CDF/PUB/TOP/PUBLIC/9369
FERMILAB-CONF-08-191-E

Monte Carlo Simulations for Top Pair and Single Top Production at the Tevatron

U. Husemann(1) on behalf of the CDF and DØ Collaborations

(1) *Deutsches Elektronen-Synchrotron, Platanenallee 6, 15738 Zeuthen, Germany*

Summary. — Monte Carlo (MC) simulations are indispensable tools for top quark physics, both at the current Tevatron collider and the upcoming Large Hadron Collider. In this paper we review how the Tevatron experiments CDF and DØ utilize MC simulations for top quark analyses. We describe the standard MC generators used to simulate top quark pair and single top quark production, followed by a discussion of methods to extract systematic uncertainties of top physics results related to the MC generator choice. The paper also shows the special MC requirements for some example top properties measurements at the Tevatron.

PACS `14.65.Ha` – Top quarks.
PACS `2.70.Uu` – Applications of Monte Carlo methods.

1. – Introduction

With more than 3.5 fb^{-1} of data recorded per Tevatron experiment to date (Summer 2008), top quark physics has entered an era of precision measurements. The latest combination of top quark mass measurements from the CDF and DØ collaborations shows a relative uncertainty of only 0.8% [1], and multivariate analysis techniques are being used in many analyses, notably the recent evidence for the electroweak production of single top quarks [2, 3]. For all these endeavors, Monte Carlo (MC) simulations are key tools, for estimating the signal acceptance and the amount of background, and for evaluating systematic uncertainties.

2. – Standard Monte Carlo Generators at the Tevatron

2˙1. *Top Pair Production.* – The Tevatron collaborations follow different approaches for their default MC generators for top quark pair production, as summarized in table I. Both collaborations use leading order (LO) MC generators with parton showering (PS), Pythia v6.2 [4] in the case of CDF, and ALPGEN v2.1 [5] (using Pythia v6.3 for PS) for DØ. As both MC generators are LO generators, the generated $t\bar{t}$ production cross

TABLE I. – *Comparison of the default MC generators used for top quark pair production by the CDF and DØ collaborations.*

	CDF	**DØ**
Generator	PYTHIA v6.2	ALPGEN v2.1 with PYTHIA v6.3
Process	$q\bar{q}, gg \to t\bar{t}$	$q\bar{q}, gg \to t\bar{t} +$ 0–2 partons
Parton Distribution Functions	CTEQ5L [6]	CTEQ6L [7]
Tunes	Tune A, W/Z p_T	—
Multiple Collisions	Minimum Bias (PYTHIA)	Zero Bias (Data)

sections must be scaled to the theoretical cross section. CDF has opted for a well established MC generator that has been carefully validated against and tuned to the Tevatron data. DØ uses a fairly recent MC generator with additional functionality compared to PYTHIA, including exact matrix elements for $2 \to n$ processes with a matching procedure between partons generated via matrix elements and parton showers, a more recent set of parton distribution functions (PDFs), and $t\bar{t}$ spin correlations. The collaborations have also chosen different ways of simulating multiple collisions in the same bunch crossing. CDF overlays minimum bias events generated with PYTHIA, with the same calibration constants applied as for the data events. DØ obtains a good description of beam and instrumental backgrounds when overlaying events recorded with a zero bias (*i.e.* random) trigger in data.

2·2. *Single Top Production.* – The Tevatron collaborations have reported evidence for electroweak single top quark production [2,3]. At $\sqrt{s} = 1.96$ TeV single top quarks are predominantly produced via the t-channel process $qb \to q't$ and the s-channel process $q\bar{q}' \to t\bar{b}$. Theoretical predictions of the production cross section are available at next-to-leading order (NLO), see *e.g.* ref. [8].

The LO kinematics of the s-channel process are unchanged by NLO corrections; therefore it is sufficient to employ a LO MC generator to simulate the s-channel process and to scale the obtained cross section to the NLO expectation. On the other hand there are important corrections to the t-channel kinematics from the $2 \to 3$ process $qg \to q't\bar{b}$, where the gluon in the initial state produces a $b\bar{b}$ pair. The soft and collinear regime of this process is well modeled by the $2 \to 2$ process available in LO MC with PS, using b quark PDFs; however, for large transverse momenta of the final state $\bar{b}$ quark, $p_T(b)$, the $2 \to 3$ process needs to be considered explicitly.

