Corfu Summer Institute 2009 - Sept 1st 2009

A Bayesian approach to supersymmetry phenomenology

Roberto Trotta
Imperial College London, Astrophysics Group

In collaboration with:
R. Ruiz de Austri, L. Roszkowski, M. Hobson, F. Feroz, J. Silk, L. Strigari, C. P. de los Heros, M. Kaplinghat, G. Martinez, J. Bullock

Imperial College London

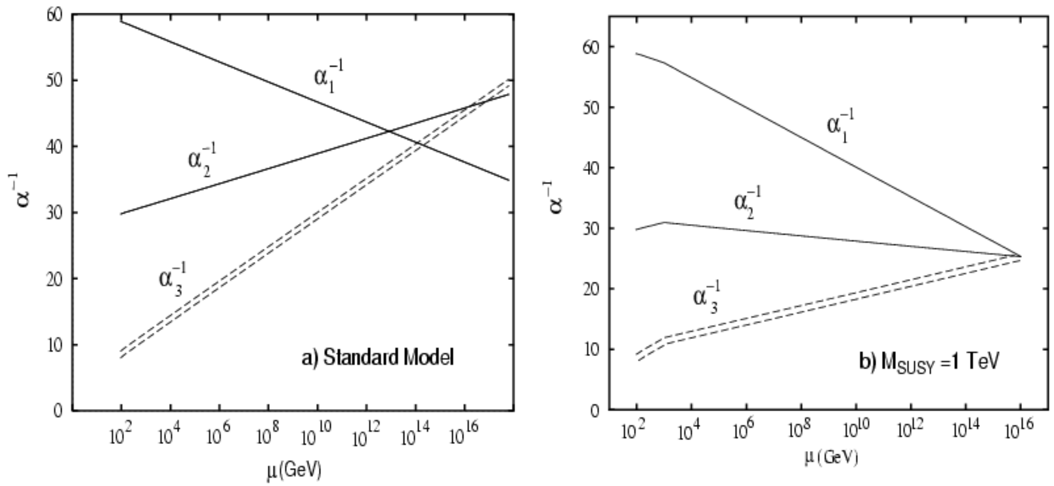
There are three kinds of lies: lies, damned lies, and statistics.
(Mark Twain, reportedly quoting Benjamin Disraeli)

The problem

Introduction

Imperial College London

- Why supersymmetry? (SUSY)



Amaldi et al (1991)

The particle content of SUSY

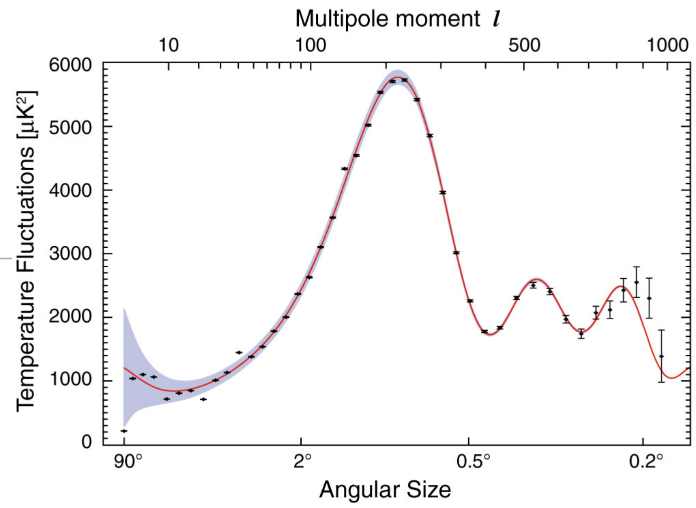
Imperial College London

Standard Model particles and fields		Supersymmetric partners			
		Interaction eigenstates		Mass eigenstates	
Symbol	Name	Symbol	Name	Symbol	Name
$q = d, c, b, u, s, t$	quark	$\tilde{q}_L, \tilde{q}_R$	squark	$\tilde{q}_1, \tilde{q}_2$	squark
$l = e, \mu, \tau$	lepton	$\tilde{l}_L, \tilde{l}_R$	slepton	$\tilde{l}_1, \tilde{l}_2$	slepton
$\nu = \nu_e, \nu_\mu, \nu_\tau$	neutrino	$\tilde{\nu}$	sneutrino	$\tilde{\nu}$	sneutrino
g	gluon	$\tilde{g}$	gluino	$\tilde{g}$	gluino
$W^\pm$	W-boson	$\tilde{W}^\pm$	wino	$\tilde{\chi}^\pm_{1,2}$	chargino
H^-	Higgs boson	$\tilde{H}_1^-$	higgsino		
H^+	Higgs boson	$\tilde{H}_2^+$	higgsino		
B	B-field	$\tilde{B}$	bino	$\tilde{\chi}^0_{1,2,3,4}$	neutralino
W^3	W^3-field	$\tilde{W}^3$	wino		
H_1^0	Higgs boson	$\tilde{H}_1^0$	higgsino		
H_2^0	Higgs boson	$\tilde{H}_2^0$	higgsino		
H_3^0	Higgs boson				

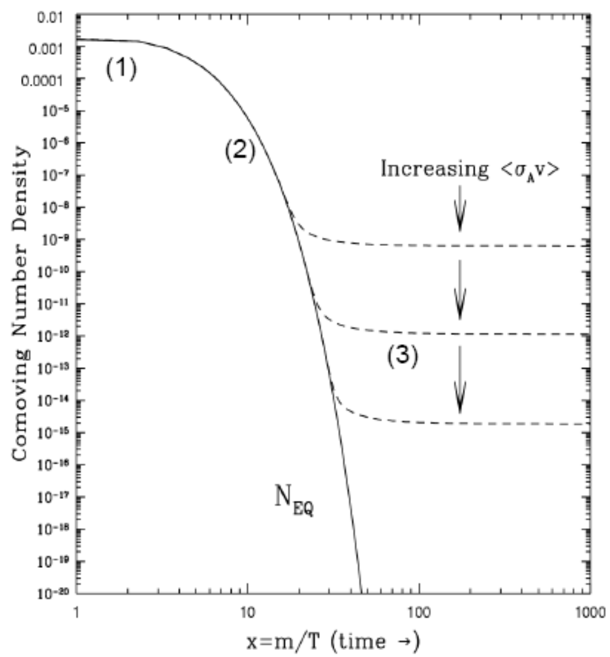
From Bertone, Hooper & Silk (2005)

Links with cosmology

- Strong cosmological and astrophysical evidence for the existence of Cold Dark Matter (CDM)
- The lightest SUSY particle (the neutralino) is a natural candidate for the CDM
- Cosmological DM abundance constrained (with a few assumptions) within ~ 10%:
  WMAP 5 yrs data:
  $\Omega_{CDM} h^2 = 0.1099 \pm 0.0062$



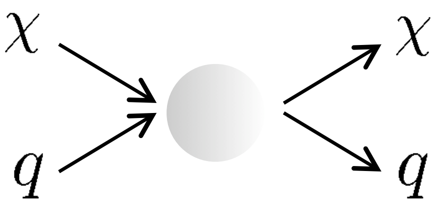
WMAP team



Courtesy Sabine Kraml

Links with astrophysics

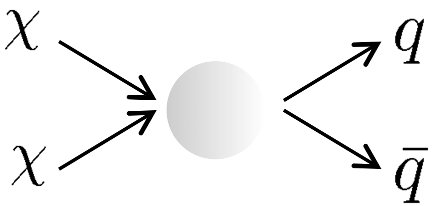
Imperial College London



Direct detection

Underground detectors looking for neutralino from local halo
scattering off nuclei
(complicated by local WIMP distribution)

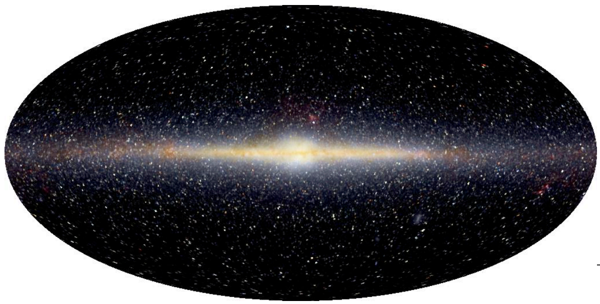
CRESST



Indirect detection

Look for neutralino-neutralino annihilation products:
gamma ray (continuum and & lines)
antimatter (e.g., positrons)
neutrinos (from Sun & Earth)
(complicated by gastrophysical nuisance parameters)

Fermi



The model & data

Imperial College London

- The general Minimal Supersymmetric Standard Model (MSSM): **105 free parameters!**
- Need some (pretty strong) simplifying assumption: **the Constrained MSSM** (CMSSM) reduces the free parameters to just **4 continous variables plus a discrete one** (sign(μ)).
- Clearly a highly constrained model (probably not the end of the story!)
- **Present-day data:** collider measurements of rare processes, CDM abundance (WMAP), sparticle masses lower limits, EW precision measurements. Soon, LHC sparticle spectrum measurements.
- **Astrophysical direct and indirect detection techniques might also be competitive:** neutrino (IceCUBE), gamma-rays (Fermi), antimatter (PAMELA), direct detection (XENON, CDMS, Eureca, Zeplin)

Analysis pipeline

Imperial College London

SCANNING ALGORITHM

4 CMSSM parameters
$\theta = \{m_0, m_{1/2}, A_0, \tan\beta\}$
(fixing sign(μ) > 0)

4 SM "nuisance parameters"
$\Psi=\{m_t, m_b, \alpha_S, \alpha_{EM}\}$

Data:
Gaussian likelihoods for each of the Ψ_j (j=1...4)

RGE

Non-linear numerical function
via SoftSusy 2.0.18
DarkSusy 4.1
MICROMEGAS 2.2
FeynHiggs 2.5.1
Hdecay 3.102

Observable quantities
$f_i(\theta, \Psi)$

CDM relic abundance
BR's
EW observables
g-2
Higgs mass
sparticle spectrum
(gamma-ray, neutrino, antimatter flux, direct detection x-section)

Physically acceptable?
EWSB, no tachyons, neutralino CDM

NO → Likelihood = 0

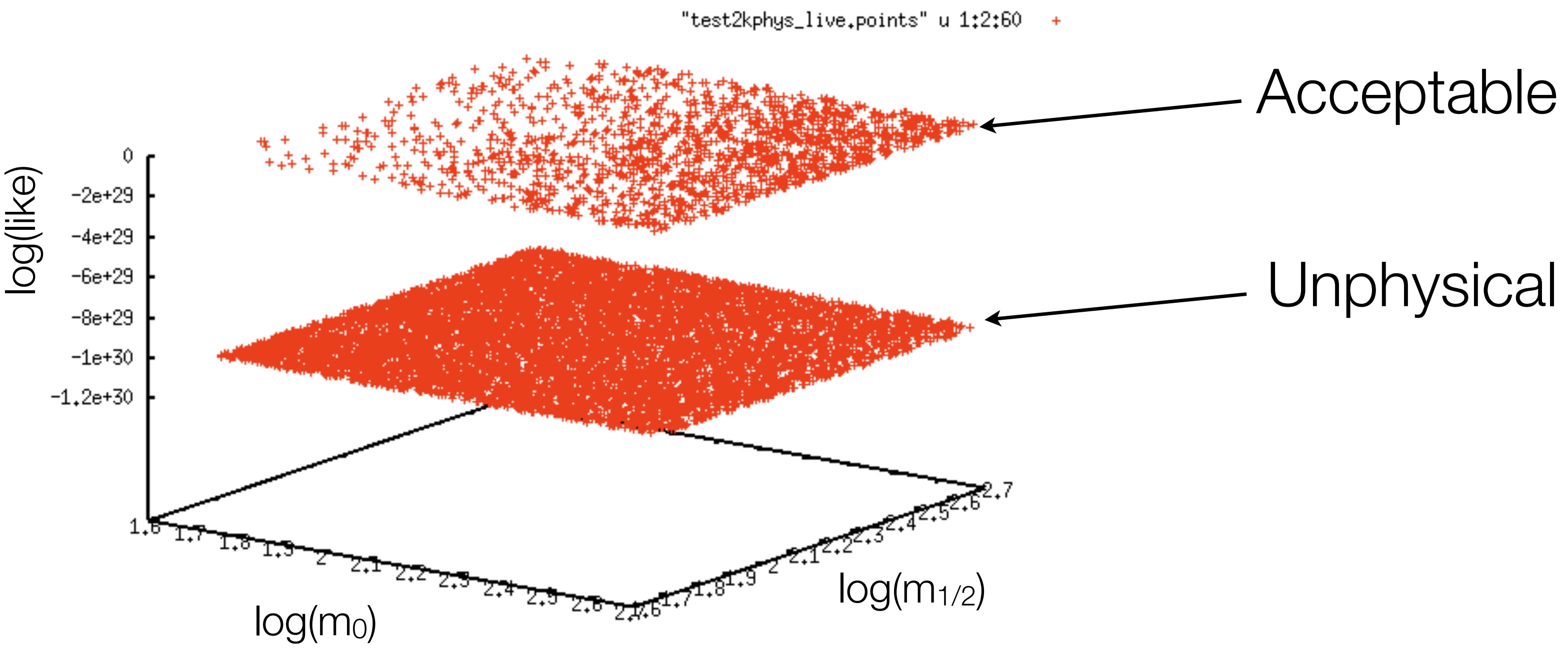
YES → Joint likelihood function

Data:
Gaussian likelihood
(CDM, EWO, g-2, $b \to s\gamma$, ΔM_{Bs})
other observables have only lower/upper limits

The accessible "surface"

Imperial College London

Scan from the prior with no likelihood except physicality constraints



Courtesy F. Feroz & M. Bridges

Data included

Imperial College London

Indirect observables

Observable	Mean value μ	Uncertainties σ (exper.)	τ (theor.)	ref.
M_W	80.398 GeV	25 MeV	15 MeV	[30]
$\sin^2\theta_{\rm eff}$	0.23153	16×10^{-5}	15×10^{-5}	[30]
$\delta a_\mu^{\rm SUSY} \times 10^{10}$	29.5	8.8	1.0	[31]
$BR(\overline{B} \to X_s\gamma) \times 10^4$	3.55	0.26	0.21	[32]
ΔM_{B_s}	17.77 ps^{-1}	0.12 ps^{-1}	2.4 ps^{-1}	[33]
$BR(\overline{B}_u \to \tau\nu) \times 10^4$	1.32	0.49	0.38	[32]
$\Omega_\chi h^2$	0.1099	0.0062	$0.1\,\Omega_\chi h^2$	[34]
	Limit (95% CL)		τ (theor.)	ref.
$BR(\overline{B}_s \to \mu^+\mu^-)$	$< 5.8 \times 10^{-8}$		14%	[35]
m_h	$>$ 114.4 GeV (SM-like Higgs)		3 GeV	[36]
ζ_h^2	$f(m_h)$ (see text)		negligible	[36]
$m_{\tilde{q}}$	$>$ 375 GeV		5%	[25]
$m_{\tilde{g}}$	$>$ 289 GeV		5%	[25]
other sparticle masses	As in table 4 of ref. [6].			

