NON-POISSONIAN TEMPLATE FITTING ON ICECUBE

Gabriel Collin

Nick Rodd

Ben Safdi

Massachusetts Institute of Technology

History of method

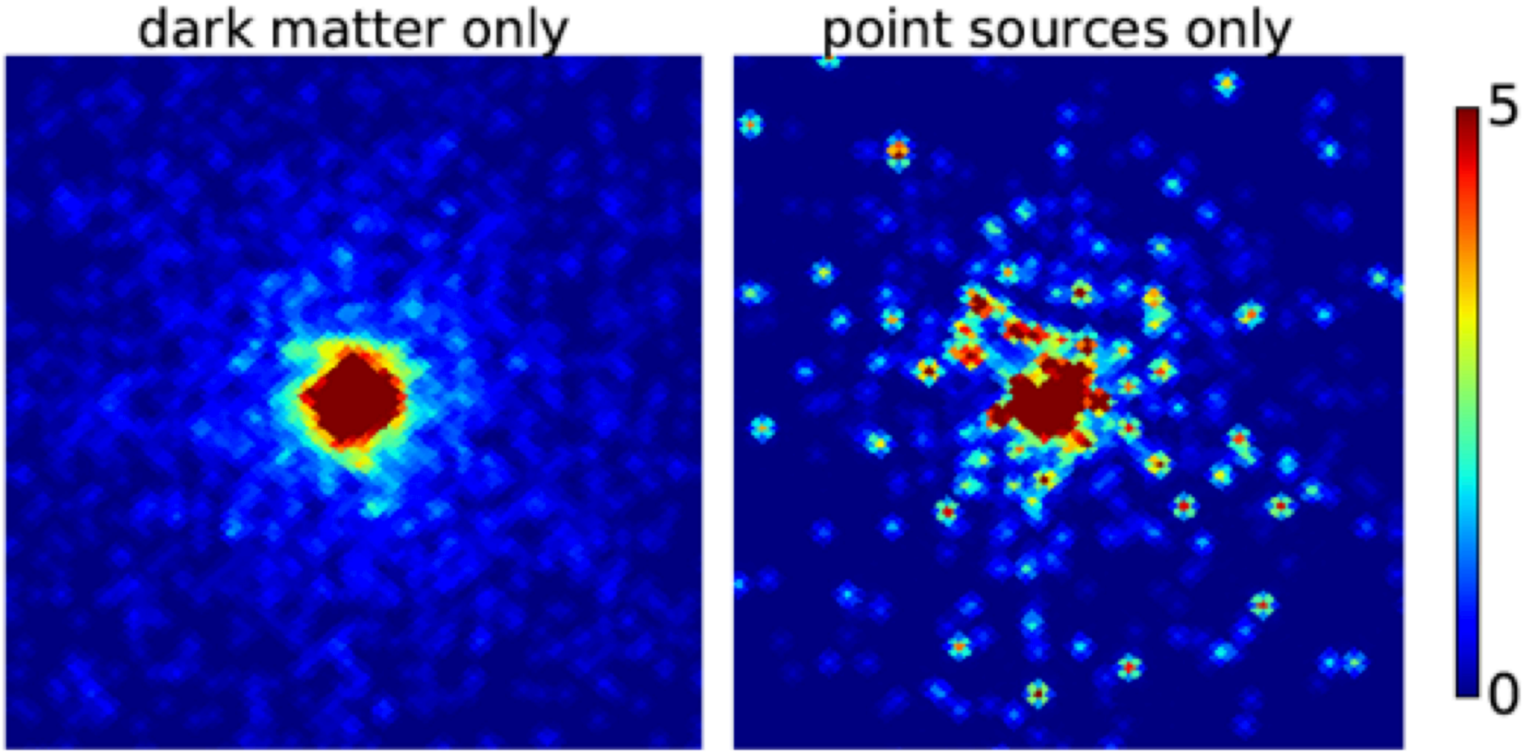
- A similar technique has a long history in radio and X-ray astronomy.
  - P(D) distributions.
  - Gaussian statistics.

- Malyshev and Hogg (arXiv:1104.0010v3) applied the P(D) distribution to Fermi data.
  - Poisson based statistics.

- Ben Safdi et. al. (arXiv:1506.05124v3) applied NPTF to the galactic center gamma ray excess.
  - Found the excess likely due to unresolved point sources.

More intuitive view

- Point sources: Many very dark and very bright pixels.
- Diffuse emission: More medium intensity pixels.



Example provided by Tracy Slatyer

More intuitive view

- I expect 10 photons per pixel, in some region of the sky. What is my probability of finding 0 photons? 12 photons? 100 photons?

