Neural Network based Jet Charge Tagger

Text for the blessed web page – CDF note 7482

The CDF Collaboration

July 11, 2005

Abstract

We describe a new version of Jet Charge algorithm for B flavor tagging. We introduce Neural Networks for track and jet probability. We develop a new selection algorithm based on jet probability. Furthermore we make use of the innermost layer of the silicon detector. We measure $\epsilon D^2 = 0.917 \pm 0.031\%$ on the e+SVT sample and $\epsilon D^2 = 0.938 \pm 0.029\%$ on the μ+SVT sample of 355 pb^{-1}.

1 Description

We document the performance of the Neural Network based Jet Charge Tagger on a sample of data containing b-hadrons with semileptonic decays. The events are triggered by a lepton with transverse momentum p_T larger than 4 GeV and a displaced track with $p_T > 2$ GeV (*SVT track*). We construct a sample with close to 100% $b\bar{b}$ purity by applying a cut on the invariant mass of the lepton and the displaced track and by performing a a background subtraction using the signed impact parameter of the displaced trigger track.

The Jet Charge Tagger is one of the tagging algorithms used in the B mixing analyses. Flavor Taggers are used to identify the type of the b-hadrons when they are produced. In this case the b-hadrons that are studied for mixing fire the CDF trigger and the tracks used for the Jet Charge flavor tag are associated with the second b quark produced in the event.

This analysis uses Monte Carlo samples to study the properties of $b\bar{b}$ events and optimize the performance of the tagger. For this procedure to be successful the physics and detector performance of the Monte Carlo must reproduce the performance seen in real data. We demonstrate the accuracy of our Monte Carlo by studying a large number of critical distributions to show the high level of data and Monte Carlo agreement. The Monte Carlo sample used is generated with Pythia [1] and includes all $b\bar{b}$ production processes. We are able to extract the b-jet properties from the Monte Carlo and use them for the jet selection in the data.

We introduce Neural Network to tag the B decay products with a track probability network. We define several jet variables based on the track probability. The jet variables are combined with a Neural Network as well to give a b-jet probability.

The track probability Neural Network (*trackNet*) has been calibrated on a ℓ+SVT Monte Carlo sample and estimates how likely it is that the track is a B decay product. The optimization has been performed with the NeuroBayes® [2] package. The set of input variables chosen for the trackNet is:

- track impact parameter magnitude, $|d_0|$
- signed impact parameter significance $d_0^{signed}/\sigma_{d_0}$, where $d_0^{signed} = |d_0| \cdot sign(\vec{d_0} \cdot \vec{p}_{jet})$
- track transverse momentum p_T
- track was used for VxPrim vertex fit (flag equal to 0 if the track was not used, 1 if the track was used, 2 if the primary vertex was not found in the event)
- cone angle $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ of track w.r.t. B candidate
- jet p_T
- rapidity with respect to jet axis $y = \frac{1}{2}\ln\frac{E+p_L}{E-p_L}$, where p_L is the projection of the track momentum on the jet axis, $p_L = \frac{\vec{p}_{track}\cdot\vec{p}_{jet}}{|\vec{p}_{jet}|}$

The tracks are reconstructed using also the hits in the closest silicon layer to the interaction point, layer 00 (*L00*). The resolution on $|d_0|$ and $d_0^{signed}/\sigma_{d_0}$ is influenced by the presence of

L00 hit on the track. The effect is taken into account by optimizing the trackNet separately for tracks with and without L00 hits. Tracks with a L00 hit have a larger spread of trackNet outputs, due to the higher resolution on $d_0^{signed}/\sigma_{d_0}$.

The jets on the opposite side are reconstructed with a simple fixed cone algorithm ($\Delta R = 1.5$) which uses tracks only. The jet probability Neural Network, *bJetNet*, is optimized on a flag that tells if the jet is a b-jet or not. The list of input variables for the bJetNet is:

- number of tracks with probability greater than 50%, N_{50}
- jet p_T
- Secondary vertex fit χ^2 probability, $P(\chi^2_{SV})$
- number of tracks in the secondary vertex, N_{SV}
- weighted number of tracks in jet $\sum_i t_i$, where the index i runs over all the tracks in the jet and t_i is the trackNet output for track i
- Jet probability J_P, where the jet probability is the likelihood combination of the probability of the tracks in the jet and it is computed as

$$J_P = P(N) \cdot \sum_{n=1}^{N} \frac{(-logP(N))^n}{n!} \tag{1}$$

$$P(N) = \prod_{m=1}^{N} t_m \tag{2}$$

- log of $J_P^{p_T}$, where $J_P^{p_T}$ is computed as in 1 and only tracks with $p_T > 1$ GeV are considered.
- Highest probability for the tracks in jet, t^{max}
- $\Delta\phi(jet, B)$
- transverse momentum of the track with the maximum momentum in the jet p_T^{max}
- $\sum_i p_{T,i}^{rel}/p_T^{jet}$, where $p_{T,i}^{rel}$ is the magnitude of the track momentum component perpendicular to the jet axis
- $\sum_i p_{L,i}$, where $p_{L,i}$ is the projection of the track momentum along the jet axis
- Neural Network tag based on invariant mass, T_{nn}
- number of tracks, N_{tracks}
- jet invariant mass, m_{jet}
- jet spread, $\frac{\sum_i cos\alpha_i}{N_{tracks}}$, where $cos\alpha_i = \frac{\vec{p}_i \cdot \vec{p}_{jet}}{|\vec{p}_i| \cdot |\vec{p}_{jet}|}$

The jet with the highest probability is selected as the tagging jet. The opposite side flavor is given by the sign of the jet charge Q_{jet} computed for the tagging jet as:

$$Q_{jet} = \frac{\sum_i Q_i \cdot p_{T,i} \cdot (1 + t_i)}{\sum_i p_{T,i} \cdot (1 + t_i)} \tag{3}$$

where the index i runs over all the tracks in the jet, Q_i is the track charge, $p_{T,i}$ is the track transverse momentum and t_i is the track probability.

The flavour on the trigger side corresponds to the charge of the trigger lepton. The dilution is scaled to take into account mixing on the same side and sequential semileptonic decays. In order to compute ϵD^2 we split the tagging jet sample in subsamples (3 bins of jet quality and 11 bins of $|Q_{jet}| \cdot P_{nn}$, where P_{nn} is the jet probability). The jet quality types are:

- **Class 1:** Jets with a secondary vertex that has a decay length significance $L_{xy}/\sigma_{L_{xy}}$ greater than 3
- **Class 2:** Jets not in Class 1 and with at least one track with probability greater than 50%
- **Class 3:** Jets not in Class 1 and with no tracks with probability greater than 50%

The classes are mutually exclusive. The total ϵD^2 is the sum of the partial tagging power evaluated in each bin. The *effective dilution* D is equal to $\sqrt{\epsilon D^2/\epsilon}$.

2 Blessed Numbers

This section summarizes the numbers which have been blessed for this analysis. The errors are statistical only.

