Nuclear Instruments and Methods in Physics Research A 775 (2015) 27–33

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Acceleration of matrix element computations for precision measurements ☆

O. Brandt [a,*], G. Gutierrez [b], M.H.L.S. Wang [b], Z. Ye [c]

[a] II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
[b] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[c] University of Illinois at Chicago, Chicago, IL 60607, USA

ARTICLE INFO

Article history:
Received 24 October 2014
Received in revised form
17 November 2014
Accepted 17 November 2014
Available online 25 November 2014

Keywords:
Matrix element
Monte Carlo integration
Low-discrepancy sequences
Hadron collider
Top quark

ABSTRACT

The matrix element technique provides a superior statistical sensitivity for precision measurements of important parameters at hadron colliders, such as the mass of the top quark or the cross-section for the production of Higgs bosons. The main practical limitation of the technique is its high computational demand. Using the concrete example of the top quark mass, we present two approaches to reduce the computation time of the technique by a factor of 90. First, we utilize low-discrepancy sequences for numerical Monte Carlo integration in conjunction with a dedicated estimator of numerical uncertainty, a novelty in the context of the matrix element technique. Second, we utilize a new approach that factorizes the overall jet energy scale from the matrix element computation, a novelty in the context of top quark mass measurements. The utilization of low-discrepancy sequences is of particular general interest, as it is universally applicable to Monte Carlo integration, and independent of the computing environment.

1. Introduction

The matrix element (ME) technique [1] is a powerful tool in experimental particle physics, especially at hadron colliders, as it provides a superior statistical sensitivity in the extraction of important parameters of the standard model. This sensitivity is achieved by taking into account the full topological and kinematic information in a given event, and determining the probabilities $P_{\rm sig}$ and $P_{\rm bkg}$ for observing each event, assuming respective signal and background hypotheses in the respective ME probabilities $|\mathcal{M}_{\rm sig}|^2$ and $|\mathcal{M}_{\rm bkg}|^2$. In the context of searches for new physics, these probabilities can be used to construct the most powerful test statistic $Q \equiv P_{\rm sig}/P_{\rm bkg}$ according to the Neyman–Pearson lemma [2]. An advantage of the ME technique is that it calculates $P_{\rm sig}$ and $P_{\rm bkg}$ *ab initio*, in contrast to multivariate methods. Furthermore, $P_{\rm sig}$ depends directly on the physical parameter of interest in a specific theoretical framework.

The ME technique was first suggested by Kondo [1] and pioneered in the context of experimental particle physics at the Tevatron in the measurement of the mass of the top quark m_t [3], in the determination of the helicity of the W boson [4], as well as for the first evidence for production of single top quarks [5,6]. Since then, the ME technique has been used in several analyses, for example in searches for the Higgs boson at the Tevatron [7] and at the LHC [8]. Recently, a general framework for the ME technique, named MadWeight [9], has become available.

Despite its superior statistical sensitivity, the ME technique is not widely applied because of its high computational demand. For example, to perform a previous measurement of m_t using 3.6 fb^{-1} of integrated luminosity [10] by the D0 Collaboration, about two million CPU-hours were required on a single core of the 64 bit XEON E5-2620 CPU, with a clock rate of 2 GHz, and a 64 bit computation. In this paper, we present two approaches that were successfully applied to reduce the computational demand of the ME technique by two orders of magnitude. First, we utilize low-discrepancy sequences (LDS) for numerical Monte Carlo (MC) integration, in conjunction with a dedicated estimator of the numerical uncertainty, which is a novelty in the context of the ME technique. Second, we factorize the overall jet energy scale (JES) from the ME computation, which was never done before in the context of m_t measurements using an *in situ* JES calibration. The use of LDS is generally applicable to MC integration. In particular, this approach is not hardware-specific, i.e., it can be used on, e.g., a graphical processing unit.

We present our results using the example of the recent measurement of the top quark mass [11], the single most precise measurement of this parameter, yielding $m_t = 174.98 \pm 0.58({\rm stat}) \pm 0.49({\rm syst})$ GeV.

☆ FNAL report number: Fermilab-Pub-14/416-E.

