Bayesian推断及其在宇宙线研究中的应用

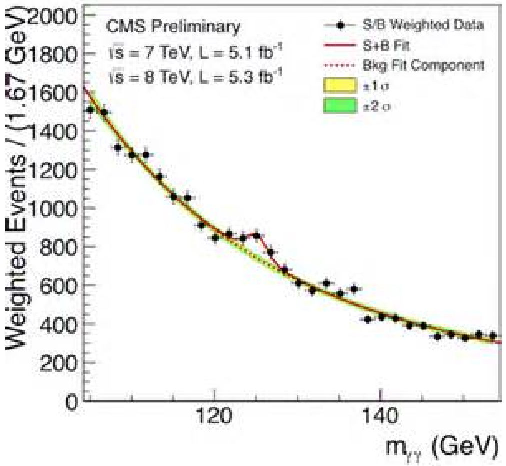
袁强

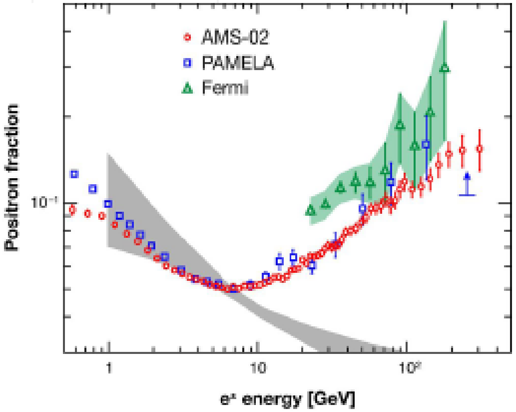
紫金山天文台

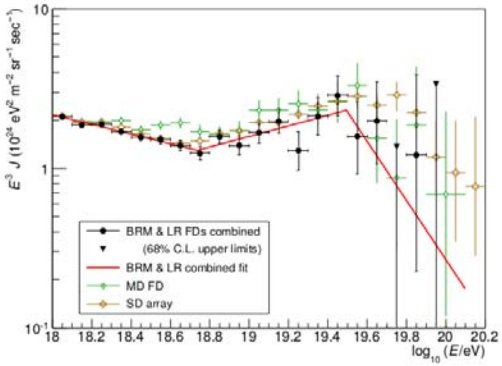
粒子物理数据分析基础和前沿研讨会
2016/09@高能所

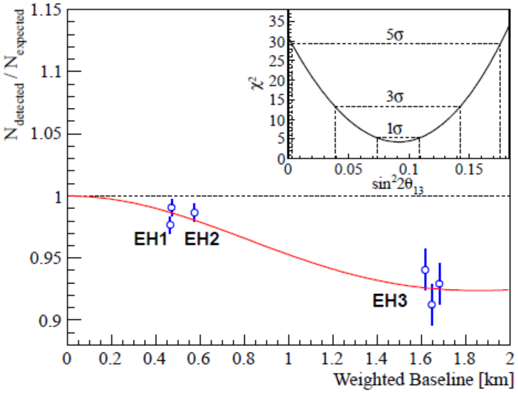
Bayesian推断及MCMC

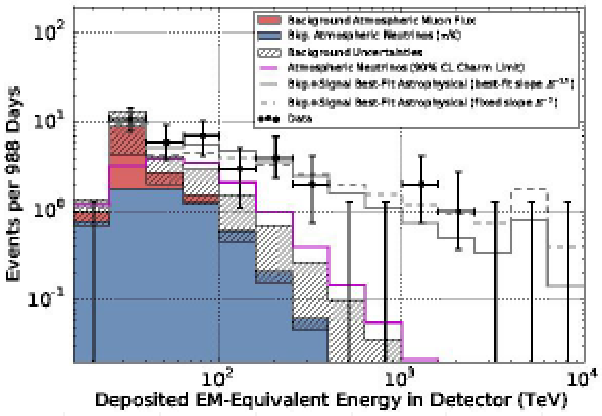
信号和背景





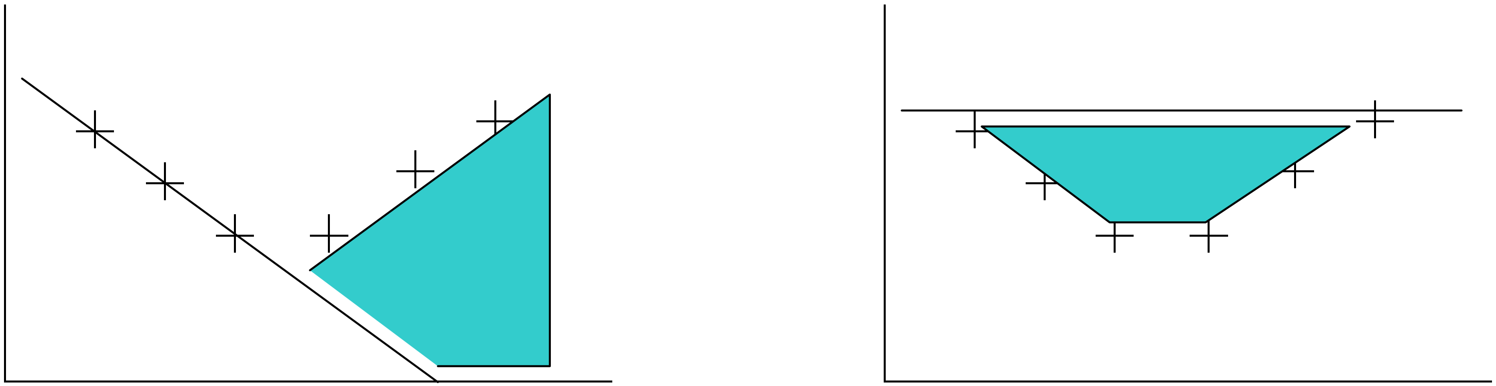






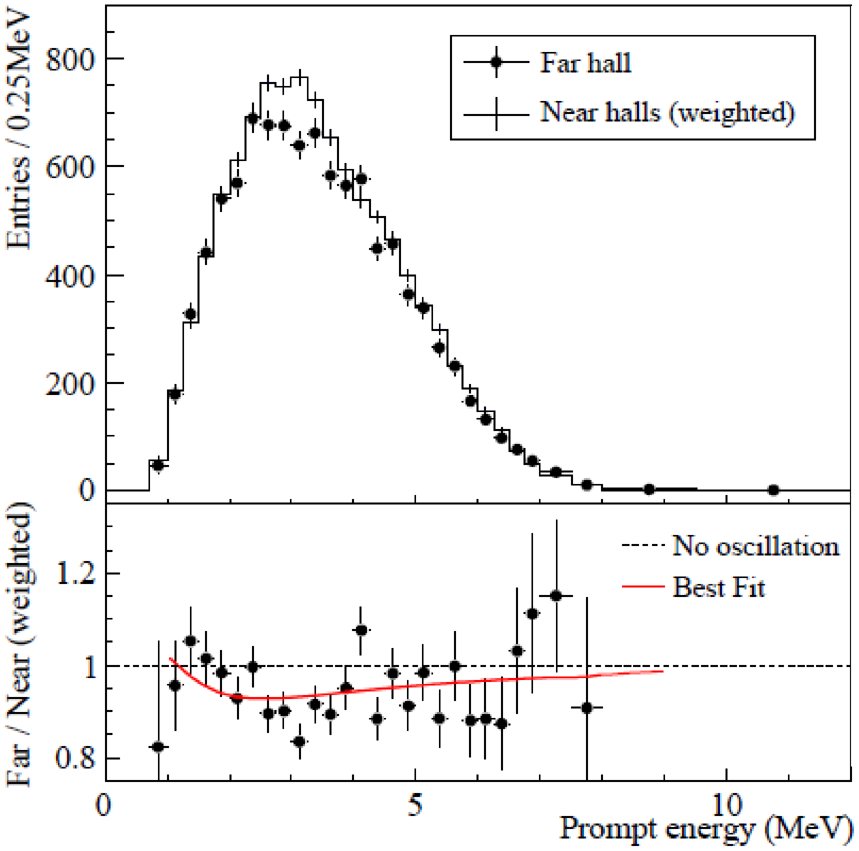
- 声称信号的时候，均是相对于背景而言
- 昨天的信号可能是今天的背景

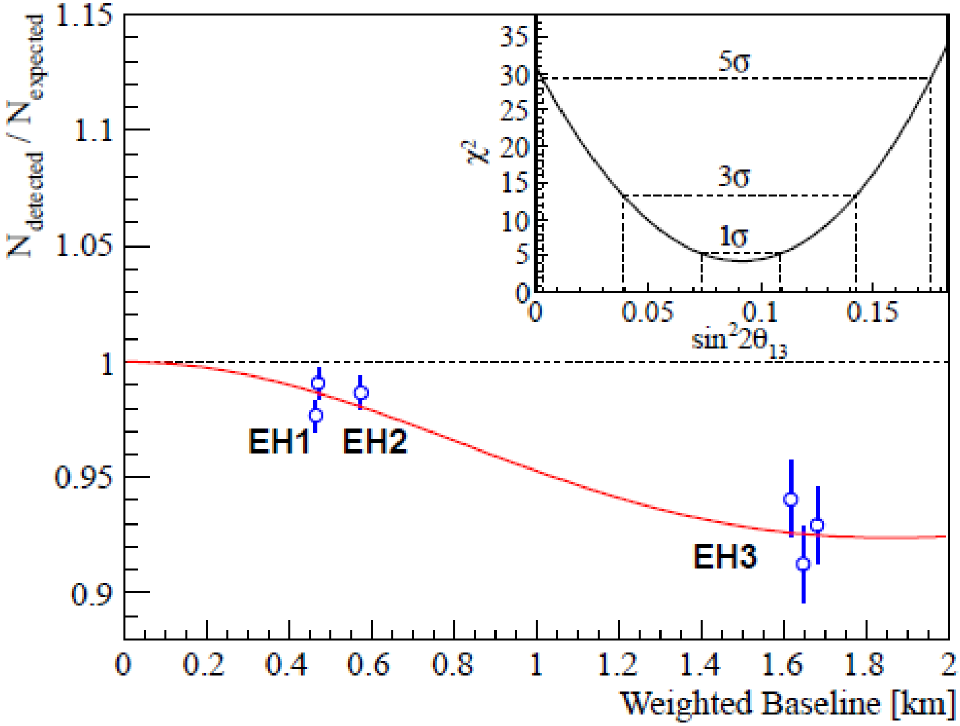
背景影响信号



- 背景会影响信号存在性的判断、显著性、信号的形式
- 背景可能是清楚地知道的，也可能是一定程度上知道

清楚的背景

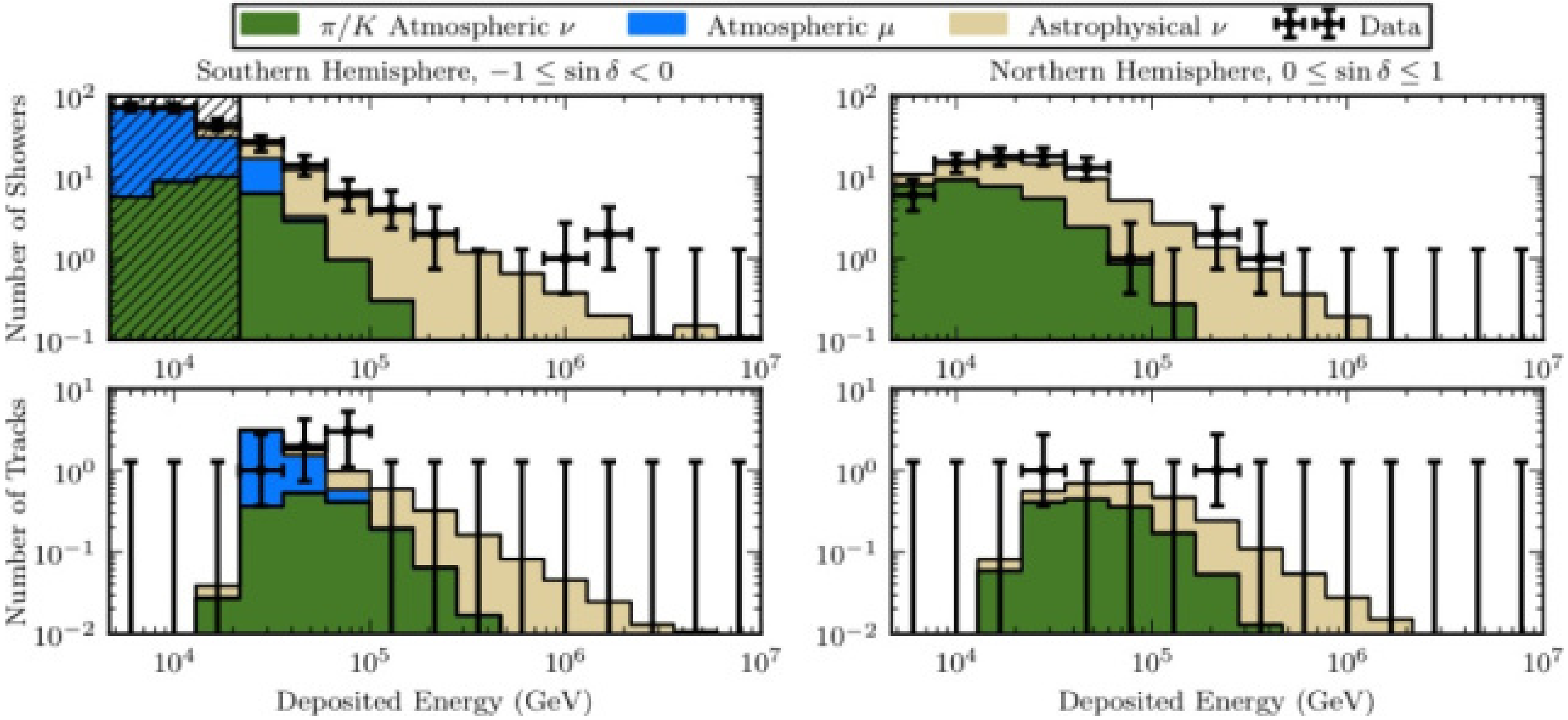




大亚湾 (2012)

