Eur. Phys. J. C 44, 233–241 (2005)
Digital Object Identifier (DOI) 10.1140/epjc/s2005-02362-2

THE EUROPEAN PHYSICAL JOURNAL C

Towards precision determination of the top-quark mass from $\mathrm{M}_{b\ell}$ distribution in semi-leptonic decays

M.L. Nekrasov[a]

Institute for High-Energy Physics, 142284 Protvino, Russia

Received: 28 December 2004 / Revised version: 26 May 2005 /
Published online: 13 September 2005 –

Abstract. We explore a possibility of optimization of the method of determination of the top-quark mass from the $M_{b\ell}$ distribution in semi-leptonic decays $t \to b\,\ell\nu$ at the LHC and a future linear collider (LC). We discover that the systematic and statistical errors of M_t determination can be diminished if considering the high moments over the distribution. In the case of LHC this allows one to reduce the errors by more than a factor of two, and in the case of LC to approach the precision expected by studying the threshold scan of the total cross section $e^+e^- \to t\bar{t}$.

1 Introduction

The precision determination of the top-quark mass is one of the major research problems at next-generation colliders [1–5]. Being a fundamental parameter of the standard model (SM), the top-quark mass is tightly constrained by quantum-level calculations with other fundamental parameters. This enables one to test the SM and/or to select the probable scenario of its extension on the basis of an independent M_t measurement.

A considerable progress in this direction is expected at Tevatron and LHC, where the accuracy of the M_t determination is anticipated of about 1–2 GeV [1]. At LHC in view of the copious production of top quarks, to increase accuracy, decreasing the systematic errors is crucial. An analysis of [1] shows that the most promising method from the point of view of optimization of the errors is based on the investigation of a distribution over the invariant mass of the observable products of semi-leptonic decays $t \to bW \to b\,\ell\nu$; more precisely of the isolated lepton ℓ and the $\mu^+\mu^-$ pair indicating a J/ψ meson produced from the decay of the b quark [6]. In this channel one can obtain experimentally very clean final states. Correspondingly, the systematic error of the M_t measurement can be made low. The evaluation made by Monte Carlo (MC) modeling gives 0.6–0.8 GeV at a statistical error of about 1 GeV for four years of LHC operation [6]. This result is recognized as the best among others obtained by various methods [1].

In the case of a future linear collider (LC) [2–5] the most promising method for precise M_t determination is based on the investigation of the threshold scan of the total cross section $e^+e^- \to t\bar{t}$. In this region the form and height of the cross section are very sensitive to the mass of the top quark. This gives an opportunity to determine M_t with very high accuracy. A serious difficulty in this approach is a precise theoretical calculation of the behavior of the cross section in the vicinity of the threshold, which becomes additionally complicated because of the resonant effects due to the strong t–$\bar{t}$ interaction. Major progress in the calculations was made by way of the summation of QCD contributions via solving the Lippmann–Schwinger equation for the Green function describing the $t\bar{t}$ production [7]. At present the theoretical value of the top mass determined by this method is estimated as 100–200 MeV [8,9], with an experimental error of about 20 MeV [10].

Alternate methods of M_t determination are based on the reconstruction of top quark decay events. Their basic features are common to LC and the hadron colliders, but at LC the precision is anticipated to be better. Thus, for example, the systematic error of M_t determination by direct reconstruction of $t\bar{t}$ events in e^+e^- collisions at $\sqrt{s} = 500$ GeV is expected [11] to be equal to 340 and 250 MeV in hadronic and semi-leptonic channels, respectively, with statistical errors of about 100 MeV for 1–2 years of LC operation [12]. Since far above the threshold one can expect very high precision of the necessary theoretical calculations, the resultant errors should be close to that expected by studying the threshold scan of the cross section. This promising anticipation again excites a question about the precision of the top mass determination by the method of [6], but this time in the LC case. Actually this method in the LC case has been discussed initially in [13] (see also the review [3]), but the errors have not been determined. So the prospect of this method at LC is still not known practically.

In this article we clear up this question. In contrast to [6], however, we consider the full reconstructed jet of the b quark instead of the J/ψ or $\mu^+\mu^-$ pair only. Such an approach has been considered in [13], and partially in [14]. We follow it by keeping in mind that the $M_{b\ell}$ distribution in any case does emerge in a certain stage of the analysis.

[a] e-mail: nekrasov@th1.ihep.su

So from the very beginning the analysis can be made in terms of the data converted to the form of $M_{b\ell}$ distribution. (Of course, the systematic errors that arise in the course of converting the data must be taken into account.) An obvious advantage of this approach is the possibility to consider the data in a uniform fashion in both the LHC and LC cases. Moreover, this allows us in a simple way to explore the possibility of optimizating the algorithm used for the extraction of the top mass from the data. The elaboration of the latter problem is actually the major purpose of the present article.

In the next section we detail the statement of the problem. In Sects. 3 and 4 we discuss a model for the calculation of the errors. The parameters of the model are fixed in Sect. 5 and the quantitative outcomes are determined. In Sect. 6 we discuss the theoretical uncertainty, and in Sect. 7 we discuss the results.

