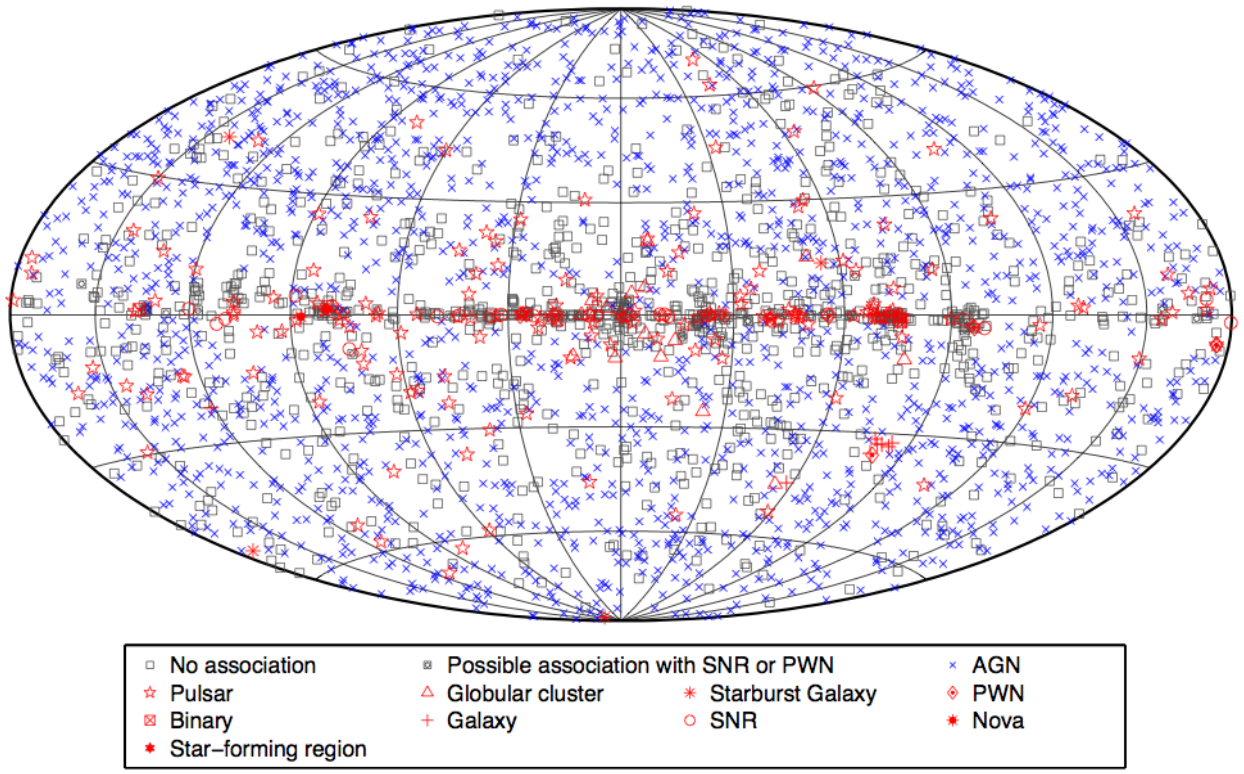
Population synthesis of Fermi LAT sources: A Bayesian analysis using posterior predictive distributions

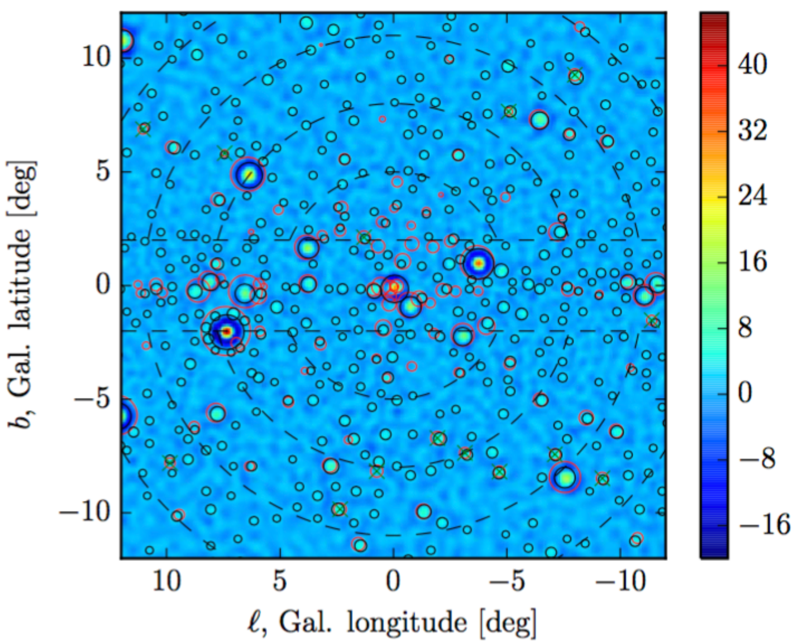
T. D. P. Edwards, F. Calore, and C. Weniger