- Case 1: diffuse emission, Poissonian statistics
  - P(12 photons) = 1012 e-10/12! ~ 0.1
    P(0 photons) ~ 5 x 10-5 ,
  - P(100 photons) ~ 5 x 10-63

- Case 2: population of rare sources.
  - Expect 100 photons/source, 0.1 sources/pixel - same expected # of photons
  - P(0 photons) ~ 0.9,
  - P(12 photons) ~ 0.1x10012 e-100/12! ~ 10-29 ,
  - P(100 photons) ~ 4 x 10-3
  - (plus terms from multiple sources/pixel, which I am not including in this quick illustration)

Non-Poissonian statistics

- Assume the differential source count as a function of flux (S) is a broken power law:

$$\frac{dN}{dS} \sim \begin{cases} S^{-n_1}, & S > S_{\text{break}} \\ S^{-n_2}, & S < S_{\text{break}}, \end{cases}$$

- For an average flux S, the probability of finding m photons in a bin is Poisson distributed:

$$p_m(S) = \frac{S^m}{m!} e^{-S}.$$

Non-Poissonian statistics

- The average number of sources that produce m detected photons (x_m) in this bin is then:

$$x_m = \frac{\Omega_{\rm pix}}{4\pi} \int_0^\infty dS \frac{dN}{dS}(S) \frac{S^m}{m!} e^{-S},$$

- The probability for k total detected photons is the sum of the random numbers with means x_1, x_2, x_3 ….

$$\mathbf{k} = \sum_{m=1}^{\infty} m x_m$$

$$p_k = p_{x_1} * p_{x_2} * \ldots * p_{x_m}$$

Non-Poissonian statistics

- A discrete probability distribution p_k can be defined in terms of a generating function over an auxiliary variable t:

$$P(t) = \sum_{k=0}^{\infty} p_k t^k.$$

$$p_k = \frac{1}{k!} \frac{d^k P(t)}{dt^k}\bigg|_{t=0}.$$

- The generating function for the sum of two random numbers is the product of their generators:

$$P(t) = A(t) \cdot B(t)$$

Non-Poissonian statistics

- Thus, the probability distribution for k total detected photons is given by the product of generating functions for the Poisson distribution with means x_m:

$$\sum_{k=0}^{\infty} p_k t^k = \exp\left(\sum_{m=1}^{\infty} (x_m t^m - x_m)\right) .$$

- Details of this derivation can be found in
  - arXiv:1104.0010v3

- This can be generalized to multiple bins in a sky map.

Non-Poissonian template fitting

- We can add templates to the model, which include:
  - Templates with Non-Poissonian statistics:

follows a spatial template (A_p)

$$\frac{dN_p(S)}{dS} = A_p \begin{cases} \left(\frac{S}{S_b}\right)^{-n_1} & S \geq S_b \\ \left(\frac{S}{S_b}\right)^{-n_2} & S < S_b \end{cases}$$

  - Templates with Poissonian statistics

- Templates are added together as generating functions:

$$\mathcal{P}^{(p)}(t) = \mathcal{D}^{(p)}(t) \cdot \mathcal{G}^{(p)}(t)$$

$\mathcal{G}^{(p)}(t)$: from non-Poissonian piece

$\mathcal{D}^{(p)}(t)$: from Poisson likelihood

Non-Poissonian template fitting

- The likelihood is then the product of probabilities for all pixels (p):

$$p(d|\theta, \mathcal{M}) = \prod_p p_{n_p}^{(p)}(\theta)\,.$$

- Likelihood is sampled using multinest.

Angular resolution

- So far we've assumed that every neutrino from a point source lands in the same bin.
  - IceCube's angular resolution means this wont be true.

- Instead, some neutrinos will land in bin colocated with the point source, and some will land in other bins.

Angular resolution

- We can simulate this by injecting a point source into an empty healpix map, and then counting how many bins receive a certain fraction of the total flux.

Number of healpix bins that have a fraction of total flux that lies in bin with edges of $(f, f + \Delta f)$

$$\rho(f) = \left.\frac{\Delta n(f)}{n \Delta f}\right|_{\Delta f \to 0, n \to \infty}$$

- This gives a probability distribution that is normalised to

$$\int_0^1 f \rho(f) df = 1$$

- To ensure that the total amount of flux is conserved.