背景(null hypothesis): 远点/近点=1

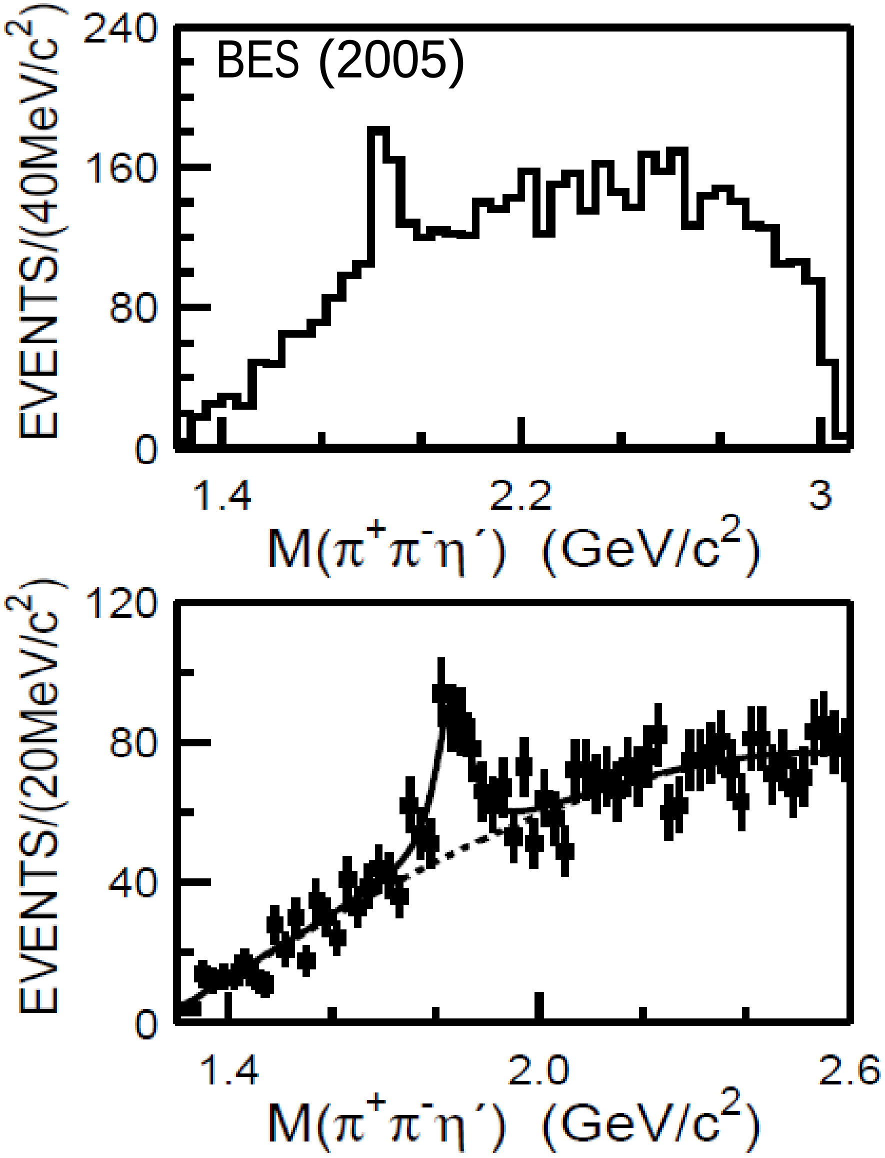
清楚的背景



IceCube neutrino

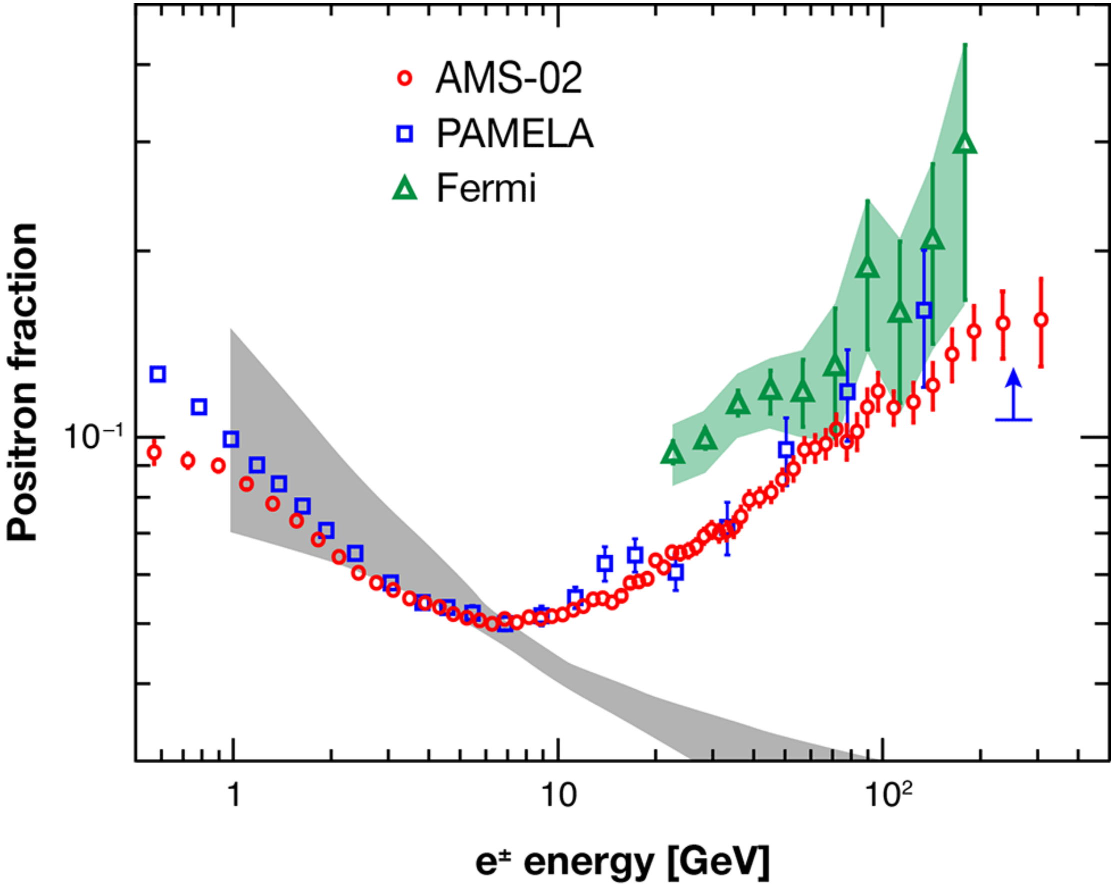
大气中微子和大气muon可以根据宇宙线测量计算给出

不完全清楚的背景



背景(null hypothesis): 多项式连续分布

不完全清楚的背景



背景(null hypothesis): 根据宇宙线测量结果和宇宙线传播模型、银河系物质分布一定程度上可以计算

不完全清楚的背景：额外的自由度

- 如果背景不完全清楚，那么在进行背景模型的时候就需要引入自由参数(冗余参数：nuisance parameters)
- 冗余参数将使得信号变得不那么显著，结论也更加保守

$$B(\mathbf{p_b})+S(\mathbf{p_s})=\mathrm{Data}$$

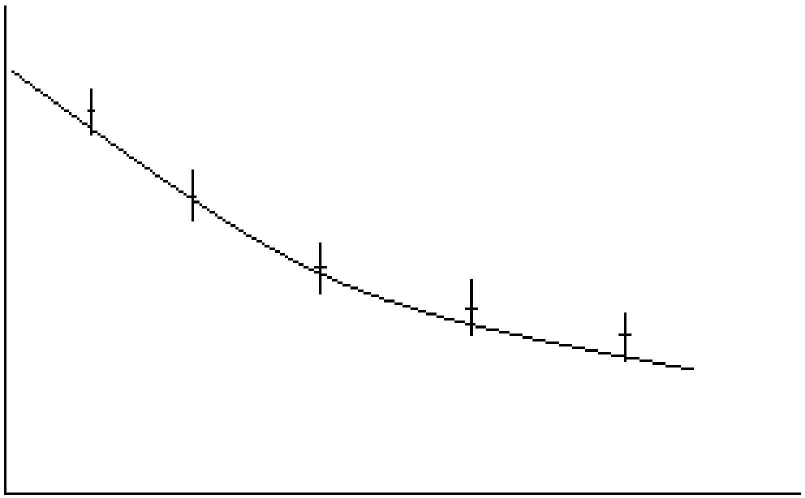
- 整体拟合得到$\mathbf{p_s}$
- $\mathbf{p_b}$的处理:
  - 固定$\mathbf{p_b}$ (conditional likelihood)
  - Profile likelihood
  - Marginal likelihood

似然函数

$$\mathcal{L}(\theta) \equiv P(\mathrm{data}|\theta)$$

- 给定观测数据data，模型(参数θ)的likelihood定义为相应模型(参数)下得到数据data的概率
- *Probability* is used before data are available to describe possible future outcomes given a fixed value for the parameter (or parameter vector).
- *Likelihood* is used after data are available to describe a function of a parameter (or parameter vector) for a given outcome.

高斯似然函数



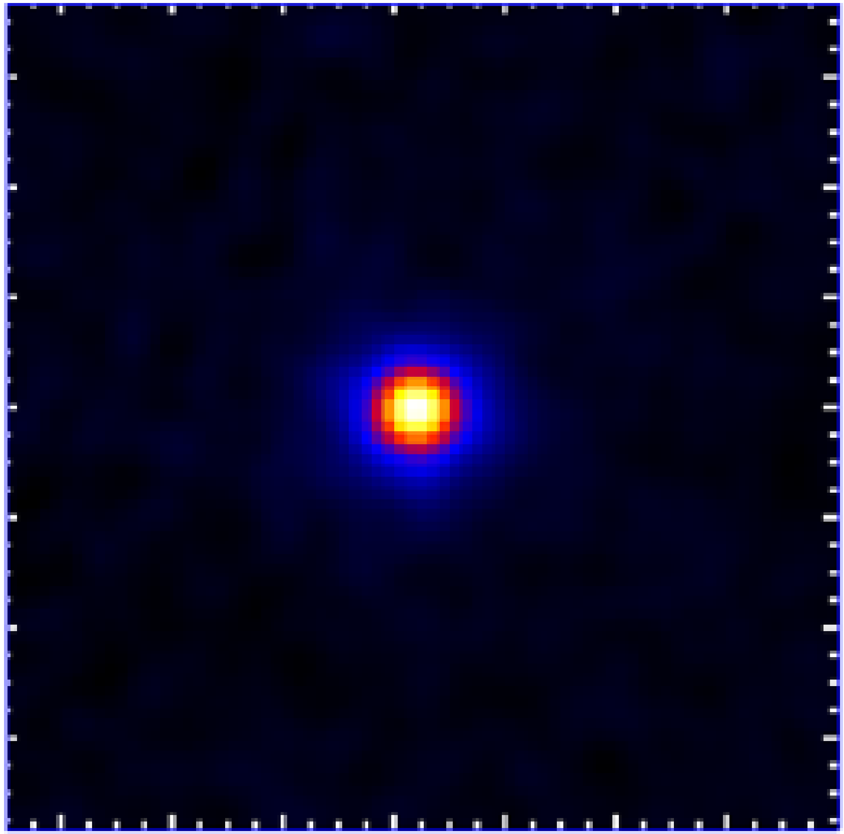
- 如果每个测量值的误差服从高斯分布

$$P(D_i|\theta) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(y(\theta)-y_i)^2}{2\sigma_i^2}\right]$$

- 似然函数为

$$\mathcal{L}(\theta) = \prod_i P(D_i|\theta)$$

泊松似然函数



- 每个bin预期e_i个光子，观测到n_i个光子的概率为

$$P_i = \frac{e_i^{n_i} \exp(-e_i)}{n_i!}$$

- 似然函数为

$$\mathcal{L} = \prod_i P_i$$

阶跃似然函数

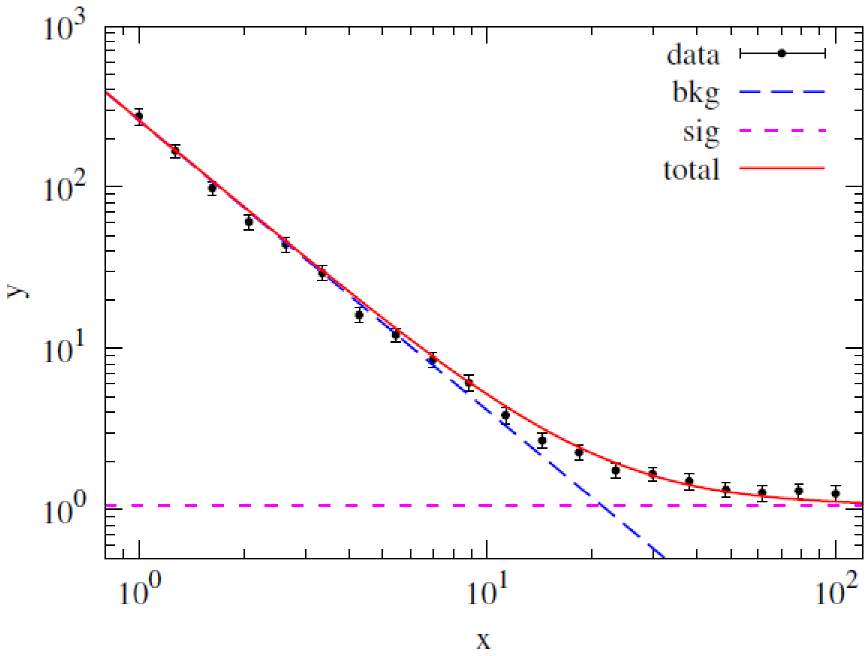
- 探测阈值

$$\mathcal{L} = H(\theta - \theta_0)$$

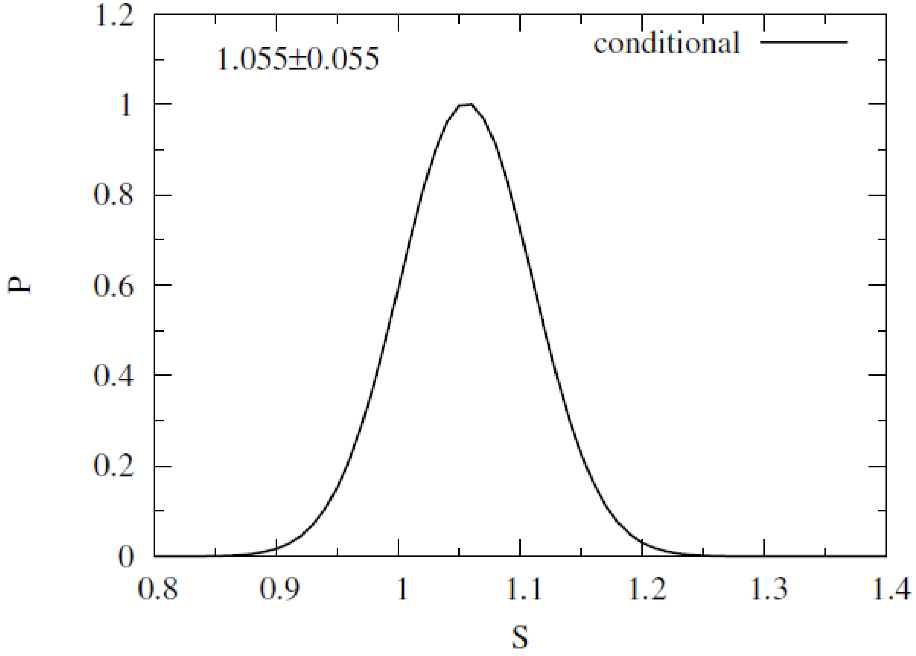
- 超对称模型中Higgsino质量参数的符号
- 理论上的上下限
- 暗物质搜寻中暗物质信号非负

Conditional likelihood

$$\mathcal{L}(\theta_i) = \mathcal{L}(\theta_i; \theta_j = \bar{\theta})$$

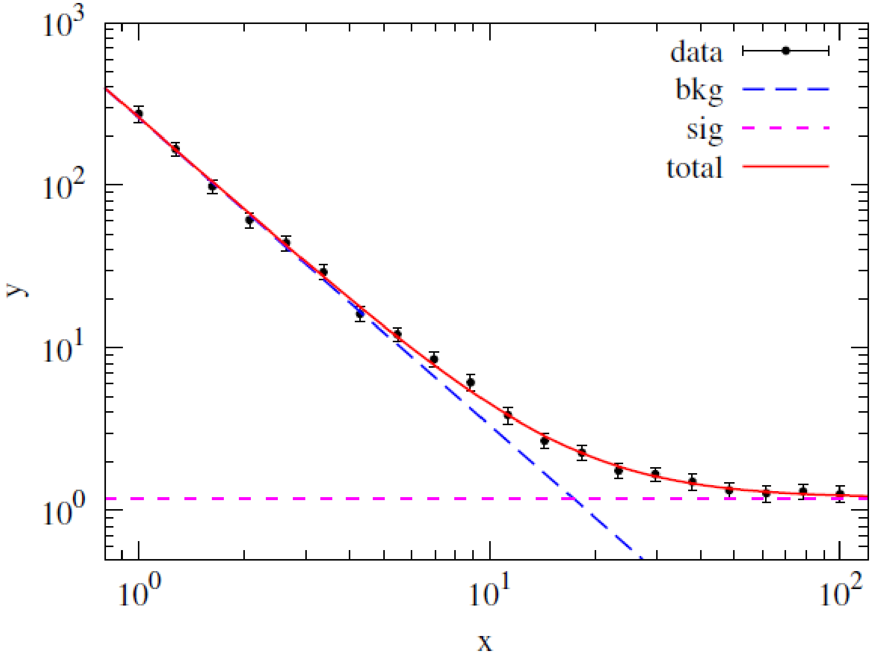


- 幂律背景y=A*x^B
- 常数信号S
- 利用x<10数据确定背景参数A,B
- 固定A,B，拟合得到S

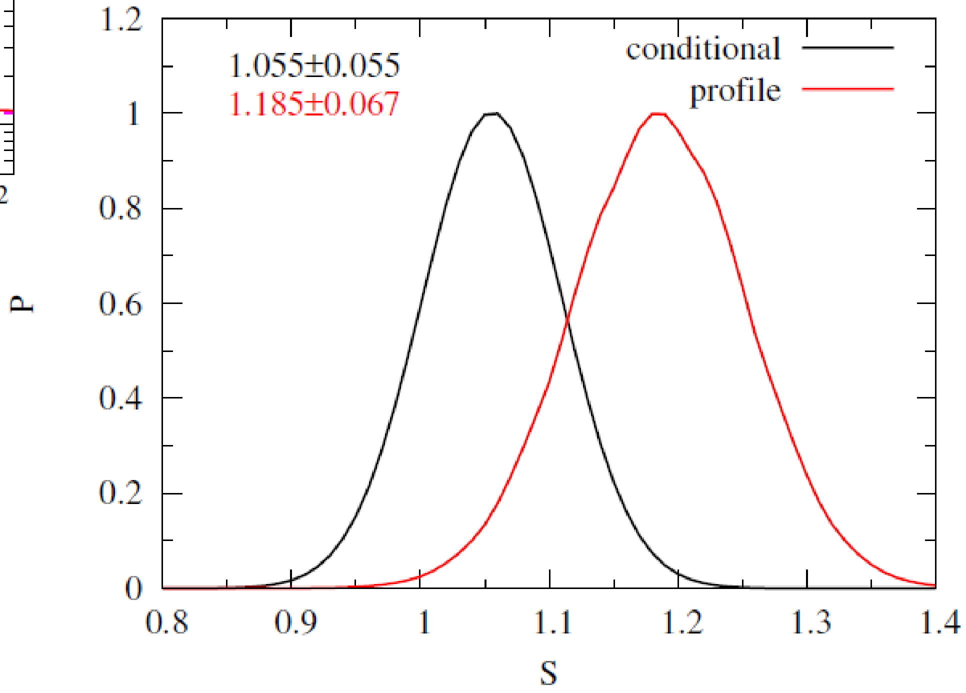


Profile likelihood

$$\mathcal{L}(\theta_i) = \max_{\theta_j} \mathcal{L}(\theta_i, \theta_j)$$



- 对每一个S值，扫描参数平面(A,B)，得到最大似然值L(S)
- 可以给出整体最佳拟合
- 参数分布更宽

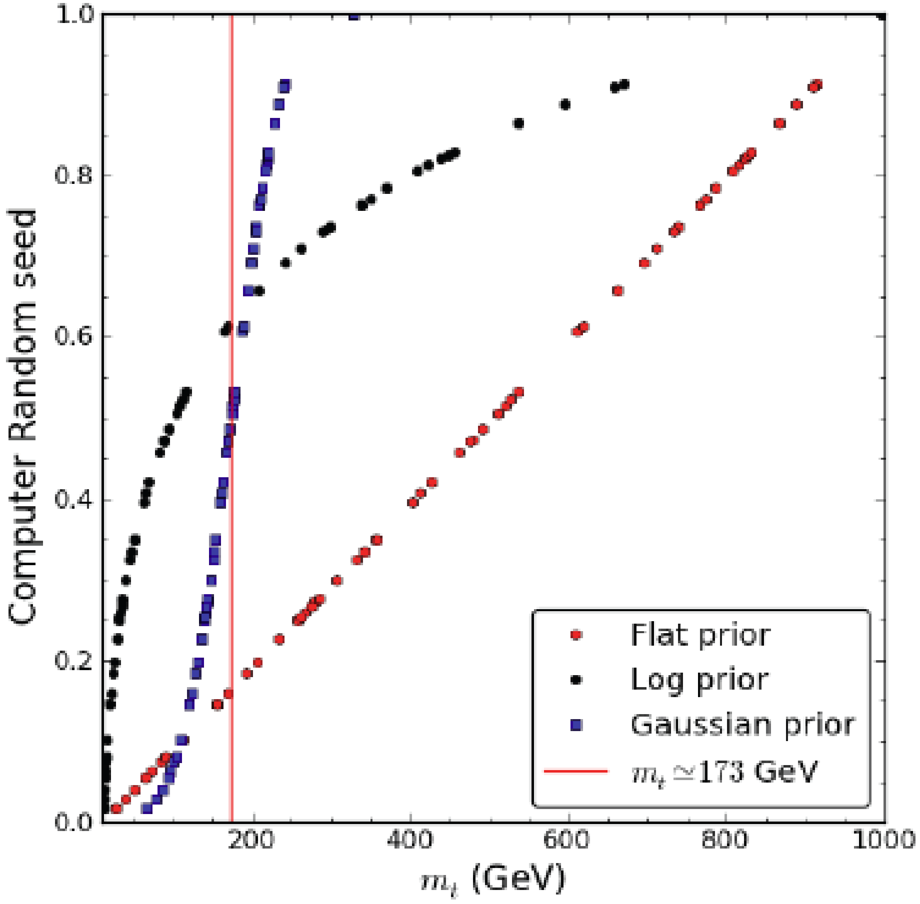


贝叶斯定理

$$P(\theta|\text{data}) = \frac{P(\text{data}|\theta) \cdot P(\theta)}{P(\text{data})}$$

