CDF note 9162

Measurement of the relative fraction of $t\bar{t}$ events produced via gluon-gluon fusion in 955 pb^{-1} using a combined fit of low p_T tracks and kinematic variables.

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: January 14, 2008)

We present a measurement of the relative fraction of $t\bar{t}$ events produced via gluon-fusion to the total number of $t\bar{t}$ events, which we call C_f^{true}, using a combined template method based on Neural Networks probabilities and low p_T track multiplicity. Using a total integrated luminosity of 955 pb^{-1} we find $C_f^{true}= 0.07^{+0.15}_{-0.07}$.

Preliminary Results for Winter 2008 Conferences

I. INTRODUCTION

The total production cross section of a given physical process can, in general, be measured simply by counting events in specific final-state channels. The contribution from different production processes to the total production cross section is, however, very hard to measure. Typically, the spin information of the decay products is washed out by the hadronization process of quarks, hence blurring out any difference in the kinematic properties of the different production processes. Additionally, the theoretical calculations of the production cross sections of gg fusion and $q\bar{q}$ annihilation to $t\bar{t}$ also have large uncertainties [1, 2]. For example, the contribution from gg fusion at the Tevatron can vary up to a factor of 2 (from 10-20%).

The measurement of the ratio of gg produced $t\bar{t}$ events to the total (C_f) has been made by two different analysis at CDF. The first method uses low p_T tracks to discriminate gg fusion $t\bar{t}$ events ($t\bar{t}^{gg}$) from $q\bar{q}$ annihilation $t\bar{t}$ events ($t\bar{t}^{qq}$). Because gluons can radiate other partons and gluons, $t\bar{t}^{gg}$ events should have more low p_T tracks. This is a data driven analysis where the shapes of the two components ($t\bar{t}^{gg}$ and $t\bar{t}^{qq}$) are derived from the inclusive dijet data samples. The background shape is then constructed as a combination of the $t\bar{t}^{gg}$ and $t\bar{t}^{qq}$ shapes. The three shapes are then fit to data. A full description of the track method can be found here [3]. The second method uses the kinematics of the production and decay of the $t\bar{t}$ to differentiate the two production mechanisms. The kinematic variables are then used to train a neural network (NN) to increase the sensitivity of the method. This analysis relies on Monte Carlo (MC). A large ensemble of pseudo-experiments (PSE) are generated to calculate the statistical and systematic uncertainties. A full description of the NN method can be found here [4].

These two methods are very complementary. Firstly, the physics is complementary in that the number of tracks is an indirect measurement where we assume $t\bar{t}^{gg}$ events have more low p_T tracks versus the neural network analysis which use very precise kinematics of the different processes. Secondly, the experimental techniques complement each other in that one is data driven and one is MC driven. We combine the two measurements to increase the sensitivity of both.

Because the two methods are already covered in detail in [3] and [4] this note will only address the combination. A brief overview is given in Section II. The correlations (both statistical and systematic) are discussed in Section III. The data with the statistical and systematic results of this analysis are shown in Section V. This note ends with final conclusions in Section VI.

II. OVERVIEW

A. Method

The track method is a purely data driven analysis where the statistical uncertainty comes straight from the fit of the templates to the data. This is very different from the way the uncertainty is handled in the NN method. The NN method uses large ensembles of PSE constructed from MC events and the Feldman-Cousins (FC) method [5] is then used to make the measurement. We choose to follow the FC method, and throw PSE for both methods and then combine the results into one FC confidence band.

B. Monte Carlo

The track method uses data to construct the template since Monte Carlo is known not to correctly model the low p_T tracks that the method uses. Since we want to be able to throw PSE for the track method with MC, we need to ensure that the low p_T tracks are modeled correctly. Figure 1 shows the data shapes. On an event by event basis we morph the MC to match the data when we generate our PSE. A similar process can be done by re-weighting the MC events, but we observed that it provides the same results.

C. Throwing of pseudo-experiments

We throw PSE as described in the NN method note [4]. We vary the value of the $t\bar{t}^{gg}$ production fraction (C_f^{true}) from 0.0 to 1.0 in steps of 0.025. For each value of C_f^{true}we obtain two distributions. One from the track method fit and one from the NN method fit.

