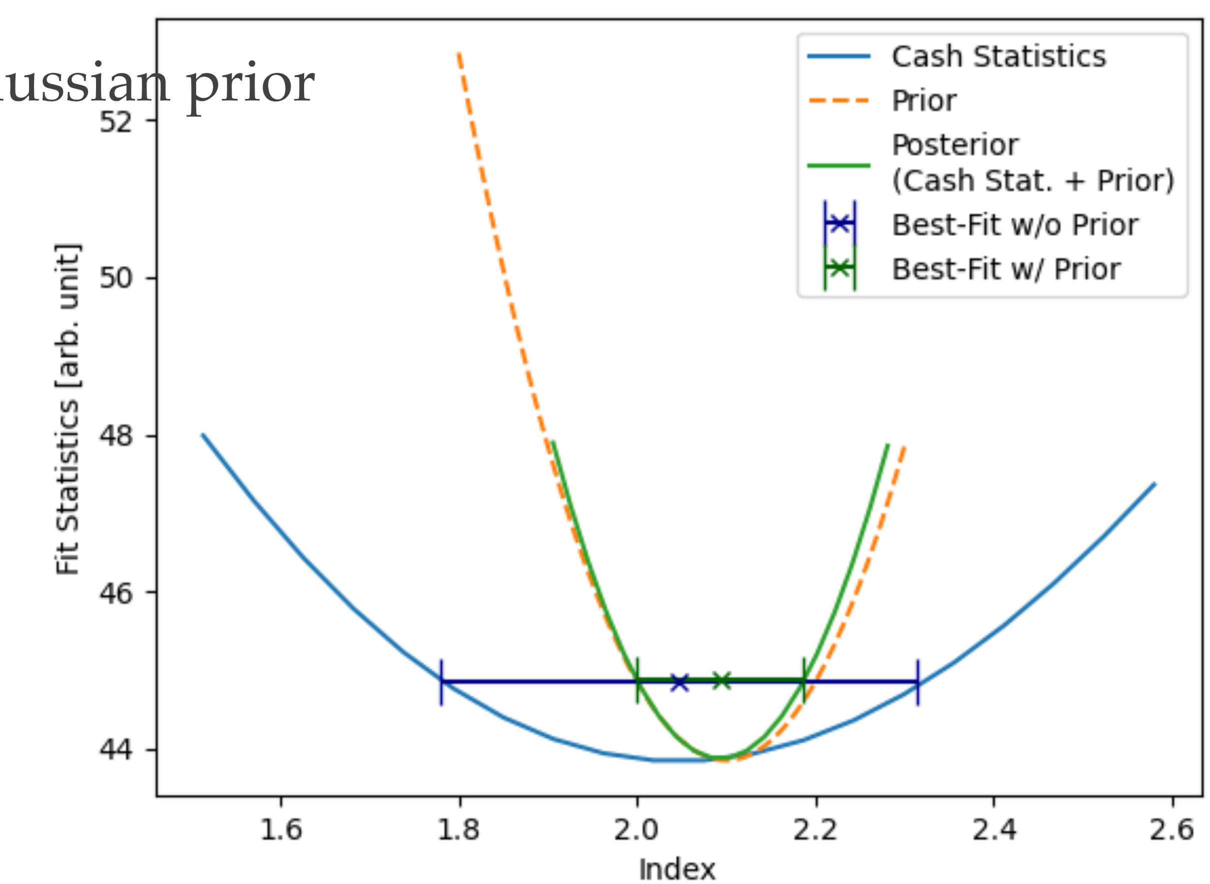
Adding priors on Parameters

- Prior: A probability density function of the model parameters

- Includes information about the parameters
- Added to the fit statistic to get the Posterior
- Possible to add Custom priors

