Statistical Applications in Dark Matter Searches

Jan Conrad
Oskar Klein Centre
Physics Department.
Stockholm University

ISAPP school,
Djurönäset, 2013

Overview

- **Hypothesis testing**
  - Look elsewhere effect → covered in particle physics section
  - Wilks theorem /Chernoffs theorem
  - **At the edge:** Separate families of hypotheses
- **Interval estimation**
  - Profile likelihood in Dark Matter searches
  - Bayesian methods in Dark Matter searches.
    - Global fit in Supersymmetry
    - Dark matter density modelling.

Issues in DM searches – overview
From raw data to physics

Instrumental background

Irreducible background

(or signal to some)

Raw detector output

digital signal

Hypothesis testing (remove instrumental background)

Multivariate (MV) classification, machine learning

Parameter estimation: (derive physical observables)

Least-squares, likelihood, more seldom (ever?) machine learning (MV **regression)**

Issues in DM searches – overview
From data to Dark Matter model

Irreducible background

Hypothetical signal

i.e. Dark matter

Physical observables

(energy, direction)

Hypothesis testing (establish Dark Matter signal)

Maximum Likelihood Ratio test

Parameter estimation/interval estimation

Profile likelihood, Neyman construction, Posterior integration

Home-made methods

Basic (unbinned) likelihoods (no nuisance parameters)

→ **Marked Poisson**

- **Indirect detection**

Parameters of interest

$$L(H;obs) = Pois(n, \sigma v) \prod_{i=1}^{n} f(E_i \mid m_{DM}, Y(E)) \prod_{i=1}^{n} f(\vec{\varphi} \mid g(\varphi_{true}))$$

- **Direct detection**

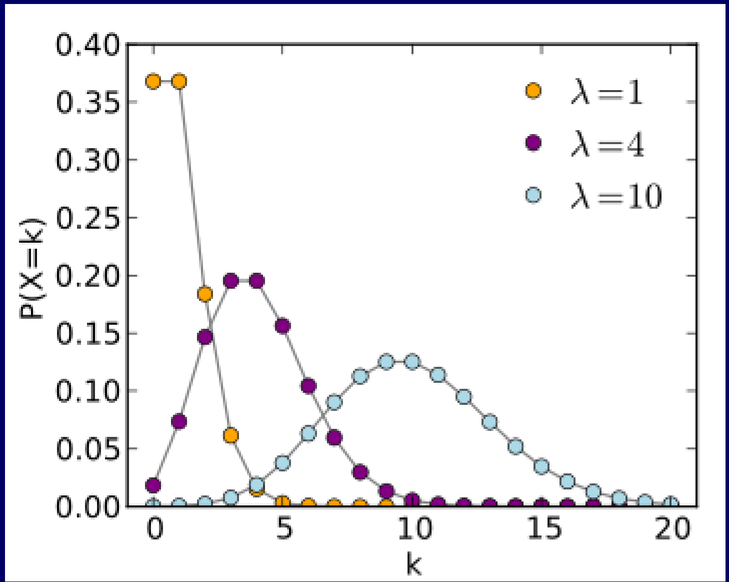
$$L(H;obs) = Pois(n, \sigma_{scat}^{SD,SI}) \prod_{i=1}^{n} f(E_{recoil} \mid m_{DM})$$

Parameters of interest

Do we need exact methods or will asymptotics work?

- **Exact methods: methods that do not rely on the large sample approximation (e.g. Neyman construction).**

- **Generically, for n > 10 or so, asymptotic method will work reasonably.**

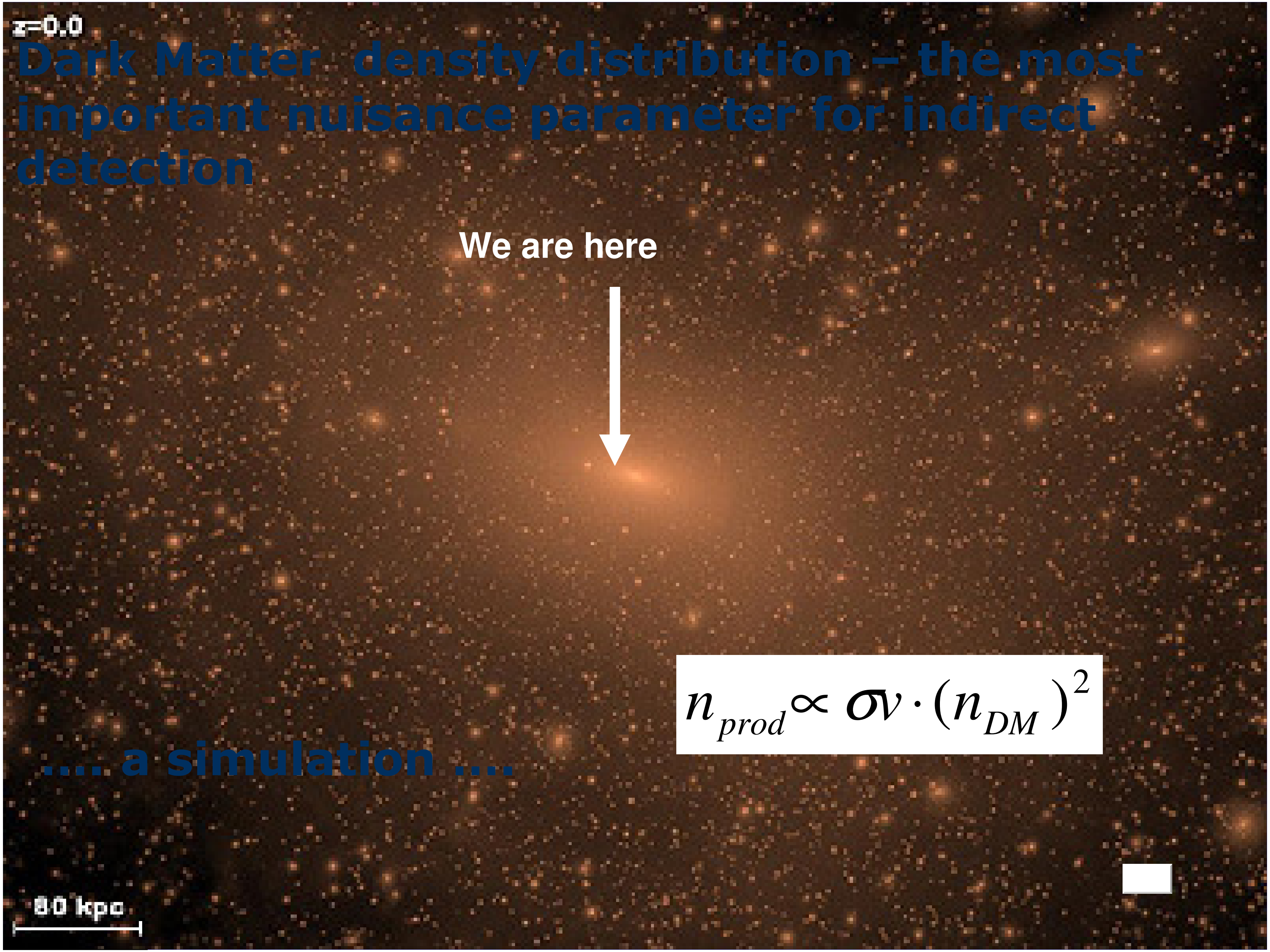


Nuisance parameters:

parameters that need to be estimated which are not of prime interest but which will affect inference on the parameters of interest.

e.g. background expectation in estimate of a signal rate.

Dark Matter density distribution – the most important nuisance parameter for indirect detection



$$n_{prod} \propto \sigma v \cdot (n_{DM})^2$$

.... a simulation ....

Nuisance parameters

Indirect detection:

$$n \sim Pois\left(n \mid \sigma v \cdot n_{DM}{}^{2}, b\right)$$

100-1000% uncertainty related to Dark Matter distribution, degenerated with parameter of interest

~20-30% uncertainty related to irreducible background (or much more, if including unknwon background sources)

N.B: due to large statistics, even modest uncertainties in nuisance parameters can dominate!

Nuisance parameters

Direct detection:

$$n \sim Pois\left(n \mid \mu(\sigma_{scatt}, astro)\varepsilon_s, \varepsilon_b b\right)$$

200-1000% uncertainty related to detector efficiency and nuclear interaction

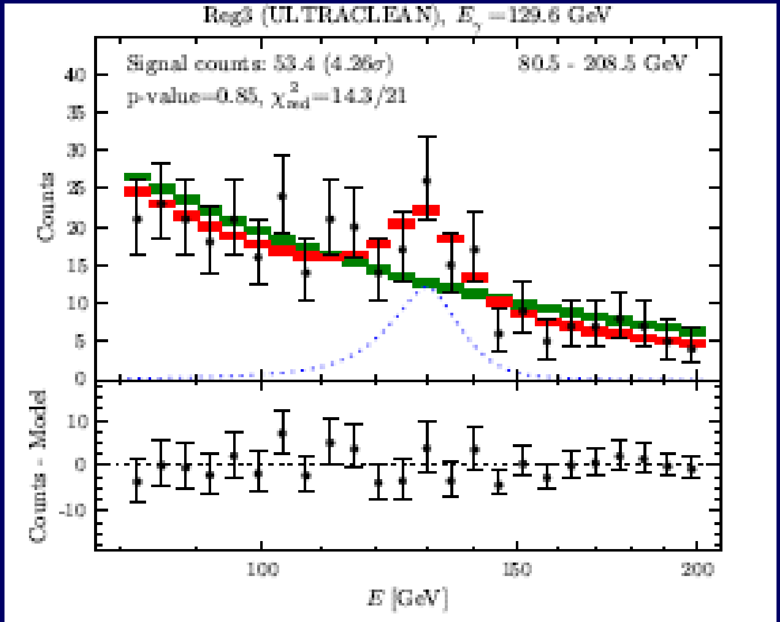
20-100% uncertainty related to astrophysics

20-100%

Worst cases for low mass WIMPs

Hypothesis testing

Applicability of Wilks/Chernoff's theorem



C. Weniger, arXiv:1204.2797

Trial factor treatment?

Separate families of hypothesis

(if compared to other features

(broken power-law)

Profumo+,arXiv:1204.6047

Hypothesis tests.

- **The goal of hypothesis tests is to test if a hypothesis is compatible to data when compared with a defined (or undefined, in this case: ”goodness-of fit”) alternative.**

- **Conventionally, one hypothesis is called the null hypothesis (H_0), the other one is called the alternative hypothesis (H_1).**

- **The test proceeds by finding the** critical region, **i.e. the region in data space for which:**

$$p(\vec{d} \in w \mid H_0) = \alpha$$

significance

- **and at the same time maximize:**

power

$$p(\vec{d} \in w \mid H_1) = 1 - \beta$$

(Maximum) likelihood ratio test statistic

- **The likelihood ratio test is the generalization of the Neyman-Pearson test to composite hypotheses.**
- **Composite hypotheses:**

$$H_0 : \sigma v \in \omega$$
$$H_1 : \sigma v \in \Theta - \omega$$

A special case would be, eg. ω= 0

- **In this situation you define the likelihood ratio:**

$$\lambda = \frac{\max_{\sigma v \in \omega} L(n \mid \sigma v)}{\max_{\sigma v \in \Theta} L(n \mid \sigma v)}$$

Null distribution: Wilks theorem

- **In hypothesis tests, traditionally, a very important property is the knowledge of the null distribution.**
- **If H_0 imposes r constraints on s+r parameters in H_1 and H_0 then under H_0. for n→infty**

$$-2\ln\lambda \sim \chi^2(r)$$

- **In this example, the hypothesis are nested. The asymptotic property for nested hypotheses is called** Wilks theorem**.**
- **Simplest example: new physics signal over known background.**

H_1: s+b, H_0: s=0.

Wilks theorem (1938)

Theorem: If a population with a variate x is distributed according to the probability function $f(x, \theta_1, \theta_2 \cdots \theta_h)$, such that optimum estimates $\bar{\theta}_i$ of the θ_i exist which are distributed in large samples according to (3), then when the hypothesis H is true that $\theta_i = \theta_{0i}$, $i = m + 1, m + 2, \cdots h$, the distribution of $-2 \log \lambda$, where λ is given by (2) is, except for terms of order $1/\sqrt{n}$, distributed like χ^2 with $h - m$ degrees of freedom.

