Resonance search in

$p\bar{p} \rightarrow X^0 \rightarrow t\bar{t}$

Yuri Oksuzian
University of Florida

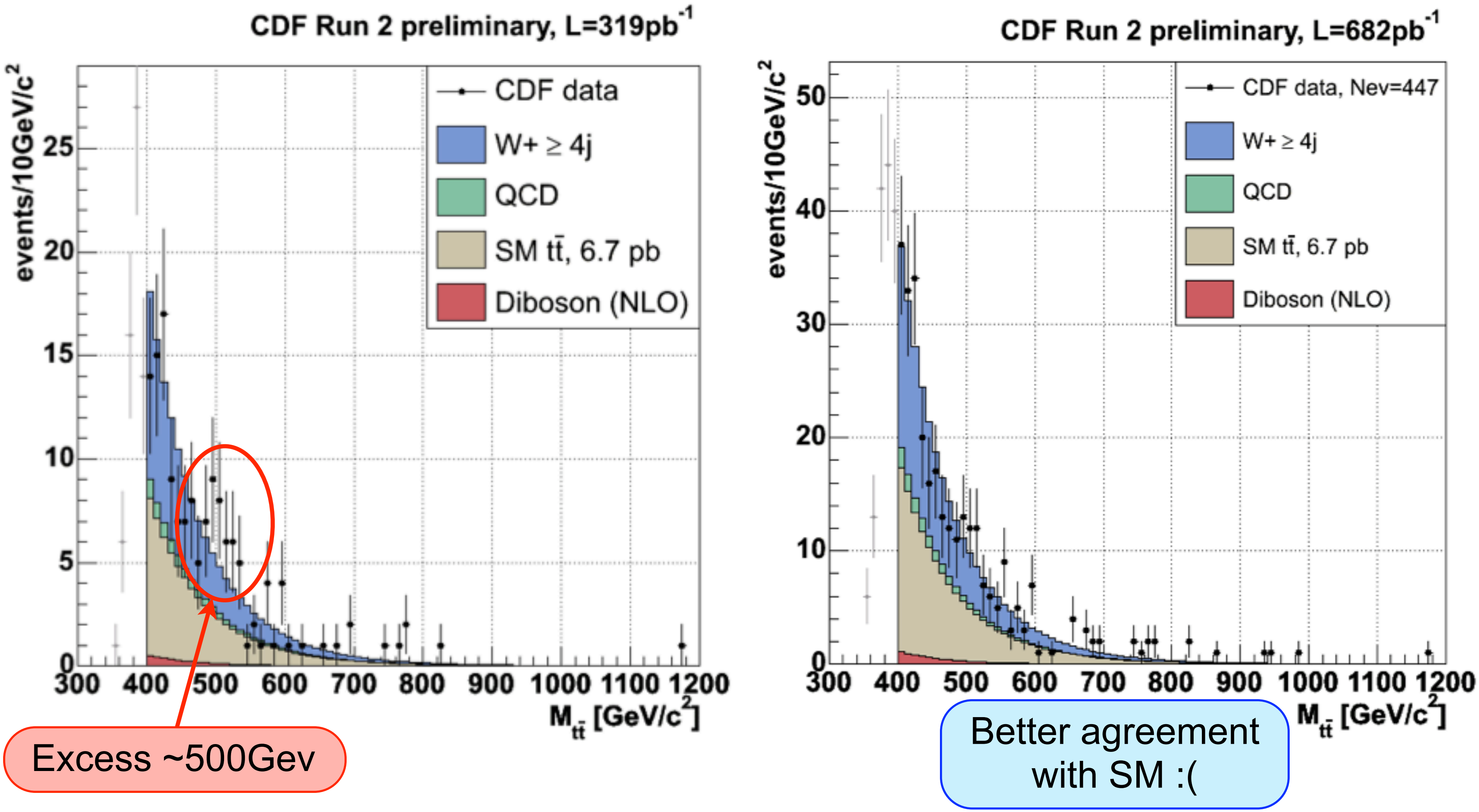
Why and How?

- Goal is to test $t\bar{t}$ production for possible new sources such as a narrow resonance
  - Top is very heavy, maybe indication of coupling to new physics
  - Top is a young particle
  - Various theoretical models predict it: technicolor, KK gluons
- Search technique:
  - M_{tt} spectrum is reconstructed, using FlaME
  - Search for a peak in M_{tt} spectrum
    - Understand SM fluctuation probabilities
    - Calculate UL(Upper Limits)
    - Compare data with our expectations(SM or with new physics)

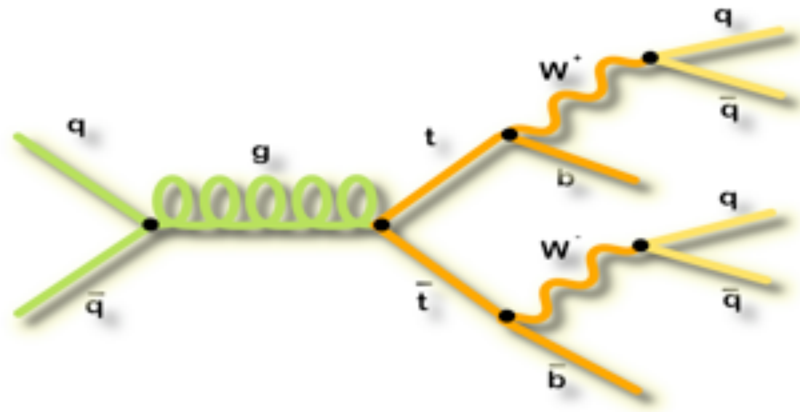
Where?

- This is the first M_{tt} analyses in All Hadronic channel
  - Disadvantages
    - Large QCD background
      - Controlled with good event selection
    - More combinations
  - Advantages
    - Highest branching ratio
      - Most $t\bar{t}$ events are here
    - No missing information like neutrino
      - Better signal templates
  - Future
    - Combined result with lepton+jets channel
      - Higher sensitivity
    - Cross-check for a possible discovery

Motivation - previous results



FlaME (Florida Matrix Element)



We calculate the *a priori* probability density for an event to be the result of Standard Model $t\bar{t}$ production and decay

$$P(j \mid M_{top}) = \frac{1}{\sigma(M_{top})\varepsilon(M_{top})N_{combi}} \sum_{combi} \int dz_b dz_b f(z_a) f(z_b) d\sigma(M_{top}, p) TF(j \mid p) P_T(p)$$

To calculate the M_{tt} probability density, we modify the integral above:

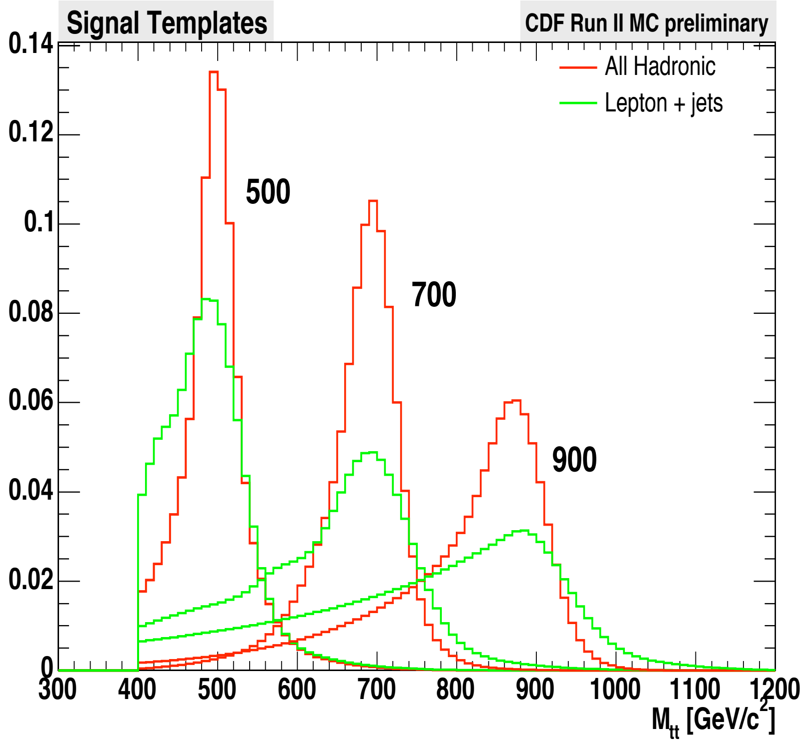
$$\rho(x \mid j) = \frac{1}{\sigma(M_{top})\varepsilon(M_{top})N_{combi}} \sum_{combi} \int dz_b dz_b f(z_a) f(z_b) d\sigma(M_{top}, p) TF(j \mid p) P_T(p) \delta(x - M_{t\bar{t}}(p))$$