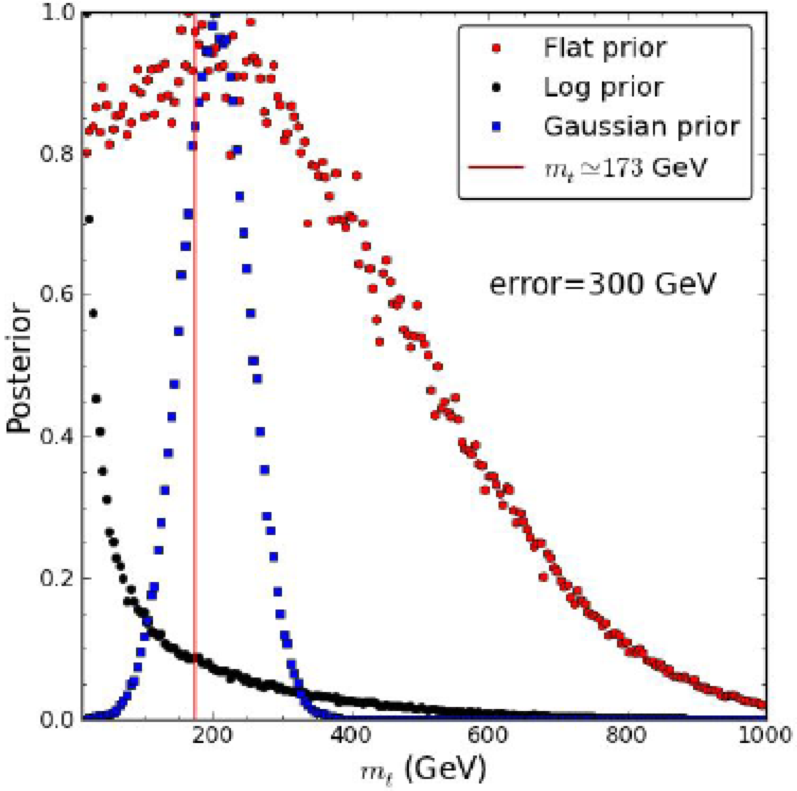
- Likelihoo $\mathcal{L}(\theta) \equiv P(\text{data}|\theta)$ ：给定假设(模型)得到观测值的概率
- Prior $P(\theta)$：观测之前对参数的知识
- Evidence $P(\text{data})$: 观测数据的概率，归一化常数
- Posterior $P(\theta|\text{data})$：观测之后参数的验后概率

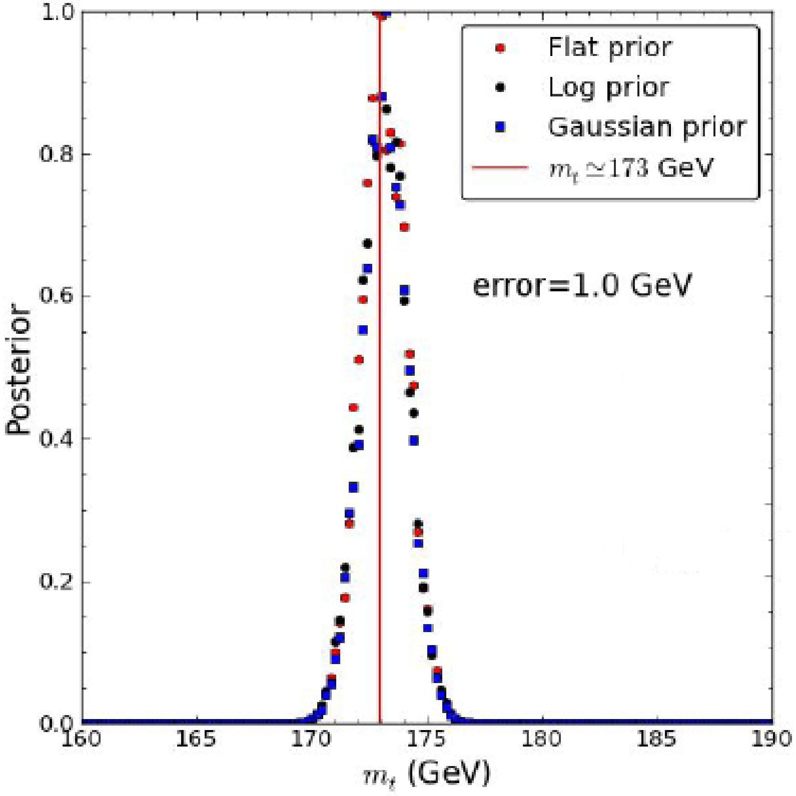
先验概率



- 具有一定的随意性
- 常见的如flat, log, gaussian
- 也可以基于独立实验结果

先验概率对验后概率的影响

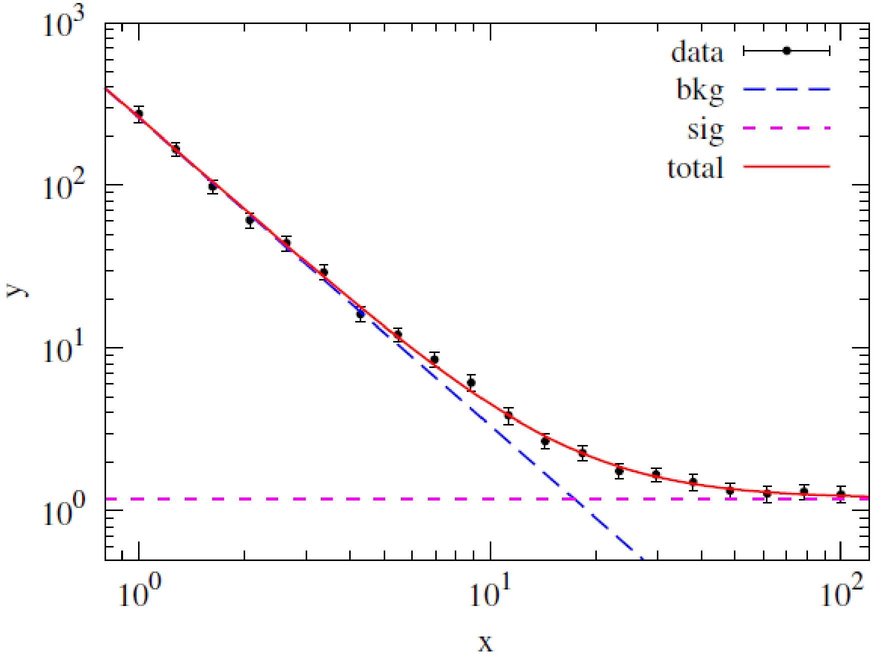




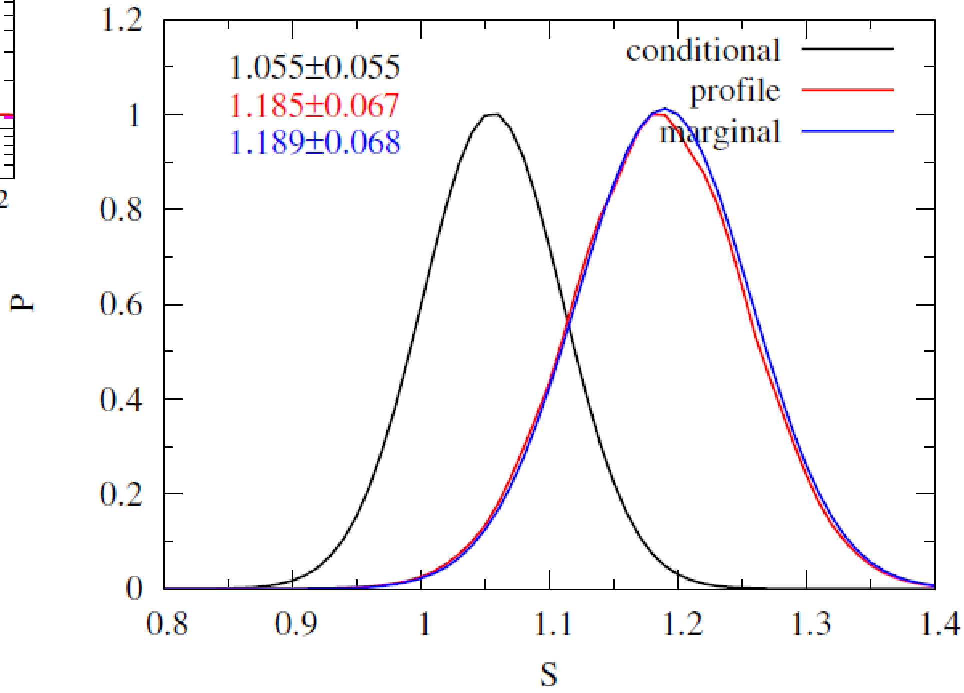
如果似然函数呈明显尖峰状，验后分布对先验依赖不大；否则先验分布决定了验后分布。

Marginal likelihood

$$\mathcal{L}(\theta_i) = \int \mathcal{L}(\theta_i, \theta_j)\mathrm{d}\theta_j$$

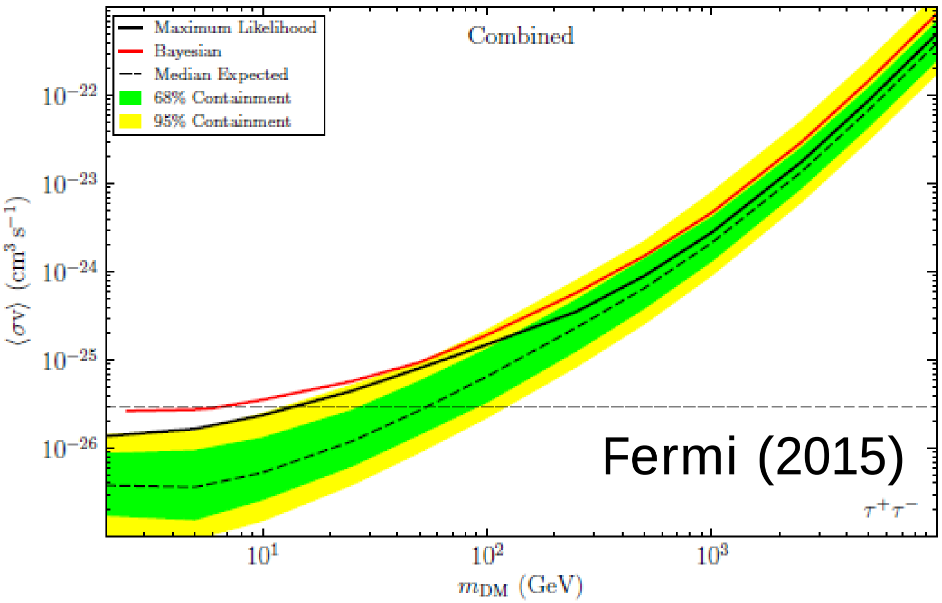


- 对每一个S值，将似然值L(S;A,B)对参数空间(A,B)积分得到L(S)
- 遍历参数空间
- 参数分布比profile likelihood略宽



	Fermi	LHC	IceCube	XENON/LUX/ PANDAX	PLANCK/WMAP
Profile	√	√	√	√	×
Marginal	√	√	×	×	√

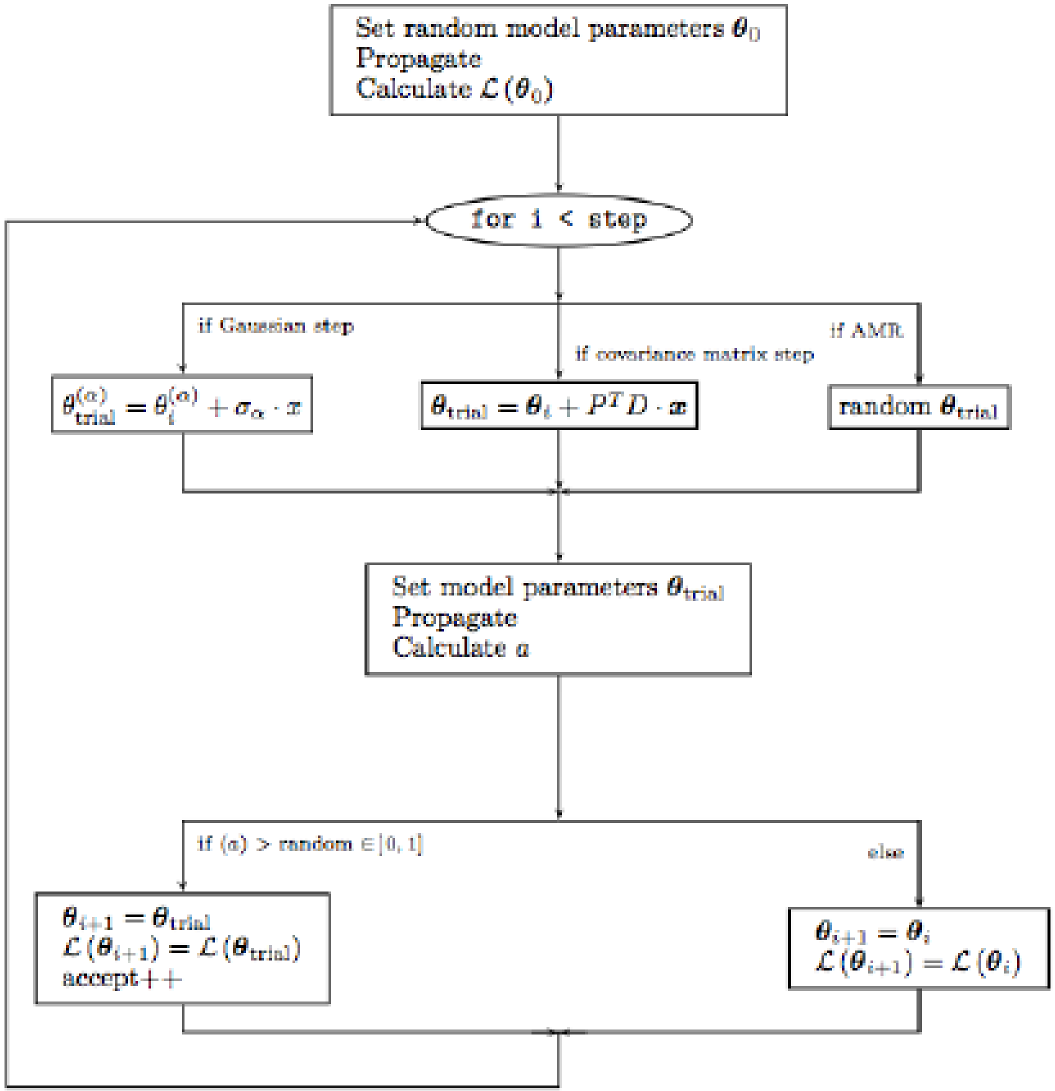
- 没有“绝对正确”的统计方法
- 越来越多的人使用Bayesian marginal profile likelihood
- Marginal profile方法结果会更加保守