Chernoff's theorem (1952)

placed by the inverse of the information matrix. In particular, if one tests whether θ is on one side or the other of a smooth $(k - 1)$-dimensional surface in k-dimensional space and θ lies on the surface, the asymptotic distribution of λ is that of a chance variable which is zero half the time and which behaves like χ^2 with one degree of freedom the other half of the time.

$$= 1/2\ \delta\ (0) + ½\ \chi^2$$

This is a common situation, e.g. if you constrain the fit to be s>0.

Applicability of Wilks/Chernoffs theorem.

- **Regularity conditions (likelihood function differentiable)**
- **Optimal (unbiased and efficient) estimator**
- **Nested hypotheses**
- **Asymptotic (distribution of estimates approximately Gaussian).**

Wilks theorem/Chernoffs theorem: an example

- **Search for Dark Matter satellites.**

  Fermi-LAT: Astrophys.J. 747 (2012) 121

- **Binned Poisson likelihood**
- **Spectral model: fit parameters**
- **Spatial model: PS/extended source**
- **Nuisance parameters: background (profile likelihood)**

Likelihood unbinned, spectral part:

$$L(H;obs) = Pois(n,\mu)\prod_{i=1}^{n} f(E_i \mid \gamma)$$

$$H_{pwl} : f(E_i \mid \gamma) \propto E_i^{-\gamma}$$

Conventional astrophysical source

$$H_{b\bar{b}} : f(E_i \mid \gamma) \propto \exp\left(-\frac{E_i}{m_{WIMP}}\right)$$

e.g. DM particle mass

Applicability of Wilks/Chernoffs theorem.

- **Regularity conditions → YES**
- **Optimal estimator → YES (MLE)**
- **Asymptotic → YES**
- **Nested hypotheses → NO**

How is this problem addressed in the paper?

$$\mathrm{TS}_{\mathrm{spec}} = -2\ln\left(\frac{\mathcal{L}(H_{\mathrm{pwl}})}{\mathcal{L}(H_{b\bar{b}})}\right)$$
$$= \mathrm{TS}_{b\bar{b}} - \mathrm{TS}_{\mathrm{pwl}}$$

Nested

(b/s+b)

Quantities calculated by the Fermi-LAT software

”These two hypotheses are not nested, and thus the significance of this test was evaluated with simulations.”

Spatial part?

$$\mathrm{TS_{ext}} = -2\ln\left(\frac{\mathcal{L}(H_{\mathrm{point}})}{\mathcal{L}(H_{\mathrm{NFW}})}\right)$$
$$= \mathrm{TS_{NFW}} - \mathrm{TS_{point}}$$

Nested (b/s+b)

"While the point and extended hypotheses are nested and TS_{ext} is cast as a likelihood ratio test, **it is unclear whether** this analysis satisfied all of the suitable conditions for the application the theorems of Wilks (Wilks 1938) or Chernoff (Chernoff 1954). **Therefore, we relied on simulations** ……"

… and indeed, it is not chi-squared …

within the ROI as free parameters in the fit. After refitting, only candidate sources with TS > 24 were accepted into the list of source candidates[4]. Finally, to avoid duplicating sources in the 1FGL,

- **TS>24 ~ 4.9 sigma detection if chi-squared, here: ~ 3 sigma´:**

[4]Monte Carlo simulations have shown that 1 in 10^4 background fluctuations will be detected at TS $\geq$ 24

Construct a comprehensive family

… see F. James "Statistical Methods in Experimental Physics"

$$h(x;\gamma_1,\gamma_2,\theta) = (1-\theta)g + \theta f$$

$H_0 : \theta = 0$	γ_1, γ_2	unspecified
$H_1 : \theta \neq 0$	γ_1, γ_2	unspecified

Cox 1961,1962 refers to this as "conventional method",

Cox proposes:

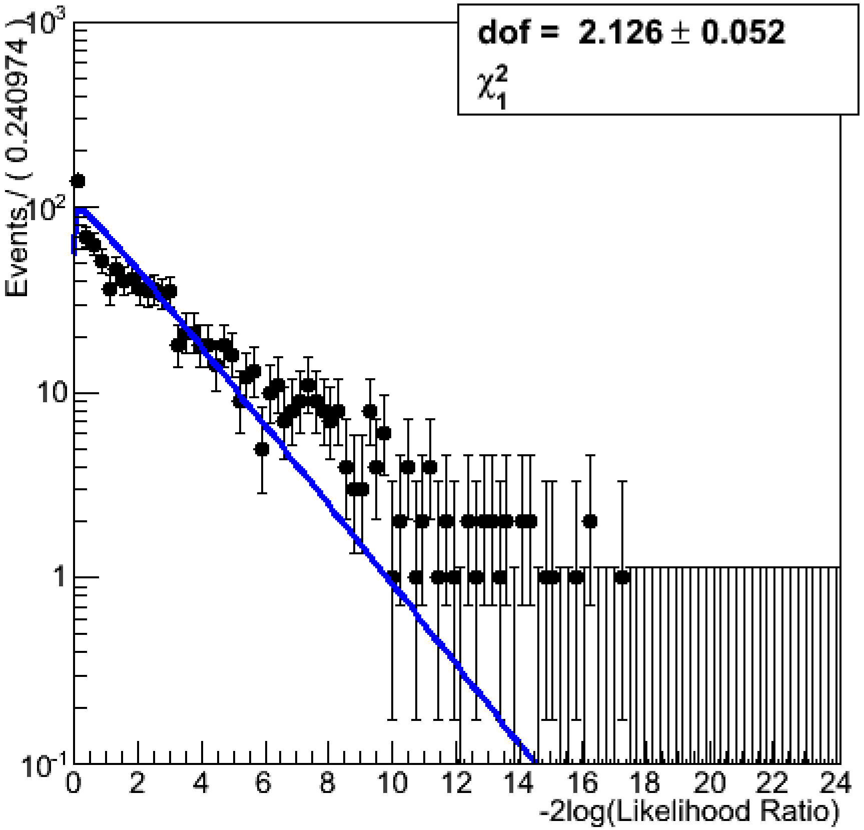
$$h(x;\psi,\gamma,\theta) = g(x;\gamma_1)^{\theta} f(x;\gamma_2)^{1-\theta}$$

… see also Atkinson (1970) and Quandt (1974) for a comparative disc.

Binned – 20 bins

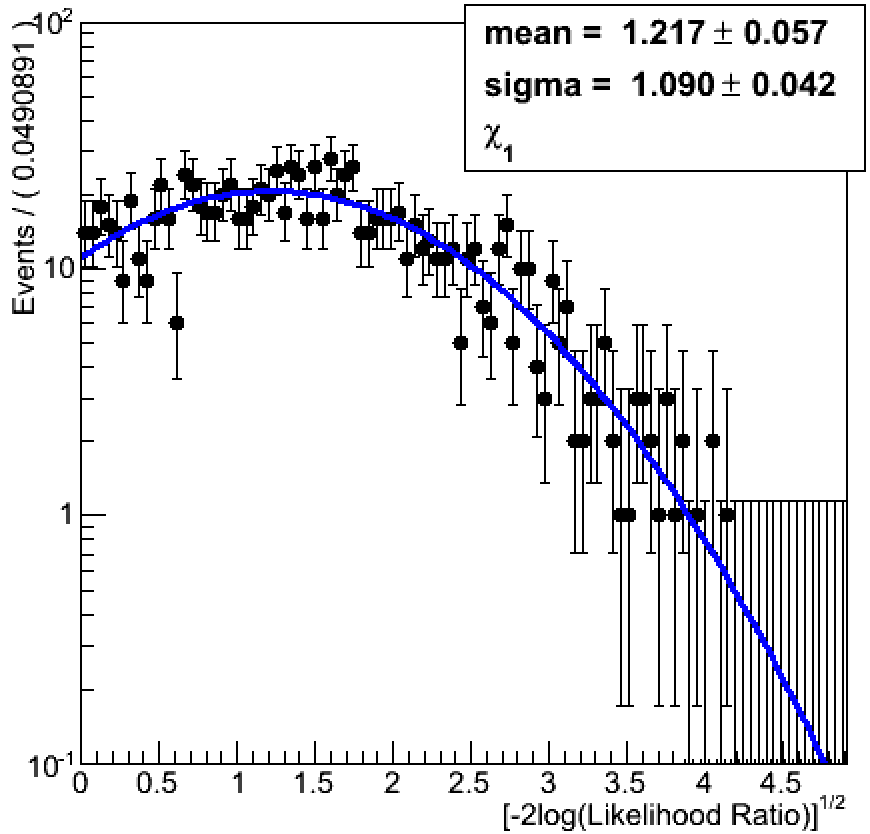
JC, H. Dickinson, 2012

20 logarithmic bins, Extended Fit



Preliminary

20 logarithmic bins, Extended Fit



Separate families of hypotheses: work in progress

- Quandt (1970) can not find chi-squared (for Gaussian variates). We do find a chi2 for special cases.

- The direct likelihood ratio (also discussed by Cox) is conservative?

- If the above approach is applicable to the Fermi-LAT case is still to be seen. Probably not for the present analysis (binned).

- In two hours of googling I went from 1961 to 1974, still 40 years of statistics research to explore ….

Interval estimation

- With nuisance parameters:
    - marginalisation
    - profiling

Nuisance parameters in interval estimation

- Consider again the case ”s”, a signal rate that you’d like to constrain
- Usually, the background ”b” will be estimated from an independent measurement, i.e. the background estimate will come with an uncertainty
- We could present constraints on ”s” and ”b” in a 2 dimensional confidence interval, but as we are not interested in ”b” we’d rather find a way to project the confidence contour on ”s” subspace.
- This is what will be the subject of the next few slides.

The Profile likelihood

→ Direct detection: Xenon-100, **Phys.Rev. D84 (2011) 052003)**

→ Indirect detection: Fermi-LAT: **Phys.Rev.Lett. 107 (2011) 241302**

→ ACT: Rolke, Lopez, JC. **Nucl.Instrum.Meth. A551 (2005) 493-503** .

$$\lambda(\sigma v) = \frac{\max_{b|\sigma v} L(\sigma v, b)}{\max_{\sigma v, b} L(\sigma v, b)}$$

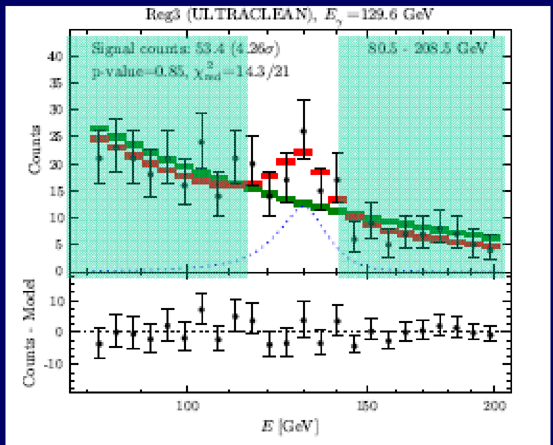
See also Cowan+, 1007.1727

Example (-2 ln L +1/2 rule, mentioned by Torsten),

- **e.g. ON-source, off-source measurement in Air Cherenkov Telescopes**

- **Sideband estimates of background in line searches (or control region estimates in e.g. Higgs search).**

$$f(x, y|\mu, b) = \frac{(\mu + b)^x}{x!} e^{-(\mu+b)} \cdot \frac{(\tau b)^y}{y!} e^{-\tau b}$$