Jet type	ϵ,%	effective D, %	ϵD^2
e+SVT data			
Class 1	10.56 ± 0.05	19.75 ± 0.43	0.412 ± 0.018
Class 2	28.20 ± 0.08	11.60 ± 0.33	0.380 ± 0.021
Class 3	56.72 ± 0.11	4.70 ± 0.29	0.125 ± 0.015
combined	95.48 ± 0.15	9.80 ± 0.16	0.917 ± 0.031
μ+SVT data			
Class 1	10.51 ± 0.04	18.35 ± 0.37	0.354 ± 0.014
Class 2	28.76 ± 0.07	12.18 ± 0.29	0.426 ± 0.020
Class 3	56.45 ± 0.09	5.29 ± 0.26	0.158 ± 0.016
combined	95.72 ± 0.12	9.90 ± 0.15	0.938 ± 0.029

3 Blessed Figures

The following figures have been blessed for the Neural Network based Jet Charge Tagger.

We define *signal tracks* to be all the B daughters, including the decay products of unstable B daughters. We define *background tracks* to be all the other tracks in the event.

The jet in the event with a high content of B daughters is the *signal jet*. All the jets that are not signal are *background jets*. More text can be found in CDF note 7482.

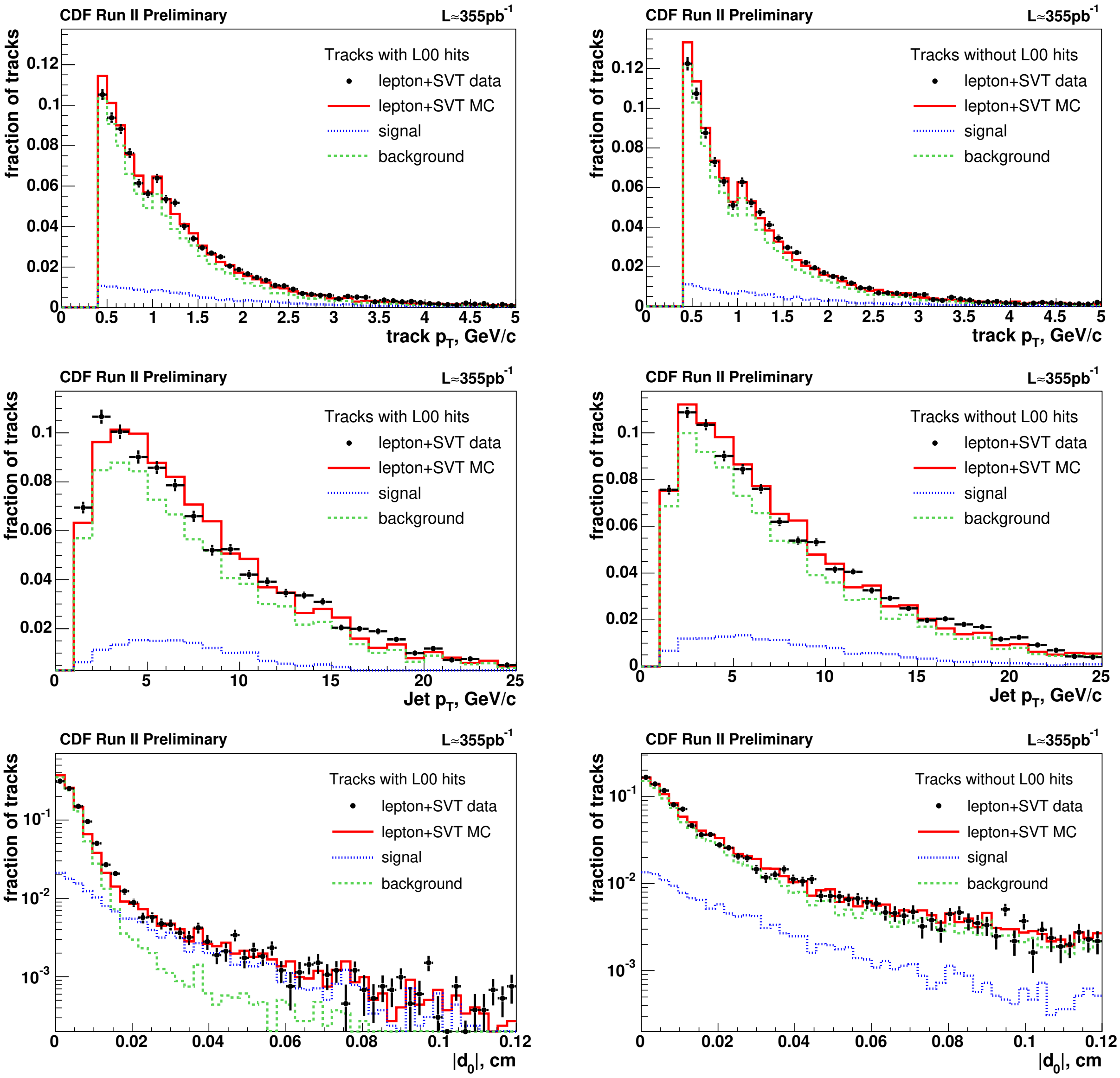

Figure 1: **Input variables for the track probability Neural Network** for tracks with a hit in the silicon layer closest to the interaction point (*left hand side*) and tracks without (*right hand side*). *Top row:* track transverse momentum. The peak at 1 GeV is given by the requirement of a minimum p_T of 1 GeV to seed jets. *Center row:* transverse momentum of the jet to which the track is assigned. *Bottom row:* track impact parameter.

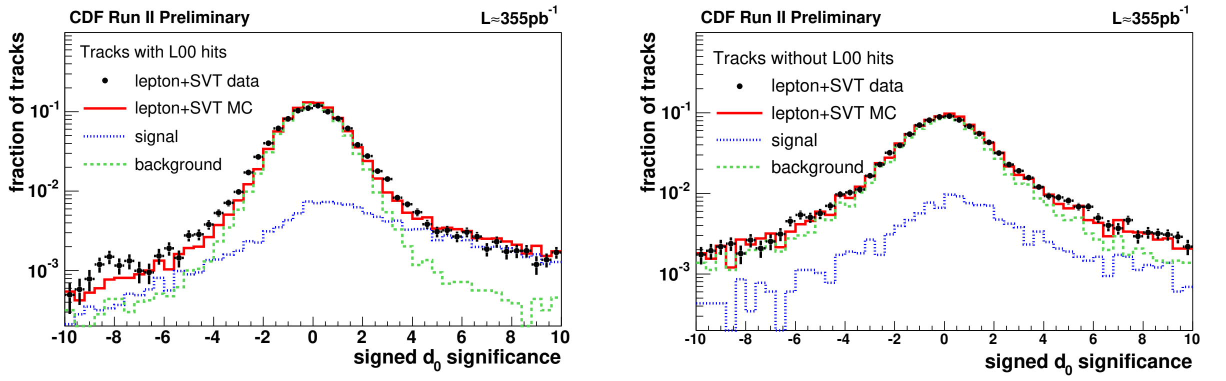

Figure 2: **Input variable for the track probability Neural Network:** Signed impact parameter significance for tracks with a hit in the silicon layer closest to the interaction point (*left*) and tracks without (*right*). The impact parameter is lifetime signed with respect to the jet momentum.