Markov Chain Monte Carlo

- 马尔可夫链: 第i+1点仅依赖于第i点的结果，跟此前的历史无关
- 满足一定条件的马尔可夫链(不可约、非周期)收敛到平稳分布
- MCMC：构建逼近目标分布的马尔可夫链，并利用蒙特卡罗办法实现
- 常见应用：高维数值积分，多参数拟合

Metropolis-Hastings sampling



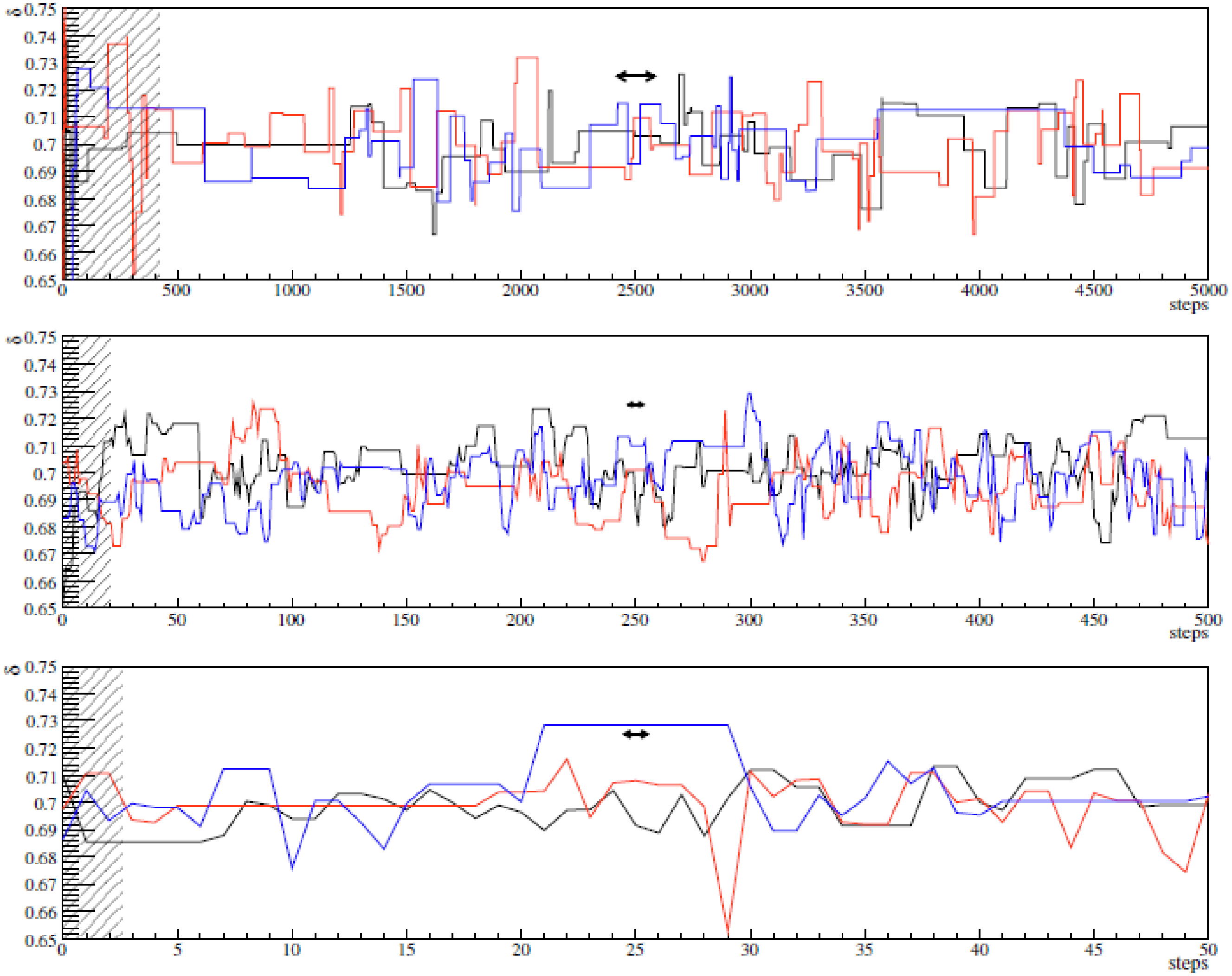
- 尝试i->i+1
- 计算接受概率

$$a = \min\left(1, \frac{p(\boldsymbol{\theta}_{\rm trial})}{p(\boldsymbol{\theta}_i)}\right)$$

- 如果a>random()接受i+1点；否则停留在i点
- 足够多步数之后样本分布逼近于目标分布

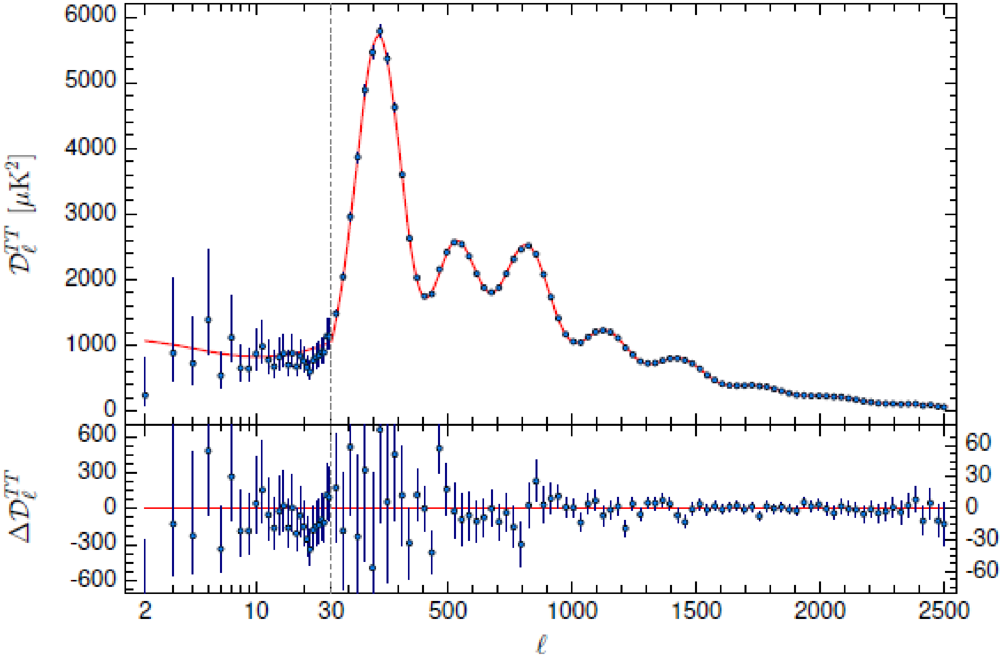
Putze et al. (2009)

Markov chains

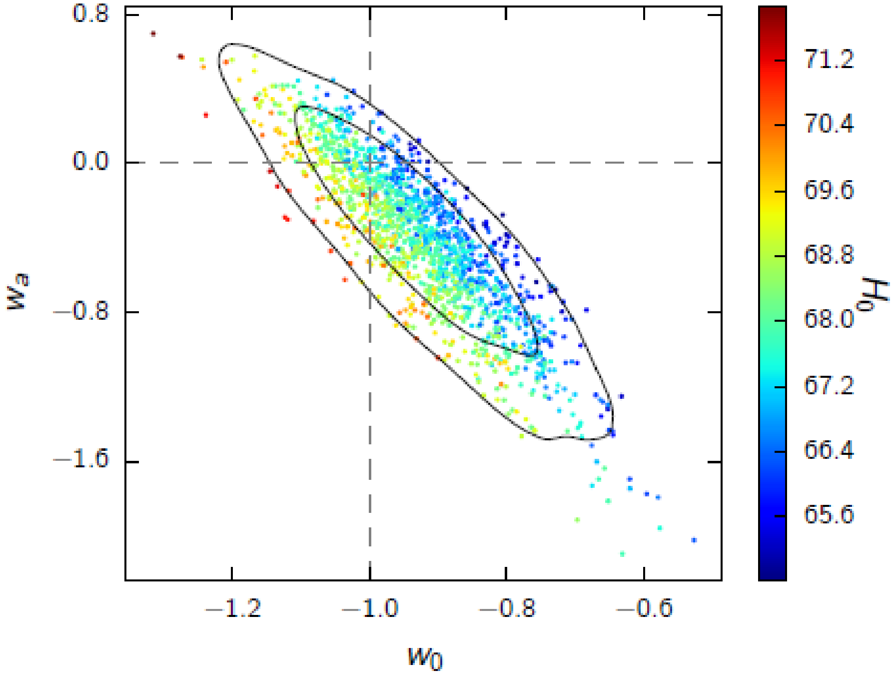


Putze et al. (2009)

宇宙学应用



CosmoMC: Lewis & Bridle (2002)
http://cosmologist.info/cosmomc/



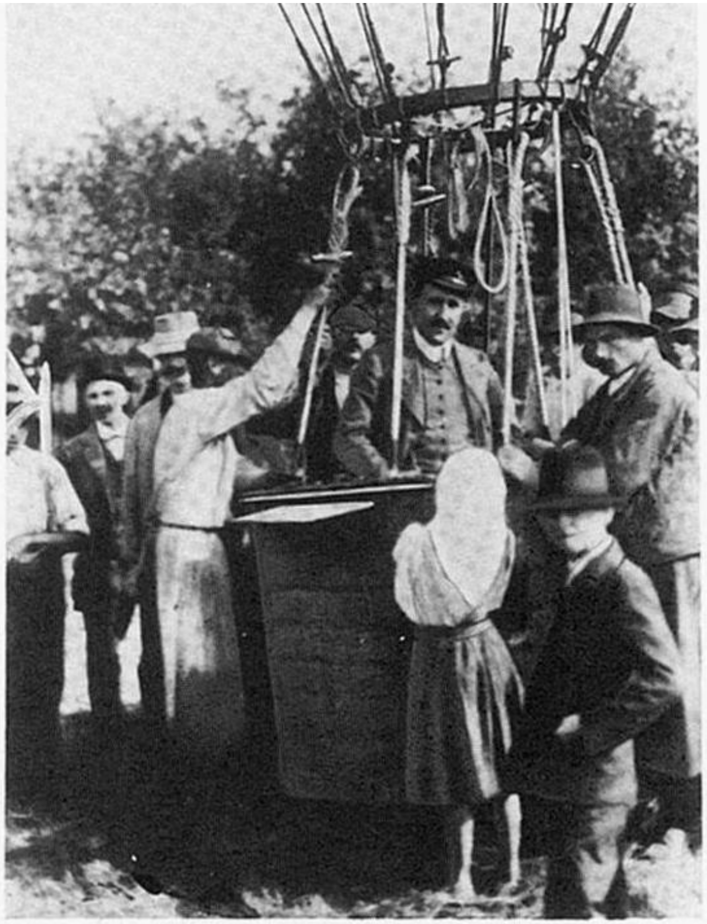
PLANCK (2016)

宇宙线物理基础

宇宙线发现

约100年前由V. Hess以及其他一些科学家发现(~1912)

V. Hess荣获1936年度诺贝尔物理学奖



Hess bei Ballonlandung (1912).

两个黄金年代

- 1930s-1950s

  1932: C. Anderson发现正电子(诺贝尔奖)

  1936: C. Anderson and S. Neddermeyer发现mu轻子

  1946: G. Rochester and C. Butler发现K介子

  1947: C. Powell发现pi介子(诺贝尔奖)

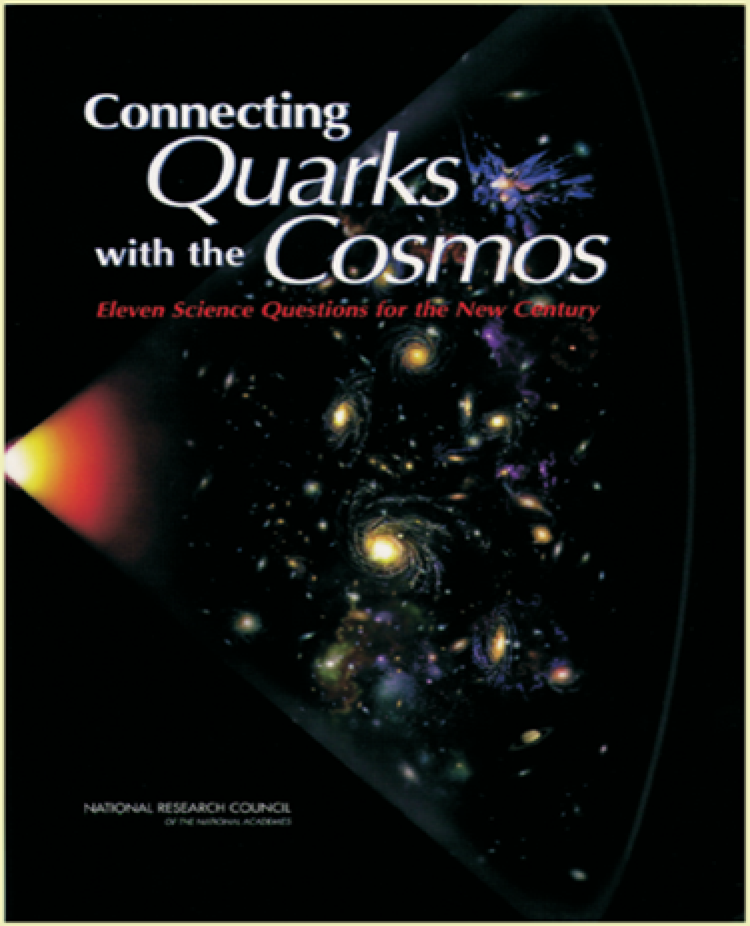
  1949: 张文裕发现mu-原子

- 今天

  高能量极端物理以及暗物质等新物理的探索

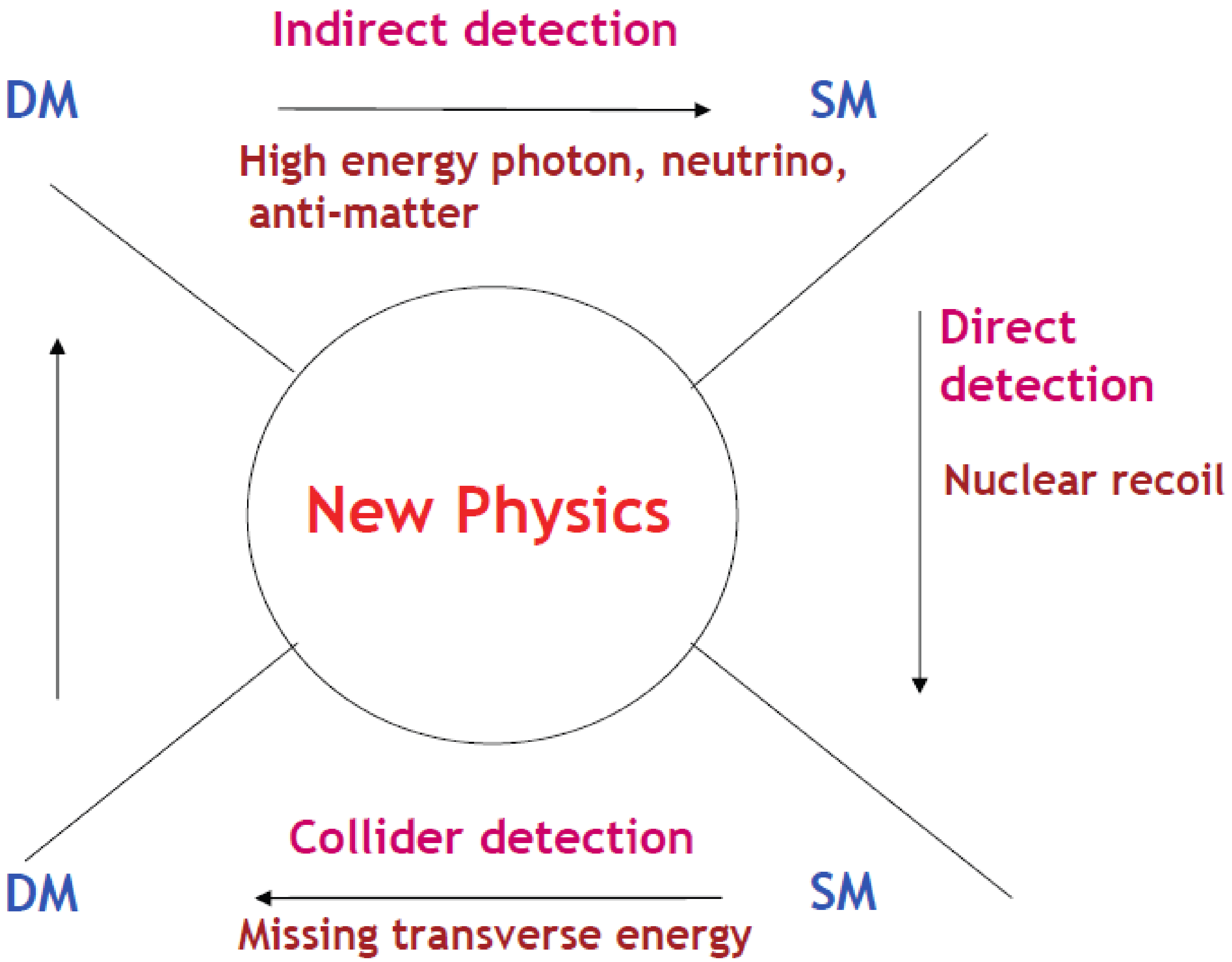
重大前沿科学问题

The Eleven Questions Identified by the *Connecting Quarks with the Cosmos* Report



1. What is Dark Matter?
2. What is the Nature of Dark Energy?
3. How Did the Universe Begin?
4. Did Einstein Have the Last Word on Gravity?
5. What are the Masses of the Neutrinos and How Have They Shaped the Evolution of the Universe?
6. How do Cosmic Accelerators Work and What are They Accelerating?
7. Are Protons Unstable?
8. What Are the New States of Matter at Exceedingly High Density and Temperature?
9. Are There Additional Space-Time Dimensions?
10. How Were the Elements from Iron to Uranium Made?
11. Is a New Theory of Light and Matter Needed at the Highest Energies?

