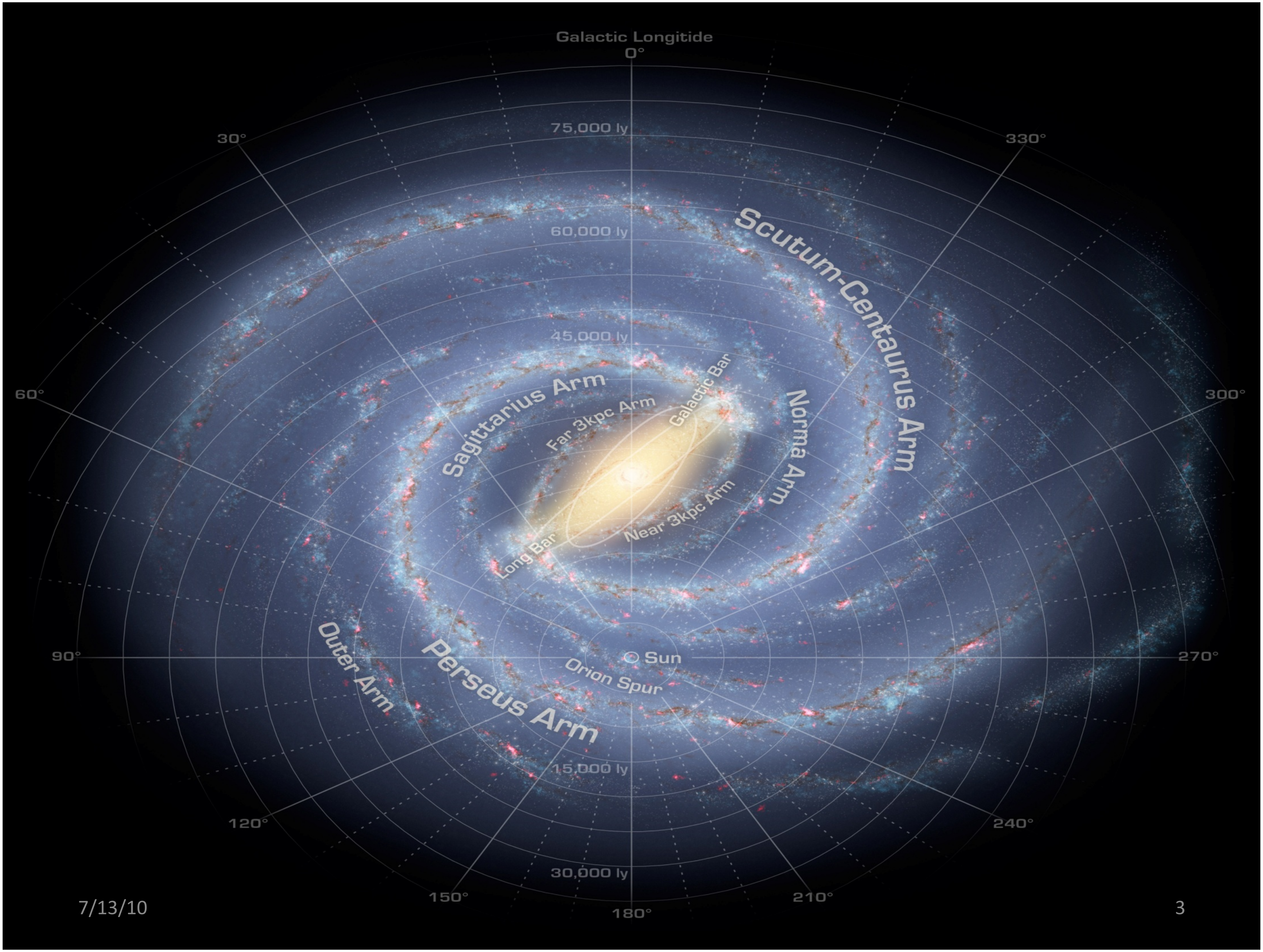
Testing Astronomical Source Extension TS_{ext} using MC. Does Wilks' Theorem Apply?

Ping Wang, Elliott Bloom, Joshua Lande
KIPAC-SLAC, Stanford University
Representing the Fermi LAT Collaboration

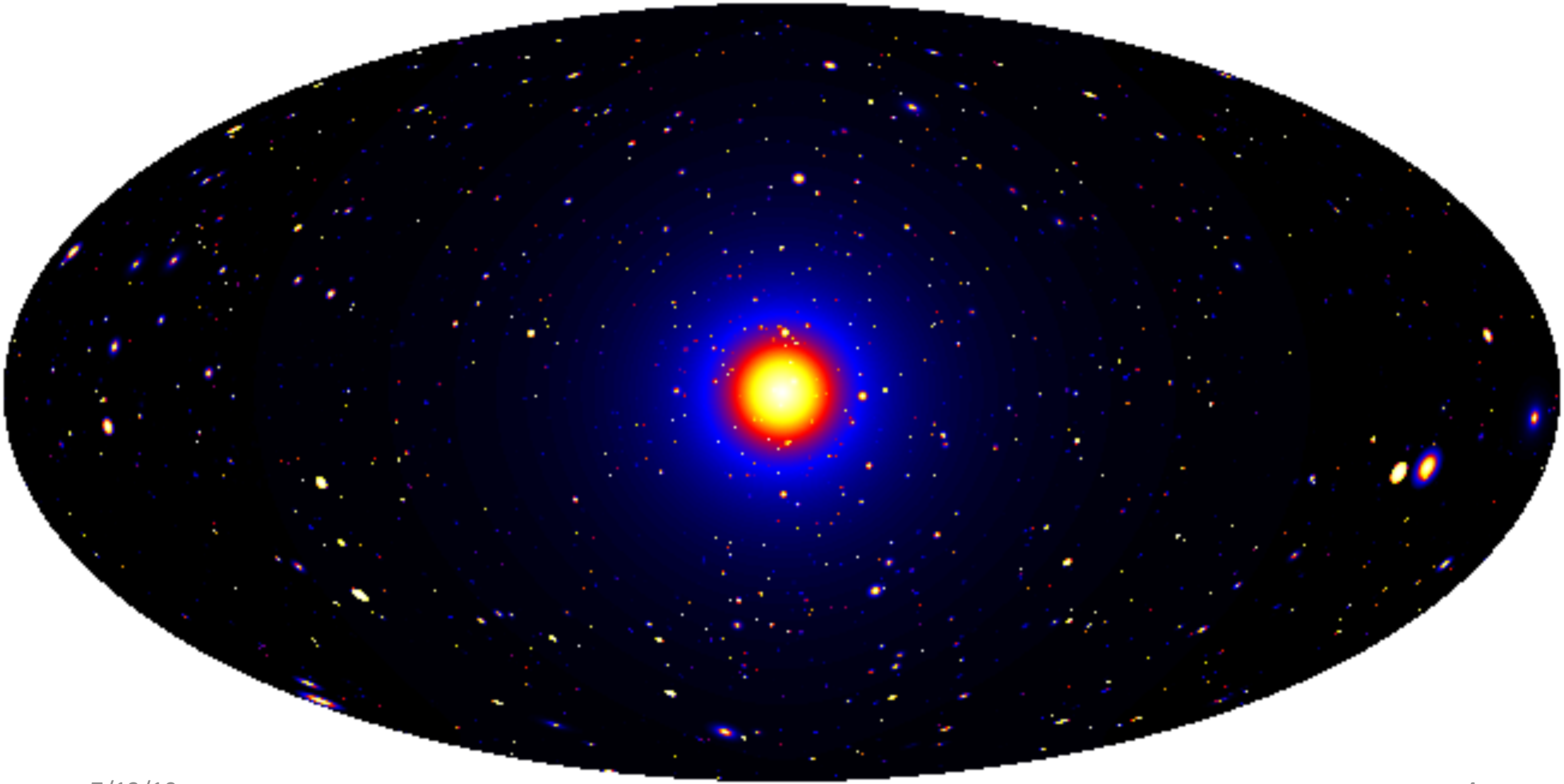
Presented by E. Bloom at Discovery Statistics Workshop
Banff, Canada
7/13/2010