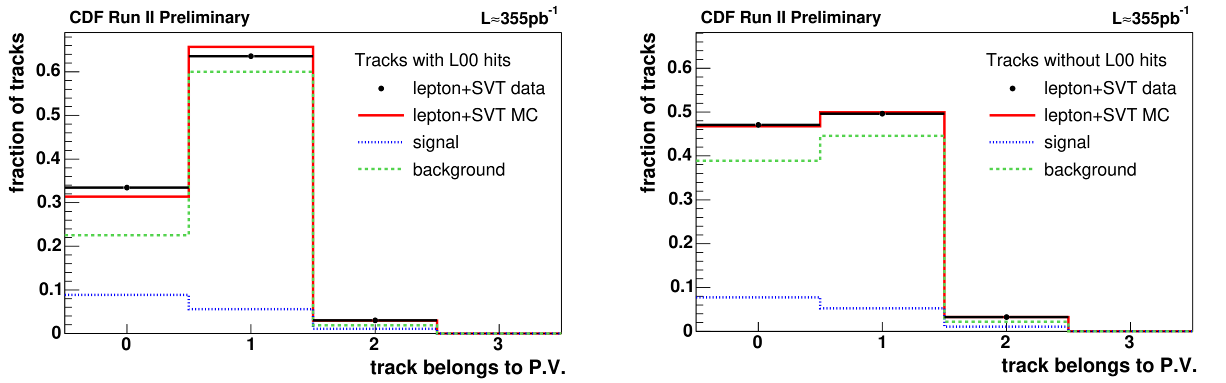

Figure 3: **Input variable for the track probability Neural Network:** Flag that tells if the track was used to fit the primary vertex (PV) of the event for tracks with a hit in the silicon layer closest to the interaction point (*left*) and tracks without (*right*). The value 0 means that the track was not used for the PV; the value 1. means that it was used and 2. represents the cases in which the PV was not found.

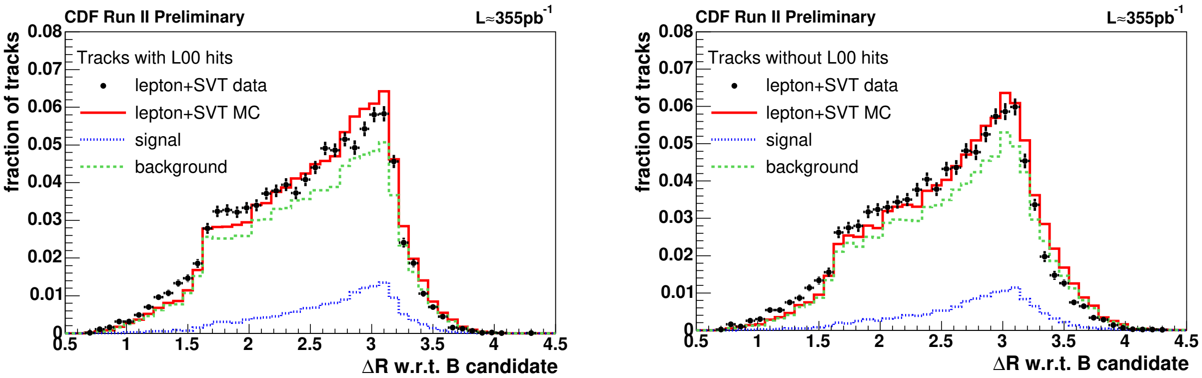

Figure 4: **Input variable for the track probability Neural Network:** cone angle $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the track and the B candidate direction, identified as the sum of the trigger lepton and displaced track momenta, for tracks with a hit in the silicon layer closest to the interaction point (*left*) and tracks without (*right*).

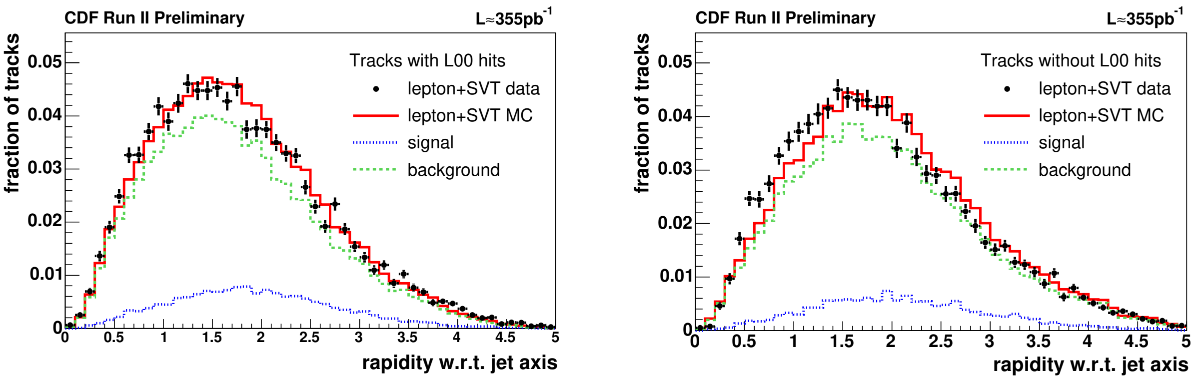

Figure 5: **Input variable for the track probability Neural Network:** rapidity of the track with respect to the jet axis for tracks with a hit in the silicon layer closest to the interaction point (*left*) and tracks without (*right*). The rapidity is defined as $y = \frac{1}{2} \ln \frac{E+p_L}{E-p_L}$, where p_L is the longitudinal component of the track with respect to the jet axis, $p_L = \frac{\vec{p}_{track} \cdot \vec{p}_{jet}}{|\vec{p}_{jet}|}$

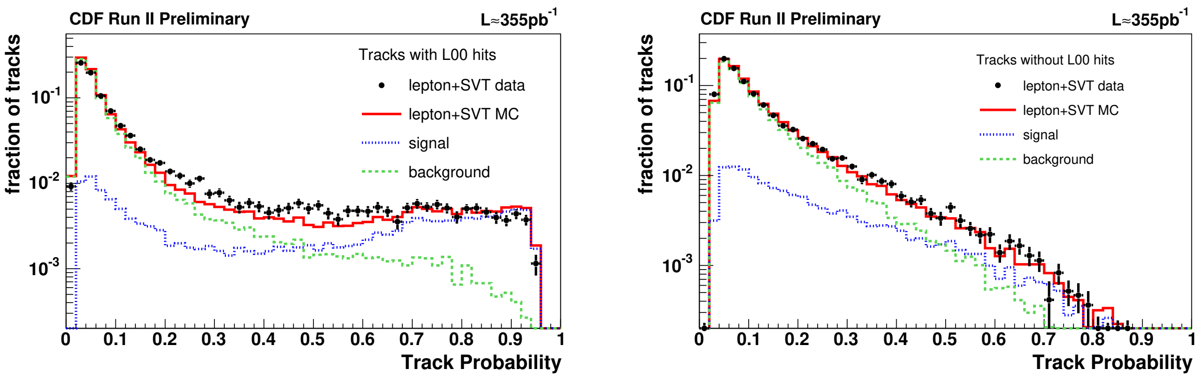

Figure 6: Output of the track probability Neural Network for tracks with a hit in the silicon layer closest to the interaction point (*left*) and tracks without (*right*).

