Real data samples: track by track evaluation

Y.Fisyak

Outlook

- Data samples
- Track by track
- Efficiencies
- Momentum reconstruction
- Conclusions

Data samples

1. 2004 AuAu200 : Sti (4.8 k), StiCA(4.8 k), Stv (2.9 k), StvCA(4.8 k)
2. 2009 pp500: Sti(82.9 k), StiCA(10k), Stv(74 k), StvCA (73.3 k)
3. 2010 AuAu200: 10k for all Sti, StiCA, Stv and StvCA

These samples were processed on the same type of CPU which allows to compare CPU usage.

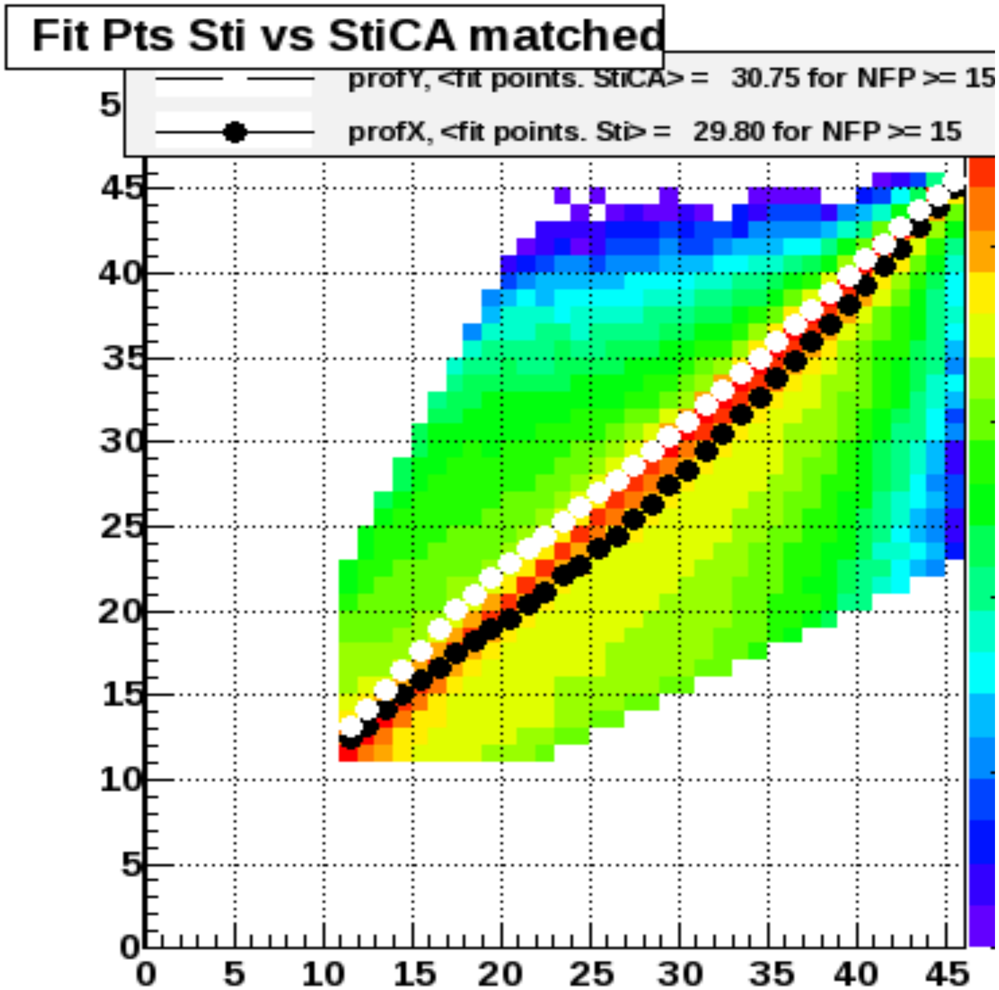
Track by Track analysis

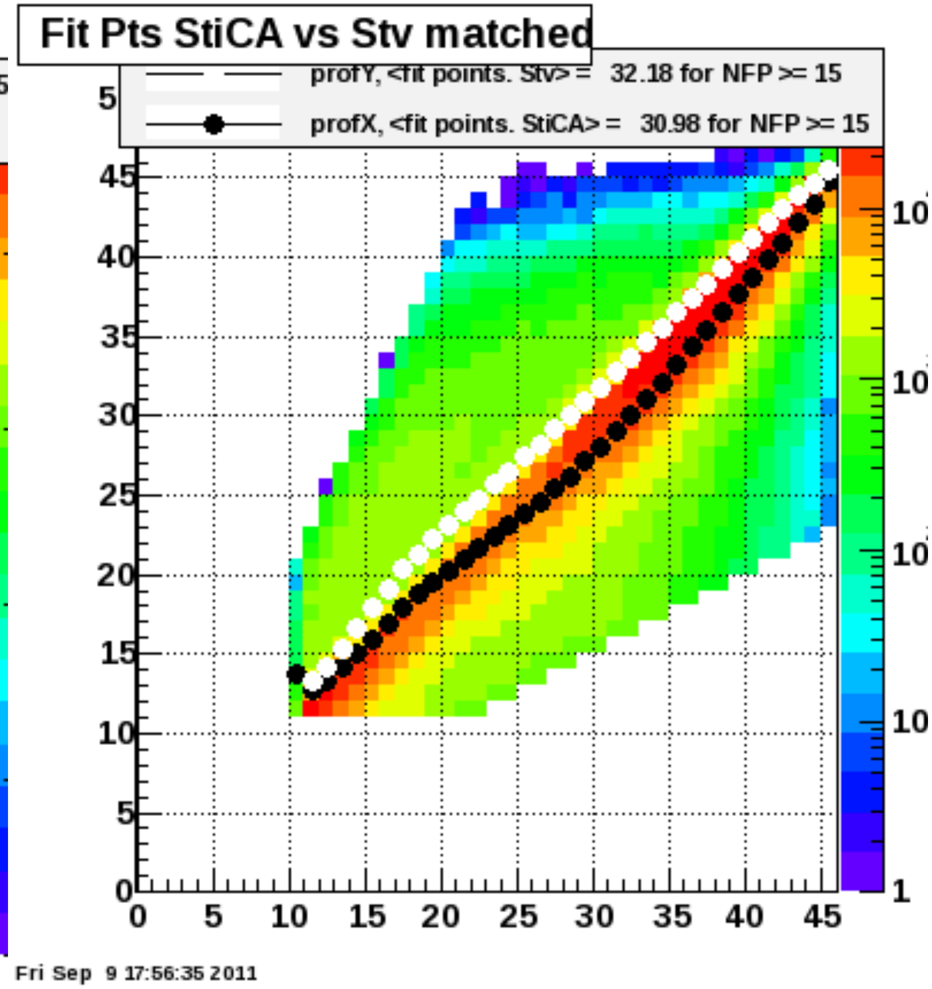
- This analysis is based on matching tracks from different reconstructions by the list of hits which were used to built these tracks.
- The algorithm has been developed ~10 years ago by Manuel.
- Thus this analysis requires full StEvent with hits because track matching is based on the hits.
- The complete list of plots obtained in this analysis can be found at

http://www4.rcf.bnl.gov/~fisyak/star/RECO/Eval/TbyT/

- In this talk I will flush the most interesting ones.