Outline

- Short Science motivation and definitions – Fermi DM targets.
  - Many Fermi Gamma-ray sources are expected to be spatially extended - Important to develop a statistical test to find them.
- The theoretical prediction for the Test Statistic (TS) for nested models.
  - Example from EGRET
  - Fermi LAT nested TS to determine the angular extension of an astronomical source, comparison to Wilk's theorem.
- Fermi LAT TS distribution for testing for source extension using independent MC simulations and comparison to Wilk's theorem.
  - Source at fixed position in sky for all tries.
  - Source at random high latitude position in sky for each try of many.
- Summary and Conclusions


Galactic Longitide
0°
30°
330°
60°
300°
90°
270°
120°
240°
150°
210°
180°
75,000 ly
60,000 ly
45,000 ly
15,000 ly
30,000 ly
Scutum-Centaurus Arm
Sagittarius Arm
Far 3kpc Arm
Galactic Bar
Norma Arm
Near 3kpc Arm
Long Bar
Outer Arm
Perseus Arm
Orion Spur
Sun

The Galaxy Shining in High Energy Gammas from the Annihilation of Dark Matter



DM realization from Taylor and Babul (2005)

Definition of Test Statistic, TS, and Application to Wilks' Theorem

- TS = 2 * (ln(L1) – ln(L0))
  - L0: null hypothesis
  - L1: alternative hypothesis
- Wilks' theorem posits that the TS distribution approximately follows a χ^2 distribution when comparing two hierarchically nested models, i.e., L0 and L1 are nested. Restricting to positive fluctuations only, the TS distribution is distributed as $\chi^2/2$ because about one-half of the samples have TS<=0.

Definition of TS –EGRET Nested Example

– *"The Likelihood Analysis of EGRET Data",* J. R. Mattox, et al. ApJ 461: 396-407 (1996)

http://articles.adsabs.harvard.edu/full/1996ApJ...461..396M

- L0: background only
- L1: background and a point source
- Restrict to positive fluctuations (physical)

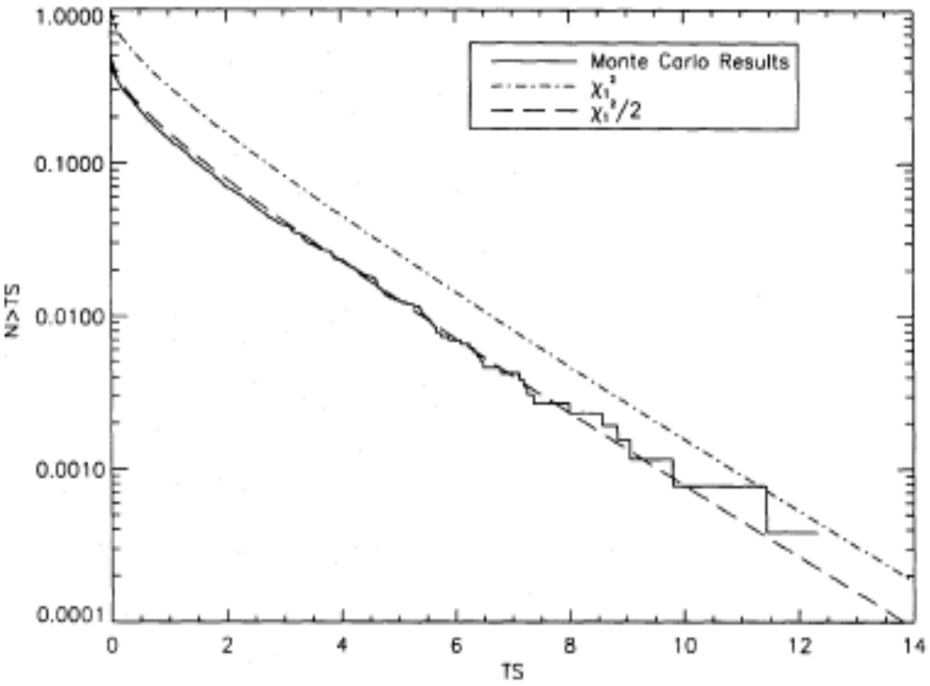


FIG. 3.—Cumulative distribution of T_s in the null hypothesis at a fixed oint. Several thousand independent likelihood ratio tests were done for imulated data with no point sources.

Definition of TS - Fermi LAT Source Extension

- Fermi LAT Likelihood-fit for extension of source: L0 = background and a “point source” (NFW distribution with 10^{-8} radian extension).
- L1= background and an extended NFW source
  - NFW is a density distribution for dark matter clumps found by Navarro, Frenk and White.
    - We also consider Gaussian extension
  - Constraint on extension of source: >=0 as source extension cannot be a negative number (different than signal for a point source over background). This may cause problems in the TS distribution.
  - Strictly speaking, Wilk's theorem should not hold in this case because the null hypothesis is on the edge of parameter space.

Fermi LAT Likelihood Extension Tests: Independent Monte Carlo Simulations

1. There are three model components to be simulated for the null hypothesis: the isotropic extragalactic diffuse, the conventional GALPROP diffuse and a power law point source at a fixed location which has a power law index -2.0 and flux 6.0e-9 /cm^2/s from 100 MeV to 300 GeV.

2. Choose a source location at (l, b) = (272.16^o, 36.13^o) as representative of a high latitude source.

3. Use gtobssim(GLAST Tools Observation Simulation) to generate 1 year of simulations of the three components (random seeds). Note that the source is always at the same (l, b).

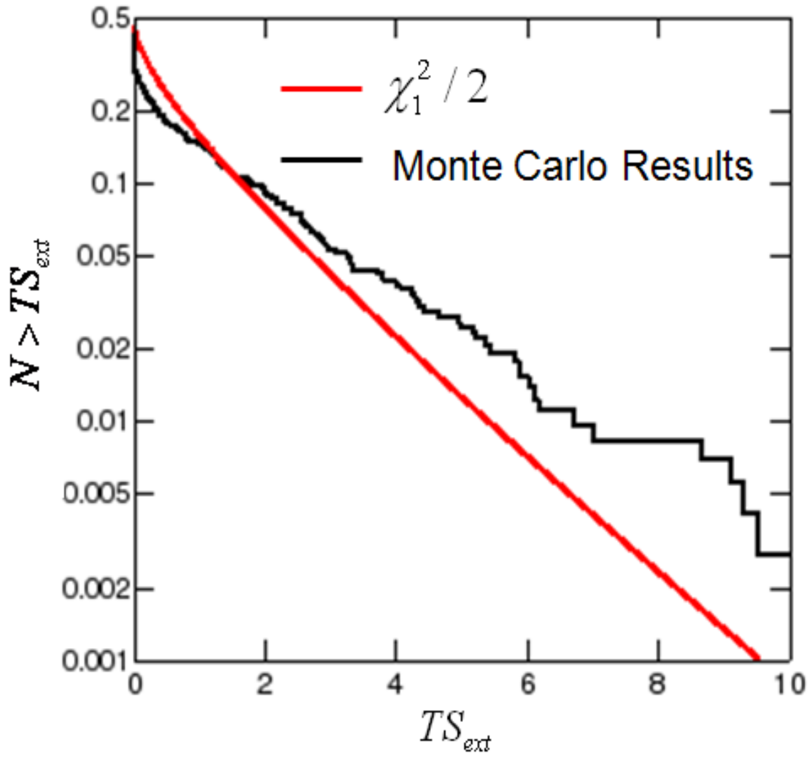
4. Use sourcelike to calculate TS_point for the null hypothesis and TS_nfw for the alternative hypothesis for the data set from step 3.