暗物质探测手段



Indirect detection

DM

SM

High energy photon, neutrino,
anti-matter

Direct
detection

Nuclear recoil

New Physics

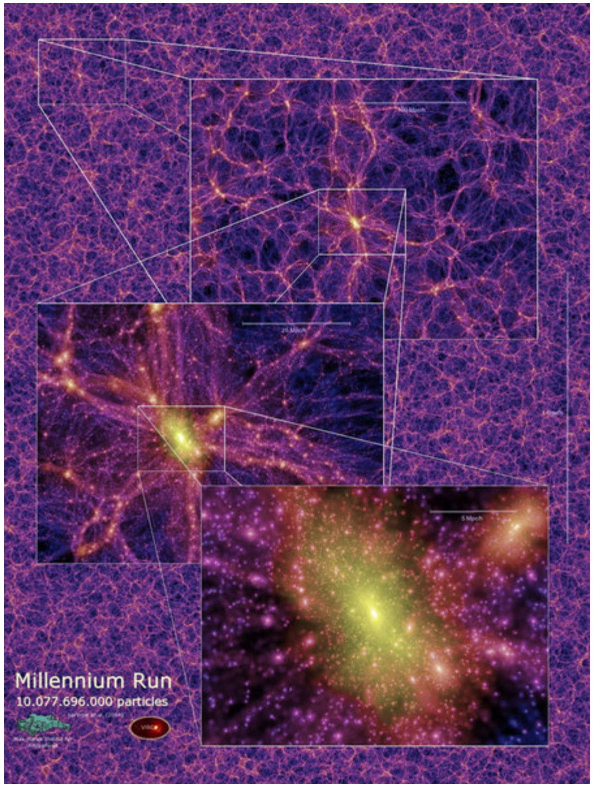
Collider detection

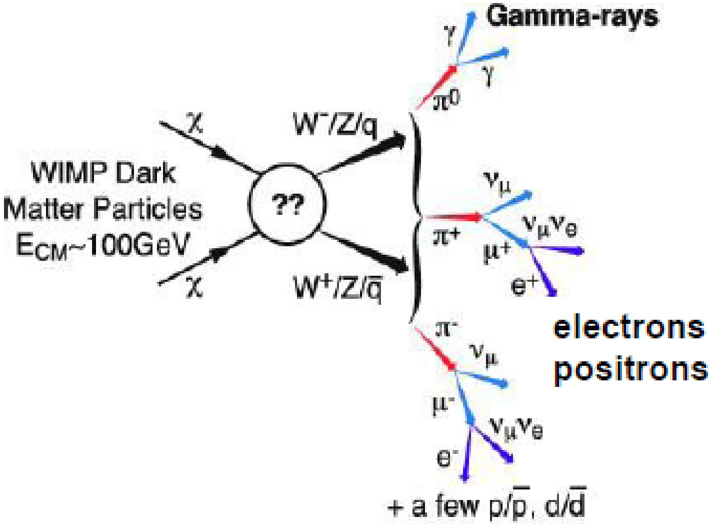
DM

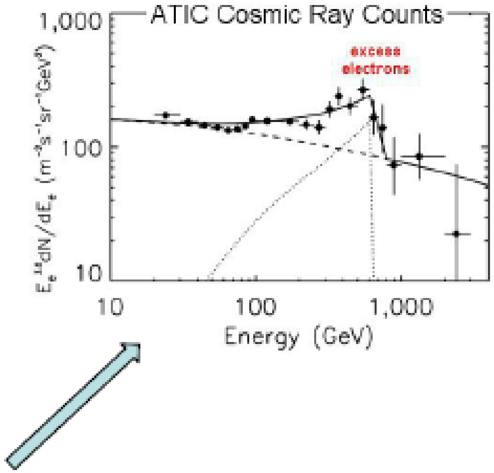
SM

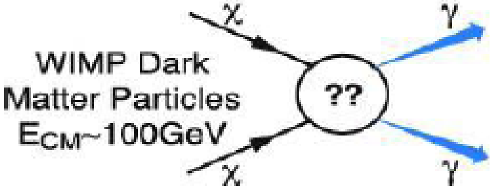
Missing transverse energy

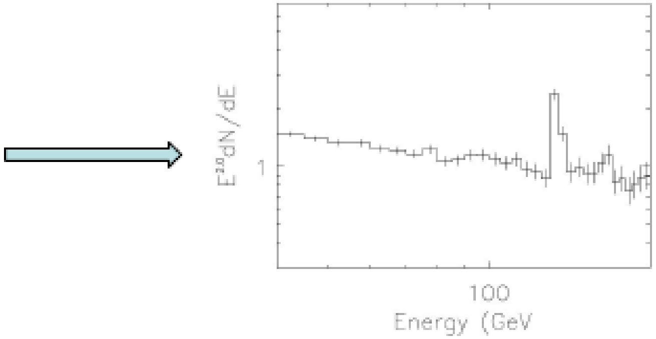
从宇宙线中探测暗物质


Millennium Run
10.077.696.000 particles

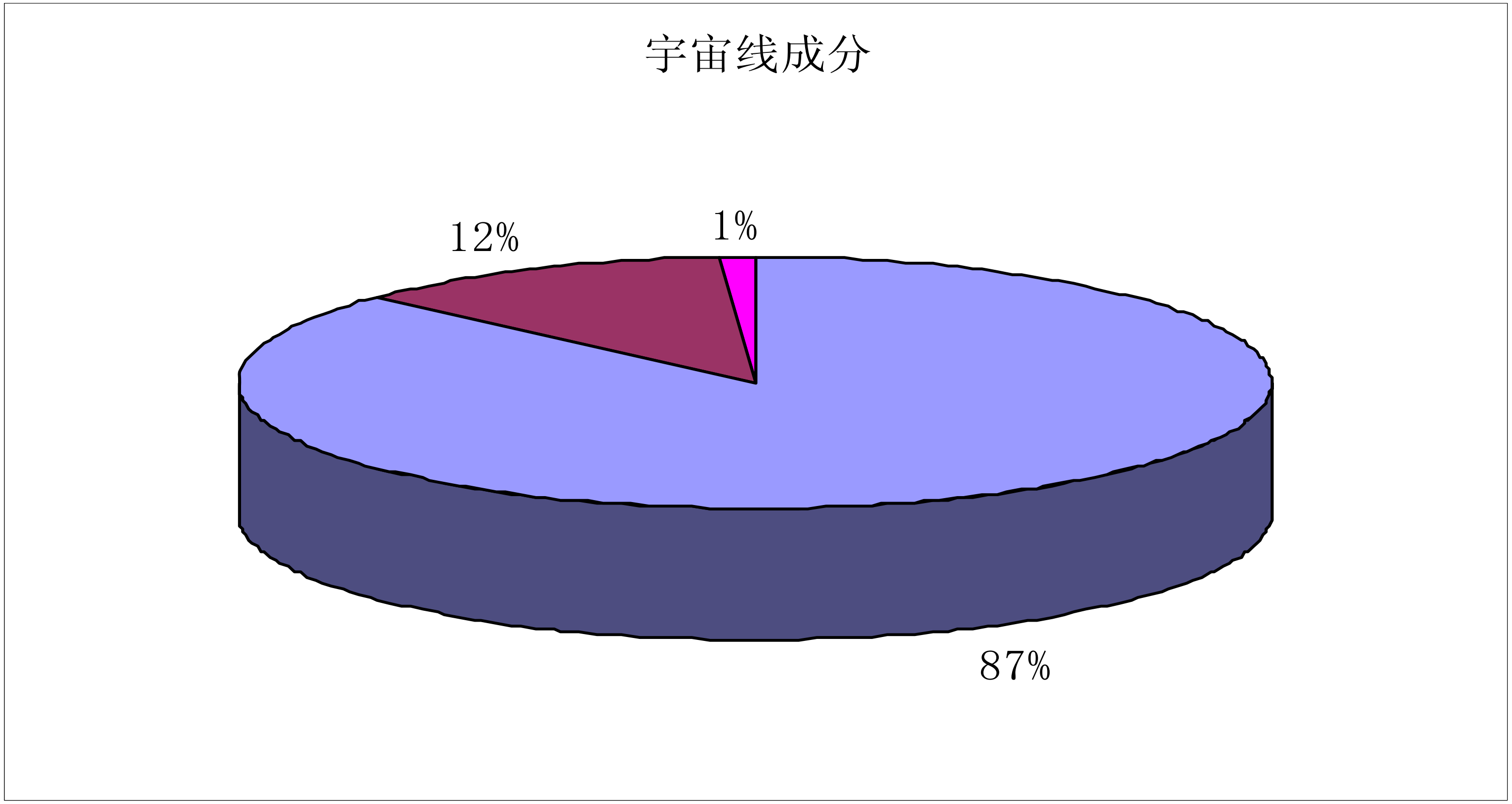

WIMP Dark
Matter Particles
E$_{CM}$~100GeV
χ
χ
??
W⁻/Z/q
W⁺/Z/q̄
Gamma-rays
γ
γ
π0
νμ
π+
νμνe
μ+
e+
π-
νμ
μ-
νμνe
e-
electrons
positrons
+ a few p/p̄, d/d̄


ATIC Cosmic Ray Counts
excess
electrons
1,000
100
10
10
100
1,000
Energy (GeV)


WIMP Dark
Matter Particles
E$_{CM}$~100GeV
χ
χ
??
γ
γ


1
100
Energy (GeV

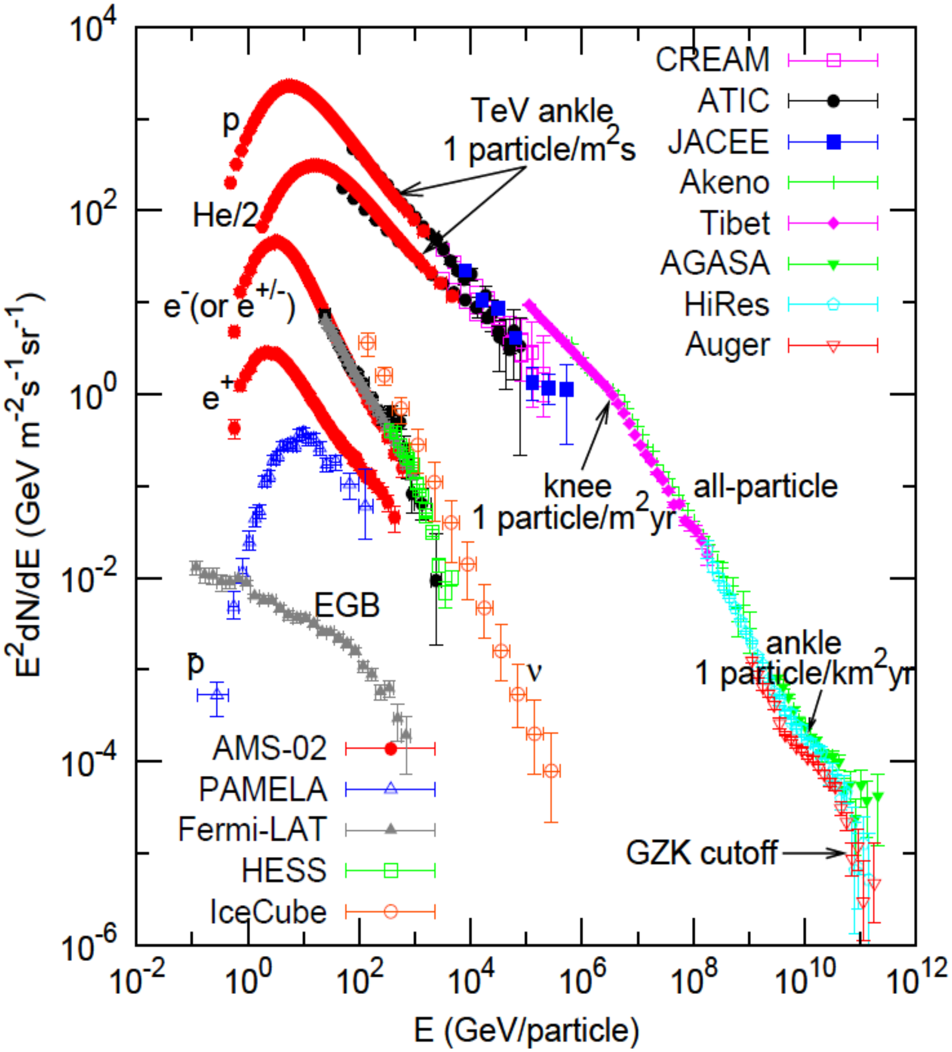
宇宙线观测结果



🟦 质子 🟥 α粒子 🟪 其它重核

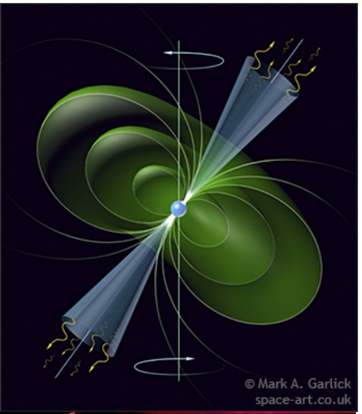
还有少量的电子，γ光子、反粒子等等

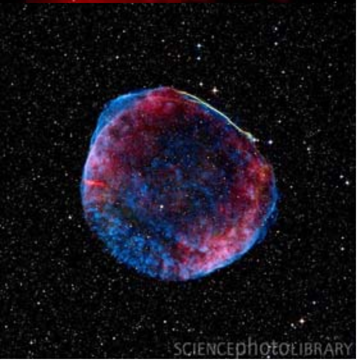
宇宙线观测结果

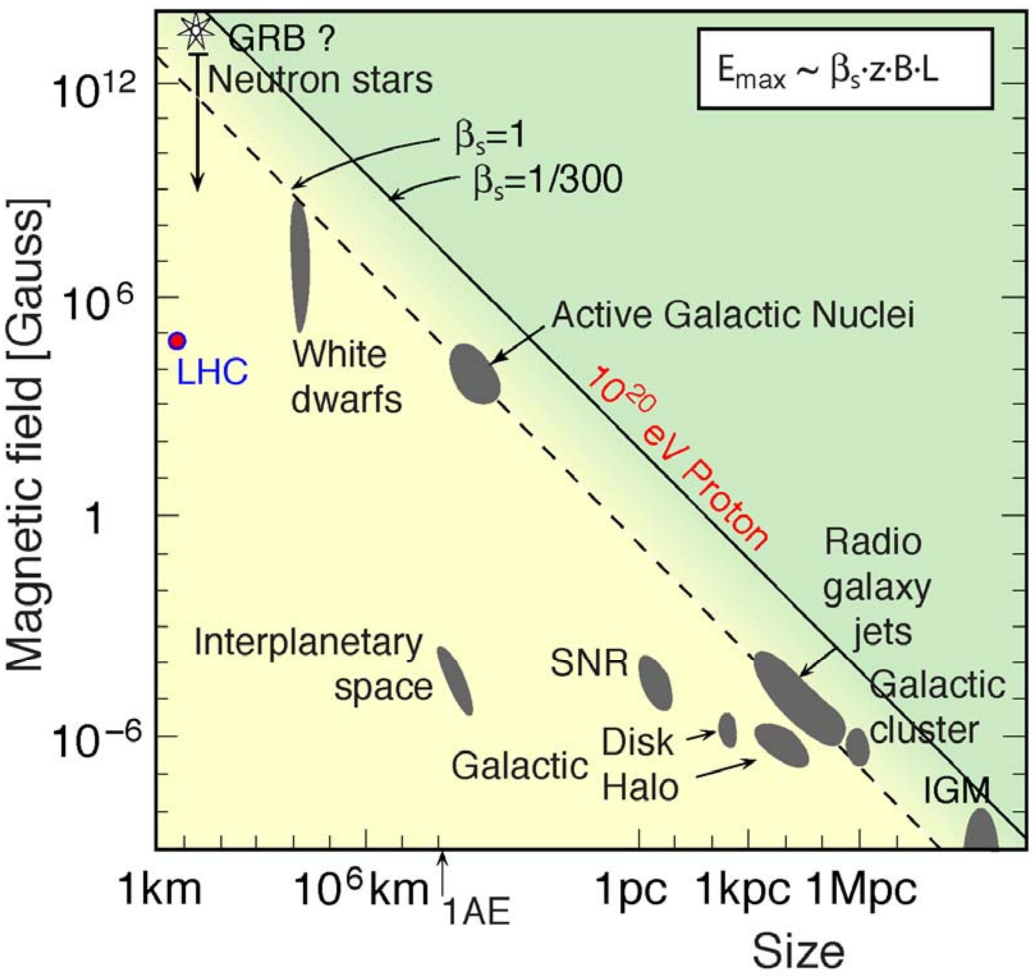


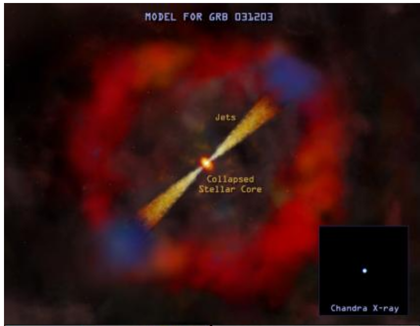
- 主要性质：幂律非热能谱
- 微小结构(膝、踝、GZK截断等)
- 最高能量宇宙线粒子3×10^{20} eV，和宏观物体能量相当
- 基本问题：起源、加速、传播
- 极端高能量前沿物理探索