Both CDF and DØ use a procedure in which the $2 \to 2$ and the $2 \to 3$ processes are generated separately and the phase space overlap between the two processes is removed by hand [9]. The NLO cross section σ_{NLO} as a function of $p_T(b)$ is then constructed as follows:

$$\sigma_{\mathrm{NLO}} = K \cdot \sigma_{2\to 2,\mathrm{PYTHIA}}\big|_{p_T(b)<p_T^0} + \sigma_{2\to 3}\big|_{p_T(b)\geq p_T^0}\,, \tag{1}$$

which is illustrated in fig. 1 and implies the following steps:

- The $2 \to 2$ and the $2 \to 3$ processes are generated separately, utilizing the `MadEvent` generator [12] in the case of CDF, and the SingleTop generator [9] at DØ.

- The $2 \to 2$ process is scaled up by a factor of K such that the cross section for both processes matches the NLO cross section.

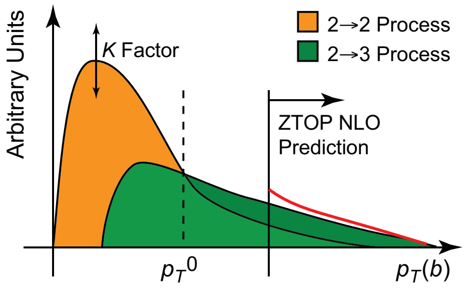

Fig. 1. – Constructing the top t-channel cross section as a function of the $\bar{b}$ quark transverse momentum from the 2 ⟶ 2 and the 2 ⟶ 3 process [9-11].

- The soft and the hard regime are separated by a simple cut on $p_T(b)$ at p_T^0, *i.e.* all events with $p_T(b) < p_T^0$ are taken from the $2 \longrightarrow 2$ process and all events with $p_T(b) > p_T^0$ are taken from the $2 \longrightarrow 3$ process. The value of p_T^0 is chosen to ensure a smooth transition of the differential cross section as a function of $p_T(b)$ between the soft and the hard regime (around 20 GeV/c) [10, 11].
- The full event kinematics are then validated against NLO predictions from the ZTOP code [8]. As shown in fig. 2, both collaborations find good agreement between the MC samples and the NLO prediction.

The above procedure represent a "pragmatic" solution that is sufficient for the current precision of the single top measurements. In the future it is desirable to treat t-channel single top production consistently in a full NLO MC simulation, *e.g.* in the framework of MC@NLO [13].

3. – Monte Carlo Simulations and Systematic Uncertainties

CDF and DØ have established their (separate) ways of assigning systematic uncertainties to top physics results over the course of Tevatron Runs I and II. A significant fraction of the systematic uncertainties is related to the MC simulation. For example, for the most recent Tevatron top mass combination, out of the total relative uncertainty of 0.8%, 0.3% can be attributed to MC related effects [1]. In the following we will review the standard treatment of MC related systematic uncertainties by CDF and DØ and comment on recent progress on this subject.

3˙1. *Signal Model.* – The two main all-purpose LO MC generators with PS, PYTHIA [4] and HERWIG [14], have different hadronization models and different tunings for the underlying event (*i.e.* interactions of partons inside the colliding proton and antiproton other than the ones involved in the hard scattering process). The treatment of initial state radiation (ISR) and final state radiation (FSR) in PYTHIA and HERWIG is similar, but not identical.

CDF extracts MC model uncertainties for many observables from a comparison of PYTHIA and HERWIG. For the first years of Tevatron Run II, the DØ collaboration did not employ the HERWIG MC generator and therefore did not compare HERWIG to PYTHIA. Instead, DØ evaluates uncertainties due to the b quark fragmentation model, a major source of MC model uncertainty in b rich top events.