No. of fit points





StiCA has ~1 extra fit point with respect to Sti.

Stv has ~1 extra fit points with respect to StiCA.

Scanning efficiencies from "Statistical methods in experimental physics" by W. T. Eadie, Frederick James.

Equation (2.5) is a necessary and sufficient condition for A and B to be independent, which happens to be quite useful. We note in passing that the correlation coefficient, to be defined later, vanishes if A and B are independent, but lack of correlation does *not* imply independence.

2.2.3. *Example of the addition law: scanning efficiency*

Suppose that data has been recorded on a photographic medium which must be scanned manually to find the events of interest. The film is scanned twice for the same kind of event. A number $(C + D_1)$ of events is found in the first scan, and $(C + D_2)$ events in the second scan. C denotes the events found in both scans, whereas D_1 and D_2 are found only in the first or second scan respectively. We wish to estimate the true number of events in the film, and the over-all scanning efficiency, under the assumptions that in each scan all events are equally likely to be found, and that the two scans are independent.

The first assumption permits us to deal with the sets 1 and 2 of events found in the first and second scans, instead of with each event separately. Thus every event on the film has probability $P(1)$ of being found in the first scan and $P(2)$ of being found in the second scan.

The second assumption implies [from Eqs. (2.3) to (2.5)], that

$$P(1|2) = P(1)$$

or

$$P(2|1) = P(2)$$

or

$$P(1 \text{ and } 2) = P(1) \cdot P(2)\,.$$

From the numbers of events observed, we can estimate the scanning efficiencies in the two scans by

$$\hat{P}(1|2) = \frac{C}{C + D_2} = \hat{P}(1)$$

and

$$\hat{P}(2|1) = \frac{C}{C + D_1} = \hat{P}(2)\,.$$

To estimate the over-all scanning efficiency, we use Eq. (2.2), to give us

$$\begin{aligned}
\hat{P}(1 \text{ or } 2) &= \hat{P}(1) + \hat{P}(2) - \hat{P}(1 \text{ and } 2) \\
&= \hat{P}(1) + \hat{P}(2) - \hat{P}(1) \cdot \hat{P}(2) \\
&= \frac{C}{C + D_2} + \frac{C}{C + D_1} - \frac{C^2}{(C + D_1)(C + D_2)} \\
&= \frac{C(C + D_1 + D_2)}{(C + D_1)(C + D_2)}\,.
\end{aligned}$$

To estimate the total number, N, of events on the film, we may use

$$\hat{P}(1) = \frac{C + D_1}{\hat{N}}\,,$$

to give

$$\hat{N} = \frac{(C + D_1)(C + D_2)}{C}\,.$$

A more detailed discussion of this problem has been given in the literature [Evans], [Knop].

2.2.4. *Bayes theorem for discrete events*

The theorem which links $P(A|B)$ to $P(B|A)$ is *Bayes theorem* [Bayes]. For sets A and B (of events X_i) it states that

$$P(A|B) = P(B|A) \cdot P(A)/P(B)\,. \tag{2.6}$$

Evidently this follows from the definition of conditional probability, Eq. (2.3). More generally, if $A_1, \ldots, A_N$ are exclusive and exhaustive sets (i.e. an observed elementary event must belong to one and only one of the sets A_i), and if B is any event, then Bayes theorem can be written as

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum P(B|A_i)P(A_i)}\,. \tag{2.7}$$

Example

An experiment is planned to study leptonic decays of the K^0 meson, and a

To calculate efficiencies I use "scanning girls" example from the above book
("Statistical methods in experimental physics" by W. T. Eadie, Frederick James).

This example considers a task to calculate efficiencies for finding some type of events by two girls scanning Bubble Chamber films.

There are three assumptions:

1.The two girls are independent i.e. one girl don't know anything about results of another.

2.The efficiencies for both girls have sense ($\varepsilon_1 > 0$ and $\varepsilon_2 > 0$) i.e. efficiencies are defined in phase space satisfied the above condition.

3.The girls don't generate ghosts.

If **N** is total events then

- $\mathbf{N_1 = N} \times \varepsilon_1$ is no. of events found by the 1-st girl,
- $\mathbf{N_2 = N} \times \varepsilon_2$ is no. of events found by the 2-nd girl, and
- $\mathbf{N_{12} = N} \times \varepsilon_1 \times \varepsilon_2$ is no. of events found by both girls.

We have 3 equations and 3 unknowns which we can resolve

- $\boldsymbol{\varepsilon_1} = \mathbf{N_{12} / N_2}$
- $\boldsymbol{\varepsilon_2} = \mathbf{N_{12} / N_1}$
- $\mathbf{N = N_1 \times N_2 / N_{12}}$

Just remember that one girls has name Sti and another has name Stv.

2010 AuAu 200, StiCA (•) versus Sti(■)

Legend:

- Color: black for all, others for different sets of ref. multiplicities.

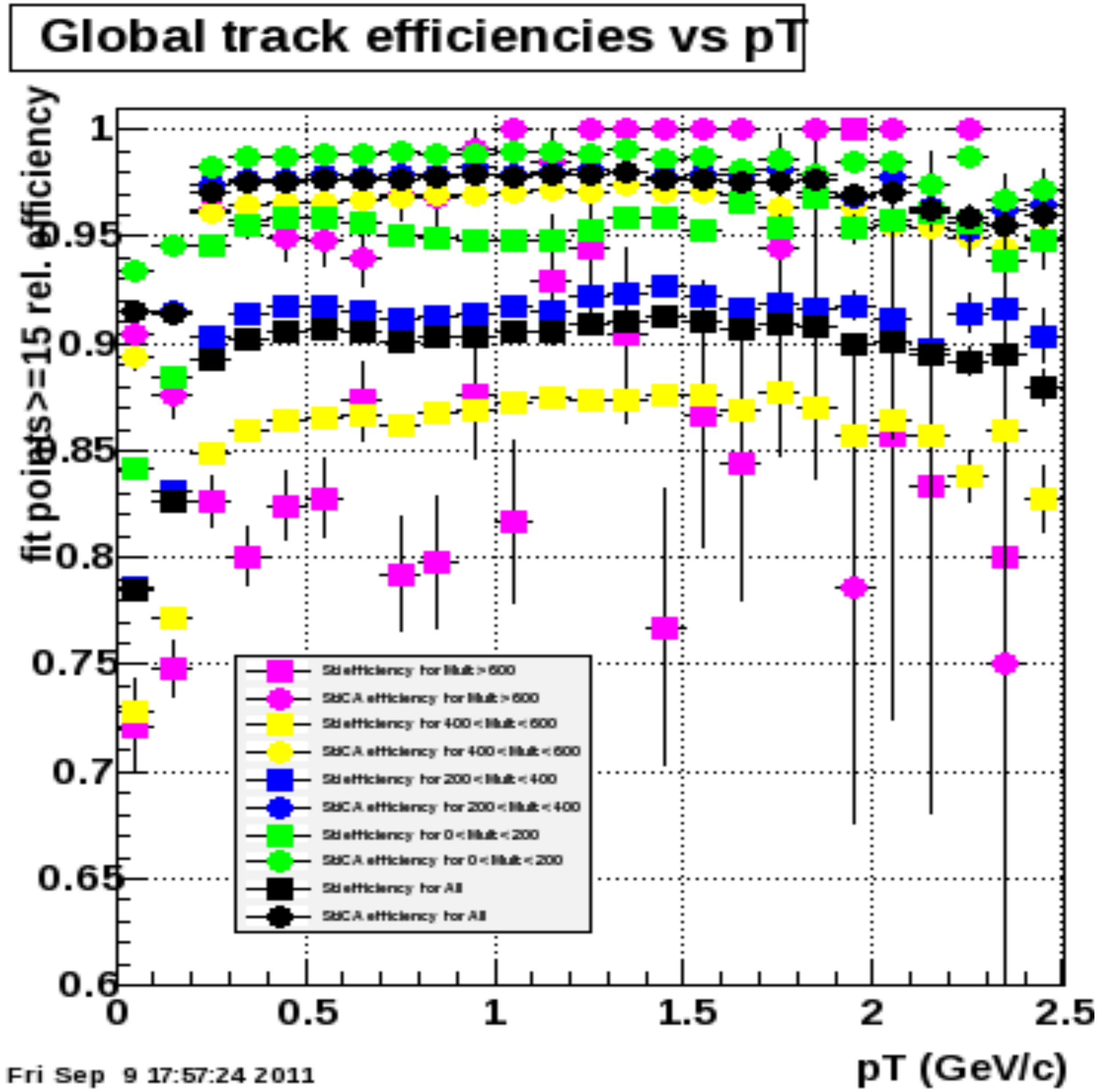
Selection:

