Dear Dr. Malenfant,

We would like to resubmit the manuscript LN13161 “Measurement of the Top Quark Mass in $p\bar{p}$ Collisions using Events with Two Leptons” for consideration in Physical Review Letters. When this manuscript was first reviewed , Referee A strongly endorsed it and recommended its publication. Referee B expressed support for the eventual publication of the manuscript, but not necessarily in PRL. The primary concerns were that the results presented have some relationship to those of Reference [7] which should be documented, and the analysis lacked innovation or interest.

As to the first point, stated in the first and last paragraphs of the comments from Referee B, we now include not only the correlations, but also a full combination of the two results. We believe its inclusion further enhances the impact of this work by reducing the total uncertainty on the top mass using dilepton events alone by more than 15%. Comparing the results of the two individual analyses and the full combination indicates that the new measurement presented in the manuscript dominates the accuracy of the combined result. More specifically, the selected approach to the calibration has substantially reduced the dilepton top quark mass uncertainty.

Concerning the second point, several innovations were omitted or unclear in the original manuscript. We edited the paper to add those descriptions and clarified the impact. We believe the paper has substantial impact because 1) it demonstrates that the dilepton events can be used to obtain competitive mass measurements, 2) resolves an experimental problem that impacts other analyses, and 3) contains the most precise mass measurement in the dilepton channel. We believe the paper has wide interest because of all these elements.

We believe the clarified text and responses will be sufficient for Referee B to reconsider his/her view.

Sincerely,

Professor Robert Kehoe
(on behalf of the D0 Collaboration)

From prl@aps.org Tue Feb 21 12:33:26 2012
Date: Tue, 21 Feb 2012 12:33:22 -0600
From: prl@aps.org
To: kehoe@physics.smu.edu
Subject: Your_manuscript LN13161 Abazov

Re: LN13161
Measurement of the top-quark mass in p bar p collisions using events with two leptons
by V. M. Abazov, B. Abbott, B. S. Acharya, et al.

Dear Dr. Kehoe,

The above manuscript has been reviewed by our referees. A critique drawn from the reports appears below. On this basis, we judge that while the work probably warrants publication in some form, it does not meet the Physical Review Letters criteria of impact, innovation, and interest.

The paper, with revision as appropriate, might be suitable for publication in Physical Review. If you submit the paper to Physical Review, the editors of that journal will make the decision on publication of the paper, and may seek further review; however, our complete file will be available.

If you submit this manuscript or a revision of it to Physical Review, be sure to respond to all referee comments and cite the code number assigned to the paper to facilitate transfer of the records.

If you feel that you can overcome or refute the criticism, you may resubmit to Physical Review Letters. Please accompany any resubmittal by a summary of the changes made and a brief response to all recommendations and criticisms.

Yours sincerely,

Jerome Malenfant
Senior Assistant Editor
Physical Review Letters
The premier APS journal for current research
Email: prl@ridge.aps.org
Fax: 631-591-4141
http://prl.aps.org/

Report of Referee A -- LN13161/Abazov

The paper "Measurement of the top quarks mass in pp collisions using events with two leptons" (by the D0 collaboration)" describes the measurement of the top mass. I do not have any significant comments and believe that the analysis, which is done well known experimental procedure and the "golden" top decay channel, is done accurately. I recommend this paper for publication

We thank the referee for the review and judgment of our paper.

Report of Referee B -- LN13161/Abazov

COMMENT: I have a major concern with this result, in that it is submitted as apparently an independent measurement of the top quark mass, when there must be significant correlations with a previously published analysis on presumably the same data set (and even the same events) [PRL 107, 082004 (2011)]. The authors reference the previous publication, but only to the extent of commenting on the size of the statistical uncertainty.

RESPONSE: At the time of the first submission, we omitted a direct comparison of the two 5 fb^{-1} measurements, and did not perform a detailed assessment of their correlation, because of the different treatment of jet energy scale calibration. The main focus of this measurement was the reduction of the systematic uncertainty, which we knew was crucial based on extrapolations from the previous analysis [Phys. Rev. D 80, 092006 (2009)] as well as from the results of the measurement based on the Matrix Element technique [Phys. Rev. Lett. 107, 082004 (2011)]) . The systematic uncertainty on the mass is limited by systematics of jet calibration. We have made the limitations and motivation more clear in Paragraph 2, noting that the dilepton channels "*remained statistically limited, unlike in channels with one lepton and four or more jets (ℓ+jets). This situation has changed recently (eg. Ref. [7]). Now, dominant systematic uncertainties from jet energy calibration, which have been larger [1] in the dilepton channel compared to ℓ+jets, are limiting accuracy of the m_t measurement*", **and the start of Paragraph 9, clarifying "***The event topology… has prevented significant progress in reducing the large standard jet energy scale uncertainties in dilepton analyses.*"