SM parameters

SM (nuisance) parameter	Mean value μ	Uncertainty σ (exper.)	Ref.
M_t	172.6 GeV	1.4 GeV	[24]
$m_b(m_b)^{\overline{MS}}$	4.20 GeV	0.07 GeV	[25]
$\alpha_s(M_Z)^{\overline{MS}}$	0.1176	0.002	[25]
$1/\alpha_{\rm em}(M_Z)^{\overline{MS}}$	127.955	0.03	[26]

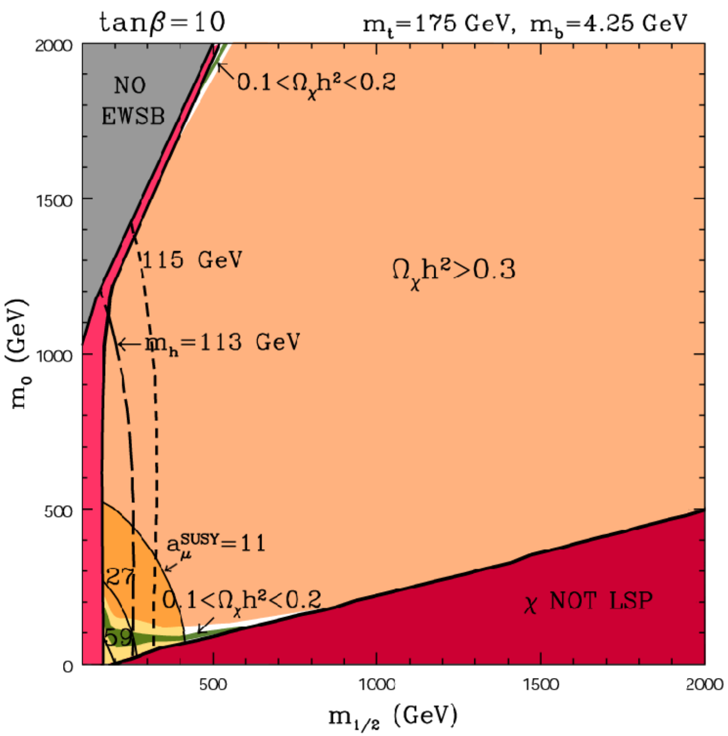
Why is this a difficult problem?

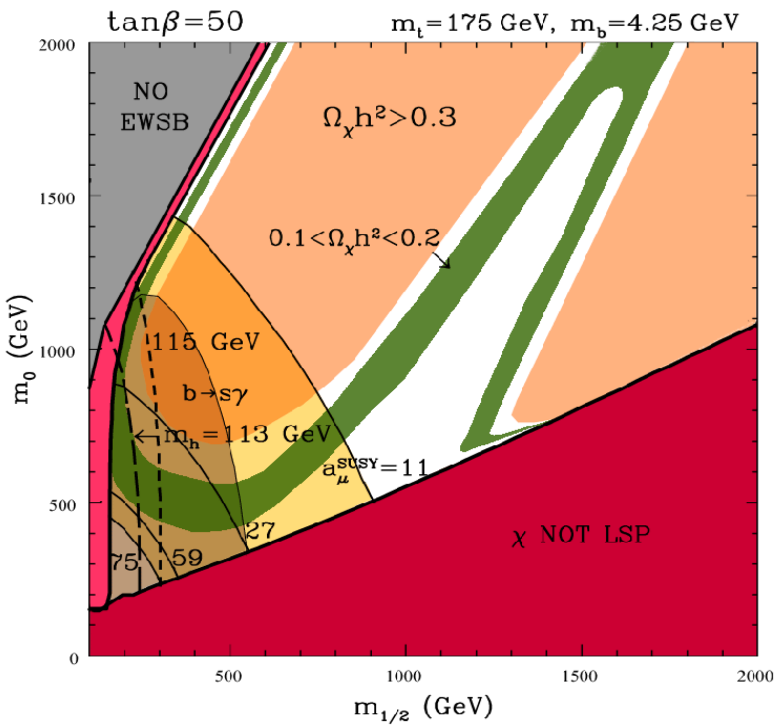
Imperial College London

- **Inherently 8-dimensional:** reducing the dimensionality over-simplifies the problem. Nuisance parameters (in particular m_t) cannot be fixed!
- **Likelihood discontinuous** and multi-modal due to physicality conditions
- RGE connect input parameters to observables in highly non-linear fashion: **only indirect (sometimes weak) constraints on the quantities of interest (-> prior volume effects are difficult to keep under control)**
- **Mild discrepancies between observables** (in particular, g-2 and $b \rightarrow s\gamma$) tend to pull constraints in different directions

2 dimensional slices

Imperial College London





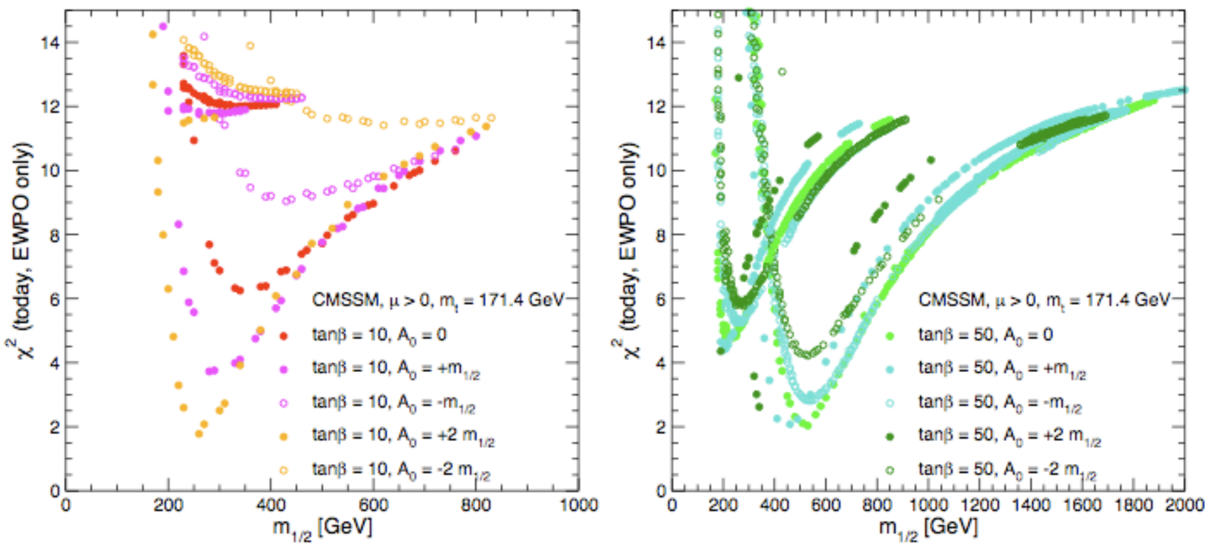
Roszkowski et al (2001)

Roberto Trotta

The “WMAP strips”

Imperial College London

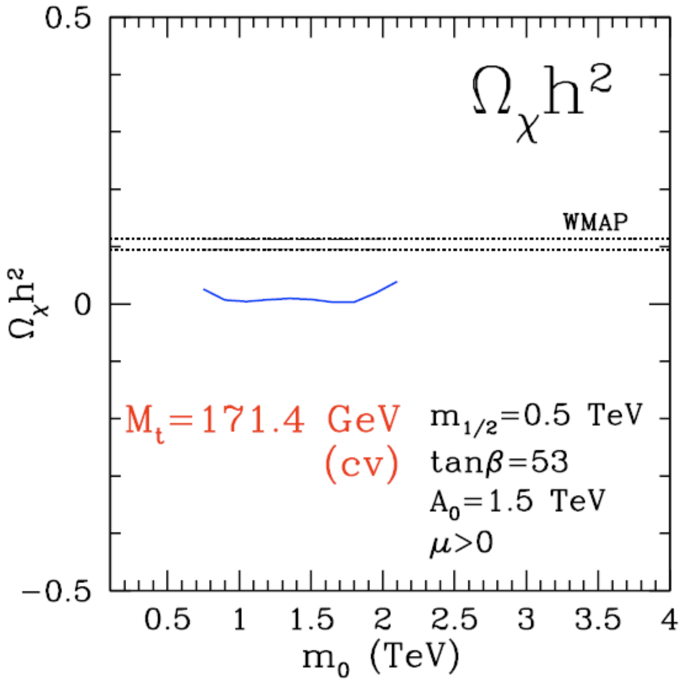
m_0 fixed by requiring that the CDM abundance matches the WMAP value.
All nuisance parameters fixed.

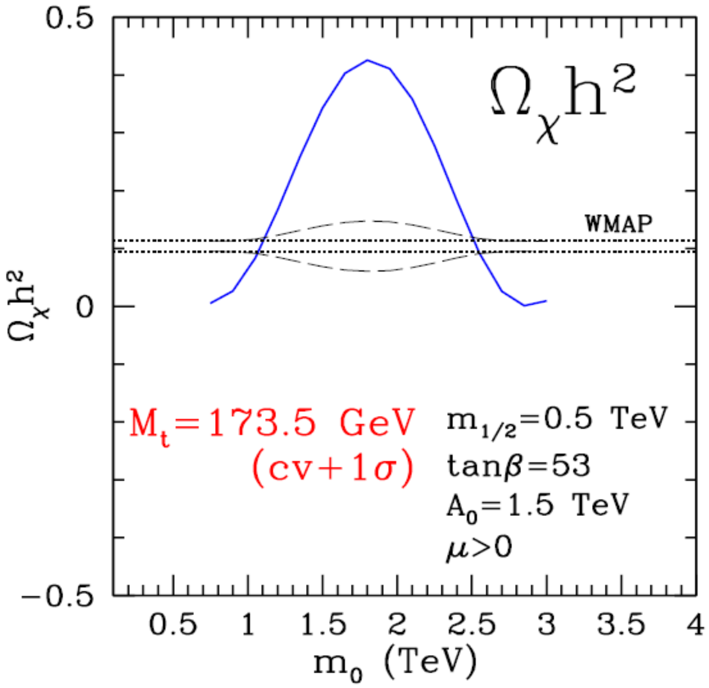


Ellis et al (2007)

Impact of nuisance parameters

Imperial College London

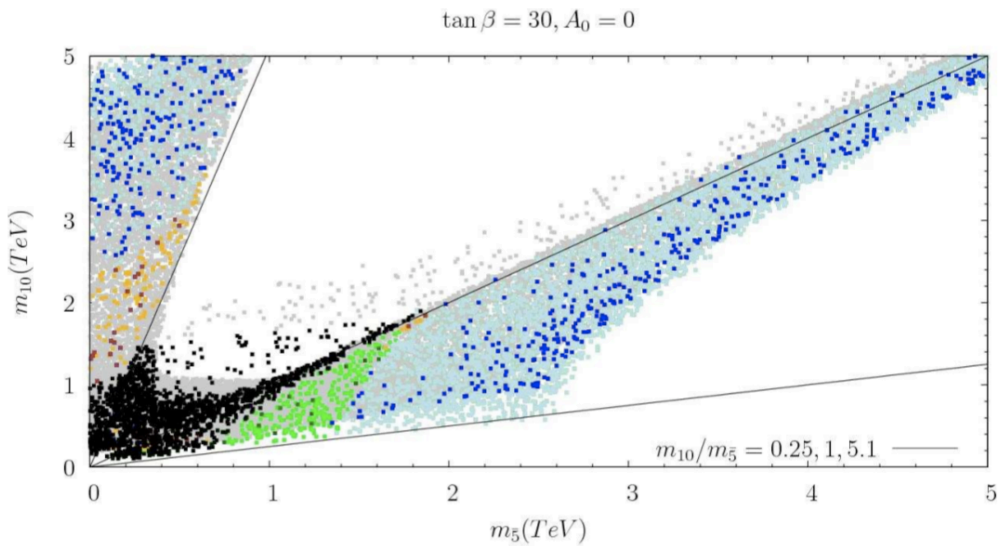




Roszkowki et al (2007)