Ratio between size of signal region and

size of background region

Example from Rolke+

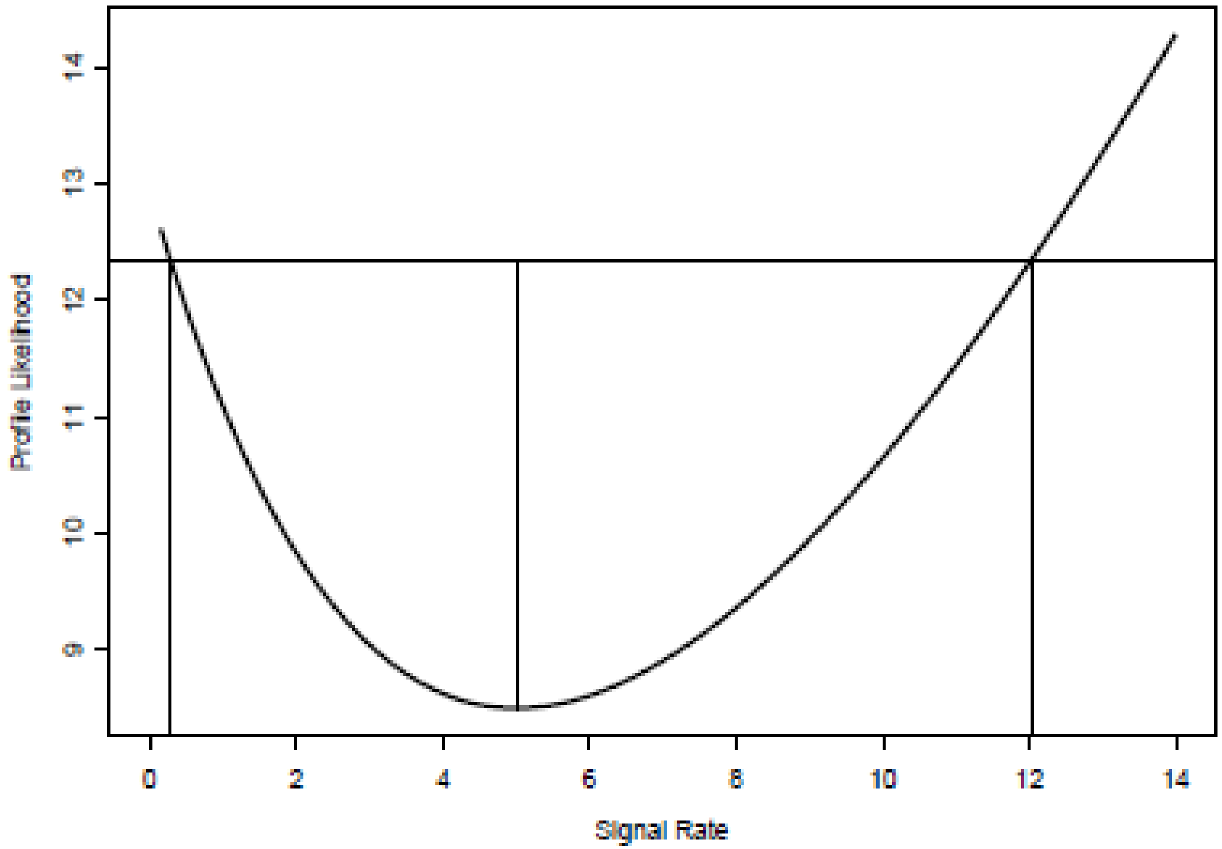


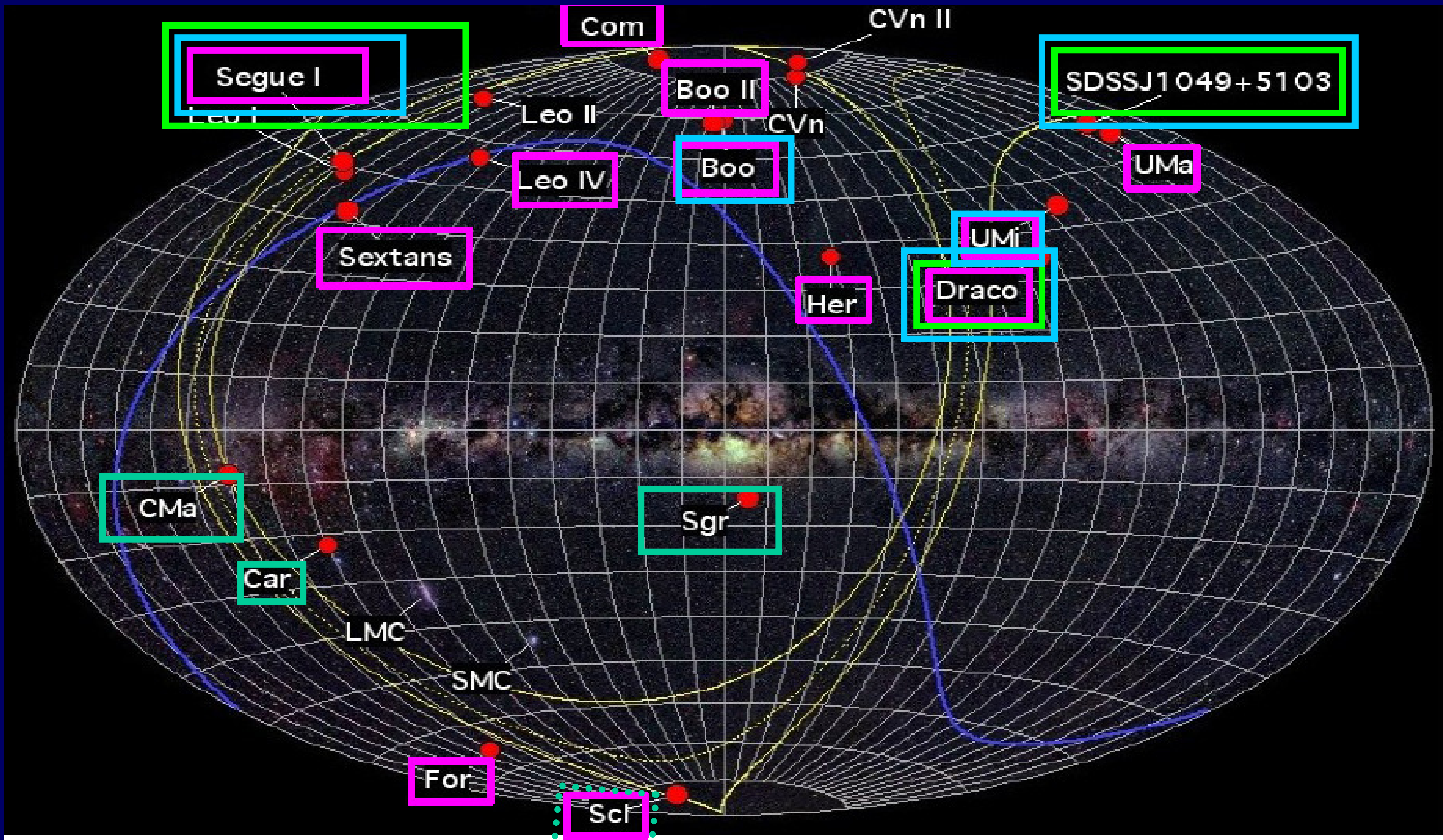
Figure 1: Example of the $-2\log\lambda$ curve. This is the case $x = 8$, $y = 15$ and $\tau = 5.0$. We find the 95% confidence interval to be $(0.28, 12.02)$.

Use in Dark matter searches?

- If you use a likelihood fit (or chi-2 fit) to obtain parameters and errors you might have used this without knowing.
- Background uncertainties have routinely been treated with profile likelihood
  - See 1205.6474 (galactic diffuse gamma-rays -- Torsten showed this constraint) for a complicated example with the background depending on 20 physics parameters, non-linearly correlated with the measurement.
- Dark Matter density → fix to benchmark values until ….. next example

Profile likelihood for DM density uncertainties: Fermi-LAT

Fermi H.E.S.S. MAGIC Veritas



Two novel approaches: likelihood combination and profiling.

Phys.Rev.Lett. 107 (2011) 241302

Universal source parameters

Product over sources

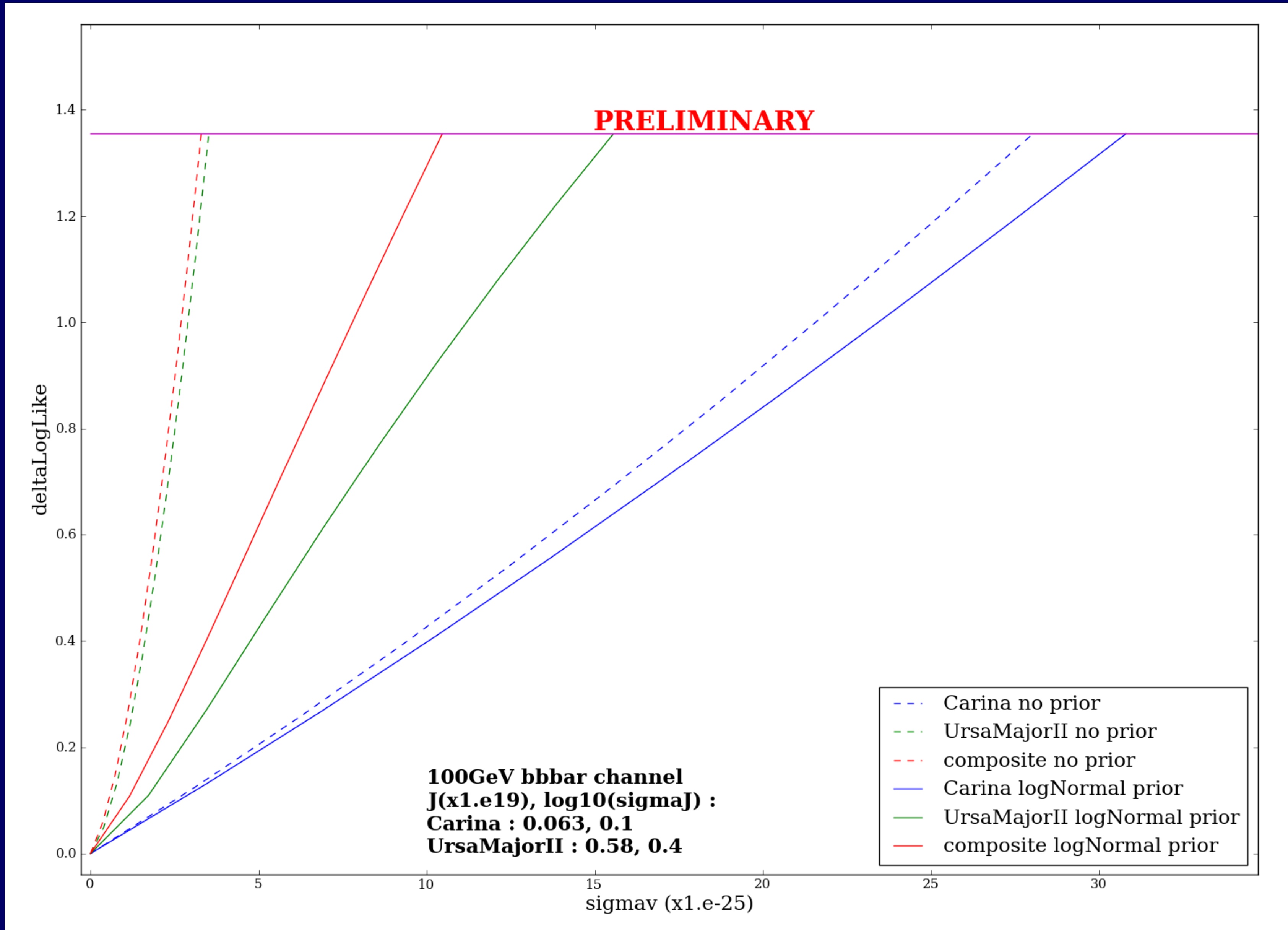
Individual source parameters

$$L(D|\mathbf{p}_W,\{\mathbf{p}\}_i) = \prod_i L_i^{\mathrm{LAT}}(D|\mathbf{p}_W,\mathbf{p}_i) \times \frac{1}{\ln(10)\, J_i \sqrt{2\pi}\sigma_i} e^{-\left[\log_{10}(J_i)-\overline{\log_{10}(J_i)}\right]^2/2\sigma_i^2}$$