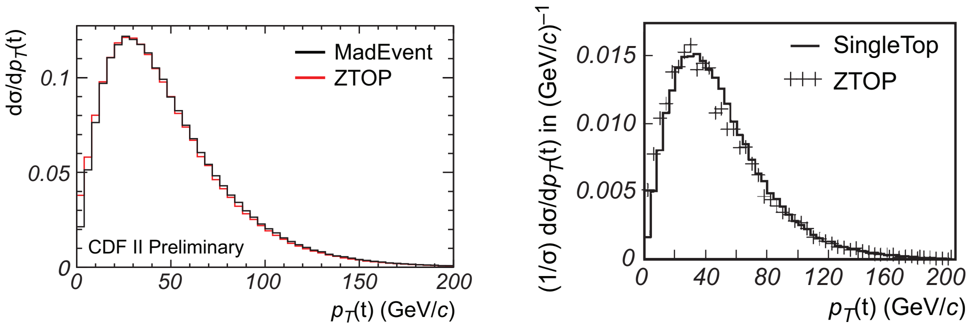

Fig. 2. – Comparison of top quark transverse momentum distributions in t-channel single top production. Left: CDF comparison of `MadGraph` and ZTOP. Right: DØ comparison of SingleTop and ZTOP (after ref. [9]).

CDF and DØ also follow different approaches to evaluating uncertainties due to ISR and FSR. CDF estimates the uncertainty from special MC samples in which the Pythia parameters controlling the amount of ISR and FSR (Λ_{QCD} and Q^2 scales) are varied. This procedure is sensitive to differences in the soft radiation. DØ extracts the systematic uncertainties due to ISR and FSR from reweighting the jet multiplicity spectrum in the MC simulation (ALPGEN $t\bar{t}$+ 0–2 partons) in a control region to the multiplicity spectrum observed in data. With this procedure, DØ mainly probes differences in hard radiation.

Both collaboration evaluate uncertainties due to the PDF choice in similar ways. For any given observable, *e.g.* the acceptance for $t\bar{t}$ events, the uncertainty is obtained from two main sources, the change in the observable when varying eigenvectors in the space of the PDF fit parameters and the difference in the observable when calculated using CTEQ [7] versus MRST [15] PDF sets.

3˙2. *Jet Energy Scale.* – In many top analyses, uncertainties in the calibration of the jet energy scale (JES) are among the leading systematic uncertainties. The CDF procedure for JES corrections is detailed in ref. [16]. The JES correction is obtained *e.g.* from dijet and photon-jet balance, and the uncertainty of the correction is estimated, among others, from a comparison of Pythia and HERWIG. Note that it is important to avoid double counting of MC uncertainties in the signal MC model and the JES, as they are both derived in part from Pythia–HERWIG comparisons.

Top quark decays are a significant source of b quarks. In general, jets containing b quarks have a different JES than light parton jets. The uncertainty of the b JES correction is larger than for light jets, because there are additional uncertainties due to fragmentation and color flow, and the small b quark jet sample sizes make a calibration in data less precise.

In recent years the JES uncertainty has been reduced significantly in top mass analyses by employing an *in situ* JES calibration technique [17]. In the $t\bar{t} \to \ell\nu b\, qqb$ decay channel, the mass of the top quarks and the mass of the hadronically decaying W boson are extracted simultaneously. The comparison of the measured W mass with the world average W mass is used to calibrate the JES. In this case, the JES uncertainty reduces to a residual uncertainty that covers the dependence on the jet kinematics.

