Introduction to Cosmic Rays
Data Analysis Issues

Nicola De Simone

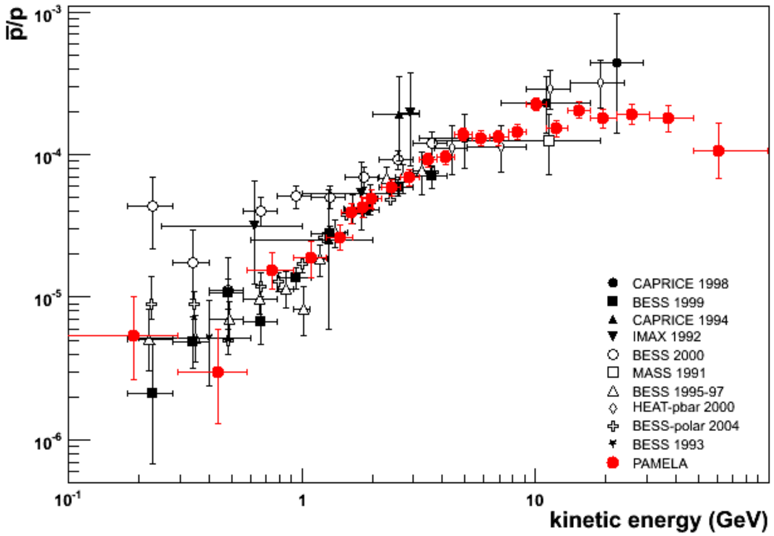
INFN and University of Rome Tor Vergata

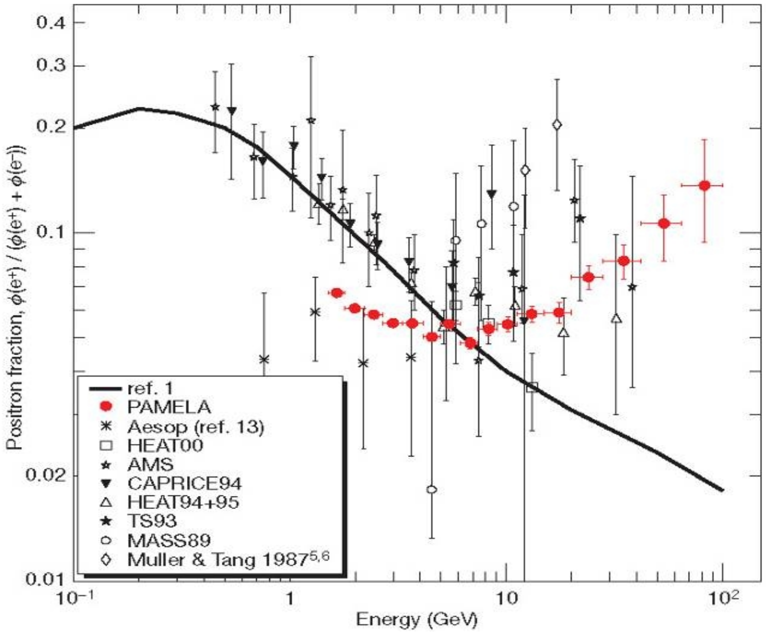
SciNeGHE 2010 - Data Analysis Tutorial

Trieste, *September 8 - 9, 2010*

The physics of PAMELA

PRL 102, (2009) 051101, Astro-ph 0810.4994



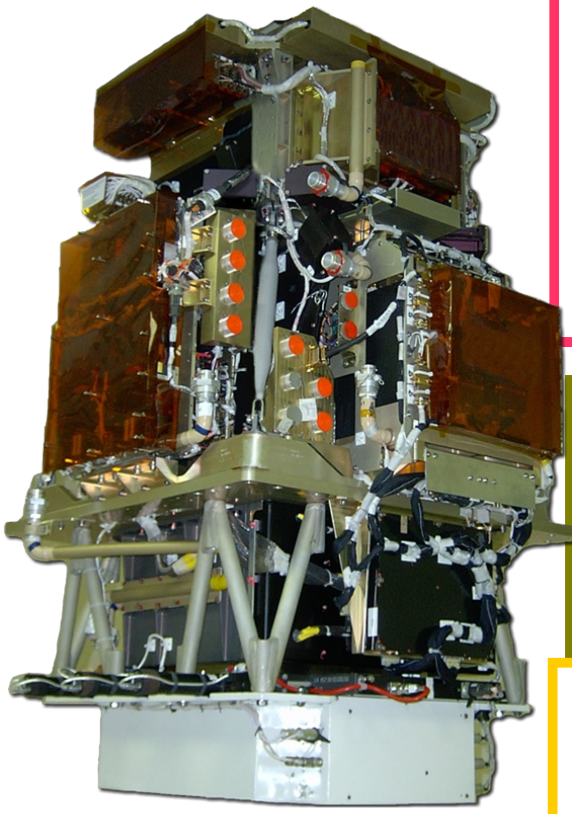


Nature 458 (2009) 607, Astro-ph 0810.4995

<u>Scientific objectives of particle detectors in space:</u>

- Study antiparticles in cosmic rays
    - Search for primordial antimatter (antihelium)
    - Search for dark matter annihilation (e^+ and pbar spectra)
- Study of cosmic-ray production, acceleration and propagation
- Study of solar physics and solar modulation
- Study of the terrestrial magnetosphere

PAMELA instrument



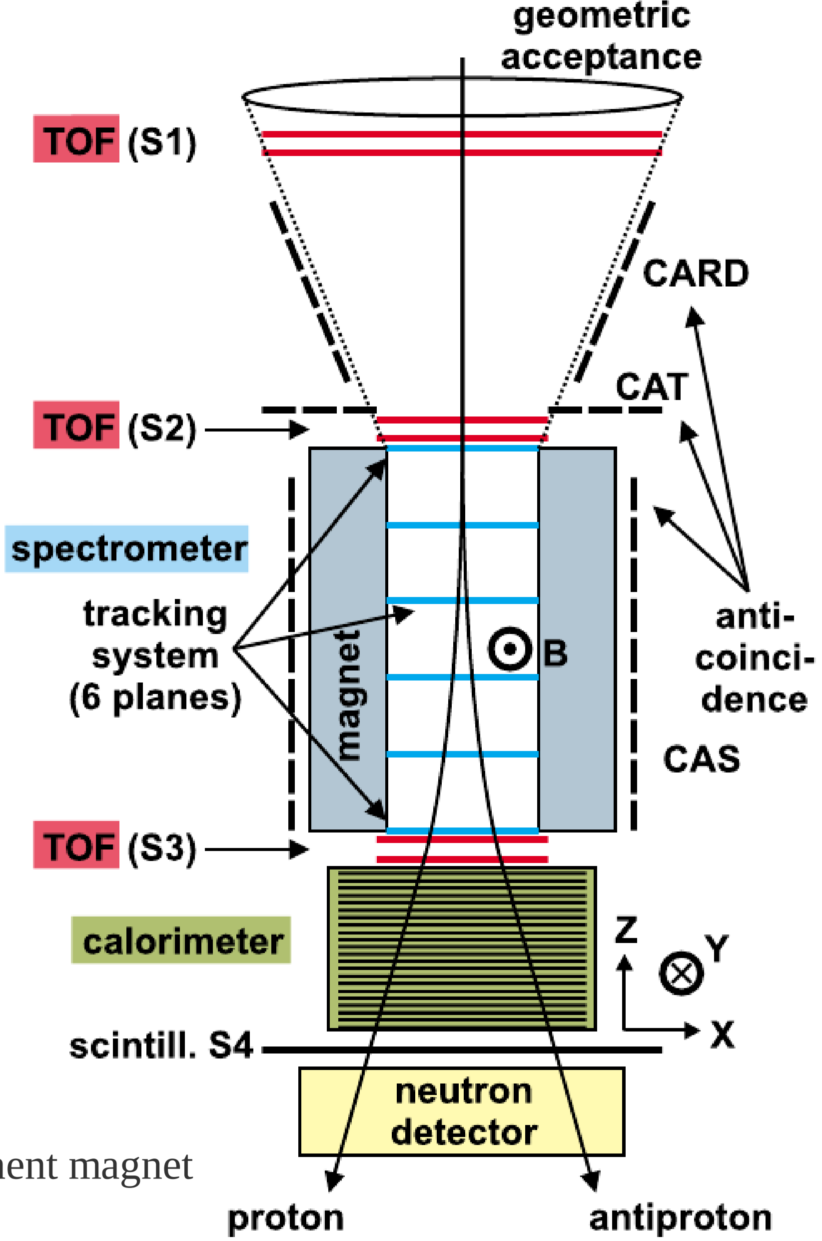
GF: 21.5 cm^2 sr
Mass: 470 kg
Size: 130x70x70 cm^3
Power Budget: 360W

Time-Of-Flight
plastic scintillators + PMT:
- Trigger
- Albedo rejection;
- Mass identification up to 1 GeV;
- Charge identification from dE/dX.

Electromagnetic calorimeter
W/Si sampling (16.3 X_0, 0.6 λI)
- Discrimination e+ / p, anti-p / e$^-$ (shower topology)
- Direct E measurement for e$^-$

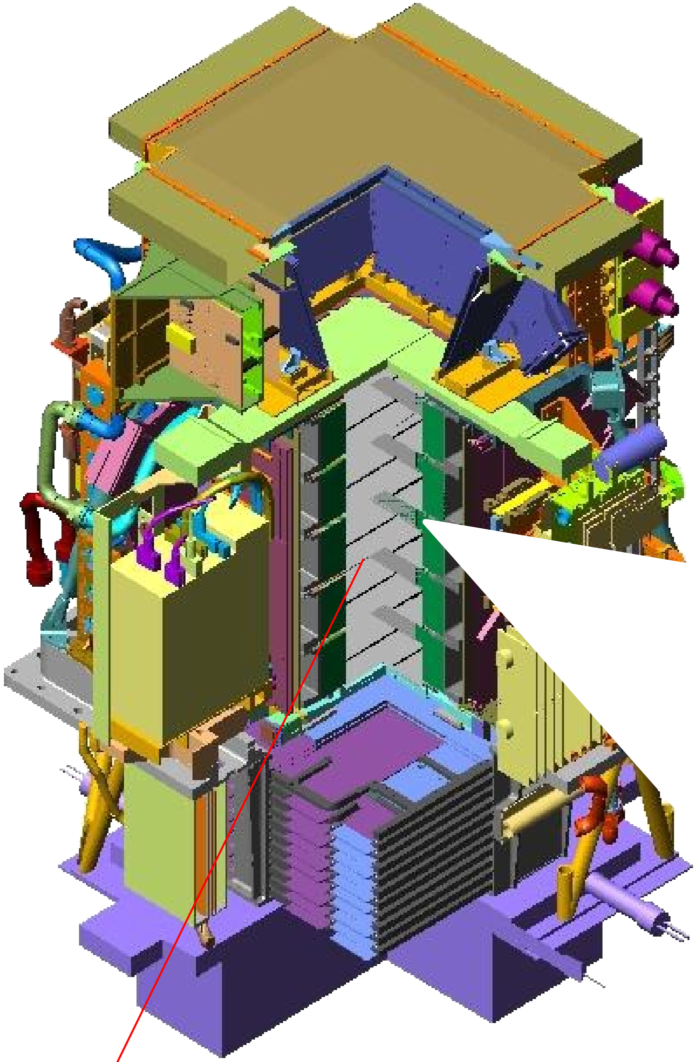
Neutron detector
^{3}He tubes + PMT:
- High-energy e/h discrimination

Spectrometer
microstrip silicon tracking system + permanent magnet
It provides:
- *Magnetic rigidity* → R = pc/Ze
  - *Charge sign*
  - *Charge value from dE/dx*

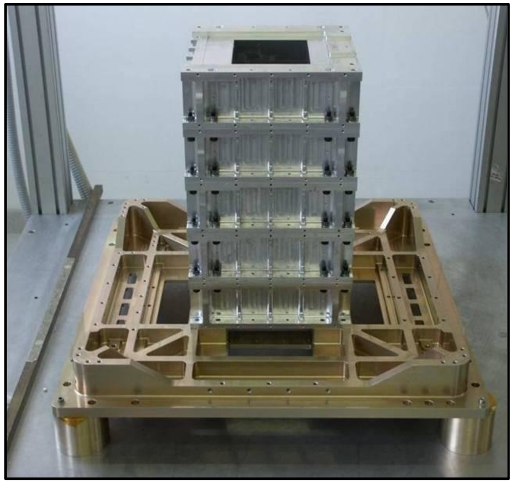


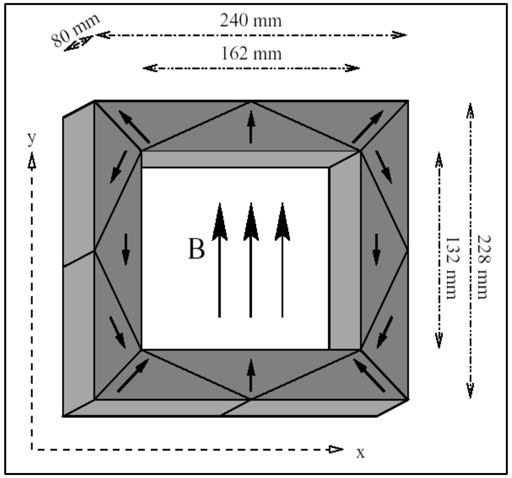
PAMELA

The magnet



SPECTROMETER

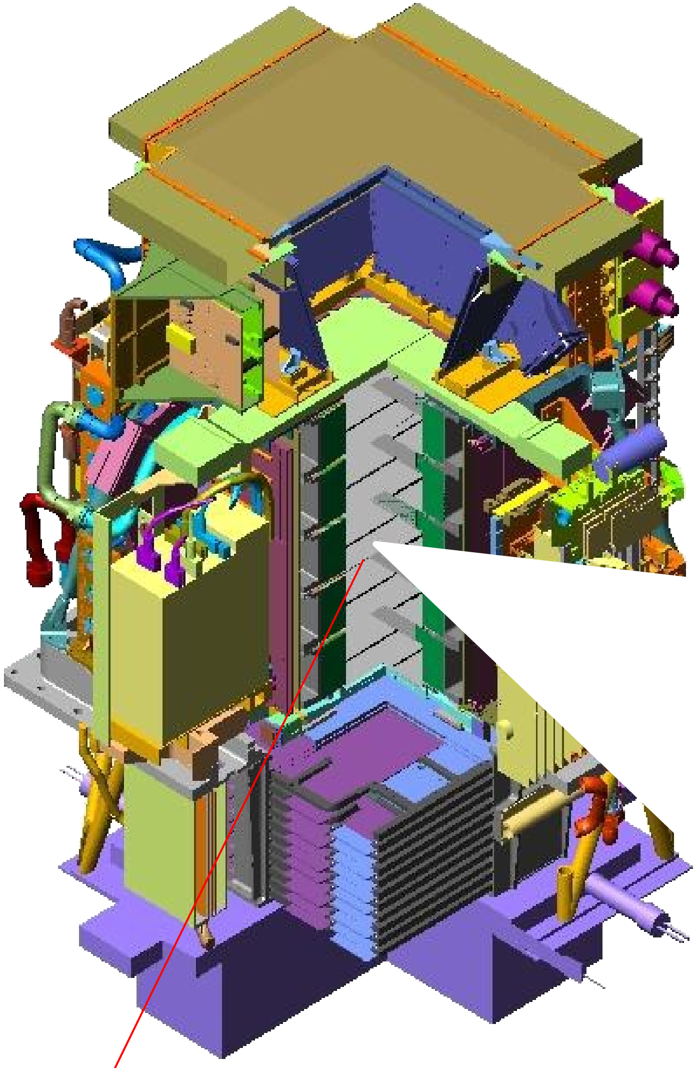




Characteristics:

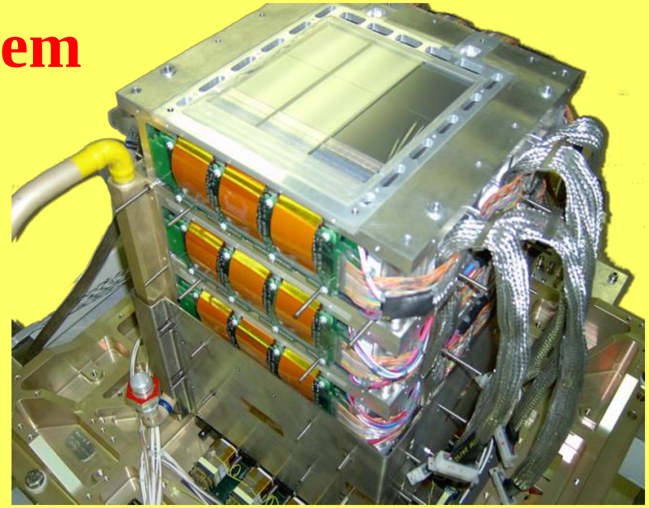
- **5 modules of permanent magnet (Nd-B-Fe alloy) in aluminum mechanics**
- Cavity dimensions 162x132x445 cm^3
  → **GF 21.5 cm^{2}sr**
- Magnetic shields
- 5mm-step field-map
- **B=0.43 T**(average along axis), B=0.48 T (@center)

PAMELA



SPECTROMETER

The tracking system



Main tasks:

- Rigidity measurement
- Sign of electric charge
- dE/dx

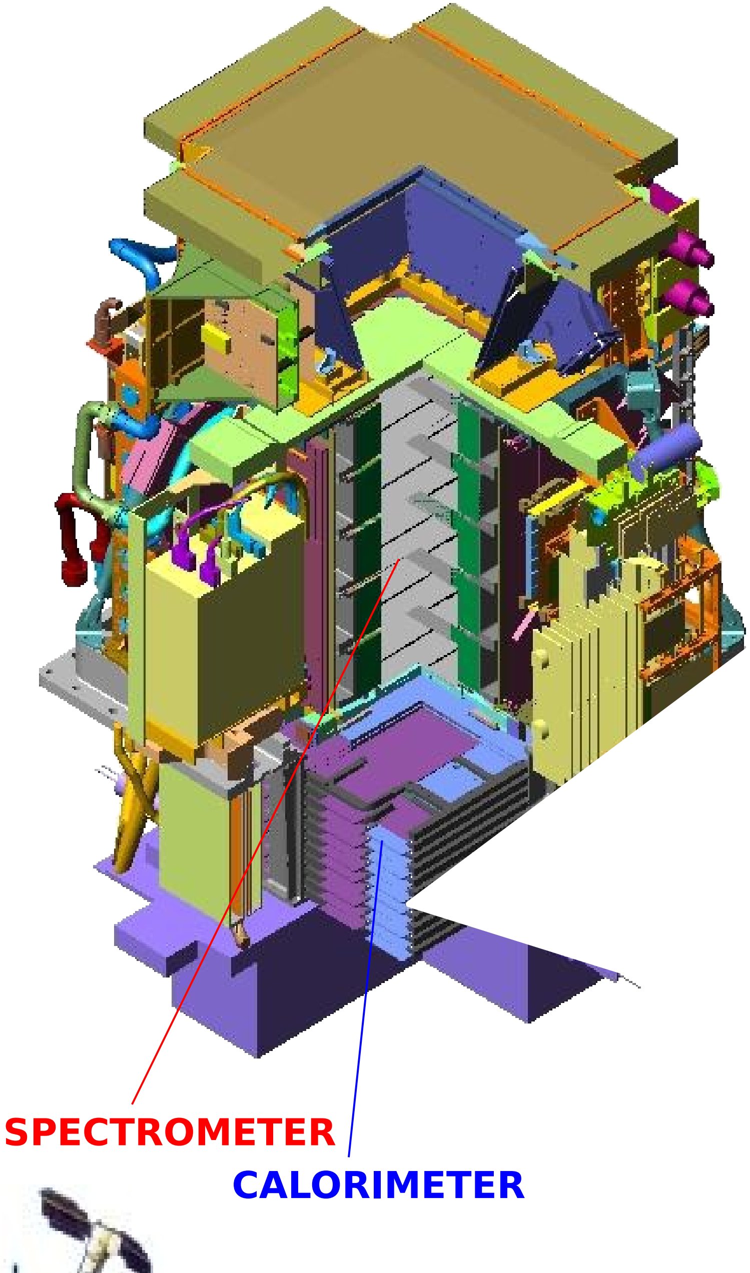
Characteristics:

- **6 planes double-side (x&yview) microstrip Si sensors**
- 36864 channels
- Dynamic range 10 MIP

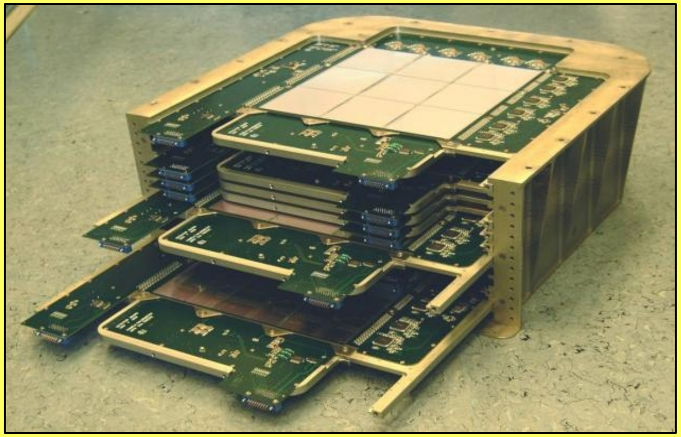
Performances:

- Spatial resolution: **3 ÷4 μm**
- **MDR ~1.2TV**(from flight data)

PAMELA



The electromagnetic calorimeter



Main tasks:

- e/h discrimination
- $e^{+/-}$ energy measurement

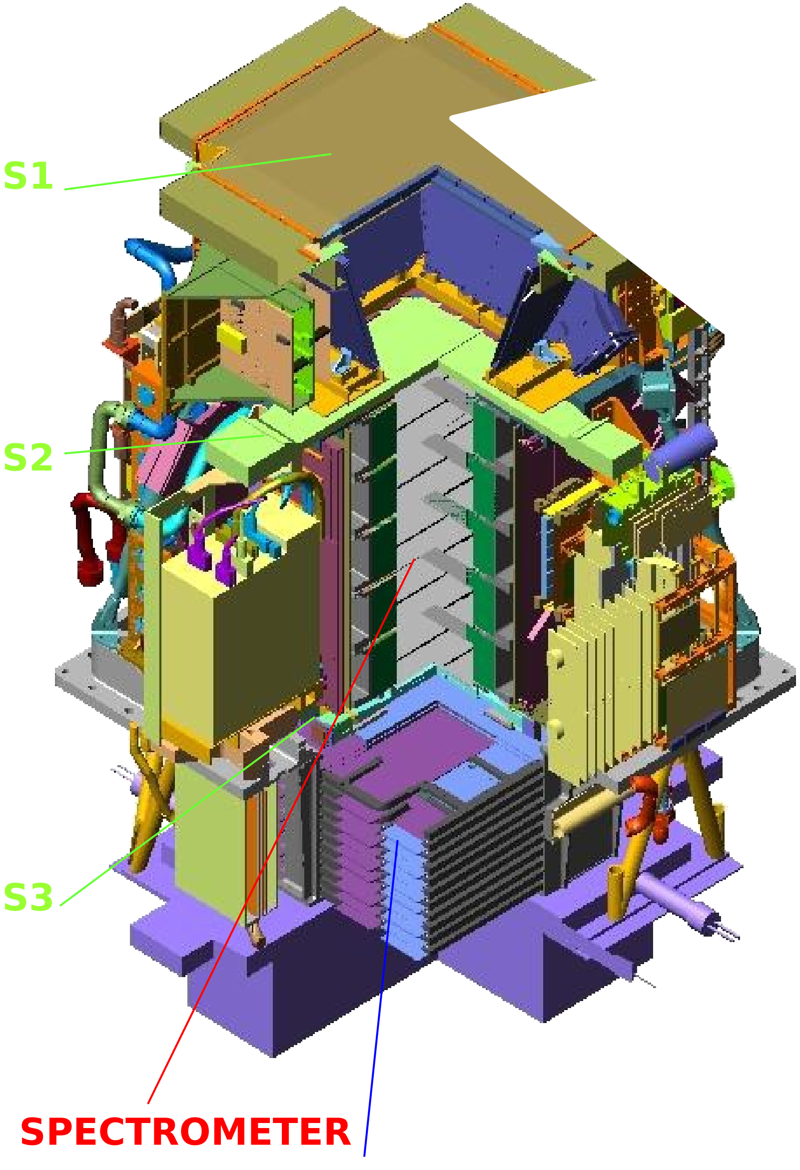
Characteristics:

- **44 Si layers (X/Y) +22 W planes**
- **$16.3\ X_o / 0.6\ l_0$**
- 4224 channels
- Dynamic range ~1100 mip
- Self-trigger mode (> 300 GeVGF~600 cm^2 sr)

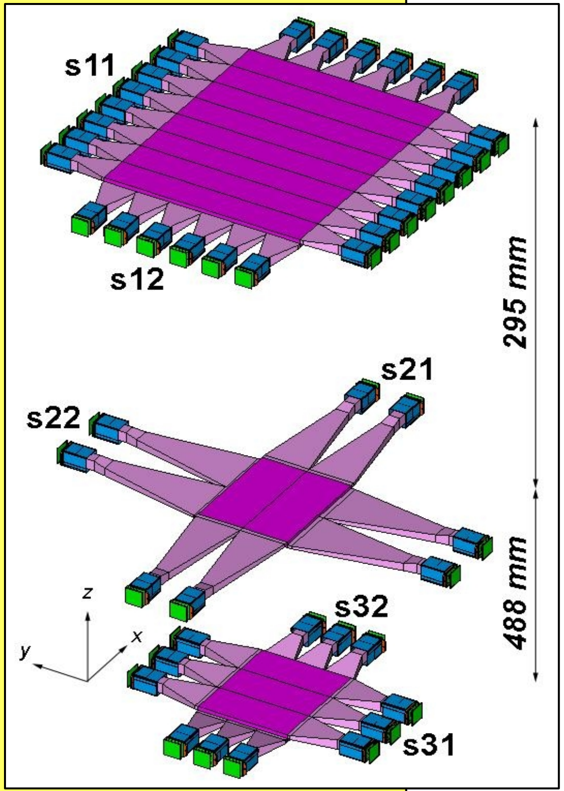
Performances:

- p/e^+ selection efficiency ~90%
- p rejection factor **10^5**
- e rejection factor **$>10^4$**
- Energy resolution **~5%**@200GeV

PAMELA



The time-of-flight system

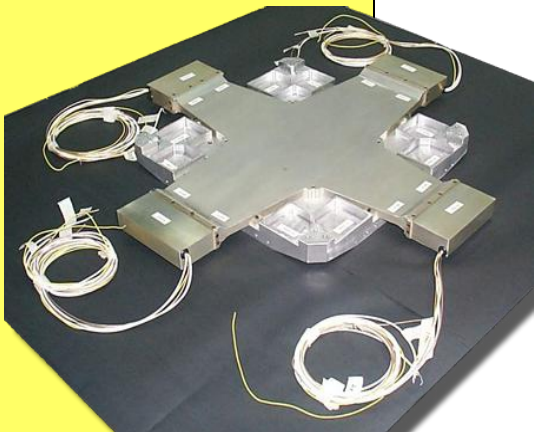


Main tasks:

- First-level trigger
- Albedo rejection
- dE/dx
- Particle identification (<1GeV/c)

Characteristics:

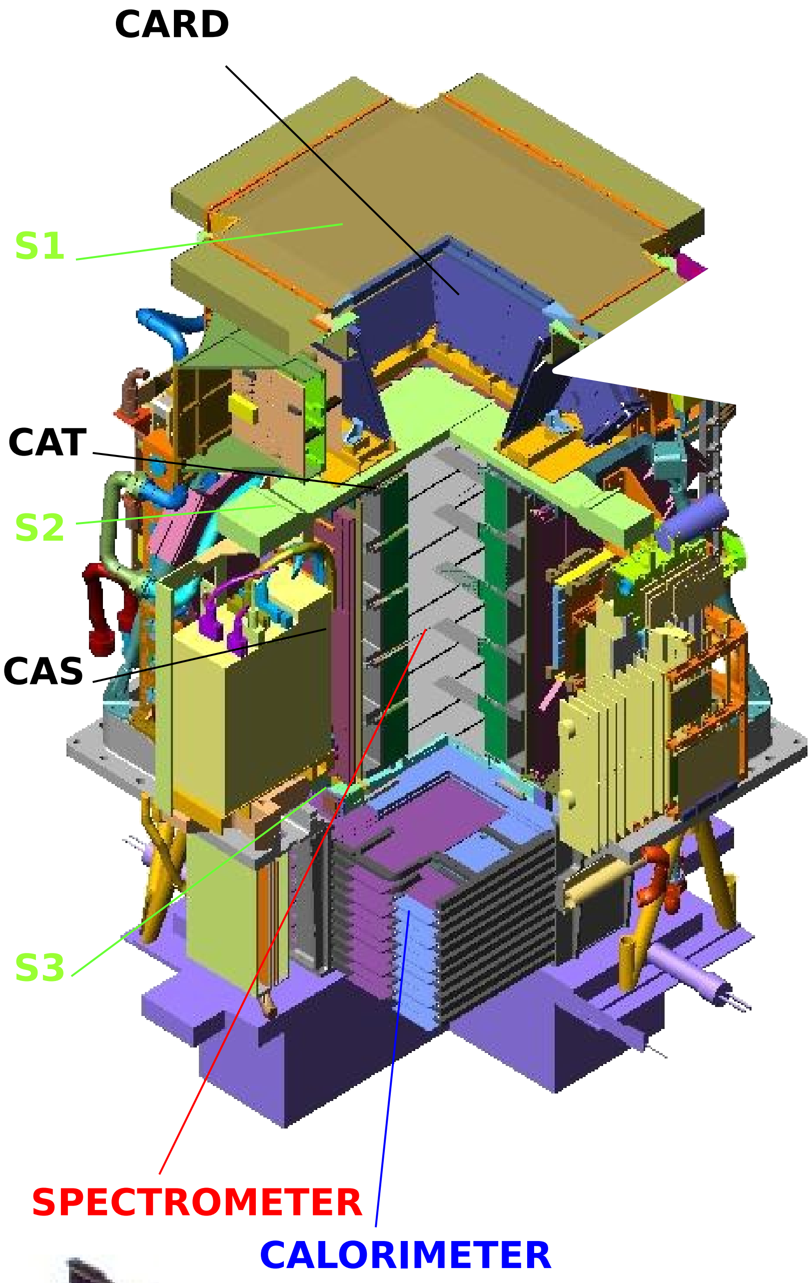
- 3 double-layer scintillator paddles
- X/Y segmentation
- Total: 48 Channels



Performances:

- σ_{paddle} ~110ps
- σ_{TOF} ~330ps(for MIPs)

PAMELA



The anticounter shields

Main tasks:

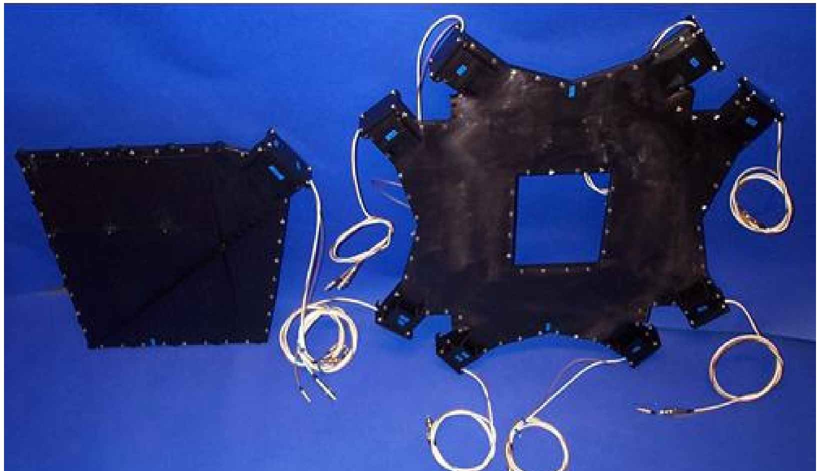
- Rejection of events with particles interacting with the apparatus (off-line and second-level trigger)

Characteristics:

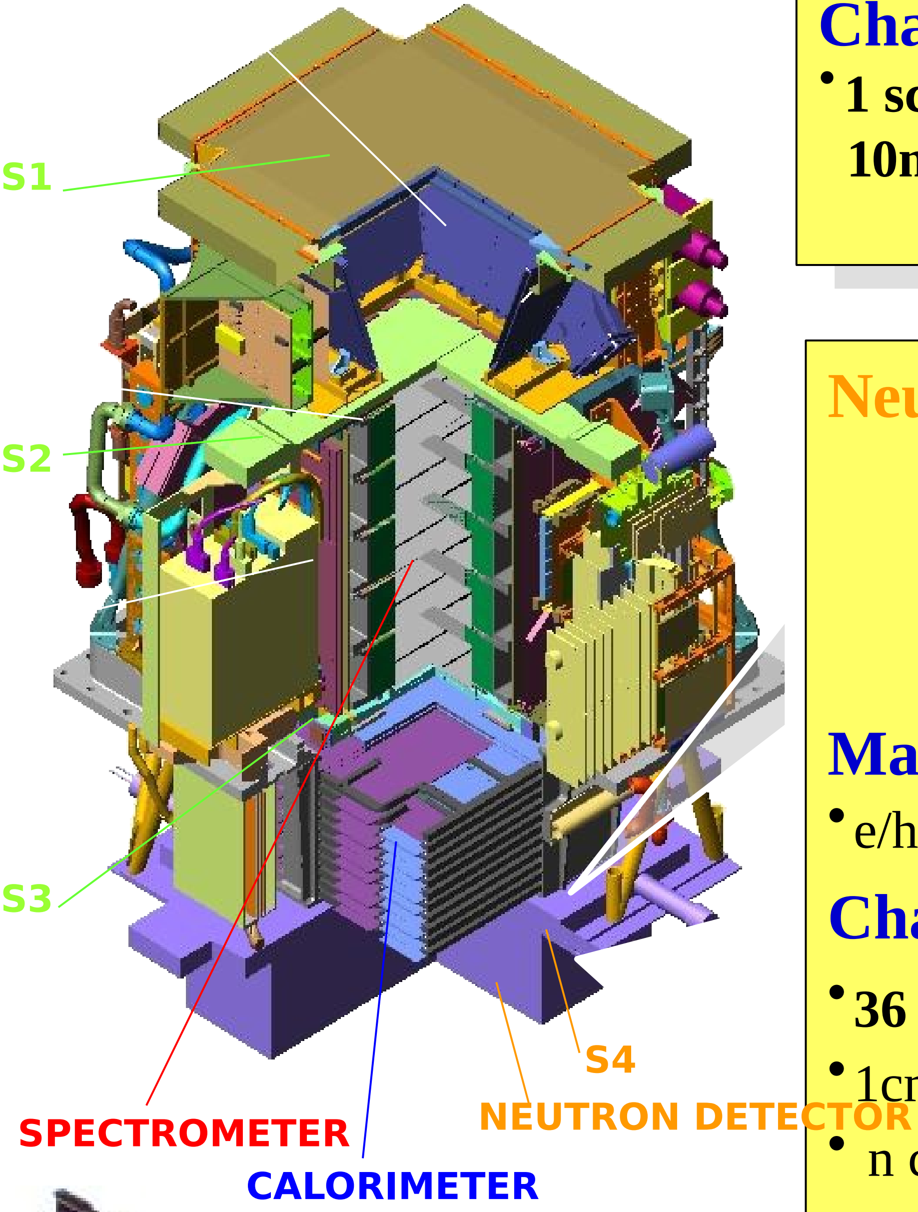
- Scintillator paddles 10mm thick
- 4 up (CARD), 1 top (CAT), 4 side (CAS)

Performances:

- Efficiency > 99.9%



PAMELA



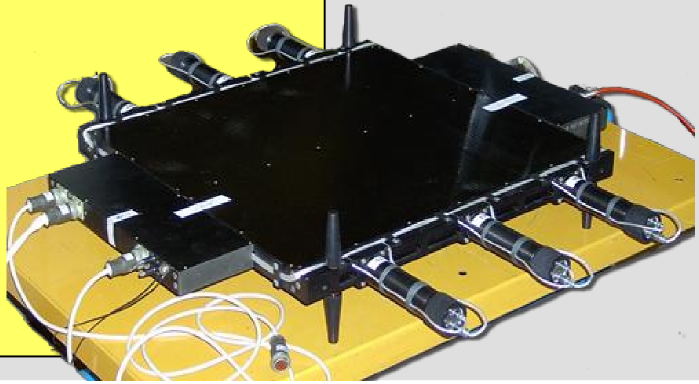
Shower-tail catcher (S4)

Main tasks:

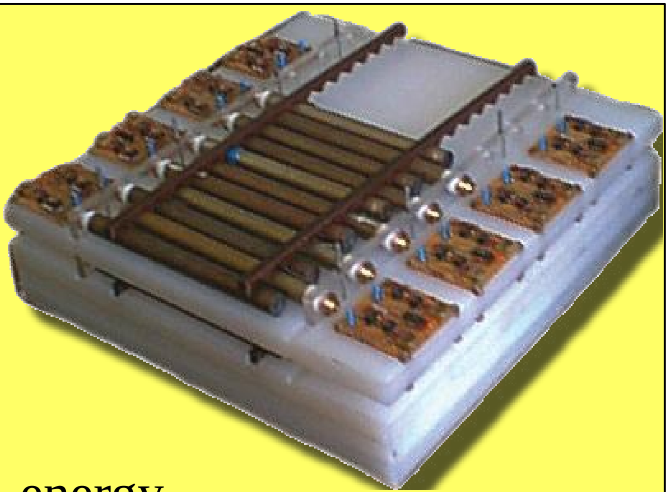
- **ND trigger**

Characteristics:

- **1 scintillator paddle 10mm thick**



Neutron detector

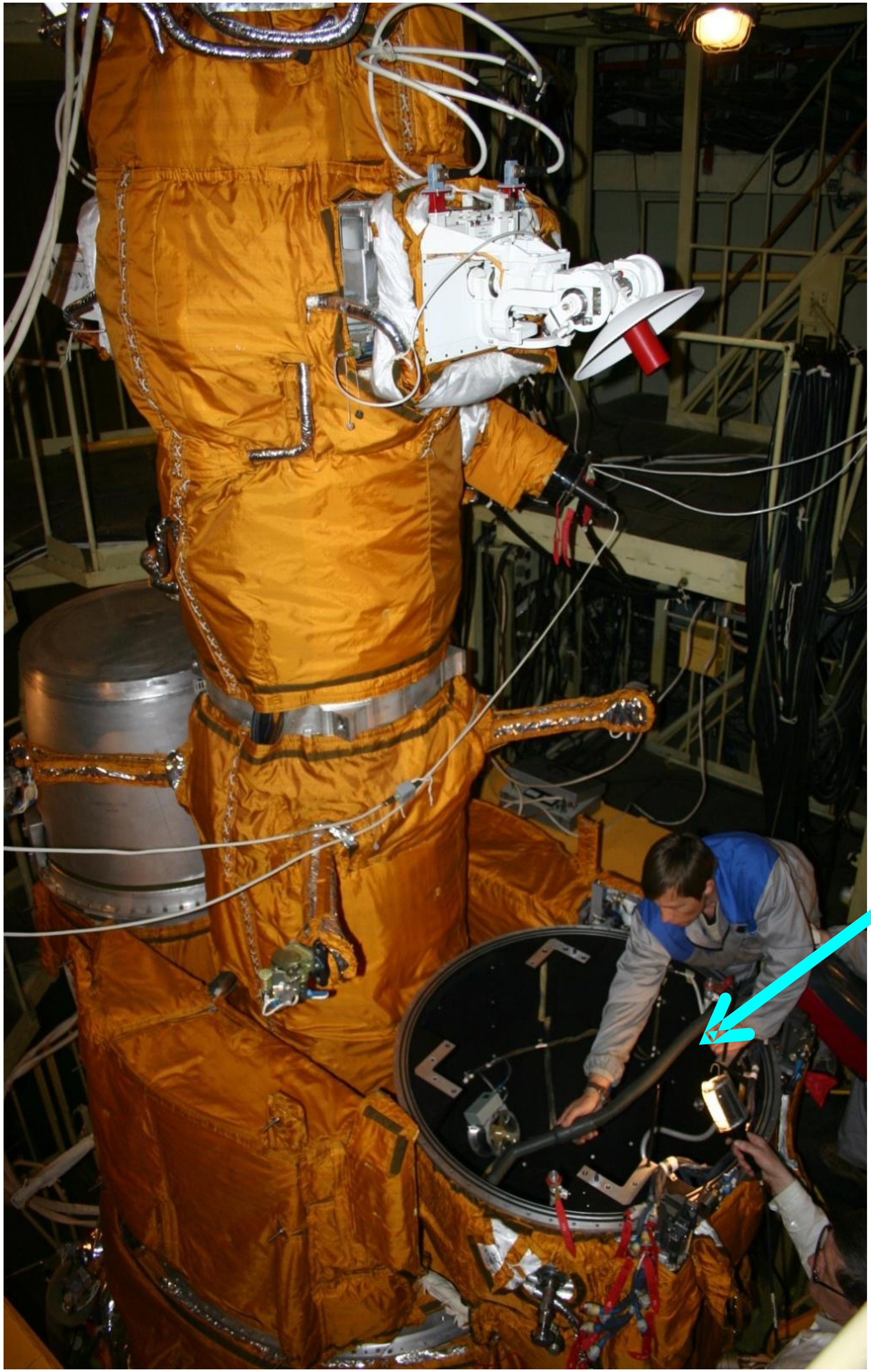


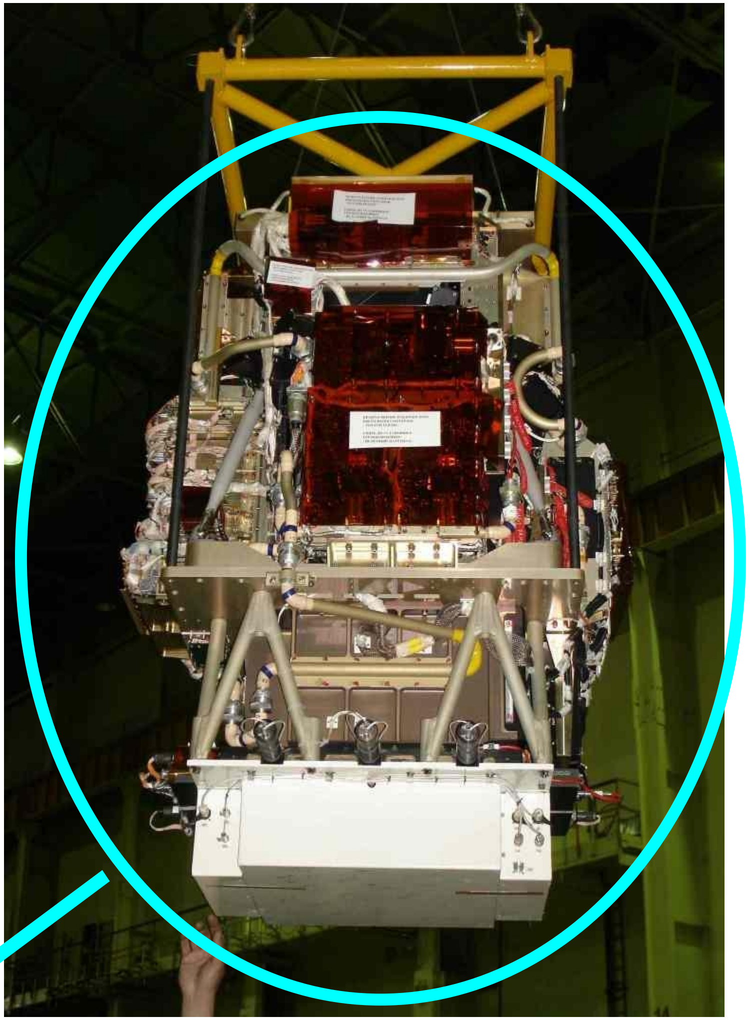
Main tasks:

- e/h discrimination @high-energy

Characteristics:

- **36 ^{3}He counters**: ^{3}He(n,p)T → Ep=780 keV
- 1cm thick polyetilenemoderators
- n collected within 200 ms time-window

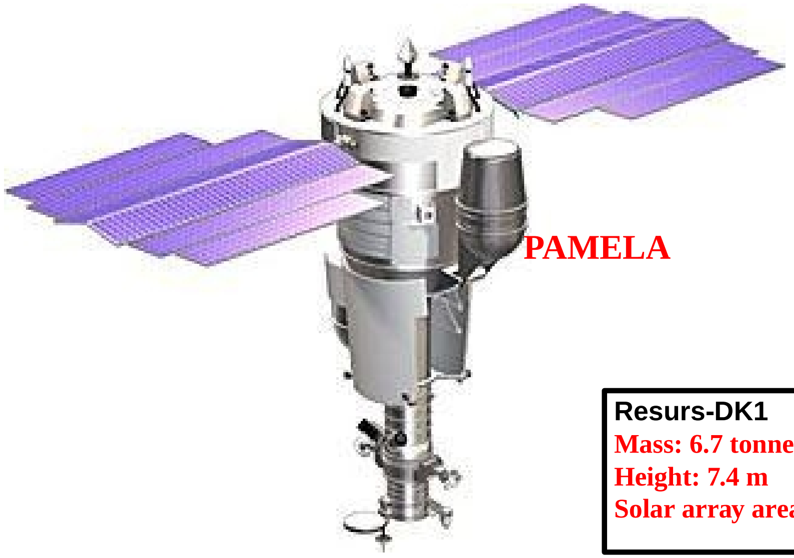




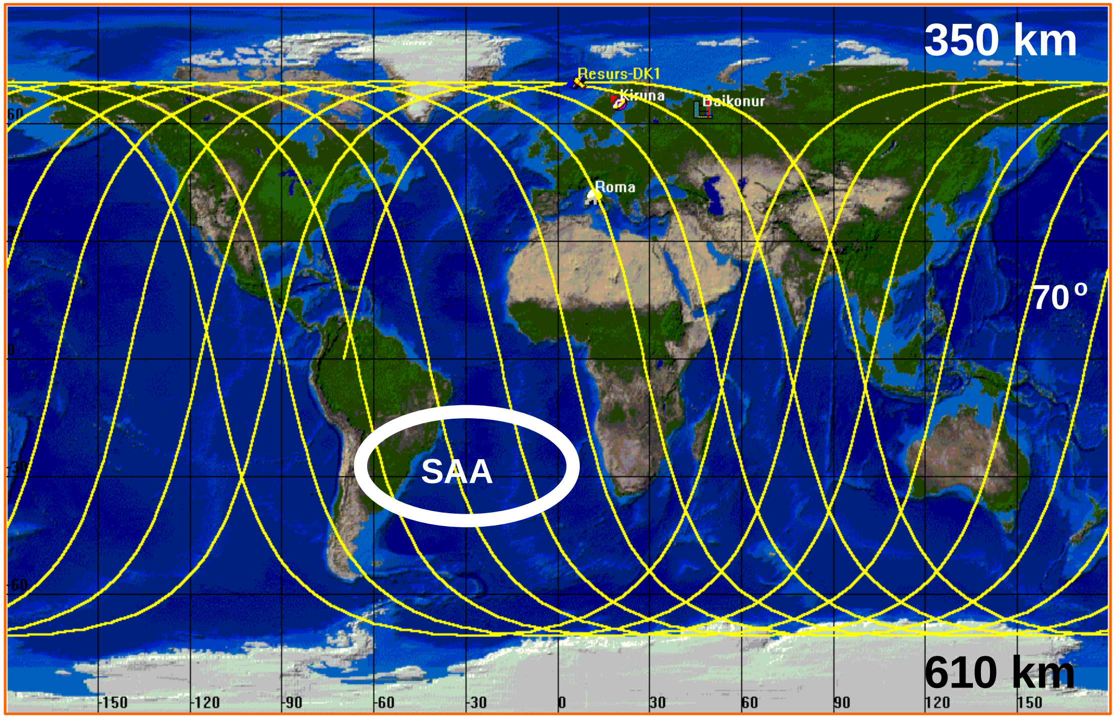
PAMELA INTEGRATION in the RESURS-DK1 satellite

The Launch: 15th June 2006

Resurs-DK1 satellite + orbit



Resurs-DK1
Mass: 6.7 tonnes
Height: 7.4 m
Solar array area: 36 m^2



- Resurs-DK1: multi-spectral imaging of earth's surface
- PAMELA mounted inside a pressurized container
- Lifetime >3 years (assisted, first time last February)
- Data transmitted to NTsOMZ, Moscow via high-speed radio downlink. ~16 GB per day
- Quasi-polar and elliptical orbit (70.0°, 350 km -600 km)
- Traverses the South Atlantic Anomaly
- Crosses the outer (electron) Van Allen belt at south pole

How many data?