- |η| < 0.5,
- No. fit points > 15

Overall efficiency is ~7% higher for StiCA than Sti.

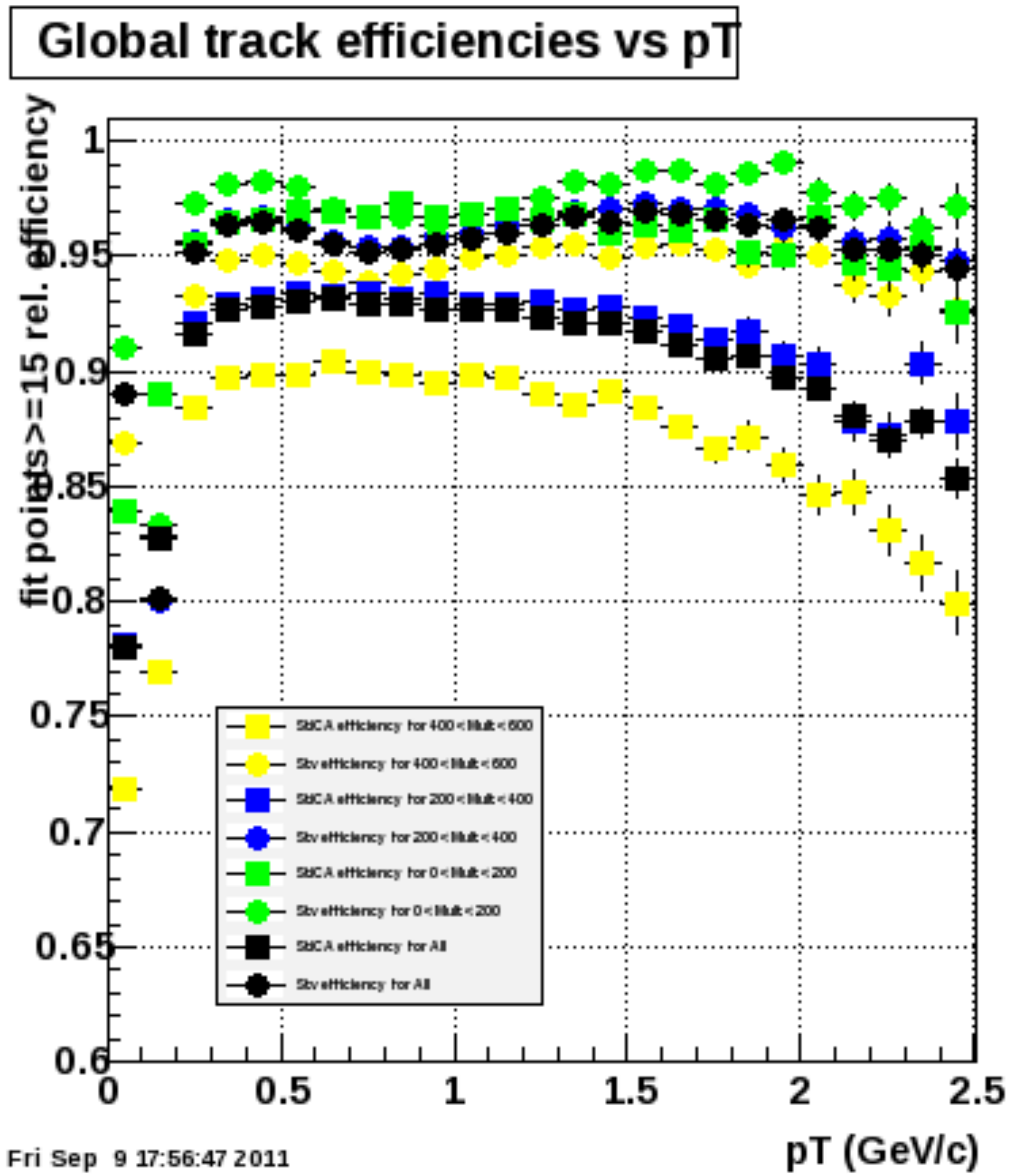
Sti efficiency does depend on ref. mult.

StiCA has much less such dependence.