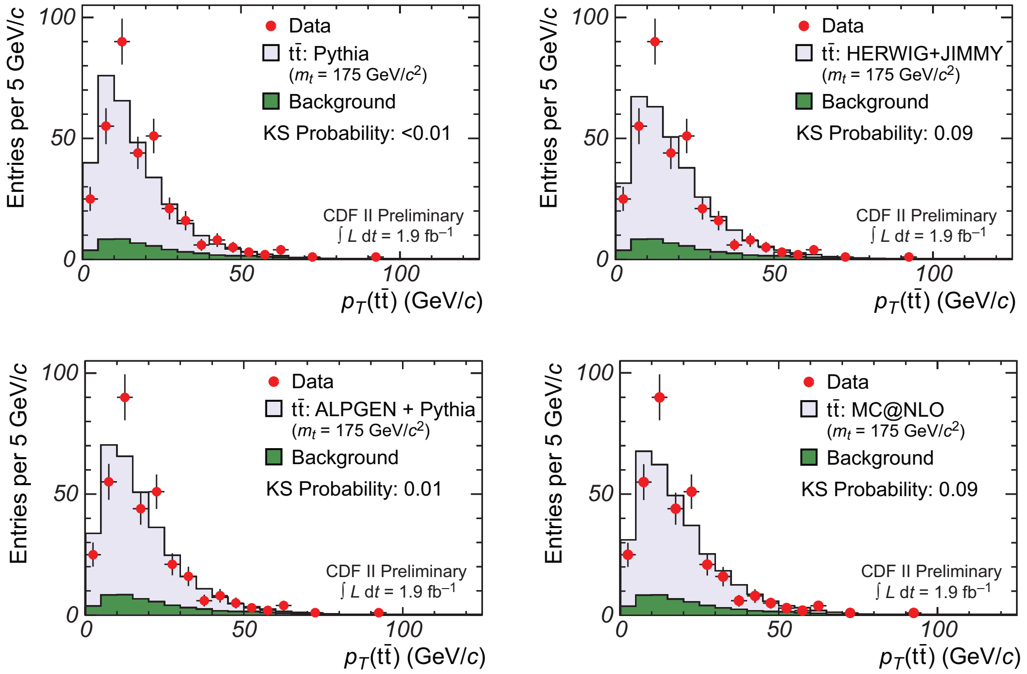

Fig. 3. – Comparison of the $t\bar{t}$ transverse momenta in 1.9 fb^{-1} of data with different MC generators. Top left: PYTHIA. Top right: HERWIG + JIMMY. Bottom left: ALPGEN + PYTHIA ($t\bar{t}$ + jets). Bottom right: MC@NLO.

3˙3. *Recent Developments.* – The Tevatron collaborations are in the process of revisiting their procedures for assigning systematic uncertainties. This becomes important especially for precision measurements of the top quark mass. Starting from two joint workshops in 2007, CDF and DØ are working towards a better understanding of each other's procedures and a common approach for assigning systematic uncertainties.

One example of a long-standing disagreement between Tevatron data and MC simulations lies in the transverse momentum distribution of $t\bar{t}$ pairs. Fig. 3 shows a comparison of CDF data with different MC generators. None of the generators shows good agreement with the data. The HERWIG-based generators (HERWIG + JIMMY [18], MC@NLO) show better agreement with the data than the PYTHIA-based generators (PYTHIA, ALPGEN + PYTHIA). Generators that include higher orders (ALPGEN + PYTHIA, MC@NLO) show the same disagreement as LO generators. These findings indicate that the source of the disagreement lies in the PS part of the MC generators. Note that the influence of the disagreement on the top quark mass uncertainty is small.

4. – Monte Carlo Simulations for Special Analyses

In many analyses of top quark properties, the standard MC tools for $t\bar{t}$ production do not provide all the features required for the analysis. This is especially true for CDF's main generator PYTHIA v6.2. In this section we will discuss examples of Tevatron top properties analyses with special MC requirements.

4˙1. *W Helicity in Top Quark Decays.* – Measurements of the helicity of W bosons in top quark decays are tests of the $V{-}A$ structure of the tWb vertex. The recent Tevatron analyses [19, 20] measure the W helicity via the distribution of the angle $\cos\theta^*$ between the charged lepton (or down-type quark) and the top boost direction in the W rest frame.

To compare the acceptances for $t\bar{t}$ processes with different W helicities, the $\cos\theta^*$ distributions obtained from ALPGEN + PYTHIA (DØ) or PYTHIA (CDF) are reweighted according to helicities other than the standard model (SM) prediction. For linearity checks of the W helicity measurement, CDF has also used the `MadEvent` generator, which changes the underlying left-handed and right-handed couplings rather than the direct observable $\cos\theta^*$. Systematic uncertainties of the W helicity measurement are obtained from comparing the SM helicities predicted by different MC generators, *e.g.* PYTHIA versus HERWIG in the case of CDF.

4˙2. *Flavor Changing Neutral Currents in the Top Sector.* – Flavor changing neutral current (FCNC) interactions of top quarks are strongly suppressed in the SM, so that any FCNC signal would be an indication for physics beyond the SM. CDF has searched for the FCNC $t \to Zq$ [21] in $t\bar{t}$ decays, and DØ has searched for single top production via the FCNC $q \to tg$ [22].