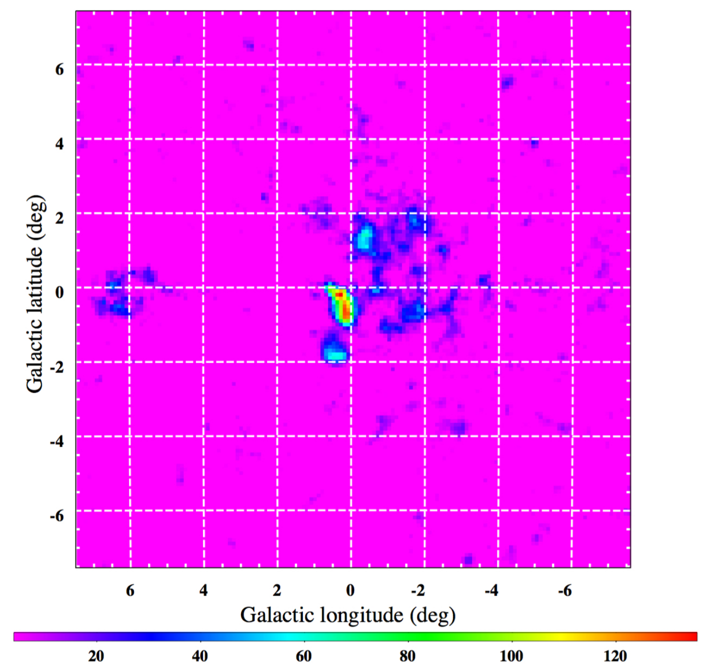
UNIVERSITY OF AMSTERDAM

GRAPPA
GRavitation AstroParticle Physics Amsterdam

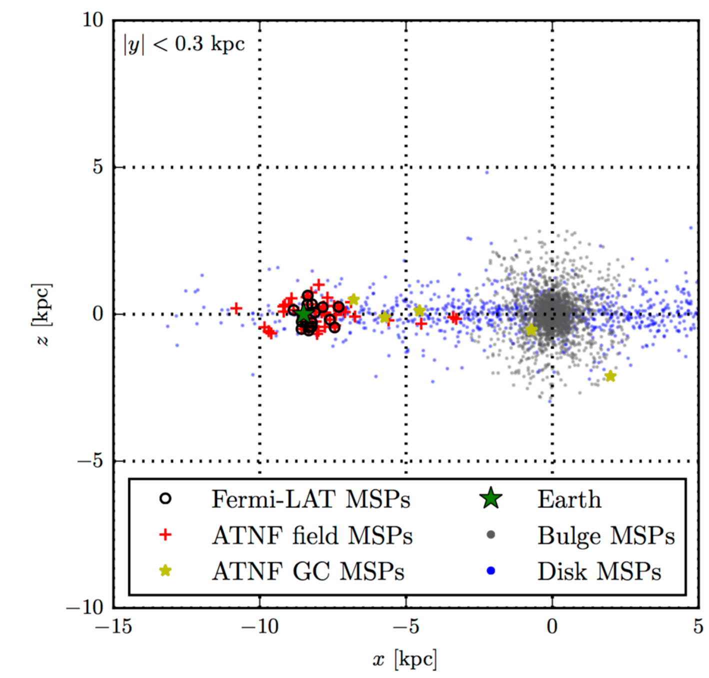
Why do we care?



Bartels et al. 2015



Fermi-LAT collaboration 2015



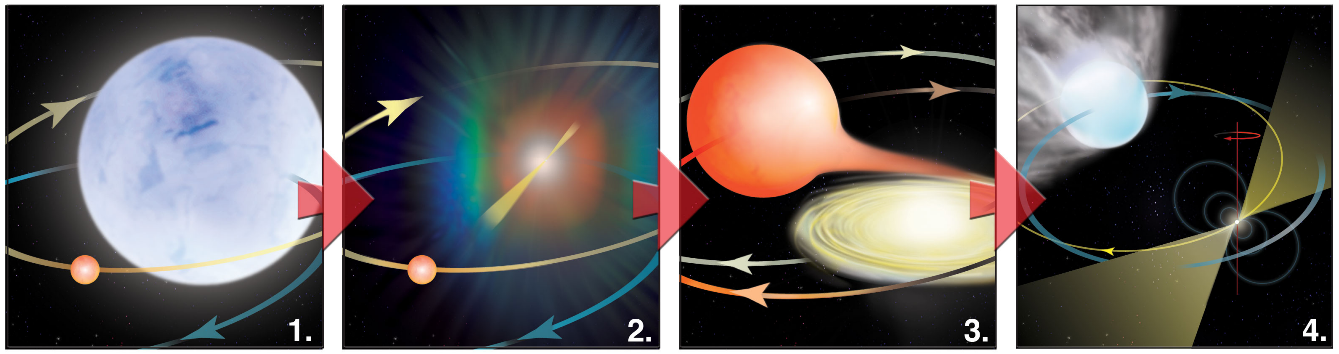
Calore et al. 2015

- The origin of the galactic centre GeV excess remains unknown
- Potential population of unknown point sources towards the galactic centre

- Realistic constraints for gamma ray point source models
- Probabilistic associations of unassociated objects to different sources classes - contribution to the millisecond pulsar interpretation of the GeV excess