5. Calculate TS_ext = TS_nfw – TS_point for the data set.

6. Repeat steps 3 – 5 a 1000 times (724 fits passed all the steps).

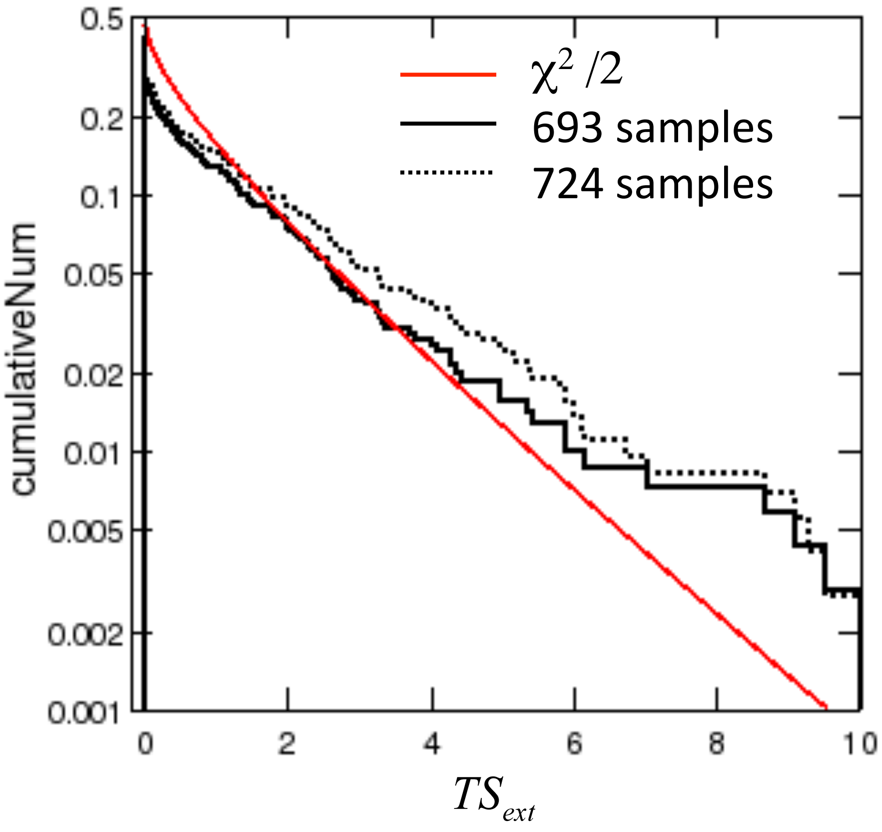
Power law index -2.0 and flux 6.0e-9 /cm^2/s: starting value 0.5°

- 724 samples
  - Starting value R0 = 0.5° : 31 samples with negative second-derivative matrix
  - Normalized cumulative distribution of TS_ext and $\chi^2/2$



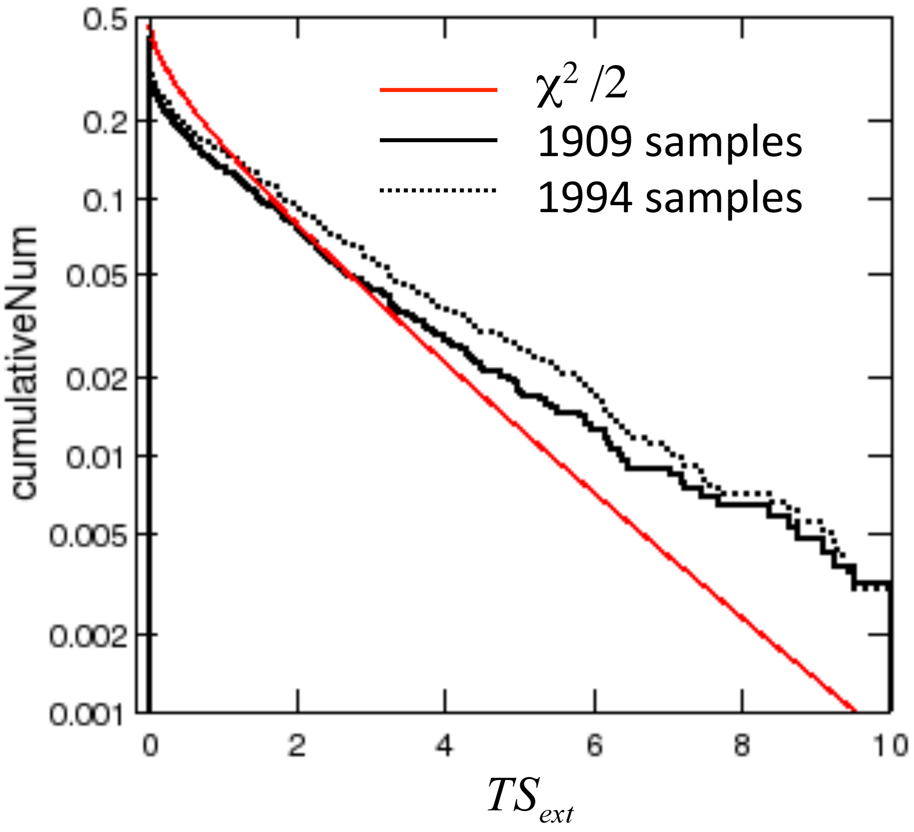
Power law index -2.0 and flux 6.0e-9 /cm^2/s: positive definite second-derivative matrix

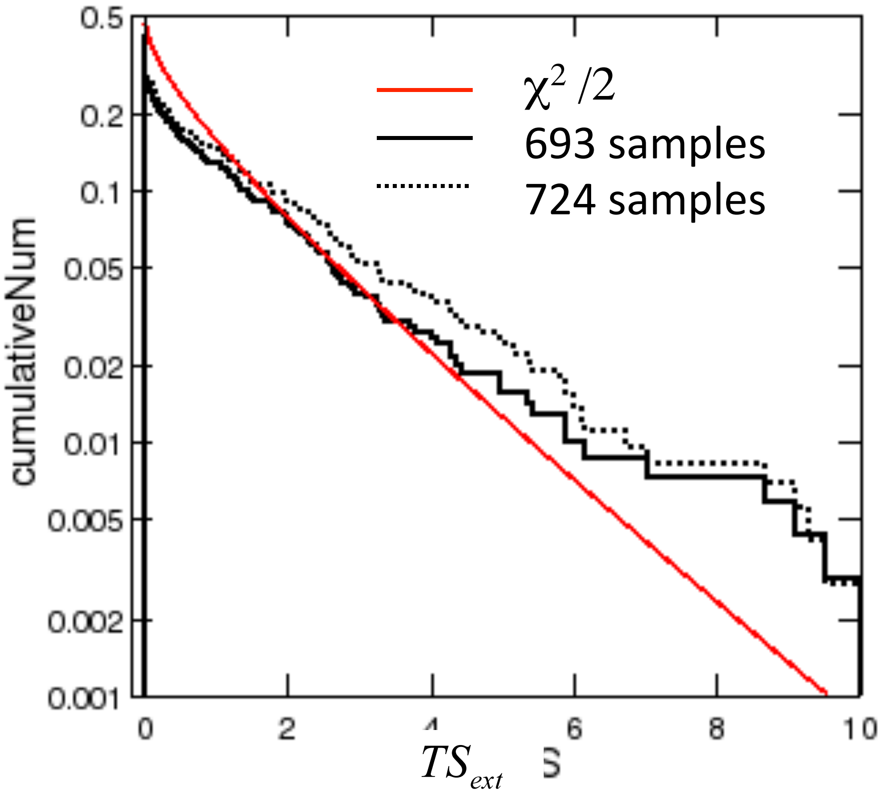
- 693 out of 724 samples with positive definite second-derivative matrix
  - Starting value R0 = 0.5° : 31 samples with negative second-derivative matrix
  - Normalized cumulative distribution of TS_ext and $\chi^2/2$