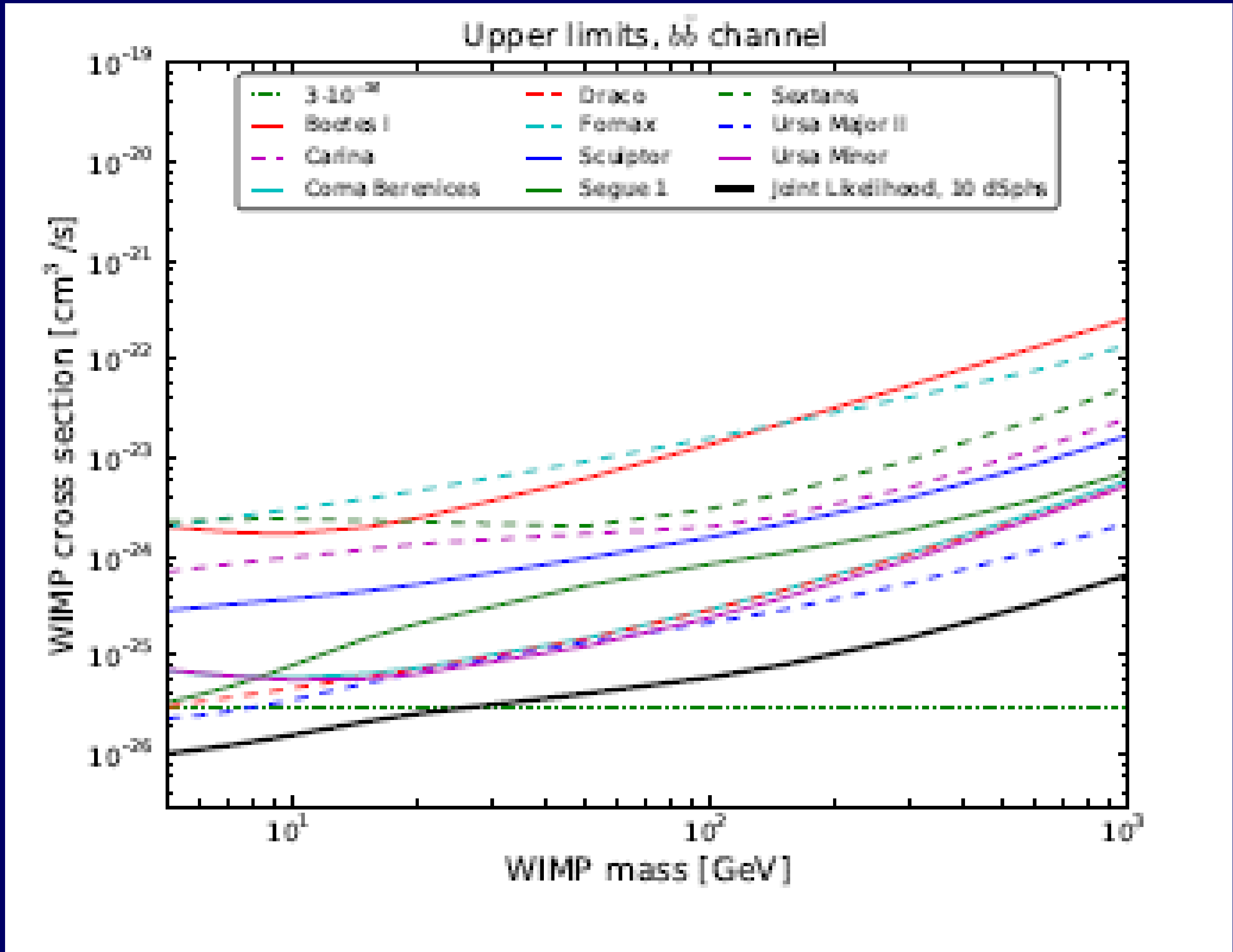
Nuisance parameter

Line of sight integral, J-factor

Upper limits from the profile likelihood



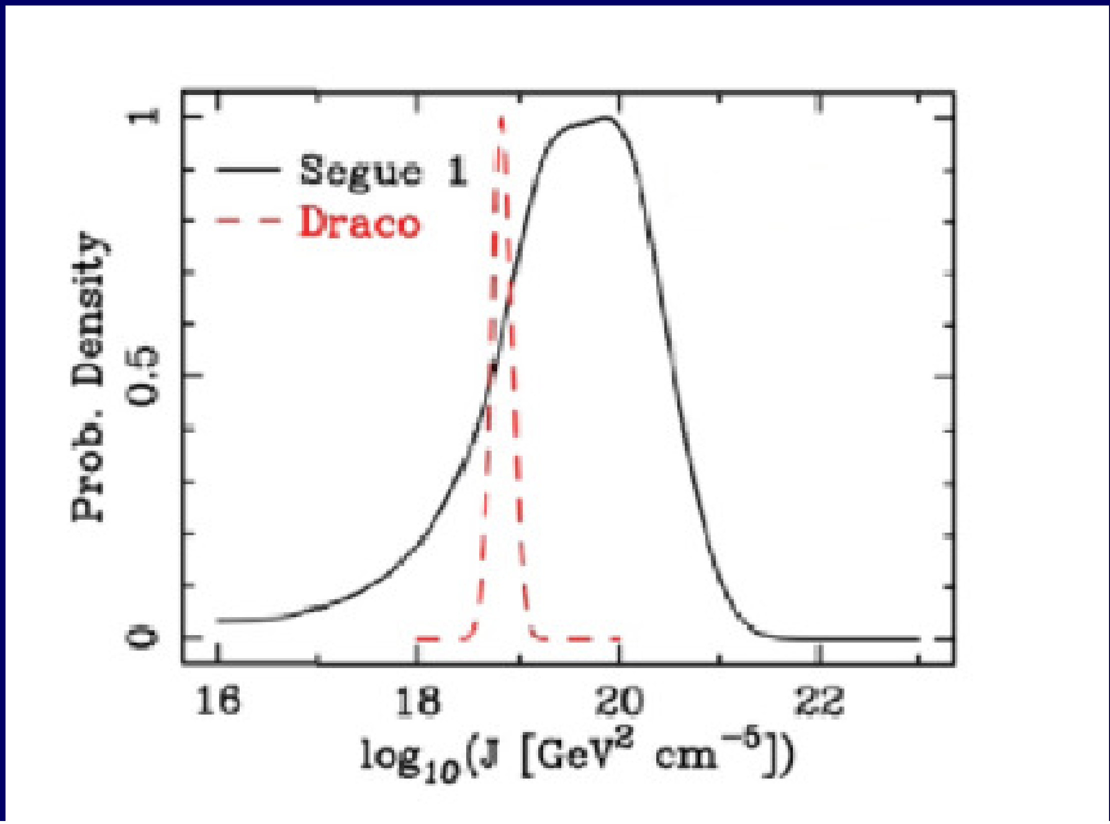
Results



Upper limits, $b\bar{b}$ channel
$3\cdot10^{-26}$
Bootes I
Carina
Coma Berenices
Draco
Fornax
Sculptor
Segue 1
Sextans
Ursa Major II
Ursa Minor
Joint Likelihood, 10 dSphs
WIMP cross section [cm^3 /s]
WIMP mass [GeV]
10^{-19}
10^{-20}
10^{-21}
10^{-22}
10^{-23}
10^{-24}
10^{-25}
10^{-26}
10^{1}
10^{2}
10^{3}

A subtlety for afficionados: posteriors and likelihoods

- **The algorithms used to infer DM density line of sight integrals delivers Bayesian posterior distributions, see later in the talk**
- **Fermi-LAT used the likelihood**
- **?**



G. Martinez, Stockholm

Log-normal likelihood → no!

$$L(D|\mathbf{pw},\{\mathbf{p}\}_i) = \prod_i L_i^{\text{LAT}}(D|\mathbf{pw},\mathbf{p}_i)$$

$$\times \frac{1}{\ln(10)\, J_i \sqrt{2\pi}\sigma_i} e^{-\left[\log_{10}(J_i) - \overline{\log_{10}(J_i)}\right]^2/2\sigma_i^2}$$

variate

parameter

The log-normal likelihood function has the variate in the denominator of the pre-factor.

We are using the likelihood corresponding to a flat prior and a log-normal posterior

Profile likelihood in direct detection

- Profile likelihood limits are presented even for cases that call for exact methods

- Coverage studies have shown that profile likelihood works satisfactorily even then.

- But eg. consider the case of no background:
  Pr(n=0|s=2) = (2)^0 * exp(-2) / 0! = 0.13.

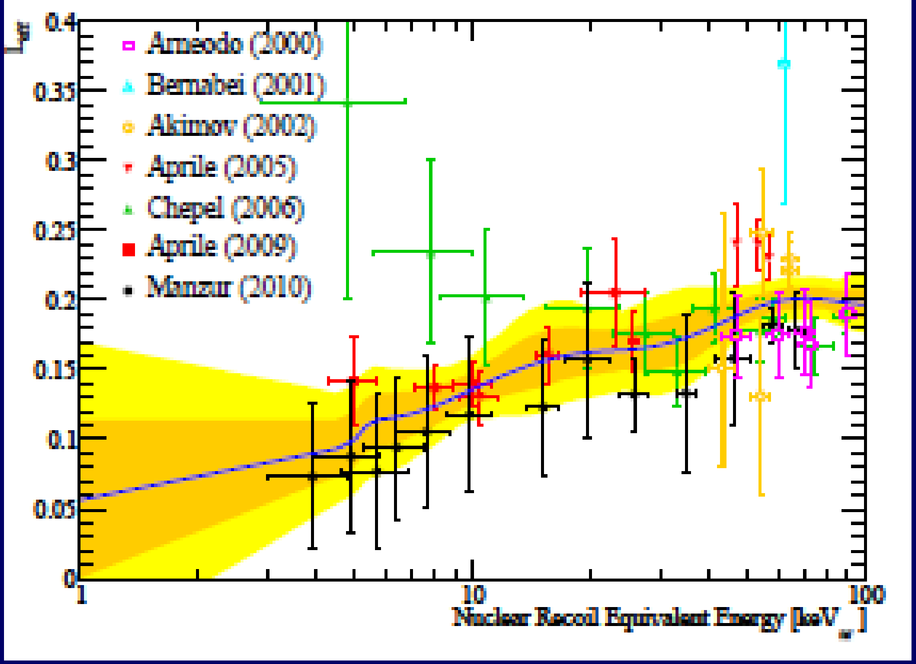
  In that case, a likelihood interval returns:
  [0,1.35]@95% , ie. undercovers by 20 % or so
  (exact method gives [0,2.44])