* Corresponding author. Now at: Kirchhoff-Institut für Physik, Universität Heidelberg, Heidelberg, Germany.
E-mail address: obrandt@fnal.gov (O. Brandt).

http://dx.doi.org/10.1016/j.nima.2014.11.063
0168-9002/

This measurement was performed in lepton+jets final states[1] with the full sample of $p\bar{p}$ collision data from the Fermilab Tevatron Collider at $\sqrt{s}=1.96$ TeV, corresponding to 9.7 fb^{-1} of integrated luminosity. The computational demand arises not so much from the number of events recorded in $p\bar{p}$ collisions, but rather from the number of the simulated MC events which are used for the calibration of the method and for the evaluation of systematic uncertainties. D0's previous measurements of m_t [10] and of the difference $\Delta m = m_t - m_{\bar{t}}$ [12], both using 3.6 fb^{-1} of integrated luminosity, were also performed with the ME technique.

This paper is structured as follows. We begin with a brief review of our previous implementation of the ME technique for the measurement of m_t [10] in 3.6 fb^{-1} of data. This analysis applies several approaches to reduce the computational demand that potentially have general interest. We follow with a discussion of our latest implementation of the ME technique, which provides further reduction in the computational demand through the use of LDS for the MC integration, presented in Section 3, and through factorization of the scale factor for jet energies $k_{\rm JES}$ from the ME computation, discussed in Section 4. Finally, we present in Section 5 the validation of our latest implementation of the ME technique with pseudo-experiments (PE), comprised of MC events fully simulated in the D0 detector, and conclude in Section 6. The MC simulations are described in Ref. [11].

2. Previous implementation of the matrix element technique

The extraction of m_t with the ME technique is performed with a likelihood that uses per-event probability densities (PD) defined by the ME of the processes contributing to the observed events. Assuming two non-interfering contributions from $t\bar{t}$ and W+jets production, the per-event PD is given by

$$P_{\rm evt} = A(\vec{x})[fP_{\rm sig}(\vec{x}; m_t, k_{\rm JES}) + (1-f)P_{\rm bkg}(\vec{x}; k_{\rm JES})] \quad (1)$$

where the observed signal fraction f, m_t, and the overall multiplicative factor $k_{\rm JES}$ adjusting the energies of jets after their default jet energy scale calibration are parameters to be determined from data. The $\vec{x}$ denotes the measured jet and lepton four-momenta, and $A(\vec{x})$ accounts for acceptance and efficiencies. The function $P_{\rm sig}$ represents the PD for $t\bar{t}$ production, and $P_{\rm bkg}$ refers to the PD for W+jets production.

In general, the measured set $\vec{x}$ will not be identical to the set of corresponding partonic variables $\vec{y}$ because of finite detector resolution and parton hadronization. Their relationship is described by a transfer function $W(\vec{x}, \vec{y}, k_{\rm JES})$. The densities $P_{\rm sig}$ and $P_{\rm bkg}$ are calculated through a convolution of the differential partonic cross-section, $d\sigma(\vec{y})$, with $W(\vec{x}, \vec{y}, k_{\rm JES})$ for the final-state partons and the PD for the initial-state partons, $f(q_i)$, where the q_i are the momenta of the colliding partons. This is done by integrating over all possible parton states that lead to $\vec{x}$:

$$P_{\rm sig}(\vec{x}; m_t, k_{\rm JES}) = \frac{1}{\sigma_{t\bar{t},{\rm obs}}(m_t, k_{\rm JES})} \int \sum d\sigma(\vec{y}, m_t) d\vec{q}_1 d\vec{q}_2 \times f(\vec{q}_1) f(\vec{q}_2) W(\vec{x}, \vec{y}; k_{\rm JES}). \quad (2)$$

The sum extends over all possible flavor combinations of the initial-state partons. The longitudinal-momentum parton density functions (PDF) $f(q_{i,z})$ are taken from the CTEQ6L1 set [13], while the dependencies $f(q_{i,x})$, $f(q_{i,y})$ on transverse momenta follow those PD obtained from the PYTHIA simulation [14,15]. The factor $\sigma_{t\bar{t},{\rm obs}}(m_t, k_{\rm JES})$, defined as the total cross-section for $t\bar{t}$ production in $p\bar{p}$ collisions to be observed in the detector, ensures that $A(\vec{x})P_{\rm sig}$ is normalized to unity. The differential cross-section, $d\sigma(\vec{y}, m_t)$, is calculated using the leading order (LO) ME for the process $q\bar{q} \to t\bar{t}$ [16,17].