Power law index -2.0 and flux 6.0e-9 /cm^2/s: more samples

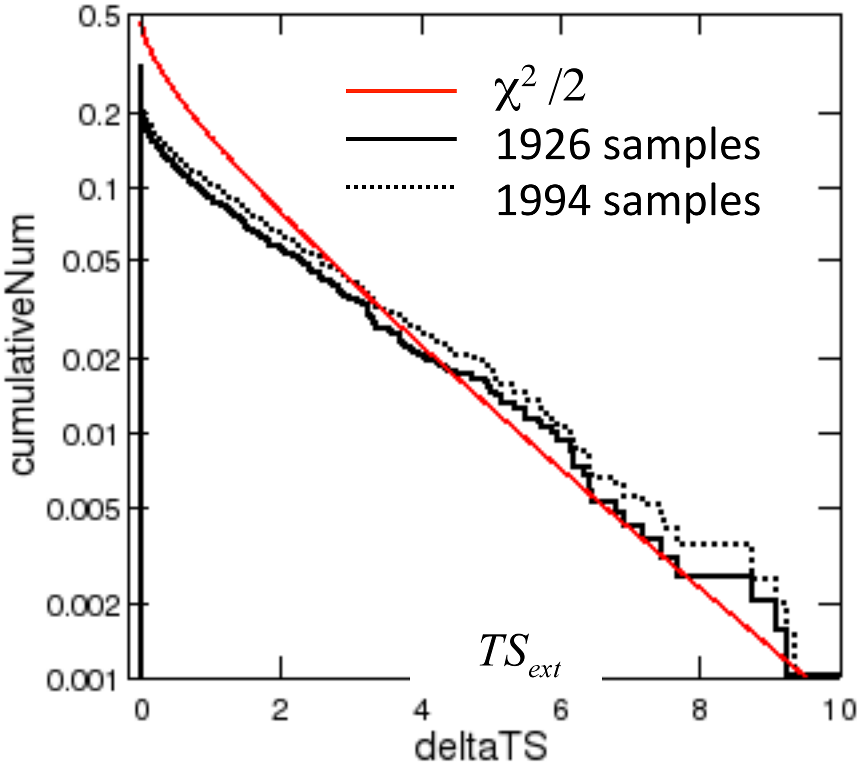
- 1994 samples
  - Starting value R0 = 0.5°: 85 samples with negative second-derivative matrix
  - 1909 samples with positive definite second-derivative matrix
  - Normalized cumulative distribution of TS_ext and $\chi^2/2$





Power law index -2.0 and flux 6.0e-9 /cm^2/s:
starting value 0.01°

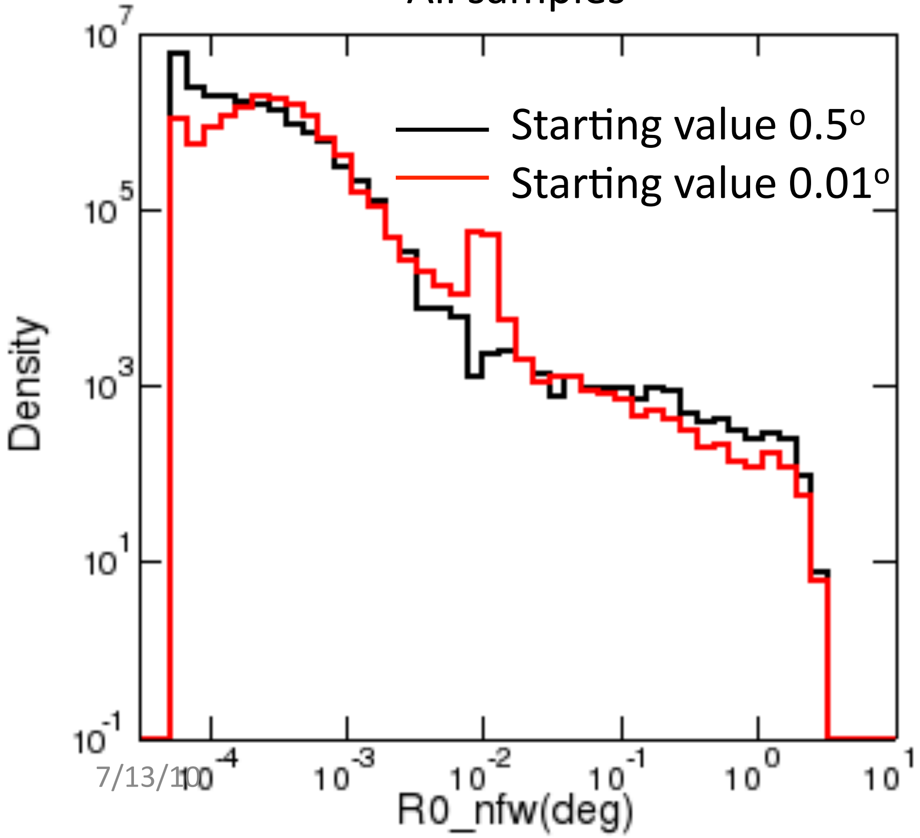
- 1994 samples
  - Starting value R0 = 0.01°: 68 samples with negative second-derivative matrix
  - 1926 samples with positive definite second-derivative matrix
  - Normalized cumulative distribution of TS_ext and $\chi^2/2$



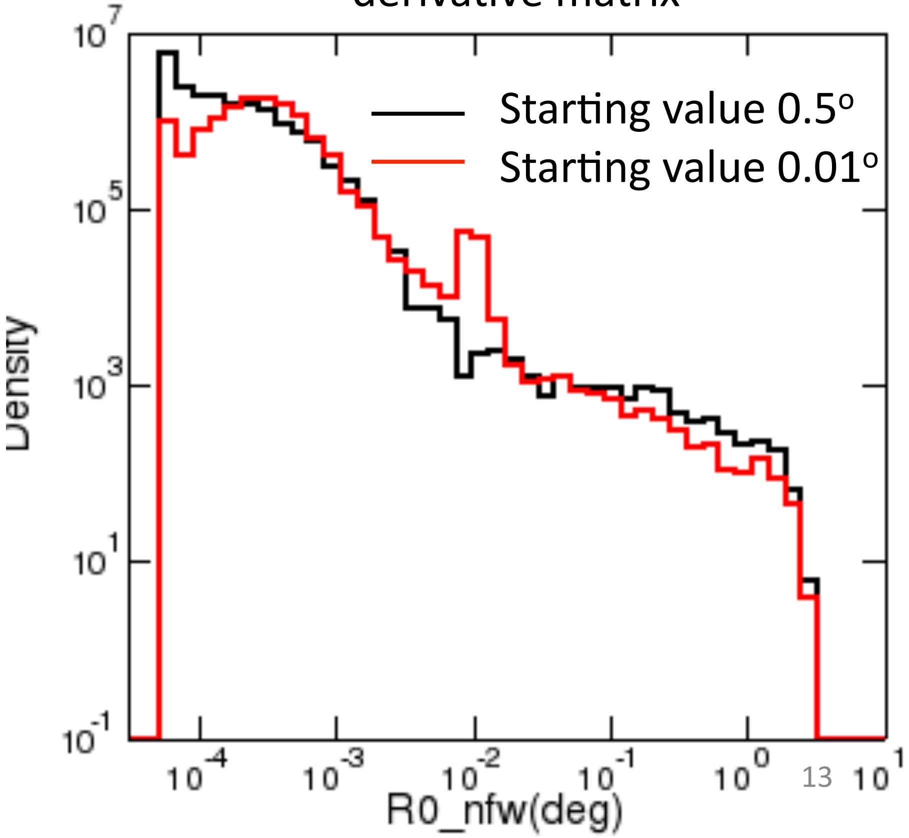
Power law index -2.0 and flux 6.0e-9 /cm^2/s: best fit R0 (angular size)

- 1994 samples
  - Starting value R0 = 0.5° : 85 samples with negative second-derivative matrix
  - Starting value R0 = 0.01° : 68 samples with negative second-derivative matrix
  - Best fit extension R0

