Nuclear Physics B Proceedings Supplement 00 (2014) 1–6

Single top quark production at CDF

Sandra Leone, on behalf of the CDF Collaboration

I.N.F.N. Sezione di Pisa, Largo B. Pontecorvo 3, 56127 Pisa Italy

Abstract

We present the most recent CDF measurements of single top quark production in proton-antiproton collisions, performed using the lepton plus jets and missing transverse energy plus jets final states. We also present the simultaneous measurements of single top quark production in s-channel and t-channel separately. The extraction of the CKM matrix element $|V_{tb}|$ from the single top quark cross section is discussed as well. A recent analysis aimed at identifying the s-channel single top quark production is described and the evidence for this process is presented.

Keywords: Single top quark, Tevatron, CDF

1. Introduction

The Fermilab Tevatron Collider provided $p\bar{p}$ collisions at a center of mass energy of $\sqrt{s}$ = 1.96 TeV until September 2011. Data corresponding to 10 fb^{-1} of integrated luminosity were recorded by CDF, a multipurpose detector which was operating at the Tevatron in several updated configurations, since 1988.

The top quark was first observed at the Tevatron in top pair production by CDF and D0 in 1995 [1]. At the Tevatron center of mass energy top quarks are produced primarily in $t\bar{t}$ pairs via the strong interaction. The standard model (SM) of fundamental interactions predicts that top (or antitop) quarks can be produced also singly, through electroweak s-channel [3] or t-channel exchange of a virtual W boson [4], with a predicted production cross section of about half the top pair production. Single top associated production with a W boson Wt is also possible [5], but at the Tevatron the expected cross section for this process is very small (approximately 0.2 pb) and is not observable by itself. Examples of SM single top quark production processes dominating at the Tevatron are shown in Figure 1.

It took 14 years after the top quark discovery for the single top quark to be observed. Single top quark production in the combined $s+t$ channels was observed for the first time independently by the CDF and D0 experiments in 2009 [2].

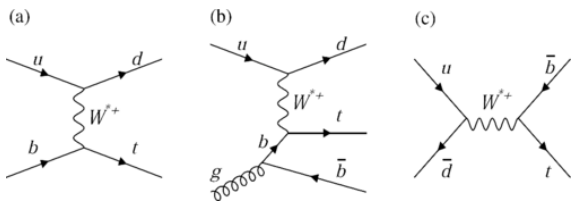

Figure 1: Feynman diagrams for electroweak single top production: (a) leading order t-channel, (b) next-to-leading-order t-channel, and (c) leading order s-channel.

If we compare the expected single top quark production cross section at the Tevatron and LHC [3, 4], we notice that the t-channel is the dominant process for both. However, the cross section of s-channel at 8 TeV only increases a factor of five compared to the Tevatron, while the signal/background ratio is lower for the s-channel at LHC than at the Tevatron. Therefore, the Tevatron is favourite with respect to LHC for the identification of the s-channel.

Single top quark production is particularly interesting because it provides direct access to the Wtb vertex: the single top quark production cross section is proportional to the square of the magnitude of the quark-mixing Cabibbo-Kobayashi-Maskawa (CKM) matrix element

V_{tb}, so its measurement provides a test of the SM via a direct determination of $|V_{tb}|$. Since this measurement assumes only $|V_{tb}|^2 \gg |V_{ts}|^2 + |V_{td}|^2$ and does not rely on the assumption of the unitarity of the CKM matrix, it can constrain various extensions of the SM like some models with fourth generation quarks which are not excluded by the precision electroweak measurements [6].

In addition each channel of the single top quark production is sensitive to different classes of SM extensions: the t-channel process is more sensitive to flavor-changing neutral currents; the s-channel is sensitive to contributions from additional heavy bosons. Therefore, independently studying the production rate of these channels provides more restrictive constraints on SM extensions than just studying the combined production rate [7].