Random scans

Imperial College London



Gogoladze et al (2008)

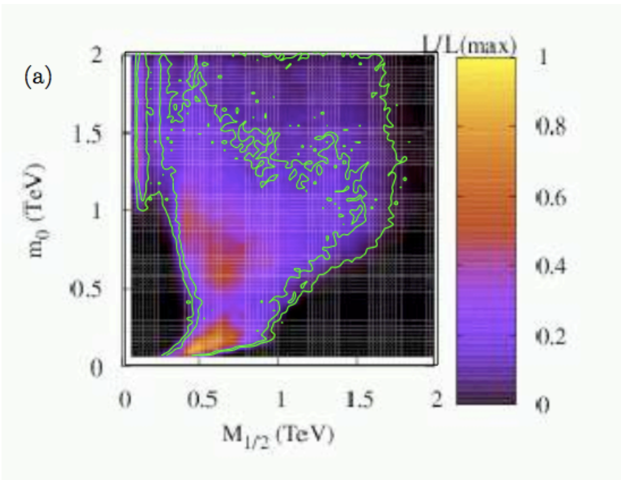
- Points accepted/rejected in a in/out fashion (e.g., 2-sigma cuts)
- No statistical measure attached to density of points
- No probabilistic interpretation of results possible
- Inefficient/Unfeasible in high dimensional parameters spaces (N>3)

Bayesian parameter inference

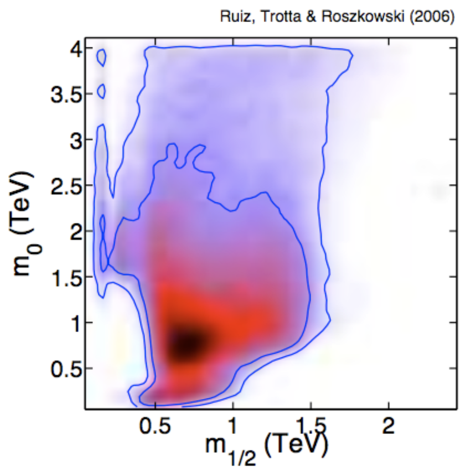
The Bayesian approach

Imperial College London

- Bayesian approach led by two groups (early work by Baltz & Gondolo, 2004):
- Ben Allanach (DAMPT) et al (Allanach & Lester, 2006 onwards, Cranmer, and others)
- Ruiz de Austri, Roszkowski & RT (2006 onwards)
  SuperBayeS public code (available from: superbayes.org)
  + Feroz & Hobson (MultiNest), + Silk (indirect detection), + Strigari (direct detection), + Martinez et al (dwarfs), + de los Heros (IceCube)



Allanach & Lester (2006)



Ruiz de Austri, Roszkowski & RT (2006)

See also, e.g.: Ellis et al (2004, 2005, 2006), Baltz & Gondolo (2004), Buchmuller et al (2008)

Roberto Trotta

Key advantages

Imperial College London

- **Efficiency:** computational effort scales ~ N rather than k^N as in grid-scanning methods. Orders of magnitude improvement over previously used techniques.
- **Marginalisation:** integration over hidden dimensions comes for free.
- **Inclusion of nuisance parameters:** simply include them in the scan and marginalise over them. **Notice:** nuisance parameters in this context must be well constrained using independent data.
- **Derived quantities**: probabilities distributions can be derived for any function of the input variables (crucial for DD/ID/LHC predictions)

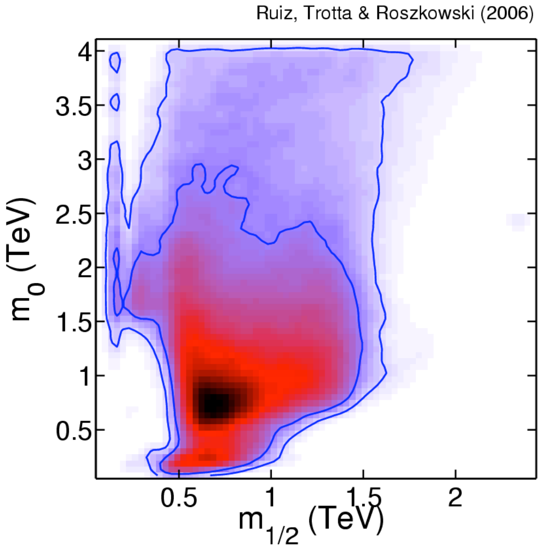
Bayesian pdf for an 8D model

Imperial College London

Experimental value: BR(b →sγ) = $(3.55 \pm 0.26)\times 10^{-4}$
Theoretical value SM: BR(b →sγ) = $(3.12 \pm 0.21)\times 10^{-4}$ (Misiak et al , 2006)
Previous SM value: BR(b →sγ) = $(3.60 \pm 0.30)\times 10^{-4}$

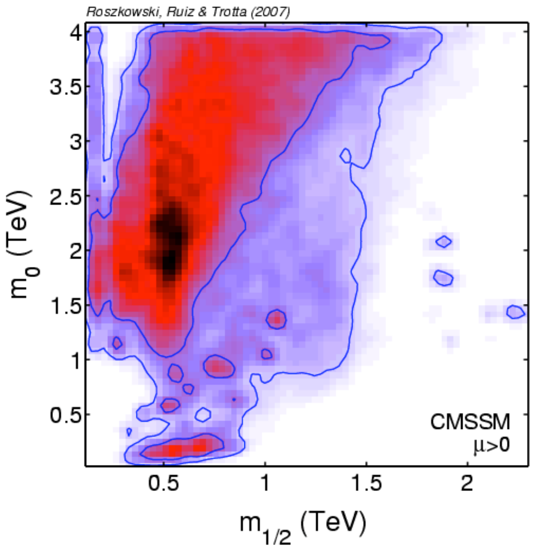
2006

(Ruiz de Austri et al 2006)

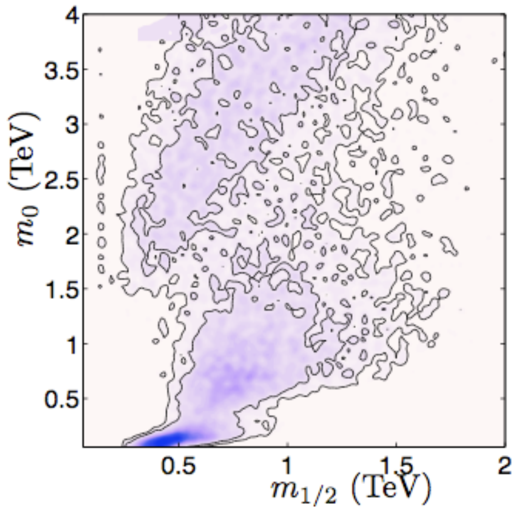


2007

(Roszkowski et al 2007)



(Feroz et al 2008)



including new SM BR(b →sγ)

Roberto Trotta

Bayes' theorem

Imperial College London

posterior

likelihood

prior

$$P(H|d,I) = \frac{P(d|H,I)P(H|I)}{P(d|I)}$$

evidence

H: hypothesis
d: data
I: external information

- **Prior**: what we know about H (given information I) before seeing the data
- **Likelihood**: the probability of obtaining data d if hypothesis H is true
- **Posterior:** our state of knowledge about H after we have seen data d
- **Evidence**: normalization constant (independent of H), crucial for model comparison

Continuous parameters

Imperial College London

$$P(\theta|d,I) = \frac{P(d|\theta,I)P(\theta|I)}{P(d|I)}$$

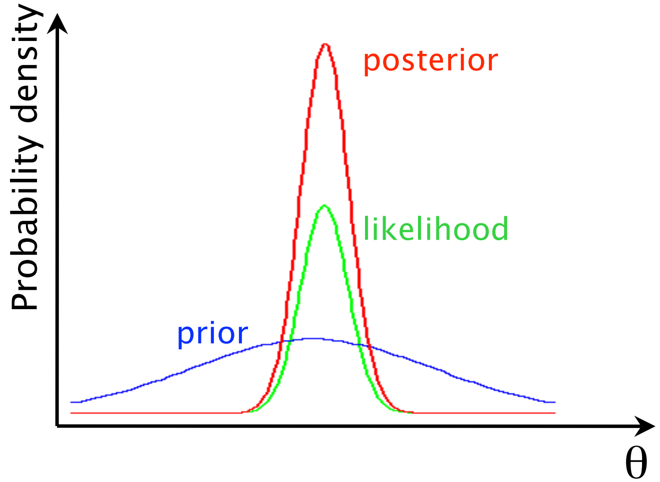
Bayesian evidence: average of the likelihood over the prior

$$P(d|I) = \int d\theta P(d|\theta,I)P(\theta|I)$$

For parameter inference it is sufficient to consider

$$P(\theta|d,I) \propto P(d|\theta,I)P(\theta|I)$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

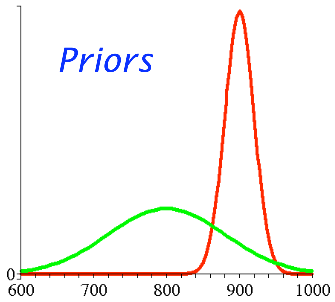


Roberto Trotta

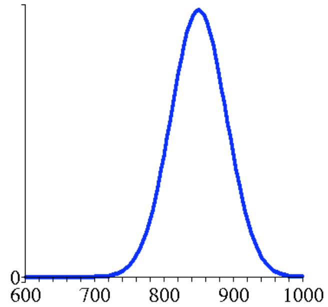
Prior dependence

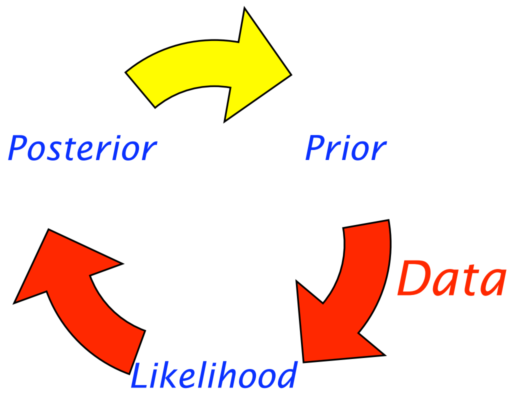
Imperial College London

- In parameter inference, prior dependence will **in principle** vanish for strongly constraining data.
  THIS IS CURRENTLY NOT THE CASE EVEN FOR THE CMSSM!

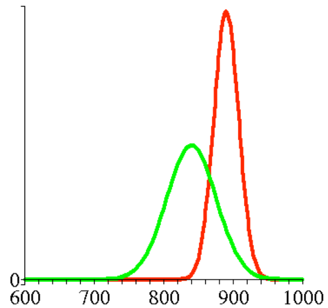


Likelihood (1 datum)

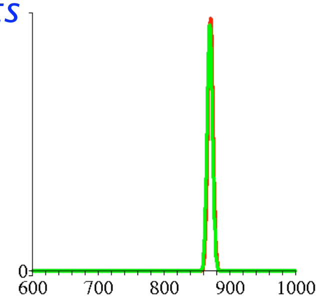




Posterior after 1 datum



Posterior after 100 data points



Roberto Trotta

The SuperBayeS package (superbayes.org)

Imperial College London

- Supersymmetry Parameters Extraction Routines for Bayesian Statistics
- Implements the CMSSM, but can be easily extended to the general MSSM
- Currently linked to SoftSusy 2.0.18, DarkSusy 4.1, MICROMEGAS 2.2, FeynHiggs 2.5.1, Hdecay 3.102. **New release (v 1.36) upcoming!**
- Includes up-to-date constraints from all observables
- Fully parallelized, MPI-ready, user-friendly interface à la cosmomc (thanks Sarah Bridle & Antony Lewis)
- Bayesian MCMC or grid scan mode, plotting routines.
  NEW: MULTI-MODAL NESTED SAMPLING (Feroz & Hobson 2008), efficiency increased by a factor 200. **A full 8D scan now takes 3 days on a single CPU** (previously: 6 weeks on 10 CPUs)

MCMC estimation

Imperial College London

$$P(\theta|d, I) \propto P(d|\theta, I)P(\theta|I)$$

- A Markov Chain is a list of samples θ_1, θ_2, θ_3,... whose density reflects the (unnormalized) value of the posterior

- A MC is a sequence of random variables whose *(n+1)*-th elements only depends on the value of the *n*-th element

- **Crucial property:** a Markov Chain converges to a stationary distribution, i.e. one that does not change with time. In our case, the posterior.