Direct detection likelihood

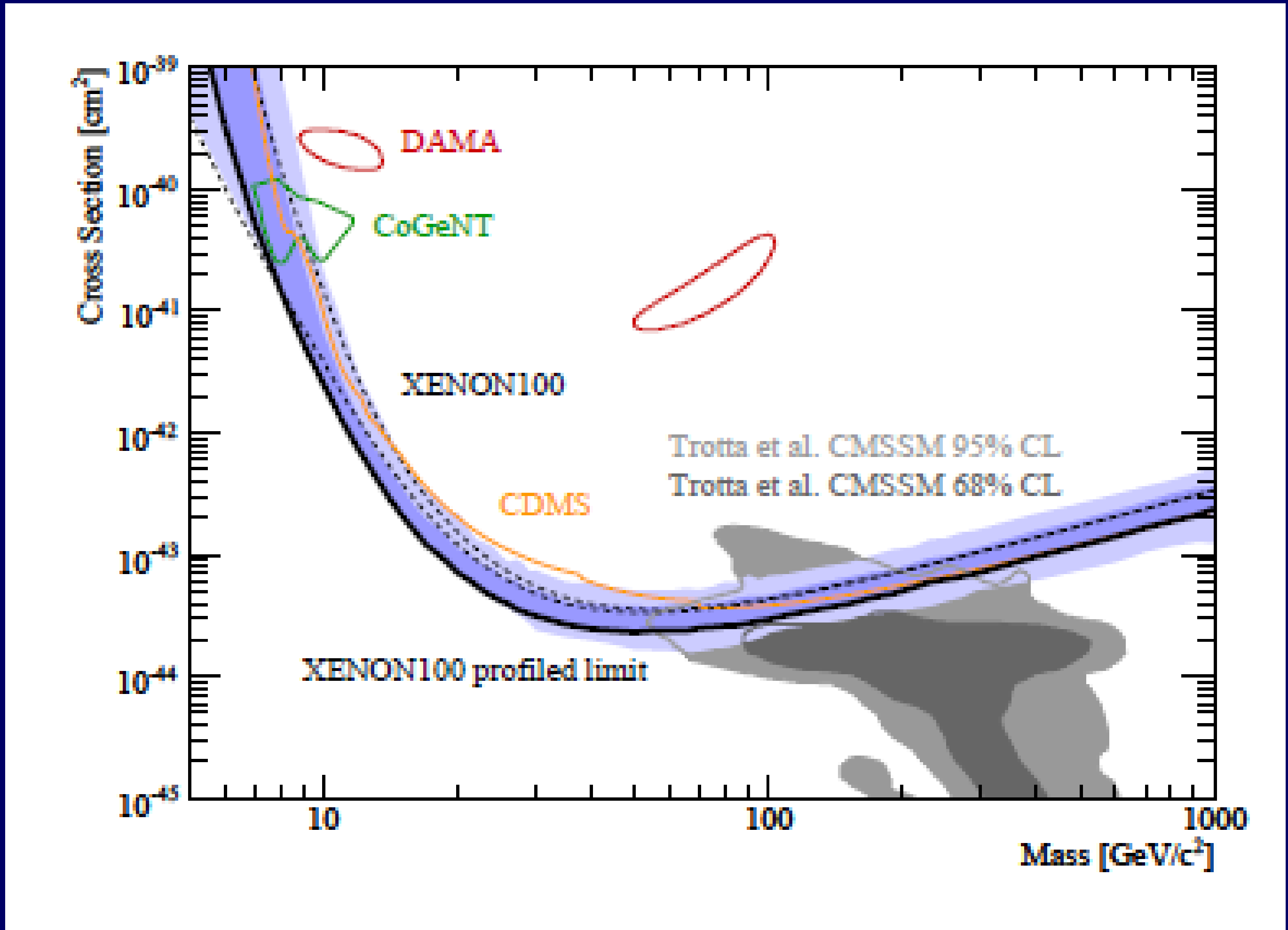
Xenon-100, Phys.Rev. D84 (2011) 052003

$$\mathscr{L} = \mathscr{L}_1(\sigma, N_b, \epsilon_s, \epsilon_b, \mathcal{L}_{\rm eff}, v_{\rm esc}; m_\chi) \times \mathscr{L}_2(\epsilon_s) \times \mathscr{L}_3(\epsilon_b) \times \mathscr{L}_4(\mathcal{L}_{\rm eff}) \times \mathscr{L}_5(v_{\rm esc}).$$

- **Likelihood terms taken from separate measurements or assumed Gaussian if not known**



Upper limits



- **Background expectation ~22 events**
- **Validity of Wilk's theorem checked.**

Bayesian methods in dark matter searches.

Bayes probability - basics

- **Bayesian probability is applicable to non-repeatable phenomena.**
- **Bayesian probability is defined as the degree of belief.**
- **It depends on the state of knowledge and beliefs of the observer.**
- **In practial applications, Bayesian probability is used for calculations invoking Bayes theorem:**

$$p(\sigma v \mid \vec{d}) = \frac{P(\vec{d} \mid \sigma v) P(\sigma v)}{\int P(\vec{d} \mid \sigma v) P(\sigma v) d\sigma v}$$

Likelihood function

Posterior distribution

Prior distribution

$$p(\sigma v \mid \vec{d}) = \frac{P(\vec{d} \mid \sigma v) P(\sigma v)}{\int P(\vec{d} \mid \sigma v) P(\sigma v) d\sigma v}$$

- **Note: σv is an hypothesis, parameter of a theory, d is the measurement.**
- **Bayes is per defintion subjective: you have to assume a prior probability distribution.**

Bayes continued

- **Bayes theorem in itself is applicable to probabilities in general.**
- **The main conceptual difference is that hypotheses are treated as a random variable. This does not make sense in frequentist statistics. For example the mass of a particle is fixed (unknown) and does not change in repeated experiments.**

Neyman construction and marginalisation

- **Neyman construction with ordering principle according to likelihood ratio:**

$$R(s,n)_{\mathcal{L}} = \frac{\mathcal{L}_{s+b}(n)}{\mathcal{L}_{s_{best}+b}(n)}$$

Feldman & Cousins,

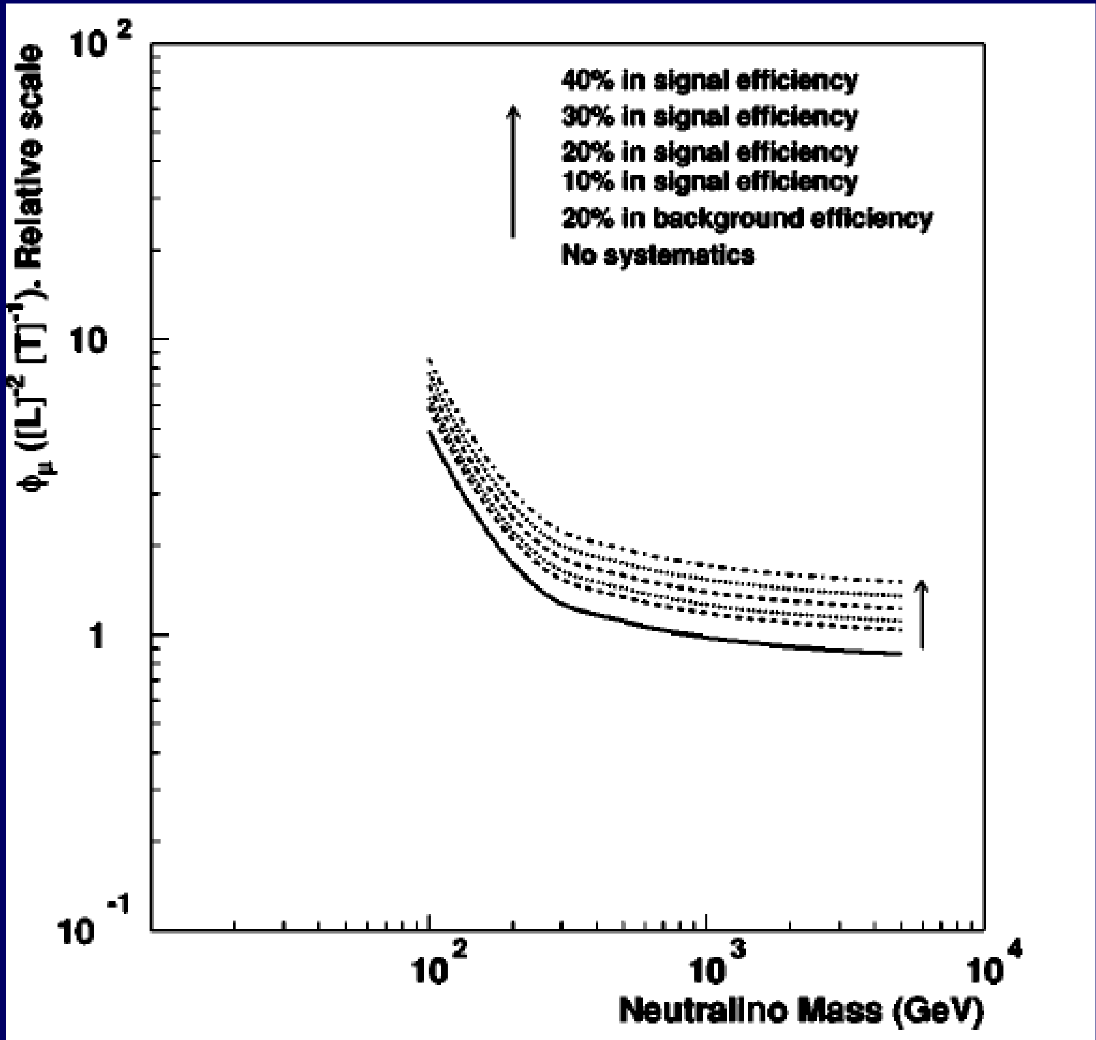
Phys. Rev.D 57, 3873(1998)

- **Treat nuisance parameter by replacing Poisson distribution with:**

$$q(n)_{s+b} = \frac{1}{2\pi\sigma_b\sigma_\epsilon}\int_0^\infty\int_0^\infty p(n)_{b'+\epsilon' s} \times e^{-(b-b')^2/2\sigma_b^2} e^{-(1-\epsilon')^2/2\sigma_\epsilon^2} db'\, d\epsilon'$$

Highland&Cousins, NIM A 320,331 (1992), JC+, Phys. Rev. D 67, 012002 (2003)

Effect of efficiency uncertainty on neutrino flux from DM in Earth



JC+, Phys. Rev. D 67, 012002 (2003)

Bayesian Inference on Supersymmetry

- **DM-Theories → Supersymmetry (O(100 free parameters), and the likelihood space is very complicated.**

- **How to estimate parameters/intervals in this set-up?**

  → Markov Chain Monte Carlo, MultiNEST deliver a direct estimate of the posterior on SUSY parameters

Feroz+, MNRAS 398, 2009, 1601

Ruiz de Austri+, JHEP 0605:002,2006

E.g. constrained supersymmetry, examples from Strege+, 1212.2636

	cMSSM Parameters	
	Flat priors	Log priors
	cMSSM parameters	
m_0 [GeV]	(50.0, 4000.0)	$(10^{1.7}, 10^{3.6})$
$m_{1/2}$ [GeV]	(50.0, 4000.0)	$(10^{1.7}, 10^{3.6})$
A_0 [GeV]	(-4000.0, 4000.0)	
$\tan\beta$	(2.0, 65.0)	
NUHM parameters as above, and additionally:		
μ [GeV]	(-2000.0, 2000.0)	
m_A [GeV]	(50.0, 4000.0)	$(10^{1.7}, 10^{3.6})$

Nuisance parameters

	Gaussian prior	Range scanned	Ref.
SM nuisance parameters			
M_t [GeV]	173.2 ± 0.9	(170.5, 175.9)	[29]
$m_b(m_b)^{\bar{MS}}$ [GeV]	4.20 ± 0.07	(3.99, 4.41)	[30]
$[\alpha_{em}(M_Z)^{\bar{MS}}]^{-1}$	127.955 ± 0.030	(127.865, 128.045)	[30]
$\alpha_s(M_Z)^{\bar{MS}}$	0.1176 ± 0.0020	(0.1116, 0.1236)	[31]
Astrophysical nuisance parameters			
$\rho_{\rm loc}$ [GeV/cm^3]	0.4 ± 0.1	(0.1, 0.7)	[32]
$v_{\rm lsr}$ [km/s]	230.0 ± 30.0	(140.0, 320.0)	[32]
$v_{\rm esc}$ [km/s]	544.0 ± 33.0	(445.0, 643.0)	[32]
v_d [km/s]	282.0 ± 37.0	(171.0, 393.0)	[32]
Hadronic nuisance parameters			
f_{Tu}	0.02698 ± 0.00395	(0.015, 0.039)	[33]
f_{Td}	0.03906 ± 0.00513	(0.023, 0.055)	[33]
f_{Ts}	0.363 ± 0.119	(0.0006, 0.72)	[33]