We have now estimated the correlation between the two measurements performed on the Run 2a data via pseudo-experiments to be 0.62. In Run 2a data, the same jet energy scale calibration was used in both measurements, so this correlation is strictly the correlation between the methods. We also calculated the correlation for measurements performed on the Run 2b data to be 0.56, taking into account the

effect of different jet energy scale calibrations. With this information, we combined the two measurements, yielding 173.9 ± 1.9 (stat) ± 1.6 (syst) GeV. We added text in Paragraph 3 of the paper to indicate that there is a correlation, and, in the final paragraph to present this combination. We now explicitly point to the correlation at the end of Paragraph 2 with "*The presented m_t measurement is performed using the same data as Ref. [7], and is correlated with it as discussed below*". **We quote the mean value and the mass combination at the end of Paragraph 20 with** "*We have also improved the precision by combining the ν WT results with the results of Ref. [7]. The statistical correlation of these two measurements is approximately 60%, calculated from pseudoexperiments. Accounting for this correlation, and correlations appropriate to each source of systematic uncertainty, we obtain m_t = 173.9 ± 1.9(stat) ± 1.6(syst) GeV.*"

COMMENT: From what I can see, the new result uses an older method to determine the top quark mass; the only innovation is the use of the jet energy scale calibration from the lepton+jets analysis.

RESPONSE: Your conclusion is understandable, given the original wording of the paper. However, we first want to point out that the method is used widely because it provides competitive precision in mass analysis, particularly in high statistics measurements, and is not computationally demanding. It is also applied for the measurement of other properties of the top quark, like the recent measurement of spin correlations in top events [Phys. Lett. B 702, 16 (2011).] In comparison with the previous publication based on the neutrino weighting technique, there are innovative modifications in three areas, (i) modeling of neutrino properties, (ii) jet energy scale calibration, and (iii) treatment of probability density, that can be summarized as follows:

(i) Two substantive improvements to the modeling of neutrino properties.

a. **The moments of the weight distribution for each event exhibit significant variations when using a coarse binning in the neutrino rapidity. This can be inferred from the difference observed for each moment of the distribution relative to its asymptotic value estimated using a factor of 400 finer granularity. Optimizing the granularity via pseudo-experiments to reduce this dependence by a factor of five also yielded an expected 4% reduction in statistical uncertainty. We have clarified this point in Paragraph 11:** "*We found in Ref. [1] that most of the statistical sensitivity to m_t is obtained from the first two moments of this weight distribution, the mean (μ_w) and RMS (σ_w). A coarse granularity of our sampling of the η distribution causes these moments to be unstable. To reduce this variation, we have increased the sampling for this integration by an order of magnitude relative to our previous analysis [1]. This improves the expected statistical uncertainty on m_t by 4%.*"

b. **A second improvement involved inclusion of a dependence of the rapidity distribution of the neutrinos over which we integrate on the mass of the top quark. We tried different functions for the rapidity distribution, including single and double Gaussians and exponentials. We checked their dependence on top mass. Eventually, we parametrized the neutrino rapidity by a single Gaussian, and we used a linear dependence of the width of this Gaussian on the top mass. Using more sophisticated parametrizations indicated a**

negligible (30 MeV) change in expected statistical uncertainty in pseudo-experiments. This discussion is now included in Paragraph 11: "*We model the neutrino η distributions by a Gaussian whose RMS is a linear function of m_t. Several more sophisticated parametrizations were tested, but provided negligible improvement in expected precision in pseudoexperiments*."