- From the chain, expectation values wrt the posterior are obtained very simply:

$$\langle\theta\rangle = \int d\theta P(\theta|d)\theta \approx \frac{1}{N}\sum_i \theta_i$$

$$\langle f(\theta)\rangle = \int d\theta P(\theta|d)f(\theta) \approx \frac{1}{N}\sum_i f(\theta_i)$$

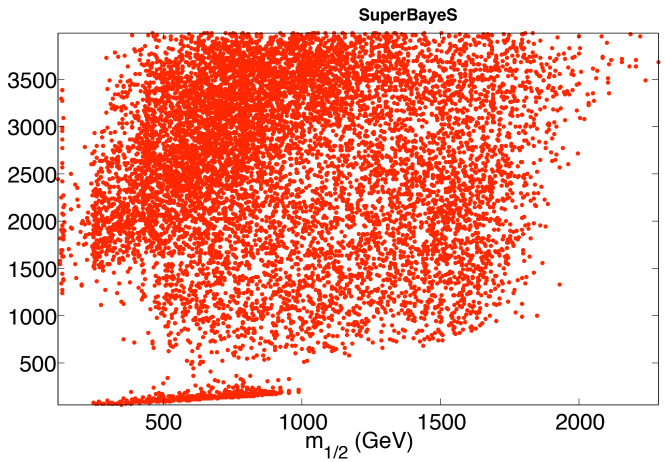
MCMC estimation

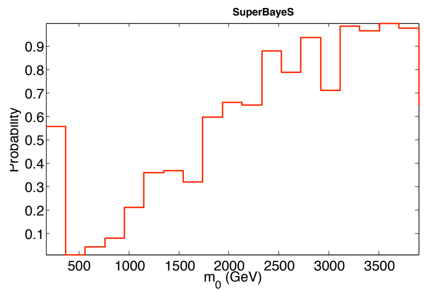
Imperial College London

- **Marginalisation becomes trivial:** create bins along the dimension of interest and simply count samples falling within each bins ignoring all other coordinates

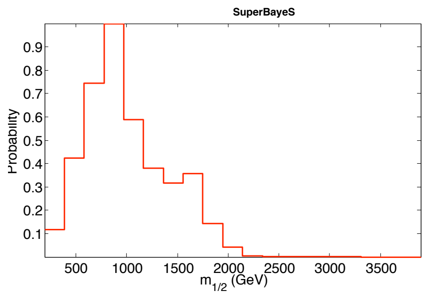
- Examples (from **superbayes.org**) :

2D distribution of samples from joint posterior





1D marginalised posterior (along y)

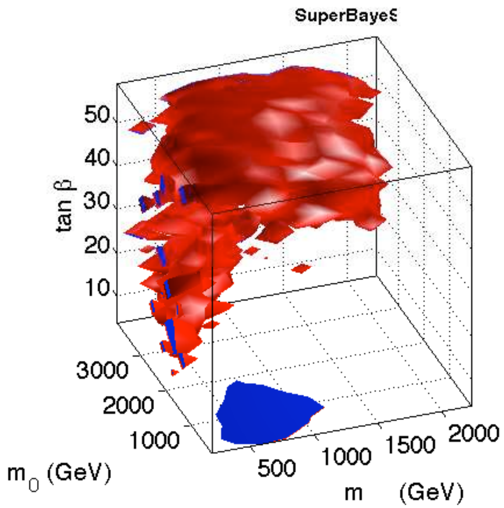


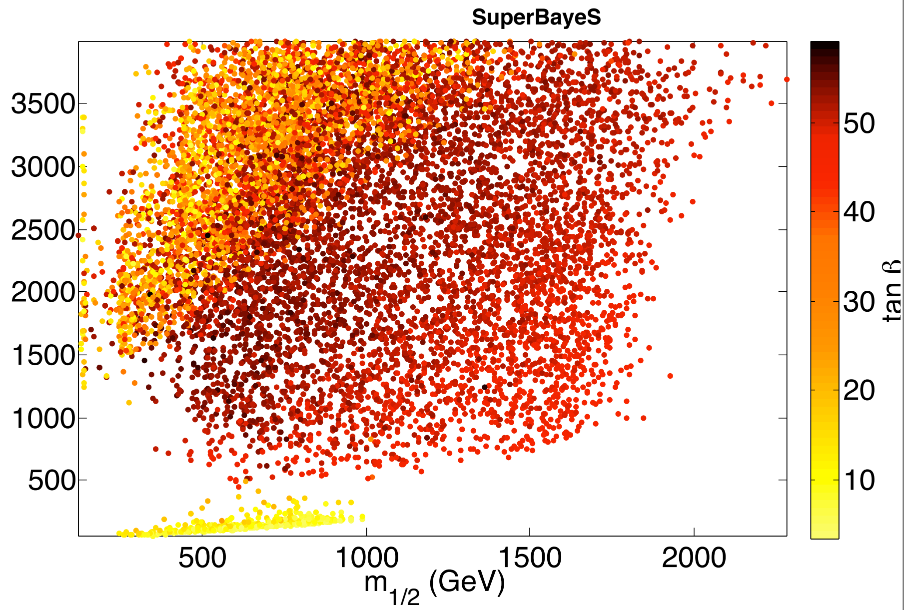
1D marginalised posterior (along x)

Roberto Trotta

Fancier stuff

Imperial College London

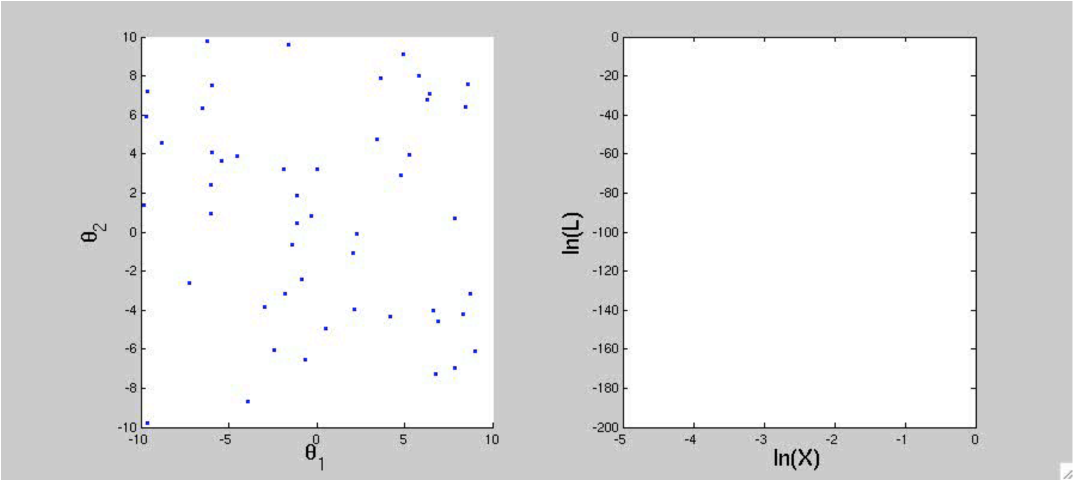




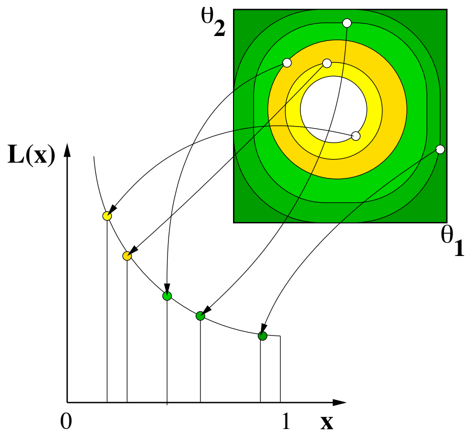
Roberto Trotta

Nested sampling

Imperial College London



(animation courtesy of David Parkinson)



An algorithm originally aimed primarily at the Bayesian evidence computation (Skilling, 2006):

$$X(\lambda) = \int_{\mathcal{L}(\theta)>\lambda} P(\theta) d\theta$$

$$P(d) = \int d\theta \mathcal{L}(\theta) P(\theta) = \int_0^1 X(\lambda) d\lambda$$

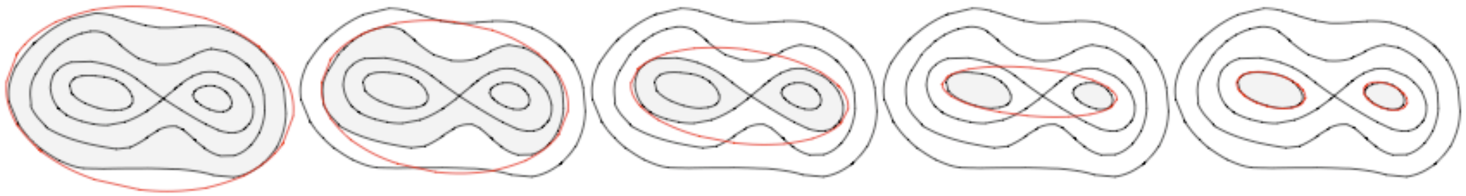
Feroz et al (2008), *arxiv: 0807.4512*, Trotta et al (2008), *arxiv: 0809.3792*

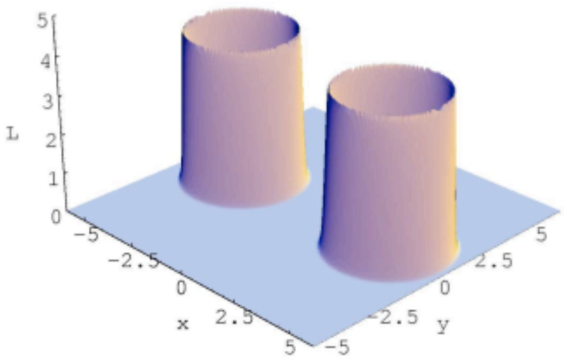
Roberto Trotta

The MultiNest algorithm

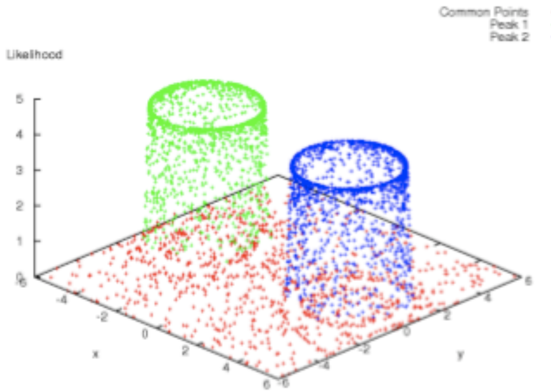
Imperial College London

- MultiNest: Also an extremely efficient sampler for multi-modal likelihoods!
  Feroz & Hobson (2007), RT et al (2008), Feroz et al (2008)





Target Likelihood



Sampled Likelihood

Roberto Trotta

Global CMSSM constraints

See also recent works by Ellis et al (2004, 2005, 2006), Baltz & Gondolo (2004), Buchmuller et al (2008), Allanach & collaborators (2006, 2007,2008)

Priors

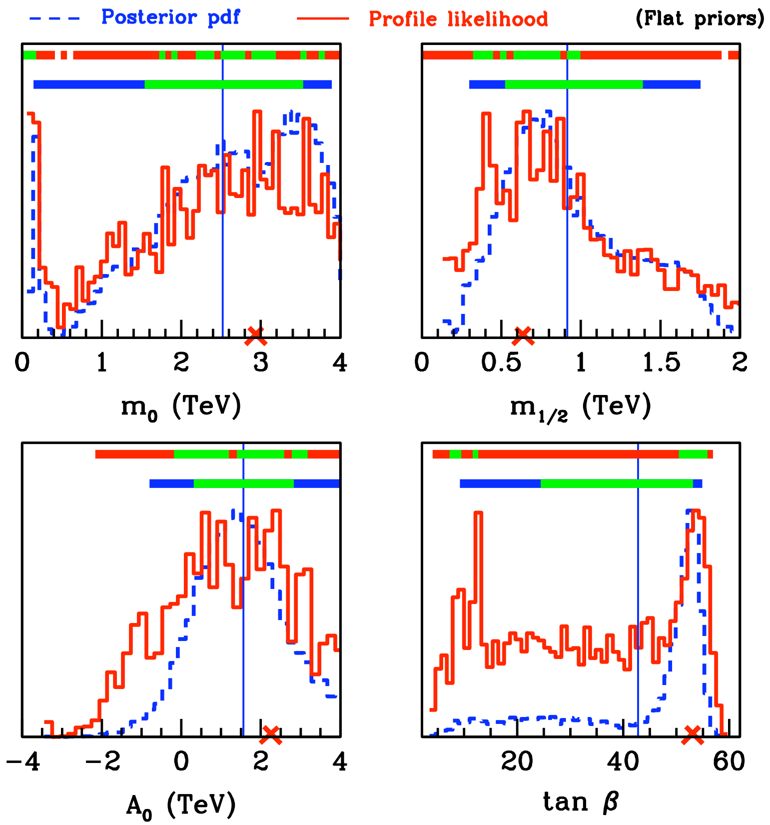
Imperial College London

- There is a vast literature on priors: Jeffreys', conjugate, non-informative, ignorance, reference, ...
- In simple problems, "good" priors are dictated by symmetry properties
- "Flat priors" (i.e., uniform in the model's parameters) are often uncritically adopted as default by cosmologists/physicists: **they do not necessarily reflect indifference/ ignorance.** Beware: in large dimensions, most of the volume of a sphere is near its surface!
- For the SM parameters we adopt flat priors (with cutoff well beyond the region where the likelihood is non-zero). This is largely unproblematic as the nuisance parameters are directly constrained by the likelihood hence the posterior is dominated by the likelihood
- Priors for the CMSSM parameters: this is a difficult issue