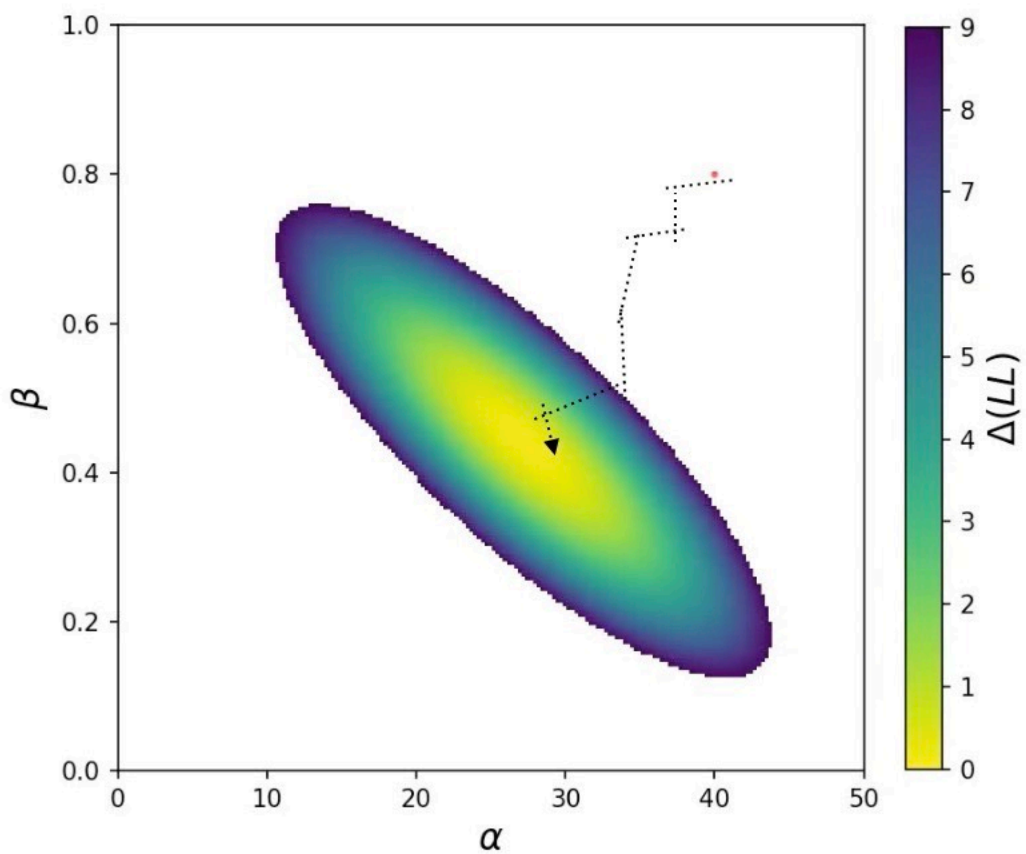
Example with a Gaussian prior





Traditional fitting method

- **Gradient descent based method**
  - **Levenberg Marquardt**
  - **Migrad in Minuit for example**
  - **Gradient is estimated numerically at each step**