2 Statement of the problem

We consider the processes

$$e^+e^-\,(q\bar{q},\,gg) \to t\bar{t} \to bW\,bW \to b\ell\nu\;bq_1q_2 \to \{b\text{-jet}+\ell\} + \{3\text{ jets}\}\,, \qquad (1)$$

with the b-jet, isolated lepton $\ell = \{e, \mu\}$, and a neutrino that is invisible in the final states coming from one t quark, and the remaining three jets coming from another t quark. In the experiment these states are registered, and measured in a distribution

$$F(q) = \frac{1}{\sigma}\,\frac{\mathrm{d}\sigma}{\mathrm{d}q}\,. \qquad (2)$$

Here σ is the cross-section of the process (1), $q \equiv M_{b\ell}$ is the reconstructed invariant mass of the system $\{b\text{-jet}+\ell\}$.

We simulate the results of the experiment under the following suppositions. First we suppose that there is a satisfactory method for extracting signal from the data. Actually this means the existence of a satisfactory model for the background processes that survive after setting of kinematic cuts.[1] Furthermore, we describe the signal in the Born approximation, identifying the b-jet with the b quark. Finally, on the basis of the results of [6], we disregard the effects of the finite width of the top quarks. The latter assumption means that $\sigma^{-1}\,\mathrm{d}\sigma/\mathrm{d}q$ is equal to $\Gamma_{b\ell\nu}^{-1}\,\mathrm{d}\Gamma_{b\ell\nu}/\mathrm{d}q$, where $\Gamma_{b\ell\nu}$ is a partial width of the decay $t \to b\ell\nu$. (Thus the distribution F becomes process-independent.)

Direct calculation gives the following formula for the distribution of the partial width:

$$\frac{\mathrm{d}\Gamma_{b\ell\nu}}{\mathrm{d}q^2} = \frac{3\,G_F|V_{tb}|^2}{4\sqrt{2}\,\pi^2}\,\frac{\Gamma_{W\to\ell\nu}\,M_W}{M_t^3}\Biggl\{q^2 - \Lambda^2 - M_W^2 + \left(\frac{\Lambda^2 - M_W^2}{2} - q^2\right)\ln\frac{(\Lambda^2-q^2)^2 + M_W^2\Gamma_W^2}{M_W^4 + M_W^2\Gamma_W^2} + \frac{(\Lambda^2-q^2)(q^2+M_W^2) + M_W^2\Gamma_W^2}{M_W\Gamma_W} \times \left[\arctan\left(\frac{\Lambda^2-q^2}{M_W\Gamma_W}\right) + \arctan\left(\frac{M_W}{\Gamma_W}\right)\right]\Biggr\}. \qquad (3)$$

[1] The set of the cuts and the background processes in the LHC case have been discussed in [6]. In the LC case that has been done in [11, 12]. At this stage we do not take the kinematic cuts manifestly into account but do so on deriving the quantitative outcomes.

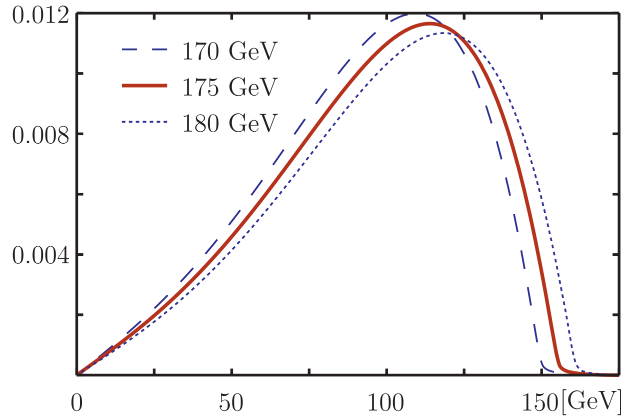

Fig. 1. The distribution $F(q) = \Gamma^{-1}\,\mathrm{d}\Gamma/\mathrm{d}q$, $q \equiv M_{b\ell}$, at $M_t =$ 170, 175, and 180 GeV

Here $\Lambda^2 = M_t^2 - M_W^2$, Γ_W is the total and $\Gamma_{W\to\ell\nu}$ is the partial width of the W boson, and we neglect the masses of the lepton ℓ and the b quark.[2] In this approximation $\Gamma_{W\to\ell\nu} = 2/9\,\Gamma_W$ and q ranges between 0 and M_t. Figure 1 shows the distribution $F(q)$ defined by formula (3) at $M_t = 170, 175, 180$ GeV. From Fig. 1 the dependence of $F(q)$ on M_t is obvious. So, by comparing the experimental distribution with a set of theoretical curves one can determine, in principle, the experimental value of M_t.