The calculation in Eq. (2) at LO involves 24 integration variables associated with the two initial-state partons and the six partons in the final state. The directions of the four jets and the charged lepton in (η, ϕ) space are well-measured, and are therefore represented by ten δ-functions. After accounting for these δ-functions, and imposing energy-momentum conservation through four additional δ-functions, ten integration variables remain.

The integration in Eq. (2) is performed numerically using the MC integration method of Ref. [18]. The pseudo-random numbers for the MC integration are generated with RANLUX [19] in a $[0,1]^{10}$ hypercube, and then transformed to the ranges of the integration variables. Importance sampling [20] is utilized to reduce the computational demand of the integration. Furthermore, we perform a Jacobian transformation of the nominal ten integration variables to variables where prior information is either known or can be easily obtained. This prior information is then used in importance sampling. The optimized integration variables are m_{W^+}, m_{W^-}, $m_t, m_{\bar{t}}$, $q_{1,x}, q_{1,y}, q_{2,x}, q_{2,y}$, $\rho = E_q/(E_q + E_{\bar{q}'})$ for the quarks from $W \to q\bar{q}'$ decay in the LO picture where E represents the particle's energy and the energy (curvature) of the electron (muon track) κ.

To integrate over m_t and $m_{\bar{t}}$, random numbers are generated according to expected Breit–Wigner distributions for each given m_t hypothesis. The constraint of $M_W = 80.4$ GeV for the *in situ* JES calibration is imposed by integrating over W boson masses using a Breit–Wigner prior. For the integration over $q_{i,x}$ and $q_{i,y}$, the ME is sampled in transverse momentum $p_T^{q_i}$ according to the distribution predicted in MC simulations, and uniformly in ϕ_i^q. To integrate over κ, random numbers are generated according to the corresponding part of the transfer function, which is defined as the probability to obtain the measured κ_x value, given a value κ_y at the parton level.

Importance sampling in ten bins is employed for the integration over ρ. The MC integration is performed iteratively with an increasing number of samplings of the integral per iteration, where each iteration uses the probability distribution in ρ from the previous one as input for importance sampling.

There are 24 possible jet-parton assignments that are summed with weights based on their consistency with b-tagging information.[2] Typically, two and sometimes four or six jet-parton assignments numerically dominate the final result for $P_{\rm sig}$. To identify them, we perform a pre-integration step, where we calculate $P^i_{\rm sig}$ for each jet-parton assignment i, until a relative numerical precision of 10% is reached, or the integral is sampled $2^{14} = 16{,}384$ times. The numerical precision of those jet-parton assignments with $P^i_{\rm sig}$ within 2% of the maximal $P^i_{\rm sig}$ value is further refined until the desired precision has been achieved, or the integral is sampled $2^{24} = 16{,}777{,}216$ times. For all other assignments $P^i_{\rm sig}$ obtained in the pre-integration step is kept.

The differential partonic cross-section for $P_{\rm bkg}$ is calculated similarly to $P_{\rm sig}$, i.e., applying MC integration and the same transfer function $W(\vec{x}, \vec{y}; k_{\rm JES})$, however using the LO $W+4$ jets ME implemented in VECBOS [21]. Here, the initial-state partons are all assumed to have no transverse momentum $p_T = 0$.

To extract m_t and $k_{\rm JES}$, we calculate $P_{\rm sig}$ and $P_{\rm bkg}$ on a grid in $(m_t, k_{\rm JES})$ with spacings of (1 GeV, 0.01). A likelihood function $\mathcal{L}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N; m_t, k_{\rm JES}, f)$ is constructed at each grid point from the product of the individual $P_{\rm evt}$ values for the measured quantities $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N$ of the selected events, and the signal fraction f is determined by maximizing $\mathcal{L}$ at that grid point. The likelihood

[1] The lepton+jets final states aim at selecting the $p\bar{p} \to t\bar{t} \to W^+ bW^- \bar{b} \to \ell^+ \nu b q \bar{q}' \bar{b}$ and its charge conjugate process, where t and b denote respectively top and bottom quarks, $W^\pm$ is the W boson, $\ell^\pm$ stands for charged leptons, and ν represents a neutrino.

[2] We identify jets from b quarks through the use of a multivariate algorithm, as discussed in Ref. [11].