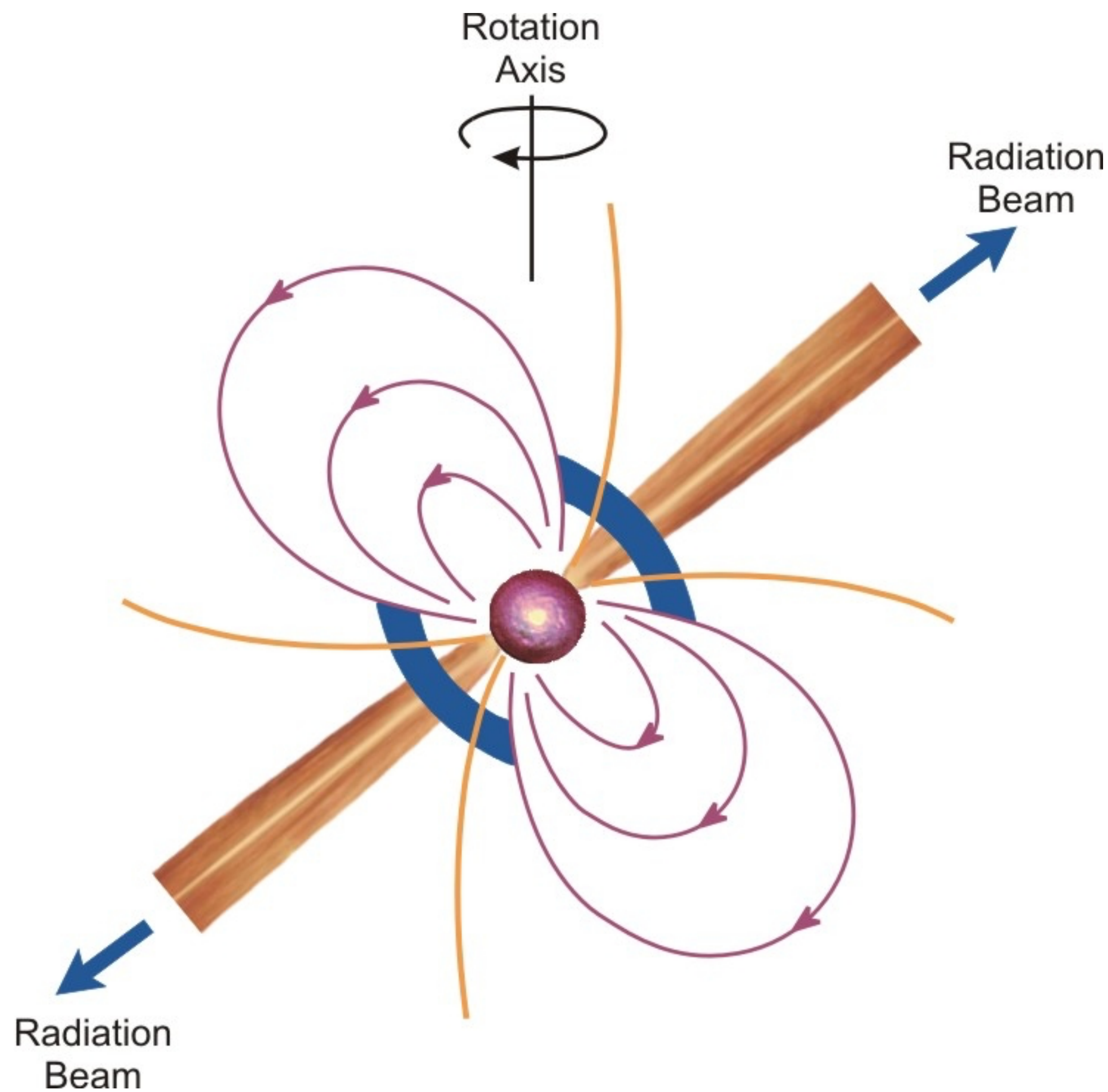
MSP (very short) Intro

$$\frac{dN}{dL} \propto L^{-\alpha}$$

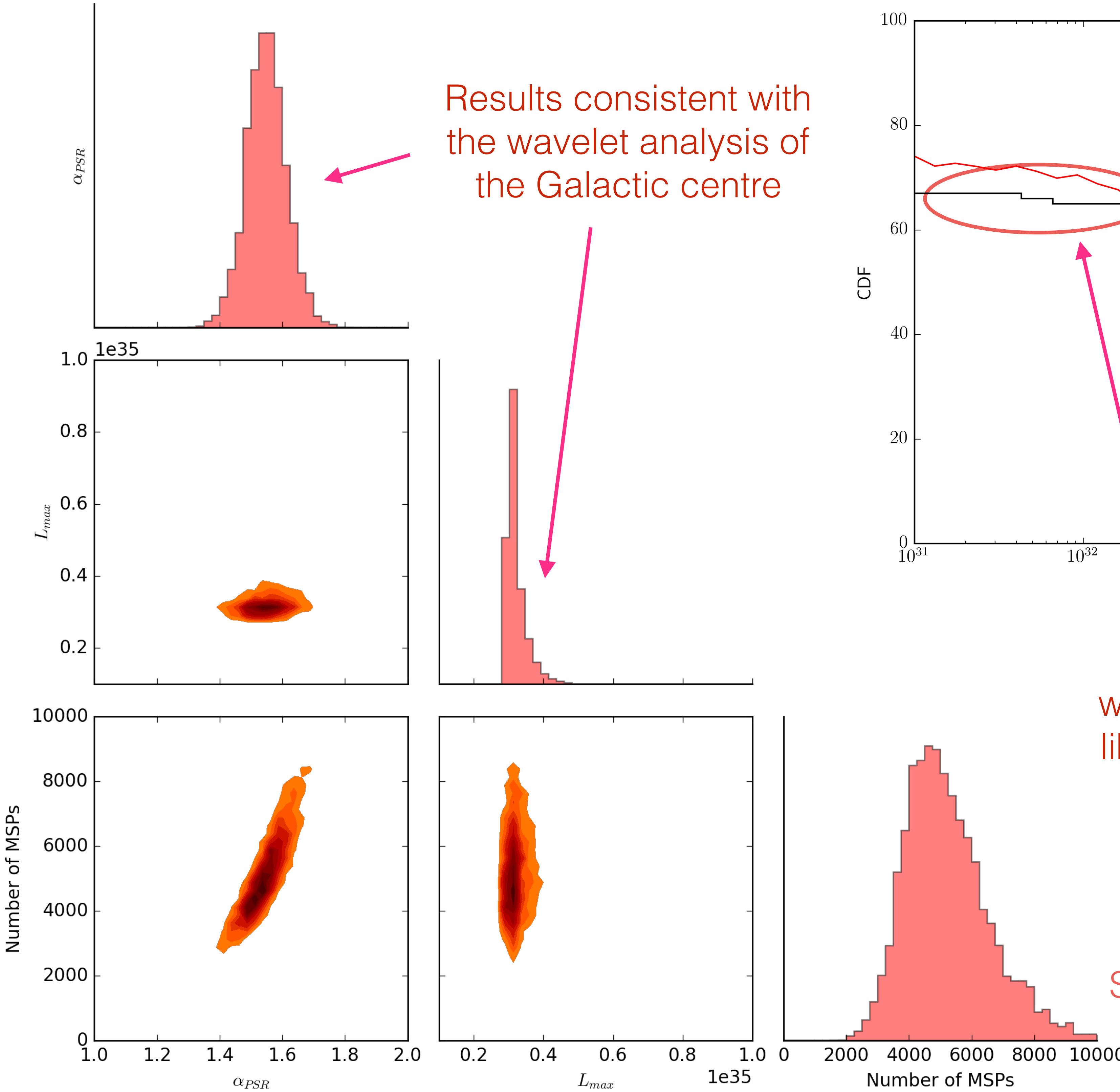


- MSP's are formed in binary pairs by accreting matter from the companion star
- They have very low spin down periods and can live for up to 10^{10} years