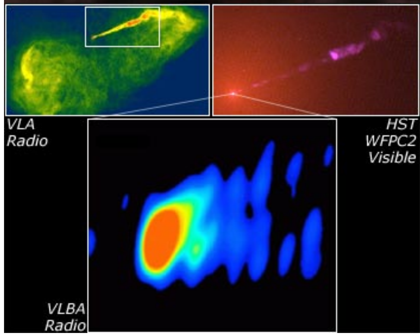
宇宙线起源和加速

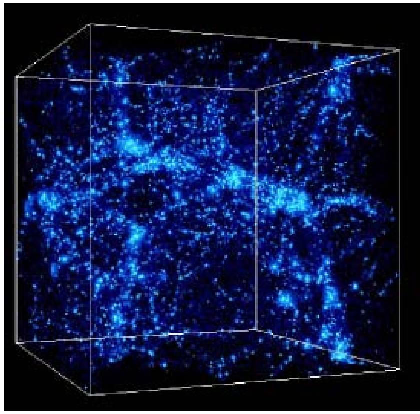

© Mark A. Garlick
space-art.co.uk


SCIENCEphotoLIBRARY

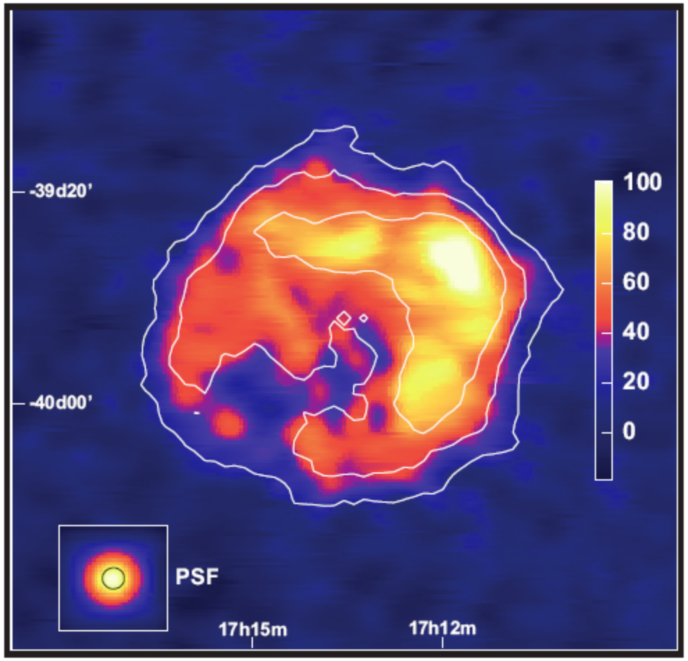

E_max ~ β_s·z·B·L
GRB ?
Neutron stars
β_s=1
β_s=1/300
10^12
10^6
1
10^-6
Magnetic field [Gauss]
LHC
White dwarfs
Active Galactic Nuclei
10^20 eV Proton
Radio galaxy jets
Galactic cluster
IGM
Interplanetary space
SNR
Galactic
Disk
Halo
1km
10^6 km
1AE
1pc
1kpc
1Mpc
Size

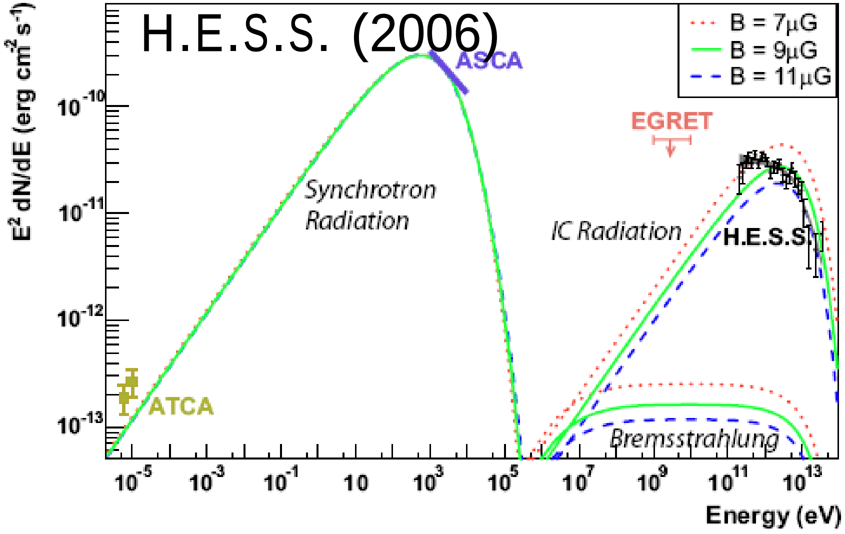

MODEL FOR GRB 031203
Jets
Collapsed Stellar Core
Chandra X-ray

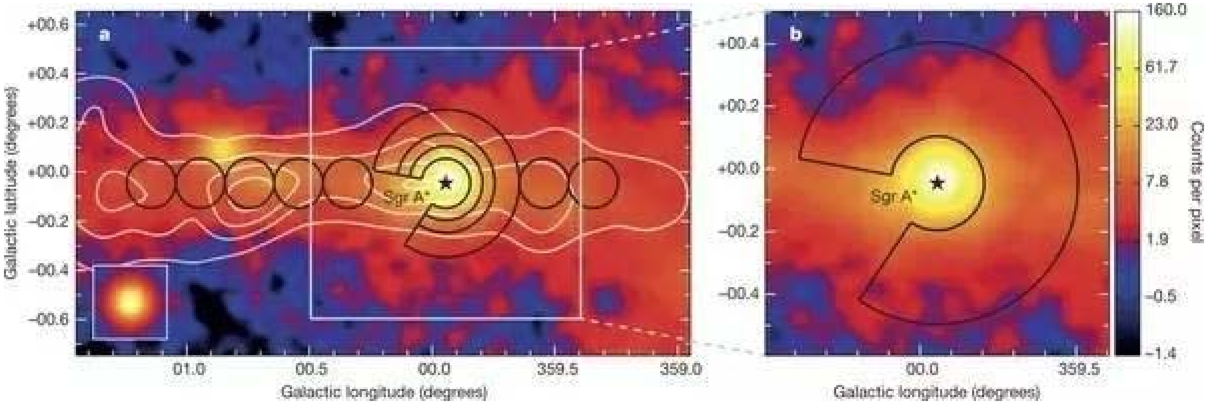

VLA Radio
HST WFPC2 Visible
VLBA Radio

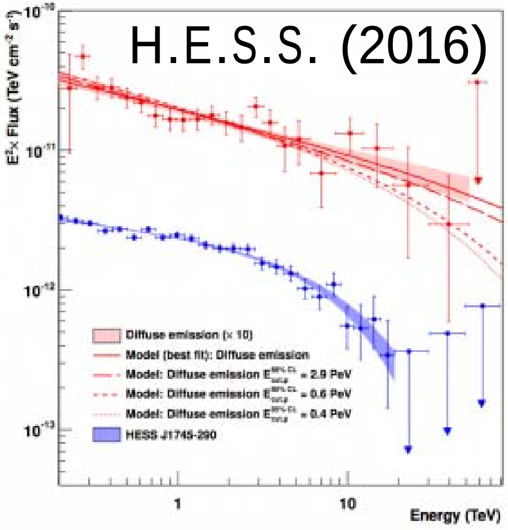


观测到的宇宙线源

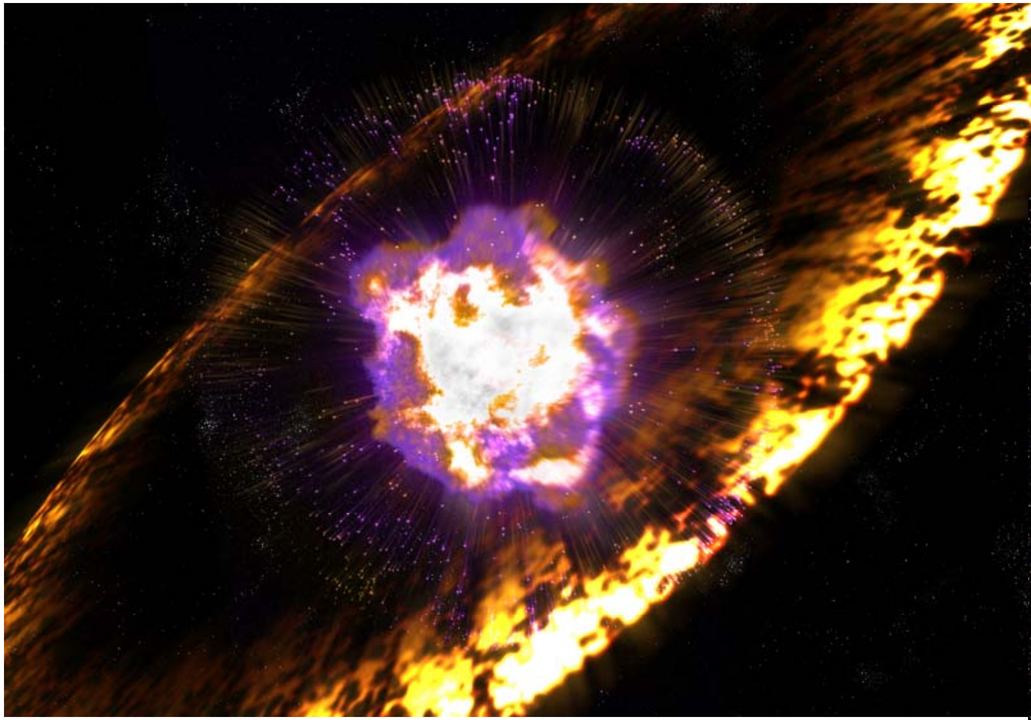

-39d20'
-40d00'
100
80
60
40
20
0
PSF
17h15m
17h12m

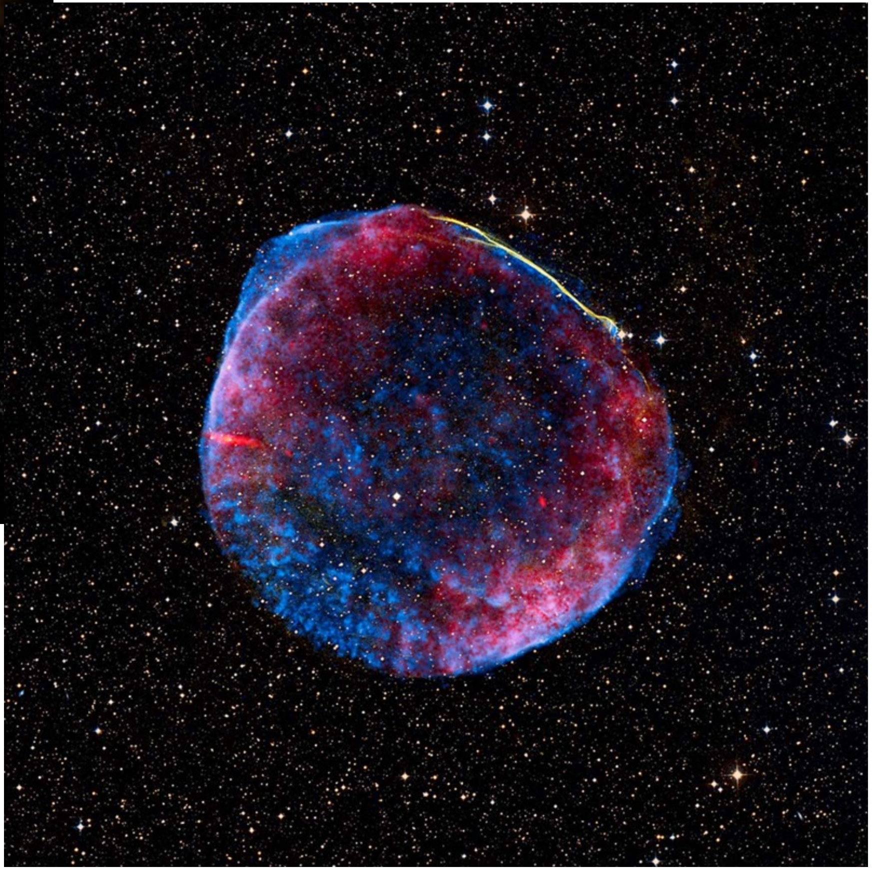

H.E.S.S. (2006)
B = 7μG
B = 9μG
B = 11μG
ASCA
EGRET
Synchrotron Radiation
IC Radiation
H.E.S.S.
ATCA
Bremsstrahlung
E^2 dN/dE (erg cm^{-2} s^{-1})
10^{-10}
10^{-11}
10^{-12}
10^{-13}
10^{-5}
10^{-3}
10^{-1}
10
10^{3}
10^{5}
10^{7}
10^{9}
10^{11}
10^{13}
Energy (eV)


a
b
Sgr A*
Sgr A*
+00.6
+00.4
+00.2
+00.0
−00.2
−00.4
−00.6
Galactic latitude (degrees)
01.0
00.5
00.0
359.5
359.0
Galactic longitude (degrees)
+00.4
+00.2
+00.0
−00.2
−00.4
00.0
359.5
Galactic longitude (degrees)
160.0
61.7
23.0
7.8
1.9
−0.5
−1.4
Counts per pixel

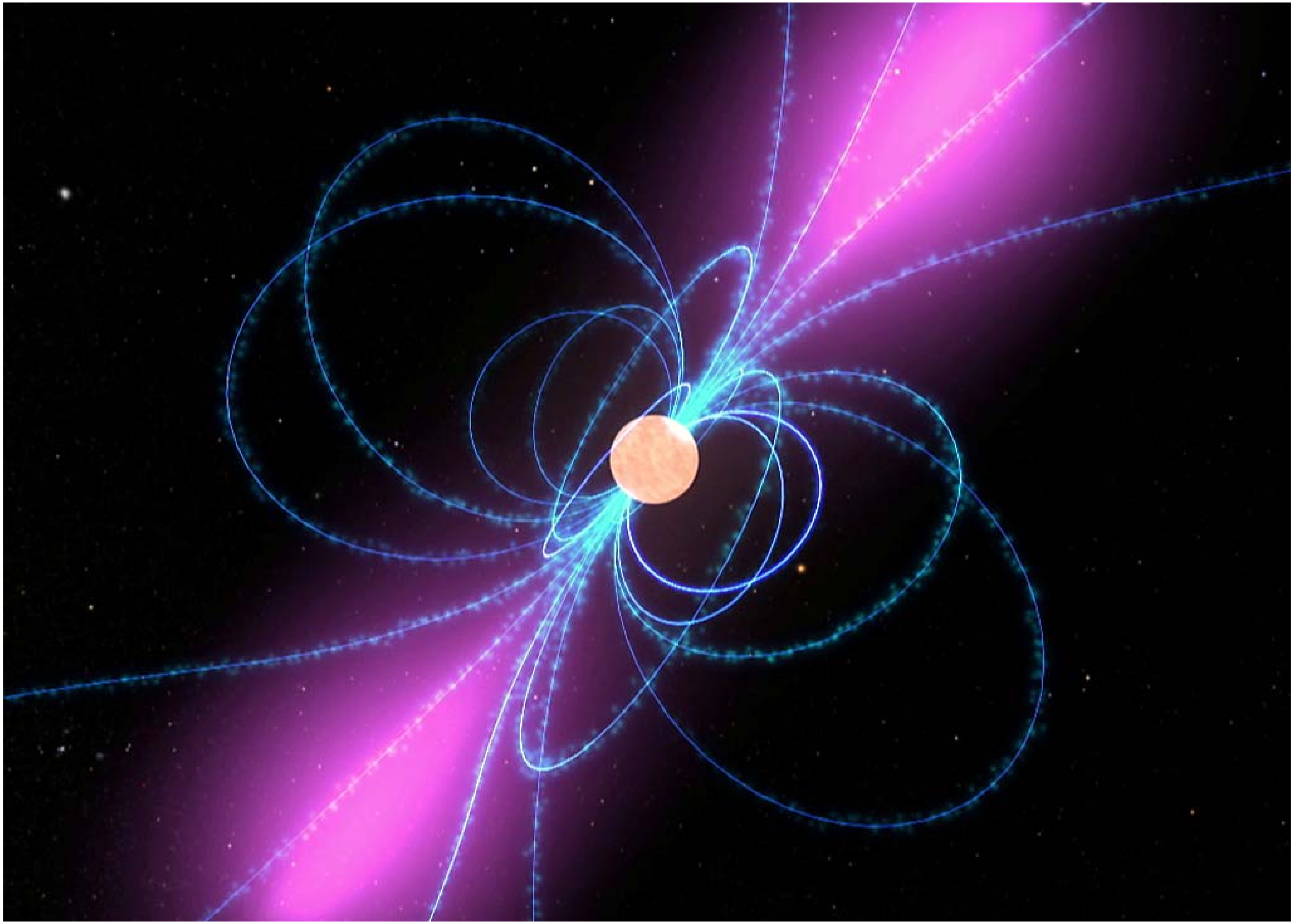

H.E.S.S. (2016)
$E^2\times$ Flux (TeV cm^{-2} s^{-1})
10^{-10}
10^{-11}
10^{-12}
10^{-13}
Diffuse emission (× 10)
Model (best fit): Diffuse emission
Model: Diffuse emission $E^{95\% CL}_{cut,p}$ = 2.9 PeV
Model: Diffuse emission $E^{95\% CL}_{cut,p}$ = 0.6 PeV
Model: Diffuse emission $E^{95\% CL}_{cut,p}$ = 0.4 PeV
HESS J1745-290
1
10
Energy (TeV)