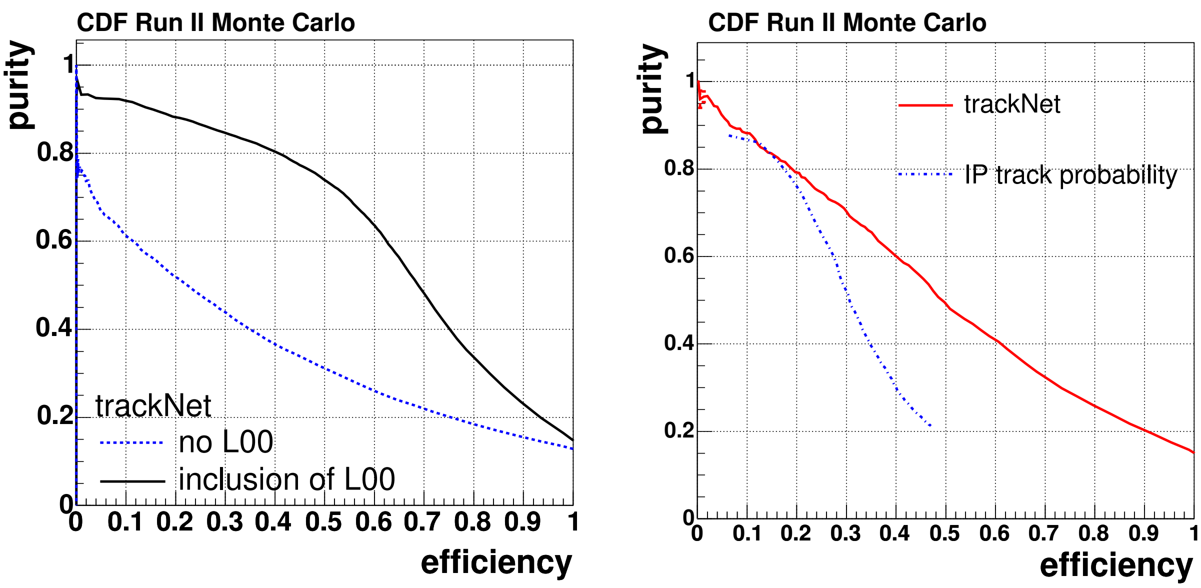

Figure 7: *Left:* Performance plot for the track probability Neural Network for tracks with and without L00 hits. Each point in the graph corresponds to a cut on the network output. The performance is better for tracks with L00 hits thanks to the higher resolution on $|d_0|$and $d_0^{signed}/\sigma_{d_0}$. *Right:* comparison between the performance of a cut on the track probability Neural Network output and a cut on the track probability based on displacement information only. Tracks with and without L00 hits are not distinguished in this case, meaning that the solid line is equivalent to a weighted sum of the curves in the left graph.

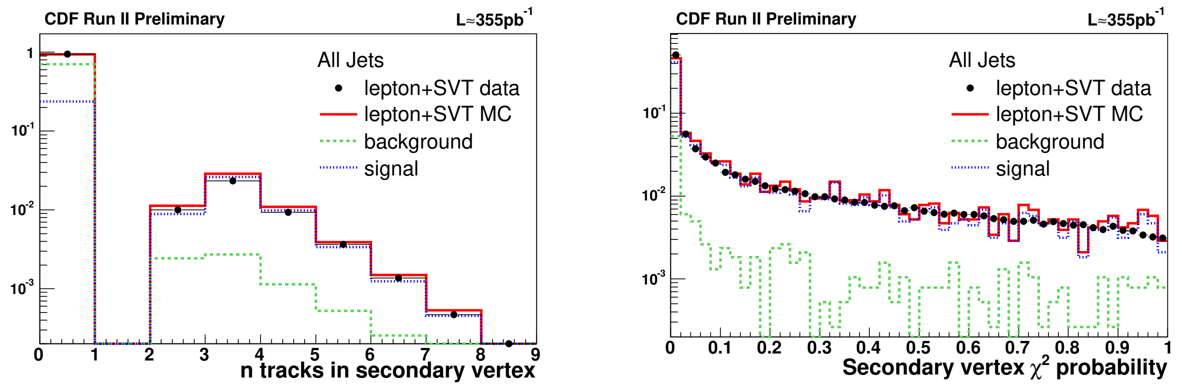

Figure 8: **Input variables for the jet probability Neural Network:** *Left:* Number of tracks in the secondary vertex. The peak at . is relative to jets for which no secondary vertex is found. *Right:* Probability of the χ^2 fit of the secondary vertex.

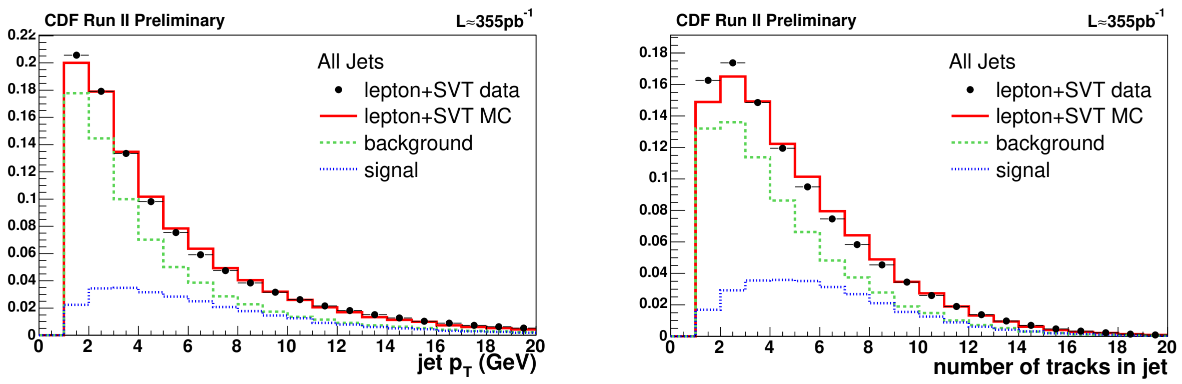

Figure 9: **Input variables for the jet probability Neural Network:** *Left:* transverse momentum of the jet. *Right:* number of tracks in the jet.

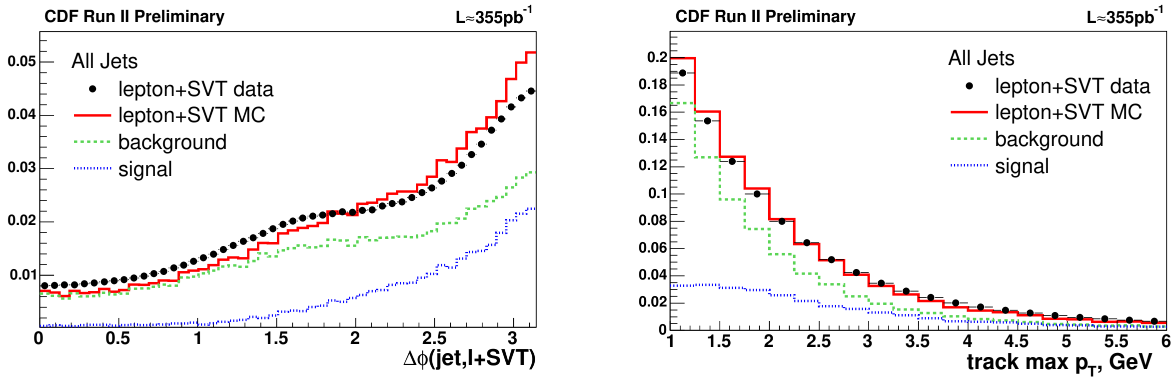

Figure 10: **Input variables for the jet probability Neural Network:** *Left:* angle between the jet and the direction of the B candidate, identified by the trigger lepton and the displaced track, in the plane transverse to the beam line. *Right:* Transverse momentum of the track with the maximum p_T in the jet.