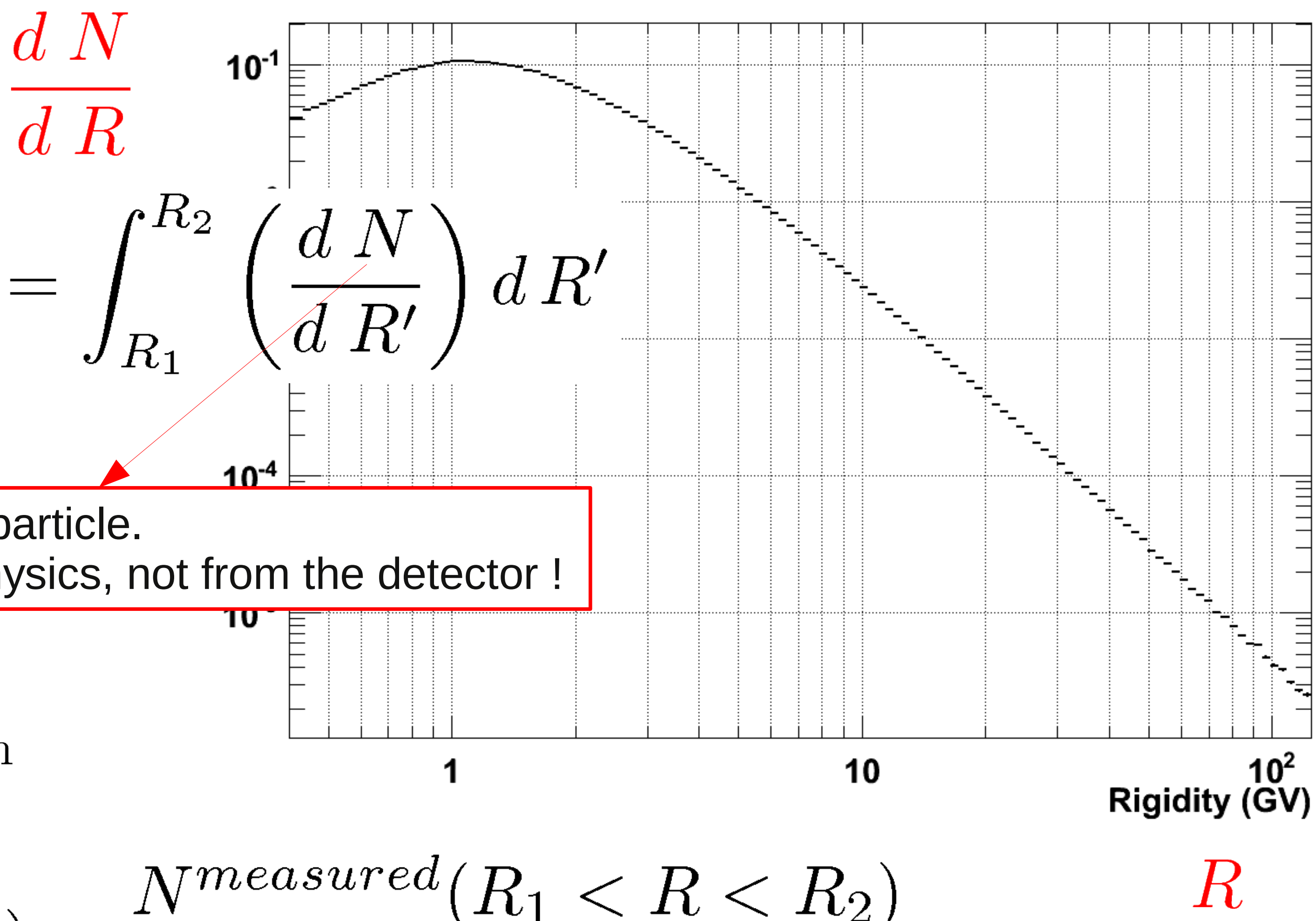
Trigger rate* ~25Hz
Fraction of live time* ~ 75%
Event size (compressed mode) ~5kB
25 Hz x 5 kB/ev → ~ 10 GB/day
(*outside radiation belts)

Till ~now:
~1400 days of data taking
~20 Tbyte of raw data downlinked
>2x10^{9} triggers recorded and analyzed
(Data till January 2010 under analysis)


Main antenna in NTsOMZ

What kind of measurements?

- Differential particle fluxes



$$N(R_1 < R < R_2) = \int_{R_1}^{R_2} \left(\frac{d\,N}{d\,R'} \right) d\,R'$$

N = **TRUE** number of particle.
It depends from the physics, not from the detector !

$$R = \text{Rigidity} = \frac{p}{q}$$

$p = \text{relativisitc momentum}$

$q = \text{charge}$

$$N(R_1 < R < R_2) = \frac{N^{measured}(R_1 < R < R_2)}{efficiency}$$

Need to know the efficiency !

What kind of measurements?

- Ratios antiparticle / particle.

  e.g. antiprotons / protons:

$$Ratio(R) = \frac{Flux_{antip}(R)}{Flux_p(R)} = \frac{N_{antip}^{meas}(R)/\varepsilon_{antip}(R)}{N_p^{meas}(R)/\varepsilon_p(R)} \underset{\underset{\varepsilon_{antip}(R) = \varepsilon_p(R)}{\Big\downarrow}}{=} \frac{N_{antip}^{meas}(R)}{N_p^{meas}(R)}$$

$\varepsilon_{antip} = \text{efficiency for antiproton selection}$

$\varepsilon_p = \text{efficiency for proton selection}$

“No need” to know the efficiency !

Some definition

- We need to develop a set of **selection criteria.**
- For each **selection S** we have to study:
  - **Efficiency**: fraction of good events selected by S

$$\varepsilon_S = \frac{N_S^{good}}{N^{good}}$$

  - **Contamination**: fraction of background events selected by S

$$C_S = \frac{N_S^{backg}}{N^{backg}}$$

Example in antiparticle analysis

good	backg
e^+	p
$\bar{p}$	p, e^-

$$\frac{signal}{noise} = \frac{N_S^{good}}{N_S^{backg}} = \boxed{\frac{\varepsilon_S}{C_S}} \frac{N^{good}}{N^{back}}$$

Rejection factor

What rejection power do we need?

The acceptable level of contaminations must be always put in relations, in flight, to the signal statistic over the background statistic.

Let's consider the case of **antiproton selection...**

What rejection power do we need?

Signal: antiprotons
Background: protons

	Number(*)	Charge	Cal. shower	Performance
antiprotons	$O(10^3)$	-	h	80 MeV - 190 GeV
protons	$O(10^7)$	+	h	80 MeV - 1.2 TeV
positrons	$O(10^4)$	+	em	50 MeV - 300 GeV
electrons	$O(10^5)$	-	em	50 MeV - 500 GeV

- ➔ Charge separation by the spectrometer.
- ➔ Contamination from spillover protons.

*: Exact numbers are energy dependent

What rejection power do we need?

Signal: antiprotons
Background: protons

	Number(*)	Charge	Cal. shower	Performance
antiprotons	$O(10^3)$	-	h	80 MeV - 190 GeV
protons	$O(10^7)$	+	h	80 MeV - 1.2 TeV
positrons	$O(10^4)$	+	em	50 MeV - 300 GeV
electrons	$O(10^5)$	-	em	50 MeV - 500 GeV

- → Charge separation by the spectrometer.
- → Contamination from spillover protons.

E.g.: Is a contamination of 0.01% of protons in the antiproton sample acceptable?

$$RF \approx 10^4 \rightarrow \frac{\text{signal}}{\text{noise}} = 10^4 \frac{10^3}{10^7} = 1$$

NO!

What rejection power do we need?

Signal: antiprotons
Background: protons

	Number(*)	Charge	Cal. shower	Performance
antiprotons	$O(10^3)$	-	h	80 MeV - 190 GeV
protons	$O(10^7)$	+	h	80 MeV - 1.2 TeV
positrons	$O(10^4)$	+	em	50 MeV - 300 GeV
electrons	$O(10^5)$	-	em	50 MeV - 500 GeV

- ➔ Charge separation by the spectrometer.
- ➔ Contamination from spillover protons.
- ➔ Rejection Factor $\gg 10^4$ needed → strong TRK quality requir
  → high energy limit 190 GeV

*: Exact numbers are energy dependent

What rejection power do we need?

Signal: antiprotons
Background: positrons

	Number(*)	Charge	Cal. shower	Performance
antiprotons	$O(10^3)$	-	h	80 MeV - 190 GeV
protons	$O(10^7)$	+	h	80 MeV - 1.2 TeV
positrons	$O(10^4)$	+	em	50 MeV - 300 GeV
electrons	$O(10^5)$	-	em	50 MeV - 500 GeV

- Charge separation by the spectrometer.
- Different calorimeter shower profile.
- Rejection Factor $\gg$ 10 is easily achieved.

*: Exact numbers are energy dependent

What rejection power do we need?

Signal: antiprotons
Background: electrons

	Number(*)	Charge	Cal. shower	Performance
antiprotons	$O(10^3)$	-	h	80 MeV - 190 GeV
protons	$O(10^7)$	+	h	80 MeV - 1.2 TeV
positrons	$O(10^4)$	+	em	50 MeV - 300 GeV
electrons	$O(10^5)$	-	em	50 MeV - 500 GeV

- Different calorimeter shower profile.
- RF $\gg$ 100 is easily achieved.

*: Exact numbers are energy dependent

Spectrometer resolution

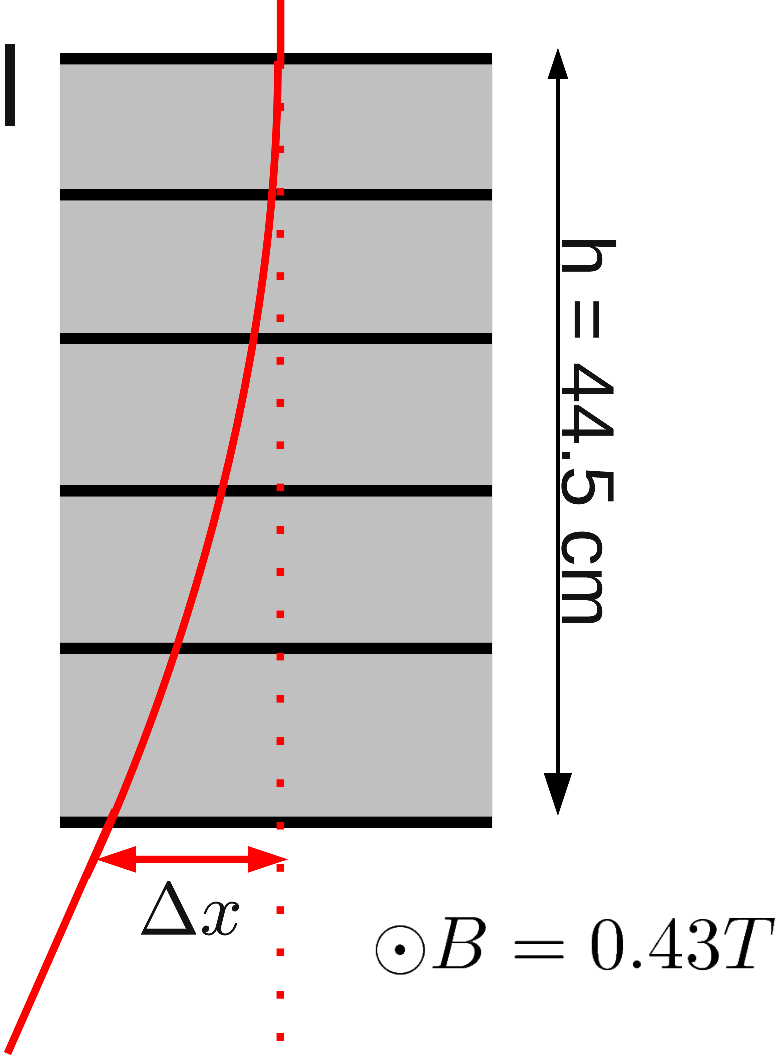
Spectrometer – Toy Model

$$\vec{F} = q\,(\vec{v} \times \vec{B})$$

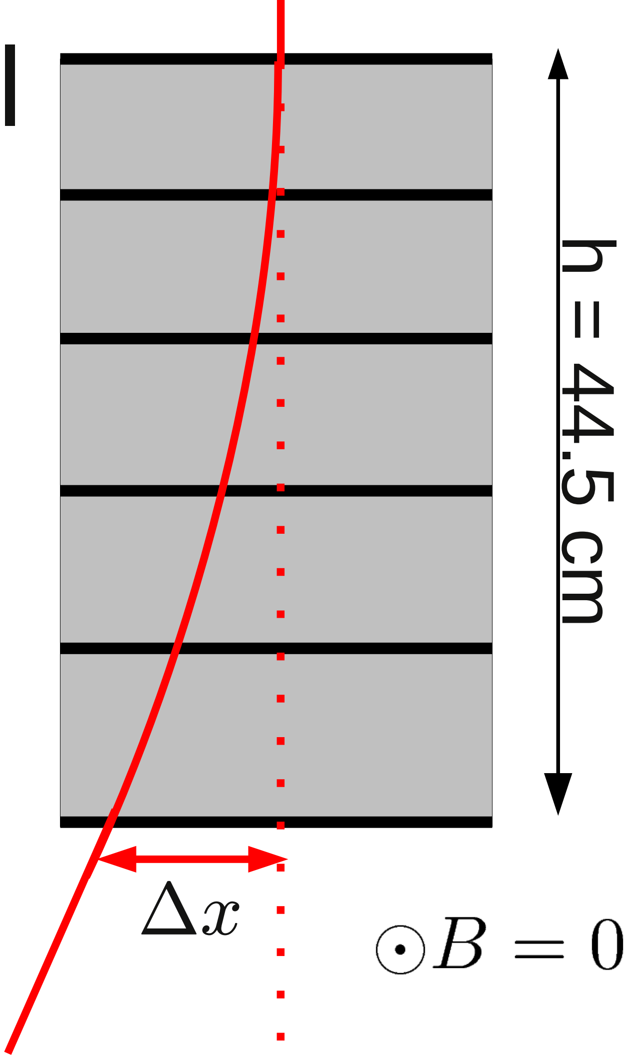
If B uniform:

$$\text{Rigidity} = \frac{p}{q} = \underset{\text{measured}}{\underset{\downarrow}{\rho}}\ \underset{\text{known}}{\underset{\searrow}{B}}$$

$p =$ relativisitc momentum
$q =$ charge
$\rho =$ Larmor radius



Spectrometer – Toy Model



$$\odot B = 0.43T$$

$$\vec{F} = q\,(\vec{v} \times \vec{B})$$

If B uniform:

$$\text{Rigidity} = \frac{p}{q} = \underset{\substack{\downarrow \\ \text{measured}}}{\rho}\ \underset{\substack{\searrow \\ \text{known}}}{B}$$

$$\rho(m) = 3.3m \times \frac{Rigidity\ (GV)/GV}{B(T)/T} = 7.7m \times Rigidity\ (GV)/GV$$

For vertically incoming particles:

$$\Delta x = \rho - \sqrt{\rho^2 - h^2} \approx \frac{h^2}{14.4}\left(\frac{1}{R}\right) = \begin{cases} 1.3\,cm & \text{at 1 GV} \\ 1.3\,mm & \text{at 10 GV} \\ 13\,\mu m & \text{at 1000 GV} \end{cases}$$

If h=44.5cm (**lever arm** = 6) !

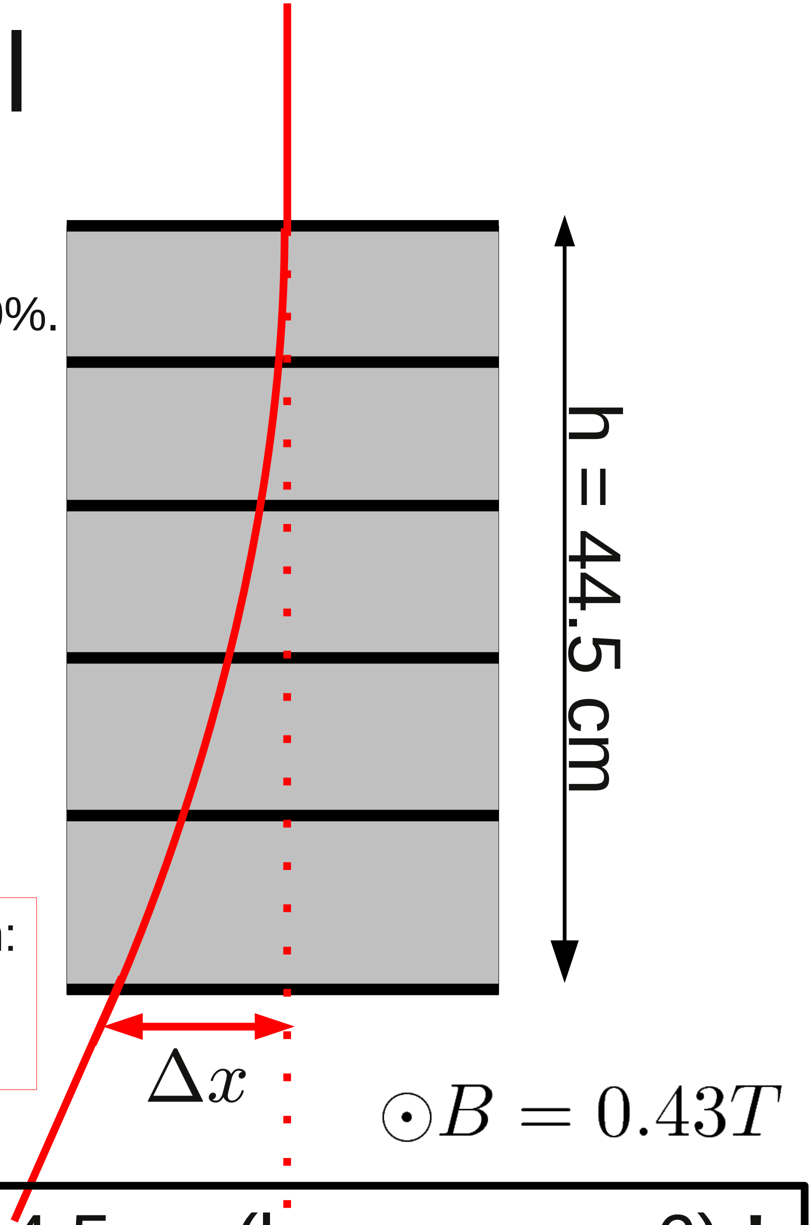
Spectrometer – Toy Model

For a given resolution we define the
Maximum Detectable Rigidity
as the rigidity where the deflection measurement error is 100%.

$$\frac{\sigma_\eta}{\eta_{\mathrm{MDR}}} = 1 \iff MDR = \frac{1}{\eta_{\mathrm{MDR}}}$$

or

$$MDR = \frac{1}{\sigma_\eta}$$



To have less than 100% error we have to select tracks with:

$$R < MDR$$

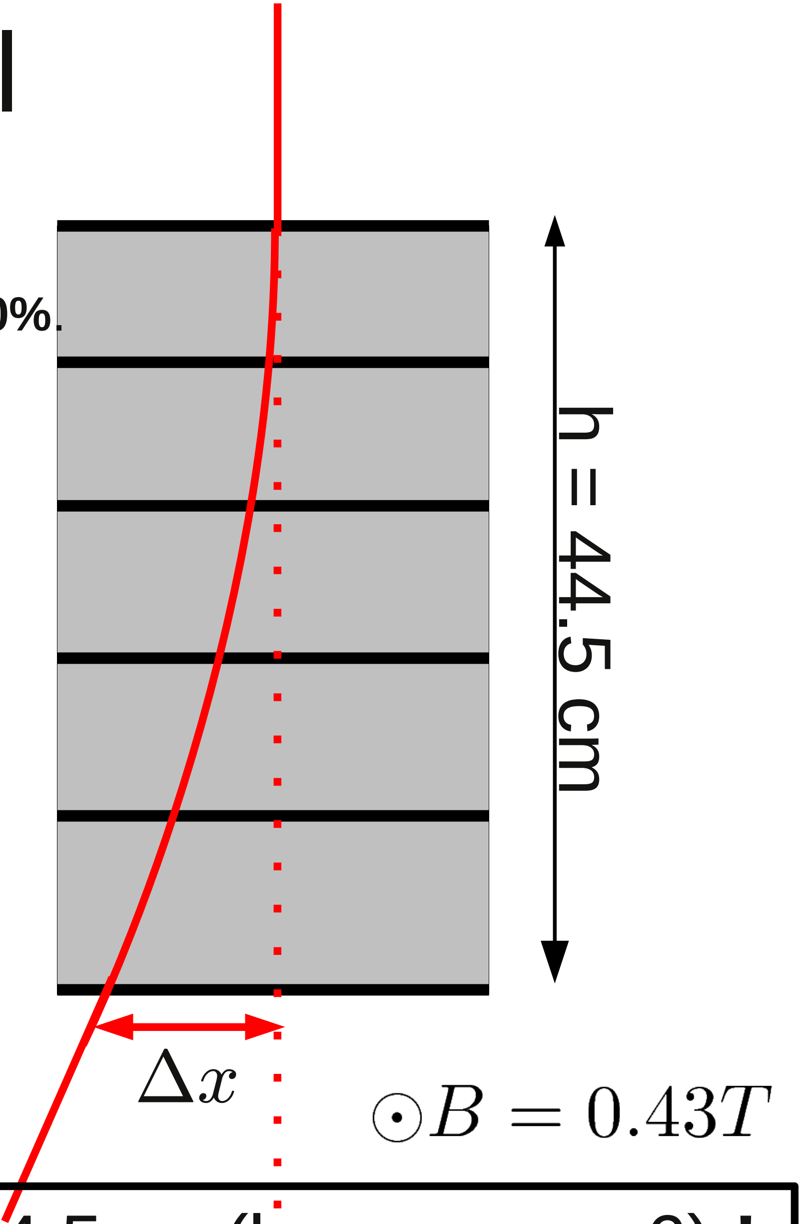
If h=44.5cm (lever arm = 6) !

$$\underset{\mathrm{m}}{\Delta x} = \rho - \sqrt{\rho^2 - h^2} \approx \frac{h^2}{14.4} \underset{\mathrm{GV}^{-1}}{\eta} = \begin{cases} 1.3\, cm & \text{at } 1 \text{ GV} \\ 1.3\, mm & \text{at } 10 \text{ GV} \\ 13\, \mu m & \text{at } 1000 \text{ GV} \end{cases}$$

Spectrometer – Toy Model

For a given resolution we define the
Maximum Detectable Rigidity
as **rigidity where the deflection measurement error is 100%**.

$$MDR = \frac{1}{\sigma_\eta}$$



In PAMELA data analysis, the MDR changes on event-by-event basis because its value depends on:

- number and distribution of fitted points along the trajectory (lever arm !)
- magnetic field intensity along the trajectory