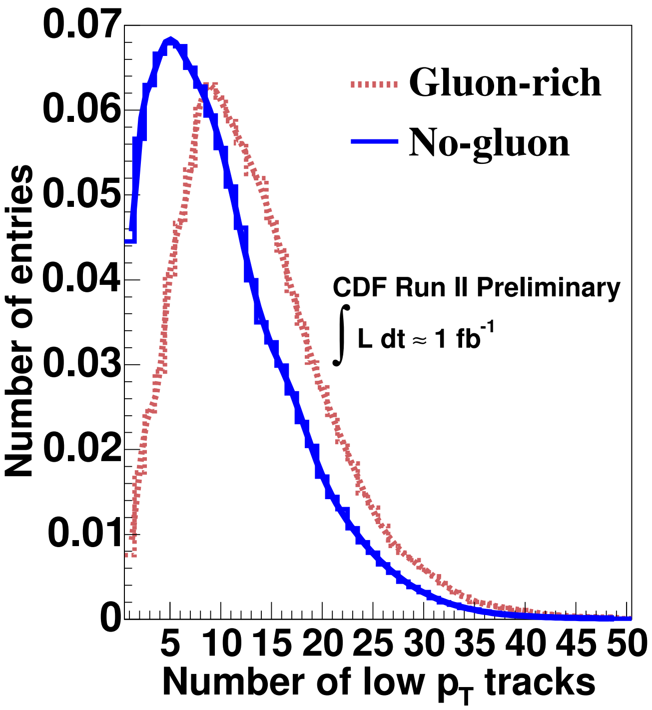

FIG. 1: Number of low p_T for the gluon-rich events and for the no-gluon events. These are fit to the data to find the gluon rich fraction. Shapes are taken from data.

III. CORRELATIONS

When combining two methods, like we are doing, we need to be careful to take into consideration any correlations between the methods.

A. Statistical Correlation

The track method and NN method use very different variables to discriminate $t\bar{t}^{gg}$ from $t\bar{t}^{qq}$. We do not expect the number of low p_T tracks to affect the kinematics of the $t\bar{t}$ pair. Because of this, we expect there to be no or little statistical correlation between the two. To check the correlation, a 10000 events PSE was generated and the results from the track fit and the NN fit are plotted and shown in Figure 2. This plot shows there are no statistical correlations between the two methods.

B. Systematic Correlation

The majority of the systematic uncertainties for the track analysis come from the definition of the low p_T track and other data issues. All the systematic uncertainties for the NN method are MC driven. Since the track method does not use MC and the NN method is not sensitive to any of the track systematics we expect that if any correlations exist they will be small, and thus assumed to be all uncorrelated. One systematic that is correlated between the two methods comes from the background estimate. Table I shows the systematic uncertainties for the track method (using the notation in [3]) and the correlations we assign. We assign the systematic uncertainties due to "non-W variation" and "modeling of the distribution" in the track analysis to be 100% correlated with the background systematic uncertainty in the NN in [4]. The rest of the systematics are taken as uncorrelated. As seen from this table, the majority of the