As M_{tt} estimator we use average of this distribution:

$$M_{t\bar{t}} = < \rho(x \mid j) >$$

MC/Data Samples

- Signal samples:
  - Pythia generated narrow resonant $t\bar{t}$ samples with masses 450, 500 ... 900 GeV
- Background Samples:
  - SM $t\bar{t}$ MC sample
  - QCD
    - Data driven

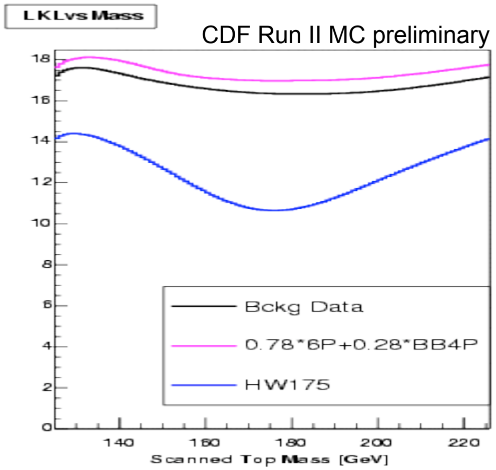


Trigger & Prerequisites

- Multijet Trigger
  - L1: ≥ 1 tower with E_T≥10 GeV
  - L2: ≥ 4 clusters with E_T^{cl}≥15 GeV, ΣE_T≥125 GeV
  - L3: N_{jet}≥4, with E_T^{jet}≥10 GeV
    - $\sigma \approx$ 14 nb, ~85% all hadronic efficiency
- Prerequisites
  - Good run list
  - Vertex: |z|<60cm & $|z-z_p|$<5cm
  - Missing Et Significance: < 3 $(\mathrm{GeV})^{1/2}$
  - Tight lepton veto
  - 6,7 tight jets - E_T^{jet} ≥ 15GeV, $|\eta|$<2.0
- After prerequisites we have $t\bar{t}$/QCD~ 1/1000!

Neural Net Idea

- Neural net event selection:
  - Uses a Root class TMultiLayerPerceptron
  - 11 inputs, 2 hidden layers with 20/10 nodes and 1 output
- SumEt - total transverse energy
- SumEt3 - sub-leading transverse energy $\sum E_T - E_{T1} - E_{T2}$
- C - centrality: $\sum E_T / \sqrt{\hat{s}}$
- A - aplanarity: 3/2*(smallest eigenvalue) of $M_{ab} = \sum_j P_a^j P_b^j / \sum_j |\vec{P}^j|$
- E^*_N - geom average of transverse energy of the N-(2 leading jets)
- E^*_{T1} - transverse energy of the leading jet
- M_{2j}^{min} - the minimum dijet mass
- M_{2j}^{max} - the maximum dijet mass
- M_{3j}^{min} - the minimum trijet mass
- M_{3j}^{max} - the maximum trijet mass
- FlaME variable, $\sum$-Log(P(M_{top}=155,160…195GeV))



QCD background

- We build tag matrix from events from 4,5 jet events.
- Each element in the matrix defined as:

$$P^{+}(\mathrm{N}v12,\ Et_{jet},\ N_{tr,jet}) \equiv \frac{N_{j,tag}(\mathrm{N}v12,\ Et_{jet},\ N_{tr,jet})}{N_{j,tgbl}(\mathrm{N}v12,\ Et_{jet},\ N_{tr,jet})}$$

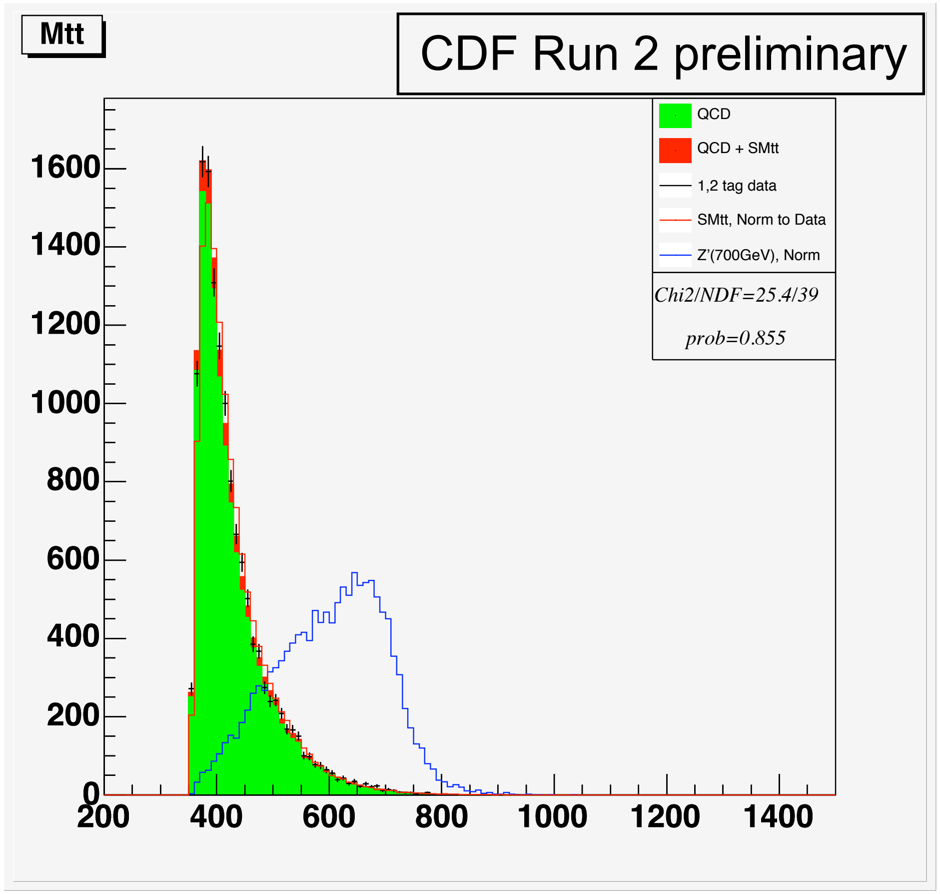
- The probability to single/double tag an event:

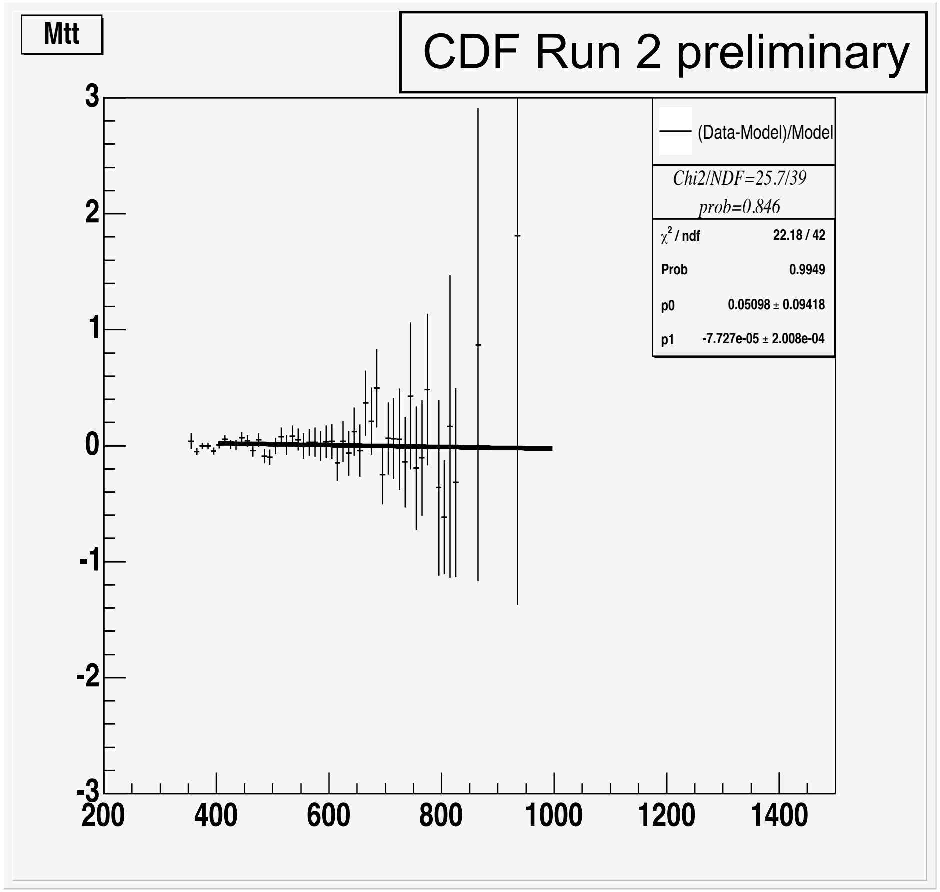
$$\sum_{i}\left[P_i^{+}\cdot\prod_{j\neq i}(1-P_j^{+})\right] \qquad \sum_{i,j\neq i}\left[P_i^{+}\cdot P_j^{+}\cdot\prod_{k\neq i,j}(1-P_k^{+})\right]$$