The collaborations have chosen different approaches to determining the FCNC signal acceptances. CDF obtains the $t \to Zq$ decay from PYTHIA and reweights the resulting isotropic $\cos\theta^*$ distribution (defined analogously to sect. **4**˙1) according to the expectations. This approach stays close to experimental observables, which can also be seen by the result, which is a limit on the branching fraction for the $t \to Zq$ decay. The DØ analysis utilizes the CompHEP MC generator [23] to modify the top quark couplings. The DØ approach is closer to theoretical calculations, and consequently, the result is quoted as a limit on the qtg coupling.

4˙3. *Fraction of $t\bar{t}$ Events Produced in Gluon-Gluon Fusion.* – NLO QCD calculations predict that at the Tevatron, 85% of all $t\bar{t}$ pairs are produced via $q\bar{q}$ annihilation and 15% are produced via gg fusion. CDF has recently published first experimental tests of this prediction [24]. One of the CDF analyses measures the fraction of $t\bar{t}$ produced via gg fusion by training an artificial neural network on the kinematics of $t\bar{t}$ production and decay. While the main sensitivity of this method comes from the production angles and velocities of the top quarks, the dependence on $t\bar{t}$ spin correlations, which is encoded in the decay kinematics, also contributes to the sensitivity.

The standard CDF MC tool PYTHIA is missing two important ingredients for this analysis. PYTHIA as a LO MC generator only generates 5% $gg \to t\bar{t}$, and it does not include $t\bar{t}$ spin correlations. Therefore the analysis makes use of the HERWIG generator, also using a CDF-internal extension of HERWIG named GGWIG that allows the user to adjust the relative fraction of $gg \to t\bar{t}$. The systematic uncertainty of the measurement due to the difference between HERWIG and PYTHIA is small; however, the systematic difference between HERWIG and the NLO generator MC@NLO results in a 20% loss in sensitivity, see fig. 4. This is one of the largest systematic uncertainties and indicates the importance of NLO effects for the analysis.

4˙4. *Charge (Forward Backward) Asymmetry.* – Both Tevatron collaboration have recently published measurements of the charge asymmetry in $t\bar{t}$ production (which translates into a forward-backward asymmetry if one assumes CPT invariance) [25, 26]. The asymmetry vanishes at LO, and NLO calculations predict an asymmetry of 4–5% [27], which is confirmed by the MC@NLO generator, which shows a 3.8% asymmetry.

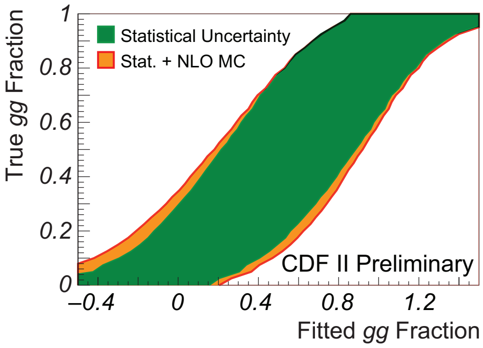

Fig. 4. – Feldman-Cousins (FC) bands at 68% confidence level for the measurement of the fraction of $t\bar{t}$ production via gg fusion. The FC band for statistical uncertainties plus uncertainties due to the comparison of HERWIG with MC@NLO is compared with the FC band for statistical uncertainties only. At the fitted gg fraction of −7.5%, the limit on the true gg fraction is deteriorated by approximately 20%.

The net asymmetry obtained from LO MC generators is zero; however, there may be asymmetries generated from color flow for specific kinematic configuration. This is illustrated in fig. 5, which shows a comparison of the asymmetry as predicted in PYTHIA, ALPGEN, and MC@NLO as a function of the rapidity difference between the t and the $\bar{t}$ quark. The asymmetries are also subject to NLO corrections. For example, for $t\bar{t}$ + jet production, a LO asymmetry of −8% is reduced to −1.5% [28]. In summary, there is currently no single MC generator that covers all features of the charge (FB) asymmetry.

5. – Summary and Outlook

The CDF and DØ collaborations have entered an era of precision top quark physics. While more and more sophisticated analysis techniques are being employed, both collaborations favor a "conservative" and "pragmatic" approach towards MC simulations:

- Prefer well validated MC tools (mainly based on PYTHIA) over "cutting edge" MC technology.
- Prefer data-driven methods whenever available.
- Adapt and re-use well understood MC samples, *e.g.* by applying reweighting techniques.

In addition there are often technical reasons for the above approach. Integrating new MC generators into the existing CDF and DØ frameworks is time intensive and error prone, despite close collaboration with the authors of the MC codes. With the current computing and person power available at the Tevatron, the cycle of software integration, event generation, validation, and tuning of a new MC generator can easily take many months.