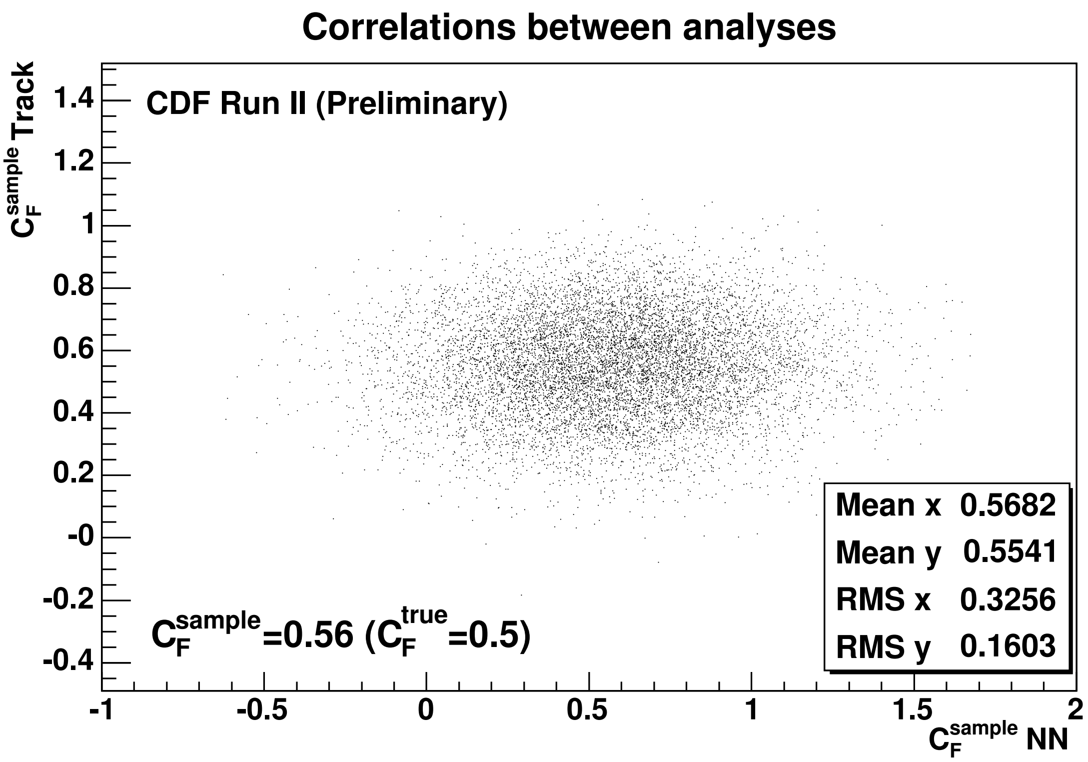

FIG. 2: 10k PSE for C_f^{true}= 0.5: Here is a scatter plot for the fitted result 10000 pseudo-experiment with C_f^{true}= 0.5. The y-axis shows the result for the track fit, while the x-axis shows the results of the NN fit. There is no statistical correlation in the two fits.

Systematic	Track Method		Correlation with NN Method
	f_g	f_g^{bkg}	
trk/jet	±0.051	±0.001	Uncorr
jet E_t cut	±0.021	±0.035	Uncorr
dijet $qq \to gg$	±0.002	±0.019	Uncorr
W+0 jet $< N_g >$	±0.039	±0.007	Uncorr
non-W variation		±0.057	Corr
Modeling the distribution		±0.089	Corr
	C_f^{true}		
f_g	±0.08		Uncorr
f_g^{bkg}	±0.02		Uncorr: ±0.0158, Corr: ±0.0060
f_b	±0.01		Corr

TABLE I: Systematic Uncertainties for the track method and the correlation with the neural network method.

systematic uncertainties are uncorrelated.

IV. COMBINATION

In this section, we describe the prescription we use to combine the track method PSE and NN method PSE into a combined method PSE that we then use to construct the FC confidence bands.

For each set of PSE we obtain a σ for each distribution of the fitted values of $C_f$for the two analyses (σ_{TK}^{stat} and σ_{NN}^{stat}). For each individual PSE we have a value for each method (C_f^{TK} and C_f^{NN}). We then apply the systematics to each pseudo-experiment:

$$\begin{aligned}
C_{f+sys}^{NN} &= C_f^{NN} + Gauss(0, \Delta_{uncorr}^{NN}) + G_{corr} * \Delta_{corr}^{NN} & (1)\\
C_{f+sys}^{TK} &= C_f^{TK} + Gauss(0, \Delta_{uncorr}^{TK}) + G_{corr} * \Delta_{corr}^{TK} & (2)\\
\text{where} & \\
G_{corr} &= Gauss(0, 1), & (3)
\end{aligned}$$