2010 AuAu 200, Stv (•) versus StiCA(■)

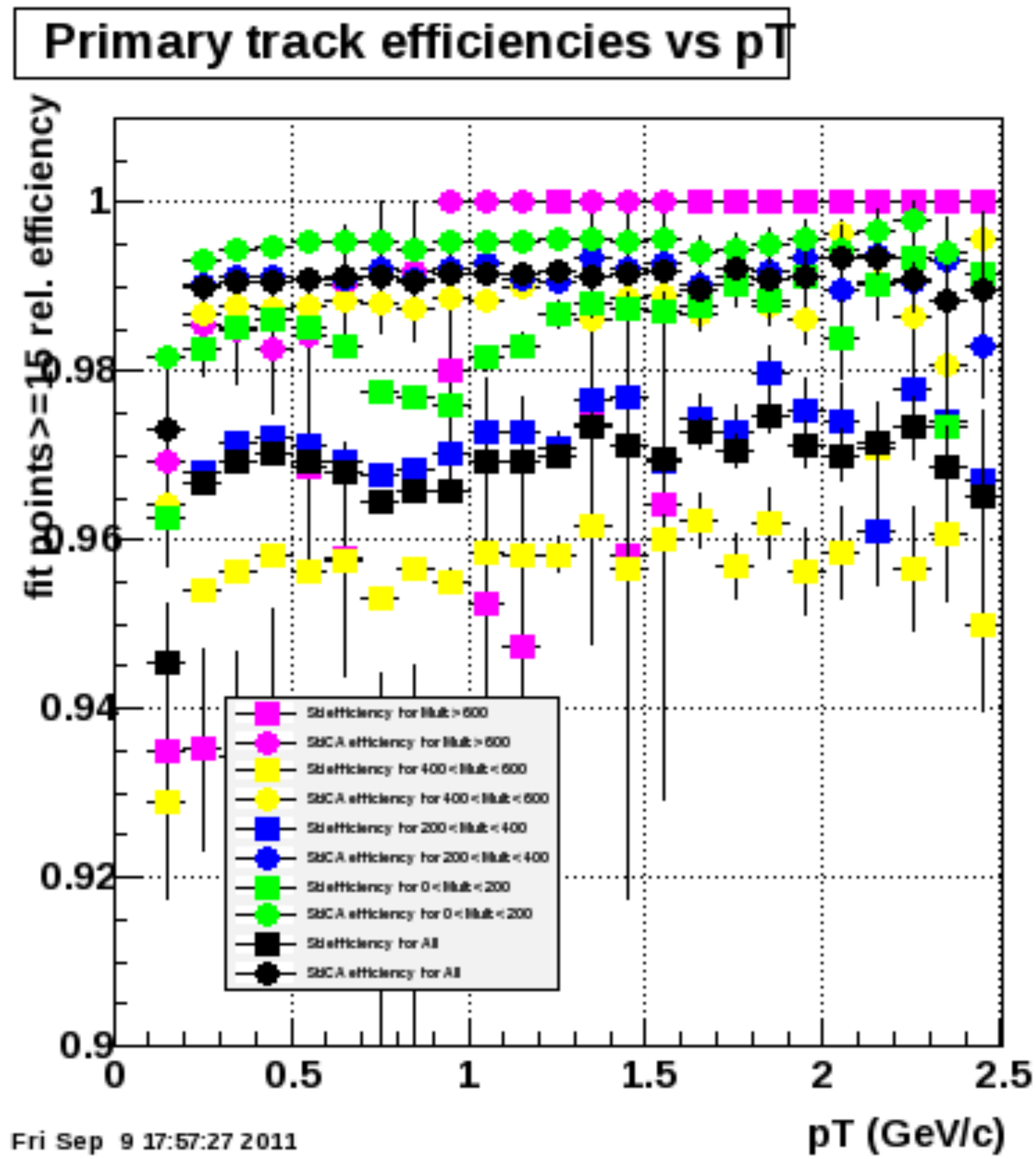
Stv efficiency is higher than StiCA (~2%) and this difference increased with p_T.



2010 AuAu 200, StiCA (•) versus Sti(■)

The same statement as before:

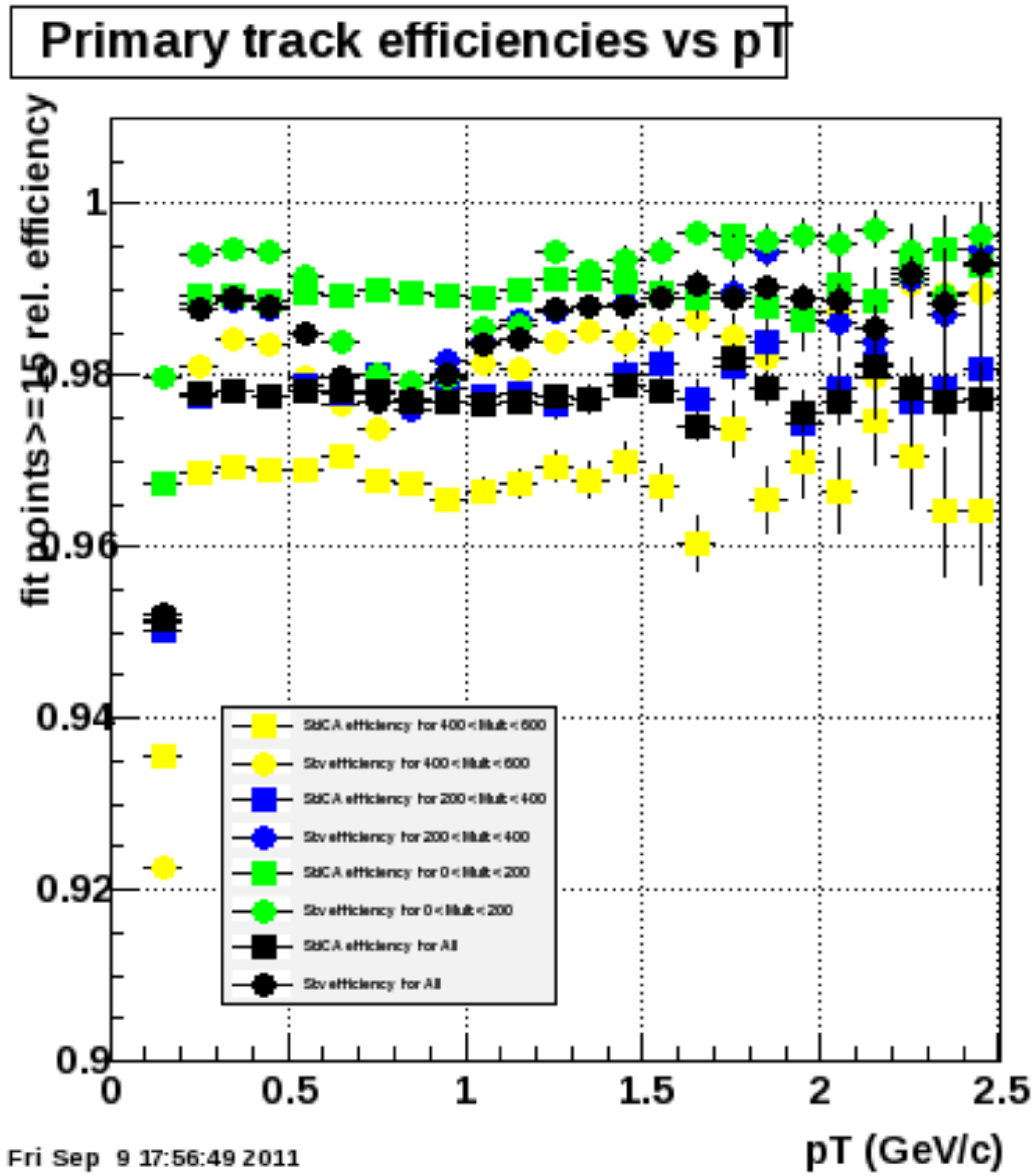
- StiCA efficiency is higher than Sti by ~2%.
- StiCA efficiency has less multiplicity dependence.