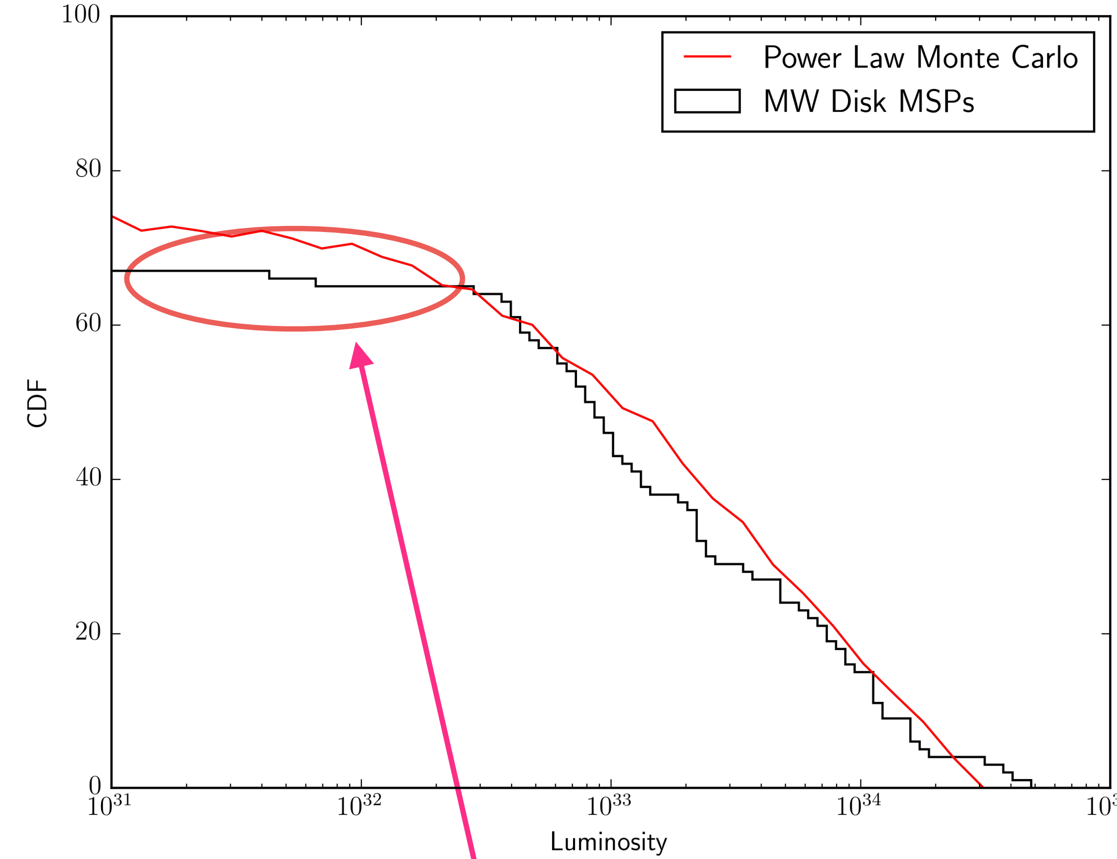
$$\mathcal{L} \propto e^{-\mu} \prod_{i=1}^{n} 4\pi D_i^4 L_i^{-\alpha} n_{\rm los} \Gamma_S$$



MSP Analysis



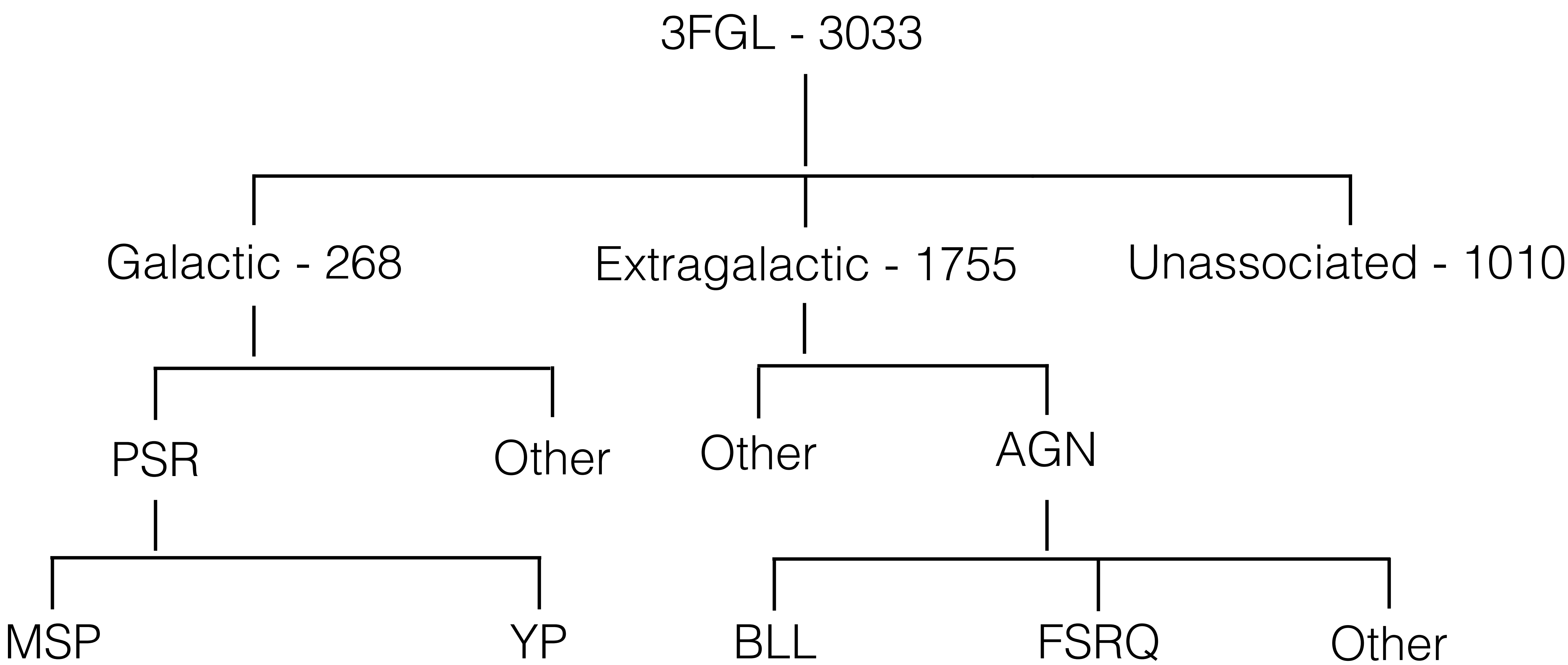
Results consistent with the wavelet analysis of the Galactic centre



The low luminosity regime is where unassociated sources are likely to play an important role. It is also important to constrain this regime as this may be where the bulge MSP's exist