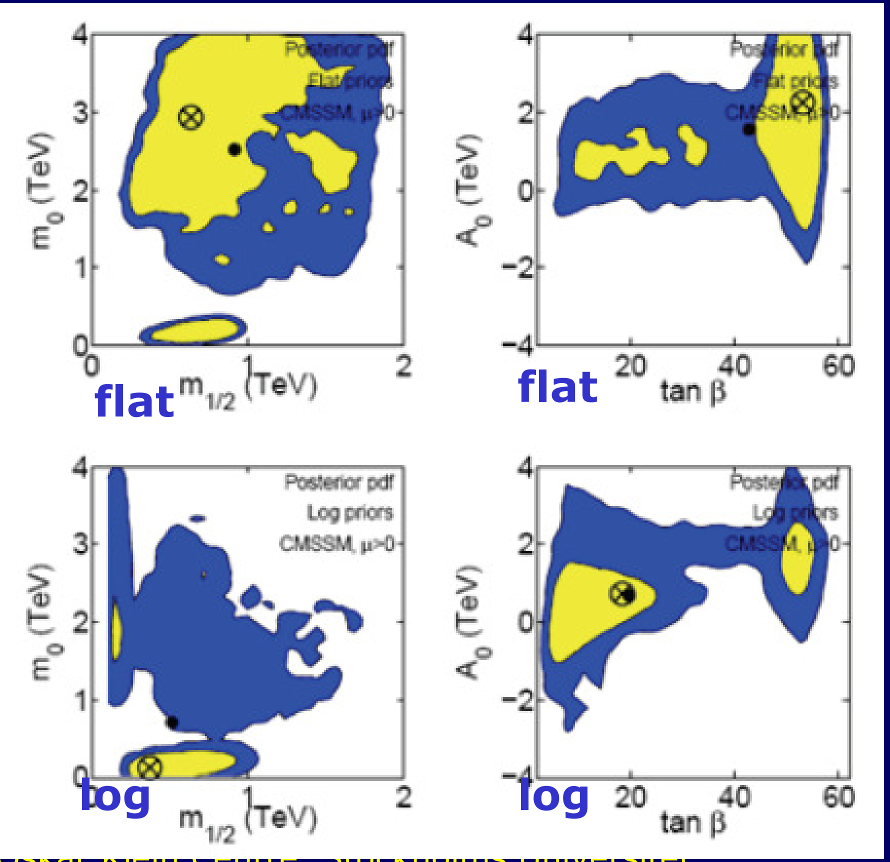
Observables

Observable	Mean value μ	Uncertainties σ (exper.)	τ (theor.)	Ref.
M_W [GeV]	80.399	0.023	0.015	[34]
$\sin^2\theta_{eff}$	0.23153	0.00016	0.00015	[34]
$\delta a_\mu^{\mathrm{SUSY}} \times 10^{10}$	28.7	8.0	2.0	[35]
$BR(\bar{B} \to X_s\gamma) \times 10^4$	3.55	0.26	0.30	[36]
$R_{\Delta M_{B_s}}$	1.04	0.11	-	[37]
$\frac{BR(B_u\to\tau\nu)}{BR(B_u\to\tau\nu)_{SM}}$	1.63	0.54	-	[36]
$\Delta_{0-} \times 10^2$	3.1	2.3	-	[38]
$\frac{BR(B\to D\tau\nu)}{BR(B\to De\nu)} \times 10^2$	41.6	12.8	3.5	[39]
R_{l23}	0.999	0.007	-	[40]
$BR(D_s \to \tau\nu) \times 10^2$	5.38	0.32	0.2	[36]
$BR(D_s \to \mu\nu) \times 10^3$	5.81	0.43	0.2	[36]
$BR(D \to \mu\nu) \times 10^4$	3.82	0.33	0.2	[36]
$\Omega_\chi h^2$	0.1109	0.0056	0.012	[41]
m_h [GeV]	125.8	0.6	2.0	[19]
$BR(\overline{B}_s \to \mu^+\mu^-)$	3.2×10^{-9}	1.5×10^{-9}	10%	[20]

	Limit (95% CL)	τ (theor.)	Ref.
Sparticle masses	As in table 4 of Ref. [42].		
$m_0, m_{1/2}$	ATLAS, $\sqrt{s} = 8$ TeV, 5.8 fb^{-1} 2012 limits		[17]
$m_A, \tan\beta$	CMS, $\sqrt{s} = 7$ TeV, 4.7 fb^{-1} 2012 limits		[18]
$m_\chi - \sigma^{\mathrm{SI}}_{\tilde{\chi}^0_1-p}$	XENON100 2012 limits (224.6×34 kg days)		[21]

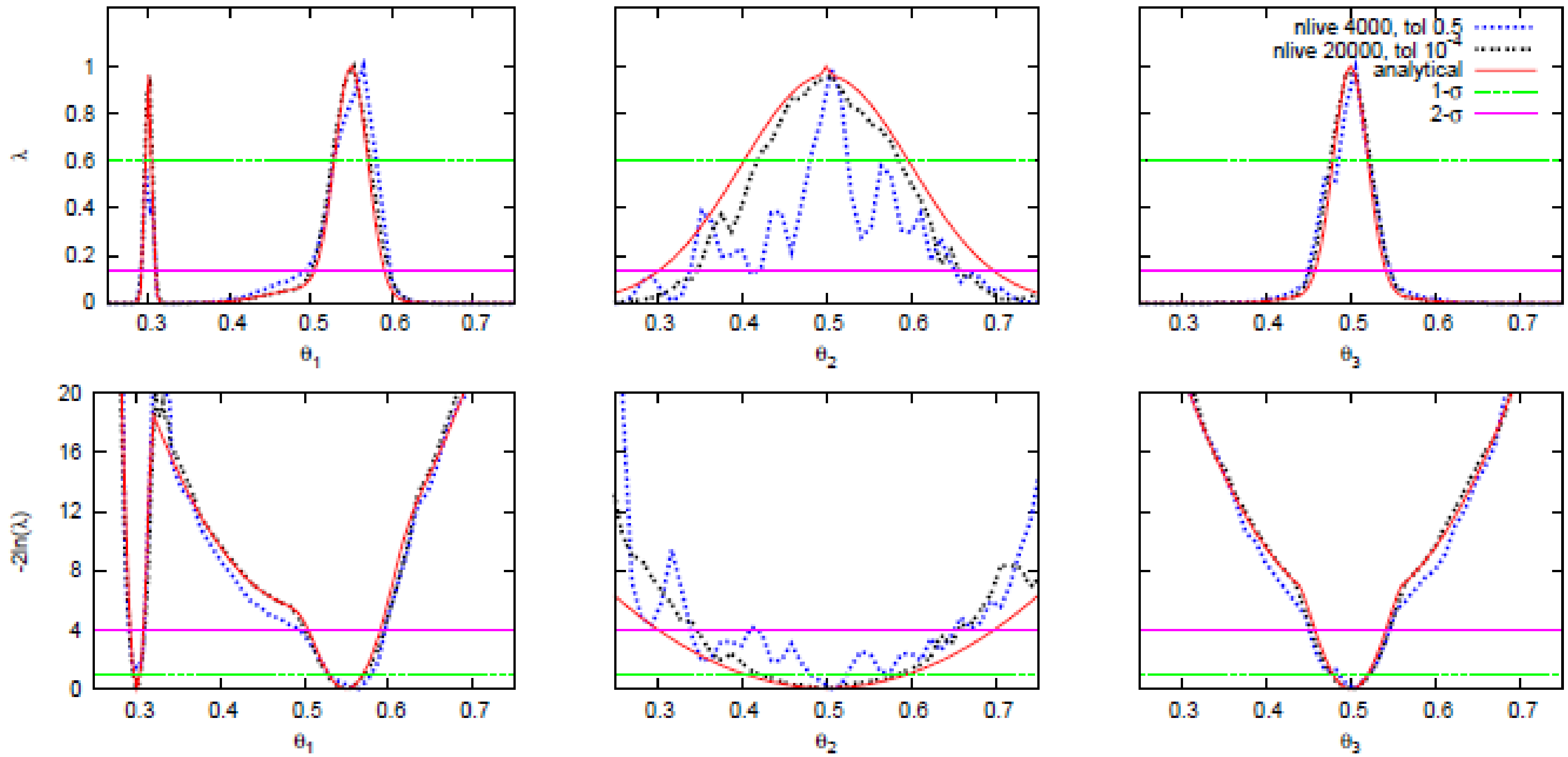
Challenges even in simplest Supersymmetric (4 parameters. CMSSM) theory

- **Prior dependence**
  - Flat vs. Log priors give signifcantly different results.
  - Remedied when including more data (LHC for CMSSM, but what happens if we have to go to 100 parameters?)



Not from the Strege paper

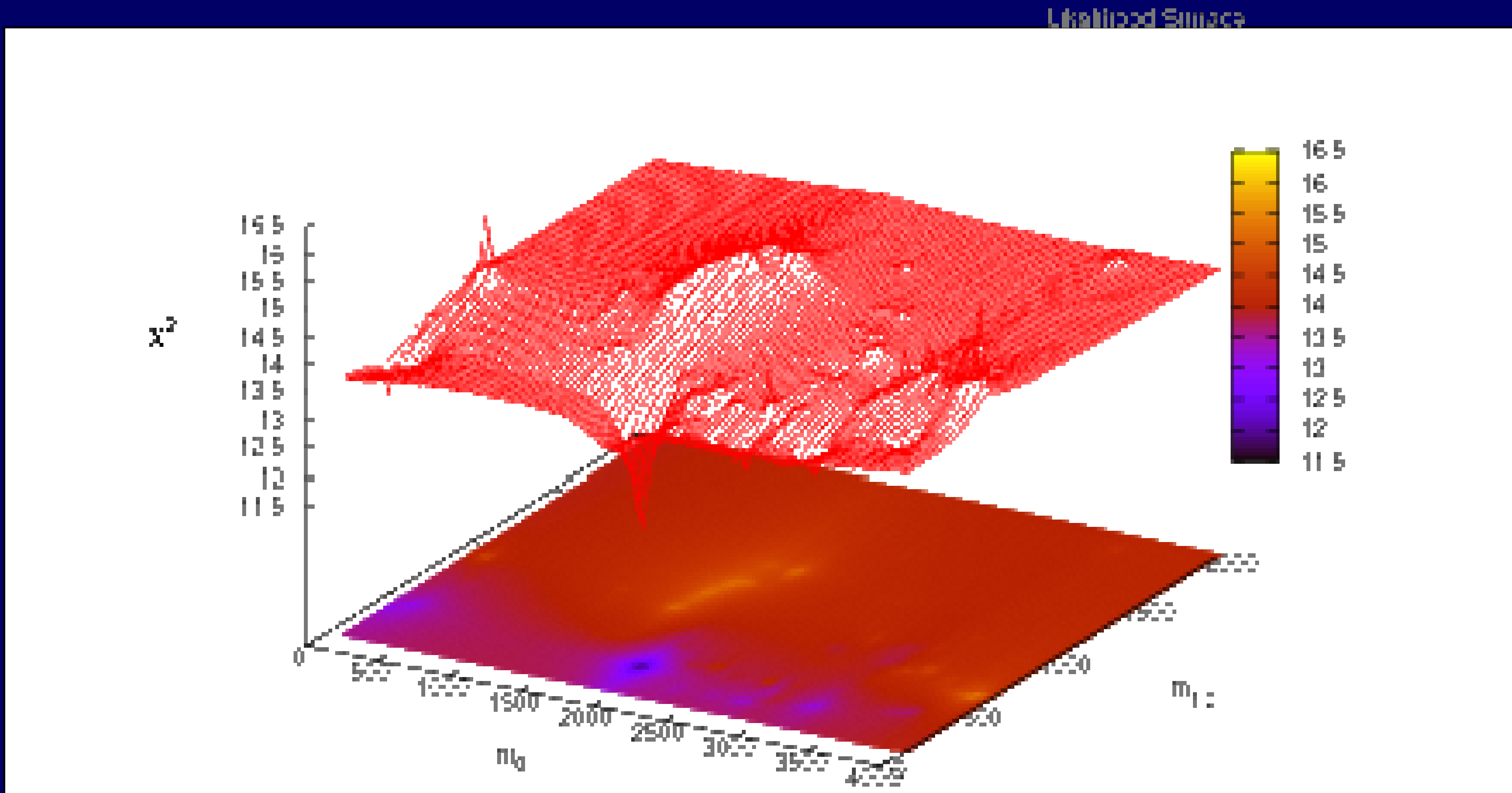
Point and intervals estimate can be made from the likelihood in this methodology – but tricky.