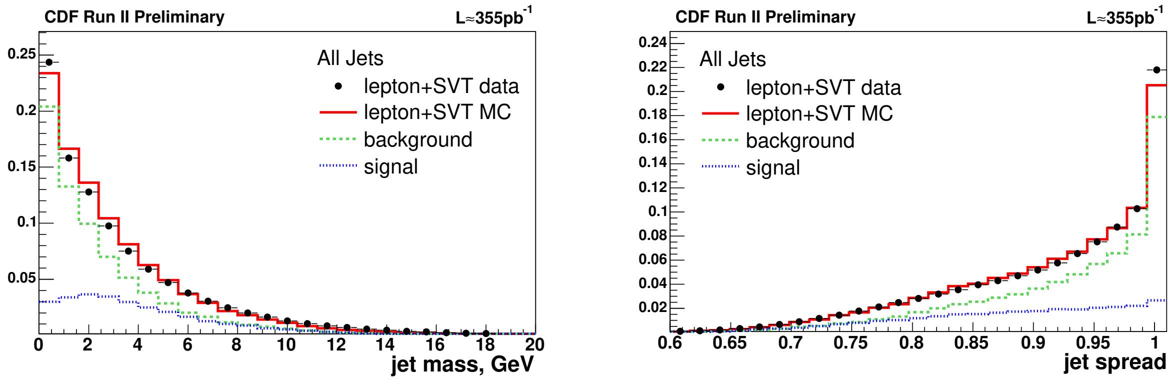

Figure 11: **Input variables for the jet probability Neural Network:** *Left:* invariant mass of the jet *Right:* jet spread, defined as the sum of the angles in space between the tracks and the jet axis. This quantity gives information on how collimated the jet is. The peak at 1. is relative to jets constituted by a single track.

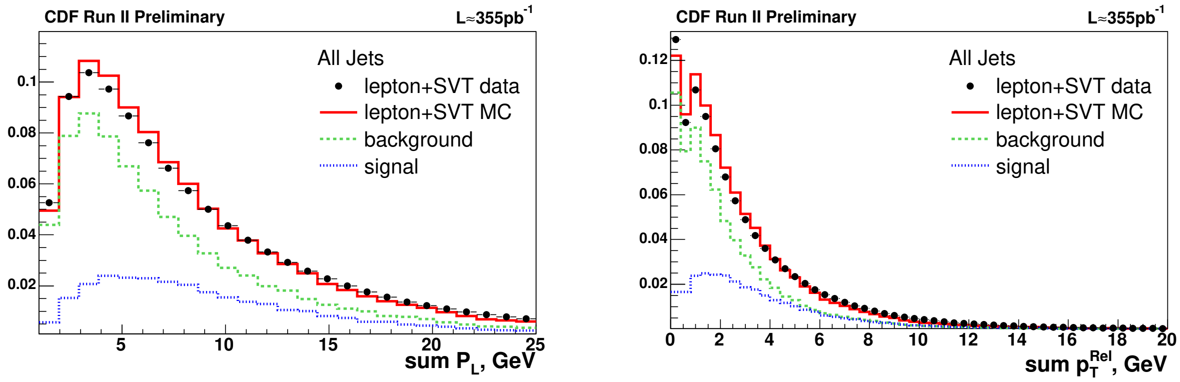

Figure 12: **Input variables for the jet probability Neural Network:** *Left:* sum of the projections of the track moments along the jet axis. *Right:* sum of the track momentum components perpendicular to the jet axis.

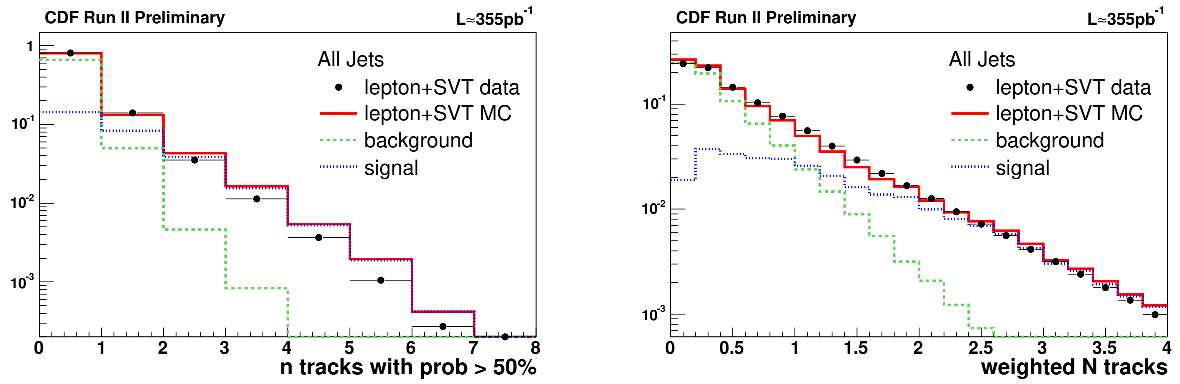

Figure 13: **Input variables for the jet probability Neural Network:** *Left:* number of tracks in the jet that have a probability greater than 50% to be a b-hadron decay product. *Right:* sum of the probabilities of the tracks in the jet.

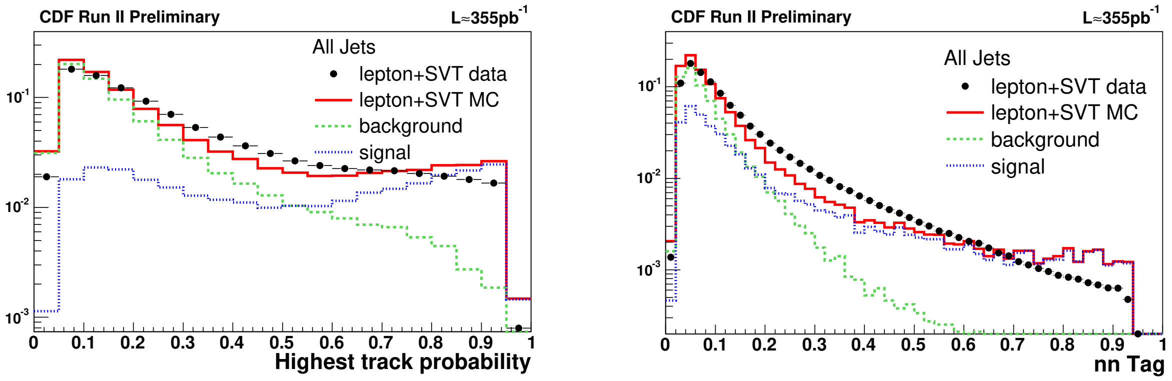

Figure 14: **Input variables for the jet probability Neural Network:** *Left:* probability of the track with the highest probability in the jet. *Right:* Neural Network tag based on invariant mass, T_{nn}. In order to compute this variable, the invariant mass of the tracks with the highest probability is computed. The tracks are progressively summed up until the invariant mass becomes larger than the D^0 mass. The tag variable is then the probability of the track that pushes the invariant mass value above the D^0 mass.