宇宙中的回旋加速器：超新星激波

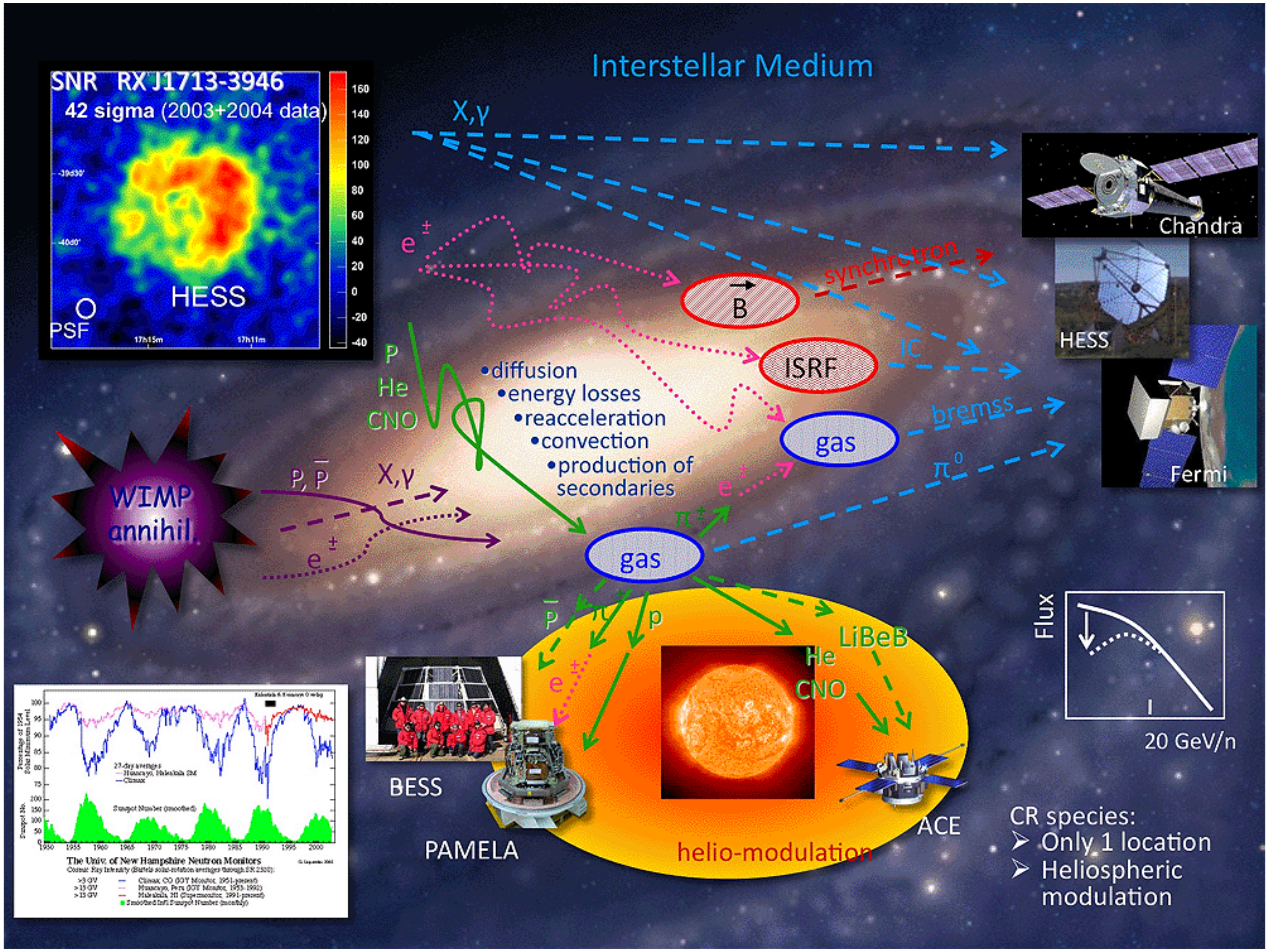




直线电场加速器：脉冲星



宇宙线传播模型



Moskalenko (2004)

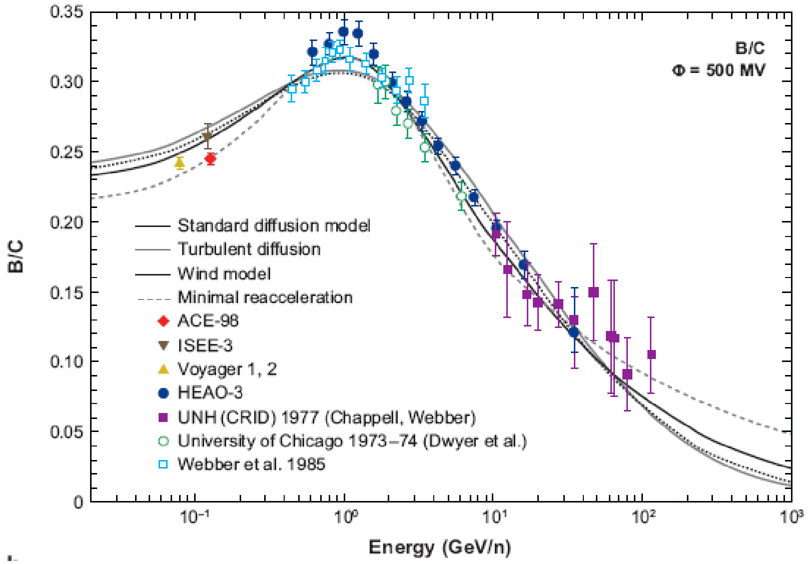
宇宙线传播模型

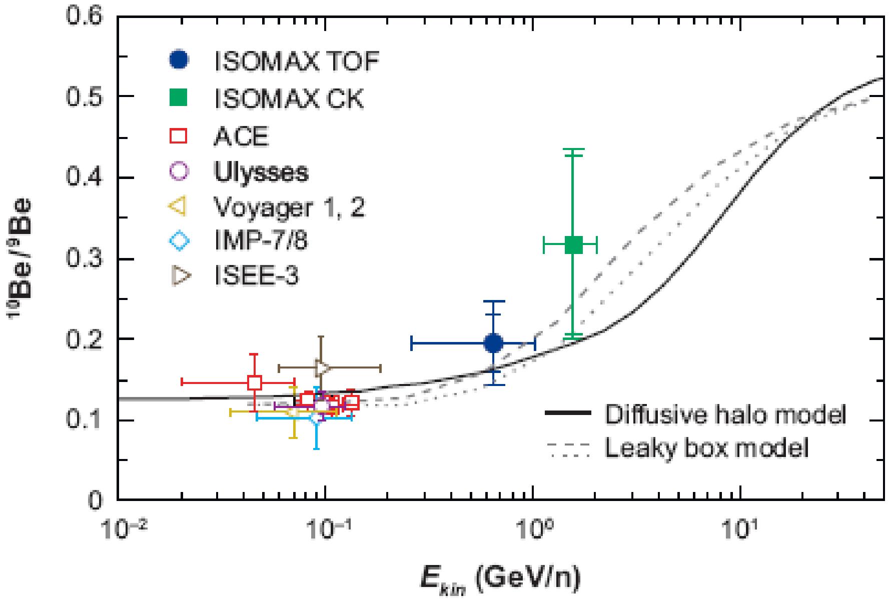
- 源函数(N,ν…)
- 随机磁场中的扩散 (D_0, δ)
- 对流(Vc)
- 重加速(V_A)
- 碎裂(次级粒子产生：B, e+, anti-p, γ, ν…)

$$\frac{\partial \psi}{\partial t} = Q(\mathbf{x}, p) + \nabla \cdot (D_{xx}\nabla\psi - \mathbf{V_c}\psi) + \frac{\partial}{\partial p} p^2 D_{pp} \frac{\partial}{\partial p} \times \frac{1}{p^2}\psi - \frac{\partial}{\partial p}\left[\dot{p}\psi - \frac{p}{3}(\nabla \cdot \mathbf{V_c}\psi)\right] - \frac{\psi}{\tau_f} - \frac{\psi}{\tau_r},$$

- ➢ 耦合方程组(O/C->Li/Be/B)
- ➢ 高维参数空间(N, ν, D_0, δ , Vc, V_A …)
- ➢ 非解析问题
- ➢ 天文输入(星际物质，辐射场，磁场等)

宇宙线传播模型





- 次级/原初粒子比值(B/C)->传播过程参数
- 原初粒子能谱->源参数

Strong et al. (2007)

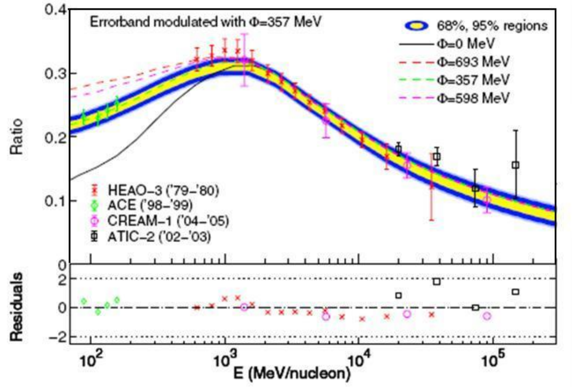
数值工具：GALPROP

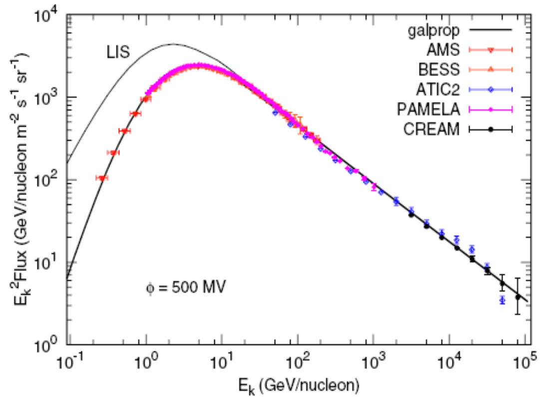
- 基于观测的星际气体分布，磁场模型，辐射场模型
- 完整的物理过程
- 完整的核碎裂链(Z=28)
- 自洽的次级粒子产生和传播、弥散伽玛射线的计算

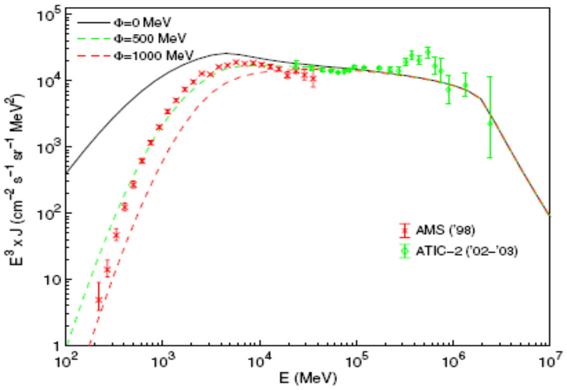
Strong & Moskalenko (1998)
Moskalenko & Strong (1998)
Strong et al. (2007)
http://galprop.stanford.edu/

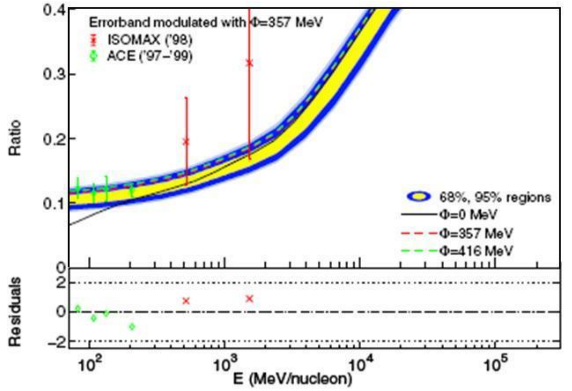
DRAGON: http://www.desy.de/~maccione/DRAGON/

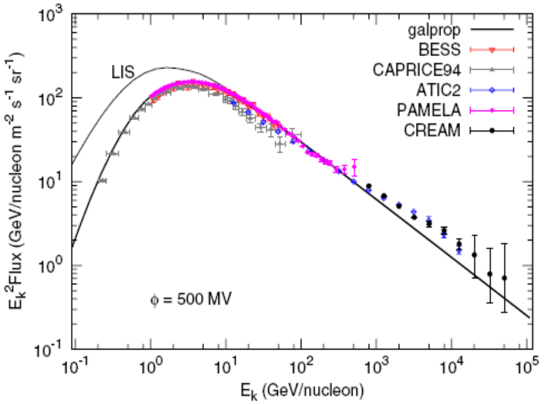
零阶近似下的成功模型

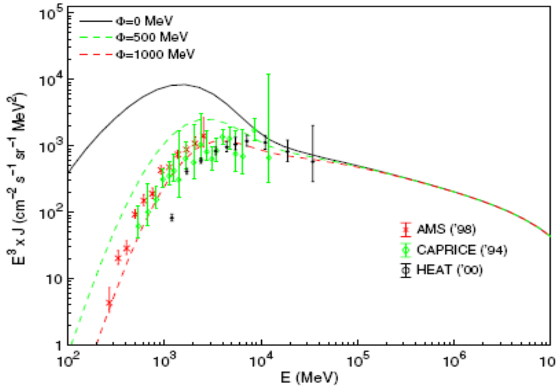


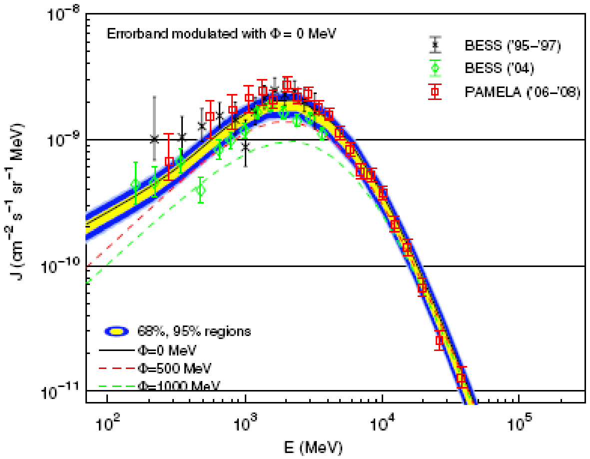


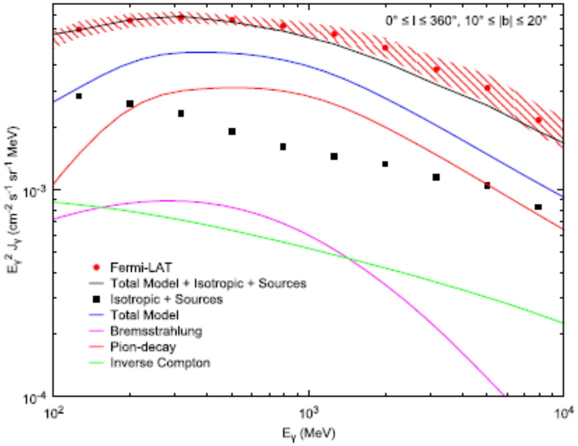






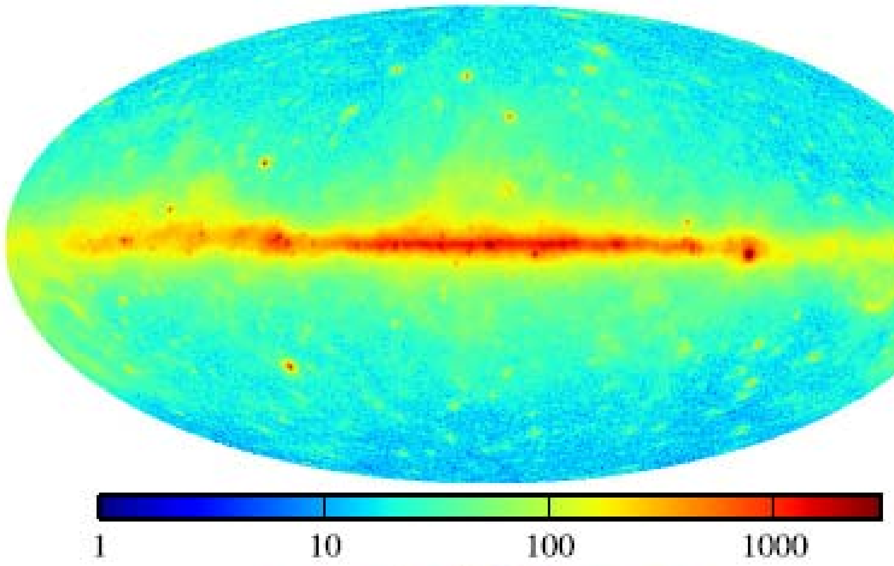




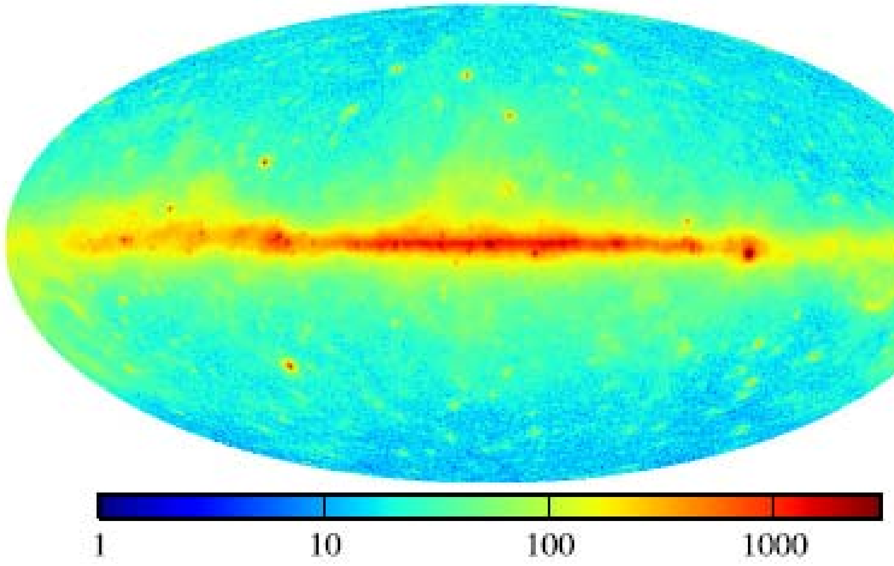


Trotta et al. (2011)