- We weight each event in pre-tagged data sample to get the prediction for 1, 2 tagged events
- Finally, we define several control region and test our modeling with observation
- For all control regions we get a very good agreement
- Biggest impact on final result comes from possible signal contamination, using this procedure

Crosscheck with data

0.75<NNetOut<0.93.





Limit Setting Methodology

- Template event weighting
  - N_{X0}: based on assumed cross-section and acceptance
  - N_{tt}: based on theoretical cross-section and acceptance
  - N_{QCD}: Balance from data

$$N_{cdf}^{tot} = \int L dt \cdot (\sigma_{X0} A_{X0} + \sigma_{t\bar{t}} A_{t\bar{t}}) + N_{QCD}$$

- Likelihood
  - N_{X0}, N_{tt}, N_{QCD} are used to compute the expected number of events in mass bin "i":

$$\mu(i) = N_{X0} T_{X0}(i) + N_{tt} T_{tt}(i) + N_{QCD} T_{QCD}(i)$$

  - Given the observed number of events n(i) and expected μ(i) in bin "i", the likelihood is equal to:

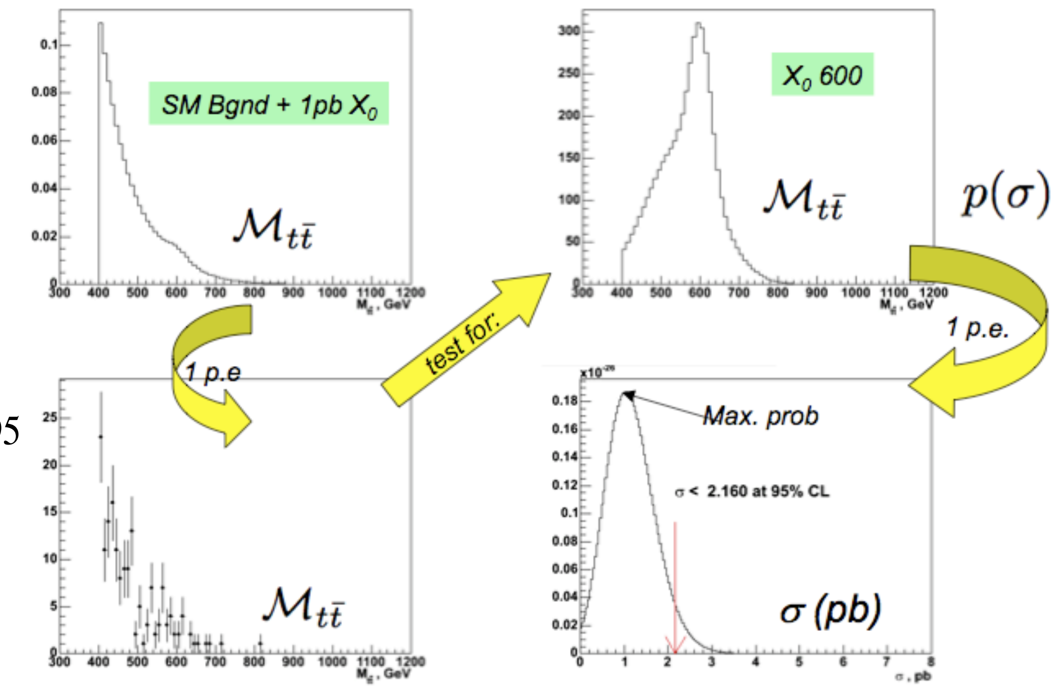
$$L(\sigma_{X0}, \vec{\nu} \mid \vec{n}) = \prod e^{-\mu_i} \frac{\mu_i^{n_i}}{n_i!}$$

Posterior density function

- Acceptance uncertainties accounting

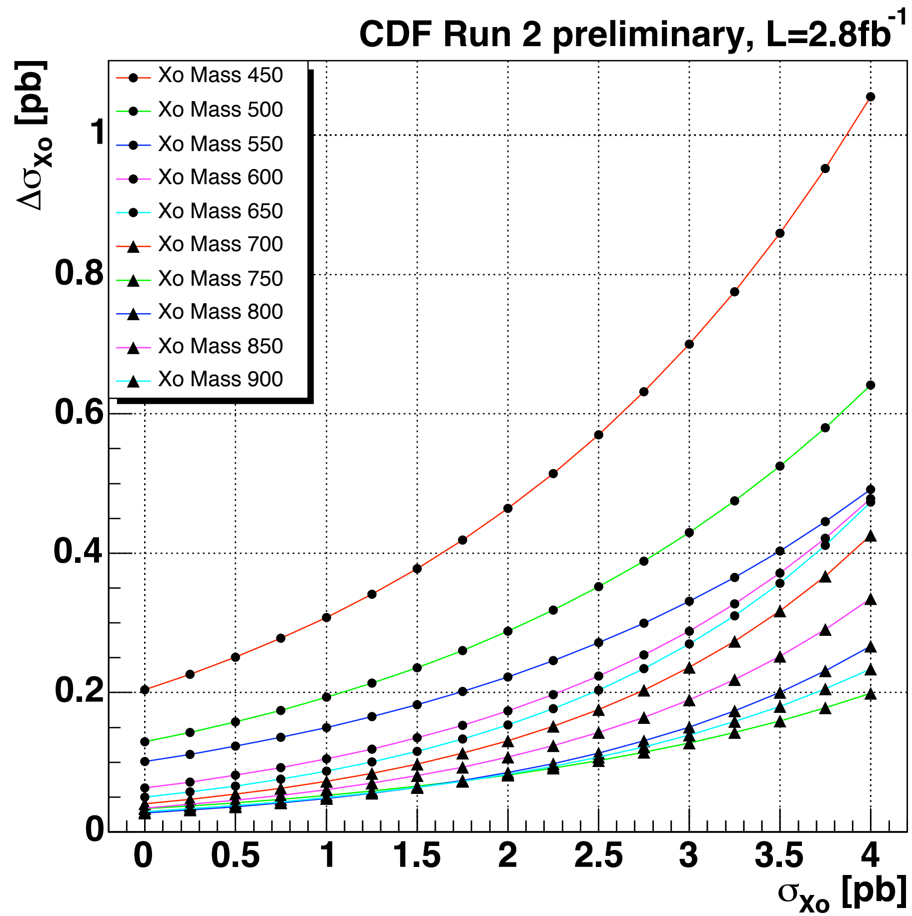
$$p(\sigma_{X0},\vec{n}) = \int d\vec{v} \cdot L(\sigma_{X0},\vec{v} \mid \vec{n}) \cdot \pi(\sigma,\vec{v})$$

- We integrate over the nuisance parameters, uncertainties for:
  - Signal acceptance
  - Background acceptance
  - Background cross-section
- Given p(σ|n) we define:
  - σ_{X0} - max of PDF
  - 95% confidence level upper limit(UL) $\frac{1}{Area}\int_0^{UL} p(\sigma \mid \vec{n})d\sigma = 0.95$
  - Values are calculated as median after 1000 PE's