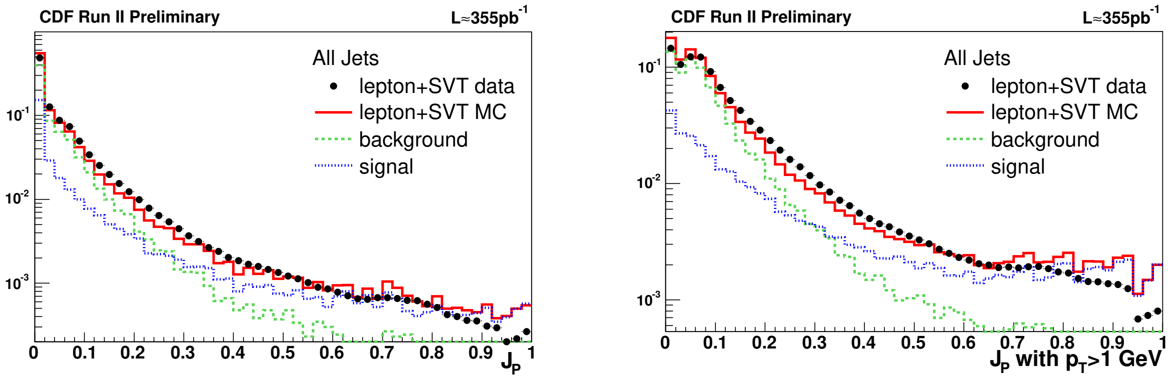

Figure 15: **Input variables for the jet probability Neural Network:** *Left:* likelihood combination of the probabilities of the tracks in the jet. *Right:* likelihood combination of the probabilities of the tracks with $p_T > 1$ GeVin the jet.

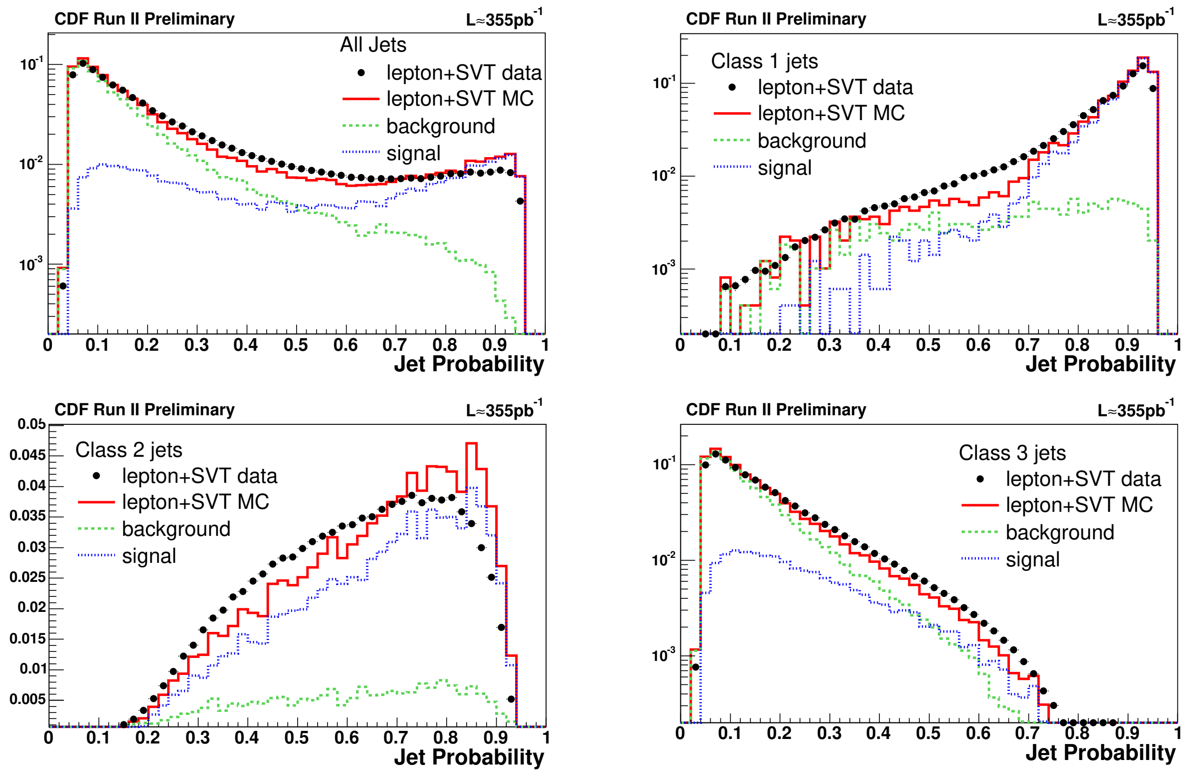

Figure 16: Output of the jet probability Neural Network for all jets (*top left*), Class 1 jets (*top right*), Class 2 jets (*bottom left*) and Class 3 jets (*bottom right*).

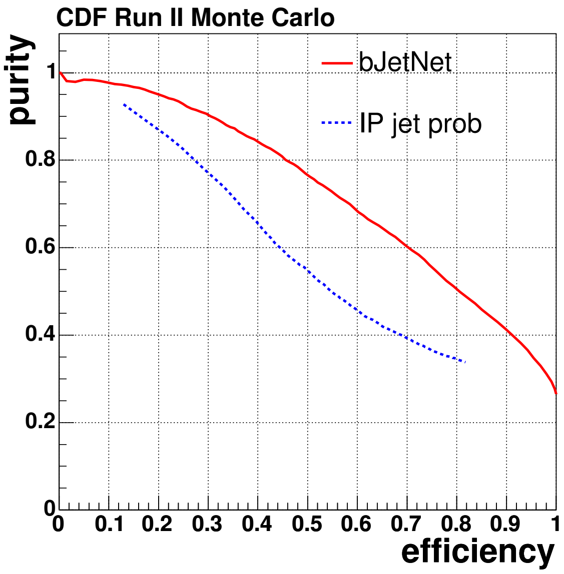

Figure 17: Performance graph for a cut on the jet Neural Network output (*solid line*) in simulated events. The comparison with the jet probability obtained by combining with a likelihood the track probability based only on impact parameter is displayed (*dashed line*).