2010 AuAu 200, Stv (•) versus StiCA(■)

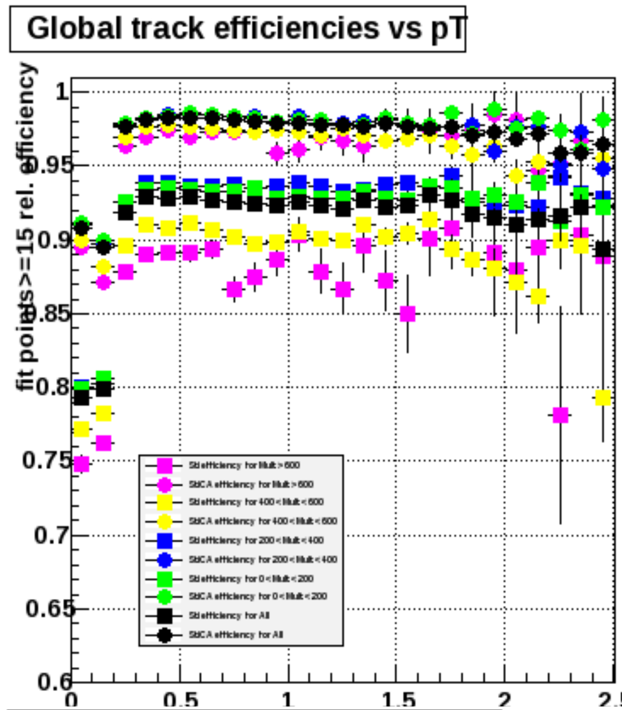
Efficiency Stv is higher by ~1%. There is strange wiggle for StiCA at p_T ~ 0.7 GeV/c.

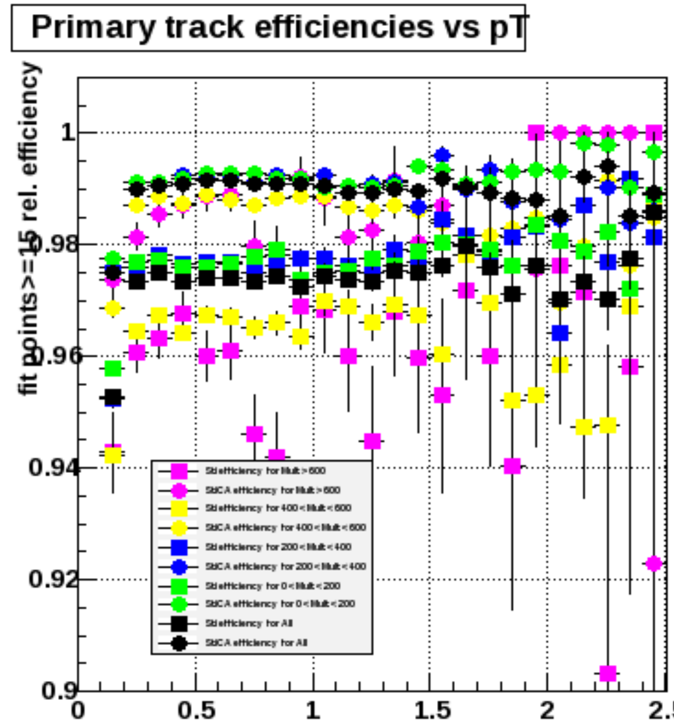
This behavior is common for all other data sets.



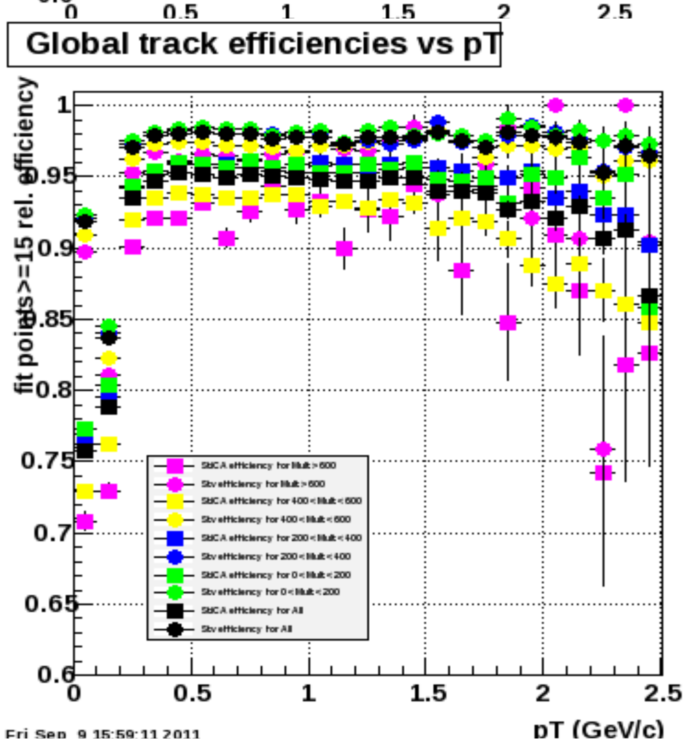
2004
AuAu200

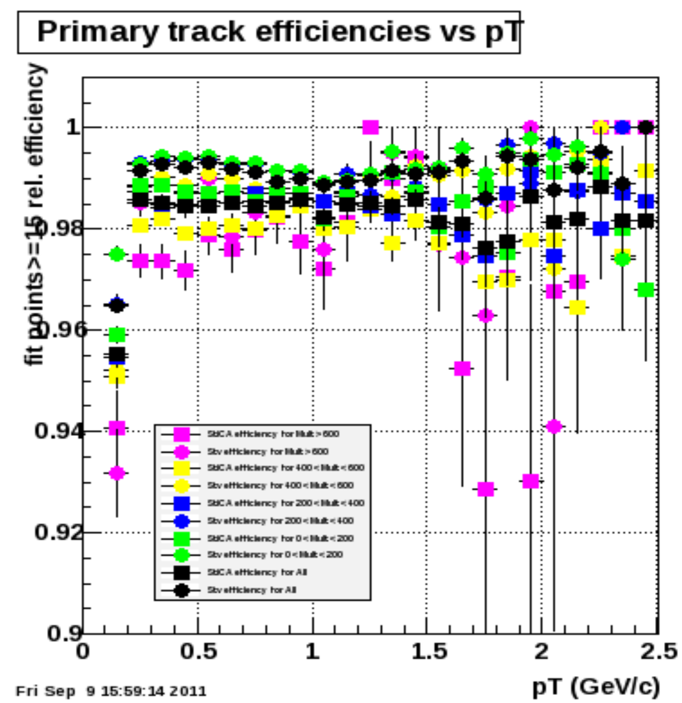
StiCA (•) versus Sti (■)





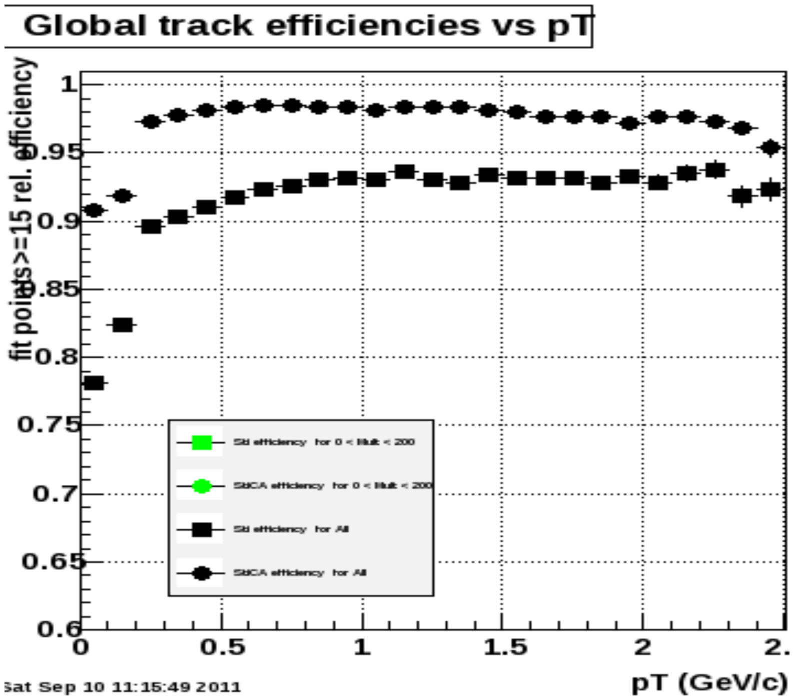
Stv(•) versus StiCA (■)