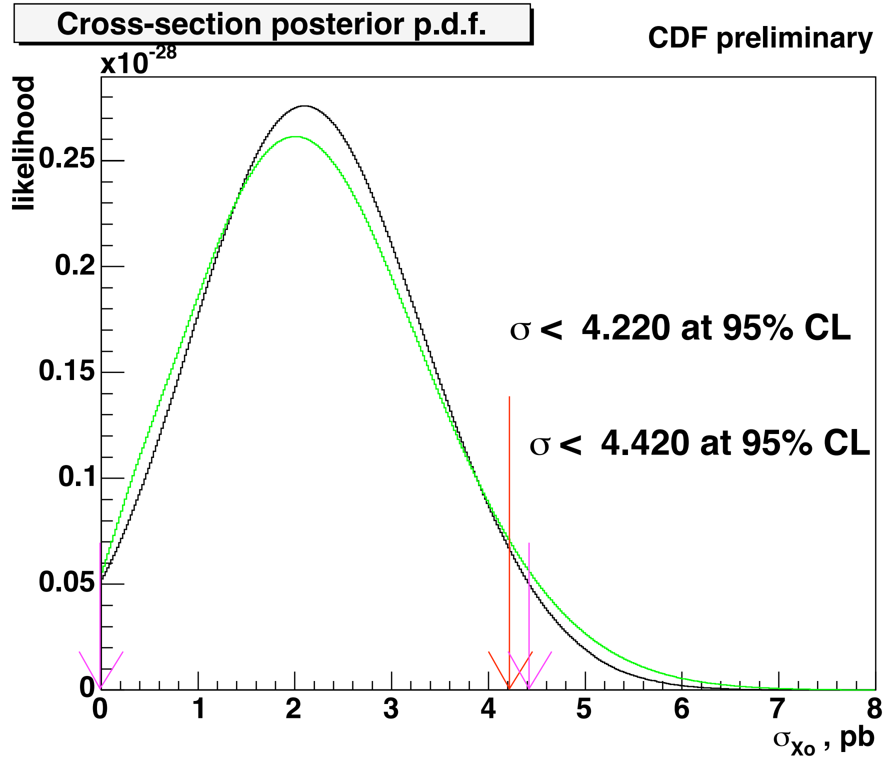
Systematics

- To consider systematics, which both affect shape and acceptances, we:
- Consider the shift on cross-section by:
- Running PE from shifted templates and fit them with nominal ones
- We considered systematics due to JES, ISR/FSR. PDF found to be negligible

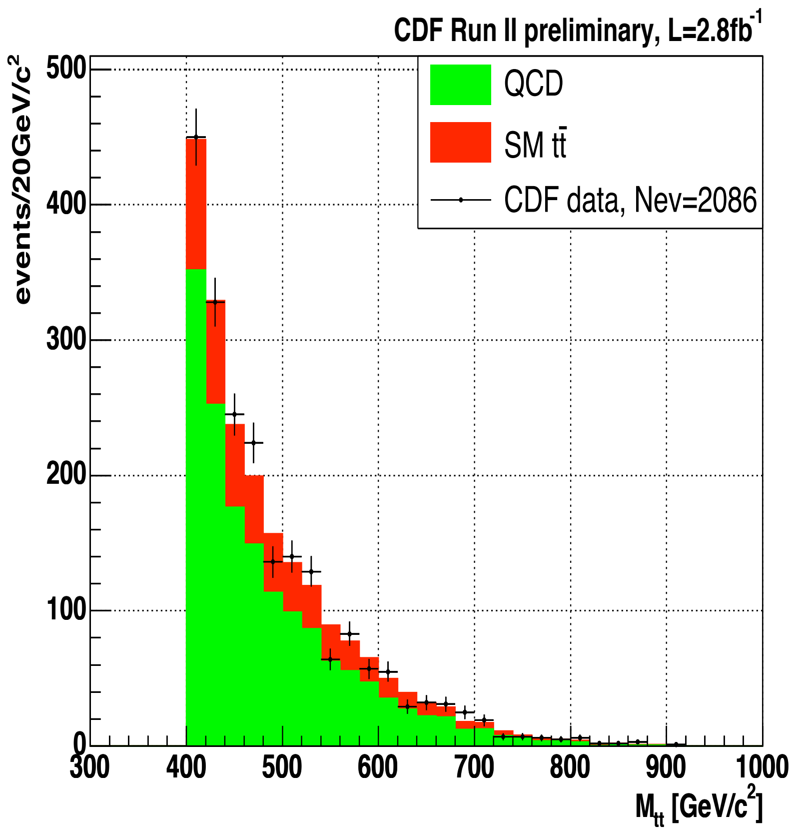


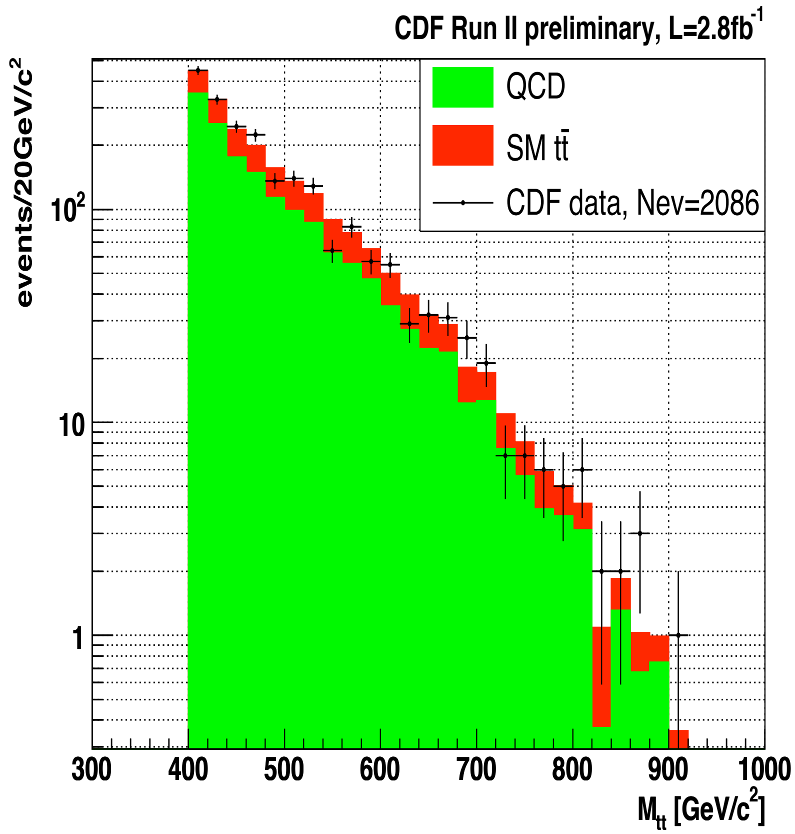
Applying systematics

$$PDF_{SYS}(\sigma_{X_0}) = \int_0^\infty \frac{1}{\delta\sigma_{X_0}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\sigma_{X_0}-\sigma'}{\delta\sigma_{X_0}}\right)^2\right) PDF(\sigma')\cdot d\sigma'$$