零阶近似下的成功模型

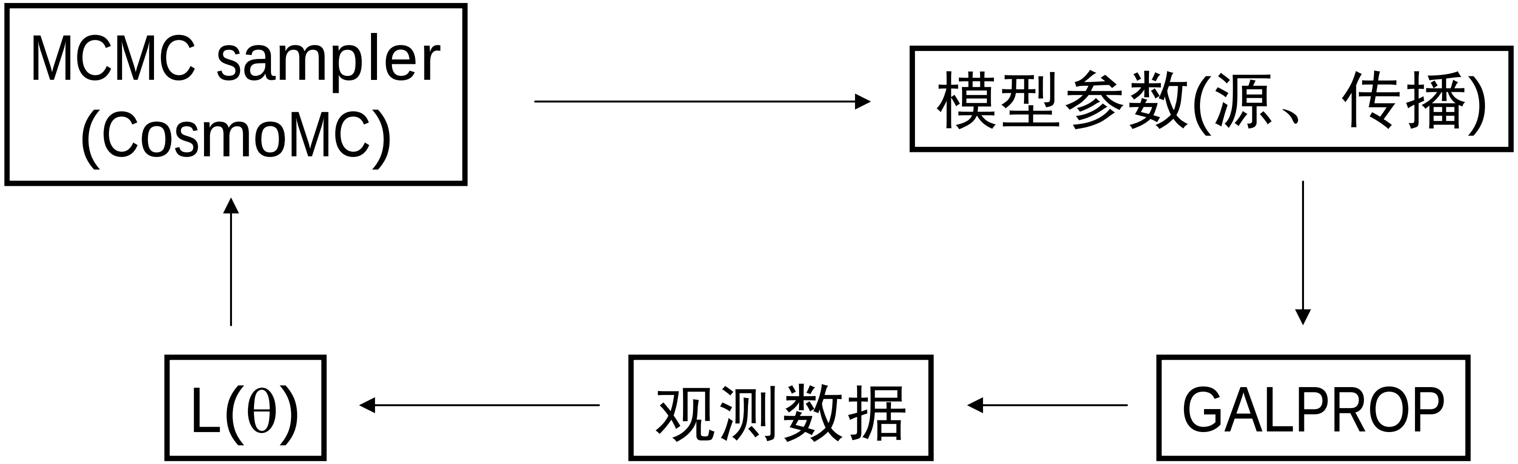


Fermi 0.2-100 GeV数据



模型预期

CosRayMC：CosmoMC+GALPROP



MCMC sampler
(CosmoMC)

模型参数(源、传播)

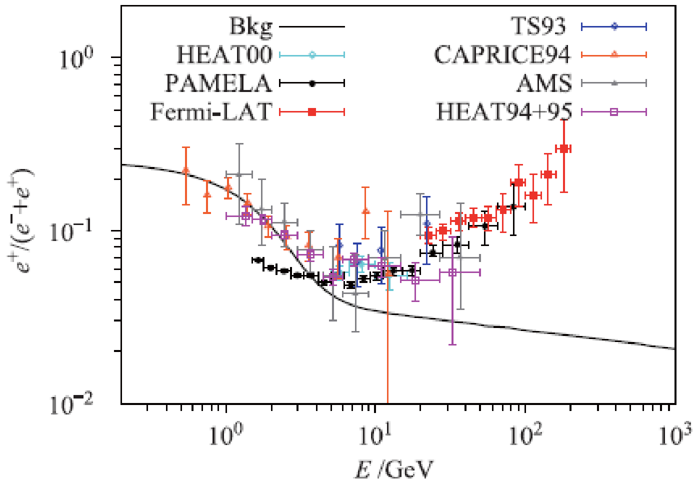
GALPROP

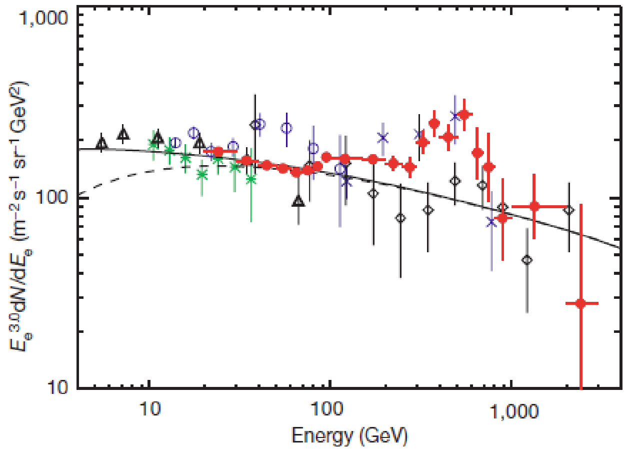
观测数据

L(θ)

Liu, Yuan et al. (2010; 2012)

宇宙线正负电子超出





- Pulsar like scenario

$$\mathrm{d}N/\mathrm{d}E \propto E^{-\alpha}\exp(-E/E_c)$$

$$f(R,z) \propto \left(\frac{R}{R_\odot}\right)^a \exp\left[-\frac{b(R-R_\odot)}{R_\odot}\right]\exp\left(-\frac{|z|}{z_s}\right)$$

γ_1, γ_2, $E_{\rm br}$, $A_{\rm bkg}$, ϕ, c_{e^+}, $A_{\rm psr}$, α and E_c

- Dark matter scenario

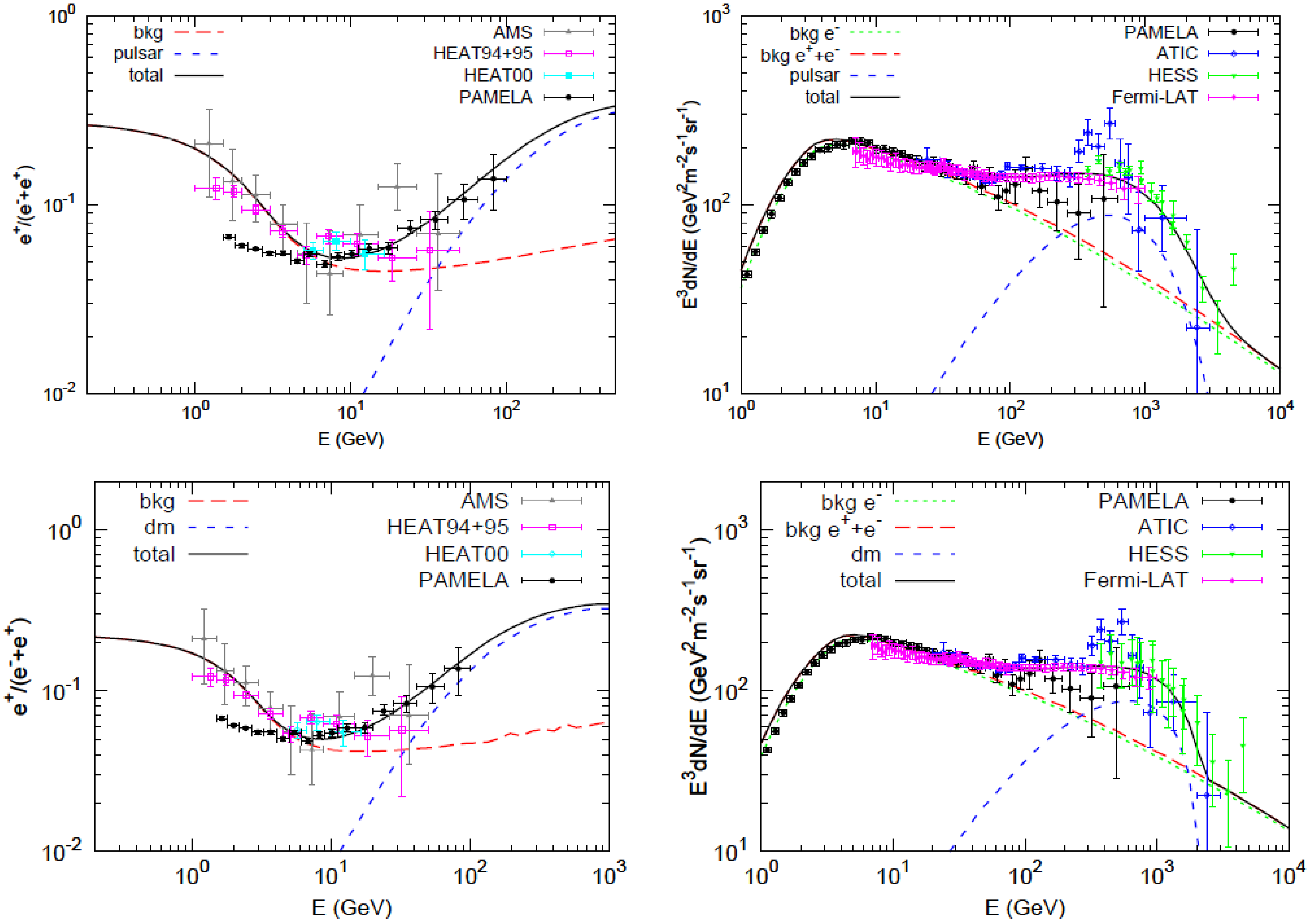
$$q(E,r) = \frac{\langle\sigma v\rangle}{2m_\chi^2}\sum_f B_f \left.\frac{\mathrm{d}N}{\mathrm{d}E}\right|_f \times \rho^2(r),$$

$$\rho(r) = \rho_{-2}\exp\left[-\frac{2}{\alpha}\left(\frac{r^\alpha}{r_{-2}^\alpha}-1\right)\right]$$

γ_1, γ_2, $E_{\rm br}$, $A_{\rm bkg}$, ϕ, c_{e^+}, m_χ, $\langle\sigma v\rangle$, B_e, B_μ, B_τ and B_q

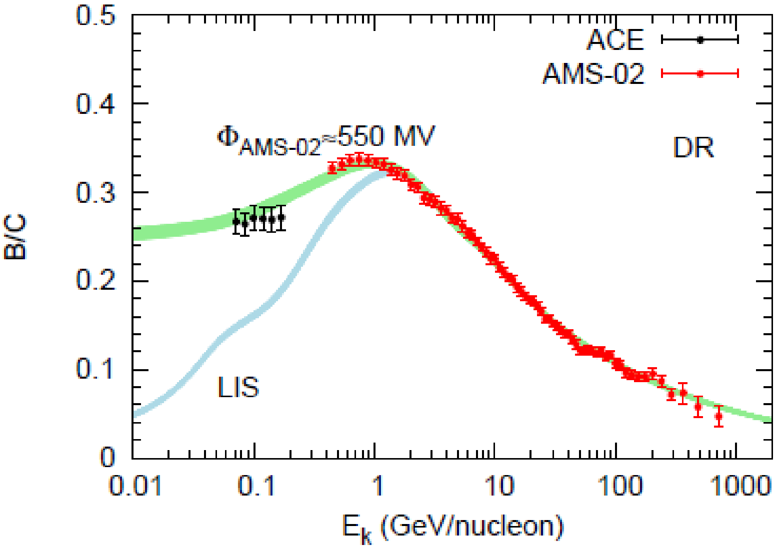
Liu, Yuan et al. (2010; 2012)

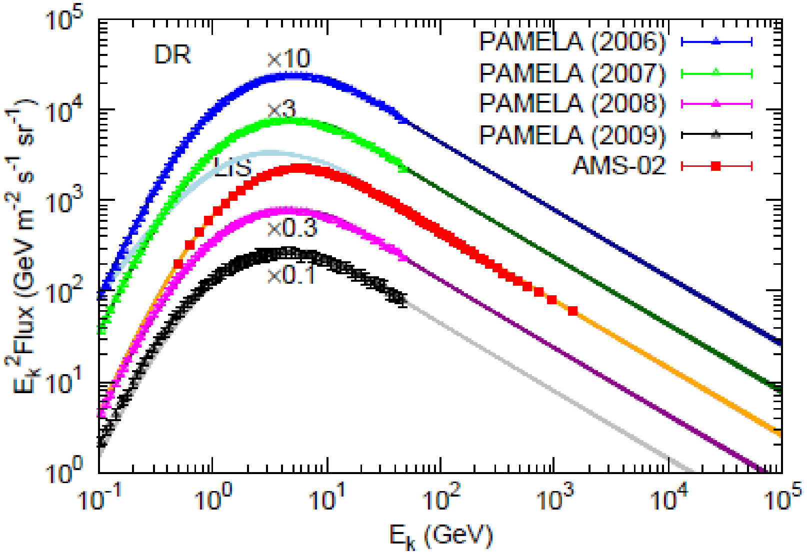
暗物质和脉冲星模型无法区分



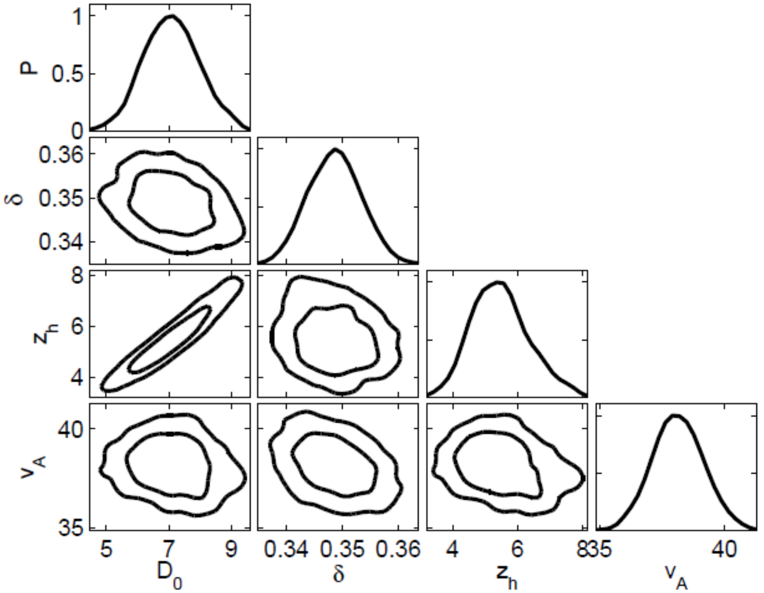
Liu, Yuan et al. (2010; 2012)

利用B/C确定传播参数





D_0	(10^{28}cm^2 s^{-1})	7.09 ± 0.87
δ		0.349 ± 0.005
z_h	(kpc)	5.47 ± 0.87
v_A	(km s^{-1})	38.14 ± 0.99
dV_c/dz	(km s^{-1} kpc^{-1})	...
R_0	(GV)	4
η		1.0
$\log(A_p^a)$		−8.364 ± 0.003
ν_1		1.702 ± 0.026
ν_2		2.399 ± 0.008
$\log(R_{br}^b)$		4.06 ± 0.03
Φ_0	(MV)	204 ± 12
Φ_1	(MV)	516 ± 21
χ^2/dof		164.0/157



结语

- Bayesian推断在处理冗余参数问题上有重要应用(警告：在弱约束情形下先验知识可能会左右结果)

- MCMC作为Bayesian框架下的一种算法实现，对高维参数空间的计算具有突出的优点，在很多领域里面获得广泛使用；在宇宙线研究中也展现出很好的潜力，特别是当数据积累越来越多，模型变得越来越复杂的情况下

中国空间科学首星——暗物质粒子卫星(DAMPE)于2015年12月17日成功发射！

- 招聘DAMPE fellow (DAMPE杰出学者、骨干学者)
- 欢迎研究生、博士后、初级、高级等各层次人才加盟！

致谢

- 报告中的图片来自网络和文献，未予一一标注，仅作教学使用
- 感谢Yue-Lin Sming Tsai提供部分素材

谢谢大家！