where $Gauss(0, X)$ is a Gaussian random number between 0 and X, Δ is the square root of the appropriate systematics added in quadrature, and the subscript *corr* stands for the correlated systematics. We then weight each result (C_{f+sys}^{NN} and C_{f+sys}^{TK}) to construct a value for the combined method (C_{f+sys}^{COM}):

$$\begin{aligned}
C_{f+sys}^{COM} &= w_{NN} * C_{f+sys}^{NN} + (1 - w_{NN}) * C_{f+sys}^{TK}, & (4)\\
\text{where} & \\
w_{NN} &= \frac{\frac{1}{\sigma_{NN}^2}}{(\frac{1}{\sigma_{NN}^2} + \frac{1}{\sigma_{TK}^2})}, & (5)\\
\text{andwhere} & \\
\sigma_{NN} &= \sqrt{(\sigma_{NN}^{stat})^2 + (\Delta_{uncorr}^{NN})^2 + (\Delta_{corr}^{NN})^2}. & (6)
\end{aligned}$$

We find that the track analysis contribution to the combined measurment is $\sim 90\%$ and the NN analysis contributes $\sim 10\%$.

V. RESULTS

The fit value we measure in the data when we combine the two analyses is 0.073. At 68% confidence level this corresponds to a purely statistical result of $C_f^{true} = 0.07_{-0.07}^{+0.13}$, and to a result of $C_f^{true} = 0.07_{-0.07}^{+0.15}$ when systematic uncertainties are taken into account. This can be seen in Figure 3.

Figure 4 shows the limits, including systematic uncertainties, for bands of 68% and 95% confidence level. At the 95% confidence level we find $C_f^{true} < 0.38$. We find that combining the two measurements we obtain an *a priori* improvement between 5-12% depending on the data result.

VI. CONCLUSIONS

We present a measurement of the relative fraction of $t\bar{t}$ events produced via gluon-fusion to the total number of $t\bar{t}$ events using a combine method based on low p_T tracks and Neural Networks probabilities. Using a total integrated luminosity of 955 pb^{-1} we find $C_f^{true} = 0.07_{-0.07}^{+0.15}$.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the

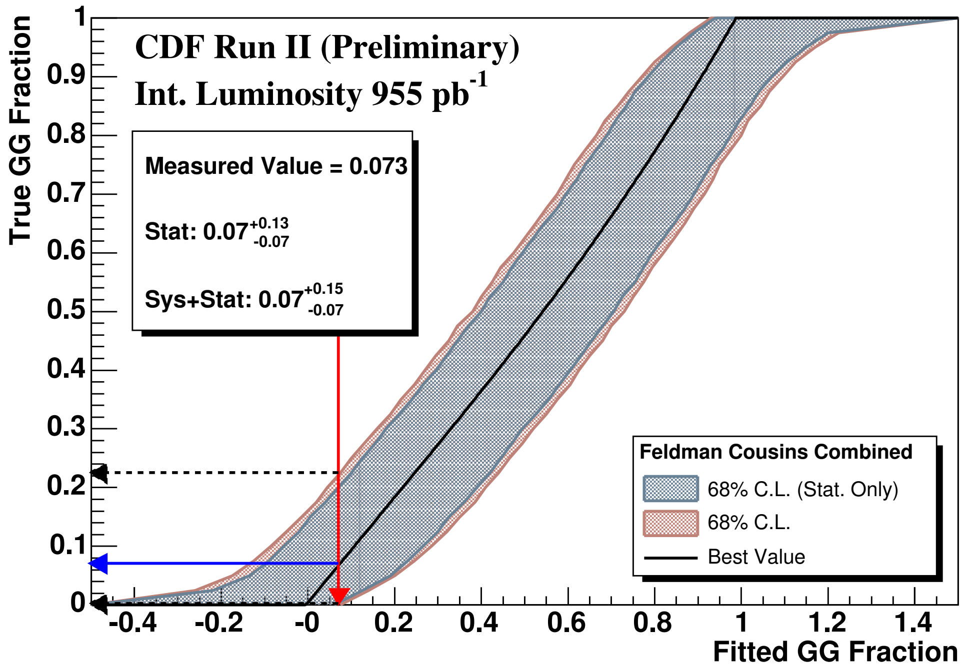

FIG. 3: Feldman-Cousins confidence band with all systematics compared to FC band with statistical only for a confidence level of 68%.