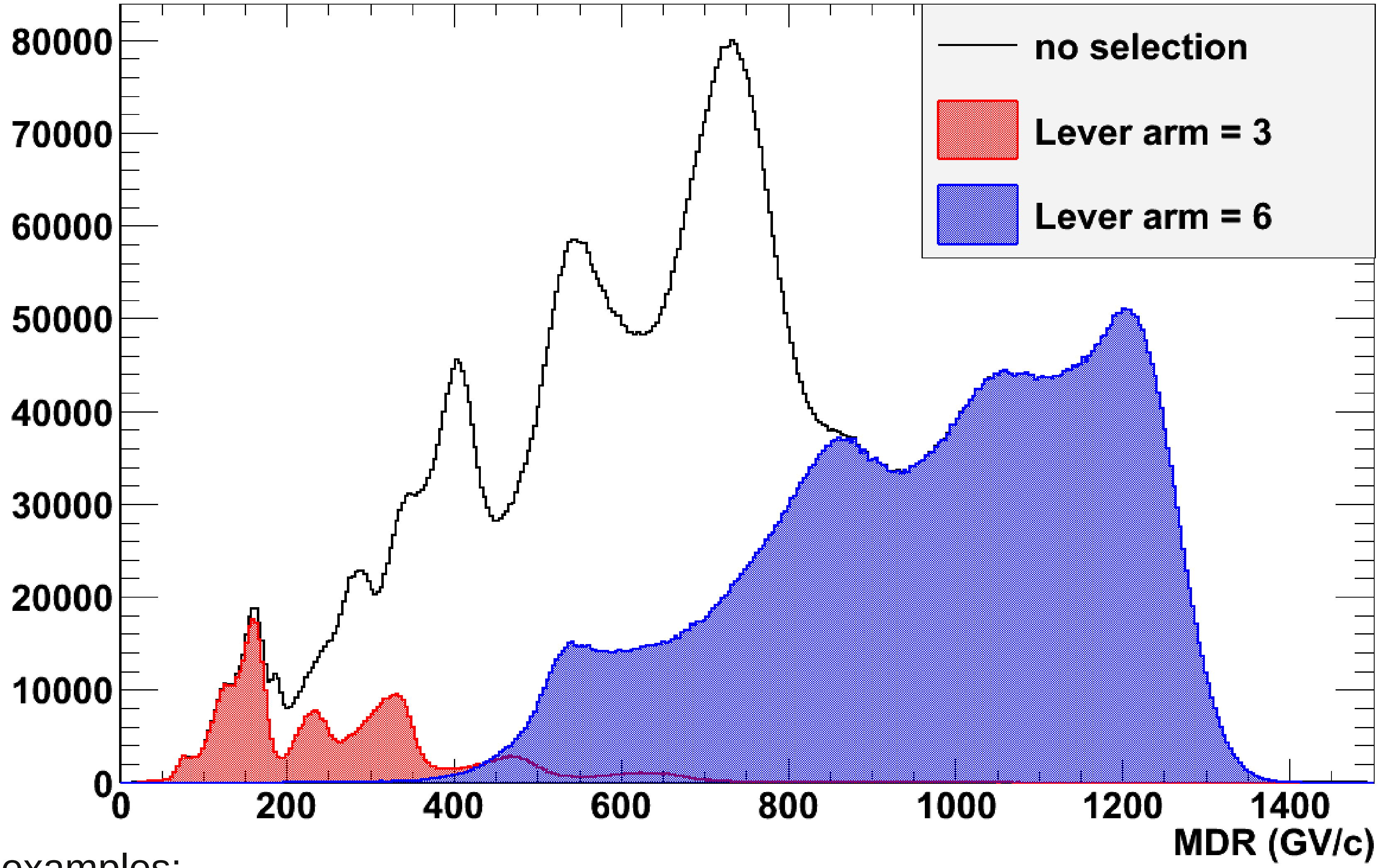
If h=44.5cm (lever arm = 6) !

$$\Delta x = \rho - \sqrt{\rho^2 - h^2} \approx \frac{h^2}{14.4}\,\eta = \begin{cases} 1.3\,cm & \text{at 1 GV} \\ 1.3\,mm & \text{at 10 GV} \\ 13\,\mu m & \text{at 1000 GV} \end{cases}$$

m

GV^{-1}

MDR and resolution



Some examples:

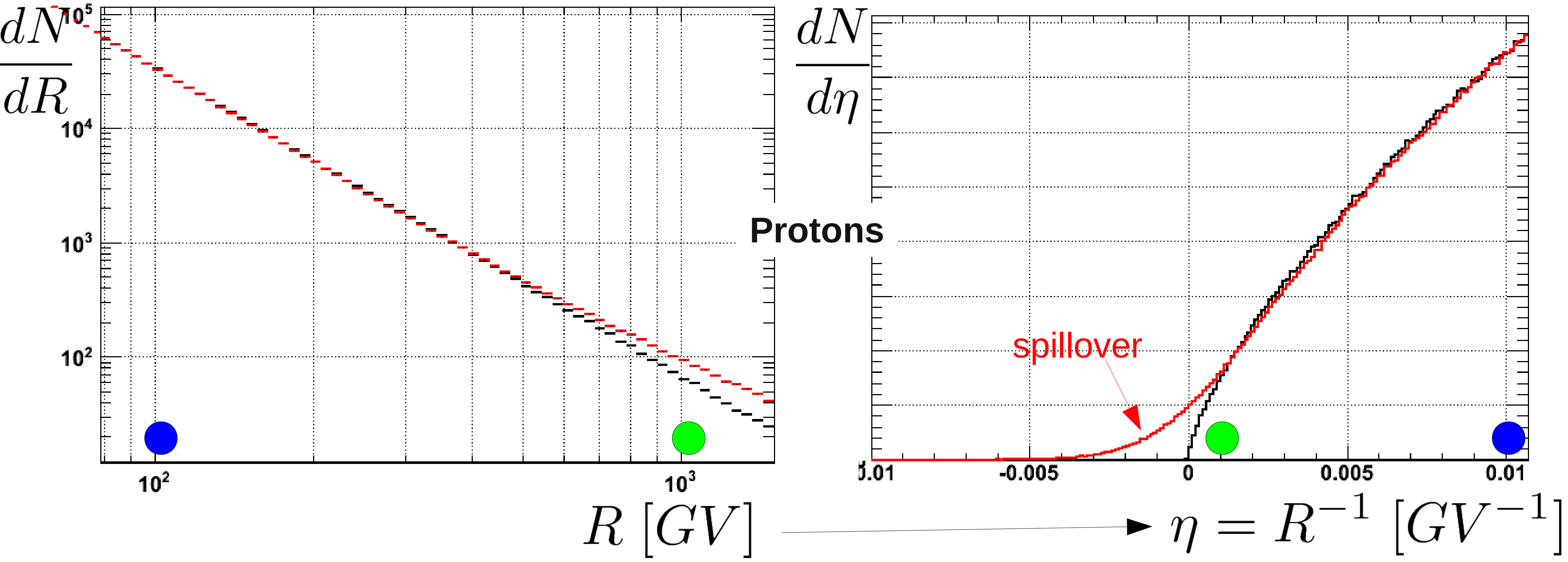
$$\begin{cases} Rig = 100GV \\ MDR = 500GV \end{cases} \Longrightarrow \frac{\delta\eta}{\eta} = 20\%$$

$$\begin{cases} Rig = 1000GV \\ MDR = 500GV \end{cases} \Longrightarrow \frac{\delta\eta}{\eta} = 200\%$$

$$\begin{cases} Rig = 100GV \\ MDR = 1000GV \end{cases} \Longrightarrow \frac{\delta\eta}{\eta} = 10\%$$

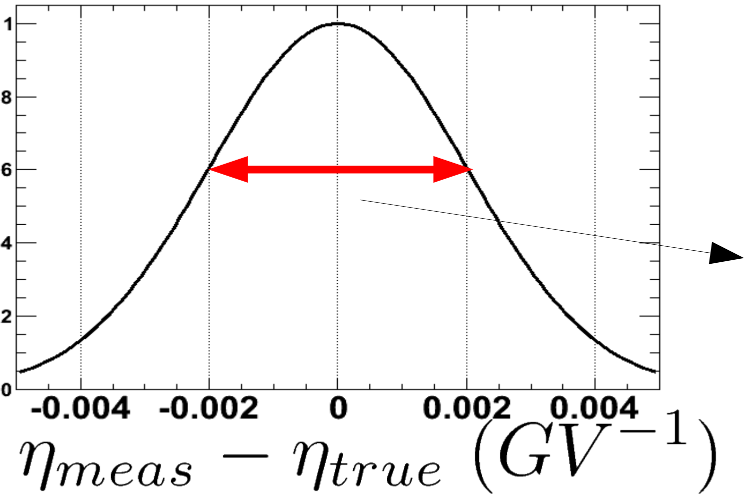
$$\begin{cases} Rig = 1000GV \\ MDR = 1300GV \end{cases} \Longrightarrow \frac{\delta\eta}{\eta} = 77\%$$

Spectrometer resolution – Toy Model



▬ Ideal tracker with infinite resolution: no spillover

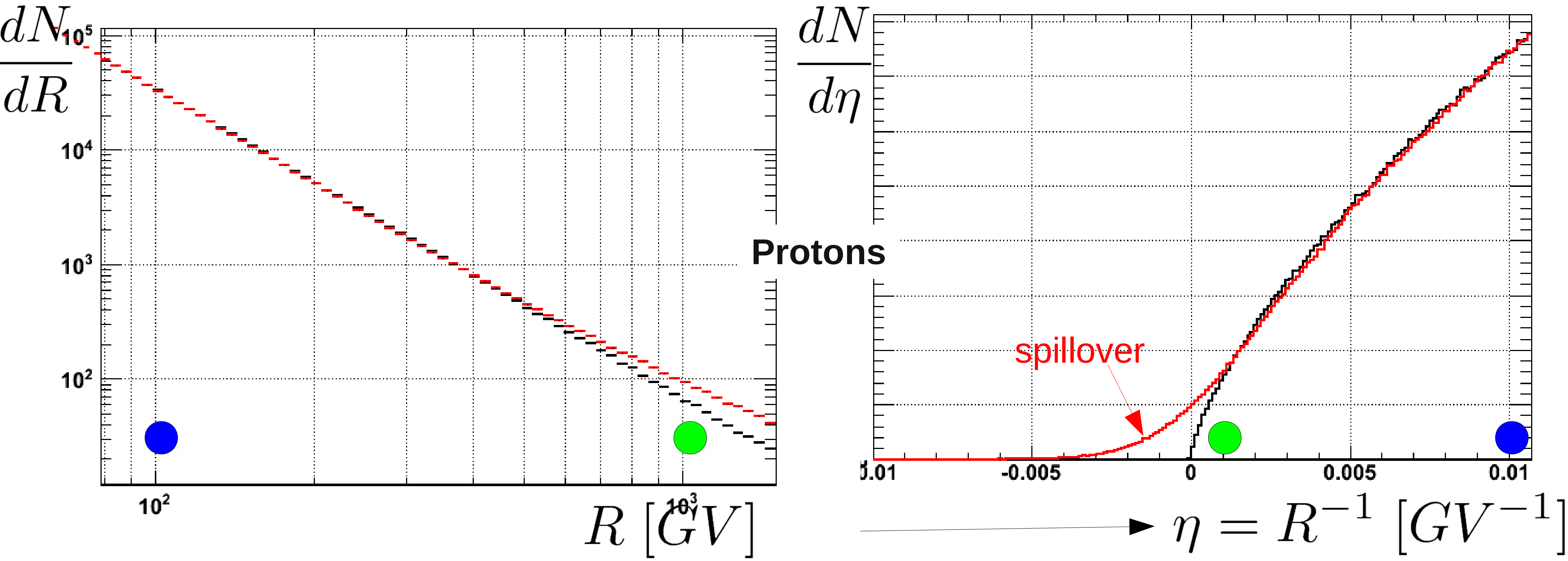
▬ Tracker with finite resolution: spillover at high energy



$$\boxed{\sigma_\eta = \frac{1}{MDR}}$$

$$[MDR] = GV$$

Spectrometer resolution – Toy Model



$$MDR = \frac{1}{\sigma_\eta}$$

The spectrometer resolution sets the high energy limit for:

- Protons (left)
- Antiprotons (because of the proton spillover) (right)

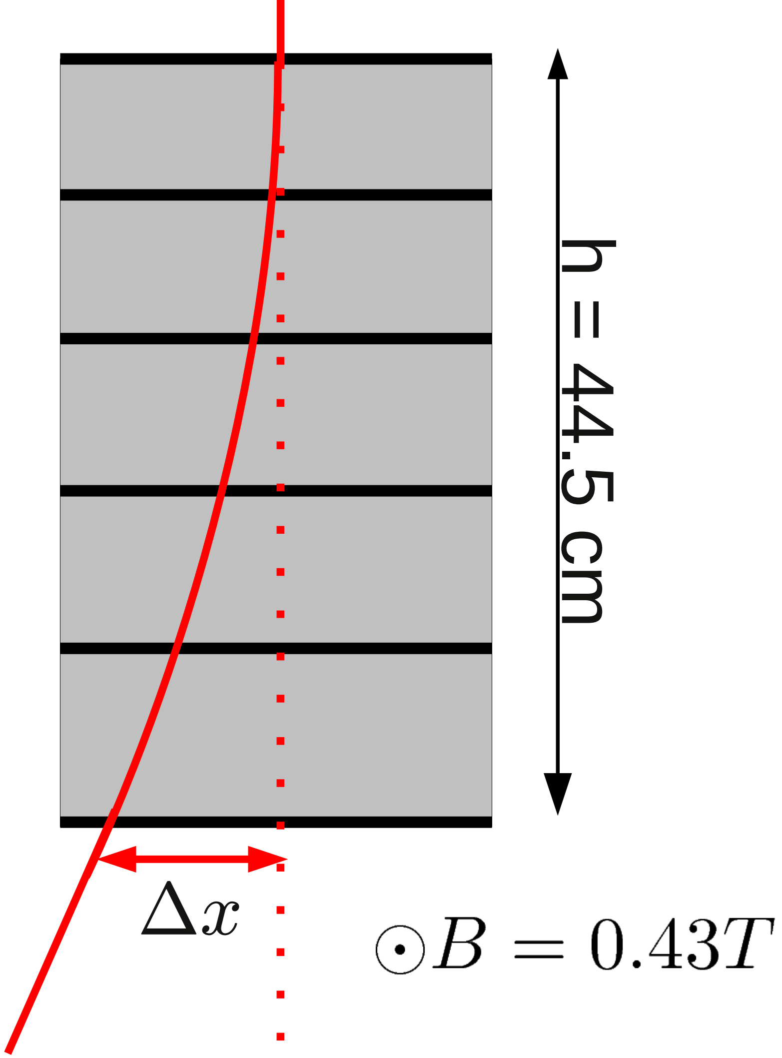
Tracker: chi2

$$\chi^2 = \frac{1}{N}\sum_i (x_i^{meas} - x_i^{fit})^2$$

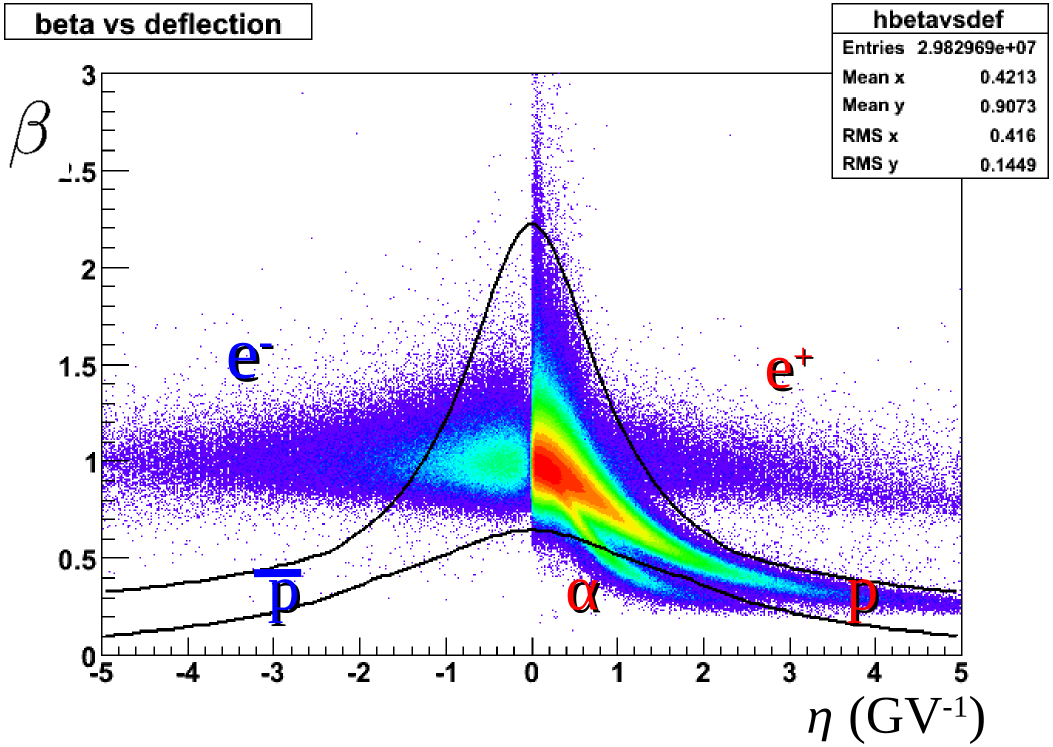
The χ² cut is needed to improve the fit quality of the selected sample.

χ² gets worse because of:

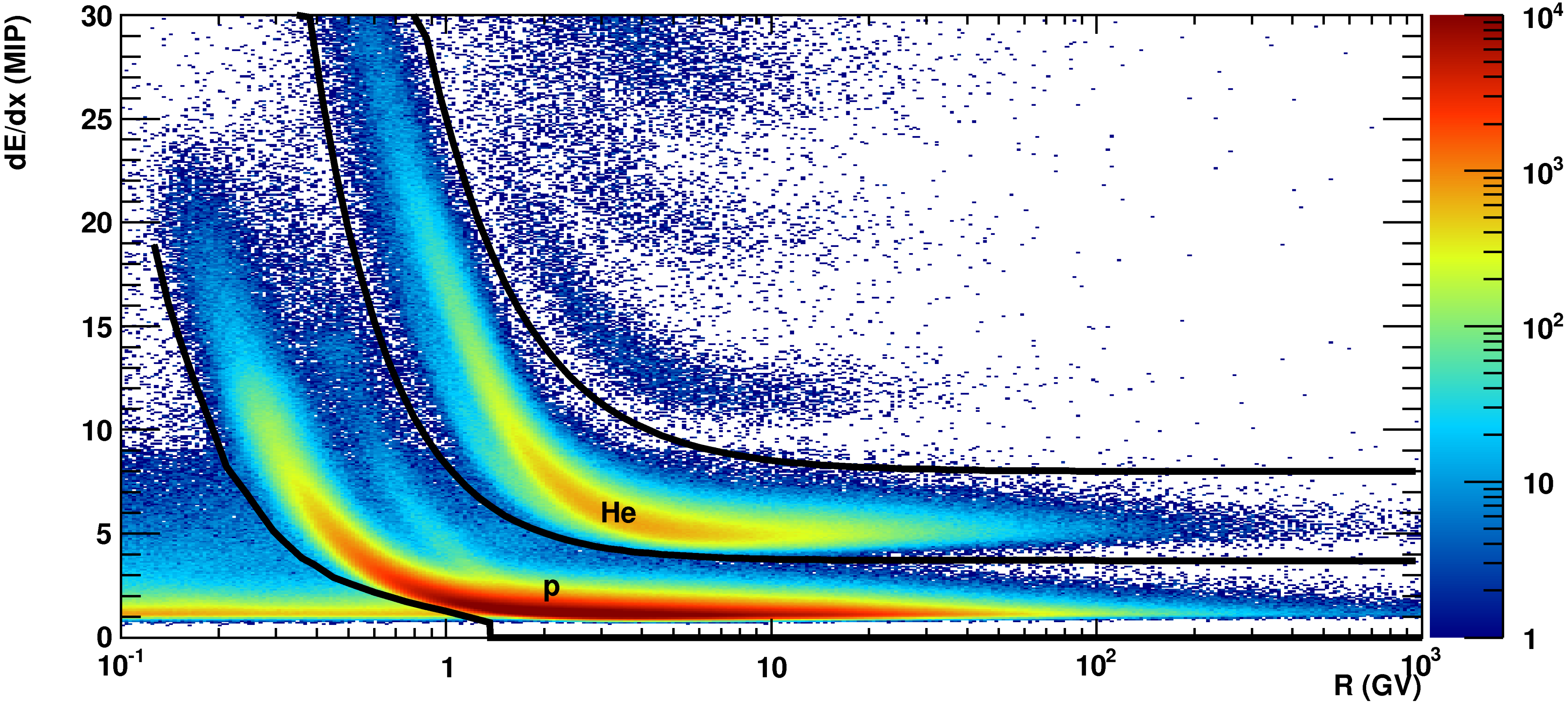
- Multiple scattering
- δ-ray emission inside the silicon
- faulty strips (high noise)

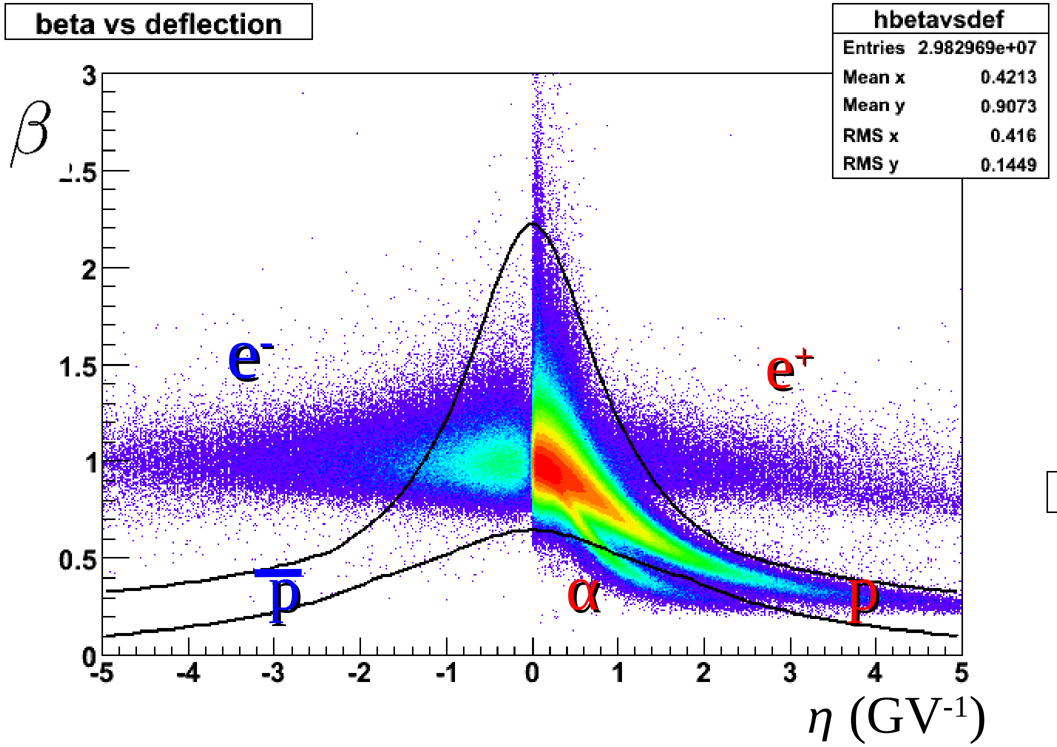


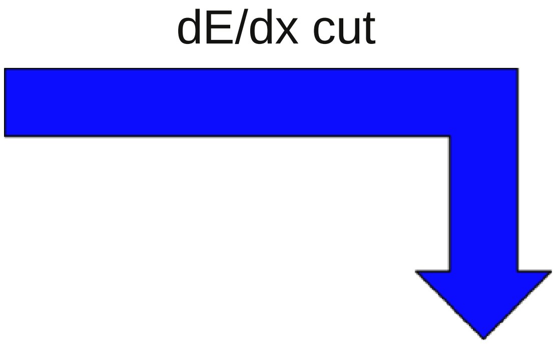
Antiproton to Proton Flux Ratio

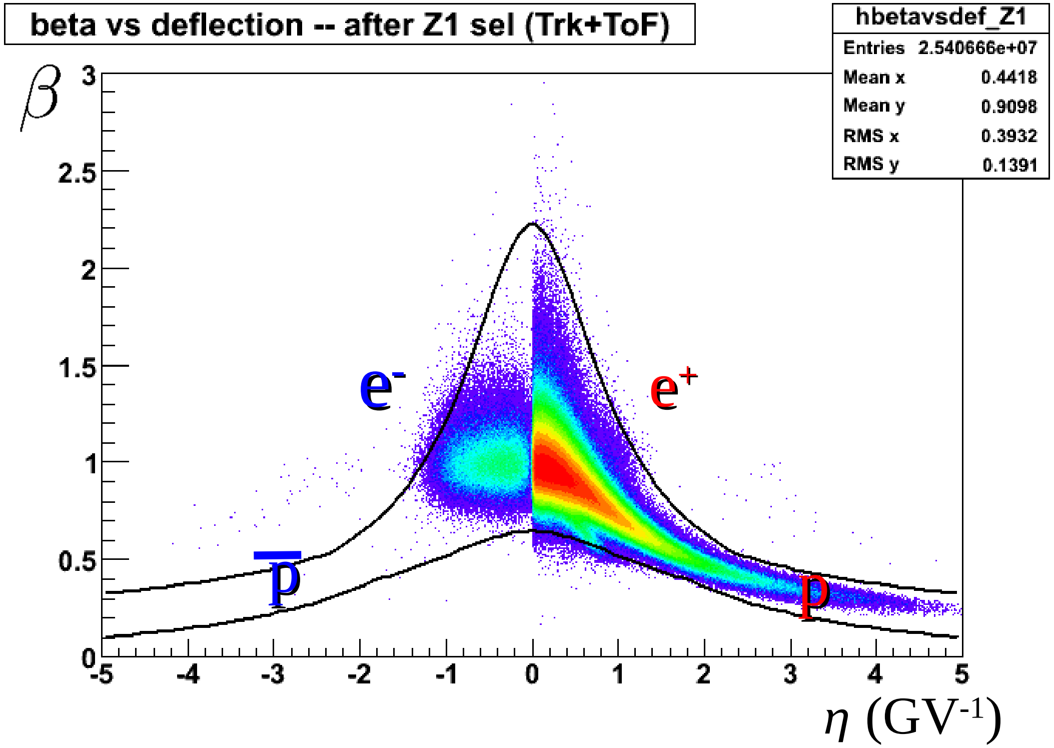

beta vs deflection
hbetavsdef
Entries 2.982969e+07
Mean x 0.4213
Mean y 0.9073
RMS x 0.416
RMS y 0.1449
β
η (GV-1)
e-
e+
p̄
α
p
0
0.5
1
1.5
2
2.5
3
-5
-4
-3
-2
-1
0
1
2
3
4
5