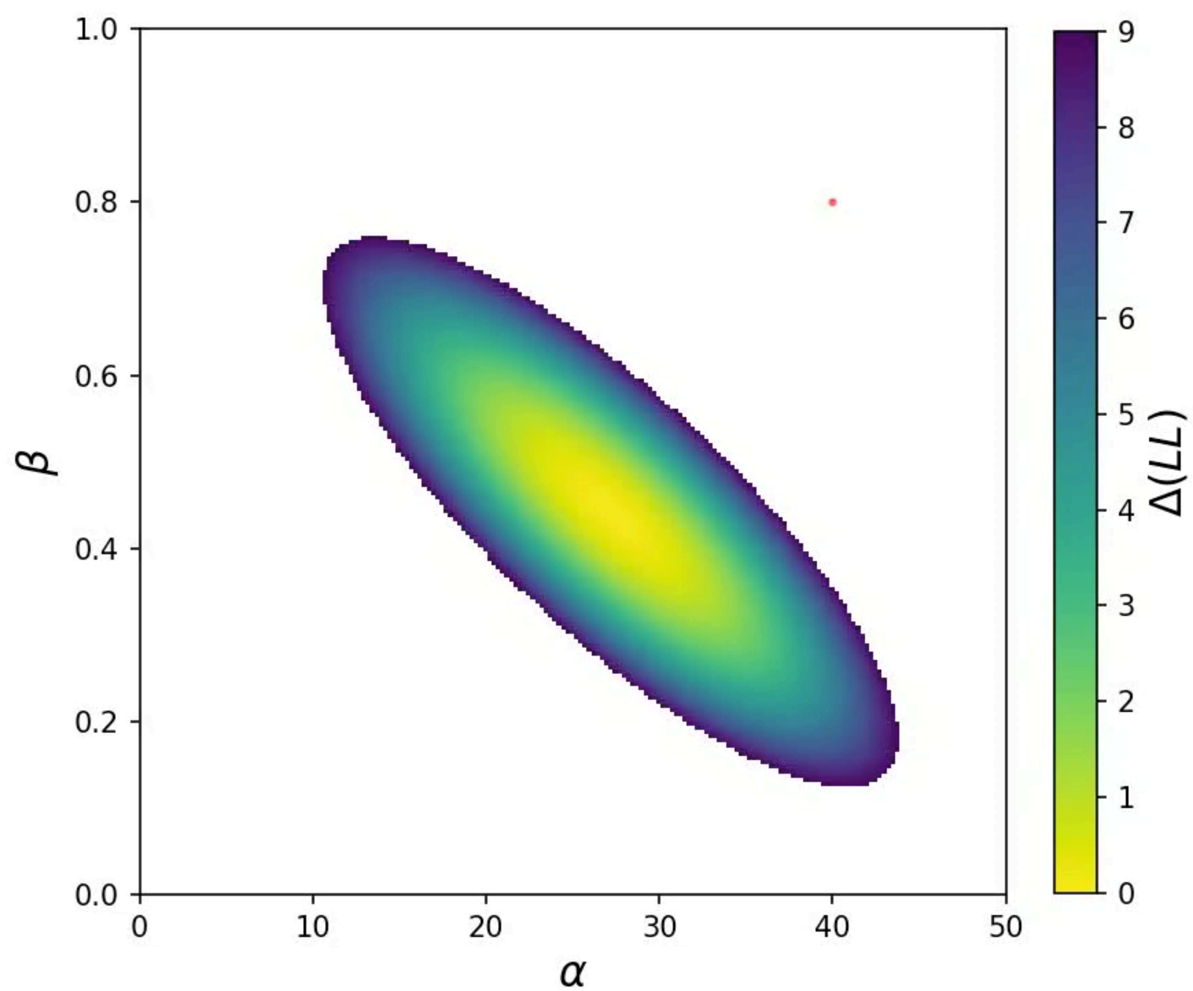
Once at best-fit stops
No information about local likelihood
Sometimes fit fails :

Migrad

FCN = -2.676e+07		Nfcn = 378		
EDM = 0.00356 (Goal: 2e-06)		time = 2.3 sec		
INVALID Minimum		No Parameters at limit		
ABOVE EDM threshold (goal x 10)		Below call limit		
Covariance	Hesse ok	APPROXIMATE	NOT pos. def.	FORCED

Markov Chain Monte Carlo (MCMC)

- **What are they:**
  - **Monte Carlo: samples are used to approximate the probability distribution**
  - **Markov Chain: semi-random walk in potential**
  - **Walkers explore the local likelihood**