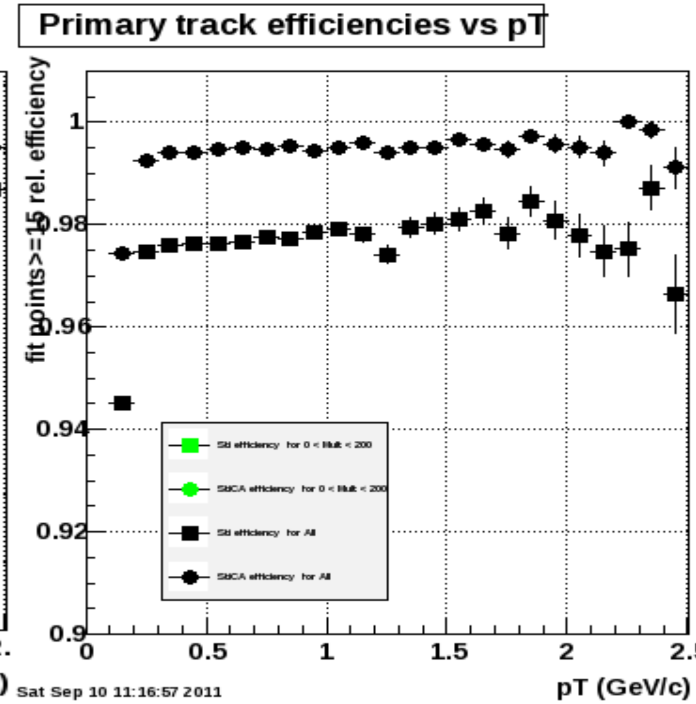




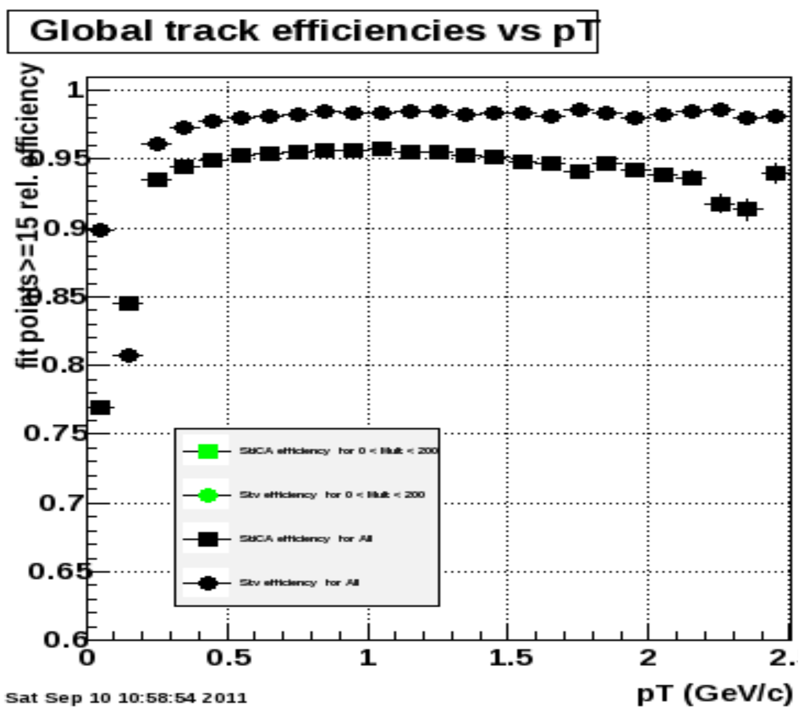
2009, pp500

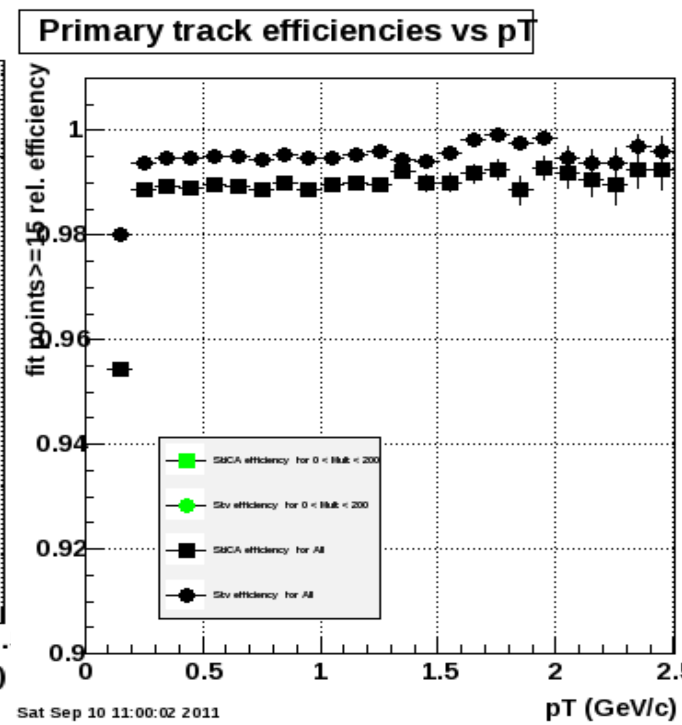
StiCA(•) versus Sti (■)





Stv(•) versus StiCA(■)