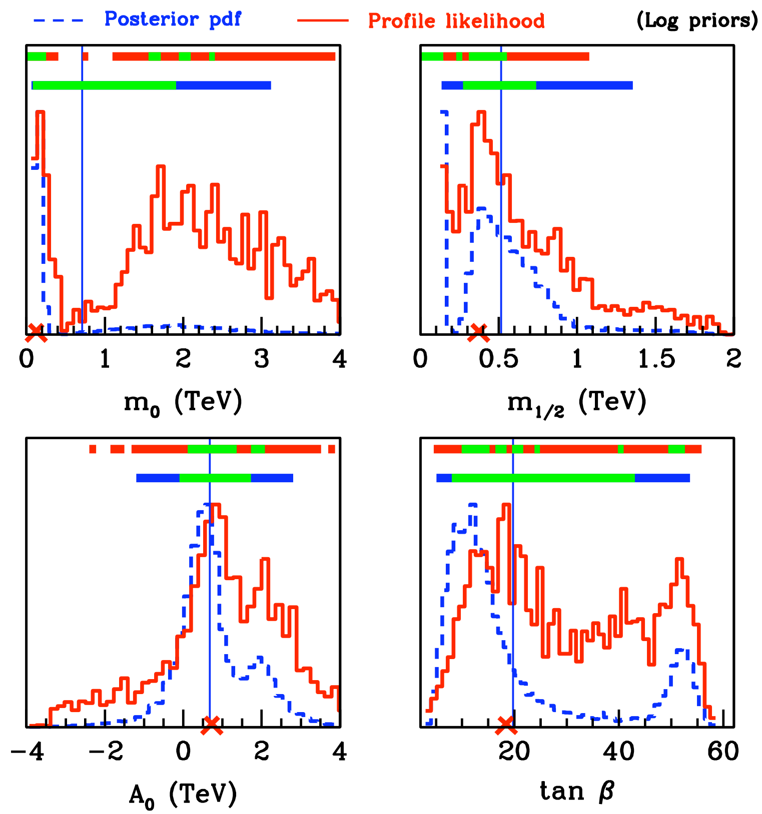
Parameter inference (all data included)

Imperial College London

Flat priors



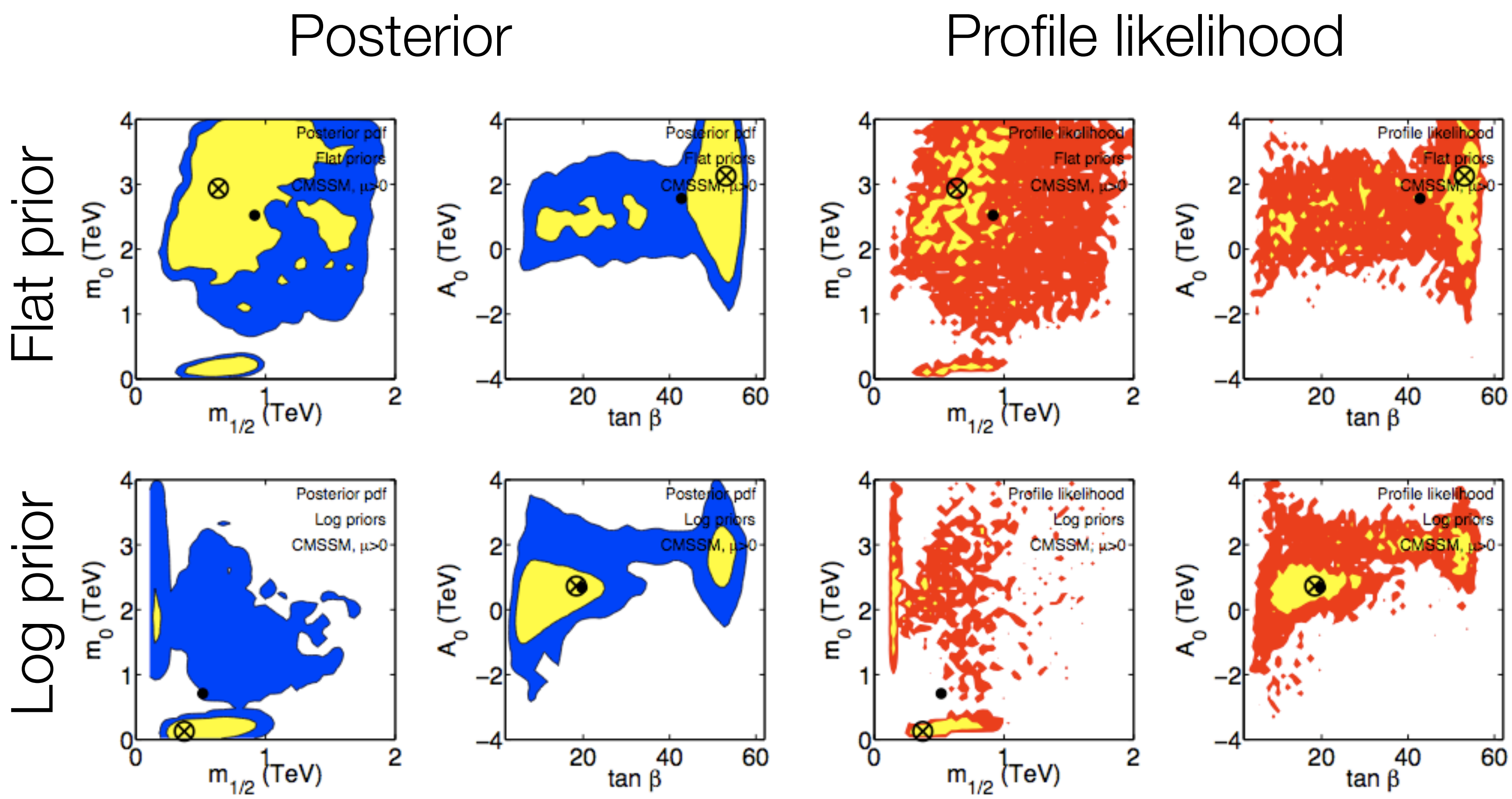
Log priors



Roberto Trotta

2D posterior vs profile likelihood

Imperial College London

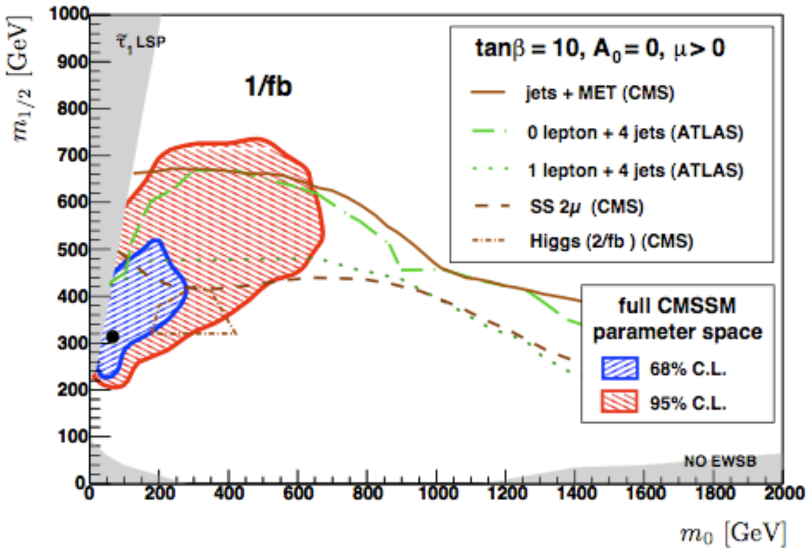


Roberto Trotta

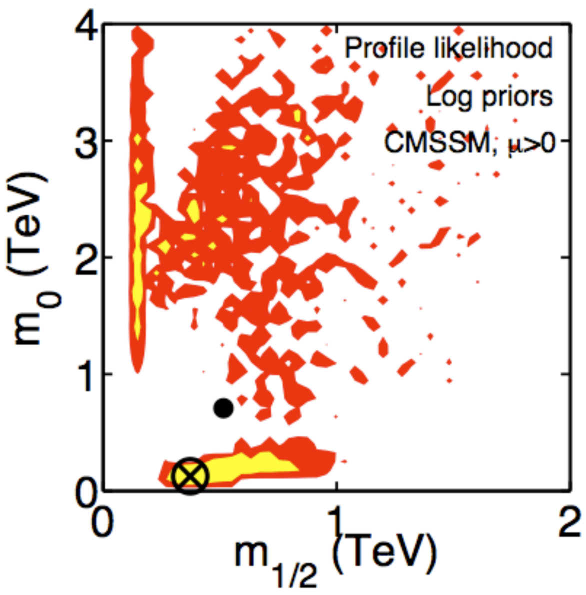
Comparison with other work

Imperial College London

Buchmuller et al (2008)



Trotta et al (2008)

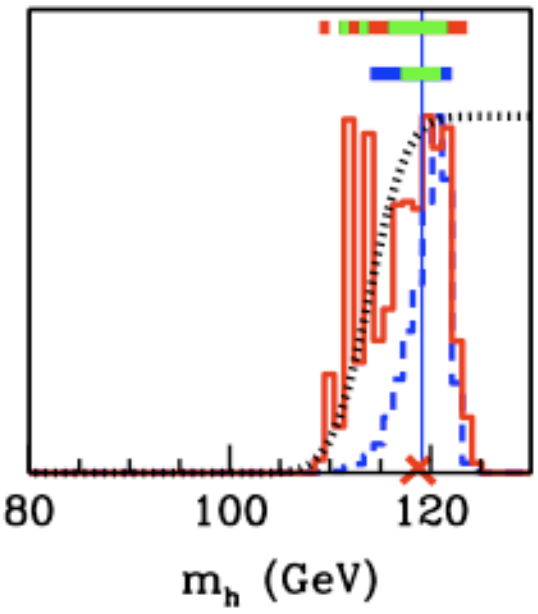


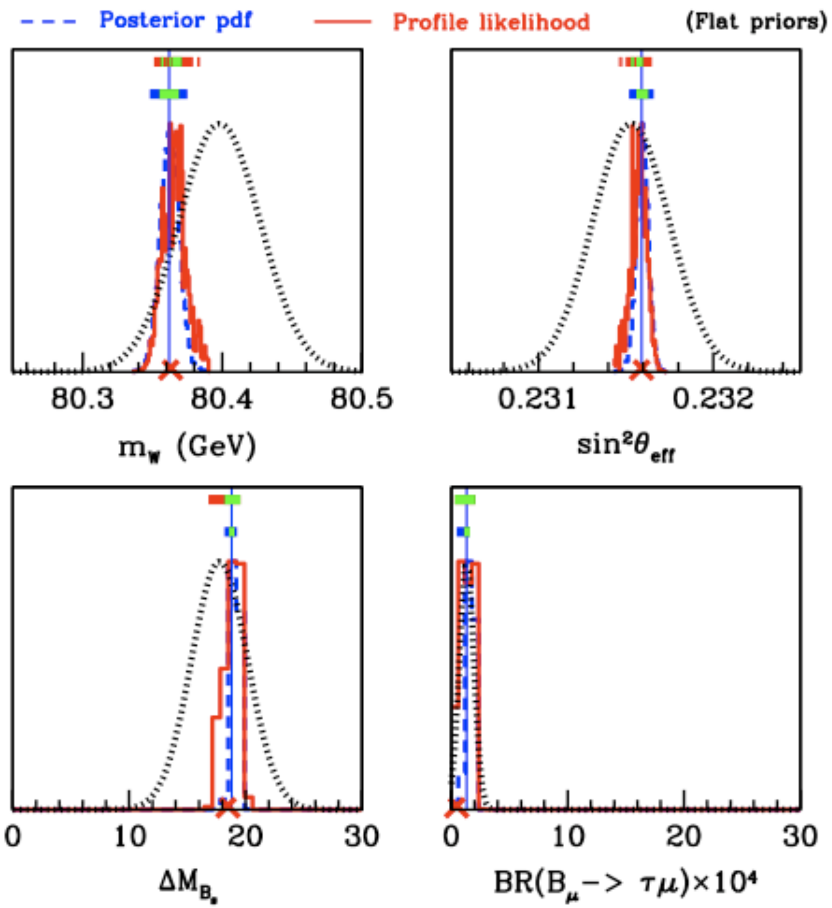
Roberto Trotta

Posterior for some observables

Imperial College London

Higgs prospects



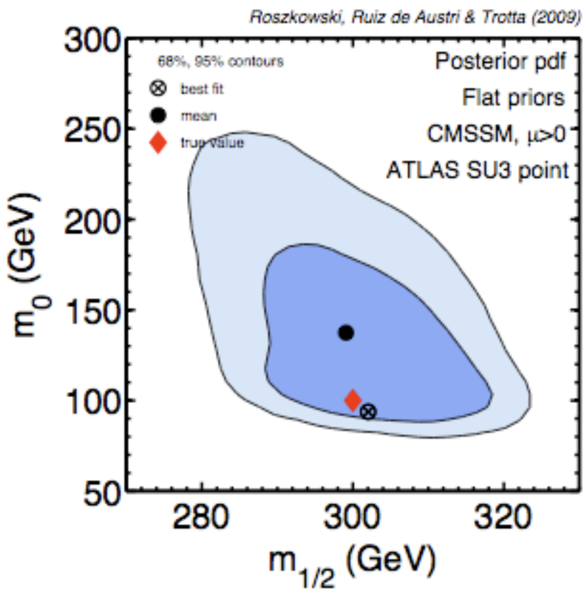


Roberto Trotta

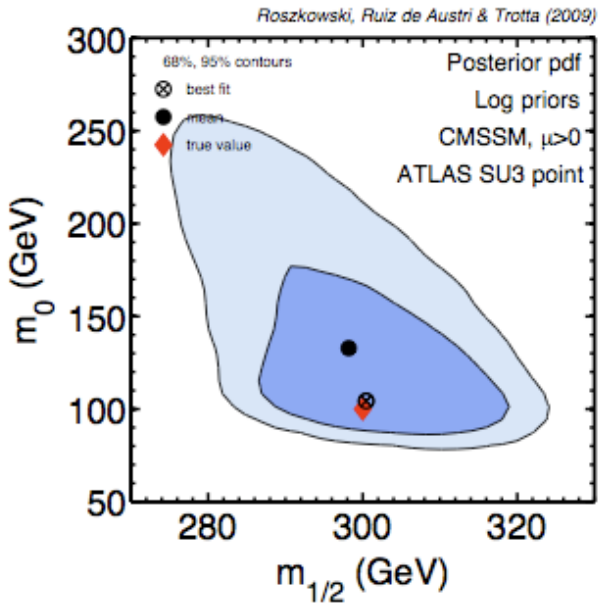
ATLAS will solve the prior dependency

- Projected constraints from ATLAS, (dilepton and lepton+jets edges, 1 fb^{-1} luminosity)