Tracker dE/dx selection

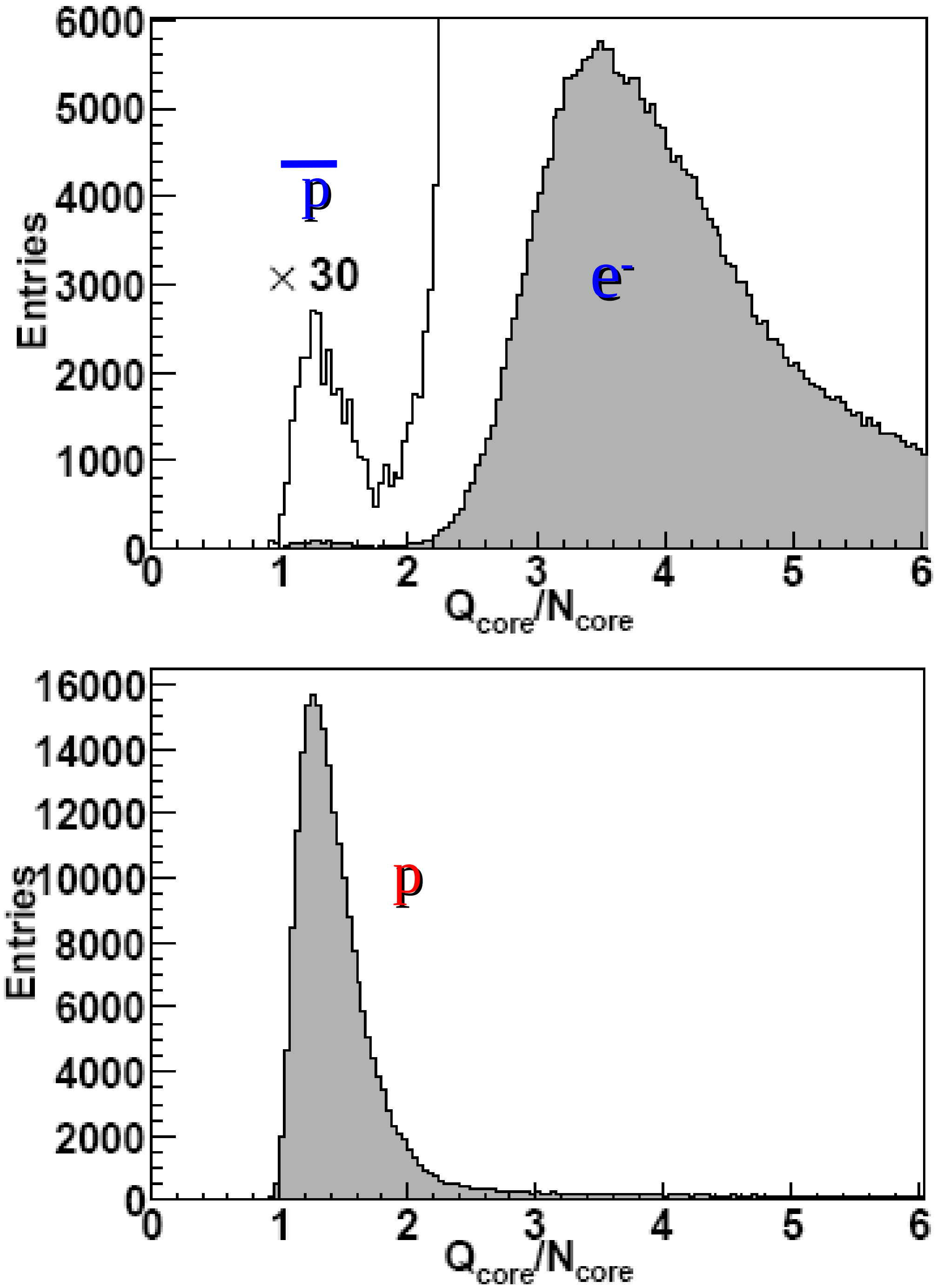


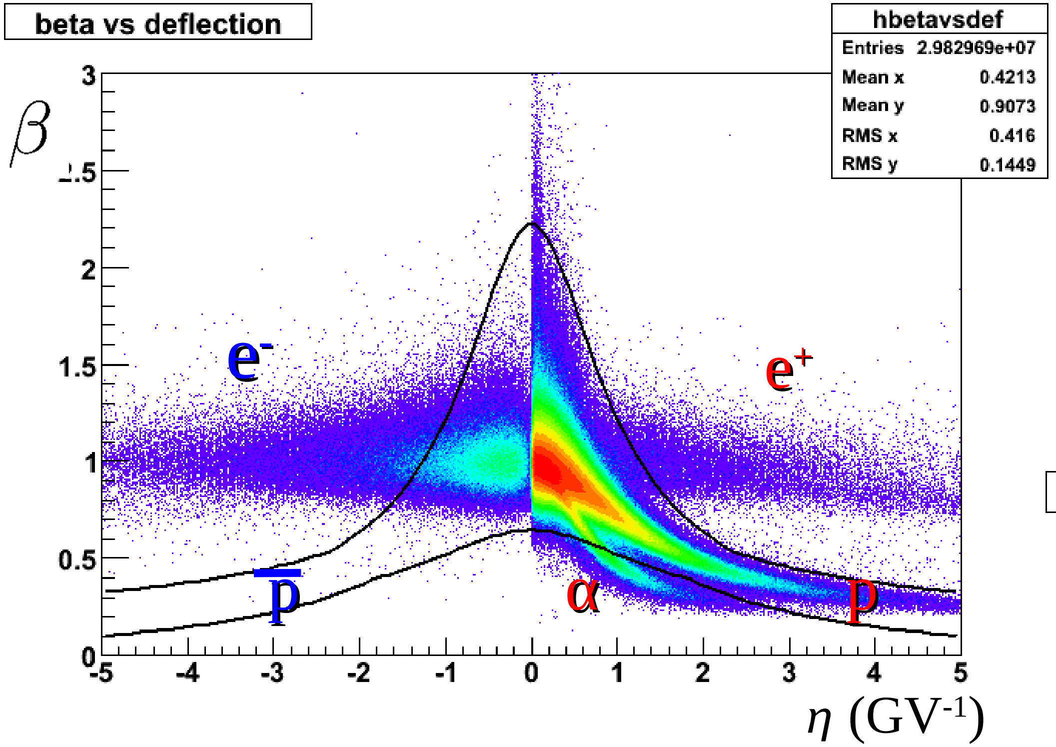

beta vs deflection
hbetavsdef
Entries 2.982969e+07
Mean x 0.4213
Mean y 0.9073
RMS x 0.416
RMS y 0.1449
β
3
2.5
2
1.5
1
0.5
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
η (GV-1)
e-
e+
p̄
α
p

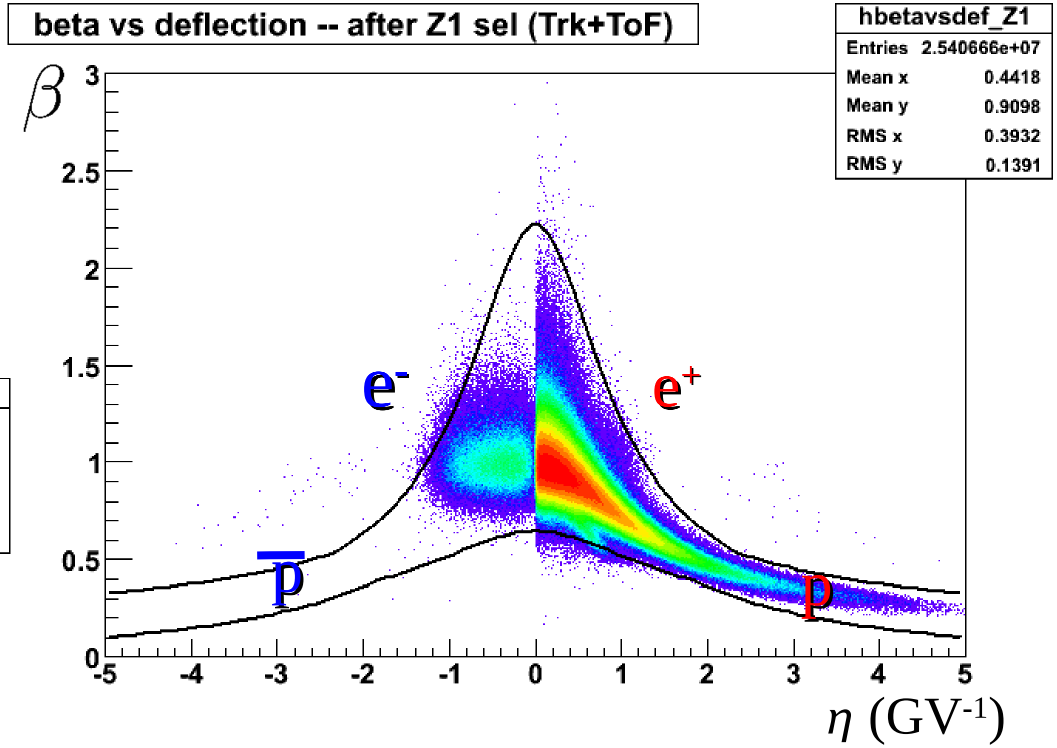

dE/dx cut


beta vs deflection -- after Z1 sel (Trk+ToF)
hbetavsdef_Z1
Entries 2.540666e+07
Mean x 0.4418
Mean y 0.9098
RMS x 0.3932
RMS y 0.1391
β
3
2.5
2
1.5
1
0.5
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
η (GV-1)
e-
e+
p̄
p

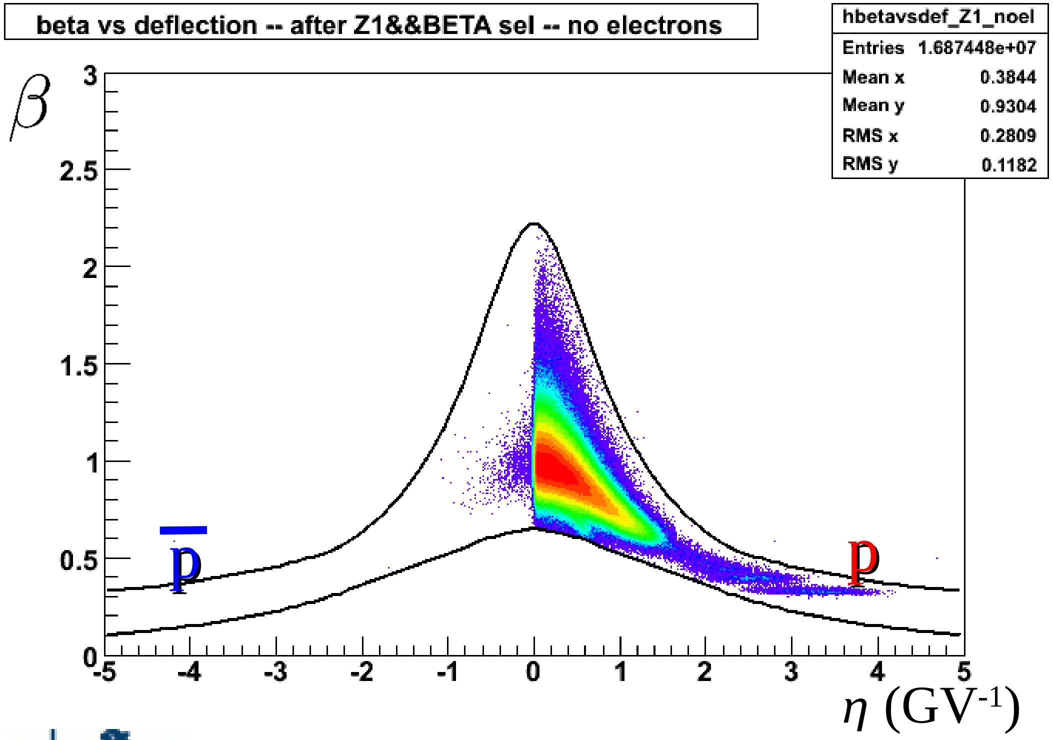
Calorimeter selection





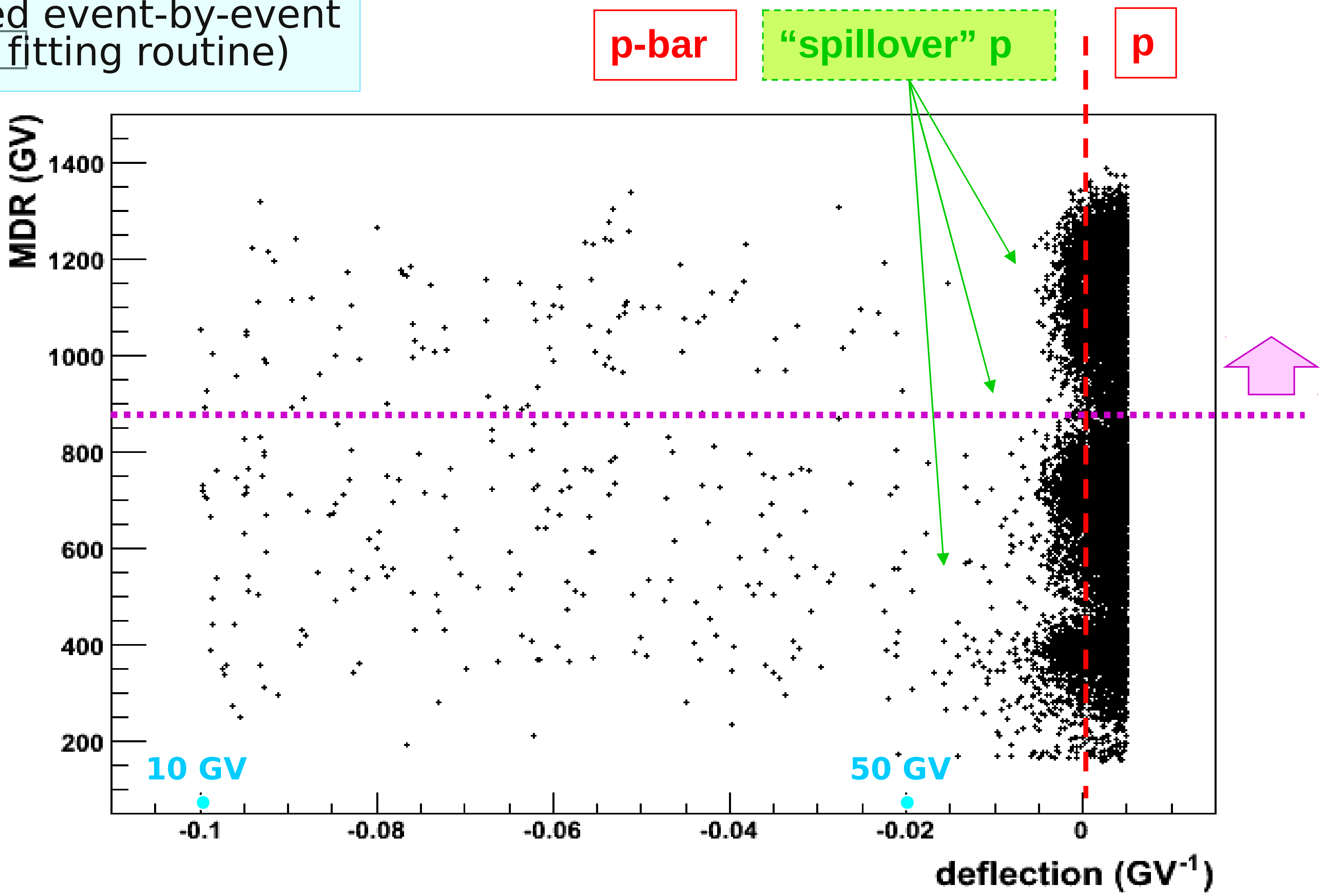


Calorimeter selection



Proton-spillover background

MDR = $1/\sigma_\eta$
(evaluated event-by-event
by the fitting routine)



p-bar

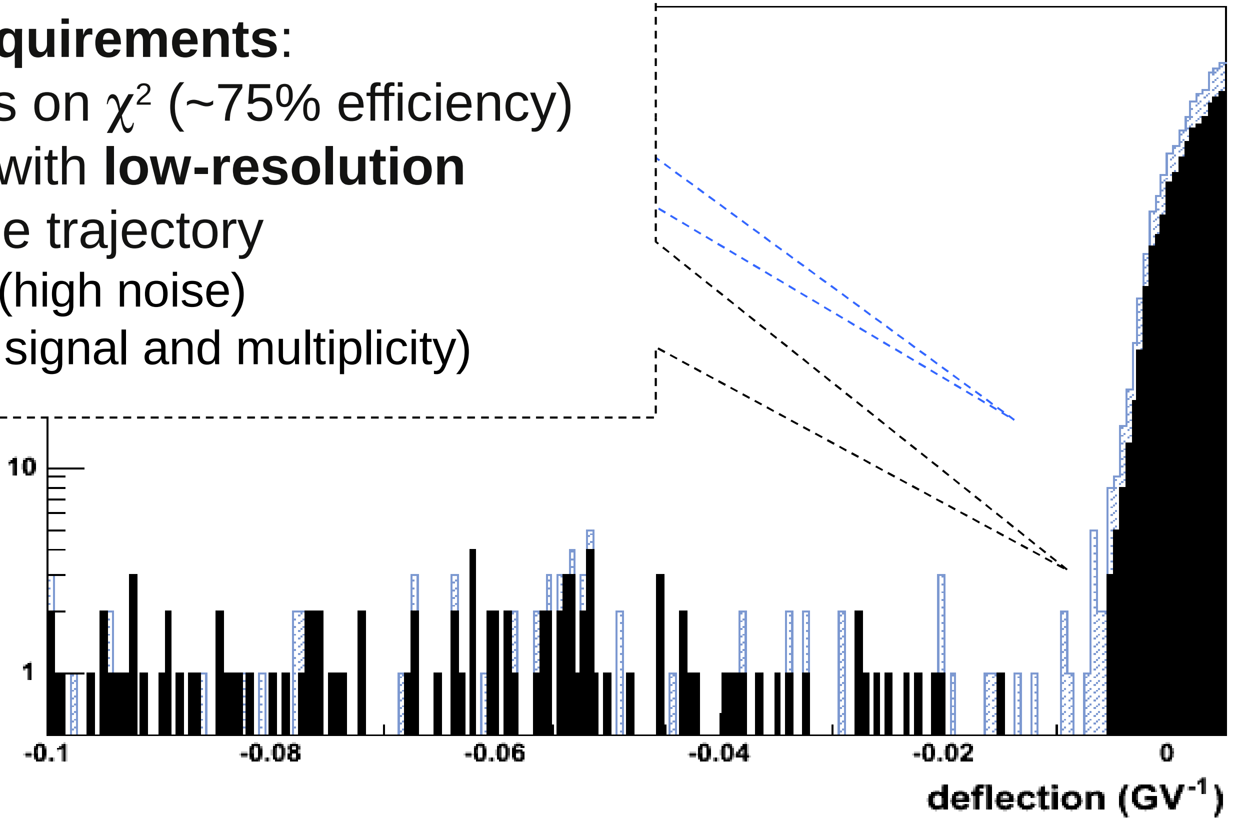
“spillover” p

p

Proton-spillover background

Minimal track requirements

MDR > 850 GV

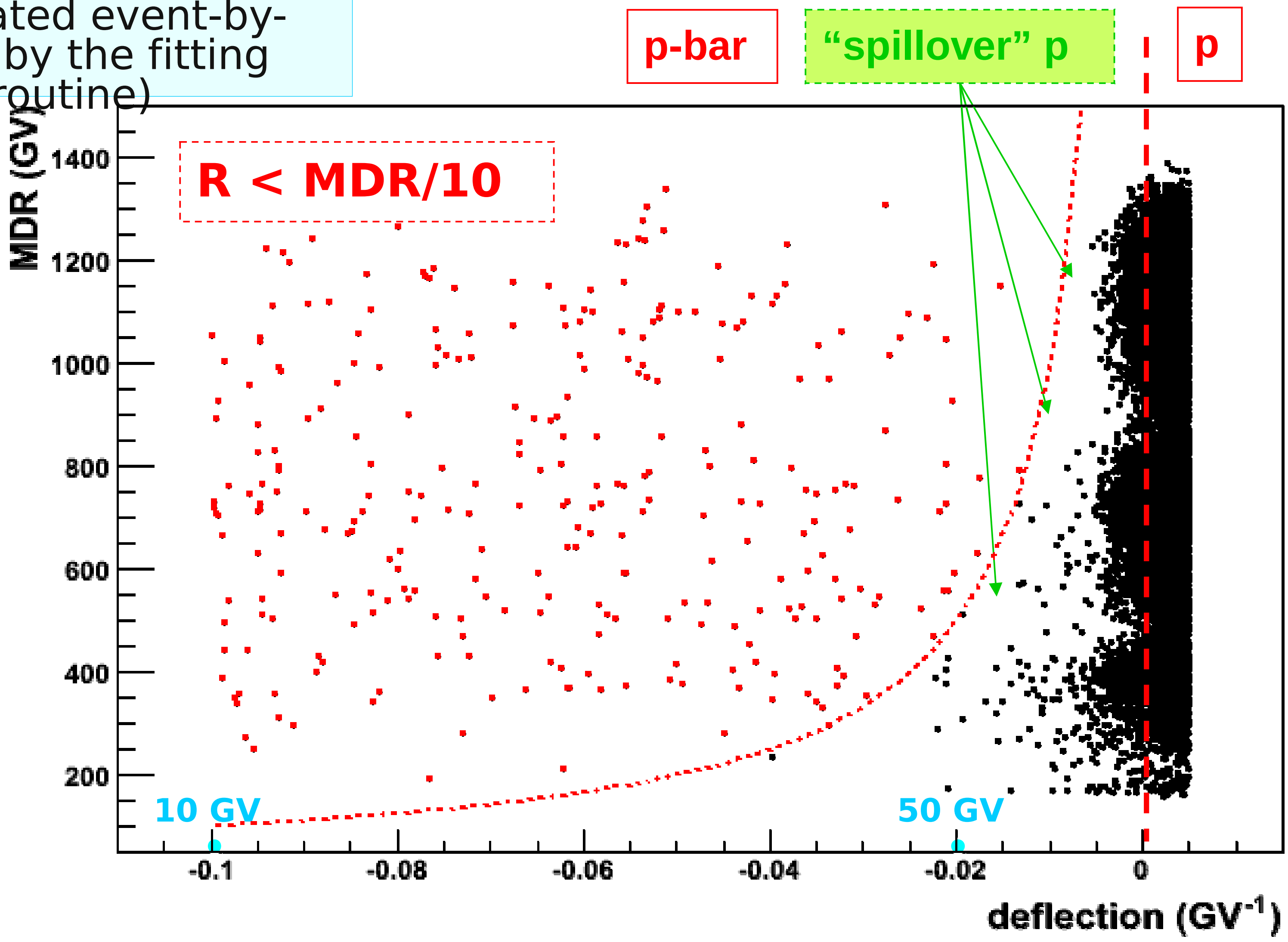


Strong track requirements:

- strict constraints on χ^2 (~75% efficiency)
- rejected tracks with **low-resolution** clusters along the trajectory
  - faulty strips (high noise)
  - δ-rays (high signal and multiplicity)

Proton-spillover background

$MDR = 1/\sigma_\eta$
(evaluated event-by-event by the fitting routine)

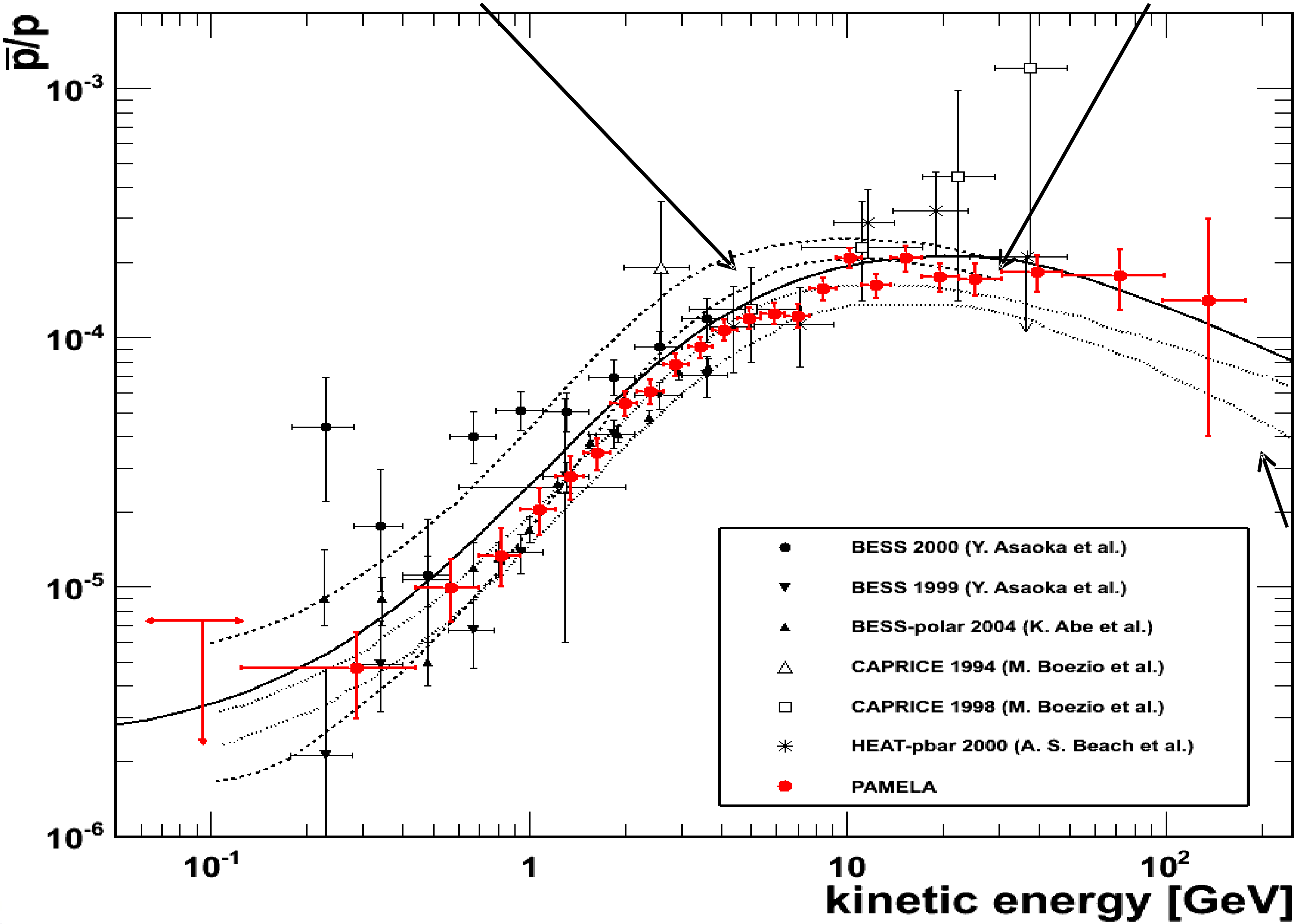

p-bar

“spillover” p

p

R < MDR/10

Antiproton to Proton Flux Ratio



Simon et al. (ApJ 499 (1998) 250)

Ptuskin et al. (ApJ 642 (2006) 902)

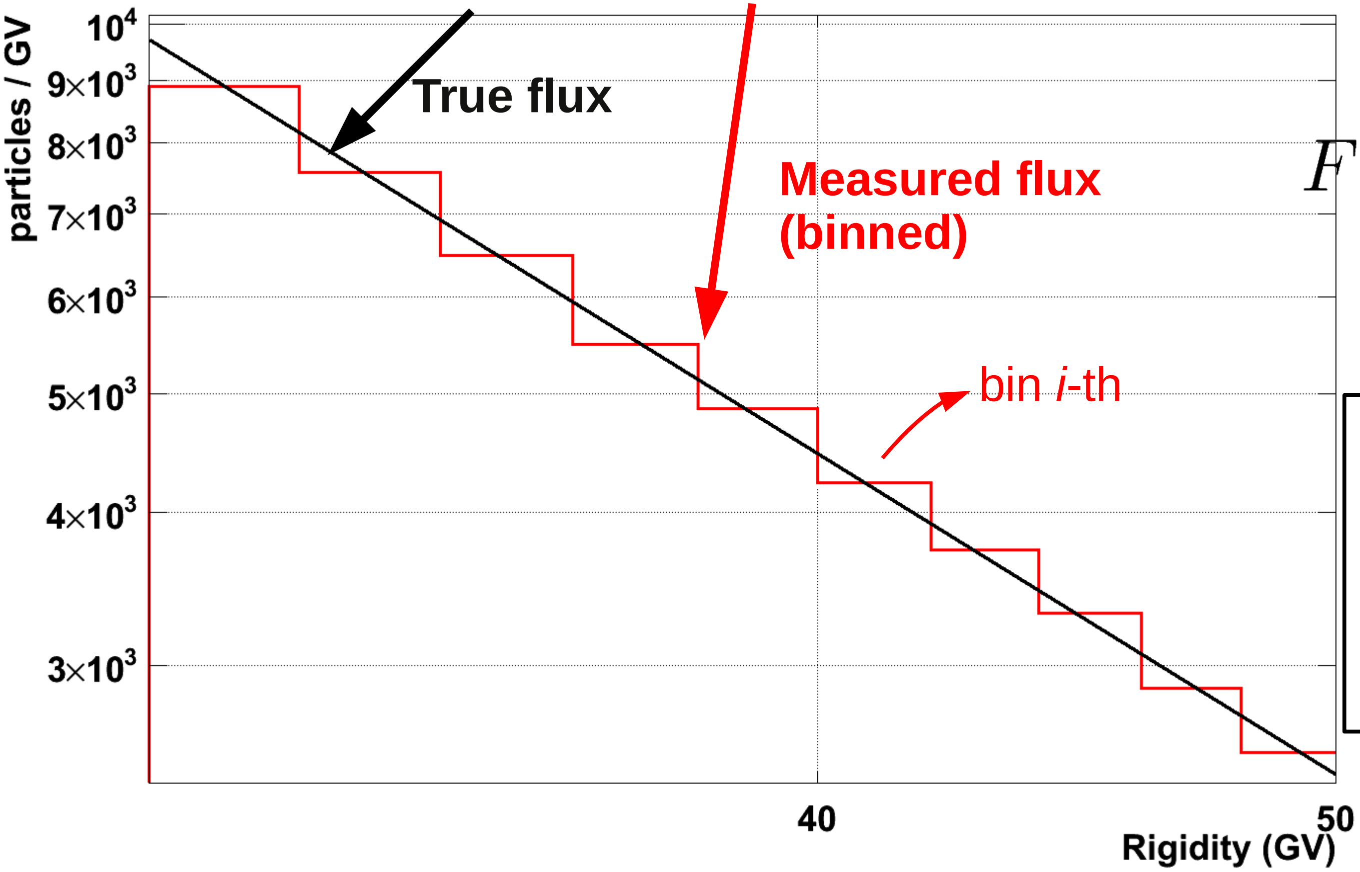
Donato et al.
(PRL 102
(2009) 071301)

Adriani et al., PRL
arXiv:1007.0821

Fluxes

Differential flux

$$\frac{d\,N}{d\,R} \longrightarrow \frac{\Delta N}{\Delta R}$$



We define the flux in the rigidity interval *i:*

$$R_{min} < R \le R_{max}$$

$$Flux_i = \frac{N_i}{G_i\ \Delta R_i\ LT_i\ \varepsilon_i}$$

- ➔ The bin width ΔR is arbitrary

A reasonable choice depends on:

- ➔ Resolution
- ➔ Statistic

Large binning kills spectral features.

Note:
in general, everything is rigidity dependent

	unit
N_i = number of selected particles	$-$
G_i = geometrical factor	$cm^2 sr$
$\Delta_i = R_{max} - R_{min}$	GV
LT_i = live time	s
ε_i = selection efficiency	$-$

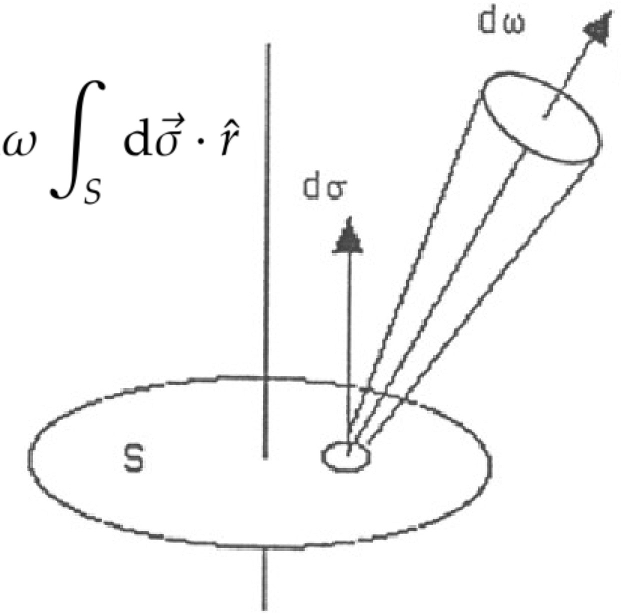
Fluxes

Geometrical factor

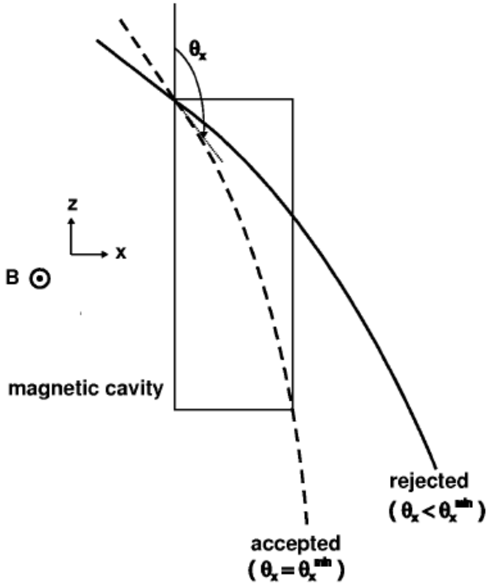
$$Flux_i = \frac{N_i}{\boxed{G_i}\ \Delta_i\ LT_i\ \varepsilon_i}$$

Geometrical factor (cm^2 sr)

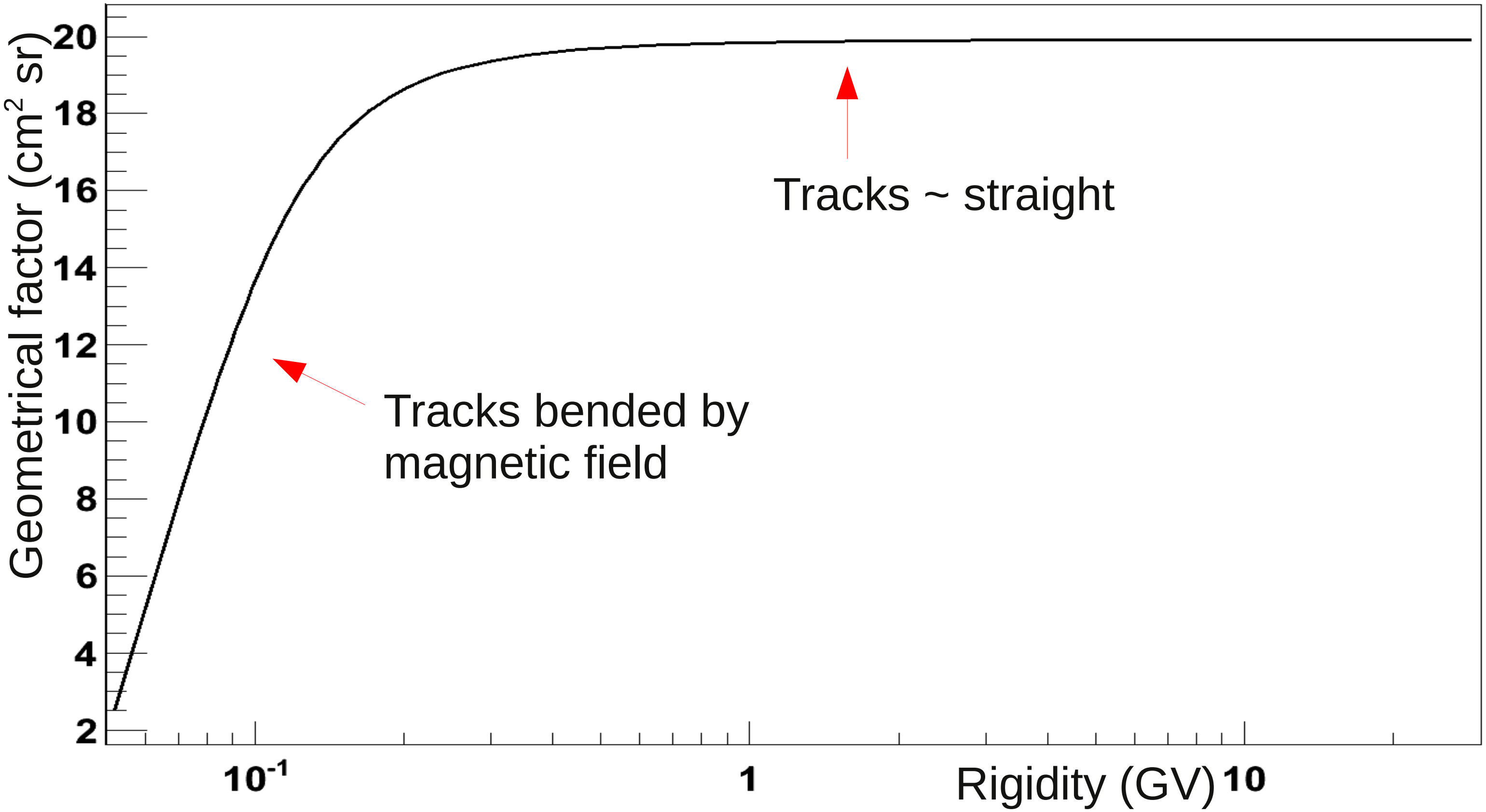
$$G = \int_\Omega d\omega \int_S d\vec{\sigma} \cdot \hat{r}$$



J. D. Sullivan. NIM, 95, 1971.



→ **It is defined for an isotropic (galactic) flux**



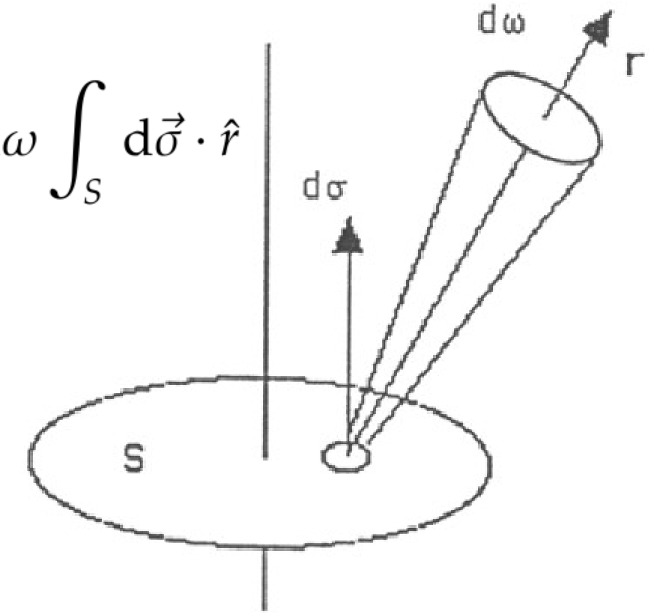
Tracks ~ straight

Tracks bended by magnetic field

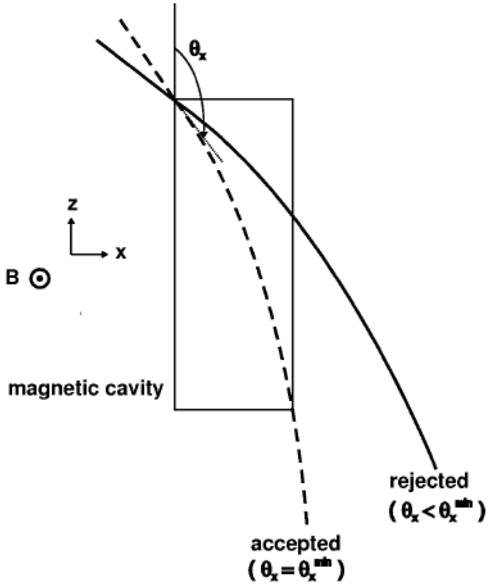
$$Flux_i = \frac{N_i}{\boxed{G_i}\ \Delta_i\ LT_i\ \varepsilon_i}$$

Geometrical factor (cm² sr)

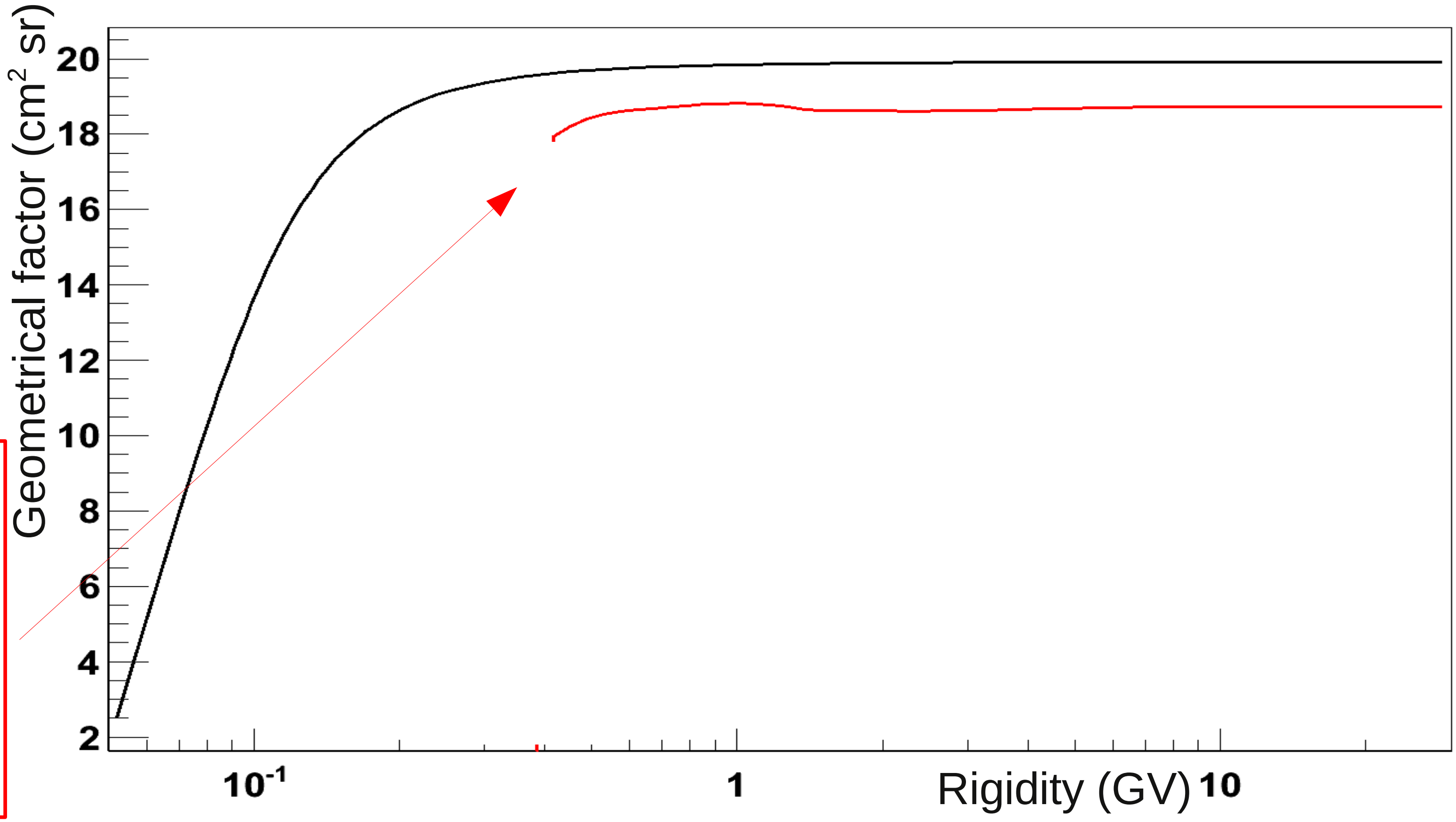
$$G = \int_\Omega d\omega \int_S d\vec{\sigma} \cdot \hat{r}$$



J. D. Sullivan. NIM, 95, 1971.



The **physical** geometrical factor includes corrections for loss/gain for elastic scattering above the tracker, loss for inelastic scattering.



Fluxes

Live Time

$$Flux_i = \frac{N_i}{G_i\ \Delta_i\ \boxed{LT_i}\ \varepsilon_i}$$

Live Time (s)

- We call ***dead time*** the time required by the apparatus to process and register each event.
- Each detector can keep the apparatus in “busy” state.
- In PAMELA the dead time is event-dependent.