2. Identification of single top quark events

In the SM the top quark decays almost exclusively into a real W and a b quark. Therefore the two selection strategies (lepton plus jets and missing transverse energy plus jets) used at CDF for single top quark selection start from this assumption, requiring (at trigger level) the presence of one high transverse energy lepton (electron or muon) and/or large missing transverse energy (MET). Backgrounds that mimic the single top quark signal originate from events in which a W boson is produced in association with one or more heavy-flavor jets ($W+HF$), events with light-flavor jets that are mistakenly b-tagged ($W+LF$), diboson (WW, WZ, ZZ) events, events with a Z boson produced in association with jets, $t\bar{t}$ events, and multijet (non-W) events, where a W candidate is falsely identified. Backgrounds are estimated using both data-driven algorithms and simulated data from Monte Carlo samples. One continues to clean up the signal sample by applying topological cuts and identifying jets resulting from the hadronization of b-quarks, but at the end of a smart conventional cut-based selection, the signal over background ratio is still of the order of 1 to 20. Moreover, the number of expected signal events is much smaller than the uncertainty on the predicted background, making the detection of this mode experimentally challenging. No single variable provides sufficient signal-to-background separation: the use of multivariate techniques is mandatory. There, multiple variables are combined into a single more powerful discriminant, to separate signal from background. Several multivariate methods have been used in the past. The analyses presented here are based on the use of Artificial Neural Networks (NN) techniques. The discriminant performance is checked using data control samples. A Bayesian approach is used to extract the cross-section, forming a binned likelihood assuming a flat prior in the cross section which one wants to measure. Each independent source of systematic uncertainty is assigned a nuisance parameter with a Gaussian prior probability distribution.

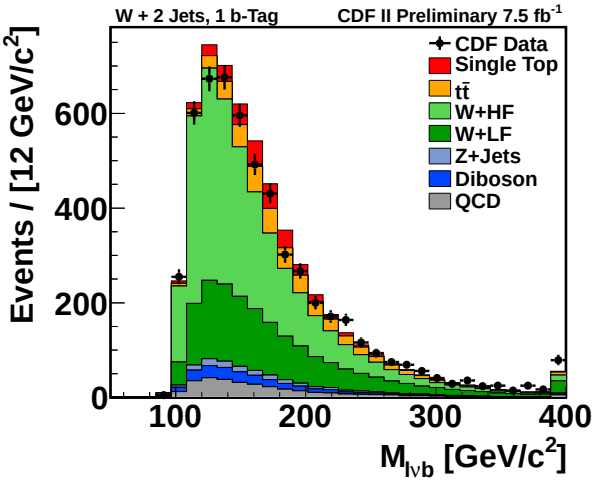

Figure 2: Comparison between data and expectations for the reconstructed lepton-neutrino-b-jet distribution.

2.1. Lepton plus jets analysis

The lepton plus jets analysis was performed using a data sample corresponding to 7.5 fb^{-1} of integrated luminosity. The final event selection requires a high transverse momentum isolated charged lepton and large missing transverse energy, consistent with the leptonic decay of a W boson, and either two or three energetic jets. Based on the type of lepton identified, events are grouped into two mutually exclusive categories called the tight lepton category and the extended muon category. At least one of the jets is required to be identified as originating from a b-quark (b-tagged), using a secondary vertex b-tagger.

The selected sample is divided into 4 statistically independent signal regions, depending on the number of jets (2 or 3) and the number of identified b-jets (1 or 2 b-tags). Events originating from s-channel single top quark frequently populate the two-tags signal region while t-channel and Wt events predominantly populate the one-tag region. A number of kinematic variables are studied and the most powerful for distinguishing signal from background are used as inputs to build the NN discriminant. Individual NNs (optimized separately) are used for each of the 4 signal regions and for each of the two identified lepton categories.

The single top quark events are modeled with POWHEG [8], a next-to-leading order accuracy in the strong coupling α_s generator. This provides an improved model with respect to the leading-order model used in the previous analyses. Figure 2 shows the reconstructed top quark mass distribution based on the charged lepton, the reconstructed neutrino, and the b-tagged jet for the 2-jets 1-b-tag subsample. This is one of the best discriminating variables.

The final NN uses from 11 to 14 variables, depending on the subsample and is trained using s-channel events as signal in the 2-jets, 2-b-tag signal region, while t-channel events are used as signal in the training process for all the other subsamples. We use 6 separate background templates for the training, in the ratios of the estimated yields. We validate data-background agreement in background-enriched control samples.

Figure 3 shows the comparison of the data with the sum of the predictions of the NN output for the combined 2-jet and 3-jet signal regions. The inset shows a magnification of the region for the NN discriminant from 0.8 to 1.0, where the single top quark contribution is larger.