Flat prior



Log prior



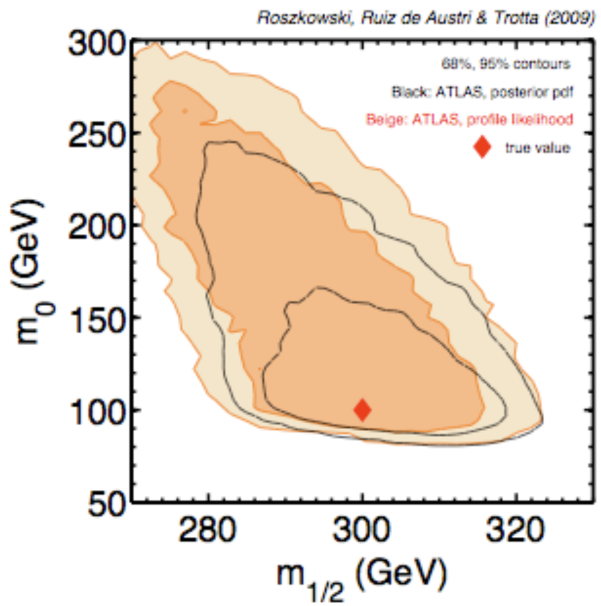
Roszkowski, Ruiz & RT (2009, 0907.0594)

Residual dependency on the statistics

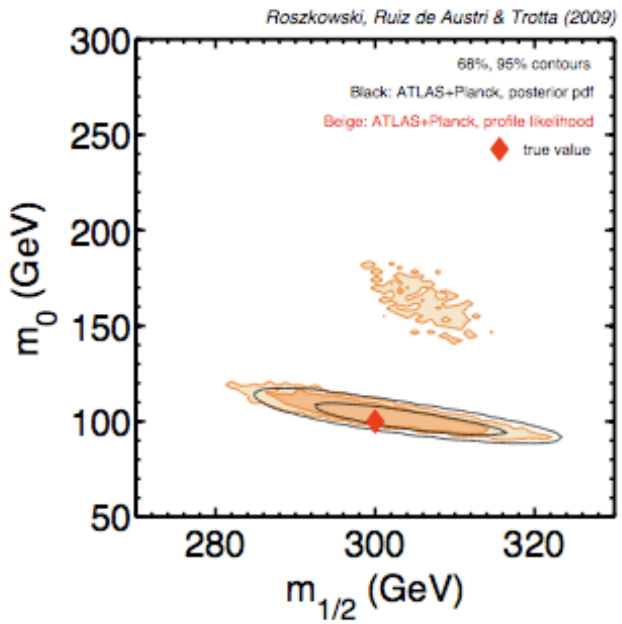
Imperial College London

- Marginal posterior and profile likelihood will remain somewhat discrepant using ATLAS alone. Much better agreement from ATLAS+Planck CDM determination.

ATLAS alone



ATLAS+Planck



Direct and indirect detection prospects

Astrophysical probes

Imperial College London

- Direct detection: underground detectors looking for nuclear recoils from WIMP scattering. **It is fundamental to account for the uncertainty in the local WIMP distribution.**

- Indirect detection: detection of annihilation products from WIMP-WIMP annihilation.

  - **Gamma ray** (galactic centre, galactic halo, diffuse extragalactic sources, nearby dwarf galaxies)

  - **Antimatter** (positrons, anti-proton) from local clumps

  - **Neutrinos** from the center of the Sun/Earth.

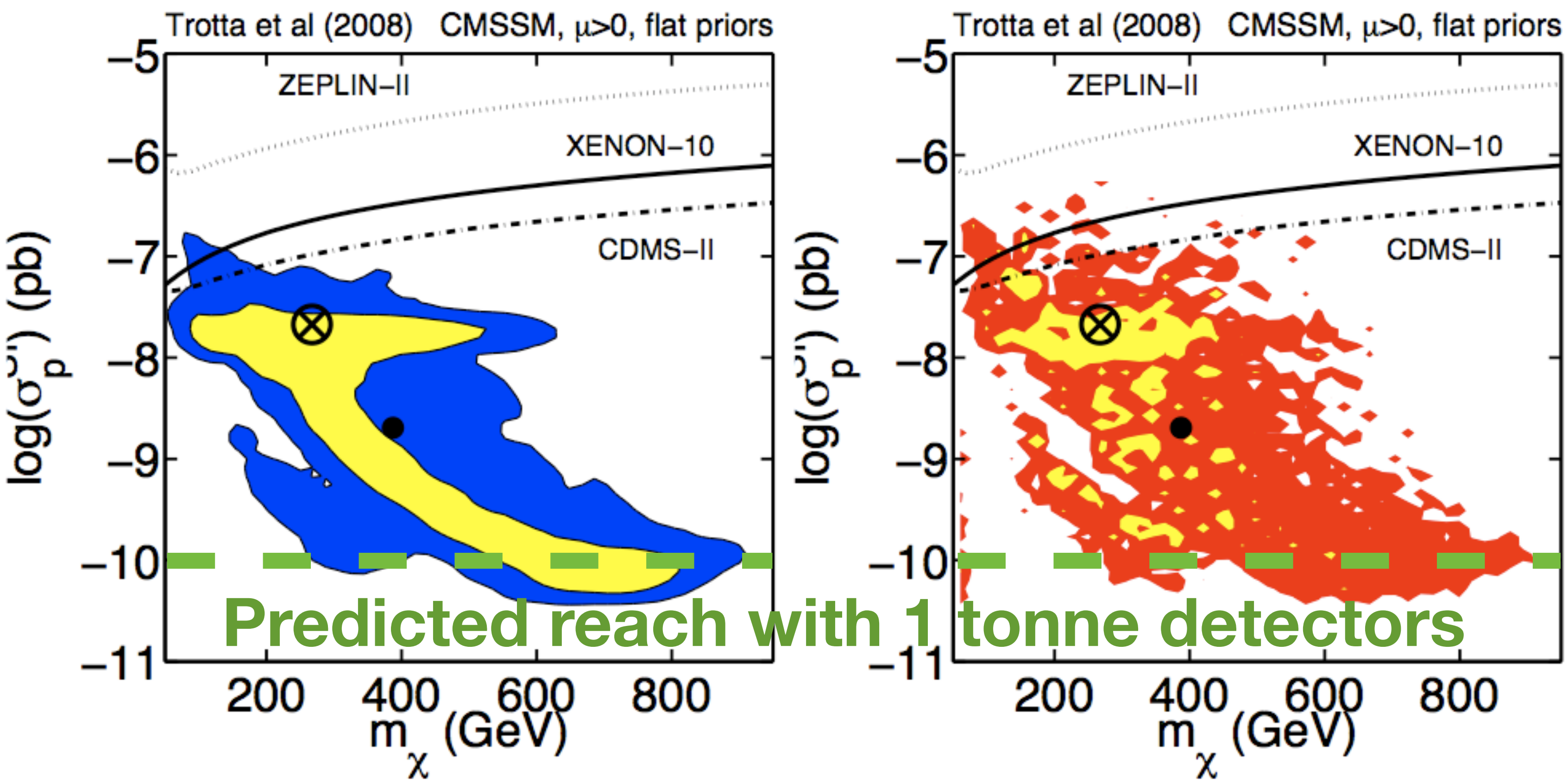
  - In all cases: it is fundamental to include a **modeling of background sources**. For gamma ray and neutrinos the unknown branching ratios have to be estimated simultaneously **(bias!).**

Direct detection prospects

Imperial College London

R. Trotta, F. Feroz, M.P. Hobson, R. Ruiz de Austri and L. Roszkowski, 0809.3792

Bayesian pdf | Profile likelihood



Predicting the gamma ray flux

Imperial College London

Differential flux: $\frac{d\Phi_\gamma}{dE_\gamma} \propto \sum_i \frac{\langle \sigma_i v \rangle}{m_\chi^2} \frac{dN_\gamma^i}{dE_\gamma} \int \rho_\chi^2 dl$

particle physics — astrophysics

DM density profile:

$$\rho_\chi(r) = \rho_0 \frac{(r/r_0)^{-\gamma}}{[1+(r/a)^\alpha]^{\frac{\beta-\gamma}{\alpha}}} [1 + (r/a)^\alpha]^{\frac{\beta-\gamma}{\alpha}}$$

$$J(\Psi) = \int_{\rm los} dl \rho_\chi^2(r(l,\Psi)) \qquad \bar{J} = \frac{1}{\Delta\Omega} \int_{\Delta\Omega} J(\Psi) d\Omega$$

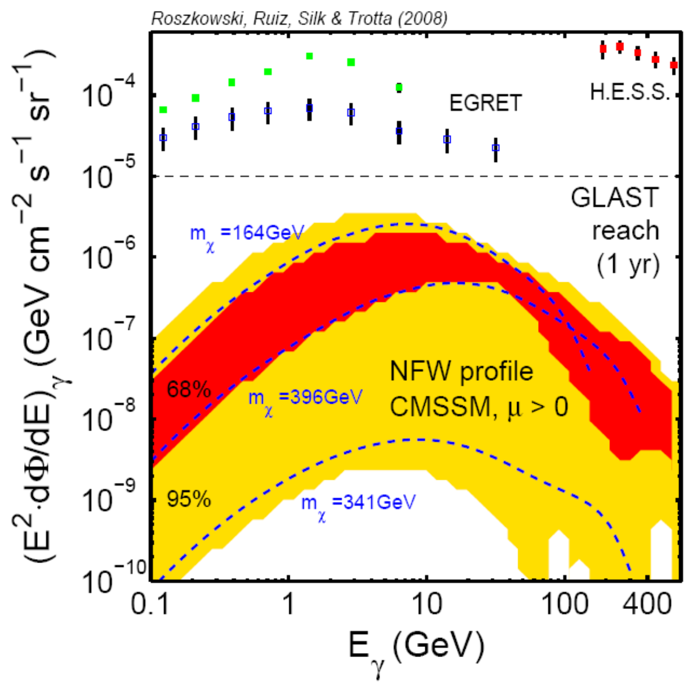
Halo model	a (kpc)	α	β	γ	$\bar{J}(10^{-3}\text{sr})$	$\bar{J}(10^{-5}\text{sr})$
isothermal cored	3.5	2	2	0	30.35	30.40
NFW	20.0	1	3	1	1.21×10^3	1.26×10^4
NFW+ac	20.0	0.8	2.7	1.45	1.25×10^5	1.02×10^7
Moore	28.0	1.5	3	1.5	1.05×10^5	9.68×10^6
Moore+ac	28.0	0.8	2.7	1.65	1.59×10^6	3.12×10^8

Trotta

Predictions for Fermi in the CMSSM

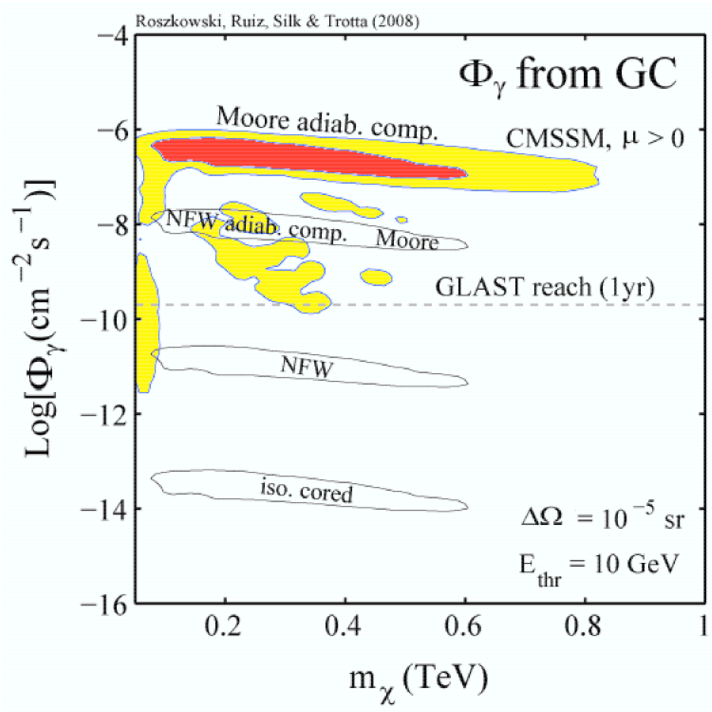
Imperial College London

Predicted gamma-ray spectrum probability distribution from the galactic center at Fermi resolution



Roszkowski, Ruiz, Silk & RT (2008)