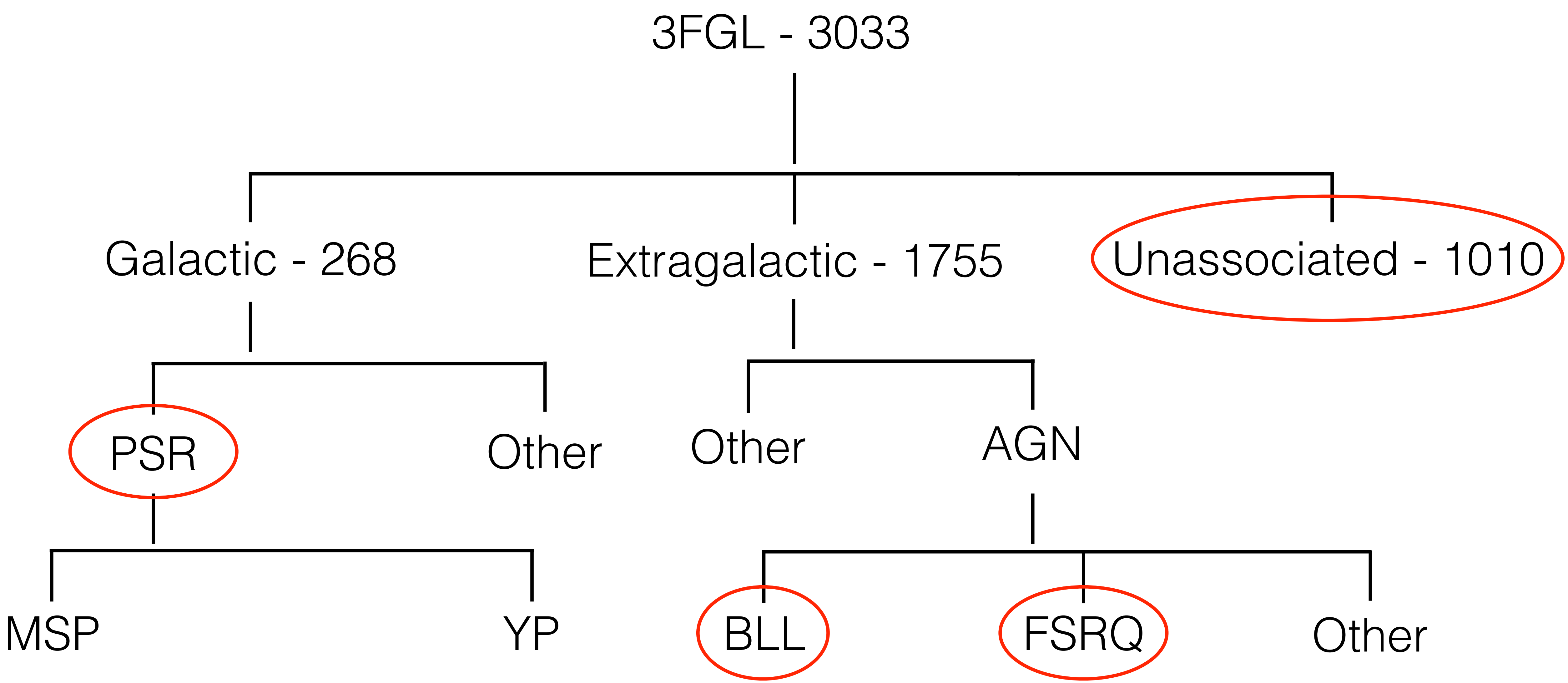
S. Liem 2016 (barrett plotting software)

3FGL Populations



- Capital letter classification means that they are confirmed in gamma rays
- Lower case classification is actually an association
  - This is done through spatial coincidence with a source known from another wavelength

3FGL Populations



- Capital letter classification means that they are confirmed in gamma rays
- Lower case classification is actually an association
  - This is done through spatial coincidence with a source known from another wavelength

Association Tables

Name	Objects[a]	Ref.
High $\dot{E}/d^2$ pulsars	213	Manchester et al. (2005)[b]
Other normal pulsars	1657	Manchester et al. (2005)[b]
Millisecond pulsars	137	Manchester et al. (2005)[b]
Pulsar wind nebulae	69	Collaboration internal
High-mass X-ray binaries	114	Liu et al. (2006)
Low-mass X-ray binaries	187	Liu et al. (2007)
Point-like SNR	157	Green (2009)
Extended SNR[†]	274	Green (2009)
O stars	378	Maíz-Apellániz et al. (2004)
WR stars	226	van der Hucht (2001)
LBV stars	35	Clark et al. (2005)
Open clusters	2140	Dias et al. (2002)
Globular clusters	160	Harris (1996)
Dwarf galaxies[†]	100	McConnachie (2012)
Nearby galaxies	276	Schmidt et al. (1993)
IRAS bright galaxies	82	Sanders et al. (2003)
BZCAT (Blazars)	3060	Massaro et al. (2009)
BL Lac	1371	Véron-Cetty & Véron (2010)
AGN	10066	Véron-Cetty & Véron (2010)
QSO	129,853	Véron-Cetty & Véron (2010)
Seyfert galaxies	27651	Véron-Cetty & Véron (2010)
Radio loud Seyfert galaxies	29	Collaboration internal
1WHSP	1000	Arsioli et al. (2014)
WISE blazar catalog	7855	D'Abrusco et al. (2014)
NRAO VLA Sky Survey (NVSS)[c]	1,773,484	Condon et al. (1998)
Sydney University Molonglo Sky Survey (SUMSS)[c]	211,050	Mauch et al. (2003)
Parkes-MIT-NRAO survey[c]	23277	Griffith & Wright (1993)
CGRaBS	1625	Healey et al. (2008)
CRATES	11499	Healey et al. (2007)
VLBA Calibrator Source List	5776	http://www.vlba.nrao.edu/astro/calib/vlbaCalib.txt
ATCA 20 GHz southern sky survey	5890	Murphy et al. (2010)
ATCA follow up of 2FGL unassociated sources	424	Petrov et al. (2013)
ROSAT All Sky Survey (RASS) Bright and Faint Source Catalogs[c]	124,735	Voges et al. (1999),[d]
58 months BAT catalog	1092	Baumgartner et al. (2010)
4^{th} IBIS catalog	723	Bird et al. (2010)
1st AGILE catalog*	47	Pittori et al. (2009)
3rd EGRET catalog*	271	Hartman et al. (1999)
EGR catalog*	189	Casandjian & Grenier (2008)
0FGL list*	205	Abdo et al. (2009d, 0FGL)
1FGL catalog*	1451	Abdo et al. (2010d, 1FGL)
2FGL catalog*	1873	Nolan et al. (2012, 2FGL)
1FHL catalog*	514	Ackermann et al. (2013a, 1FHL)
TeV point-like source catalog*	82	http://tevcat.uchicago.edu/
TeV extended source catalog[†]	66	http://tevcat.uchicago.edu/
LAT pulsars	147	Collaboration internal
LAT identified	137	Collaboration internal

Extremely thorough list of complementary catalogs at other wavelengths. Association is made if the spatial coincidence probability posterior peak reaches above 80% - We assume these associations are concrete

Fermi-LAT collaboration 2015

Theory: Inference

Model parameters

$$P(\theta, \vec{k}|D) = \frac{\prod_i \mathcal{L}(D|d,\theta)P(k)P(\theta)}{P(D)}$$