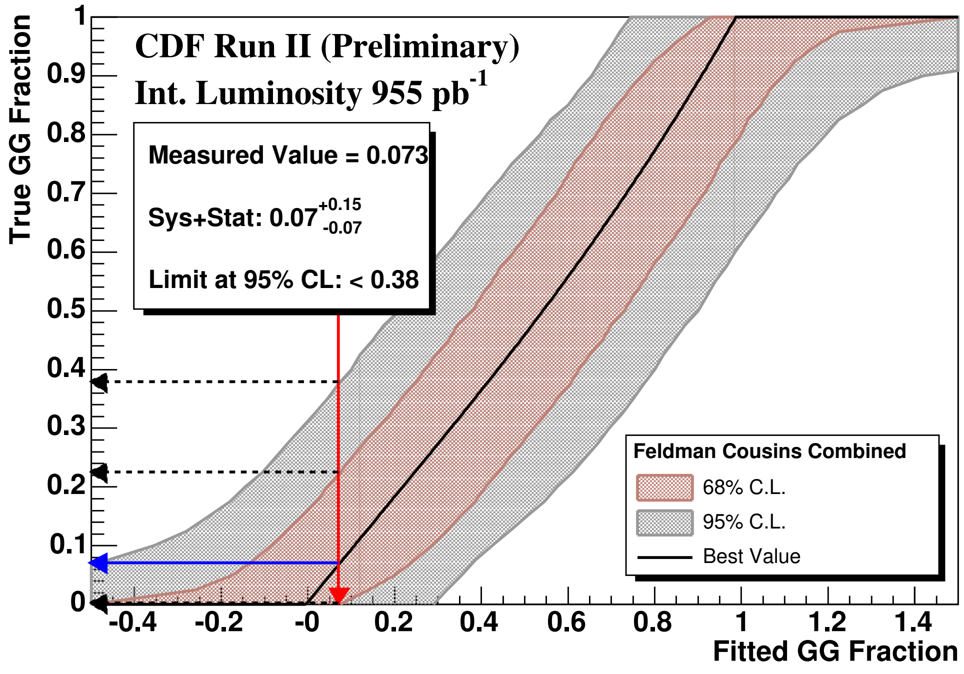

FIG. 4: Final FC with all systematics.

Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

[1] M. Cacciari, *et al.*, "The t anti-t cross-section at 1.8-TeV and 1.96-TeV: A Study of the systematics due to parton densities and scale dependence", hep-ph/0303085.
[2] N. Kidonakis, *et al.*, "Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction", Phys. Rev. D 68 114014 (2003).
[3] CDF Collaboration, "First Measurement of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$, CDF public note 8724.
[4] CDF Collaboration, "Measurement of the relative fraction of ttbar events produced via gluon-gluon fusion in 955 pb^{-1}, CDF public note 8811.
[5] G. Feldman and R. Cousins, Phys. Rev. D57, 3873 (1998).

APPENDIX A: FELDMAN-COUSINS PLOTS FOR THE DIFFERENT METHODS.