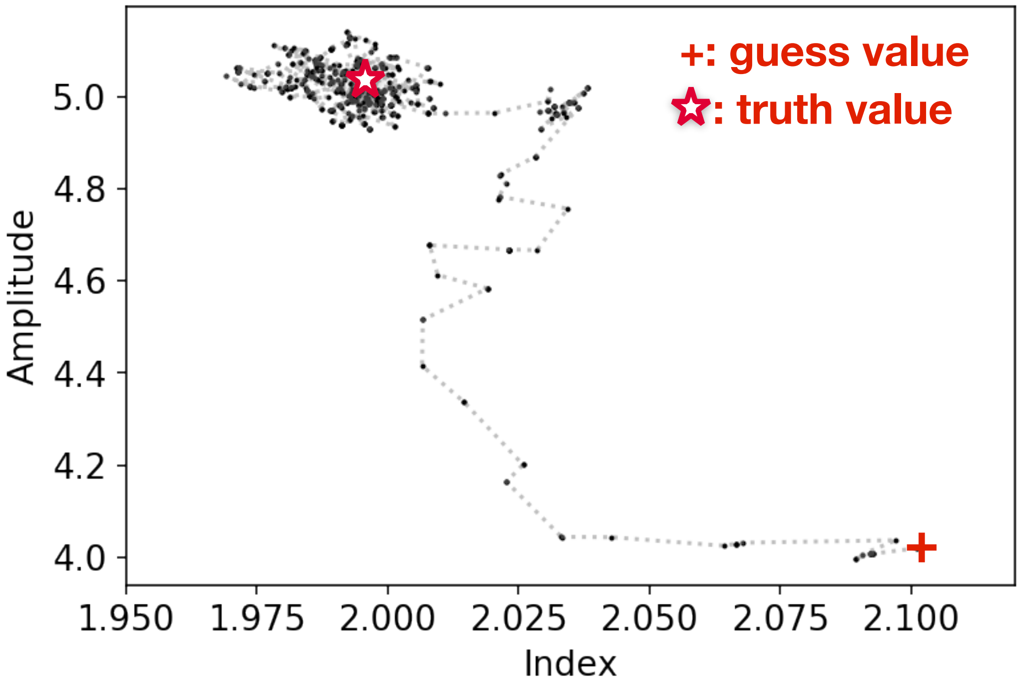


Random walk directed by potential (likelihood) spend most of their time in interesting region

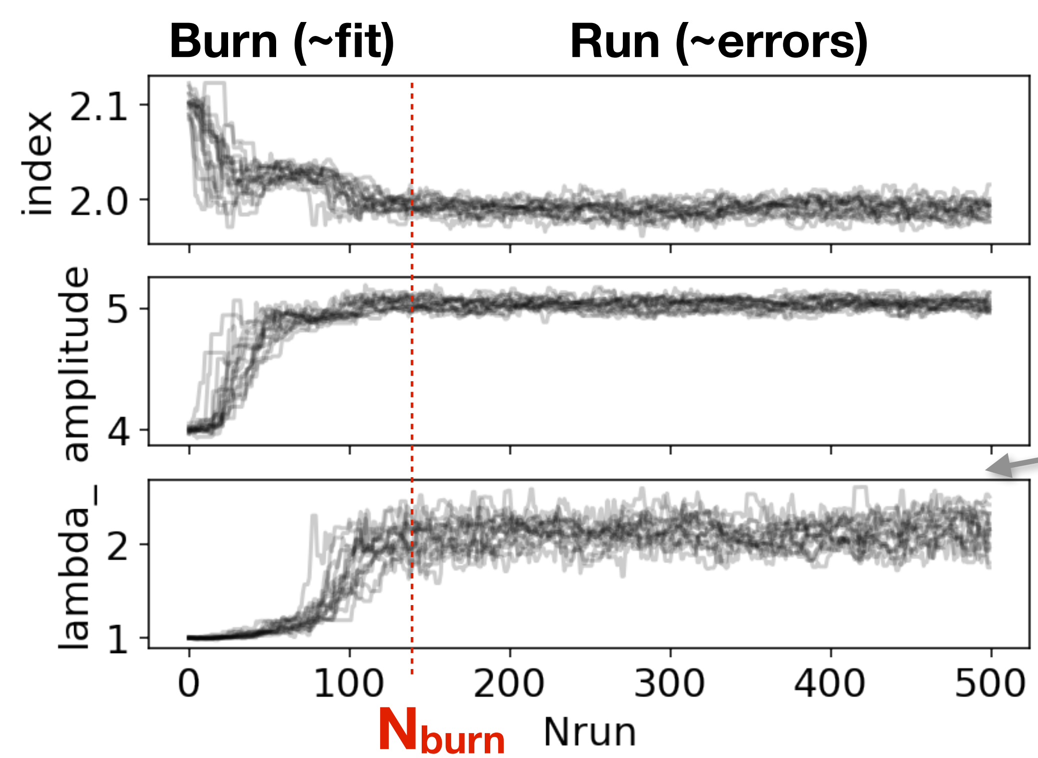
Technicaly not a fit (no convergence) it’s a phase space parameter exploration