Angular resolution

- Now, the average number of sources that produce m counts in a bin is

$$x_m = \frac{\Omega_{\text{pix}}}{4\pi} \int_0^\infty dS \frac{dN}{dS}(S) \int_0^1 df \rho(f) \frac{(fS)^m}{m!} e^{-fS}$$

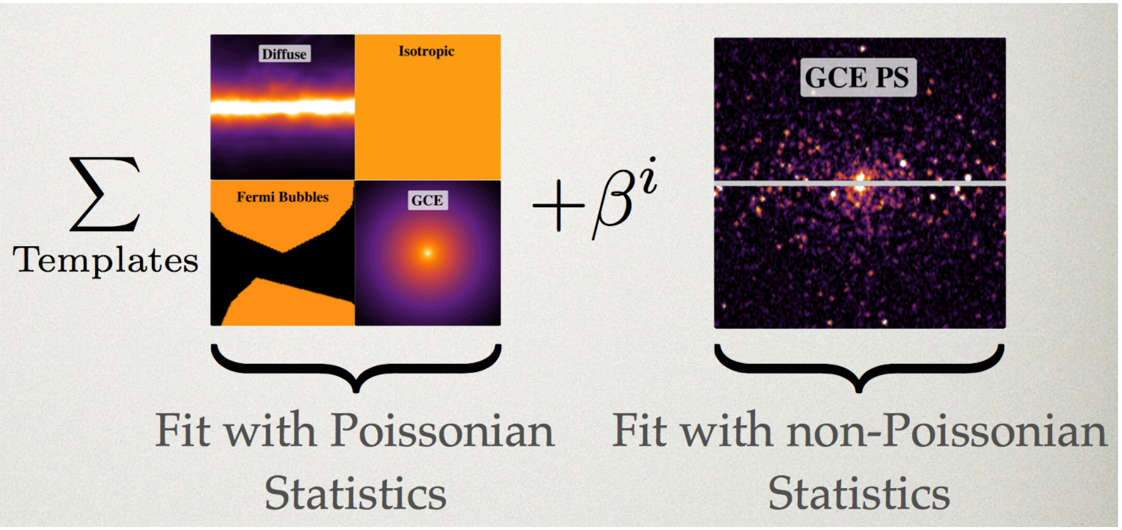
- We integrate over this fractional flux distribution to account for the fact that not all neutrinos land in one bin.
- This can also be seen as a transformation of the source count function:

$$\frac{dN}{dS}(S) \rightarrow \frac{d\tilde{N}}{dS}(S) = \int_0^1 df \frac{\rho(f)}{f} \frac{dN}{dS}(S/f)$$

Application to IceCube

- We think that the NPTF technique can be directly applied to the question of below threshold point sources in IceCube data.
  - The galactic center gamma-ray excess problem that NPTF was applied to at FermiLAT is similar in nature.

- Power of NPTF is in the templates. Eg, point sources distributed around the galactic plane:

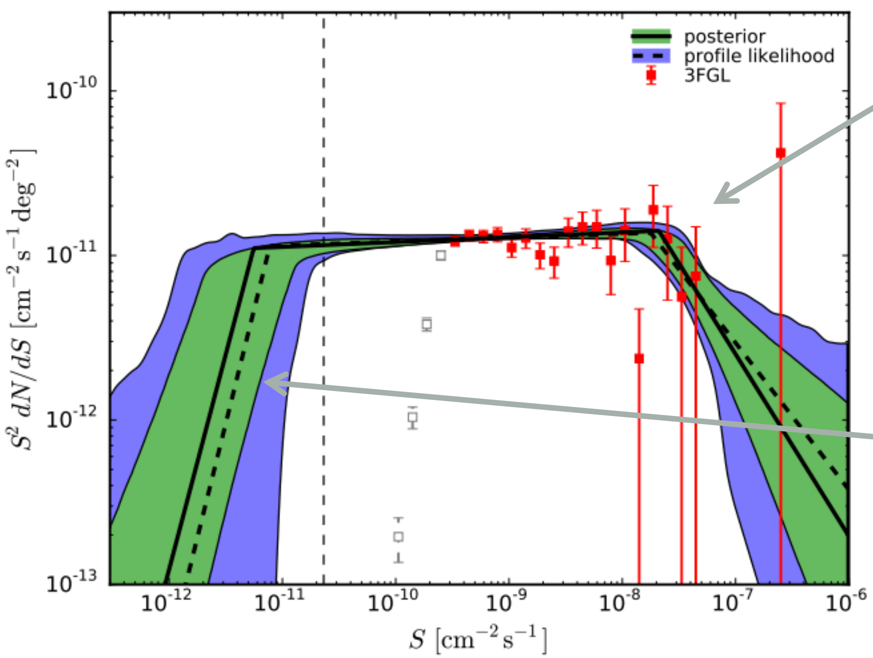


Example from Fermi

Comparison to other methods

- Comparison of non-Poissonian distribution with a catalog search:

Source count function



(a) MBPL, $N_{\rm b} = 2$

Red points: derived from 3FGL catalog.

Green at solid black: derived from non-Poissonian distribution.

arXiv:1512.07190

Conclusion

- The NPTF method has been used successfully at Fermi-LAT to identify a population of point sources.

- This method complements existing techniques at IceCube by:
  - Fitting to a spatial distribution (template).
  - Fitting a source count function (dN/dS)

- We are currently developing a 7 year NPTF analysis on IceCube.