Predicted gamma-ray flux above 10 GeV at Fermi resolution



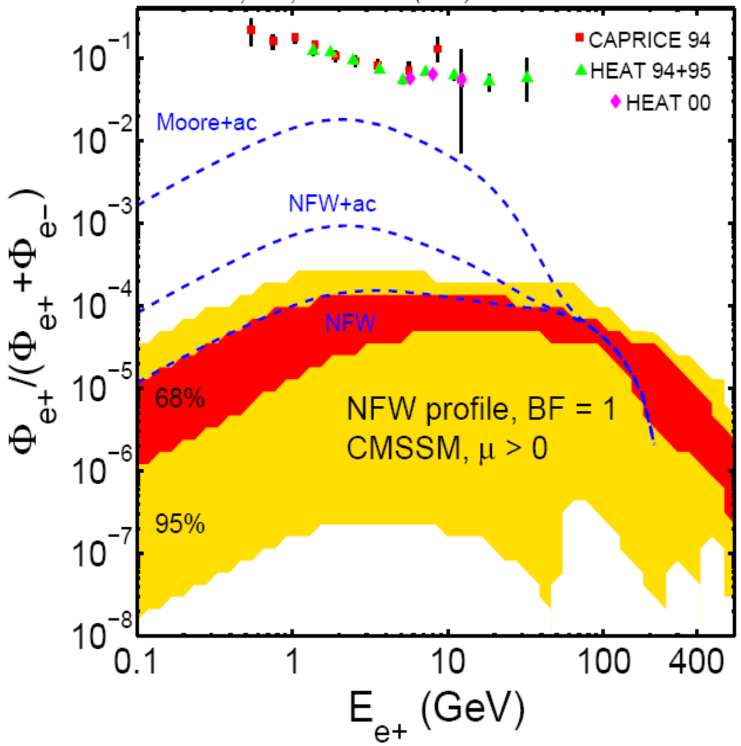
Roberto Trotta

Predictions for the positrons spectrum

Imperial College London

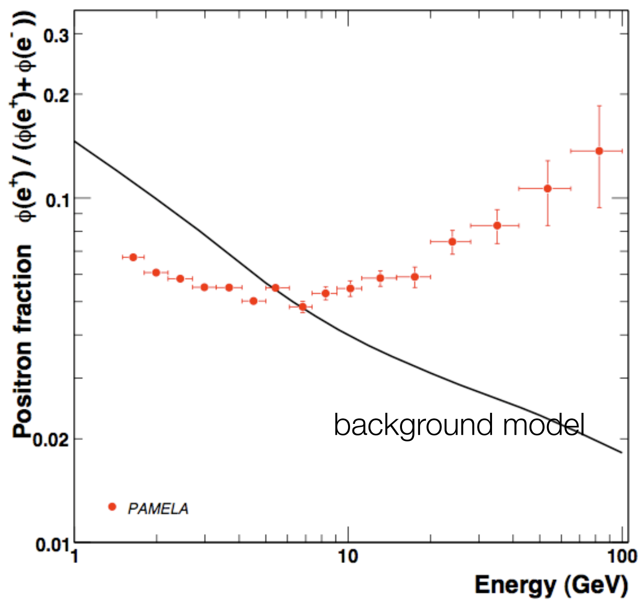
Notice: this is for a fixed choice of propagation parameters!

Roszkowski, Ruiz, Silk & Trotta (2008)



PAMELA data

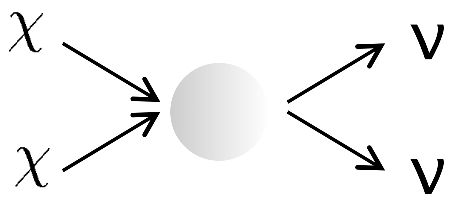
(Adriani et al 2008)

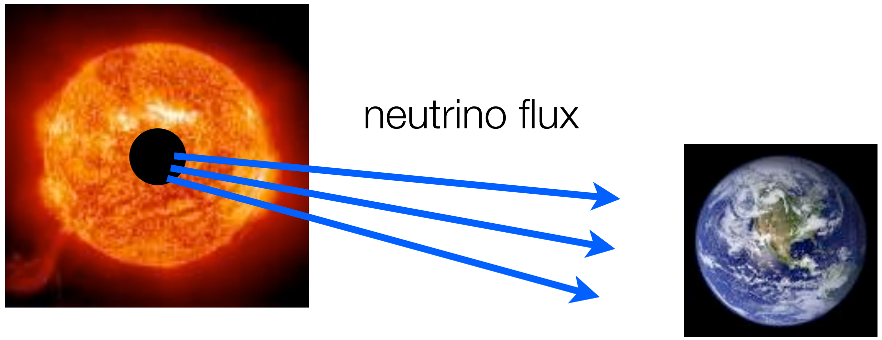


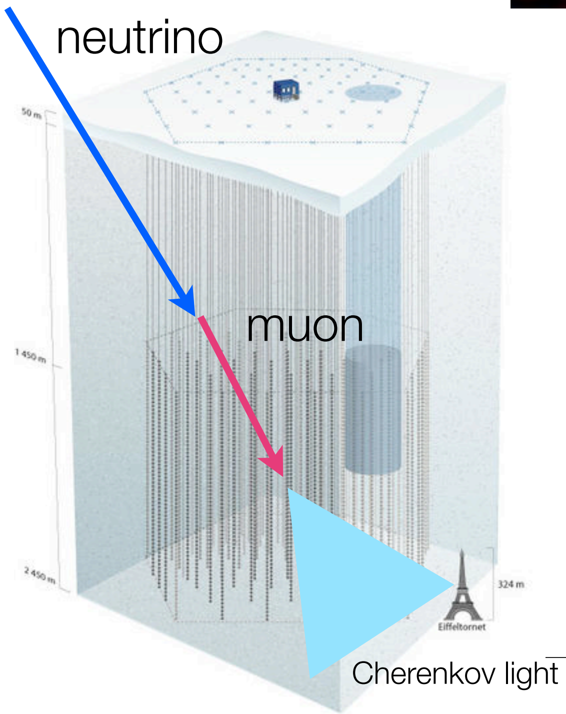
Roberto Trotta

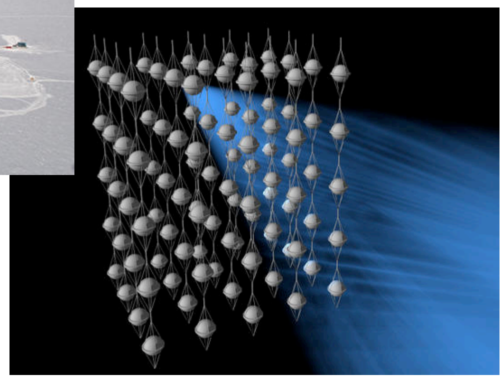
Neutrinos from WIMP annihilations in the Sun

Imperial College London





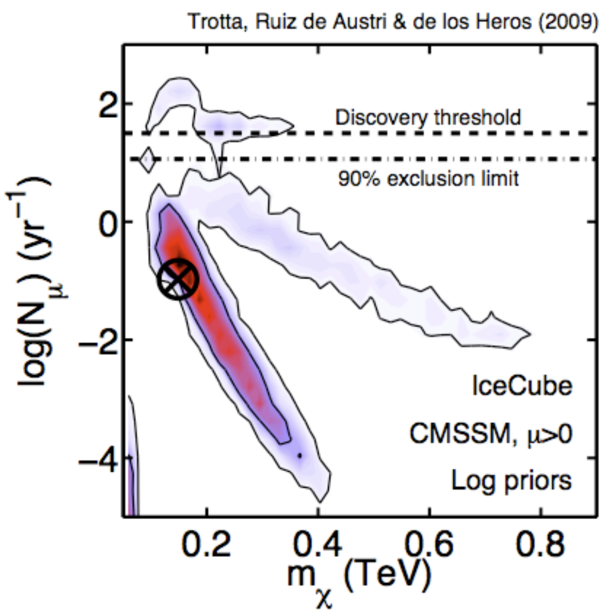


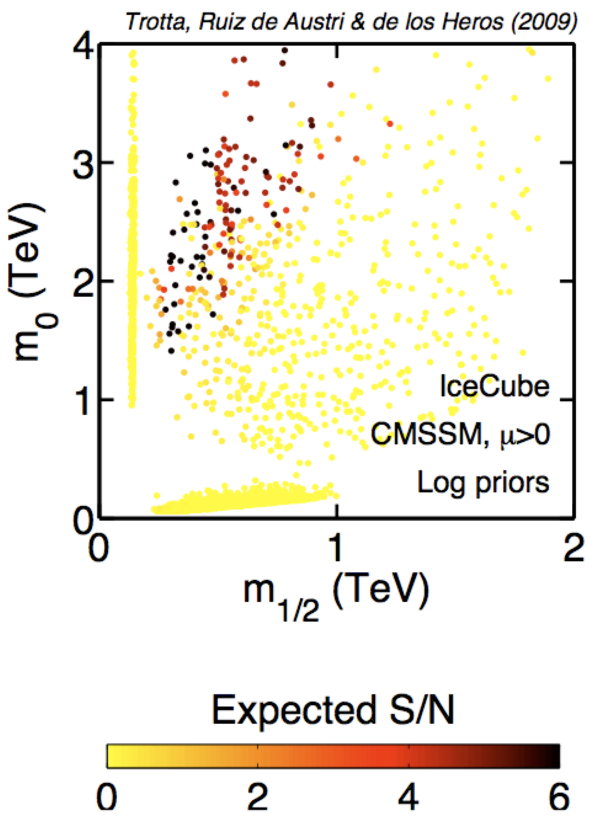


Detection prospects with IceCube

Imperial College London

- In the context of the CMSSM, the final configuration of IceCube (with 80 strings) has between 2% and 12% probability of achieving a 5-sigma detection





Trotta, Ruiz de Austri & de los Heros (0906.0366)

otta

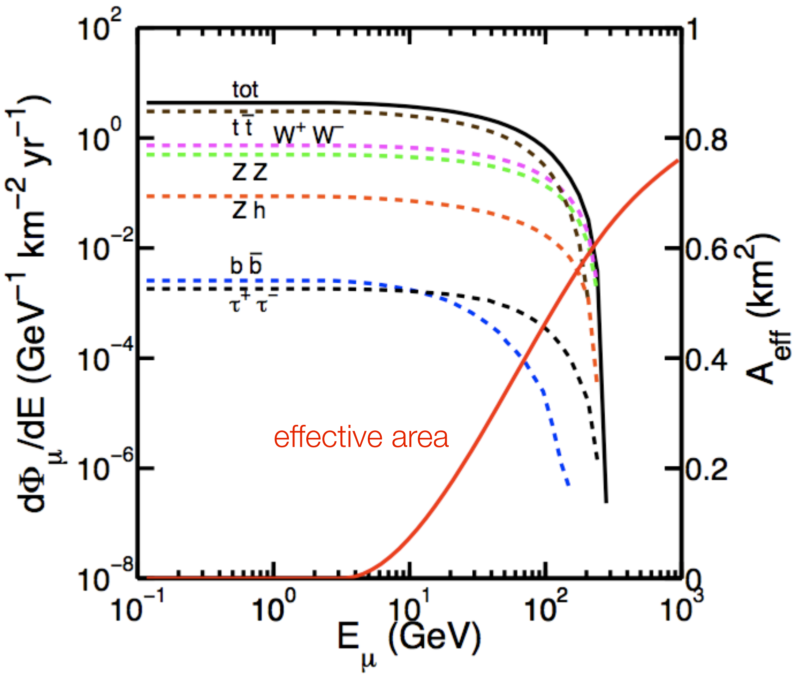
Bias from assuming the wrong final state

Imperial College London

Muon events rate = $N_\mu = \int_{E_\mu^{\rm th}}^{+\infty} dE_\mu A_\mu^{\rm eff}(E_\mu) \frac{d\Phi_\mu}{dE_\mu}$

effective area

total spectrum - sum over final states



Total events = 34 1/yr

BR's for dominant channels		Events/yr assuming only 1 channel
$\mathrm{BR}(\chi\chi \to b\bar{b})$	2.1×10^{-3}	
$\mathrm{BR}(\chi\chi \to t\bar{t})$	0.717	
$\mathrm{BR}(\chi\chi \to W^+W^-)$	0.162	
$\mathrm{BR}(\chi\chi \to ZZ)$	0.094	
Muon events fraction for IceCube		
$N_{b\bar{b}}/N_\mu$	5×10^{-5}	0.8
$N_{t\bar{t}}/N_\mu$	0.46	
$N_{W^+W^-}/N_\mu$	0.30	64
N_{ZZ}/N_μ	0.21	