All samples

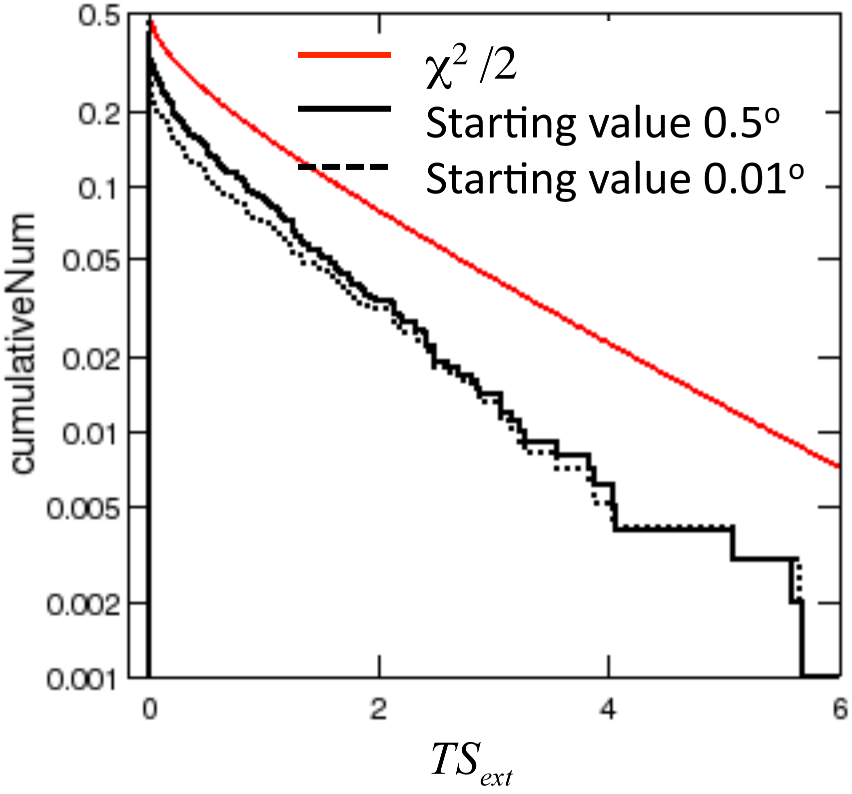


Samples with positive second derivative matrix



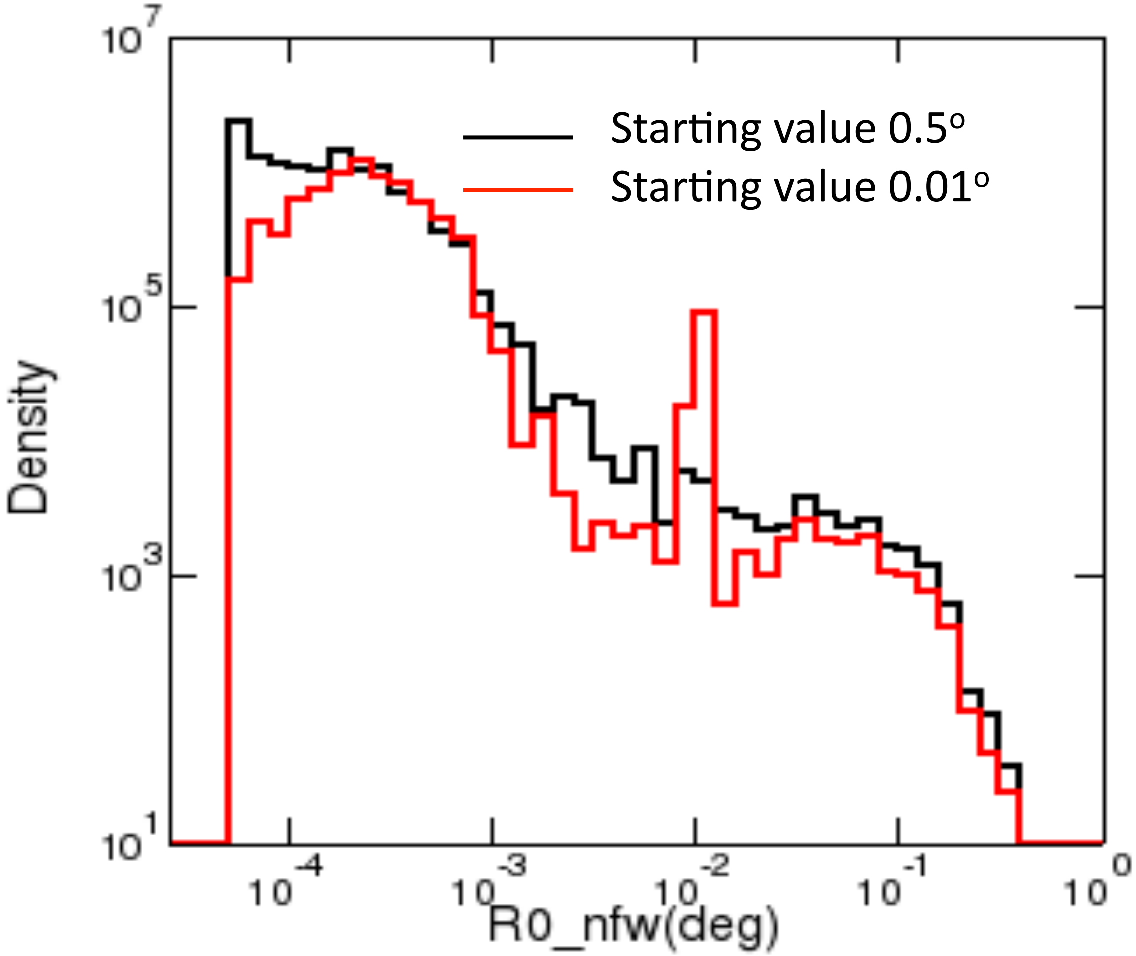
Power law index -2.0 and flux 6.0e-8 /cm^2/s

- 1000 samples
  - Starting value R0 = 0.5°: no sample with negative second-derivative matrix
  - Starting value R0 = 0.01°: 10 samples with negative second-derivative matrix
  - Only consider the samples with positive definite second-derivative matrix
  - Normalized cumulative distribution of TS_ext and $\chi^2/2$