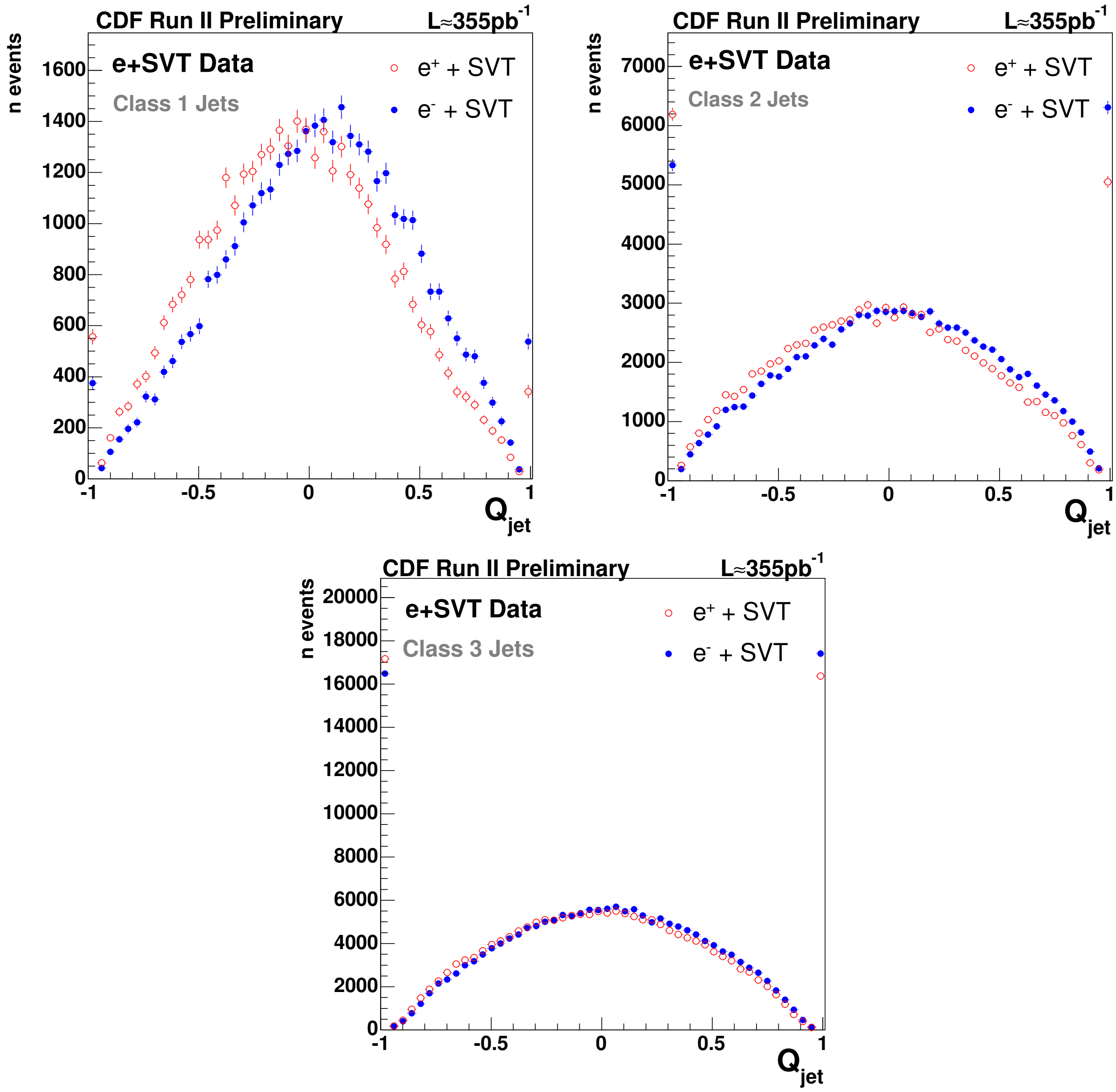

Figure 18: Jet charge for the different types of jet for the e+SVT sample, shown separately for positive and negative electrons. The shift between the two distributions is more evident for Class 1 jets.

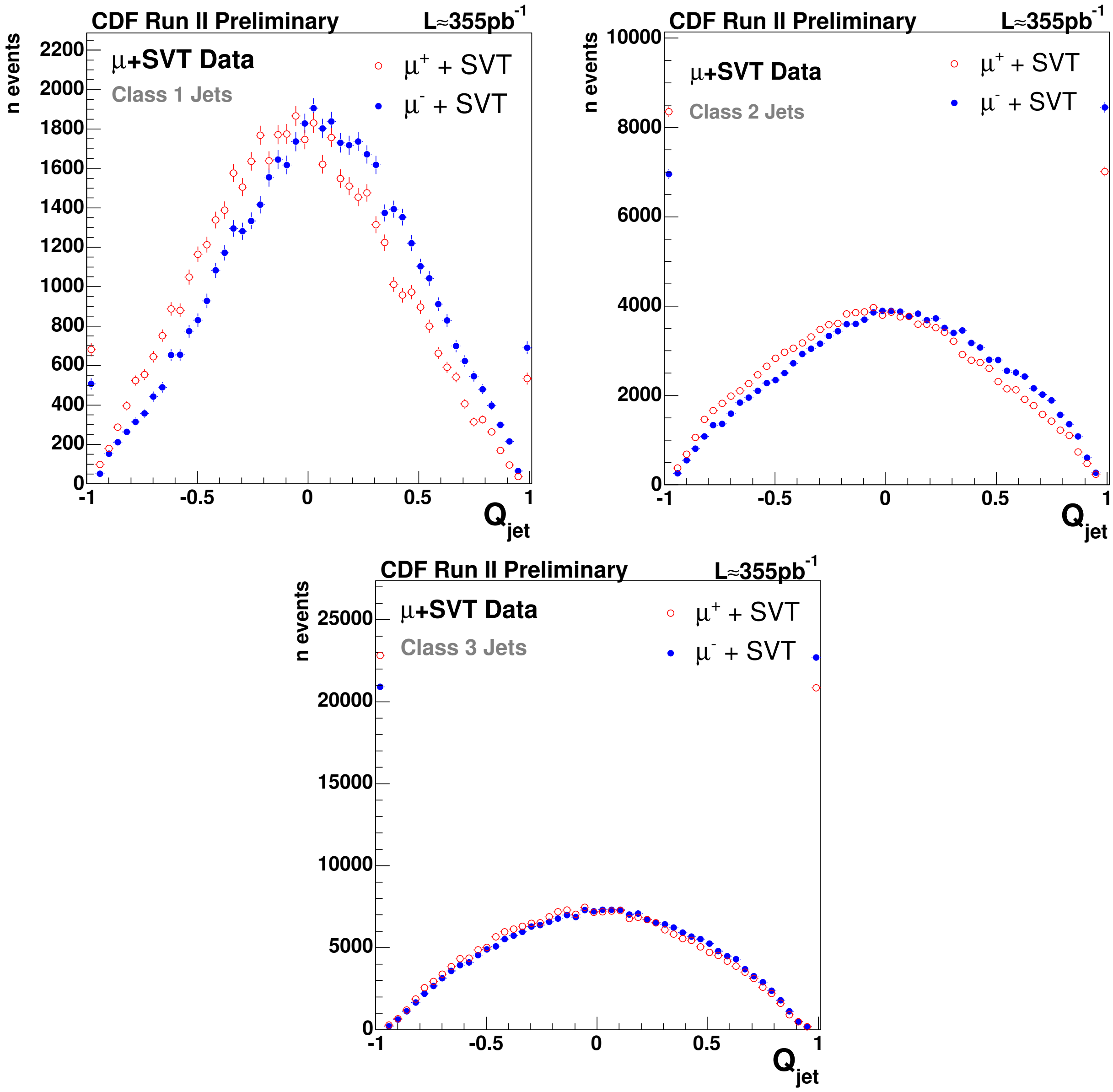

Figure 19: Jet charge for the different types of jet for the μ+SVT sample, shown separately for positive and negative muons. The shift between the two distributions is more evident for Class 1 jets.