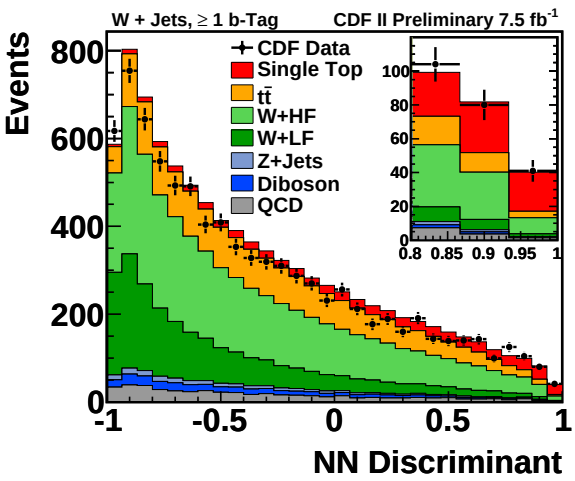

Figure 3: NN output distribution for the combined 2-jet and 3-jet signal regions. The inset shows a magnification of the NN region from 0.8 to 1.

The measurement of the total single top quark cross section σ_{s+t+Wt} is performed using a maximum likelihood fit to the binned NN output distributions [9]. Assuming the SM ratio among the s-channel, t-channel, and Wt production rates we measure $\sigma_{s+t+Wt} = 3.04^{+0.57}_{-0.53}$ pb, for a top quark mass of 172.5 GeV/c^2.

To extract the single top quark cross sections for s-channel and t-channel + Wt production separately,

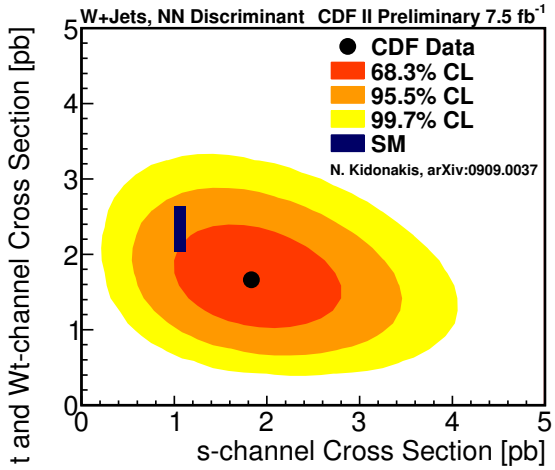

Figure 4: Results of the two-dimentional fit for σ_s and σ_{t+Wt}.

we take advantage of the differences in jet multiplicity and final-state kinematics between the two production mechanisms. We assume a uniform prior-probability density distribution in the two-dimensional plane (σ_s; σ_{t+Wt}) and determine the cross sections that maximize the posterior-probability density distribution. The t-channel and Wt process are combined as they share the same final-state topology. The best-fit cross sections correspond to $\sigma_s = 1.81^{+0.63}_{-0.58}$ pb and $\sigma_{t+Wt} = 1.66^{+0.53}_{-0.47}$ pb, with a correlation factor of -24%. The best fit values, the credibility regions of 68, 95 and 99 %, and the SM predictions are shown in Figure 4.

2.2. Extraction of the CKM matrix element $|V_{tb}|$

To extract the CKM matrix element $|V_{tb}|$ we exploit the direct proportionality between the single top quark production cross section and the square of the magnitude of $|V_{tb}|$: $|V_{tb}|^2 = |V_{tb}^{SM}|^2 \times \frac{\sigma^{obs}}{\sigma^{SM}}$. In this formula we use the SM predicted single top production cross section σ^{SM}, the measured value σ^{obs} and the nearly unit value of $|V_{tb}^{SM}|^2$ obtained in SM.

Under the assumption that the top quark mostly decays to a W boson and b quark but no assumption on number of families or CKM unitarity we obtain a 95% Bayesian credibility level lower limit of $|V_{tb}| > 0.78$ and extract $|V_{tb}| = 0.95 \pm 0.09$ (stat + syst) $\pm$ 0.05 (theory) which includes the theoretical uncertainty on the predicted production rate, which is not part of the cross section posterior [9].

Figure 5 shows the joint posterior distribution of all the independent subsamples as a function of $|V_{tb}|^2$.