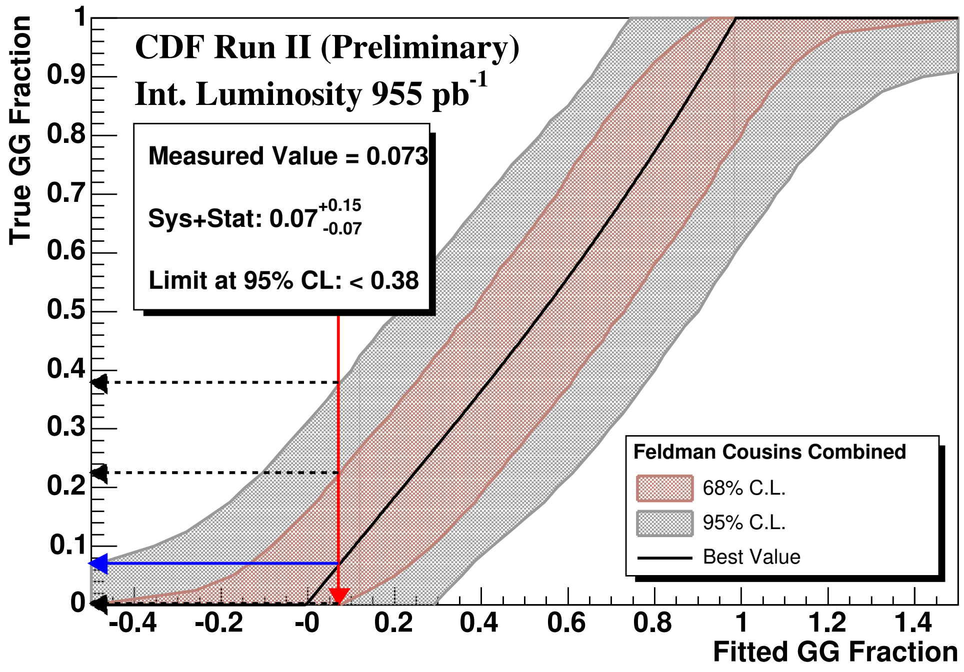

FIG. 5: Final Feldman-Cousins plot of the combination. It shows the statistical+systematic 68% and 95% bands, the central value and its limits.

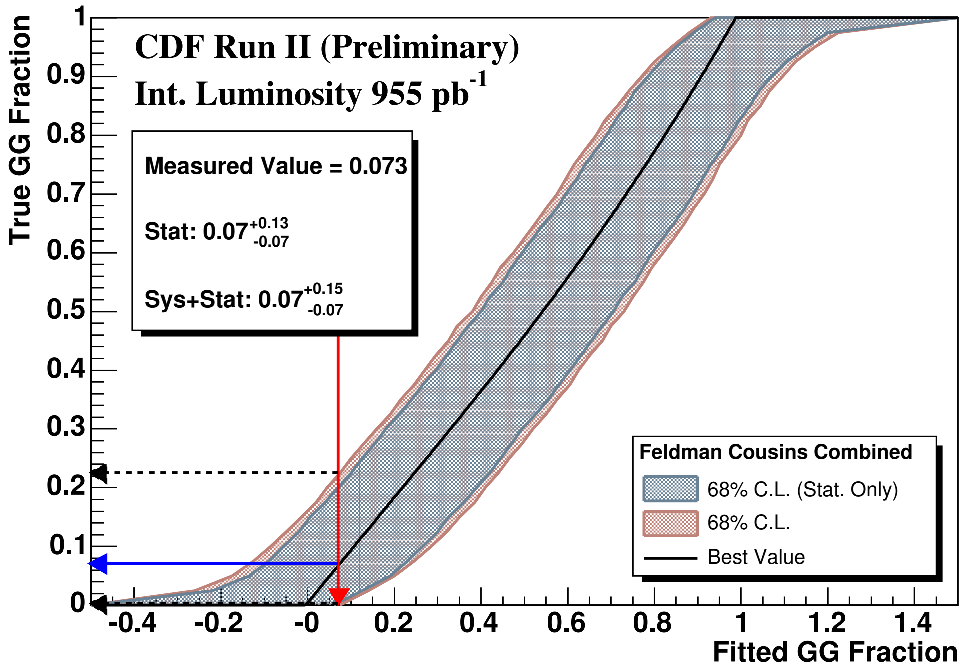

FIG. 6: Feldman-Cousins plot of the combination showing the statistical-only as well as the statistical+systematic band at the 68%.