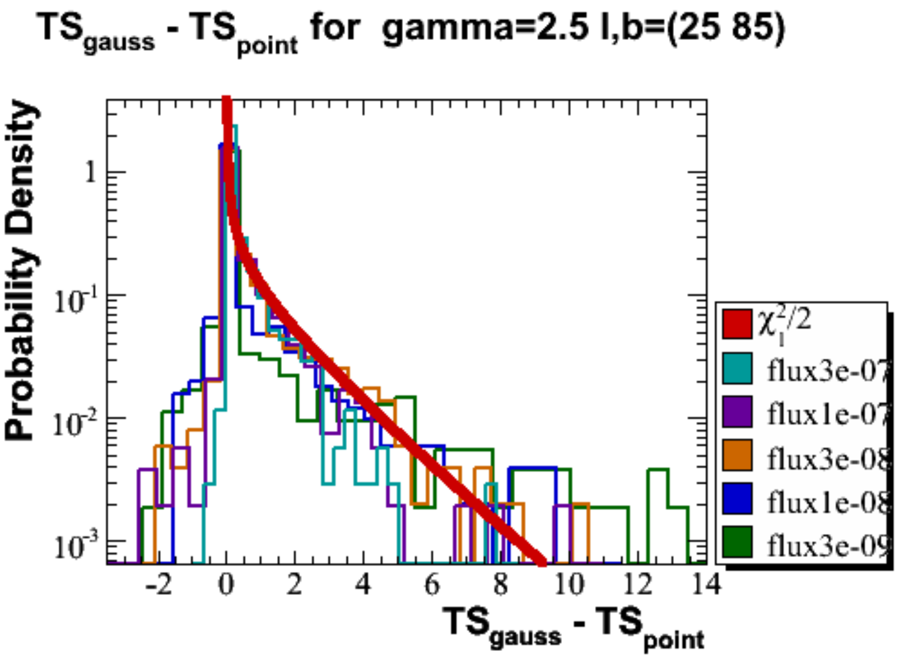


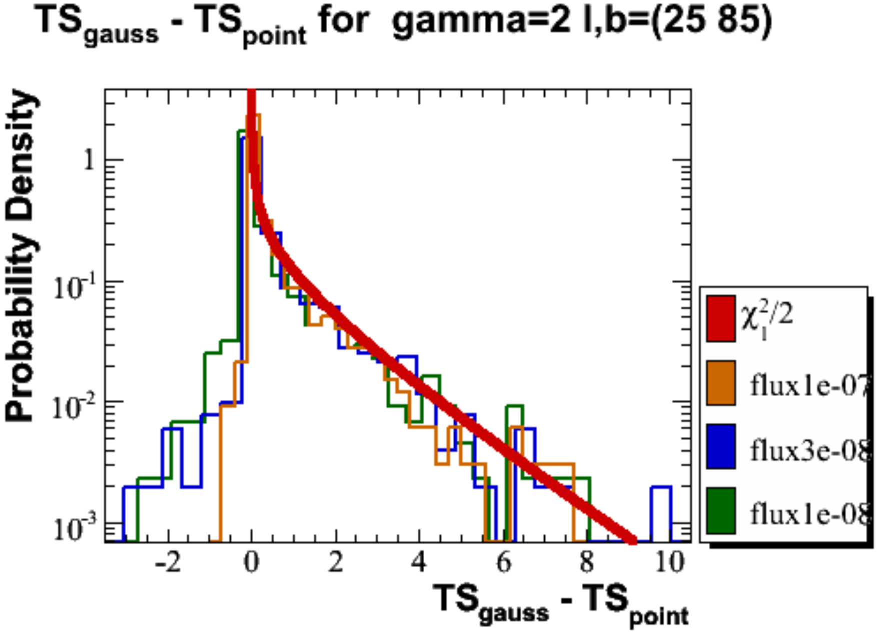
Power law index -2.0 and flux 6.0e-8 /cm^2/s: best fit R0

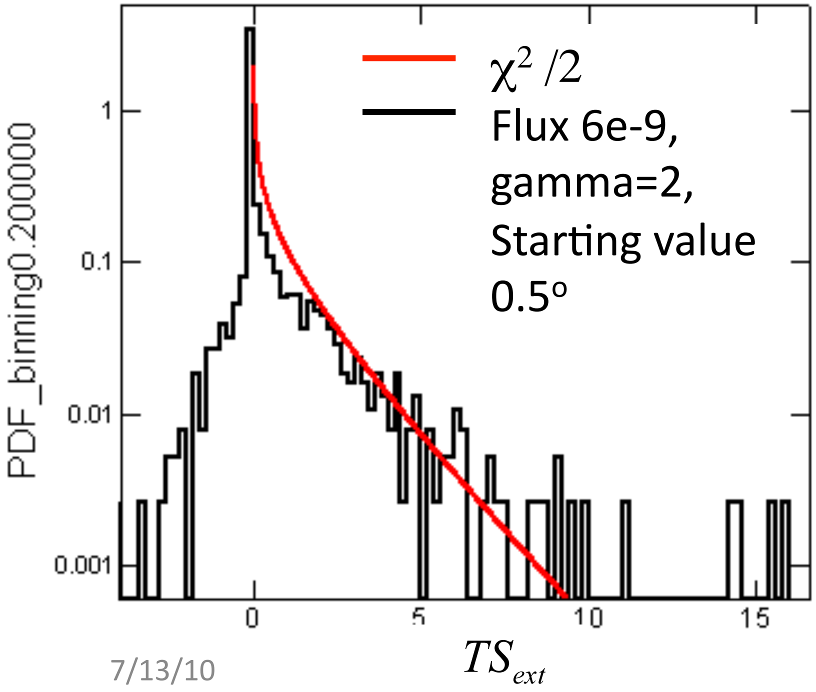
- 1000 samples
- Only consider the samples with positive definite second-derivative matrix
- Best fit extension R0

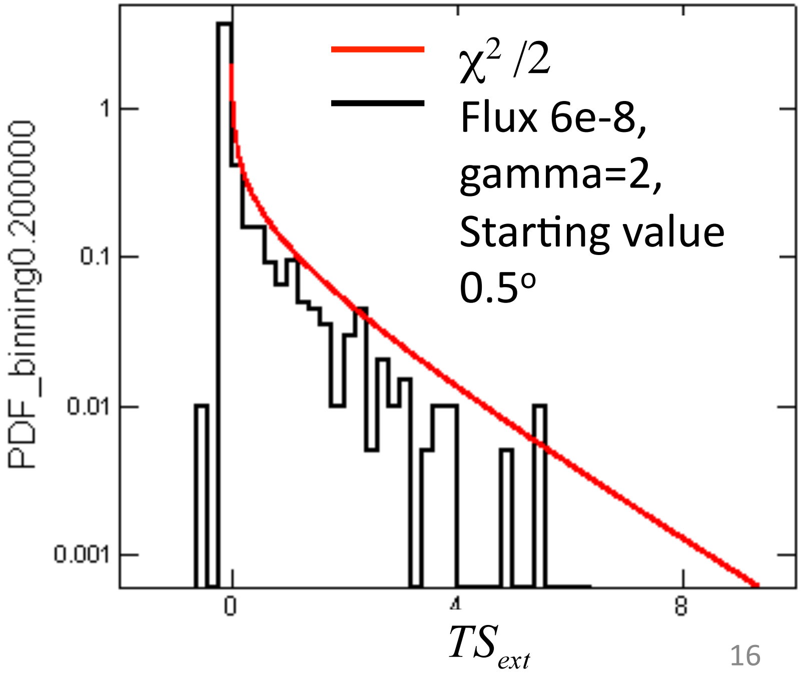


Gaussian angular size









Test source extension Likelihood TS with source at random |b| >20 deg.

- Methodology
  - Instead of simulating and fitting a point source in one location for many times, we throw a MC point source to the MC diffuse background at a random locations for |b|>20deg for 1000 times. The distribution of the randomly chosen locations on the sky for |b|>20 deg is uniform.
  - This method can average the systematic fitting fluctuations for the high latitude sky, and does not bias for any specific location.
  - 12-month Diffuse background simulation: isotropic diffuse and Ring Galactic Diffuse model, P6_V3
    - The background simulation is from 30 MeV to 100 GeV, and the data used in this test is from 200 MeV to 100 GeV.
  - 12-month point source simulation: different power law spectrum models
  - Obtain the value of TS_ext for the significance level (false detection probability) 0.01 (TS_ext_0.01)

Pwl 2.0: 1000 samples, starting value R0=0.5°

TS = 23.8 ⇔ 4 sigma significance
TS = 33.8 ⇔ 5 sigma significance

The average values of TS_nfw and TS_point for 1000 samples

Model	Power law index	Flux from 100 MeV to 300 GeV (cm^{-2}s-^{1})	<TS$_{nfw}$>	<TS$_{point}$>	TS$_{ext}$ for significance level 0.01 (5.41 if chi2/2)
1	-2.0	1.0E-8	73.46	73.16	5.45
2	-2.0	6.0E-9	39.27	38.52	7.22
*	-2.0	3.0E-9	21.70	21.16	12.07