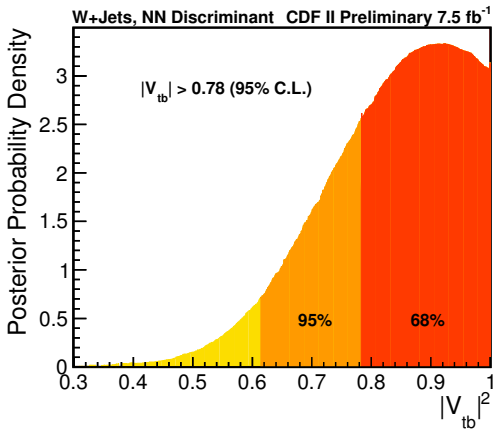

Figure 5: Posterior probability distribution as a function of $|V_{tb}|^2$.

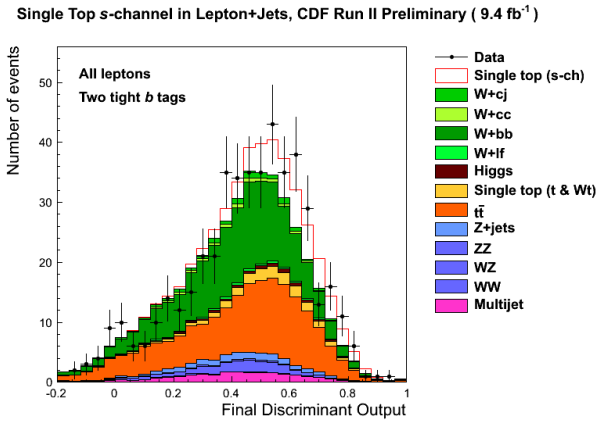

Figure 6: Lepton plus jets analysis for s-channel search: final discriminant for the events in the TT b-tags category.

2.3. *Missing transverse energy plus jets analysis*

The MET plus jets analysis was performed on a sample orthogonal to the one used by the previously described lepton plus jets analysis, using the full CDF Run II data set corresponding to 9.5 fb^{-1} of integrated luminosity. This analysis recovers events with non-reconstructed electrons and muons and Ws decaying to tau leptons which decay hadronically, and adds approximately a 30% of acceptance to the lepton plus jets selection strategy.

Events are selected to have large missing transverse energy and at least two energetic jets. A multivariate b-tagging algorithm which was developed for the Higgs search is used in this analysis. The selected sample is divided into six subsamples depending on the number of jets (2 or 3) and the number and quality of b-tags: only one tightly tagged jet and no other tag (1T), two tightly tagged jets (TT), and one tightly tagged jet and one loosely tagged jet (TL). Several NNs are used to discriminate against QCD background, vector bosons plus jets, and $t\bar{t}$ processes, for s- and t-channels separately. A NN^{sig}_{s+t} final discriminant is used to simultaneously separate both s- and t-channel signal processes from the remaining background. To measure the single top quark production cross section, a combined likelihood is formed, which is the product of Poisson probabilities for each bin of the NN^{sig}_{s+t} discriminants for the six signal regions. We find $\sigma_{s+t} = 3.53^{+1.25}_{-1.16}$ pb and a $|V_{tb}|$ lower limit of 0.63 is obtained at the 95% credibility level [10].

2.4. *Combination of lepton plus jets and missing energy plus jets analyses*

The results of the two analyses described above (lepton plus jets and MET plus jets) are combined by taking the product of their likelihoods and simultaneously varying the correlated uncertainties.

The combined measurement results in an electroweak single top quark production cross section $\sigma_{s+t} = 3.04^{+0.57}_{-0.53}$ pb, consistent with the SM prediction [10]. A t-channel single top quark production cross section, considering the s-channel production as a background constrained to the theoretical prediction, is measured to be $\sigma_t = 1.65^{+0.38}_{-0.36}$ (stat+syst) pb. From the posterior probability density on $|V_{tb}|^2$, a 95% credibility level lower limit of $|V_{tb}| > 0.84$ is obtained.

3. Evidence for s-channel single top production

A different analysis optimized for s-channel search has been performed in the lepton plus jets [11] and MET plus jets final states [12], using the full CDF Run II dataset corresponding to 9.5 fb^{-1} of integrated luminosity.