Source class vector

$$\vec{k} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$$\sum_{\vec{k}} P(\theta, \vec{k}|D) = P(\theta|D) = \frac{\sum_{\vec{k}} \prod_i \mathcal{L}(D|d,\theta)P(k)P(\theta)}{P(D)}$$

$$= \frac{\prod_i \sum_k \mathcal{L}(D|d,\theta)P(k)P(\theta)}{P(D)}$$

(blue) = Associated object

(green) = Unassociated object

Theory: Association

$$P(k_i|D) \propto \int d\theta \sum_{\substack{\vec{k} \\ k_i \text{ fixed}}} \prod_j \mathcal{L}_j(D_j|k_j,\theta) P_j(k_j) P(\theta)$$

$$P(k_i|D) \propto \int d\theta \mathcal{L}_i(D_i|k_i,\theta) P_i(k_i) \underbrace{\sum_{\substack{\vec{k} \\ k_i \text{ fixed}}} \prod_{j \neq i} \mathcal{L}_j(D_j|k_j,\theta) P_j(k_j) P(\theta)}_{\approx \sum_{\vec{k}} \prod_j \mathcal{L}_j(D_j|k_j,\theta) P_j(k_j) P(\theta)}$$

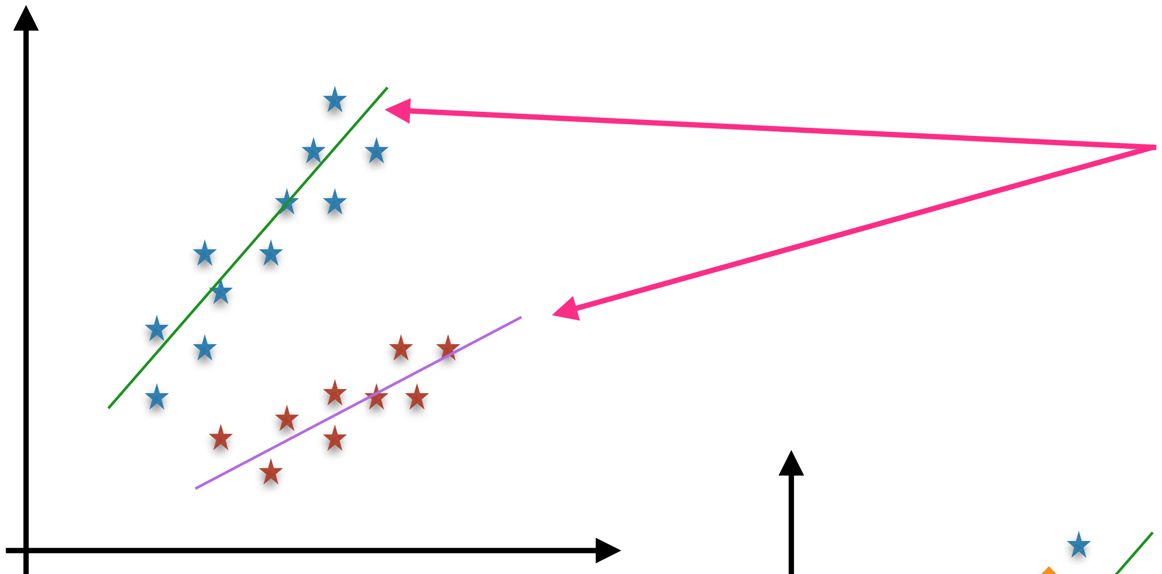
Approximation: Fit with n-1 sources is well approximated by a fit to the entire data set

Posterior for the entire data set

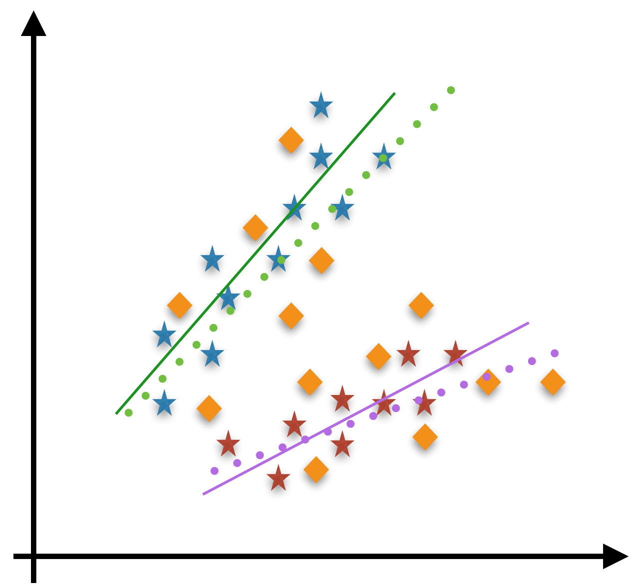
GRAPPA
GRavitation AstroParticle Physics Amsterdam

Cartoon Example

Standard fitting routines only take into account the data known to be part of a particular class. In this example we fit two linear models to some data

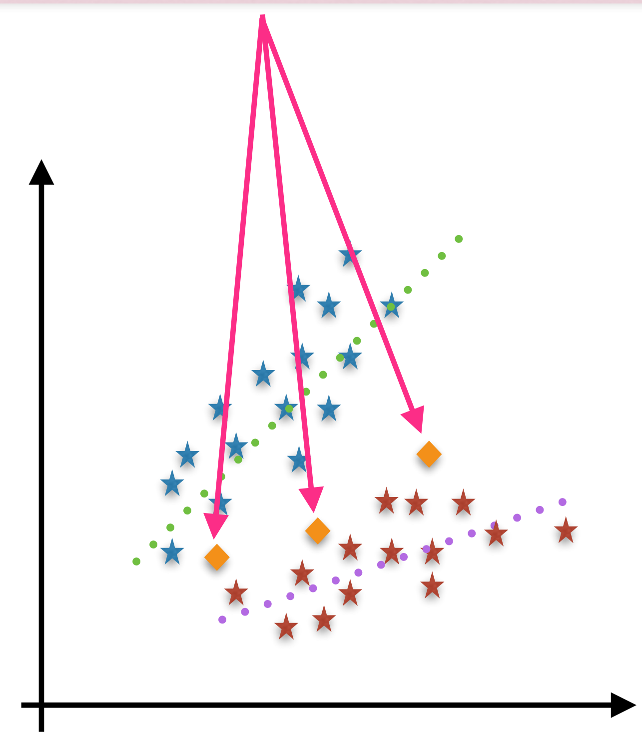


Now we can add unassociated sources



Probabilistic statements about the remaining sources

Perform Association procedure



Likelihood Function

$$\mathcal{L} = \frac{dN}{dS}\frac{dN}{dE}\Gamma(S|l,b)VC_sS(l,b)$$

Instead of latitude cuts we use full galactic spatial distributions to build probability density functions for galactic coordinates l and b on the sky