< 4 sigma, very weak source

Pwl 2.0: 1000 samples, starting value R0=0.1°

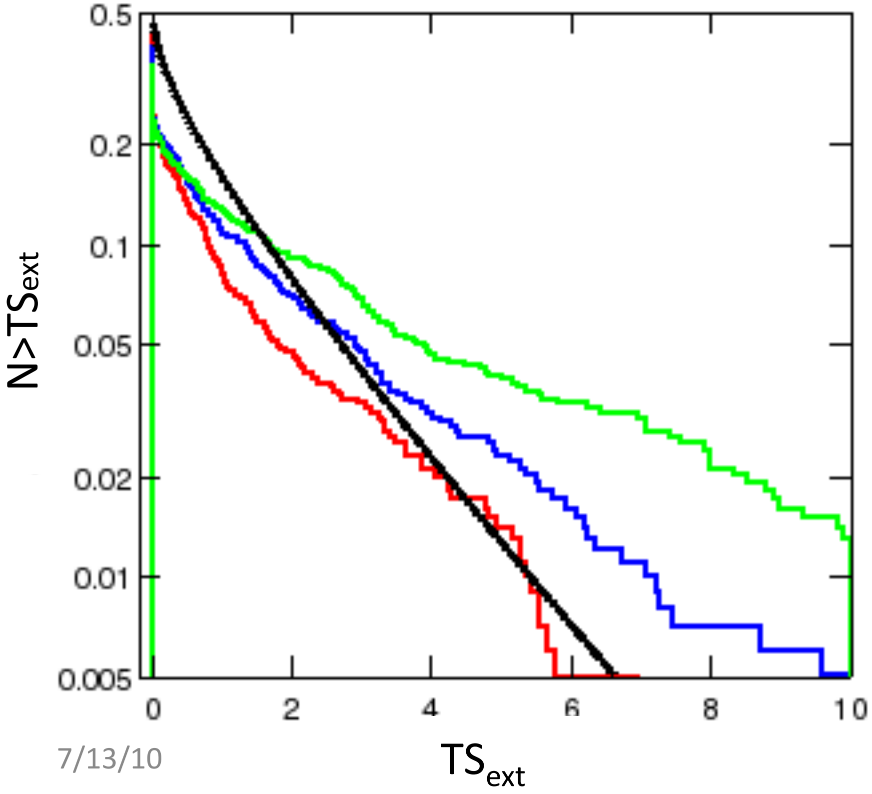
TS = 23.8 ⇔ 4 sigma significance
TS = 33.8 ⇔ 5 sigma significance

Model	Power law index	Flux from 100 MeV to 300 GeV ($cm^{-2}s^{-1}$)	$<TS_{nfw}>$	$<TS_{point}>$	TS_{ext} for significance level 0.01 (5.41 if chi2/2)
1	-2.0	1.0E-8	73.45	73.16	5.30
2	-2.0	6.0E-9	38.89	38.52	6.21
*	-2.0	3.0E-9	21.53	21.16	9.03

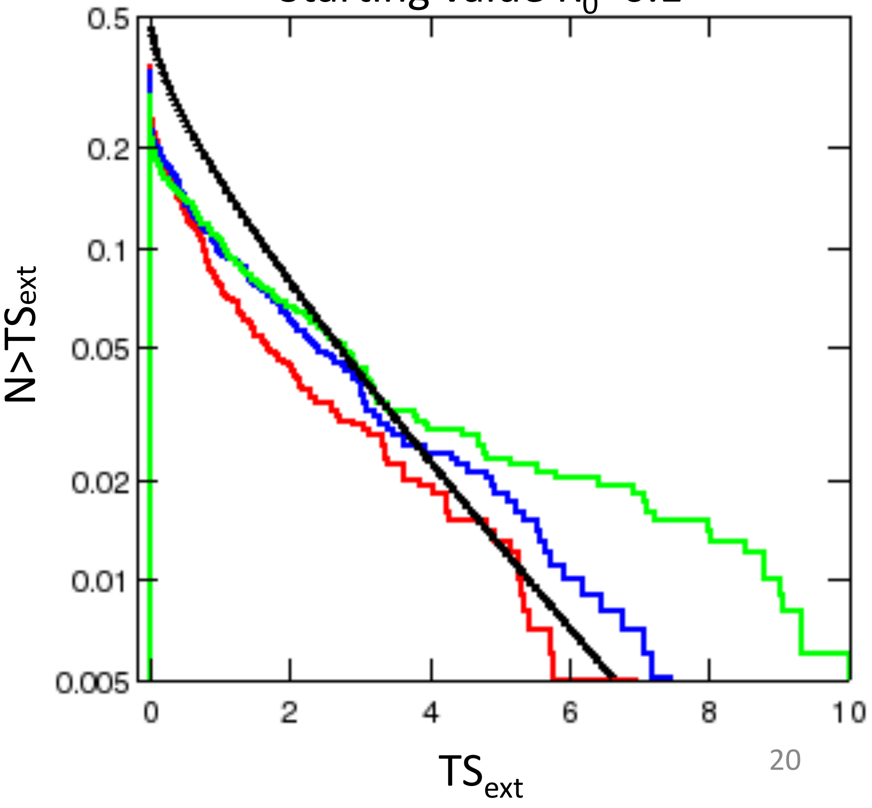
Pwl 2.0: TS_ext, 1000 samples

Black: chi2/2
Red: pwl2.0_flux1.0e-8
Blue: pwl2.0_flux6.0e-9
Green: pwl2.0_flux3.0e-9