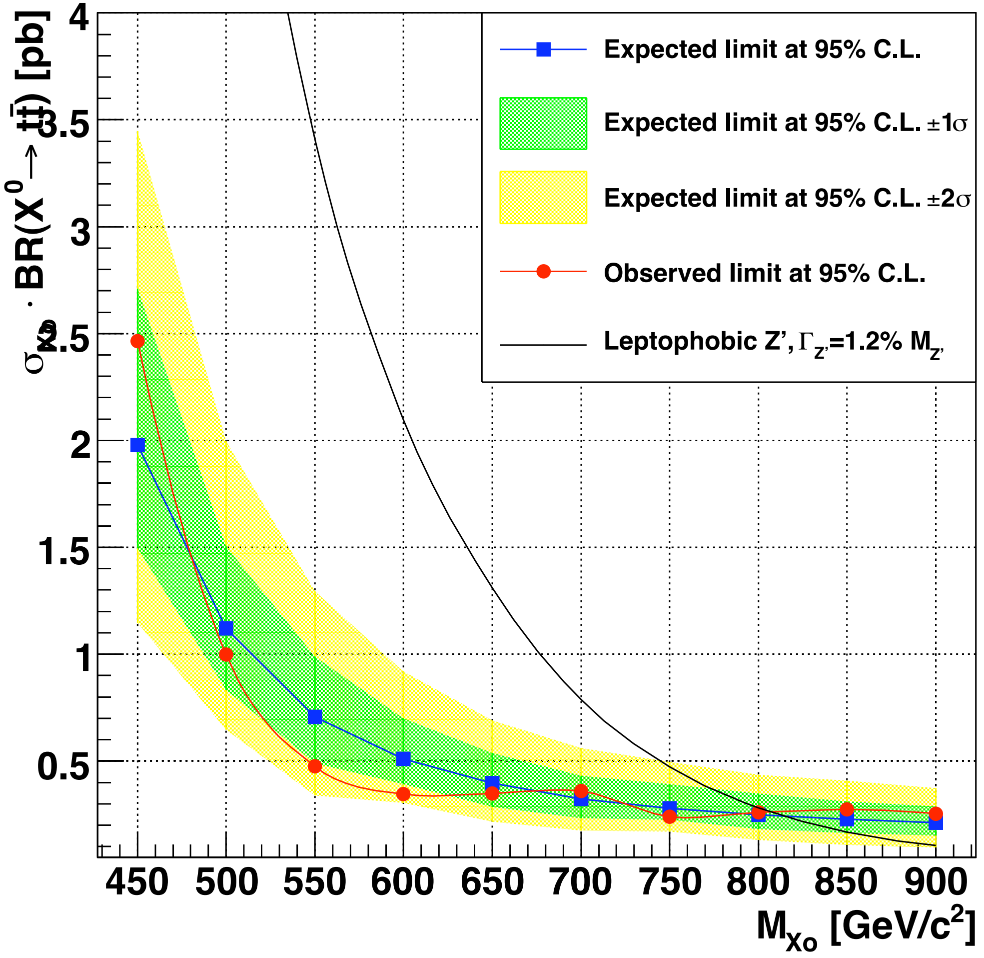


Data/BG prediction





Upper Limits



Conclusions&Plans

- First search for narrow ttbar resonance in all jets final state
  - No excess found in 2.8/fb of CDF data
  - We set observed upper limit on leptophobic Z’ mass up to 805 GeV
  - Used various tools to test SM for very small % "contaminations of new physics"
- Analysis has been reviewed and no unresolved issues were found
- Plan
  - PRD publication
  - Updated result with improvements in lepton plus jets

Backup slides

Signal contamination

- Signal contribution to QCD shape will be treated as following:

From Equtaion 4 we have, number of events in bin "i":

$$\mu = \sigma_s A_s T_s + \sigma_{tt} A_{tt} T_{tt} + N_{QCD}^{pure} T_{QCD}^{pure}$$

$$N_{QCD}^{pure} T_{QCD}^{pure} = N_{QCD}^{cont} T_{QCD}^{cont} - \sigma_s A_s^{cont} T_s^{cont} - \sigma_{tt} A_{tt}^{cont} T_{tt}^{cont}$$

Comparing signal templates of predicted and observed values we can assume:

$$T_s = T_s^{cont}$$

So, finally we get:

$$\mu = \sigma_s (A_s - A_s^{cont}) T_s + \sigma_{tt} A_{tt} T_{tt} + N_{QCD}^{cont} T_{QCD}^{cont} - \sigma_{tt} A_{tt}^{cont} T_{tt}^{cont}$$

- In the end, it decreases signal acceptances by the values we get from TRM, which is about 1-1.5%
- It will obviously result in the worse sensitivity.

Simplifications

- To calculate that probability we need to compute 28 integrals:
  - P_t and P_z of incoming partons
  - 4-momenta of 6 final partons
- To reduce CPU time, we made some assumptions:
  - Pt of the incoming partons is 0. -2 integrals
  - All quarks except top are massless. -8 integrals
  - Partons and jets have the same direction. -12 integrals
  - W's and top's are on shell. -4 integrals
- Only 2 integrals in total. We'll do more in the future.

Improvements

- Some of the improvements made:
  - Added events with 7 jets, considering last 3 jets in Et as extra jets from radiation
    - results in better signal acceptance
  - Used refined binning for transfer functions, both in Eta in Et
    - results in better signal templates
- Both improvements should result in better sensitivity

Details

Jet-parton assignments

Differential xsection

P_t of ttbar system

$$P(j \mid M_{top}) = \frac{1}{\sigma(M_{top})\varepsilon(M_{top})N_{combi}} \sum_{combi} \int dz_b dz_b f(z_a) f(z_b) d\sigma(M_{top}, p) TF(j \mid p) P_T(p)$$

Normalization factor

Integration over PDF's

Transfer functions

dσ calculation

$$d\sigma(M_{top},p) = \frac{|\mathcal{M}(M_{top},p)|^2}{4E_aE_b|v_a - v_b|}(2\pi)^4\delta^4(p_a + p_b - \sum_{i=1}^{6} p_i)\prod_{i=1}^{6}\frac{d^3p_i}{(2\pi)^3 2E_i}$$

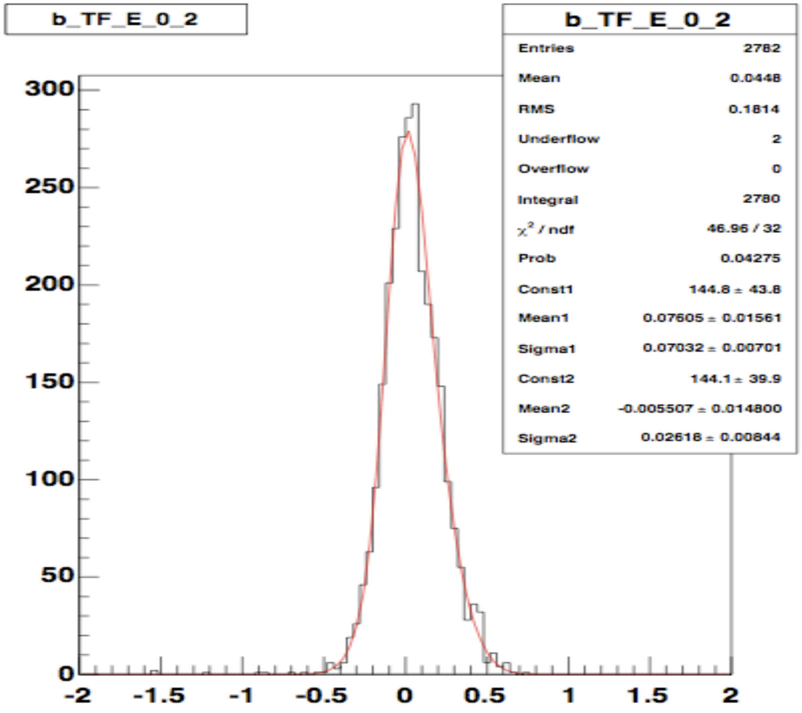
$$\mathcal{M}(M_{top},p) \approx \frac{\overline{v}(p_{\overline{q}})\gamma^\mu u(p_q)}{(p_q + p_{\overline{q}})^2}\cdot\frac{\overline{u}(p_u)\gamma^\alpha(1-\gamma^5)v(p_{\overline{d}})}{P^2_{W+} - m^2_W + im_W\Gamma_W}\cdot\frac{\overline{u}(p_d)\gamma^\sigma(1-\gamma^5)v(p_{\overline{u}})}{P^2_{W-} - m^2_W + im_W\Gamma_W}\cdot$$
$$\overline{u}(p_b)\gamma_\alpha(1-\gamma^5)\frac{\not{p}_t + m_t}{p_t^2 - m_t^2 + im_t\Gamma_t}\gamma_\mu\frac{\not{p}_{\overline{t}} + m_t}{p_{\overline{t}}^2 - m_t^2 + im_t\Gamma_t}\gamma_\sigma(1-\gamma^5)v(p_{\overline{b}})$$

- We use uubar --> 6 exact tree level ME
- Spin-correlations are included
- We compute the amplitudes directly using explicit Dirac matrices and spinors

Transfer functions

- From MC calculate the probability density function TF($E_j|E_p$)
  - ➢ $\xi = 1 - E_{jet} / E_{parton}$
- Use differnet TF's for different regions in η, energy, quark types