Additionally we compare the integrated flux in different energy bins provided by the 3FGL meaning that the spectra normalisations are not free parameters but derived from the energy flux S we get for each source

(Broken) Power laws for the AGN source count distributions. For pulsars we marginalise over the distance and luminosity to construct the directional independent source count distribution

Parameter	Min.	Median	Max.
Spectral_Index	0.5	2.2	3.1
Variability_Index[b]	3.0	4.0	11.0
Flux_Density[c]	-35.4	-28.2	-19.9
Unc_Energy_Flux100[d]	-28.5	-27.6	-24.8
Signif_Curve[b]	-5.8	0.4	4.4
hr_{12}	-1	-0.1	1
hr_{23}	-1	-0.1	1
hr_{34}	-1	-0.2	1
hr_{45}	-1	-0.3	1

[a]For the YNG vs MSP models we also used the Galactic latitude (GLAT) of the source, as a predictor parameter.

[b]Number represents the log of the original value contained in the catalog.

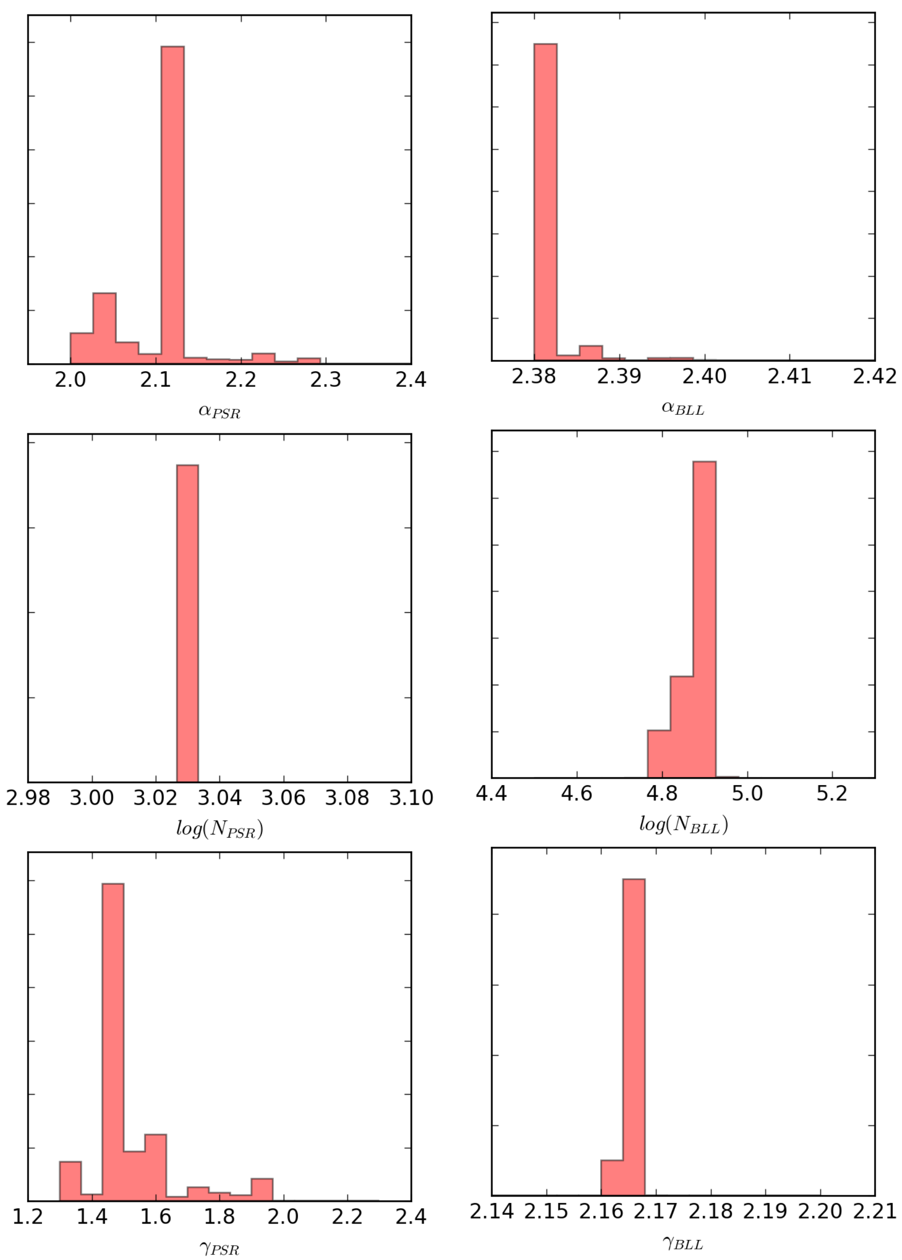
[c]In photon cm^{-2} MeV^{-1} s^{-1} (log of the original value contained in the catalog).

[d]In erg cm^{-2} s^{-1} (log of the original value contained in the catalog).

Saz Parkinson et al. 2016

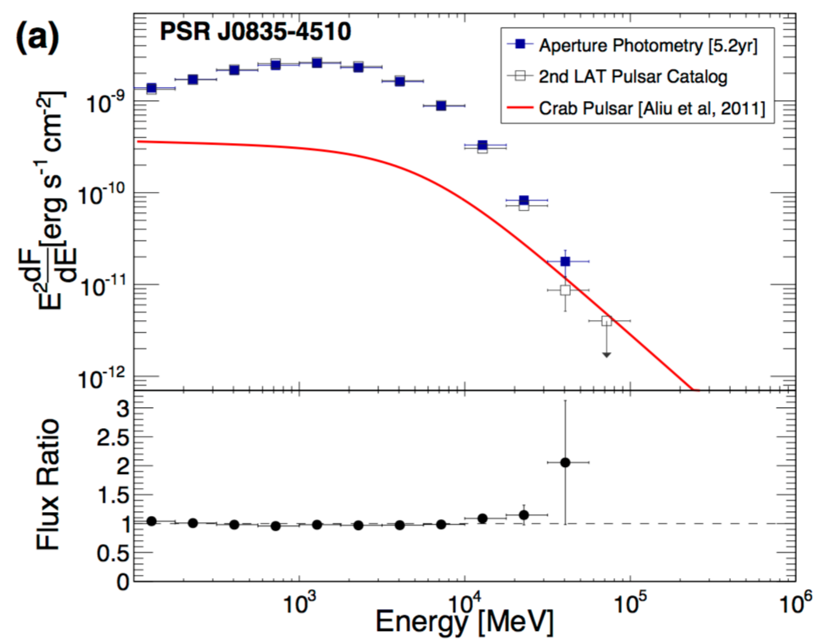
$$\frac{dN}{dS} = \int dL dD \delta\left(S - \frac{L}{4\pi D^2}\right)\frac{dN}{dL}P(D|l,b)$$