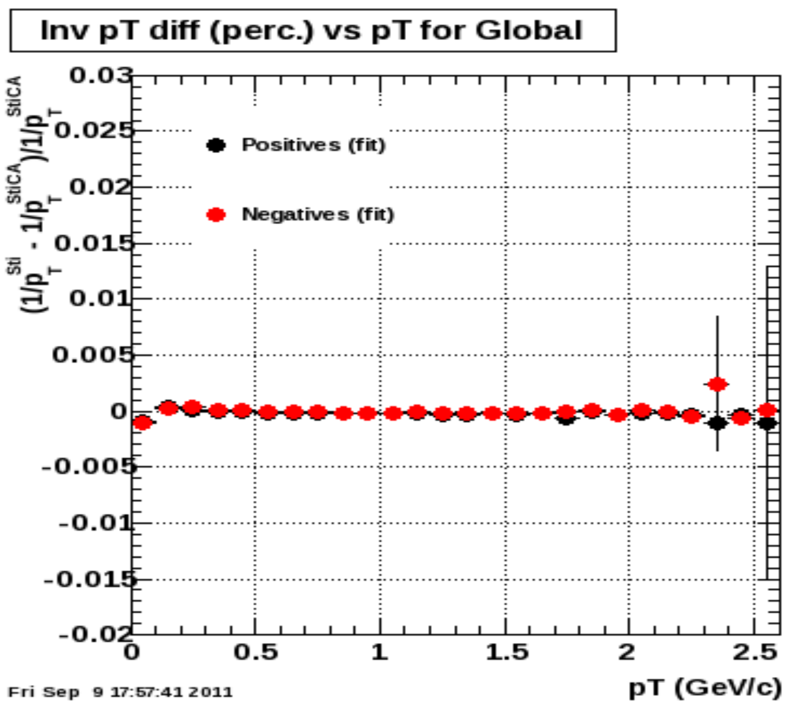


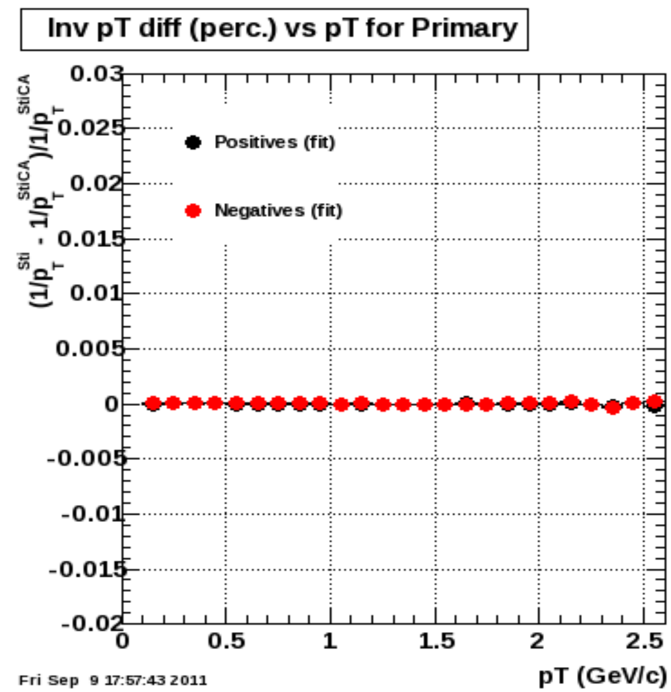


Momentum difference, 2010 AuAu200

Sti versus StiCA:

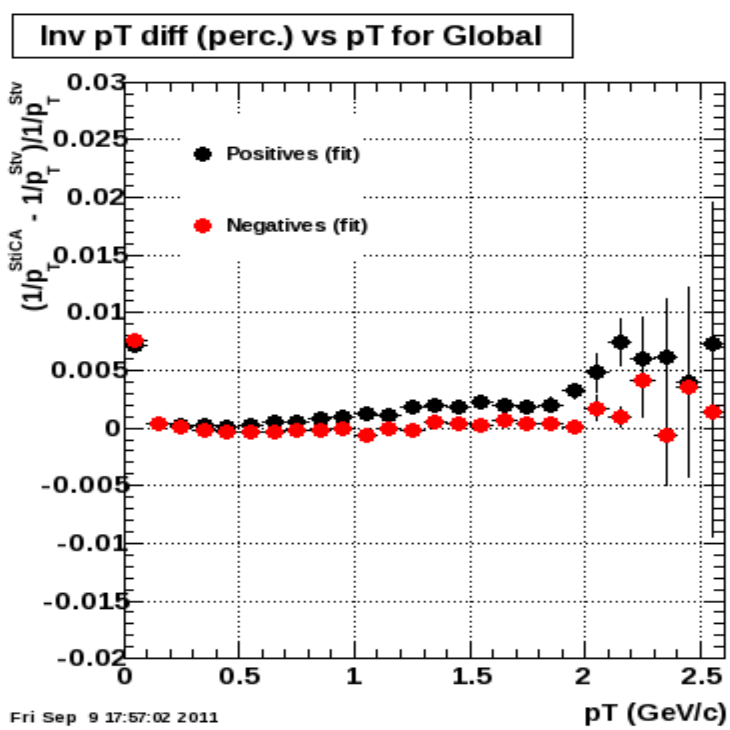
No difference

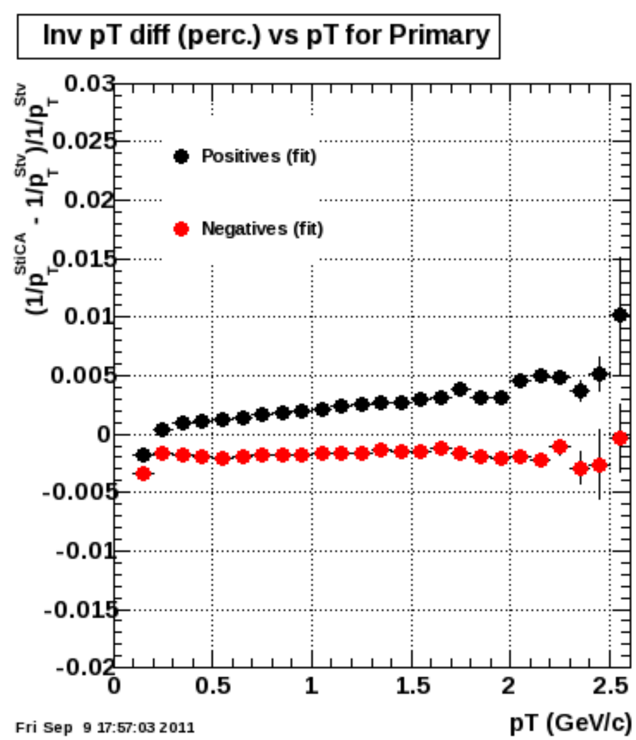




Stv versus StiCA:

- Small difference for global tracks
- For primary tracks sign splitting on the level a few per mills.