Challenges even in simplest Supersymmetric (4 parameters. CMSSM) theory

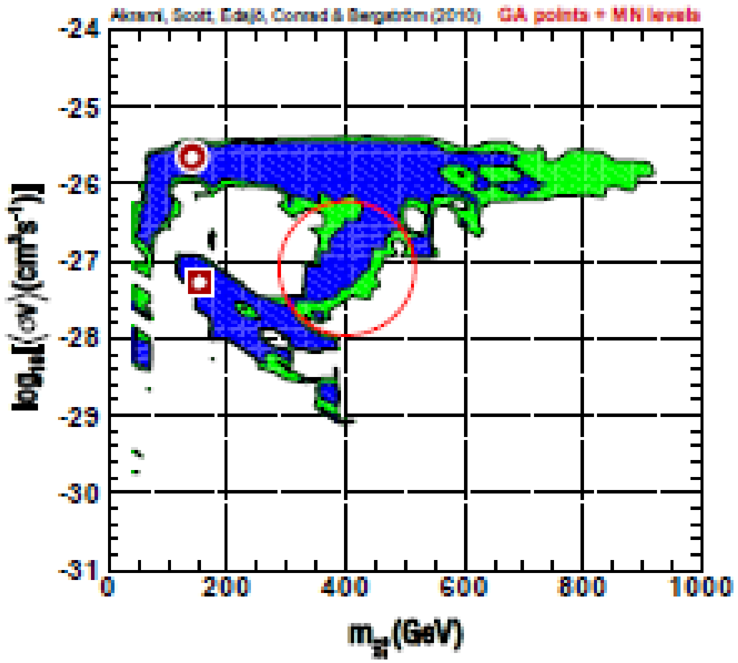
- **Sensitivity to fine-tuning (especially for profile likelihood)**
  - **PL picks "false" or "true" likelihood peaks**
  - **PL much more sensitive to adequate sampling of the likelihood**

e.g. Feroz+, **JHEP 1106:042,2011**

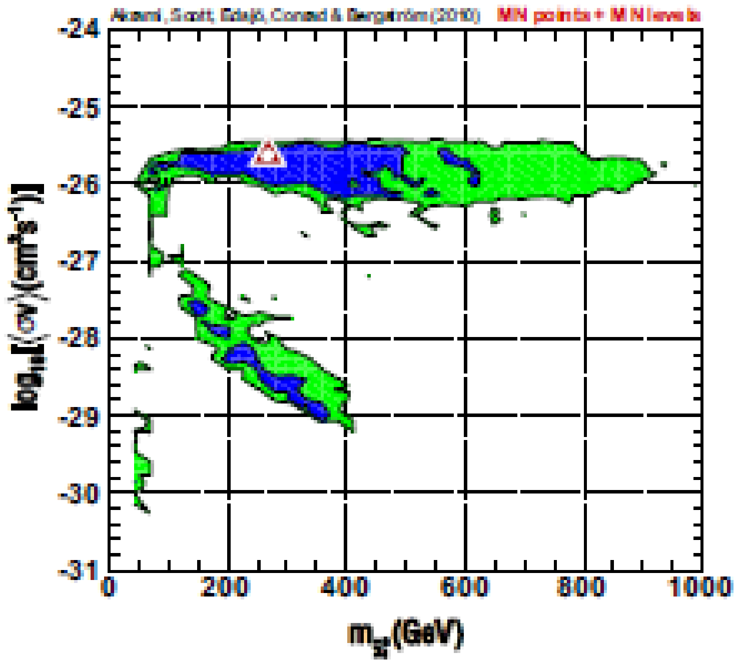


Other scanning algorithms: Genetic algorithms

Genetic algorithm



Multinest trapping into a local minimum



Akrami+, JHEP 1004:057,2010

Coverage

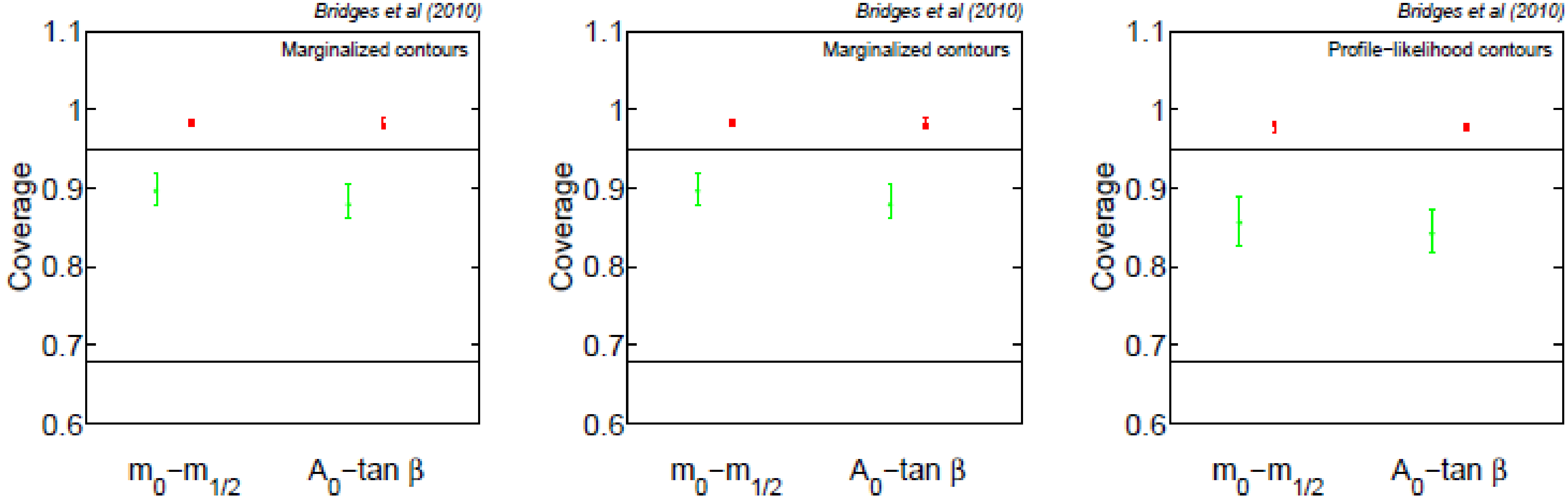
In a very large number of experiments, each providing a confidence interval [θ_{ll}, θ^{ul}], the fraction of intervals that contain the true value is 1- α, independent of what the true value is.

Challenges even in simplest Supersymmetric (4 parameters. CMSSM) theory

- **Frequentist properties**
  - Both over and undercoverage

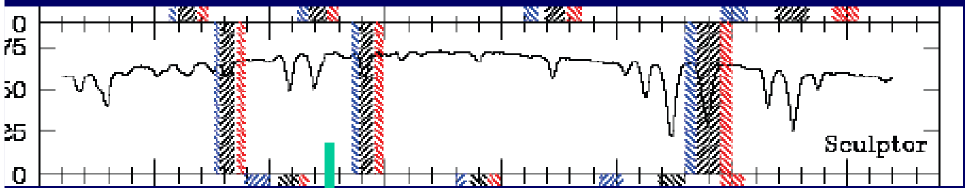
    Bridges+, JHEP 1103(2011) 012, LHC
    Akrami+, JCAP 1107 (2011) 002

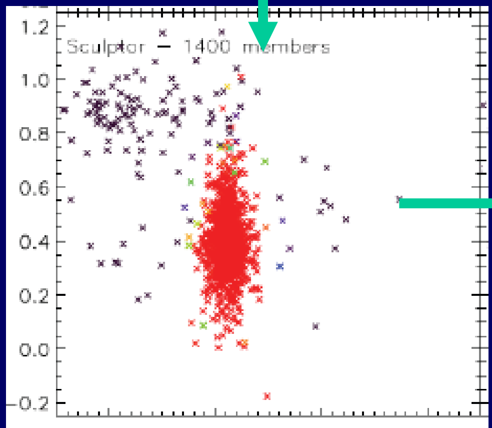
  - Bad sampling of the likelihood, boundaries on the parameters, flat prior in many dimensions (my guess) ….

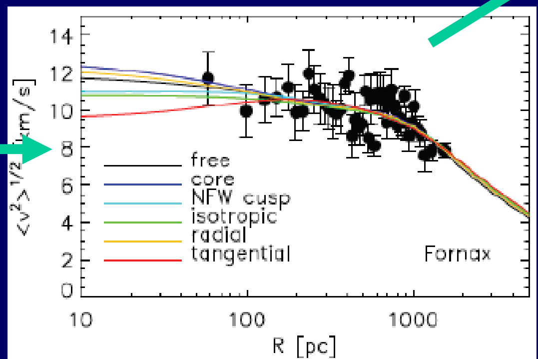


Dwarfs galaxies again – cleanest target

- **Stellar velocities can be used to measure DM density (error can be propagated to particle constraints, as we have seen)**







Name	l	b	d	$\overline{\log_{10}(J)}$	σ	ref.
	deg.	deg.	kpc	$\log_{10}[\mathrm{GeV}^2\mathrm{cm}^{-5}]$		
Bootes I	358.08	69.62	60	17.7	0.34	[15]
Carina	260.11	−22.22	101	18.0	0.13	[16]
Coma Berenices	241.9	83.6	44	19.0	0.37	[17]
Draco	86.37	34.72	80	18.8	0.13	[16]
Fornax	237.1	−65.7	138	17.7	0.23	[16]
Sculptor	287.15	−83.16	80	18.4	0.13	[16]
Segue 1	220.48	50.42	23	19.6	0.53	[18]
Sextans	243.4	42.2	86	17.8	0.23	[16]
Ursa Major II	152.46	37.44	32	19.6	0.40	[17]
Ursa Minor	104.95	44.80	66	18.5	0.18	[16]

e.g:
Charbonnier+, MNRAS 418 (2011) 1526
Strigari+,Phys. Rev. D, 75, 083526
Evans+, Phys. Rev., D69, 123501, (2004)

Posterior for the mass, strong prior dependence

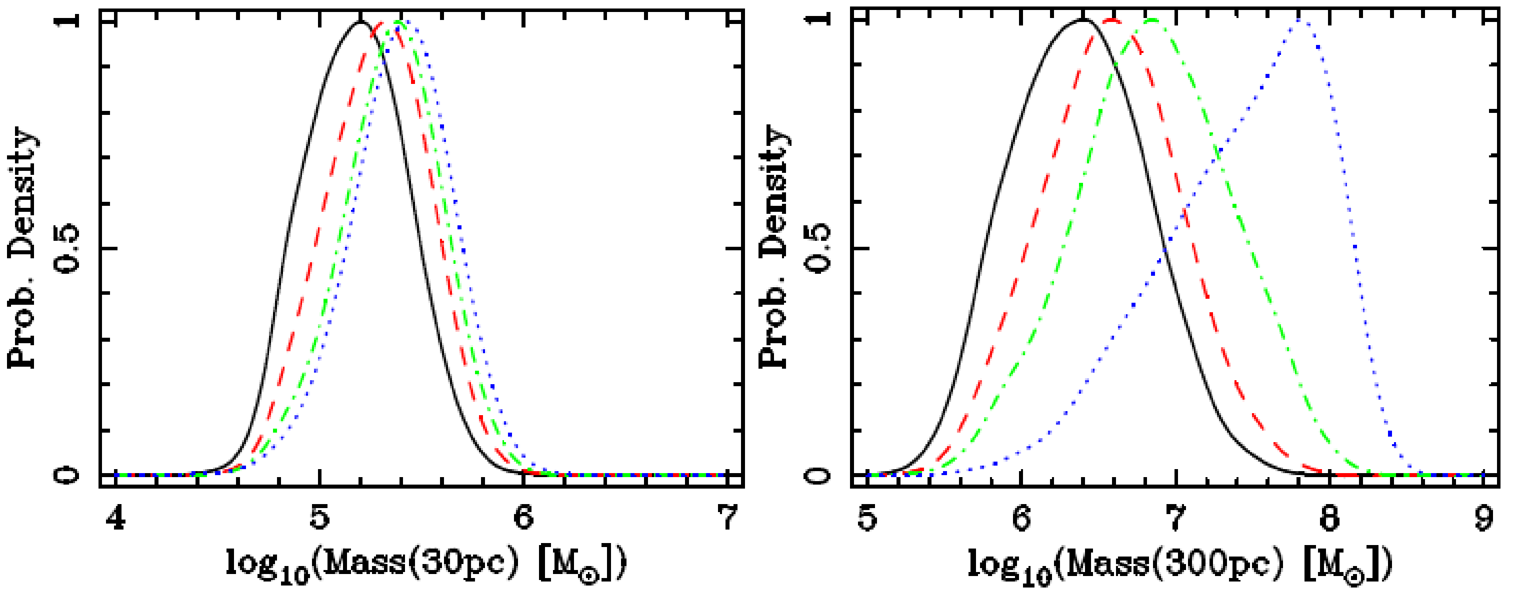


Figure 1. Posterior probability distribution for the mass within 30 pc for Segue 1(*left panel*) and the mass within 300 pc for Segue 1 (*right panel*). The four curves in each panel assume different Bayesian priors: a uniform prior in V_{max}^{-3} (black, solid) V_{max}^{-2} (red, dashed), V_{max}^{-1} (green, dot-dashed), and $\ln(V_{max})$ (blue, dotted). The prior distributions are truncated at $V_{max} = 3\text{km s}^{-1}$ as described in the text. Increasing negative powers of V_{max} causes the posterior to be more "biased" toward lower mass solutions. As a result, the posterior corresponding to these different priors differ.