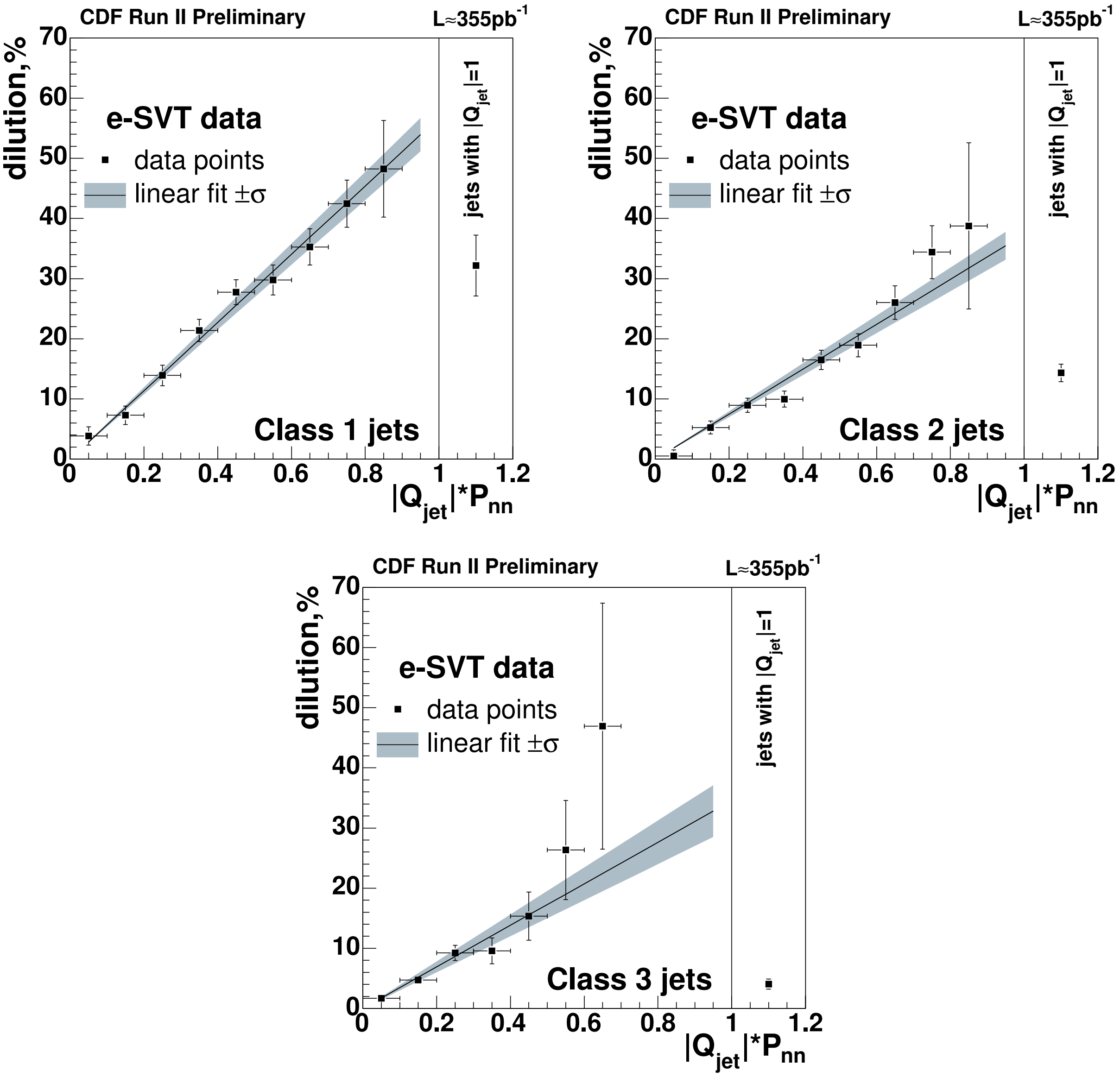

Figure 20: Dependency of dilution on $|Q_{jet}|P_{nn}$, where Q_{jet} is the jet charge and P_{nn} is the jet probability. The plots are relative to the e+SVT sample.

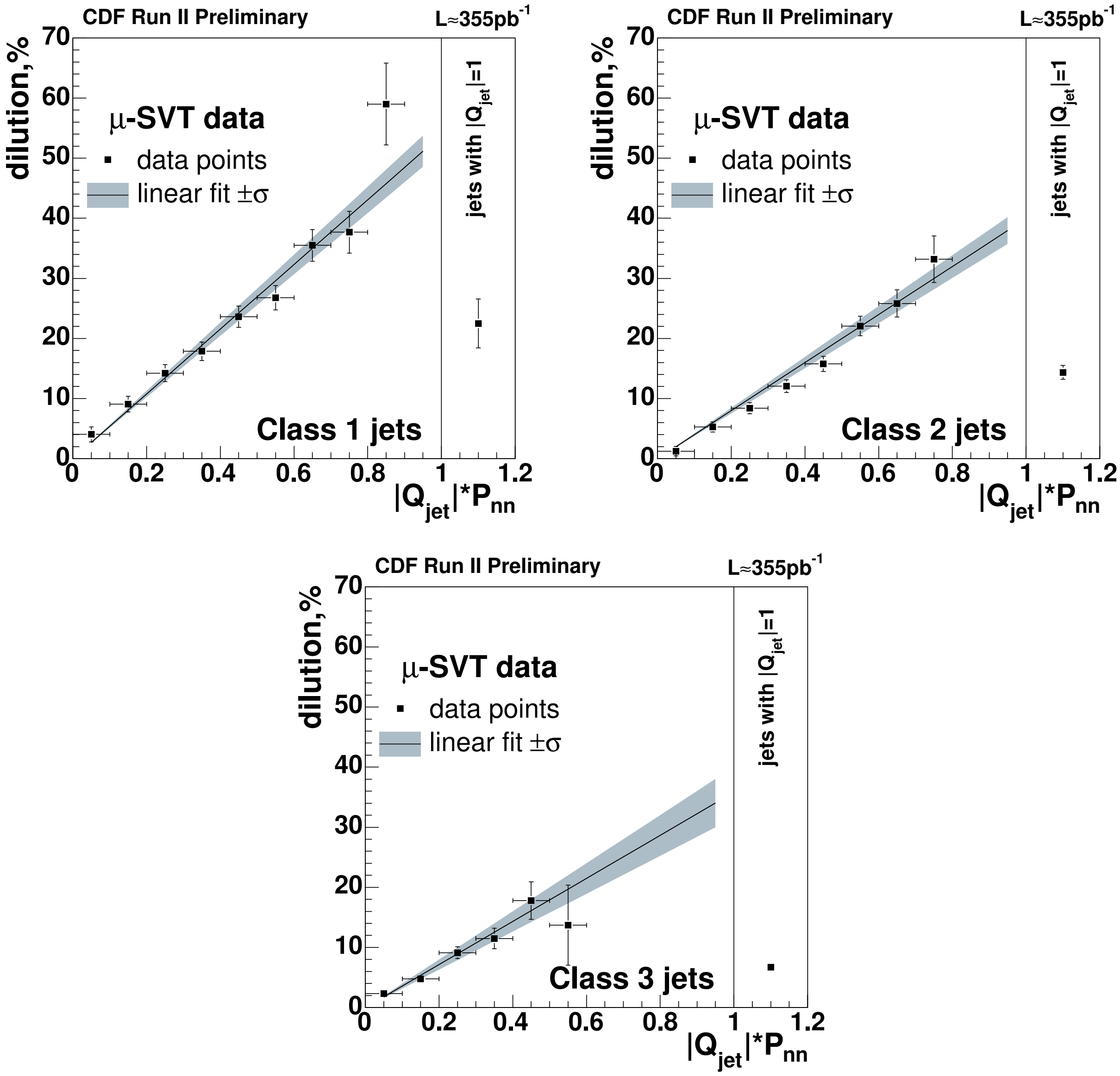

Figure 21: Dependency of dilution on $|Q_{jet}|P_{nn}$, where Q_{jet} is the jet charge and P_{nn} is the jet probability. The plots are relative to the μ+SVT sample

References

[1] T. Sjöstrand, Comput. Phys. Commun. **82** (1994) 74

[2] M. Feindt, *"A Neural Bayesian Estimator for Conditional Probability Densities"*, physics/0402093