Example of b-quark Transfer Function for $1.3 \leq |\eta| \leq 2$



Prereqs effects

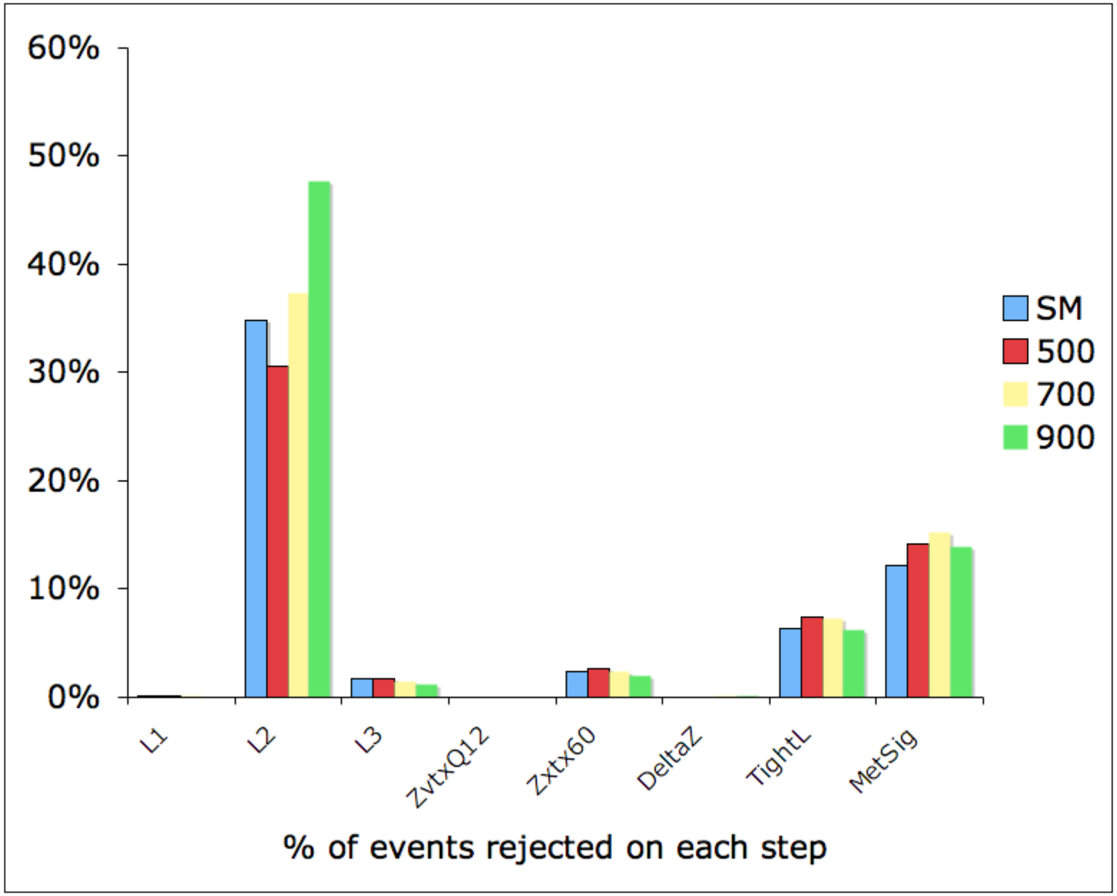
- Total efficiencies: ε_{tt}=42%, ε_{500}=43%, ε_{700}=36%, ε_{900}=28%
- Where do we lose events for high masses?

More interesting: Why?

Most of the events are lost on L2, which requires at least 4 clusters

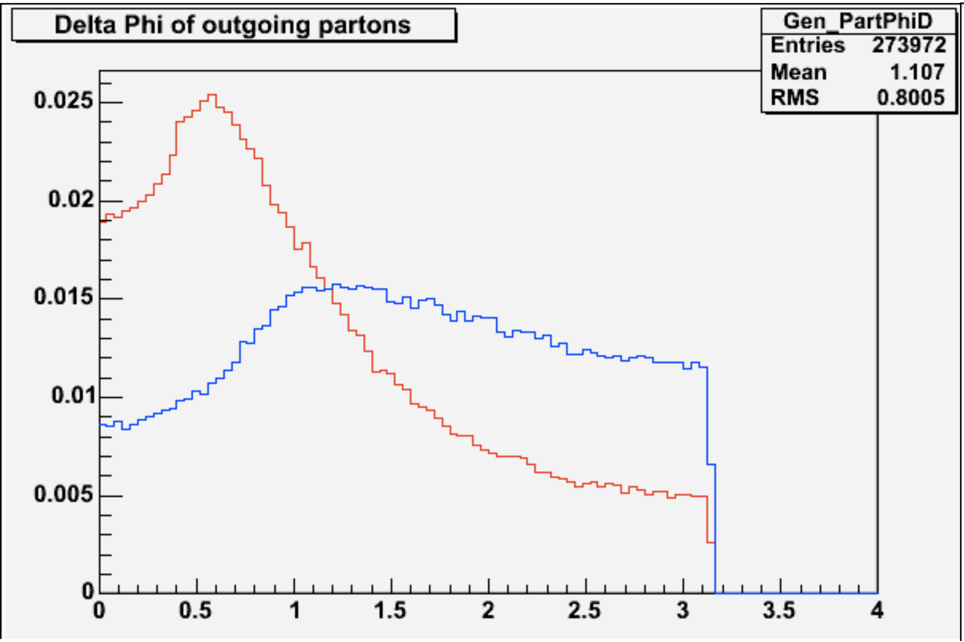
For higher resonance masses, decay products are boosted more=> higher chance to merge in one cluster