Starting value R_0=0.5°



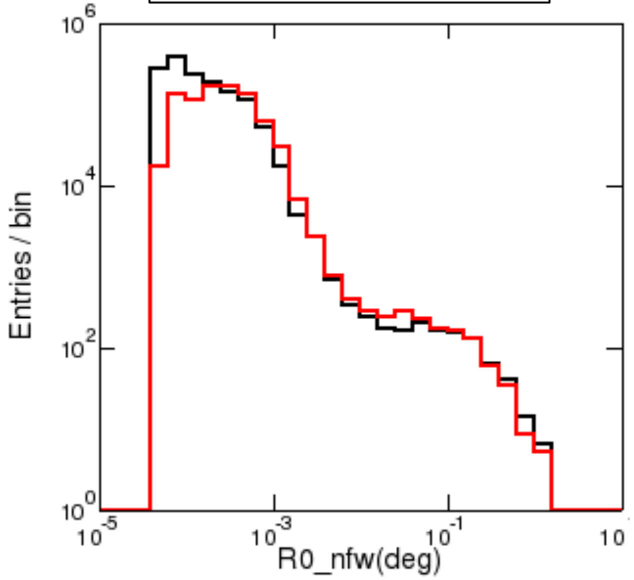
Starting value R_0=0.1°



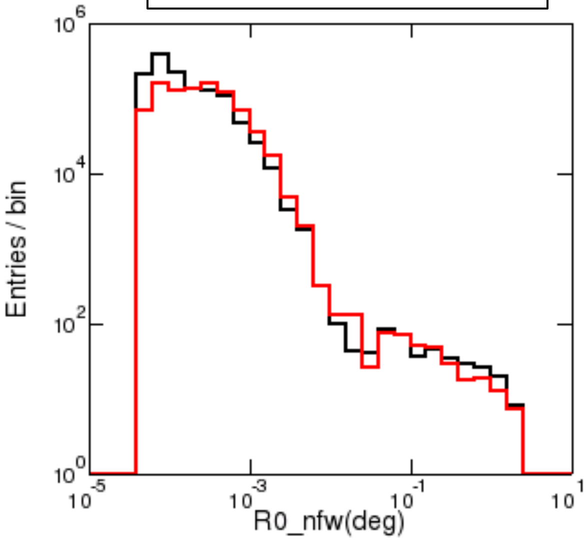
Pwl 2.0: fitted R0, 1000 samples

Black: Starting value R_0=0.5°
Red: Starting value R_0=0.1°

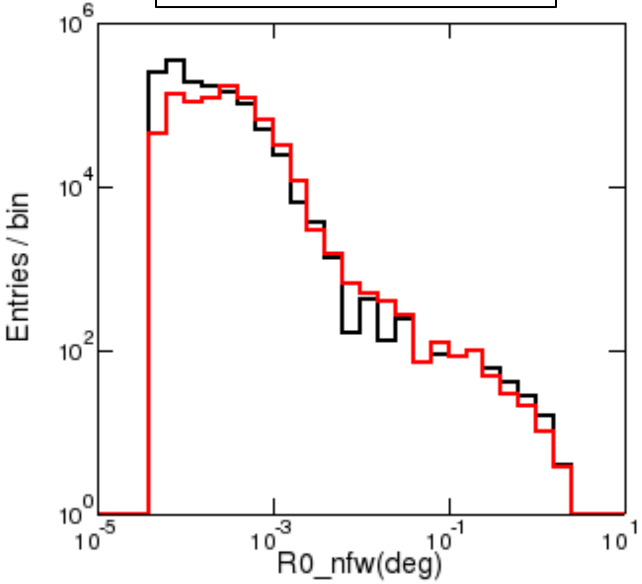
pwl2.0_flux1.0e-8



pwl2.0_flux3.0e-9



pwl2.0_flux6.0e-9

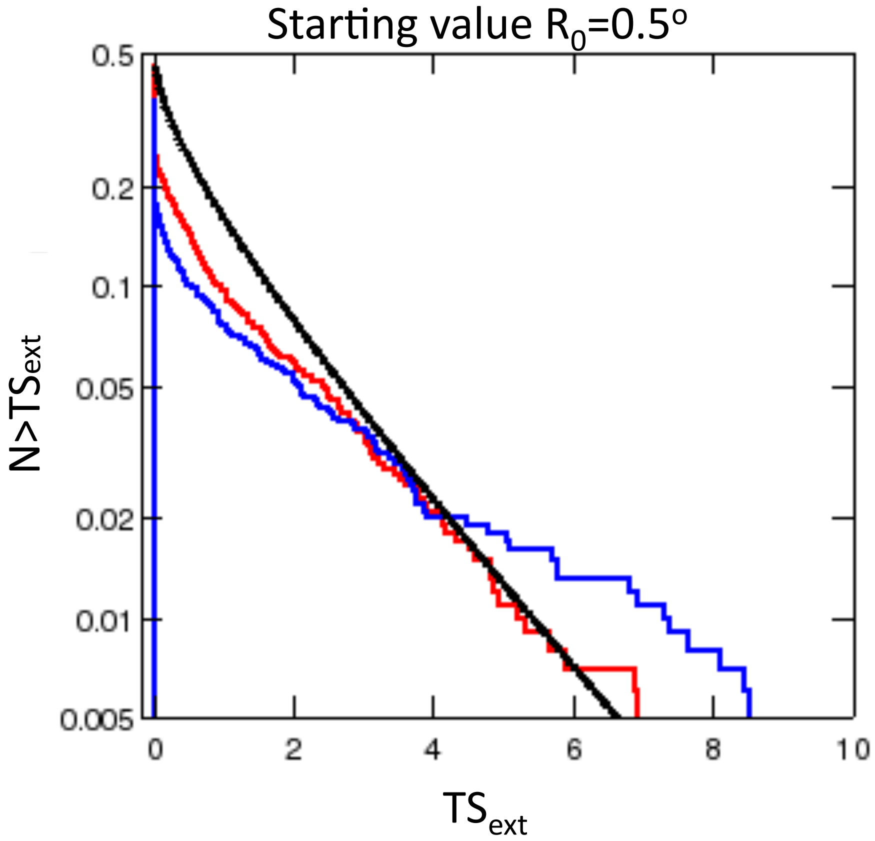


Pwl 1.5: 1000 samples, starting value R0=0.5^o

TS = 23.8 ⇔ 4 sigma significance
TS = 33.8 ⇔ 5 sigma significance

Model	Power law index	Flux from 100 MeV to 300 GeV (cm^{-2}s-^{1})	<TS$_{nfw}$>	<TS$_{point}$>	TS$_{ext}$ for significance level 0.01 (5.41 if chi2/2)
3	-1.5	2.2E-9	67.66	67.38	5.31
4	-1.5	1.3E-9	37.83	37.58	7.40

Pwl 1.5: TS_ext, 1000 samples



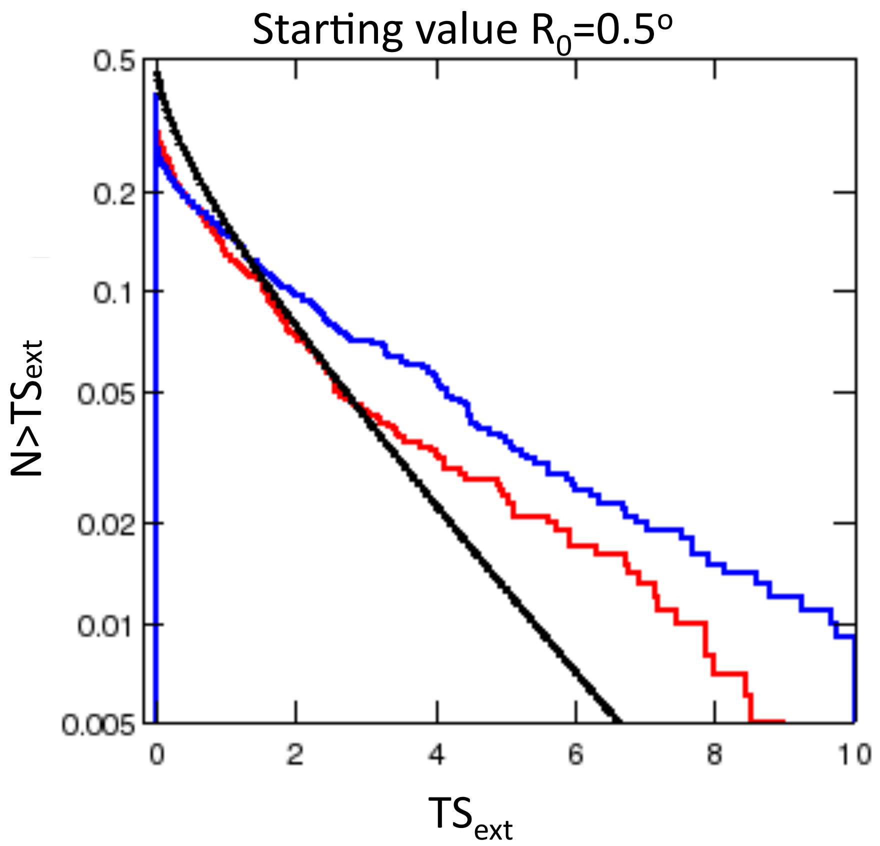
Black: chi2/2
Red: pwl1.5_flux2.2e-9
Blue: pwl1.5_flux1.3e-9

Pwl 2.5: 1000 samples, starting value R0=0.5°

TS = 23.8 ⇔ 4 sigma significance
TS = 33.8 ⇔ 5 sigma significance

Model	Power law index	Flux from 100 MeV to 300 GeV ($cm^{-2}s^{-1}$)	$<TS_{nfw}>$	$<TS_{point}>$	TS_{ext} for significance level 0.01 (5.41 if chi2/2)
5	-2.5	3.2E-8	80.57	80.12	7.88
6	-2.5	1.6E-8	32.62	32.04	9.73

Pwl 2.5: TS_ext, 1000 samples



Black: chi2/2
Red: pwl2.5_flux3.2e-8
Blue: pwl2.5_flux1.6e-8

Summary and Conclusions

- Power law index and flux affect the distribution of the TS_ext.
- Starting value of the extension also affects the distribution of the TS_ext.
- For power law index -2.0, TS_ext distribution is more consistent with chi2/2 for the starting R0=0.1deg, except for the very faint source (<4 sigma).
- We conclude that Wilks' Theorem is not useful for this particular application as its use typically systematically underestimates the TS_ext significance level (false detection probability) at 0.01 (and lower).
- We are studying the issue further to try to find a root cause for this problem, and maybe find an approach that will fix it.

Guesses from Josh Besides Wilks' Theorem Does Not Apply

- May be due to the way Sourcelike (Fermi LAT fitter) does energy bin independent fits. It may be the case that for a point source, one of the energy bins fails to converge on the true value even though for the extended source (even with a small extension) there is correct convergence in the bin. In this case, you would see a Gaussian extension fit significantly better than a NFW extension, even for the same physical extension.

- Another possible cause may be due to not including the energy dispersion in the analysis.

- Another possible cause may be due to assuming a spectral index of 2 inside of each energy bin in sourcelike to calculate the exposure across the bin when doing the fit, although this wouldn't cause a problem for the simulated index 2 sources.