Since the final state of this process is the same one as in the search for a Higgs boson (H) produced in association with a W boson this analysis benefited from tools and techniques developed in the framework of the CDF analysis searching for the WH process [13], but with a final discriminant optimized for the single top s-channel. The same leptonic selection as in the inclusive single top quark search is used, while there are differences in the jet selection strategy and exactly two central jets are required. Since events with two b-jets provide the most sensitivity for the s-channel process, the

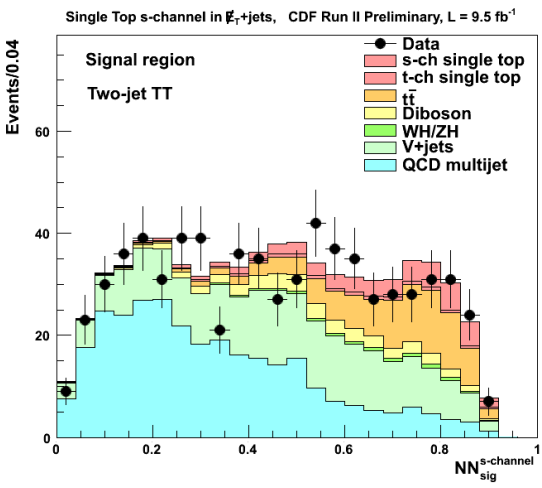

Figure 7: MET plus jets analysis for s-channel search: final discriminant for the events in the 2 jets, TT b-tags category.

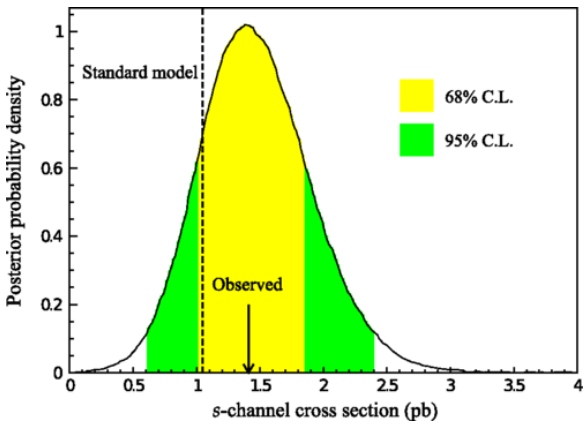

Figure 8: Lepton plus jets analysis: posterior probability density distribution for the s-channel cross section measurement, with the SM prediction shown as the vertical dashed line.

identification of b-jets has been improved with a more efficient, neural network based b-jet tagging algorithm which exploits several observables that can be used to discriminate b-jets from other jets. By applying these tagging requirements to each jet in an event, we construct several non overlapping tagging categories subsamples obtained with different combinations of tight (T) and loose (L) b-tags: TT, TL, T, and LL. For the 2 b-tag categories, the category with the highest signal-to-background ratio is chosen if an event satisfies more than one category; for the single-tag category, one jet of the event is required to be tight tagged, and the other one is untagged. We use a set of artificial NNs trained having the s-channel as signal, separately for each tagging category and for different lepton categories, using different input variables. Figure 6 shows the NN output for the lepton plus jets analysis for the TT b-tagging category.

In the MET plus jets analysis the jet selection strategy is the same as in the inclusive single top quark search. Events are classified based on the number of jets and b-tagging category: T, TL, and TT. Figure 7 shows the NN output for the MET plus jets analysis for the 2 jets TT b-tagging category.

Background and signal models for both analyses are the same as in the inclusive search for single top quark. We use a binned-likelihood technique to extract the single top quark s-channel cross section from the NN output distribution. We measure $\sigma_s = 1.41^{+0.44}_{-0.42}$ pb for the lepton plus jets analysis and $\sigma_s = 1.12^{+0.61}_{-0.57}$ pb for the MET plus jets analysis, assuming that the top quark mass is 172.5 GeV/c^2. The posterior probability density of the s-channel cross section in the lepton plus jets analysis is shown in Figure 8.

As a cross check, we also measured the single top s-channel cross section in separate lepton and b-tagging categories. As an example, Figure 9 shows the s-channel cross section measured in each lepton category. The results in each independent measurement are all consistent with each other and with the theory prediction.

The lepton plus jets and MET plus jets analyses are combined by taking the product of their likelihoods and simultaneously varying the correlated uncertainties. We measure a combined s-channel cross section $\sigma_s = 1.36^{+0.37}_{-0.32}$ pb. The background-only p-value corresponds to a signal significance of 4.2 standard deviations [12].

By combining this result with the D0 analysis [14], the Tevatron experiments reported the observation of the single top s-channel production at 6.3 standard deviations, measuring $\sigma_s = 1.29^{+0.26}_{-0.24}$ [15].

Figure 10 shows a summary of the Tevatron s-channel cross section measurements.