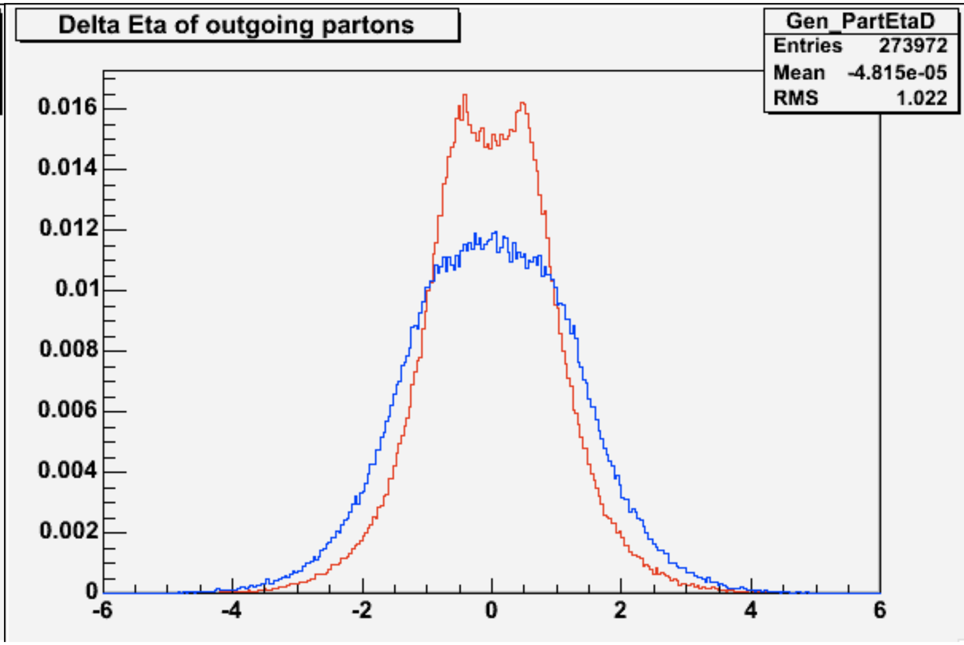
See backup slides for details



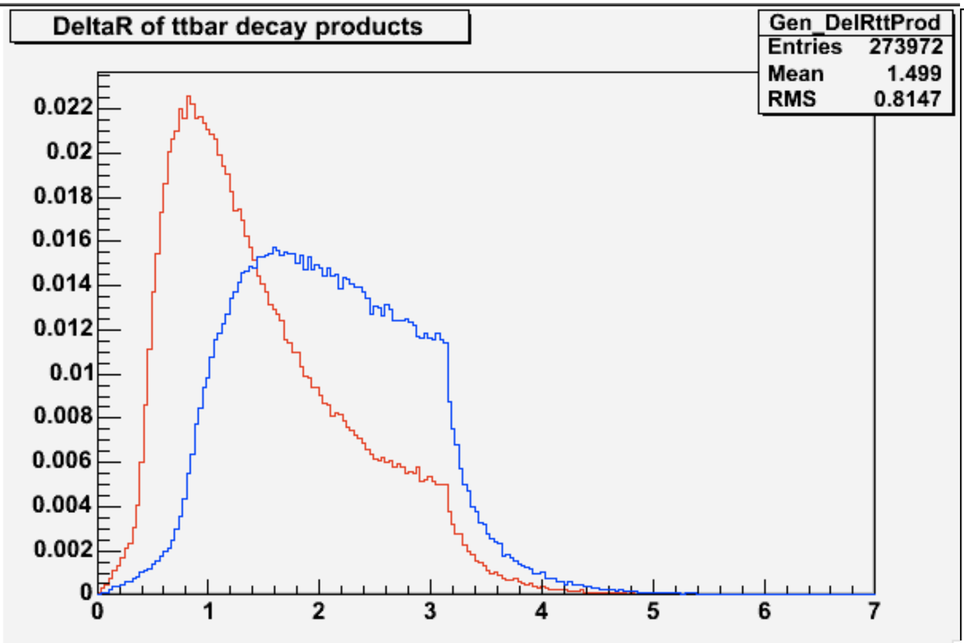
Support for L2 issue

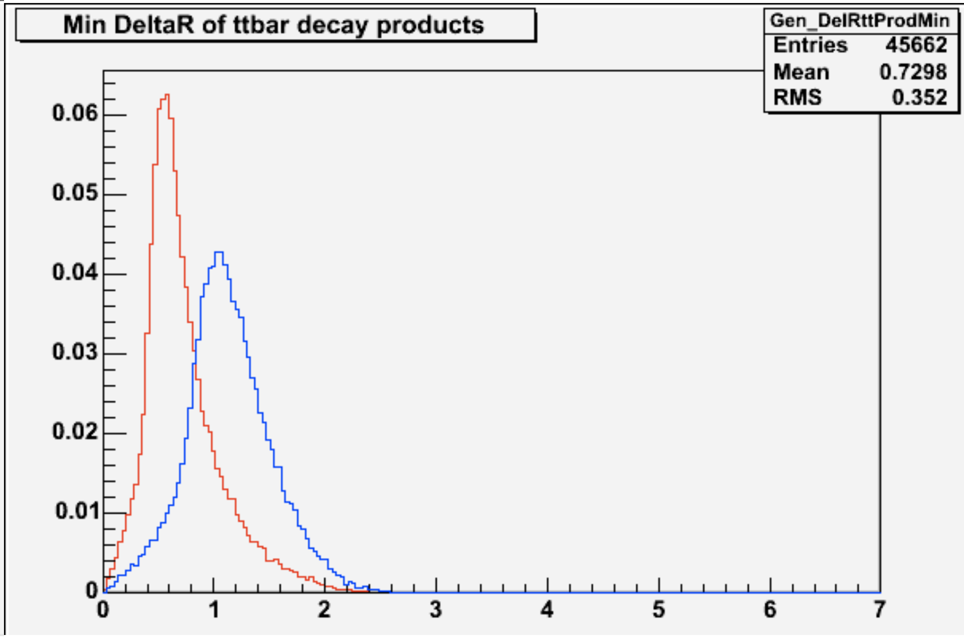
Blue 500 GeV resonance
Red 900 GeV resonance.