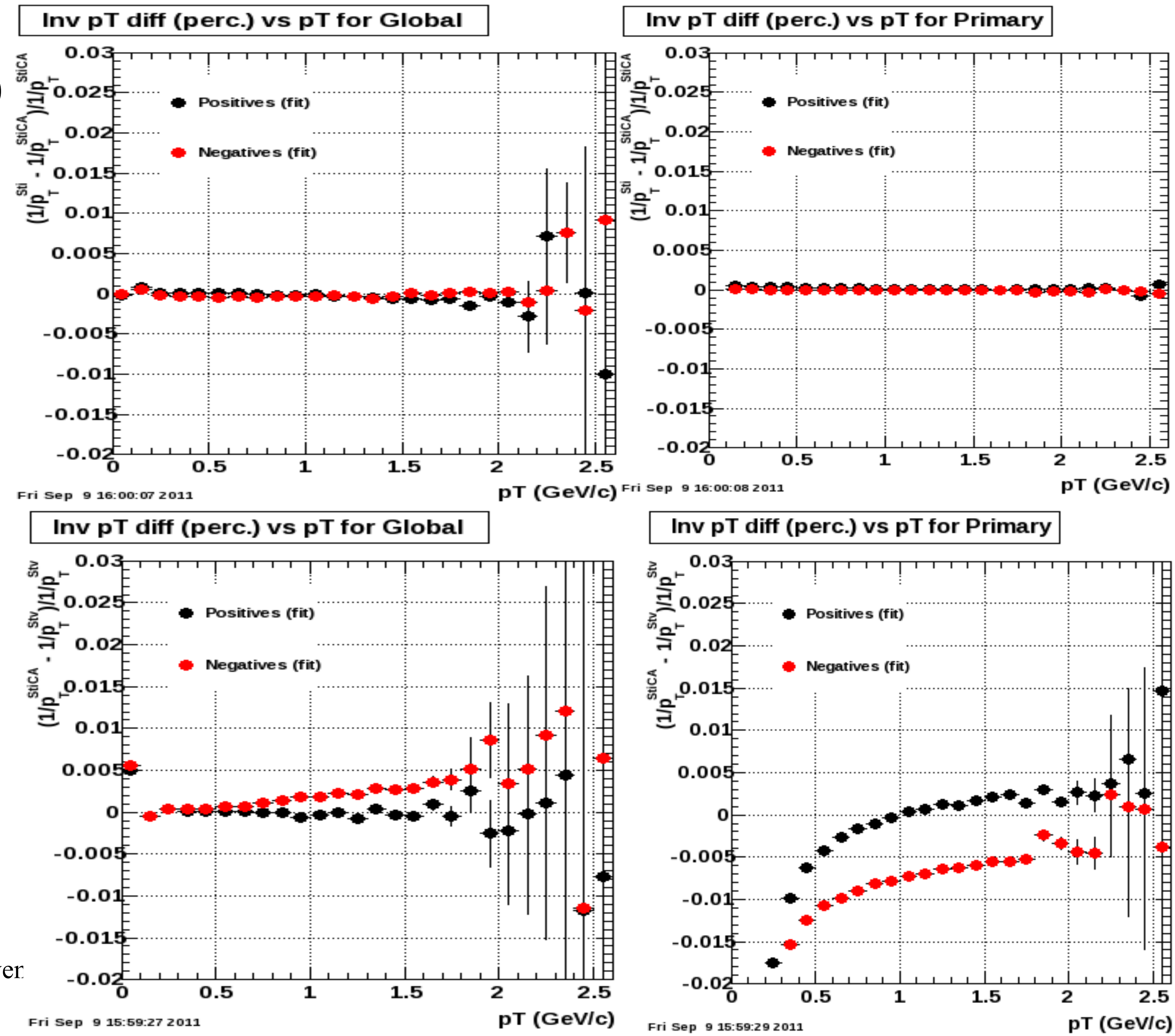




2004, AuAu200

StiCA versus Sti

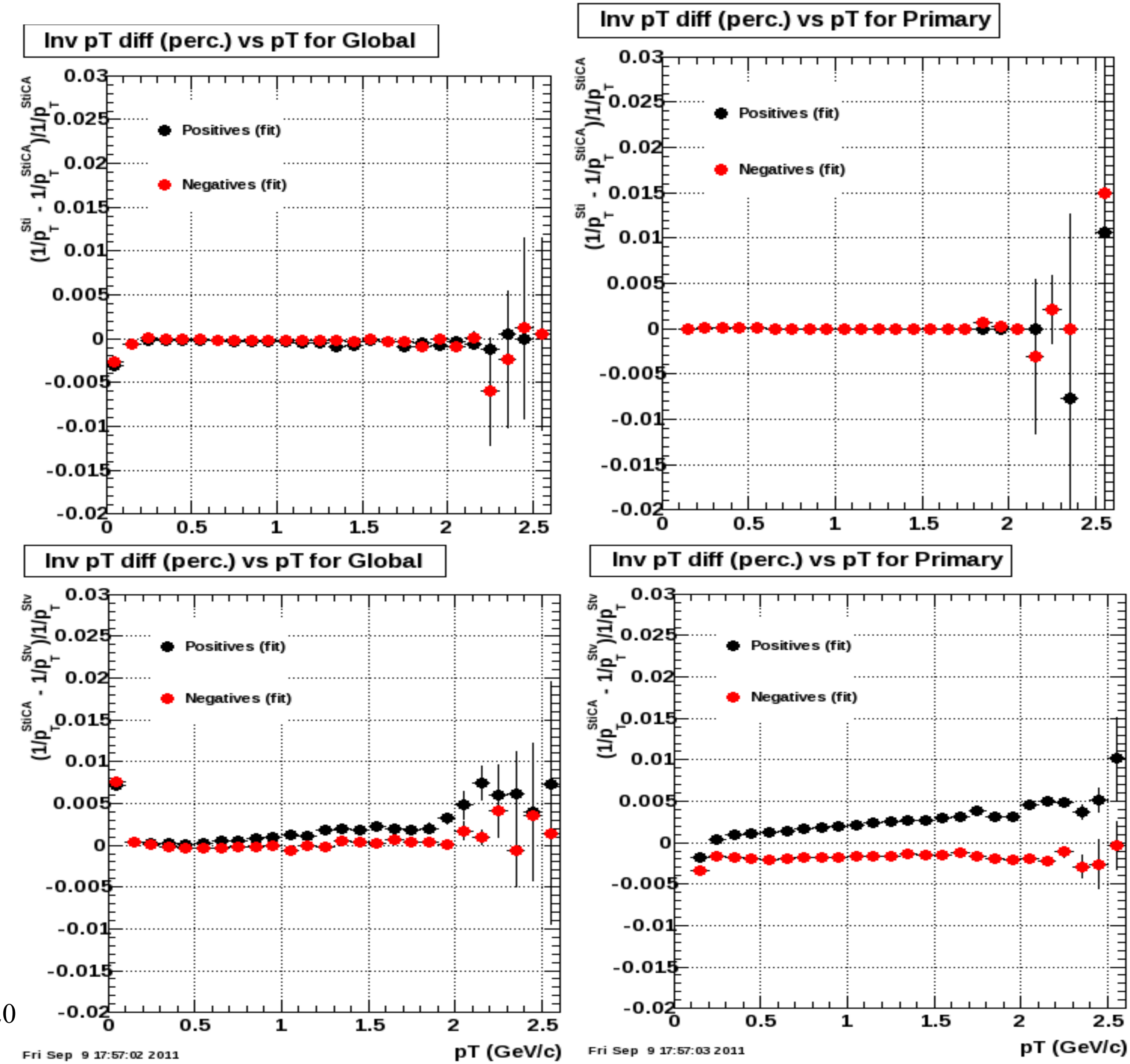
Stv versus StiCA: Sign splitting and shift for primaries on the level ~1%



2010, pp500

StiCA versus Sti:

No differences



Stv versus StiCA:

The same as for

2010 AuAu200

Conclusions

	StiCA/Sti	Stv/StiCA
No. fit points	(+) 1	(+) 1
Efficiency for global tracks	(+) 7%	(+) 2%
Efficiency for primary tracks	(+) 2%	(+) 1%
$1/p_T$ difference for globals	(+) No	(+) < 0.1 %
$1/p_T$ difference for primaries	(+) No	(?) ~0.5 %

Sti to StiCA conclusions are the same as ones presented at S&C meeting in August 2010 and CHEP 2010 in October 2010.