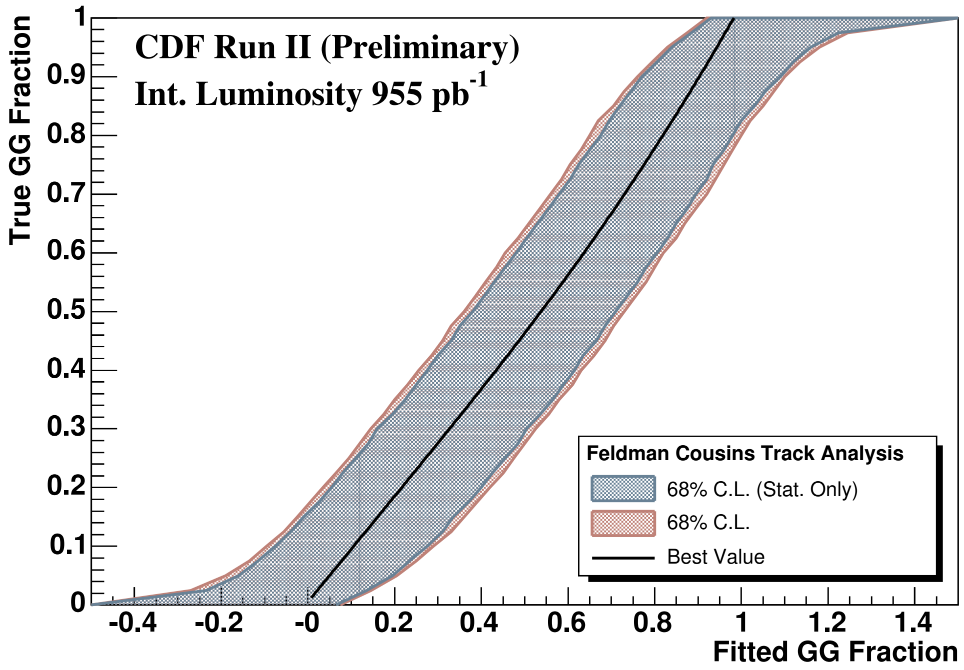

FIG. 7: Feldman-Cousins plot obtained for the track analysis alone and shows the statistical-only as well as the statistical+systematic band at the 68%.

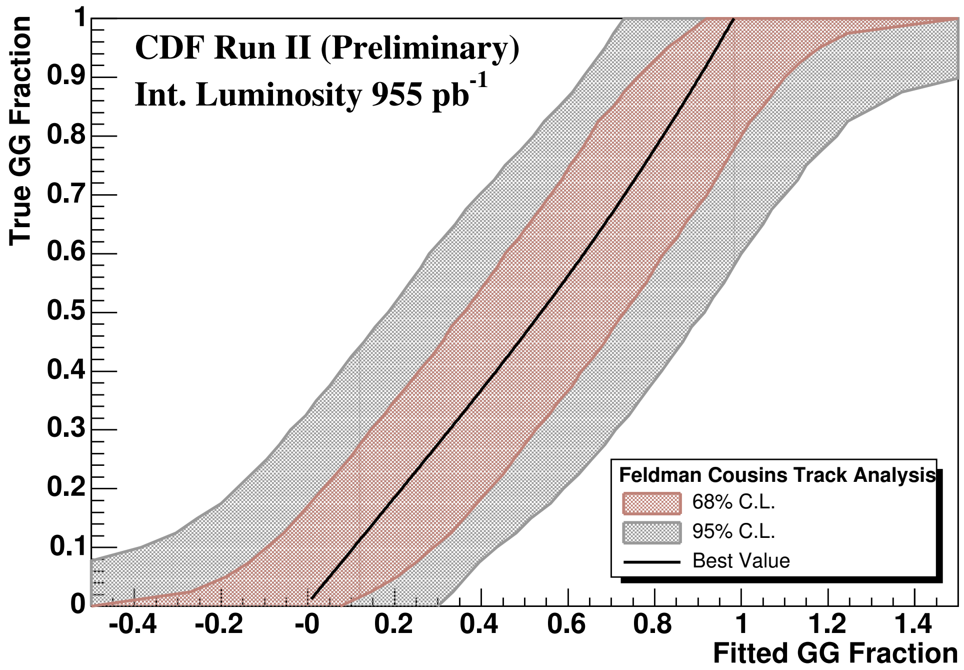

FIG. 8: Feldman-Cousins plot obtained for the track analysis alone, and shows the statistical+systematic 68% and 95% bands.