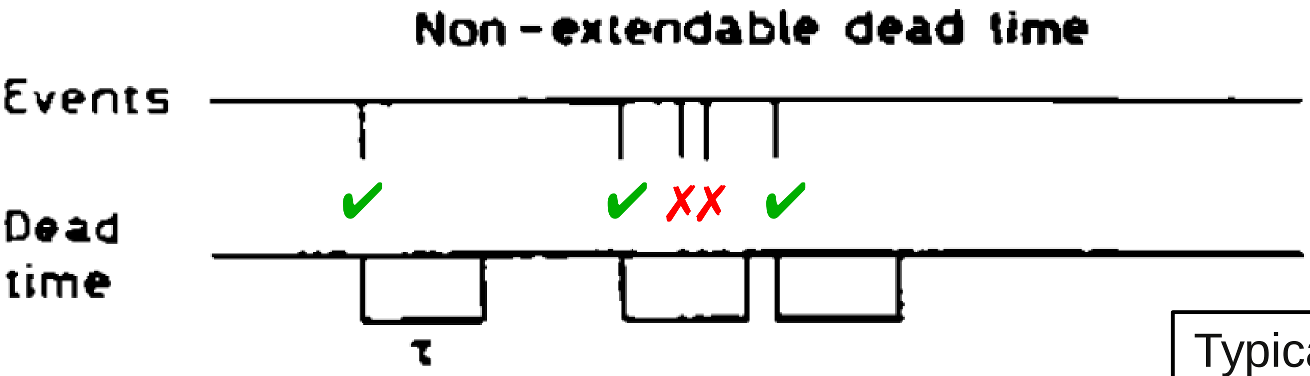


Fig. from
W.R. Leo, Springer (1987)

Typical rate: 25 Hz
Live time: 73% clock time

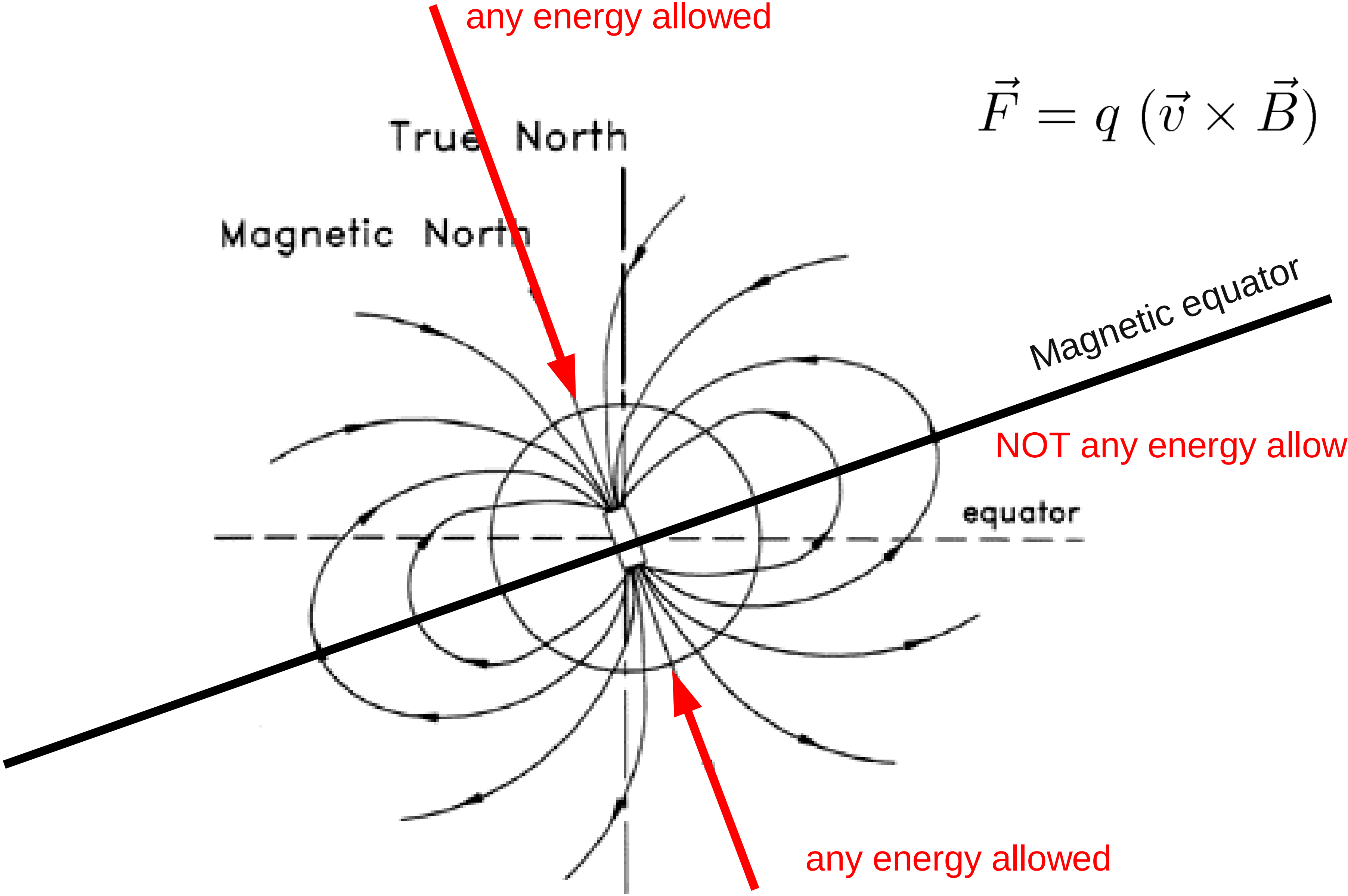
clock time = dead time + live time

In PAMELA, the trigger board counts the dead and the live time for each event.

Because there is a **geomagnetic field**, the live time LT for **galactic particles** is not rigidity independent.

CR trajectories in the Earth's magnetic field

$$\vec{F} = q\,(\vec{v} \times \vec{B})$$



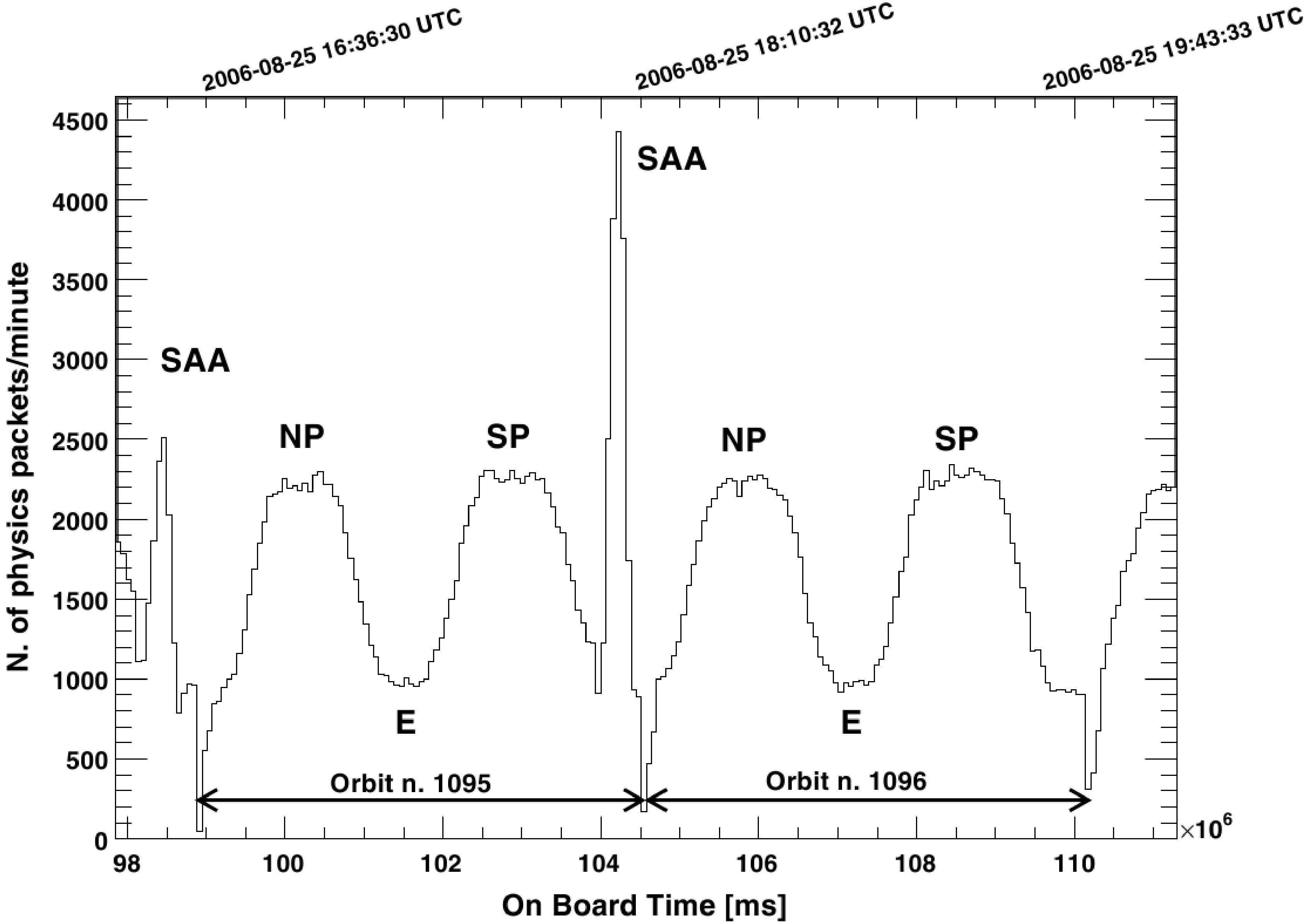
any energy allowed

Magnetic equator

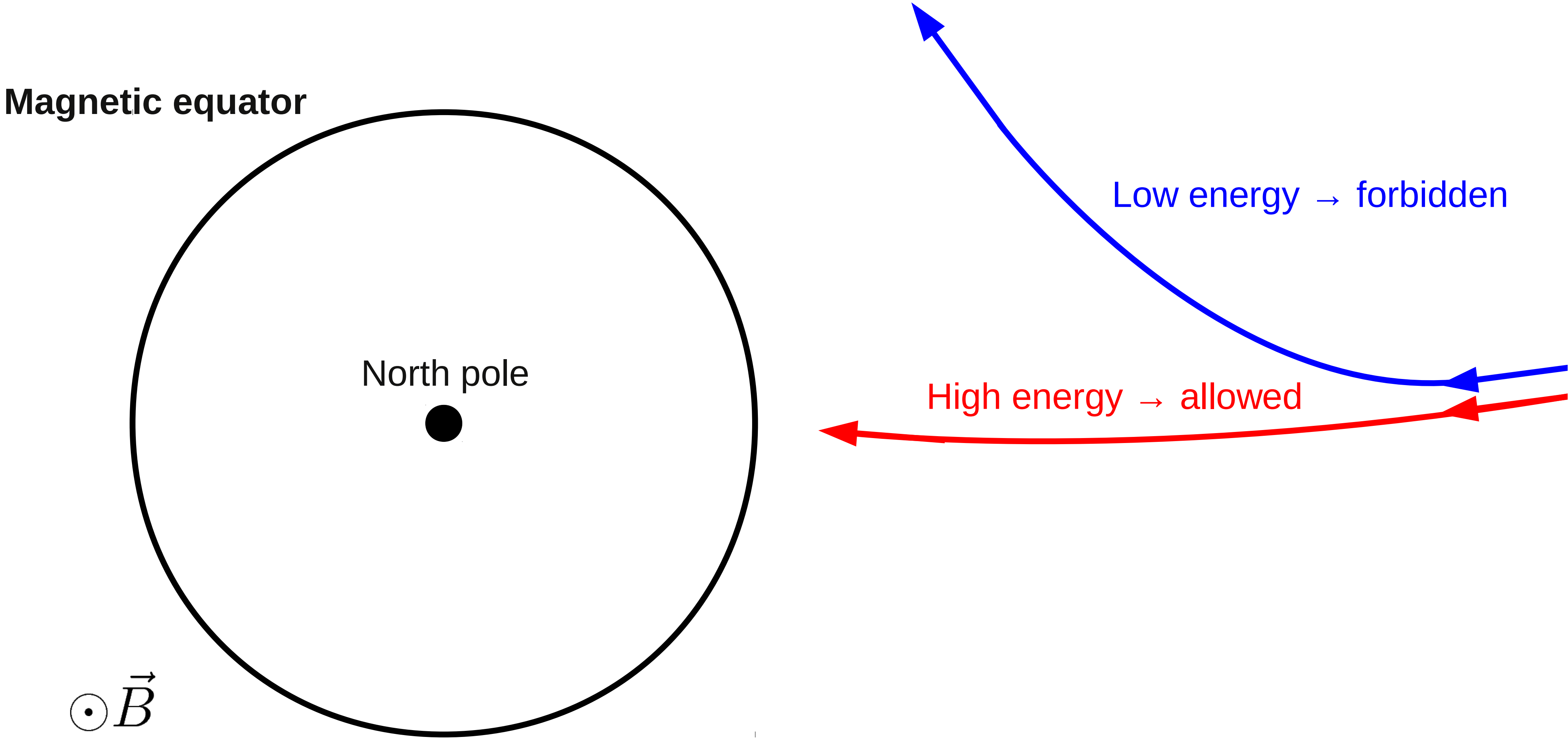
NOT any energy allowed

any energy allowed

The orbit


2006-08-25 16:36:30 UTC
2006-08-25 18:10:32 UTC
2006-08-25 19:43:33 UTC
SAA
SAA
NP
SP
NP
SP
E
E
Orbit n. 1095
Orbit n. 1096
N. of physics packets/minute
0
500
1000
1500
2000
2500
3000
3500
4000
4500
98
100
102
104
106
108
110
×10^6
On Board Time [ms]

CR trajectories in the Earth's magnetic field



Magnetic equator

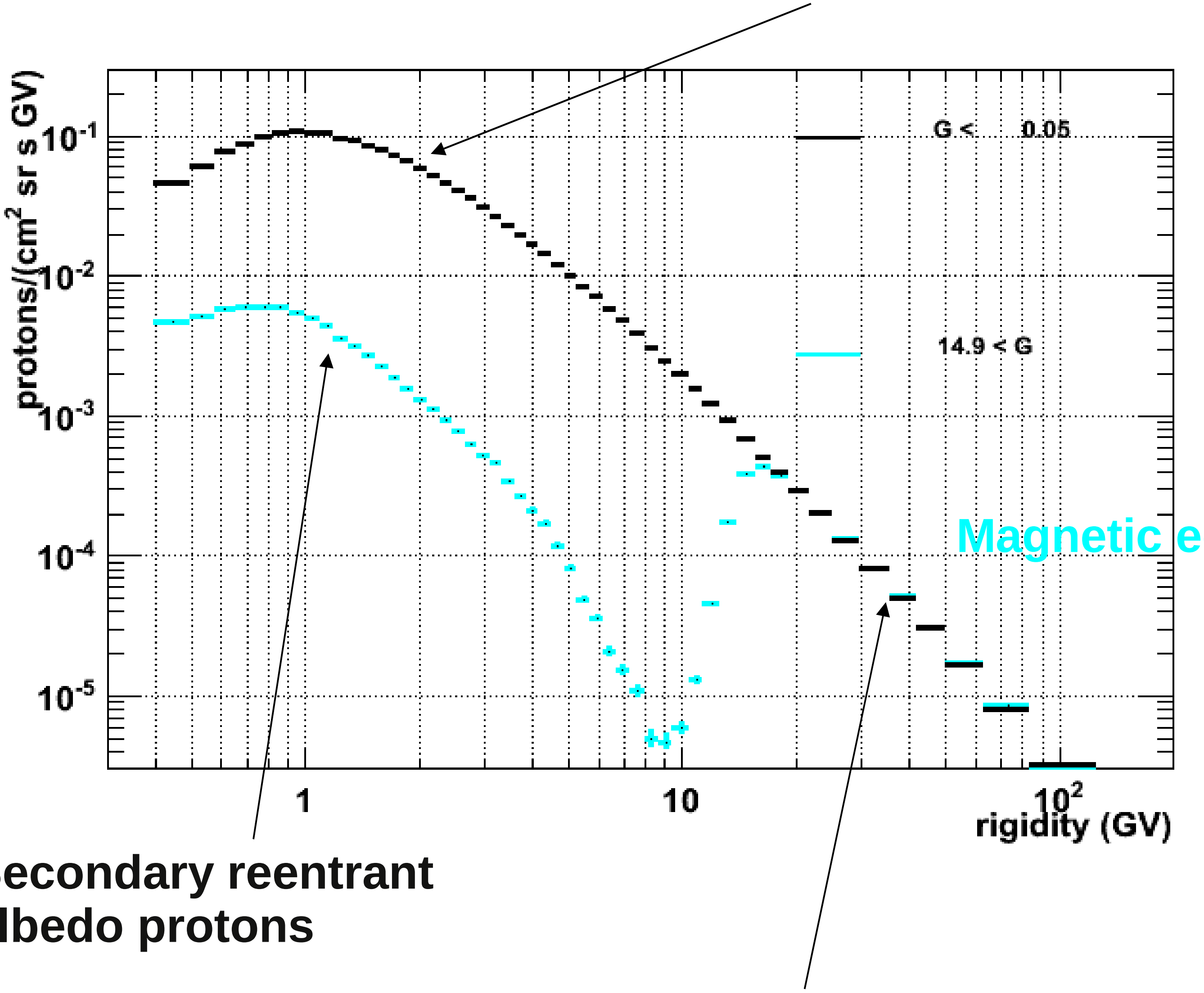
Low energy → forbidden

North pole

High energy → allowed

$\odot \vec{B}$

Proton Flux at various cutoff



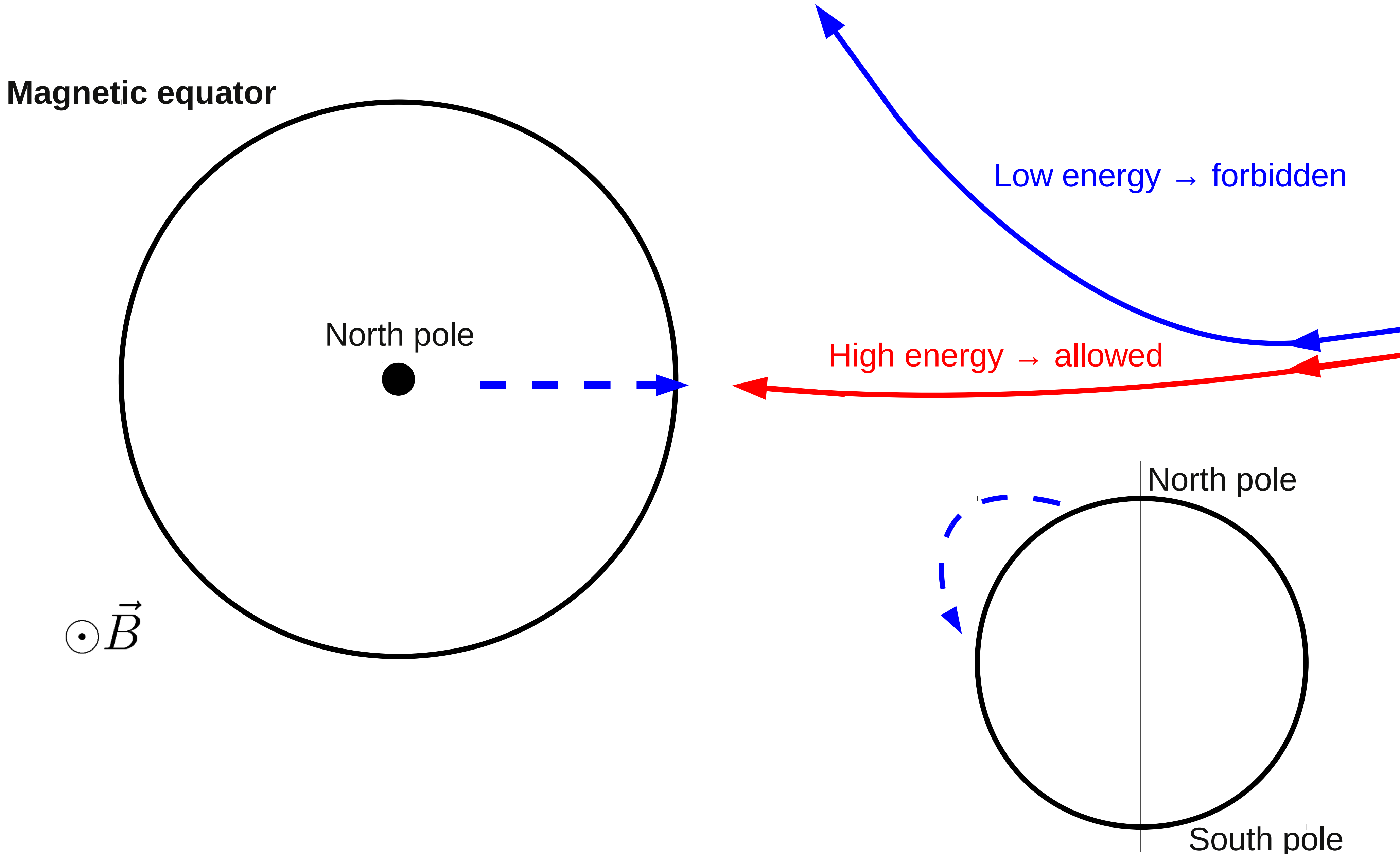
Magnetic poles (galactic protons)

Magnetic equator

Secondary reentrant albedo protons

Galactic protons

CR trajectories in the Earth's magnetic field



Magnetic equator

North pole

Low energy → forbidden

High energy → allowed

$\odot\vec{B}$

North pole

South pole

The **cutoff rigidity R_c** in a given position and for a given direction of the incoming particle is the minimum rigidity that a galactic charged particle needs to reach that position from that direction.

In dipole approximation, the relation is analytical:

$$\boxed{R_c(GV) = \frac{59.6\,GV\ cos^4\lambda}{r^2\ [1+\sqrt{1-cos^3\lambda\ cos\epsilon\ sin\ \zeta}]^2}}$$

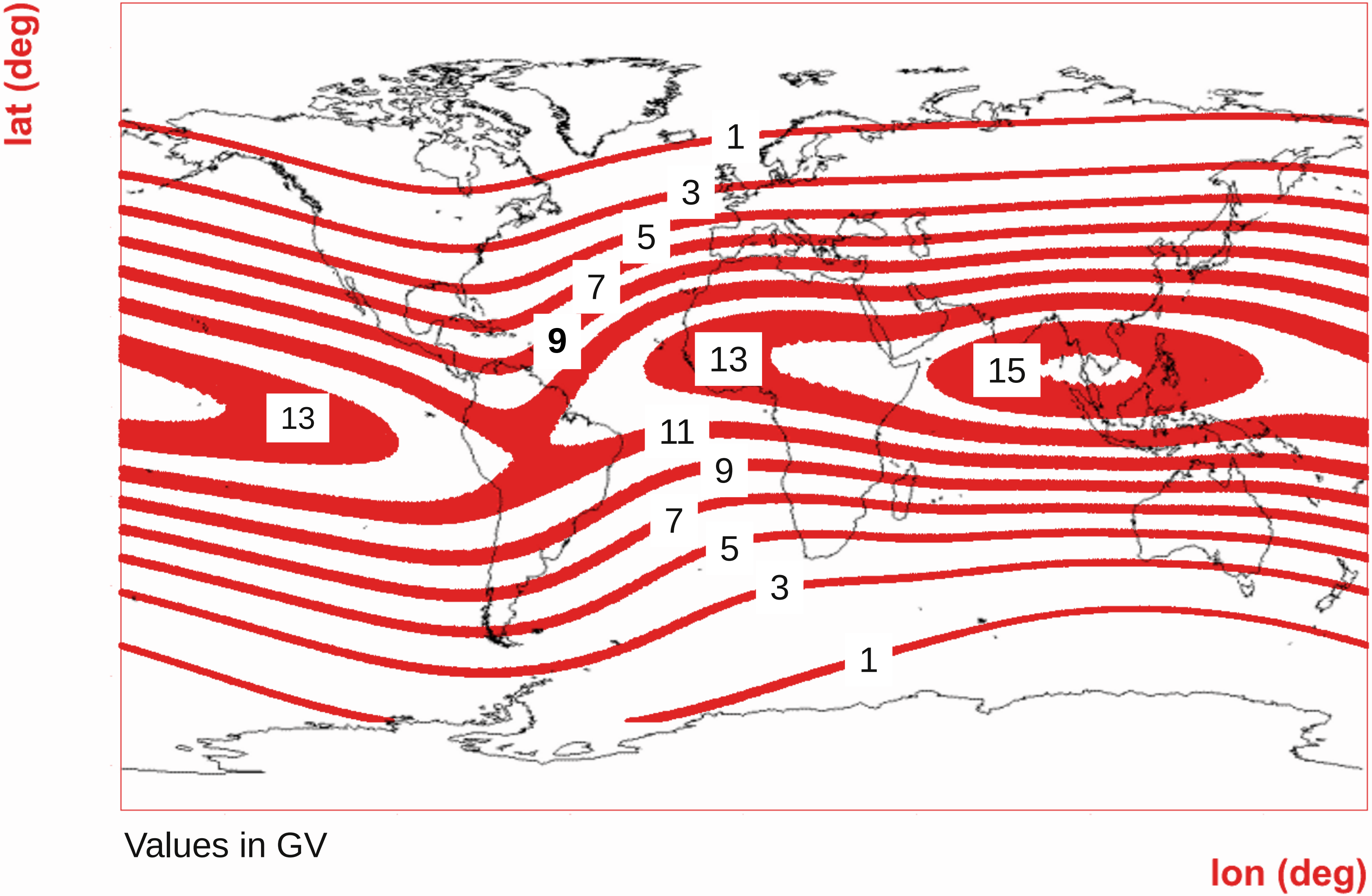
$$\left.\begin{array}{l}\lambda = \text{latitude}\\ r = \text{radius in Earth radii}\end{array}\right\}\quad \begin{array}{c}\text{satellite position}\\ \text{(magnetic coordinates)}\end{array}$$

$$\left.\begin{array}{l}\epsilon = \text{zenithal angle of the particle}\\ \zeta = \text{azimuthal angle of the particle (east} = 0^\circ)\end{array}\right\}\quad \begin{array}{c}\text{particle direction}\\ \text{(local ref. frame)}\end{array}$$