4. Conclusion

Single top quark production was observed by CDF and D0 in 2009. Since then, single top quark measurements have been refined and benefited from the tools and analysis techniques developed for the search of the Higgs boson at CDF. Single top quark s-channel production evidence was announced by CDF in 2014, and the combination with the previous D0 result led to the s-channel observation. All single top quark measurements are in agreement with each other and with the SM prediction. Now the CDF single top quark program is

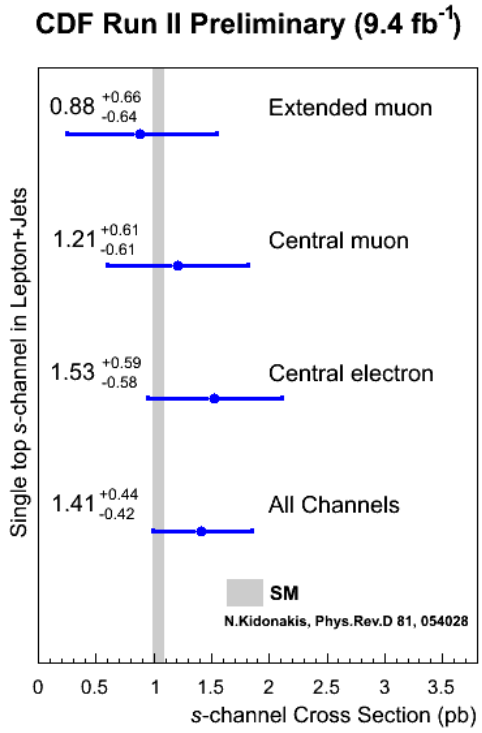

Figure 9: Lepton plus jets analysis: s-channel cross section in each lepton category. The grey vertical line is the theoretical cross section value with uncertainties.

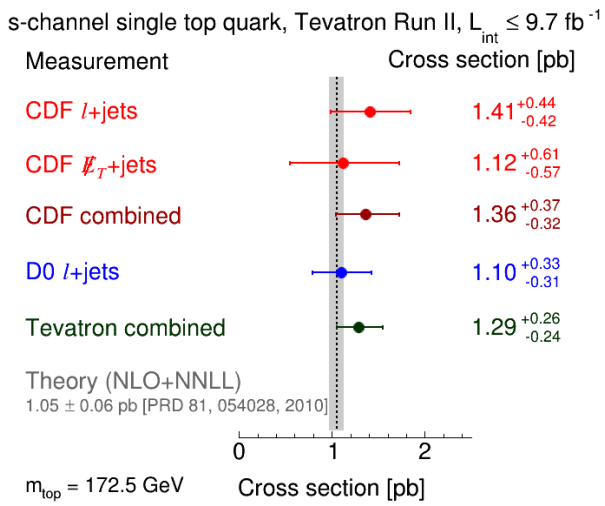

Figure 10: Summary of the Tevatron single top s-channel cross section measurements. The grey vertical line is the theoretical cross section value with uncertainties.

almost complete. The Tevatron combined $s+t$ channels production cross section will be released soon. Three years after the end of Run II CDF continues producing valuable top physics results.

References

[1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, (1995) 2626; S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, (1995) 2632.

[2] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **103**, (2009) 092001; T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, (2009) 092002.

[3] N. Kidonakis, Phys. Rev. **D81**, (2010) 054028.

[4] N. Kidonakis, Phys. Rev. **D83**, (2011) 091503.

[5] N. Kidonakis, Phys. Rev. **D82**, (2010) 054018.

[6] T. M. P. Tait and C.-P. Yuan, Phys. Rev. **D63**, (2000) 014018.

[7] J. Alwal *et al.*, Eur. Phys. J. C **49**, (2007) 791.

[8] S. Alioli, P. Nason, C. Oleari, and E. Re, J. High Energy Phys. 09 (2009) 111.

[9] T. Aaltonen *et al.* (CDF Collaboration), accepted by Phys. Rev. Lett., arXiv:1407.4031.

[10] T. Aaltonen *et al.* (CDF Collaboration), submited to Phys. Rev. Lett., arXiv:1410.4909v2.

[11] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **112**, (2014) 231804.

[12] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **112**, (2014) 231805.

[13] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **109**, (2012) 111804.

[14] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **726**, (2013) 656.

[15] T. Aaltonen *et al.* (CDF and D0 Collaborations), Phys. Rev. Lett. **112**, (2014) 231803.