Preliminary Results

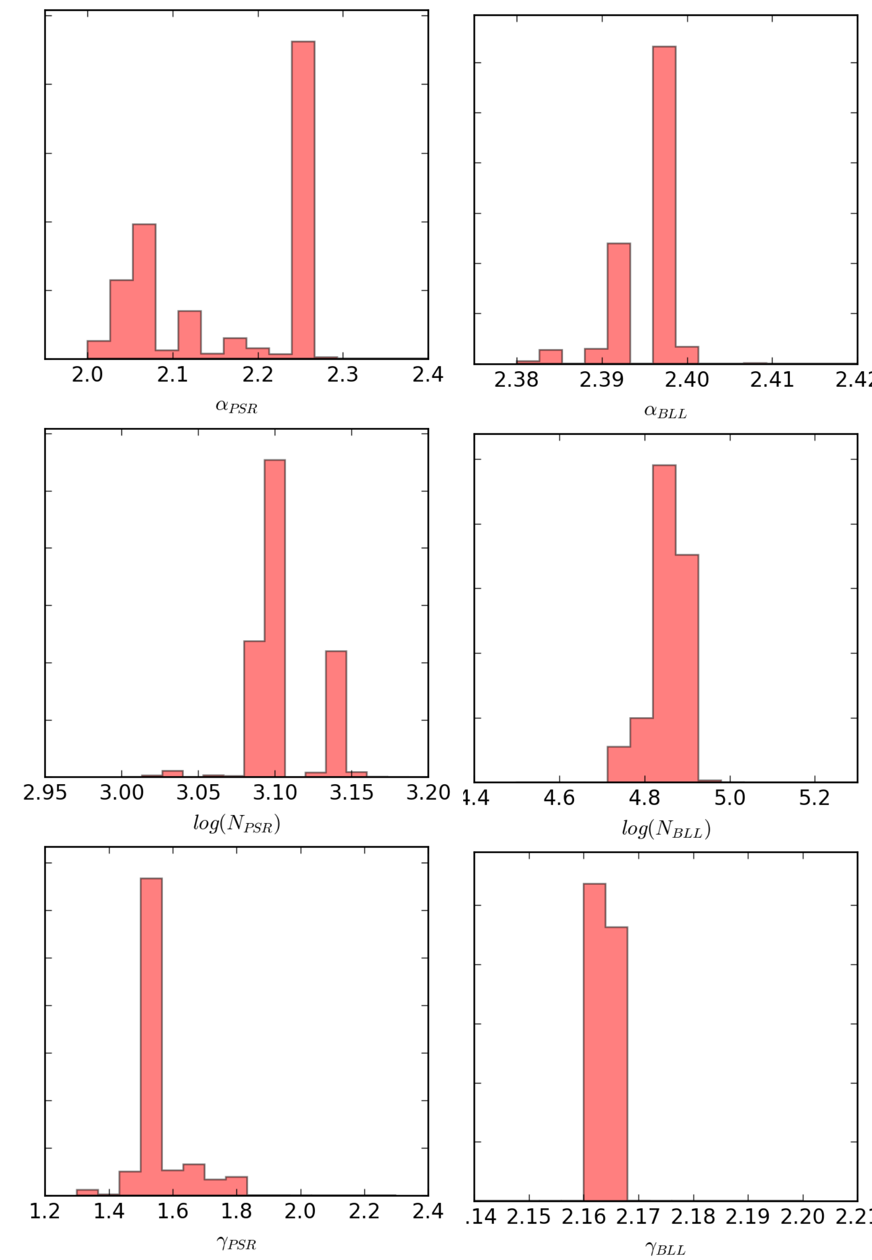


Now we can add unassociated sources

Posteriors tend to become broader but roughly keep the best fit value



A. McCann 2015



Good and Bad news

The associations are self consistent and associate known sources to the correct class in all preliminary tests

Still validating the accuracy of the method. Next steps will be applying to unassociated sources

Source Name	Association	PSR Probability	BLL Probability
3FGL J0002.2-4152	bll	2.84989433905e-58	1.0
3FGL J0008.6-2340	bll	2.43544493636e-63	1.0
3FGL J0009.6-3211	bll	1.11973314841e-57	1.0
3FGL J0010.5-1425	bll	8.497178032e-64	1.0
3FGL J0013.2-3954	bll	6.74432806546e-70	1.0
3FGL J0013.9-1853	bll	2.65181900573e-45	1.0
3FGL J1136.9+2551	PSR	1.0	7.32996794332e-21
3FGL J0022.1-1855	bll	9.92644557986e-57	1.0
3FGL J1100.5+4020	PSR	1.0	1.04224954955e-17

Advantages of the method

Posterior predictive distributions

- Ability to take into account unassociated sources into likelihood scans
- Self consistent probability estimates for source class associations
- Uses models which can be derived from theory or a fit to alternative data sets
- Fit to multiple populations simultaneously

Random Forest (ML)

- Forms decision trees based off the training data i.e. can only be trained on associated sources
- Probabilities can be estimated but loss of interpretability when forming a decision forest
- Only uses the data to form the optimal decision cuts
- Provides only associations - no fits

Summary and Future Work

- The 3FGL presents the largest sample of point sources in the gamma ray sky. By using only the associated sources we are ignoring a wealth of data that are likely to contribute to different source classes. We present here a systematic way to take into account these unassociated sources
- As a side benefit the statistical technique we are able to create probabilistic associations of the unassociated sources to different classes
- Our method is complementary to previous attempts to use machine learning classifiers (random forests) to assign classes to objects

- With association of high latitude sources we can dissect the composition of the high latitude gamma-ray emission (IGRB) through population synthesis studies, also combining with recent results from 1point-pdf of gamma-ray counts
- Associations in the disk: study contribution from "new" source populations such as bulge MSP's
- Reducing the number of unassociated sources might improve significantly the limit on DM annihilation from DM subhalos