Roberto Trotta

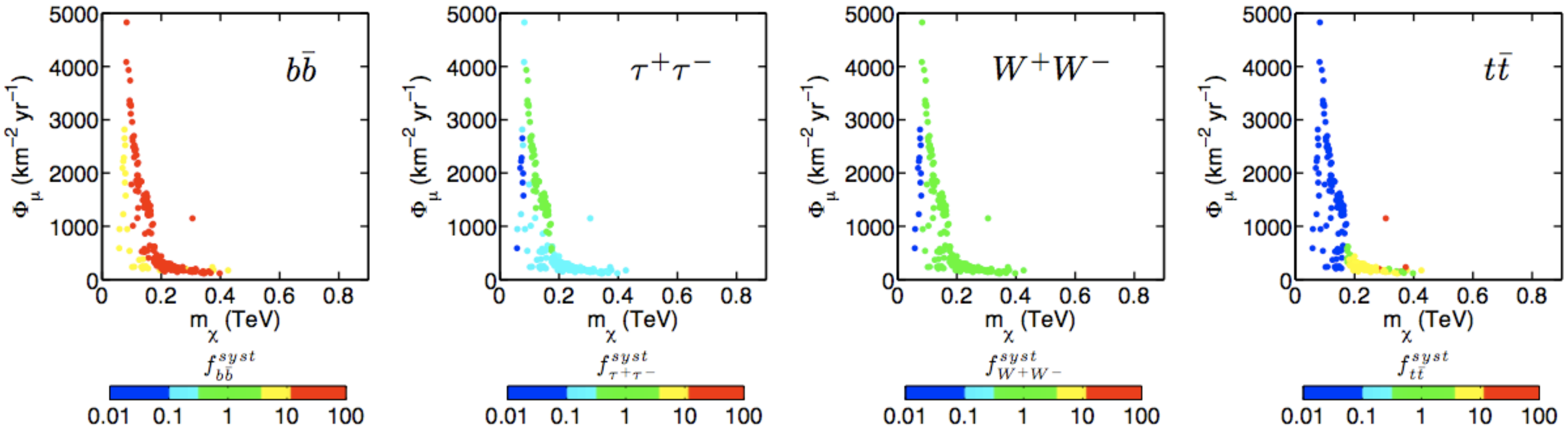
Bias from assuming the wrong final state

Imperial College London

- In general, the systematic error from assuming only 1 dominating channel is given by

$$f_i^{\rm syst} = \frac{{\rm BR}(\chi\chi \to i)}{N_i/N_\mu}$$

Systematic bias from assuming single-channel domination (IceCube)



Trotta, Ruiz de Austri & de los Heros (0906.0366)

Roberto Trotta

Some uses for the Bayesian evidence

Model comparison

Imperial College London

- The Bayesian evidence is the prime tool for Bayesian model comparison. It automatically includes the notion of "Occam's razor" (see RT, 0803.4089)

$$P(d|M) = \int_\Omega d\theta P(d|\theta, M)P(\theta|M)$$

Model's posterior probability : $$P(M|d) = \frac{P(d|M)P(M)}{P(d)}$$

Two models (e.g. sgn(μ)=±1): $$\frac{P(M_0|d)}{P(M_1|d)} = \frac{P(d|M_0)}{P(d|M_1)}\frac{P(M_0)}{P(M_1)}$$

Posterior odds = Bayes factor × prior odds

Feroz et al (2008)

Prior	"2 TeV"		"4 TeV"	
	flat	log	flat	log
$\log \Delta E$ (our determination)	2.7 ± 0.1	4.1 ± 0.1	1.8 ± 0.1	3.2 ± 0.1
P_+/P_- (our determination)	15.6 ± 1.1	61.6 ± 1.1	5.9 ± 1.1	24.0 ± 1.1

sgn(μ)=+1 between 6 and 60 times more probable than sgn(μ)=-1

Information content of observables

Imperial College London

- Which observable is the most constraining for the CMSSM?

- The information content (i.e., constraining power) of each observable with respect to the model and prior can be quantified using the Kullback-Leibler divergence between prior and posterior:

$$D_{KL} = \int d\theta P(\theta|d) \ln \frac{P(\theta|d)}{P(\theta)} = -\ln P(d) - \langle \chi^2/2 \rangle$$

Information content = -ln(Bayesian evidence) - average chi-square/2

RT et al (2008)

Roberto Trotta

Constraining power of observables

Imperial College London

Constraints	Data points	Flat priors χ^2_{min}	Flat priors $\langle\chi^2\rangle$	Flat priors D_{KL}	Log priors χ^2_{min}	Log priors $\langle\chi^2\rangle$	Log priors D_{KL}
PHYS+NUIS	4	0.06	3.89	1.00	0.02	3.88	1.00
+CDM	5	0.05	4.36	3.22	0.10	4.32	2.59
+BSG	5	0.31	6.48	1.11	0.10	5.48	1.21
+GM2	5	0.27	11.55	1.35	0.13	6.38	1.20
+COLL+CDM	5+	0.28	4.60	3.20	0.15	5.04	2.98
+COLL+BSG	5+	0.99	6.82	1.11	0.45	6.54	1.24
+COLL+GM2	5+	1.79	13.43	1.10	0.17	9.92	1.49
+COLL+CDM+BSG	6+	0.75	7.15	3.36	0.68	7.72	3.29
+COLL+CDM+GM2	6+	0.62	9.24	2.90	0.43	7.49	3.23
+COLL+CDM+BSG+GM2	7+	6.27	15.83	3.48	4.67	14.89	3.39
ALL but GM2	10+	3.51	9.45	3.42	3.22	9.51	3.28
ALL but CDM	10+	12.17	18.86	1.10	4.14	18.30	1.24
ALL	11+	13.51	19.29	3.38	11.90	18.41	3.26

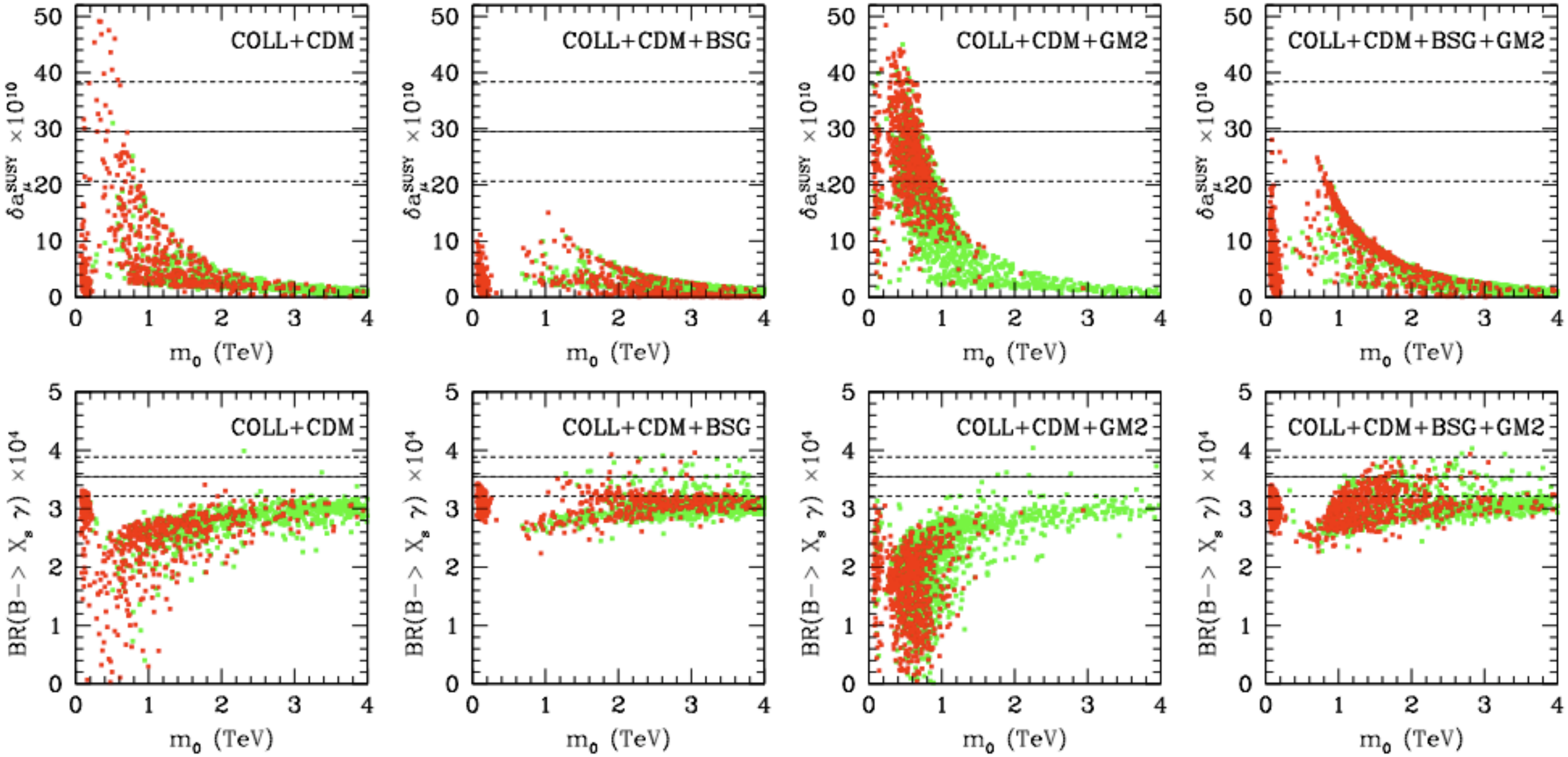
Cosmology provides 80% for flat priors (95% for log priors) of the total constraining power on the CMSSM

Roberto Trotta

Data consistency

Imperial College London

Tension between the anomalous magnetic moment and b -> sγ



RT et al (2008)

Roberto Trotta

Testing for data consistency

Imperial College London

- Evaluate the probability of “systematic inconsistency” between g-2 and and b → sγ within the CMSSM:

- Baysian model comparison:
  H_0: (g-2, b → sγ) compatible and described by a unique set of CMSSM parameters
  vs
  H_1: (g-2, b → sγ) systematically incompatible

$$R = \frac{\Pr(\mathbf{D}|H_1)}{\Pr(\mathbf{D}|H_0)} = \frac{\Pr(\mathbf{D}|H_1)}{\prod_i \Pr(D_i|H_0)}.$$

Feroz et al (2008) find that $|\ln R| < 1.0$, hence no significant evidence for tension (but no conclusive evidence for H_0, either)

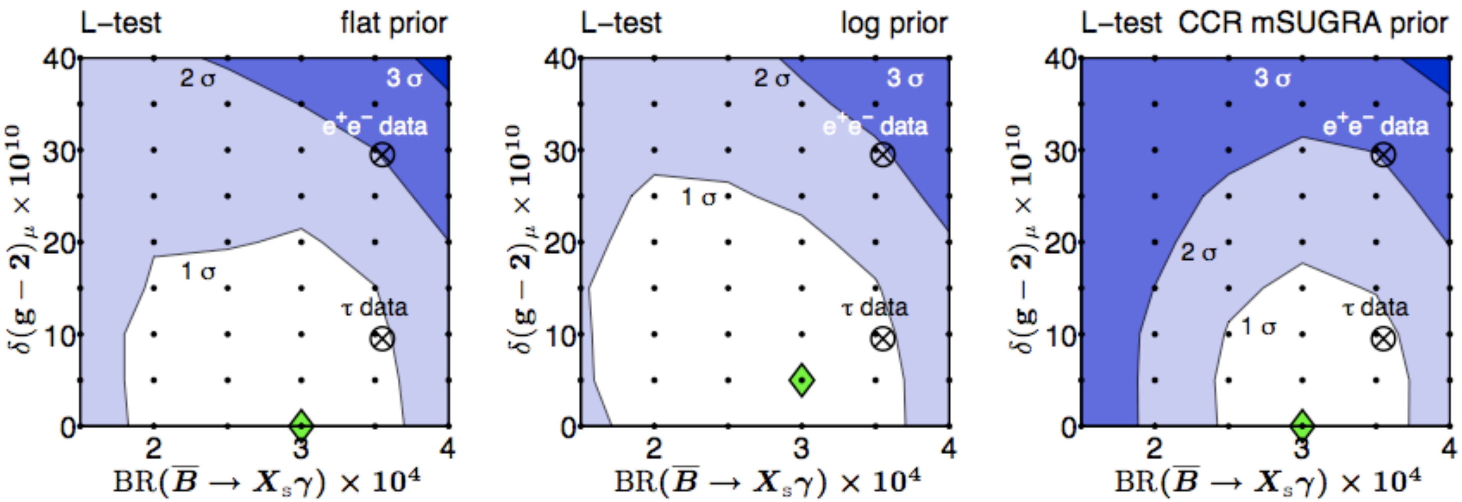
Roberto Trotta

Predictive data distributions

Imperial College London

- What is the probability of observing (g-2, b → sγ) given the CMSSM and all other constraints? Compute the predictive data distribution (i.e., conditional evidence) for D^{obs} given the other observations D

$$\mathscr{L}(\mathscr{D}^{\rm obs}|D) \equiv \frac{p(\mathscr{D}^{\rm obs}|D)}{p(\mathscr{D}^{\rm max}|D)} = \frac{p(\mathscr{D}^{\rm obs}, D)}{p(\mathscr{D}^{\rm max}, D)}$$



Feroz, Hobson, Ruiz, Roszkowski & RT, 2009, 0903.2487

Roberto Trotta

Conclusions

Imperial College London

- SUSY phenomenology provides a timely and challenging problem for parameter inference and model selection. A considerably harder problem than cosmological parameter extraction!
- Bayesian advantages: higher efficiency, inclusion of nuisance parameters, predictions for derived quantities, model comparison
- CMSSM only a case study. There are several models around (NUHM, pMSSM, ... -> more free parameters) that will need to be analyzed as soon as new data flow in
- Currently, even the CMSSM is somewhat underconstrained: ATLAS+Planck will take us to “statistics nirvana” (Bob Cousins)
- The Bayesian evidence can be used for model selection, data consistency checks and prediction. Also useful to quantify constraining power of the data.