MCMC issues : inside out approach

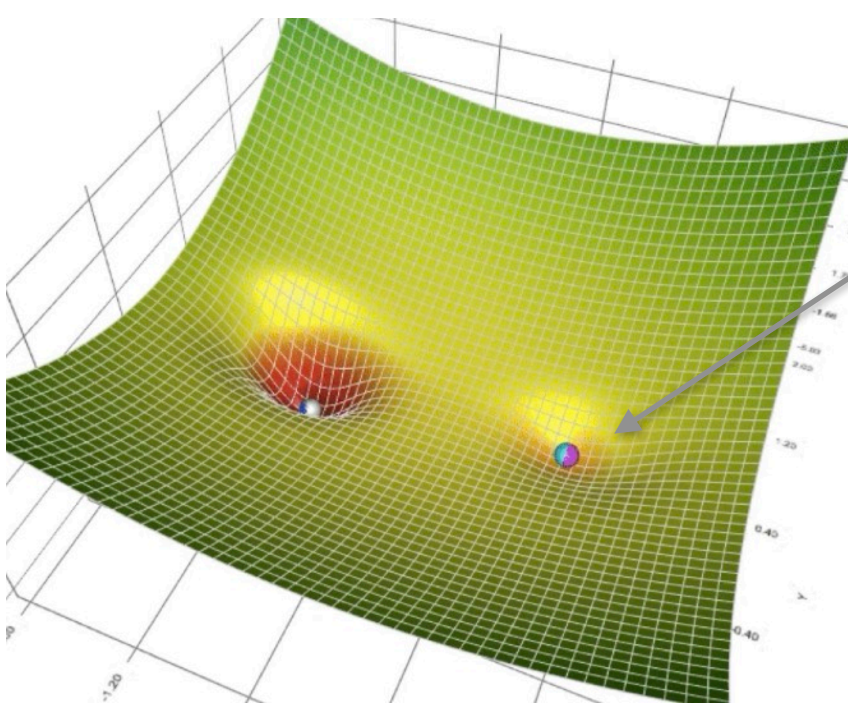
1 walker evolving for 500 steps



10 walkers evolving for 500 steps



each element of the chain are samples of the target posterior distribution.



But what if all your walkers end up here ?

Tips & tricks for fitting

- Choose a reasonable starting point
- Plot your N_{pred} counts map to investigate issues
- Set some boundaries (min, max)
    - Goal is to avoid unphysical values:
        - Negative fluxes, positions outside box, too large size
        - But be careful for upper-limit then
- Start with a simpler model and add complexity if needed:
    - Start with frozen spatial positions
    - PL first then ExpCutOff PL, source extension, etc
    - Start with the brightest sources, then fainter
    - Mask regions that are too complex
- Freeze some parameters that cannot be constrained
- Plot your spectral & spatial residuals
- Plot 1D or 2D likelihood profiles

Pray for the minimizer god
Or change religion