Strigari+,Phys. Rev. D, 75, 083526

Possible solutions: hierarchical Bayes (multi-level modelling)?

Bayesian methods necessary?

- **Estimation of J factors and inference in SUSY does not per-se require Bayesian methods.**
- **The main motivation for these methods is ease of use, which to my mind is not sufficient.**
- **There are however, conceptually, a few arguments valid in DM searches:**
  - Systematic uncertainties (e.g. theoretical) are sometimes beliefs anyway, there is little reason to pretend these estimates are statistically distributed.
  - ”Naturalness” (see SUSY likelihood space) is very conveniently implemented in a Bayesian framework.

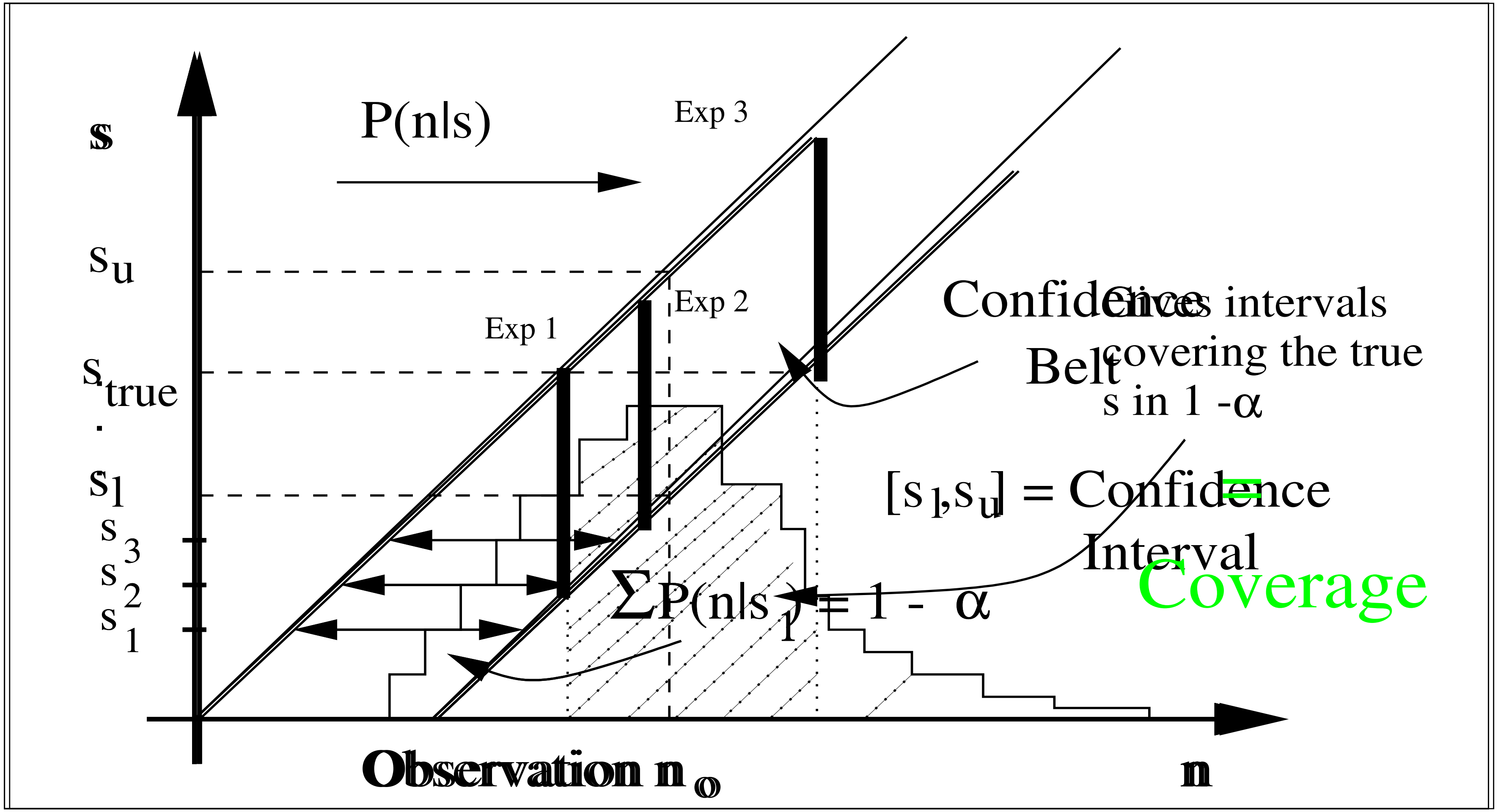
Summary

- **Astrophyiscal searches for dark matter require modern statistical methods for parameter estimation and hypothesis testing.**

- **Hypothesis testing: cases were the Wilks/Chernoff theorem are not applicable (in worst case non-nested hypotheses), trial factor correction**

- **Interval estimation: nasty nuisance parameters, to be treated by Profile likelihood/marginalisation, inference in complicated/highly structured likelihood spaces.**

Exact frequentist intervals- The Neyman construction



Coverage studies

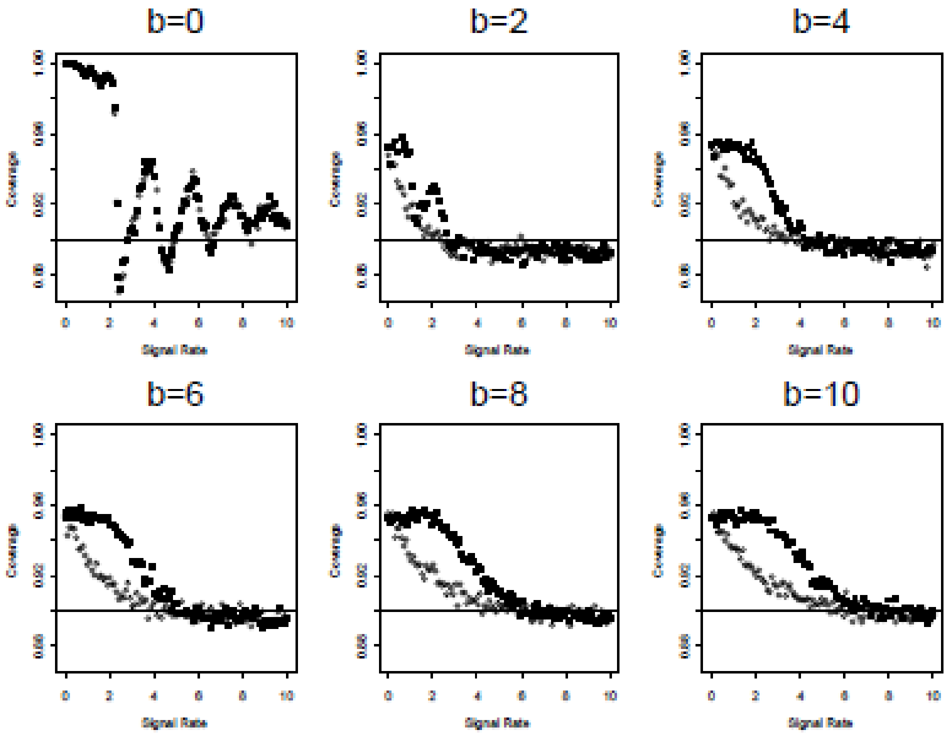


Figure 5: 90% coverage graphs when the signal is modeled as a Poisson and the background and the efficiency are modeled as Gaussians with $\sigma_b = 0.5$, $\epsilon = 0.85$ and $\sigma_\epsilon = 0.075$. The empty circles show the coverage using the unbounded likelihood method and the solid squares show the coverage using the bounded likelihood method.

eg. Rolke, Lopez, JC , NiM A551 (2005) 493-503

Estimate of DM density.

$$\mathcal{L}(\mathcal{H},\mathcal{V}) \equiv P(d|\mathcal{H},\mathcal{V}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(\sigma^2_{th,i}+\sigma^2_{m,i})}} \exp\left[-\frac{1}{2}\frac{(d_i-u)^2}{\sigma^2_{th,i}+\sigma^2_{m,i}}\right]$$

$n, r_{max}, V_{max},$

anisotropy parameters

Slope index of density profile
Radius of maximum radial velocity
Maximum radial velocity

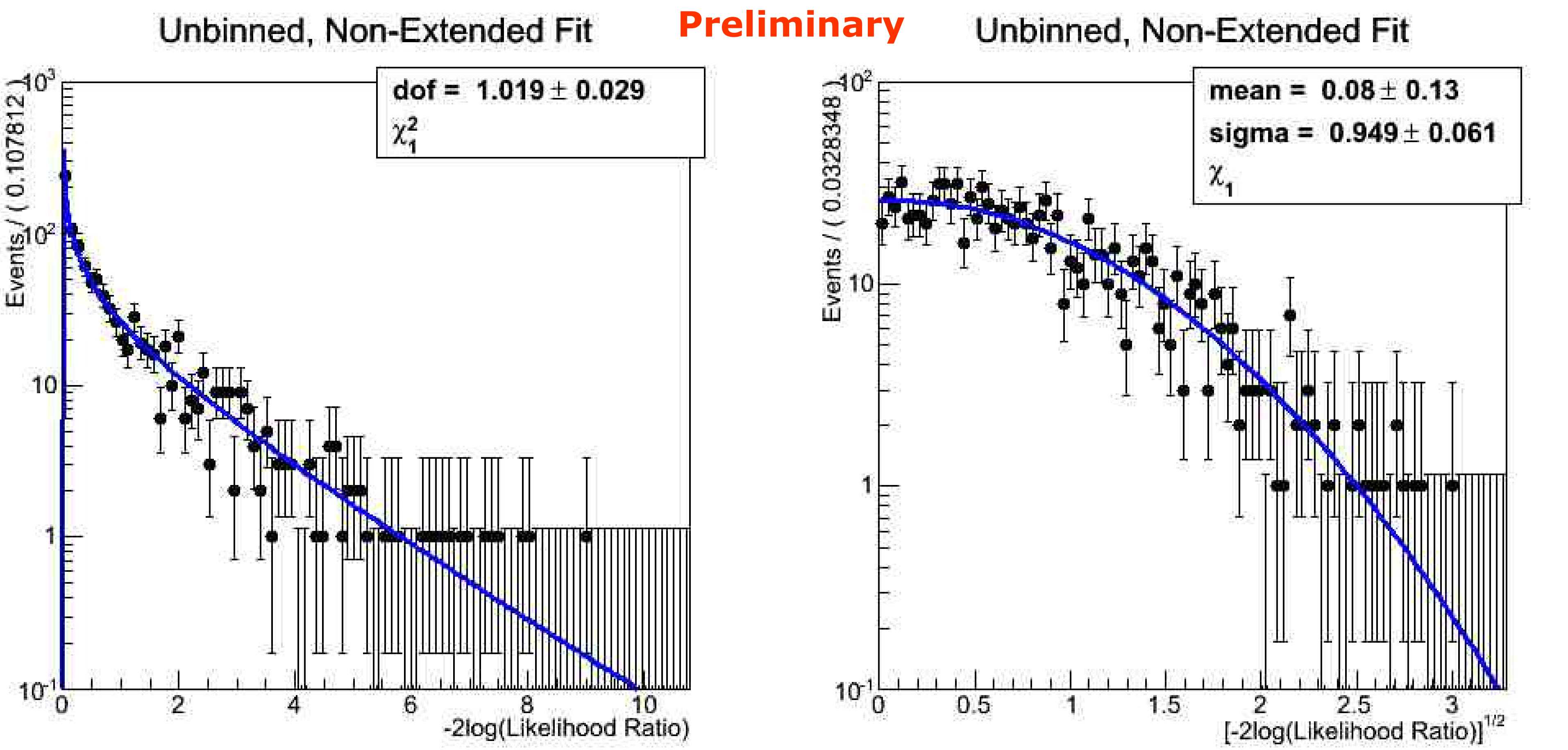
Los velocity dispersion, intrinsic
Los velocity dispersion, measurement

$$P(\mathcal{H},\mathcal{V}|d) \propto P(\mathcal{H})P(\mathcal{V})\mathcal{L}(\mathcal{H},\mathcal{V})$$

Martinez+, JCAP 0906 (2009) 014

Null distribution, unbinned, no Poisson term

JC, H. Dickinson, 2013



Extended likelihood (with Poisson term)

JC, H. Dickinson, 2012