For vertically incoming particles:

$$\epsilon = 0^\circ \quad \Longrightarrow \quad \boxed{R_c(GV) = \frac{14.9\,GV\ cos^4\lambda}{r^2}}$$

Isolines of cutoff rigidities



Values in GV

Geomagnetic cut for galactic particles

For vertically incoming particle:

$$\text{satellite position } \rightarrow R_c$$

To take into account a) non vertical particle, b) non-dipolarity of the magnetic field, we use:

$$\boxed{R > 1.3 \times R_c}$$

Examples:

$$\left\{ \begin{array}{l} \lambda = 80^\circ \\ altitude = 500km \end{array} \right. \Longrightarrow R_c \approx 12MV \Longrightarrow \begin{array}{l} \text{Events are accepted if:} \\ R > 16MV \end{array}$$

(below PAMELA acceptance!)

$$\left\{ \begin{array}{l} \lambda = 60^\circ \\ altitude = 500km \end{array} \right. \Longrightarrow R_c \approx 800MV \Longrightarrow \begin{array}{l} \text{Events are accepted if:} \\ R > 1050MV \end{array}$$

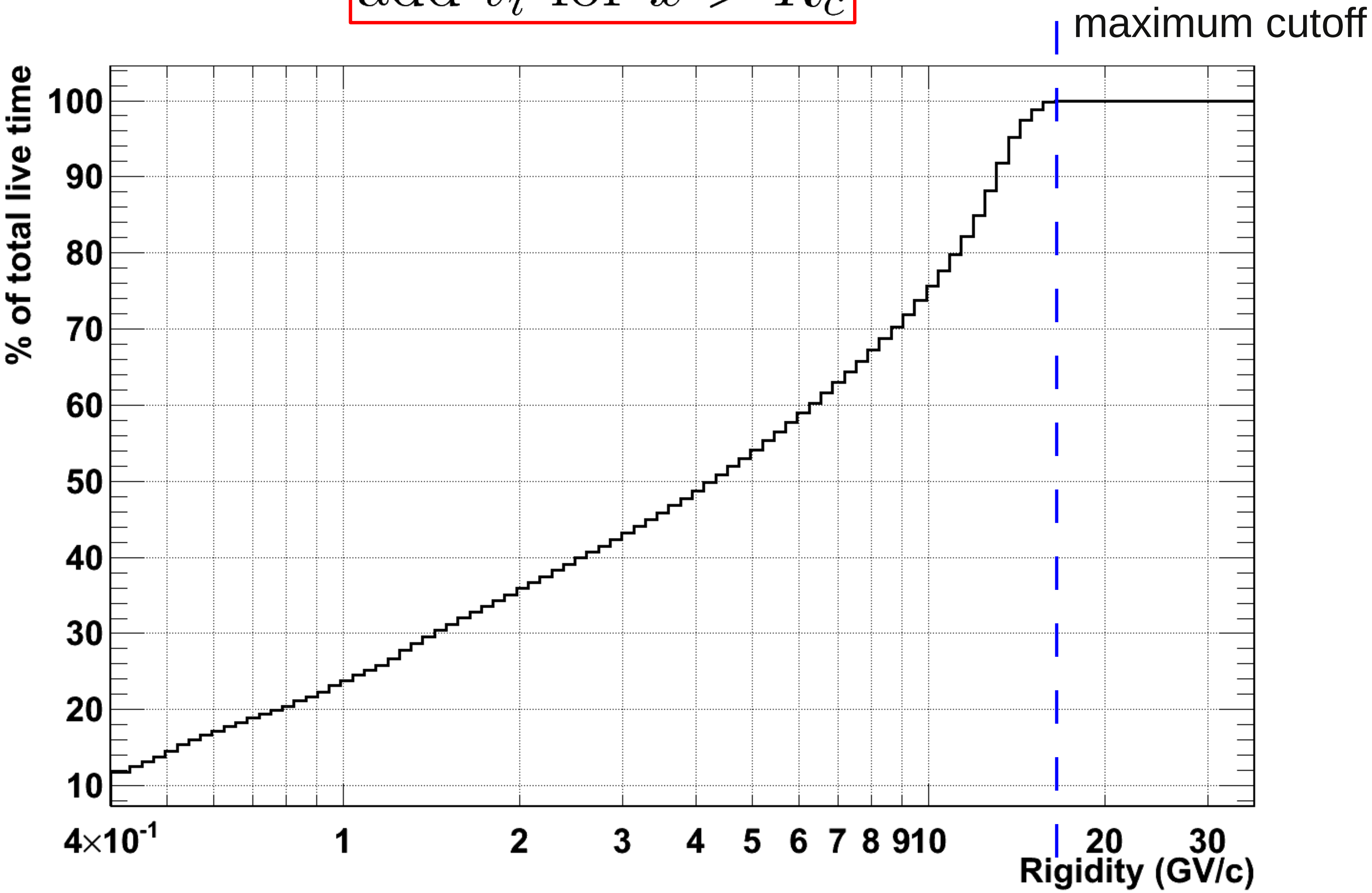
Live time for galactic fluxes

trigger i-th

satellite position $\rightarrow R_c$

live time t_i from trigger board

add t_i for $x > R_c$



Fluxes

Efficiencies

Efficiency of the selection S: fraction of events selected by S on
a sample of good (not background) events

$$\varepsilon_S = \frac{N_S^{good}}{\boxed{N^{good}}} \longrightarrow$$

Efficiency sample
Must be **unbiased** from S !

Efficiency samples can be obtained by:

- Simulations
- Test beams
- Flight data

Each method has issues:

- In-flight conditions might vary with time, bringing to time-dependent efficiencies (launch);
- Simulation cannot reproduce everything;
- Beam test data are not isotropic

The total efficiency is the combination of the efficiencies of every selection cut used:

$$\varepsilon = \varepsilon_{TRIGGER} \times \varepsilon_{TRK} \times \varepsilon_{TOF+ANTI} \times \varepsilon_{dE/dx} \times \varepsilon_{MDR}$$

➔ Each efficiency sample must representative of the flux sample.
➔ All the terms of this product must be uncorrelated to each other. High risk to count an efficiency more than once.
➔ Each term groups together correlated cuts.

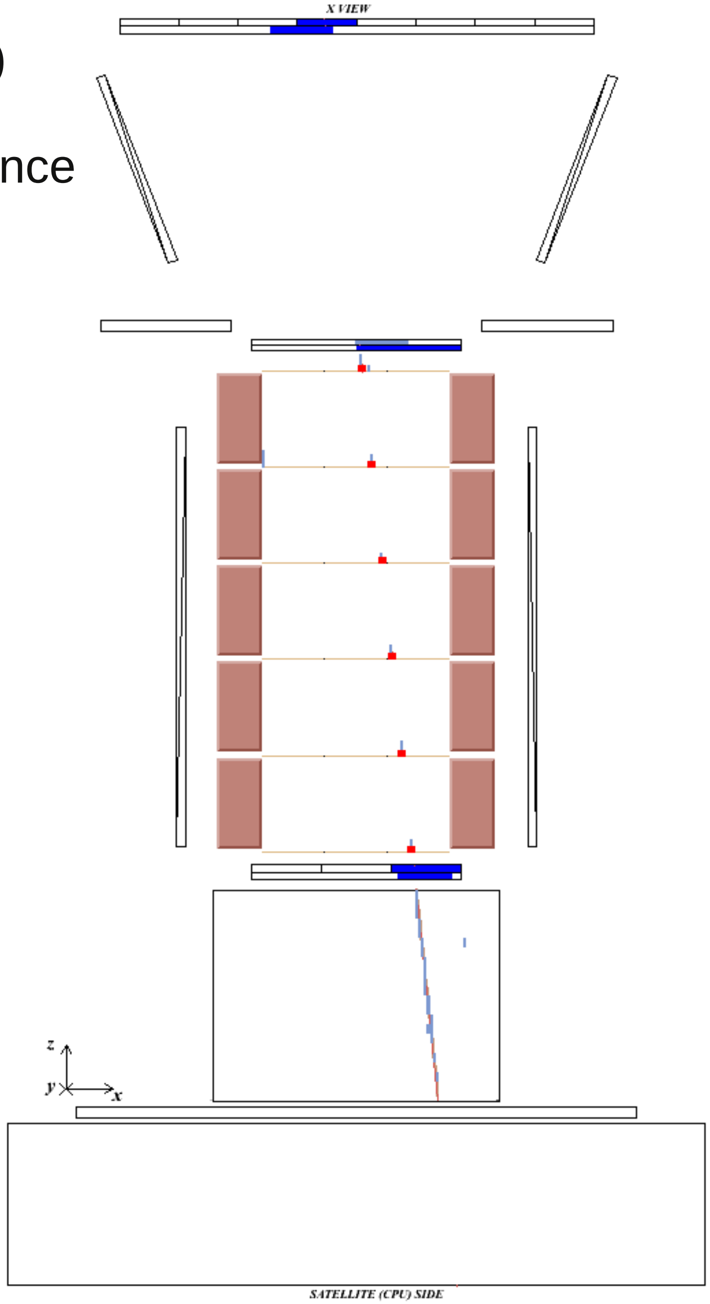
Additional steps:
- consider time dependence of the efficiencies;
- consider contaminations in the flux and efficiency samples;
- estimate systematics.

Tracker efficiency

Study: efficiency of the fitting algorithm (there is one fitted track)

Efficiency sample: **primary** downward-going protons in acceptance

Clean TOF and anticoincidence pattern

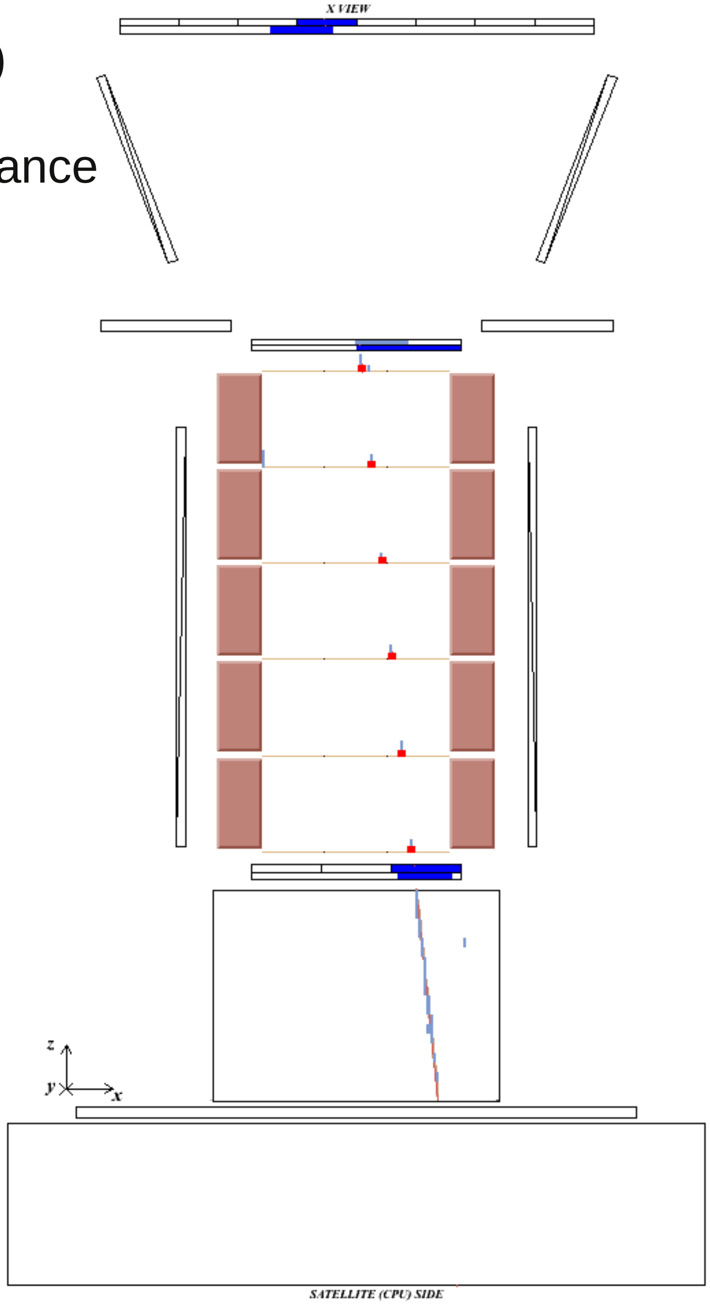


Tracker efficiency

Study: efficiency of the fitting algorithm (there is one fitted track)

Efficiency sample: primary **downward-going** protons in acceptance

Velocity from TOF



Tracker efficiency

Study: efficiency of the fitting algorithm (there is one fitted track)

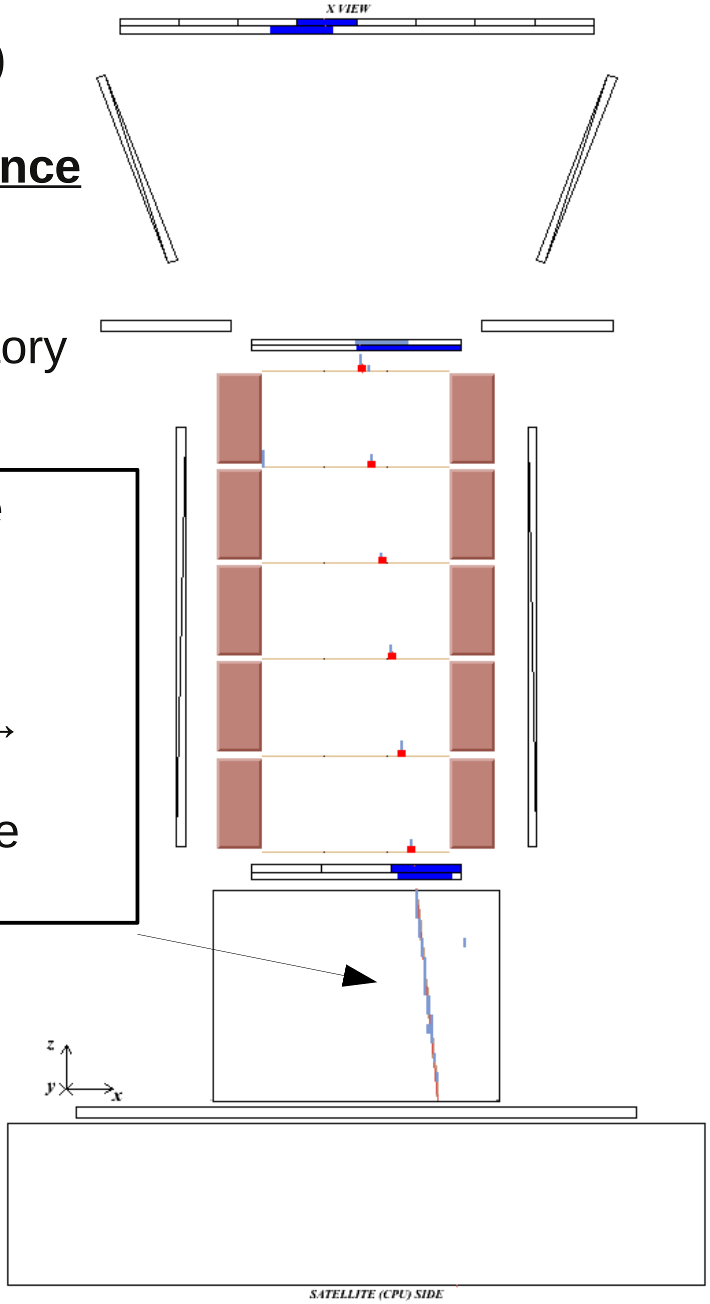
Efficiency sample: primary downward-going protons **in acceptance**

How? We need an independent reliable tool to check the trajectory

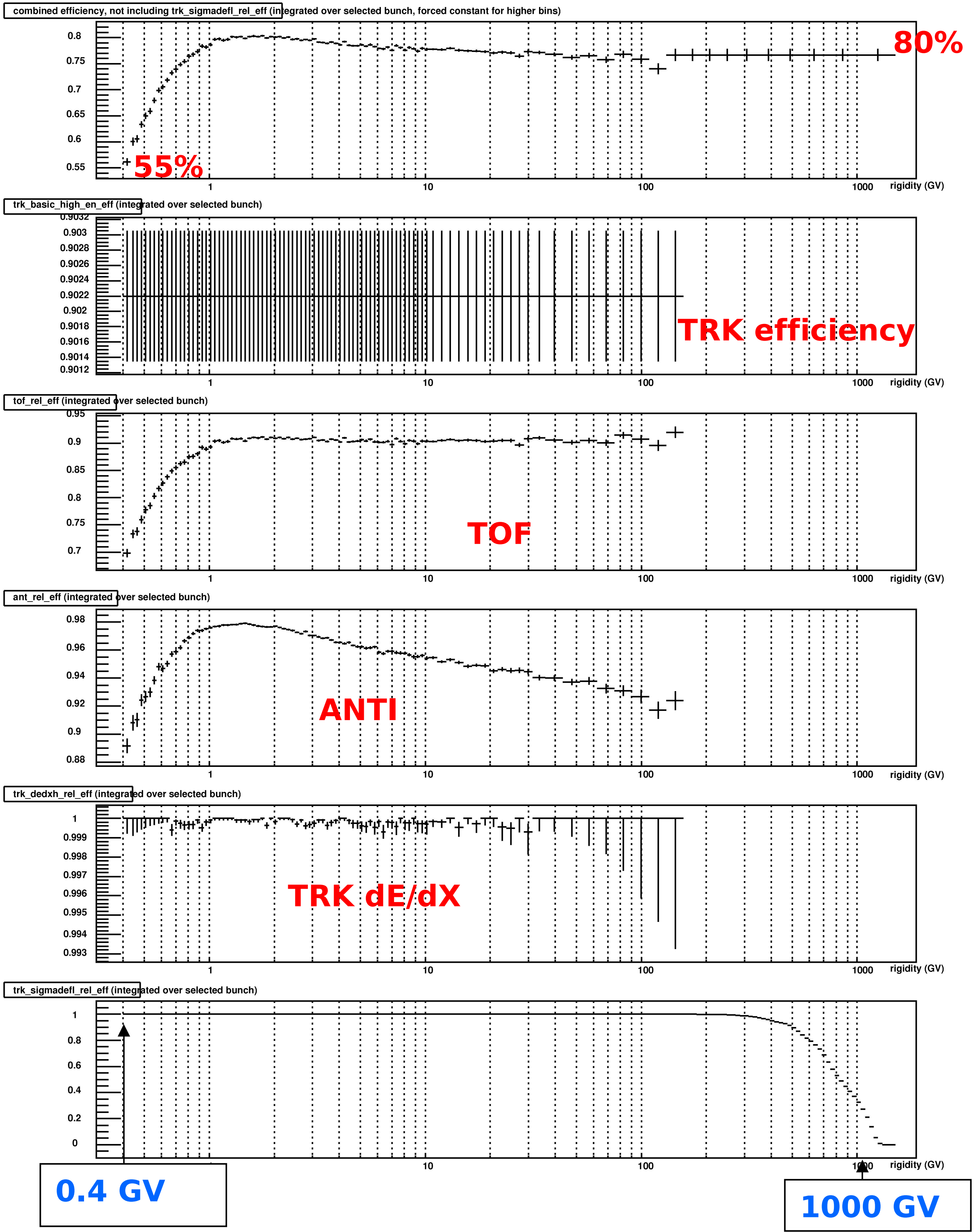
The **calorimeter** (44 silicon planes) provides a track that can be extrapolated up to S1. Requirements:
1) Energy release compatible with a proton.
2) Not interacting.
3) Crossing all the calorimeter (this select high energy protons → straight tracks → linear extrapolation possible).
4) Low resolution (in comparison with the tracker): safe cut at the edges of the acceptance planes is needed.

Drawback: no rigidity information. But the efficiency is ~ flat (small correction with simulation).

At **low energy** (< 1 GV)**:** the TOF provides topology velocity (→ momentum if select protons)



Combined efficiency



80%

55%

TRK efficiency

TOF

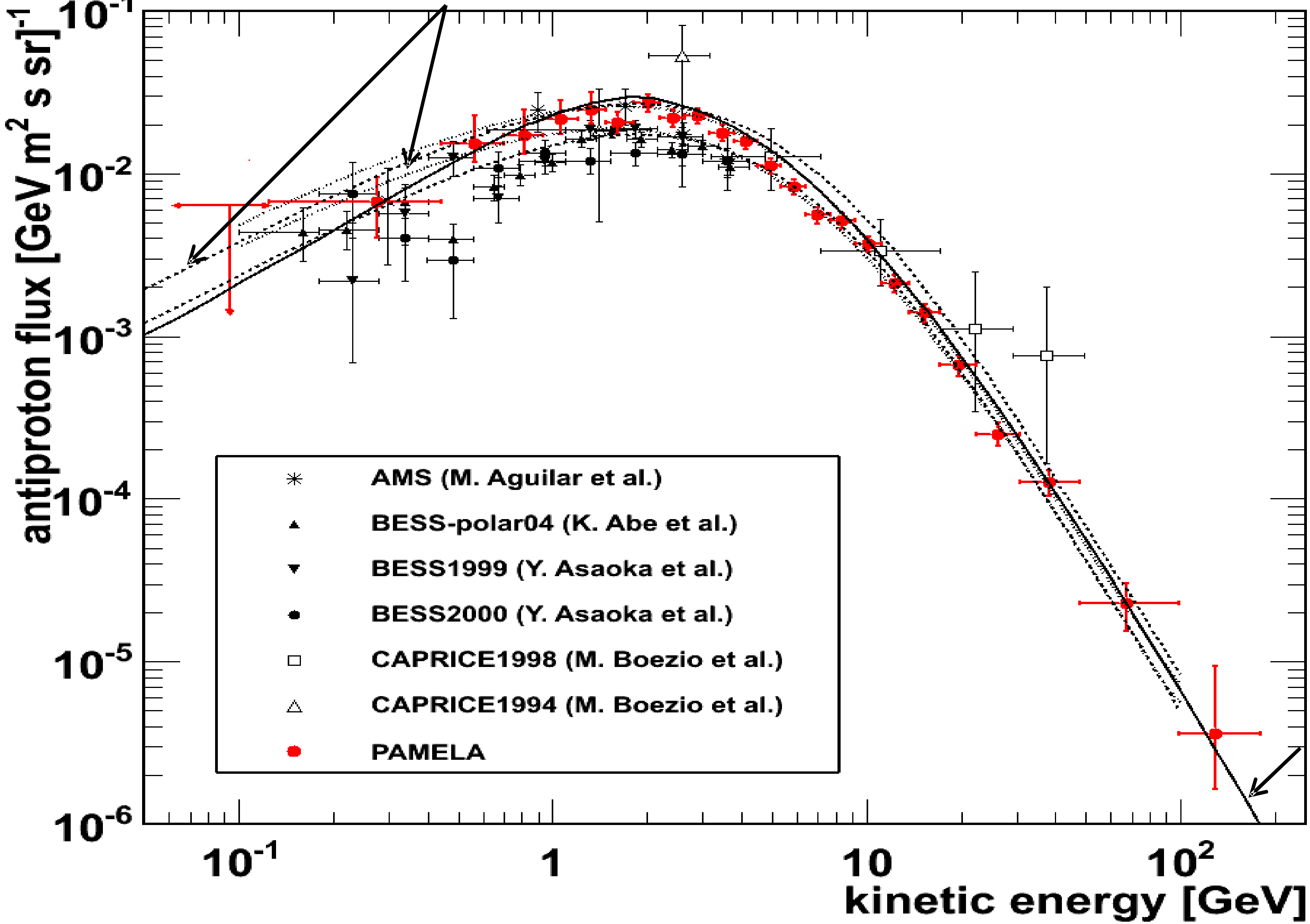
ANTI

TRK dE/dX

0.4 GV

1000 GV

Antiproton Spectrum



Donato et al. (ApJ 563 (2001) 172)

Ptuskin et al.
(ApJ 642
(2006) 902)

O. Adriani et al., Physical Review Letters 105, no. 12 (2010), 121101

Thanks!