(ii) There are two major improvements in the treatment of jet energy scale calibration. Only one is based on the W→jj analysis in lepton+jets events.

a. **The carry-over of the dijet mass calibration from lepton+jets, including the determination of systematic uncertainty, has been a goal of many analyses for several years. The main obstacle in the transfer of the calibration from the lepton+jets events to the dilepton events is caused by the difference in the jet multiplicity, which is then reflected in a large systematic uncertainty. For example, in the most recent CDF measurement in the dilepton final state [Phys. Rev. D 83, 111101 (2011)], the benefit from the use of the W→jj constraint on the jet energy scale is limited by the lack of improvements in the systematic uncertainties. The "double ratio" presented in our article, while numerically small, addresses the issue of the difference in event topology between l+jets and dilepton events, and thereby provides a reduction of uncertainty through the W→jj mass calibration. For the dilepton analysis, this calibration reduces the systematic uncertainty from ~2% to 1.1%. (from 1.35 to 0.85 GeV). With this improvement, the dominant uncertainties in the dilepton mass measurement continue to scale with the size of the data sample. In addition to the changes to the text that have already been mentioned, we clarified Paragraph 3 to state that we "…***estimate the uncertainties of transferring that calibration to the dilepton event topology. This procedure demonstrates how the calibration obtained using the dijet constraint from M_W can be applied to different final states, and has wide applicability beyond the measurement of m_t in 2ℓ events*." **We have clarified how this is done in Paragraph 9:** *"We therefore take 0.3%, the maximum excursion of R_b from unity, as a systematic uncertainty on carrying over the ℓ+jets scale to the jets in our dilepton sample. This is applied as a direct correction to the standard calibration."* **We also clarified how the residual uncertainty is estimated in Paragraph 14:** "*We calculate the average of the energy scale uncertainty for jets in the W→jj sample. For each jet in the dilepton sample, we apply a shift corresponding to the difference between its uncertainty in energy scale and the W→jj sample's average uncertainty in energy scale*."

b. **The use of flavor corrections to the MC jet energy scale is also implemented for the first time in a dilepton mass analysis. This is described in Paragraph 6, and the resulting systematic uncertainty is described in Paragraph 15. This has significant impact on reducing the measurement's remaining largest systematic uncertainty, from 0.8 GeV to 0.5 GeV. We now introduce this improvement in Paragraph 3 with "***We also employ flavor-dependent corrections to jet energies for the first time in a dilepton analysis that substantially reduce the uncertainties on jet energy resulting from jet flavor*." **We added text to Paragraph 6 to motivate this correction and lead into the description already in the paper:** "*Because the flavor dependence of jet energy calibration can yield one of the largest systematic uncertainties on our measurement [1], we have improved our analysis by accounting for this dependence*." **We clarified the discussion of the uncertainty**

on this correction in Paragraph 15: **"***The flavor-dependent jet energy corrections described earlier provide MC-based mass templates that accurately reflect the data. As in[9], we propagate the uncertainty in these corrections and obtain a systematic uncertainty on m_t of 0.5GeV*."

(iii) The treatment of probability density also has two major improvements.

a. **The use of the moments of the weight distribution was an innovation we reported in [Phys. Rev. D 80, 092006 (2009)]. For the current publication, we extended this method for a much larger data sample, and use a bin size for the weight distribution that optimizes the expected statistical uncertainty. This innovation was not discussed in the previously submitted version of the paper, and we have added a brief description in Paragraph 12: "***We use a binning that provides the minimum expected statistical uncertainty, as checked in pseudoexperiments*."

b. **Maximum likelihood fits as a function of mass can occasionally fail to converge. This is a particular problem in channels with relatively large background, such as that involving dimuon events. To address this issue, we implemented an iterative fitting procedure in pseudo-experiments that rectified the maximum likelihood calculation. Text describing this is added to Paragraph 12:** *"Point-to-point fluctuations mean that the initial placement of the window may result in an unconverging fit. We therefore iterate the fit around the current fit minimum. This results in a significant improvement in fitting efficiency, particularly in the dimuon channel."*

COMMENT: This is a detail, that while reducing the systematic uncertainty on the result, is presumably only of interest to specialists in the field. And by specialists, I do not mean high-energy physicists, but top quark mass analysts.

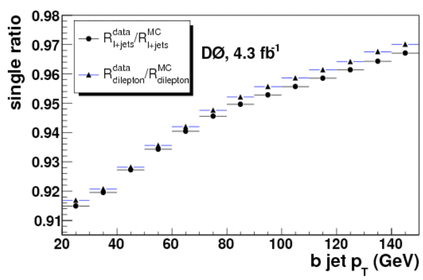
RESPONSE: We respectfully disagree for several reasons. The calibration methodology in this paper provides the first evidence that dilepton channels can ultimately achieve competitive sensitivity to the top quark mass as available in lepton+jets channels. This means that in the future these events will allow a further significant reduction on the uncertainty of the top mass. The approach in this paper applies to any measurement in top dilepton events where jet calibration is a significant uncertainty. As a paper on an improved technique to measure a fundamental parameter constraining electroweak fits, this paper would attract those interested in Higgs and other electroweak physics. We have clarified the relationship of top mass and Higgs in Paragraph 1 with: "*It is therefore important to sharpen the measurement of m_t as its precise value, and with the mass of the W boson (M), constrain Mh through radiative corrections*."