The Tevatron collaborations are actively working on common approaches to top MC generators. The final goal is a common treatment of MC-related systematic uncertainties, primarily for the top quark mass measurements, but also for other top analyses.

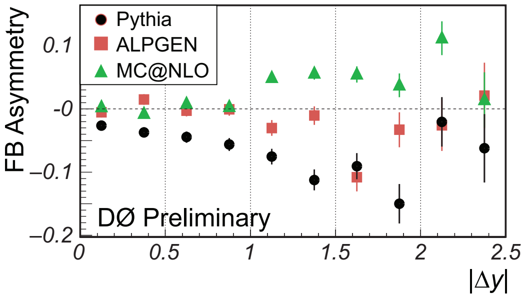

Fig. 5. – Charge (forward-backward) asymmetry as a function of the rapidity difference between the t and the $\bar{t}$ quark for the Pythia, ALPGEN, and MC@NLO generators.

* * *

It is a pleasure to thank the organizers of TOP2008 for an extremely productive workshop at the perfect venue. This work was supported by the Helmholtz Association under contract number VH-NG-400.

REFERENCES

[1] Tevatron Electroweak Working Group, arXiv:0803.1683 [hep-ex].
[2] Abazov V. M. *et al.* (DØ Collaboration), *Phys. Rev. Lett.*, **98** (2007) 181802.
[3] CDF Collaboration, CDF Public Notes 9217, 9221, 9223, 9252.
[4] Sjöstrand T. *et al.*, *Comput. Phys. Commun.*, **135** (2001) 238.
[5] Mangano M. L. *et al.*, *JHEP*, **07** (2003) 001.
[6] Lai H. L. *et al.* (CTEQ Collaboration), *Eur. Phys. J. C*, **12** (2000) 375.
[7] Pumplin J. *et al.* (CTEQ Collaboration), *JHEP*, **07** (2002) 012.
[8] Sullivan Z., *Phys. Rev. D*, **70** (2004) 114012.
[9] Boos E. E. *et al.*, *Phys. Atom. Nucl.*, **69** (2006) 1317.
[10] Lück J., FERMILAB-MASTERS-2006-01.
[11] Gerber C. E. *et al.* (TeV4LHC-Top and Electroweak Working Group), arXiv:0705.3251 [hep-ph].
[12] Maltoni F. and Stelzer T., *JHEP*, **02** (2003) 027.
[13] Frixione S., Nason P. and Webber B. R., *JHEP*, **03** (2003) 007.
[14] Corcella G. *et al.*, *JHEP*, **01** (2001) 010.
[15] Martin A. D., Roberts R. G., Stirling W. J. and Thorne R. S., *Eur. Phys. J. C*, **4** (1998) 463.
[16] Bhatti A. *et al.*, *Nucl. Instrum. Meth. A*, **566** (2006) 375.
[17] Abulencia A. *et al.* (CDF Collaboration), *Phys. Rev. D*, **73** (2006) 032003.
[18] Butterworth J. M., Forshaw J. R. and Seymour M. H., *Z. Phys. C*, **72** (1996) 637.
[19] Abazov V. M. *et al.* (DØ Collaboration), *Phys. Rev. Lett.*, **100** (2008) 062004.
[20] CDF Collaboration, CDF Public Notes 9114, 9144, 9215.
[21] Aaltonen T. *et al.* (CDF Collaboration), arXiv:0805.2109 [hep-ex].
[22] Abazov V. M. *et al.* (DØ Collaboration), *Phys. Rev. Lett.*, **99** (2007) 191802.
[23] Pukhov A. *et al.*, hep-ph/9908288.
[24] CDF Collaboration, CDF Public Note 8811.
[25] Abazov V. M. *et al.* (DØ Collaboration), *Phys. Rev. Lett.*, **100** (2008) 142002.
[26] Aaltonen T. *et al.* (CDF Collaboration), arXiv:0806.2472 [hep-ex].
[27] Kühn J. H. and Rodrigo G., *Phys. Rev. Lett.*, **81** (1998) 49.
[28] Dittmaier S., Uwer P. and Weinzierl S., *Phys. Rev. Lett.*, **98** (2007) 262002.