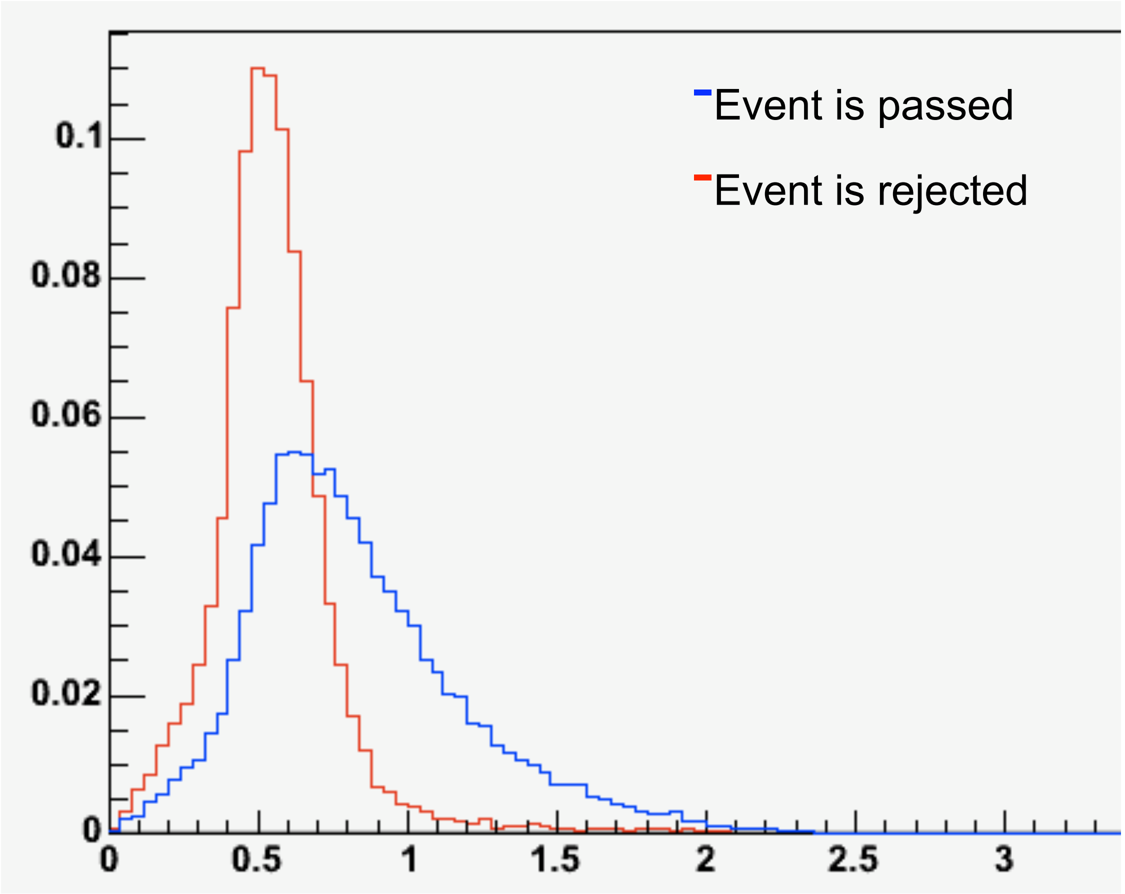


L2 continued



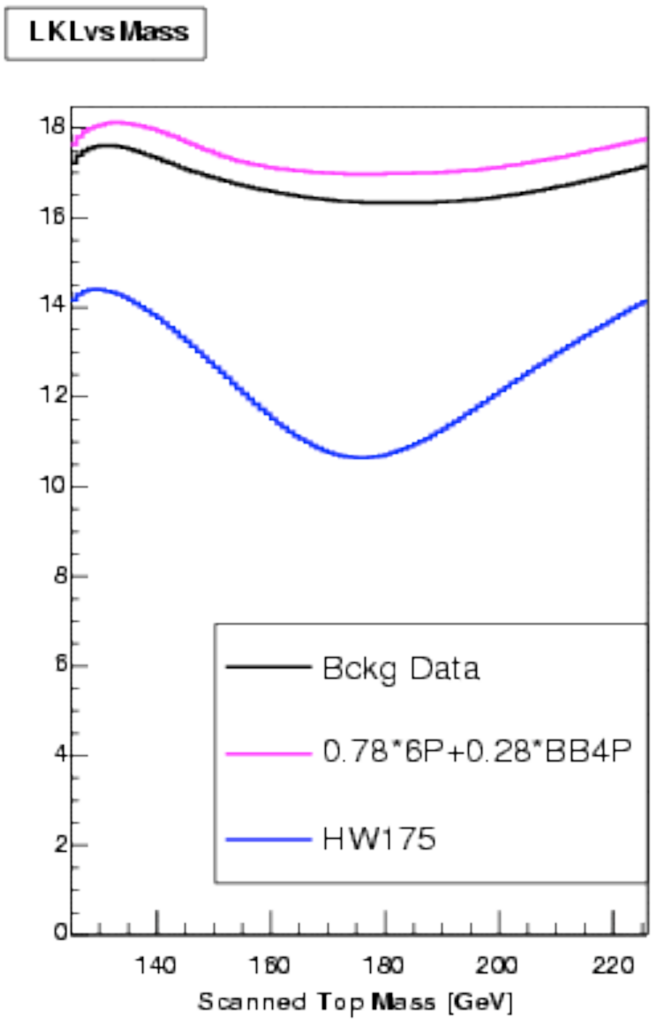


L2 final



Event is passed
Event is rejected
0.1
0.08
0.06
0.04
0.02
0
0
0.5
1
1.5
2
2.5
3

FlaME Variable



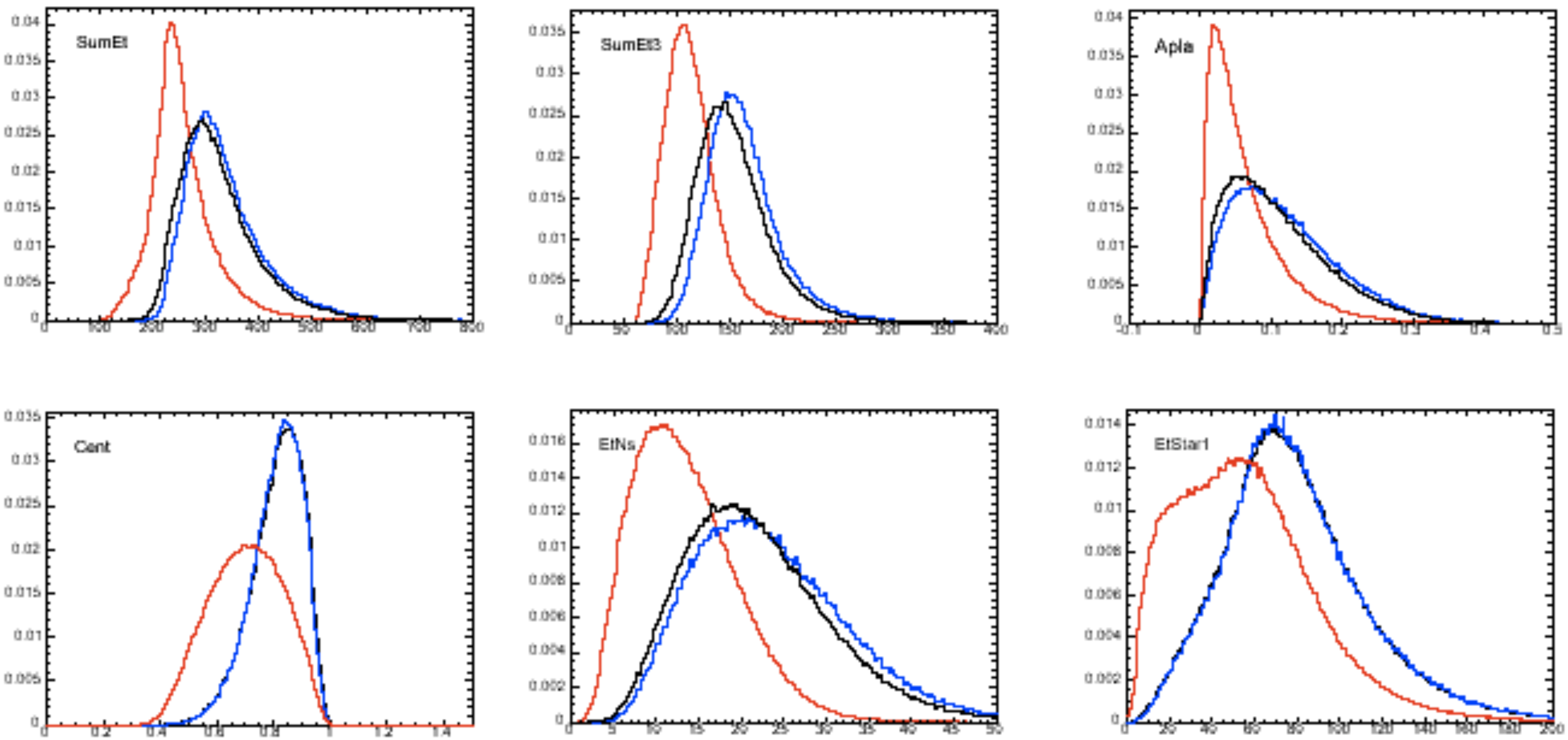
FlaME gives the probability of an event to come from SM ttbar. Let's take advantage of it!

Here we plot -log(P) vs top mass for various samples. As you see there is a difference between ttbar and QCD

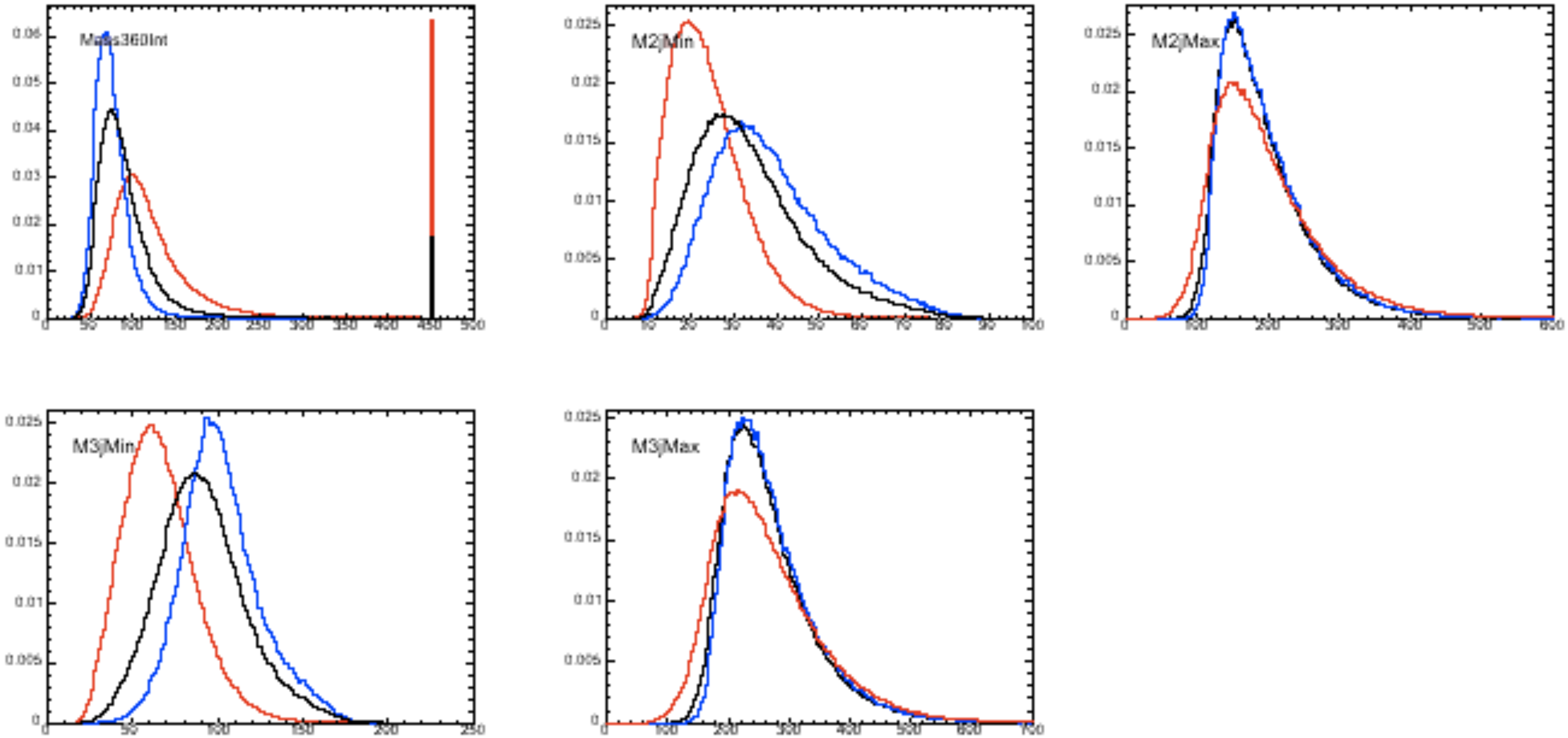
Lets calculate -log(P) for 9 mass points: 155,160…195GeV. Decided to use their sum

…and their distributions

Red lines correspond to data.
Black lines correspond to SMtt
Blue lines correspond to SMtt matched only



…and their distributions



Plug&Play

Black with FlaME, Red without FlaME, green kin. ev. sel.