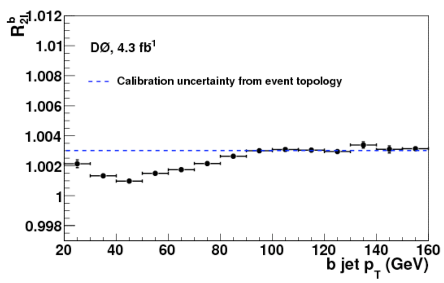
Beyond top quark physics, the successful isolation of the effect of event topology relative to effects from flavor and kinematics, in the transfer of the W→jj mass calibration, is a conceptual breakthrough that has potentially wide applicability in analyses limited by systematics of jet calibration. Our edits described above make this point in the paper clear.

COMMENT: That said, I do have some specific comments on the paper as written.

o One of the major results is that the energy scale can be transferred from the lepton+jets to the dilepton sample. This is shown by looking at the double ratio (top of page 5). It would be illuminating to see how well the data and Monte Carlo agree by showing the single ratios as well.

RESPONSE: Please see the attached plots of the single ratio and double ratio.





o In the same paragraph, it is stated that the particle multiplicity in b jets in lepton+jets events is different than in dilepton events. I can't see of a good reason for the other than through reconstruction effects, so it should say that the *reconstructed* particle multiplicity is different.

RESPONSE: The change in particle multiplicity is not at the 'reconstructed' event level, but at the generator level in MC events. We apply a jet cone algorithm at this level, and, due to the larger number of colored objects in l+jets events compared to dilepton events, the generator level b-quark jets have higher particle multiplicities. We have changed the wording in paragraph 9 to: "*The multiplicity of particles in b-quark jets in l+jets events at the MC generator level is, after the application of the offline jet reconstruction algorithm, a few percent higher than in the dilepton sample …*".**

o Irrespective of whether the paper is published in PRL, I think that a comparison of this result with the previously reported result is merited.

RESPONSE: Due to space limitations we cannot give a complete detail of the differences between the two analyses. However, as mentioned above, we now include a combination of the two results taking into account all the correlations, including those between systematic uncertainties. We believe this combination efficiently facilitates a comparison of the methods. In particular, the 2.5 GeV total uncertainty of the combination can be compared to 2.6 GeV for the current result, and 3.0 GeV for the matrix-element result.

<u>Further Changes to the text of the paper:</u>

While working on the manuscript, we have modified the text in few other places. In response to Referee B's comments, we concluded that three substantive changes

Robert Kehoe 6/7/12 10:45 AM
Formatted: Space Before: 12 pt, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tabs:Not at 0.39" + 0.78" + 1.17" + 1.56" + 1.94" + 2.33" + 2.72" + 3.11" + 3.5" + 3.89" + 4.28" + 4.67"

Robert Kehoe 6/7/12 10:45 AM
Deleted: **, as discussed below**

Robert Kehoe 6/7/12 10:45 AM
Deleted:

were needed. The discussion of numerical integration was clarified in Paragraph 11: "*We integrate over the η distributions predicted for both neutrinos, solve the event kinematics, and calculate missing E_T from the neutrino momentum solutions***." The description of uncertainties in our method are improved in Paragraph 18:** *"Pseudoexperiments are used similarly to account for the uncertainty in the method that arises from the uncertainties on the offset and slope in the calibration of the fitted m_t."* **We have also clarified the discussion of jet energy calibration in Paragraph 2: "***Jets are calibrated with the standard D0 jet energy correction which is derived from data [17]. The method corrects the measured jet energy to the value obtained by applying the reconstruction cone algorithm to particles from jet fragmentation before they interact with the detector. We establish the efficacy of the method in the MC, where we compare the measured jet and the jet reconstructed from fragmentation particles.*"

One textual error was fixed in Paragraph 6: *'imbalance in missing transverse energy'* **was changed to** *'imbalance in transverse momentum'.* **Several edits sought to slightly improve the readability or clarity of the text without changing its meaning, such as** *'We fit a parabola to the dependence of –ln L…"* **in Paragraph 12. Beyond these edits, the text was edited to fit within the 3500 word PRL limit.**