In a practical respect, however, it is convenient to compare integrated parameters of the distributions. For instance, in [6] the M_t was extracted from the mean value (position of the maximum) of the Gaussian distribution approximating the measured distribution. [13,14] determined M_t by the first moment $\langle\, q\,\rangle$ over the distribution. In the present article we consider a method of M_t determination by the higher moments

$$\langle q^n \rangle = \int_0^M \mathrm{d}q\; q^n F(q)\,. \qquad (4)$$

Here M is a fixed quantity close to M_t, see below for details. In fact this method means the matching of the experimental distribution $q^n F(q)$ with the corresponding theoretical distribution which depends on the parameter M_t.

As we will see below, the insertion of the q^n factor will significantly increase the precision of the M_t determination. Eventually this can be checked by a quantitative analysis. Nevertheless some hints on this can be seen a priori. Really, with increasing n the moment $\langle q^n\rangle$ becomes increasingly

[2] The influence of the mass of the b quark is noticeable at very small q, but this region is inessential when considering the moments over the distribution.

dependent on the behavior of $F(q)$ in a region located between the position of its maximum and a large-q tail where $F(q)$ almost vanishes. (More precisely, by the tail we mean the range $\Lambda < q < M_t$, where in the limit $\Gamma_W = 0$ the distribution identically vanishes for kinematic reasons.) Further, in the mentioned region the behavior of $F(q)$ in the greatest measure is sensitive to the value of M_t, which is seen from Fig. 1. As a result, with increasing n the sensitivity of $\langle q^n \rangle$ with respect to M_t is increasing. That is why one can expect increased precision for the value of M_t extracted from the higher moments.

Now let us dwell on the details of the definition (4). The point of the discussion is the upper limit in the integral. We set it to M instead of conventional M_t, meaning the upper bound of the region allowed by kinematics, because M_t is also a parameter, which is subject to determination. To avoid an inconvenience, we use for the upper limit a certain predetermined value M fixed close to M_t. Simultaneously we adjust the normalization of the distribution $F(q)$ so as to satisfy the equality $\langle 1 \rangle = 1$. The moments $\langle q^n \rangle$ at $n \geq 1$ after this redefinition practically do not change (for values of n that are not large) in view of the almost vanishing F in the tail at large q.

So, we define the experimentally measured value of the top-quark mass as a solution to the equation

$$\langle q^n \rangle = \langle q^n \rangle_{\rm exp} \,. \qquad (5)$$

Here in the right-hand side the moment is determined (at a given M) on the basis of the experimental data, and that in the left-hand side on the basis of the theoretical distribution, which depends on the parameter M_t. Let, for a given n, a solution to (5) be $M_t = M_{t(n)}$. Then the error of the solution can be determined as

$$\Delta M_{t(n)} = \Delta \langle q^n \rangle_{\rm exp} \Bigg/ \frac{{\rm d}\langle q^n \rangle}{{\rm d}M_t}\Bigg|_{M_t = M_{t(n)}} . \qquad (6)$$

Our aim is to estimate $\Delta M_{t(n)}$ and find an optimal value of n which would minimize $\Delta M_{t(n)}$. Since by virtue of (3) the derivative ${\rm d}\langle q^n \rangle / {\rm d}M_t$ is known, the problem is reduced to the determination of the statistical and systematic errors, the components of the experimental error $\Delta \langle q^n \rangle_{\rm exp}$.

3 Statistical errors

We determine the statistical errors of the moments on the supposition that the data averaged over ensemble are described by $F(q) = \Gamma_{b\ell\nu}^{-1}\, {\rm d}\Gamma_{b\ell\nu}/{\rm d}q$ with $\Gamma_{b\ell\nu}$ determined by formula (3) at $M_t = 175\,{\rm GeV}$.

Let δq_i be the size of a bin, within which the i-th element of the distribution is measured, and let $\overline{N_i}$ be the number of events counted in this bin on average. Then

$$F(q_i)\delta q_i = \overline{N_i} \, / \, \overline{N} \,, \qquad (7)$$

where $\overline{N}$ is the total number of events counted in all the bins on average. Further, we do not distinguish between $\overline{N}$ and $N = \sum_i N_i$, the total number of events counted in all bins in the given experiment. The experimentally measured n-th moment is

$$\langle q^n \rangle_{\rm exp} = \sum_i q_i^n \, \frac{N_i}{N} \,. \qquad (8)$$

By virtue of (7) the averaged experimental moment $\overline{\langle q^n \rangle}_{\rm exp}$ is found by formula (4). Since N_i is distributed by the Poisson law with parameter $\overline{N_i}$, the variance of $\langle q^n \rangle_{\rm exp}$ is

$$D\langle q^n \rangle_{\rm exp} = \sum_i q_i^{2n} \, \frac{\overline{N_i}}{N^2} \equiv \frac{1}{N} \langle q^{2n} \rangle \,. \qquad (9)$$

Formula (9) implies the following estimation for the statistical error:

$$\Delta^{\rm stat} \langle q^n \rangle_{\rm exp} = \sqrt{\frac{1}{N} \langle q^{2n} \rangle} \,. \qquad (10)$$

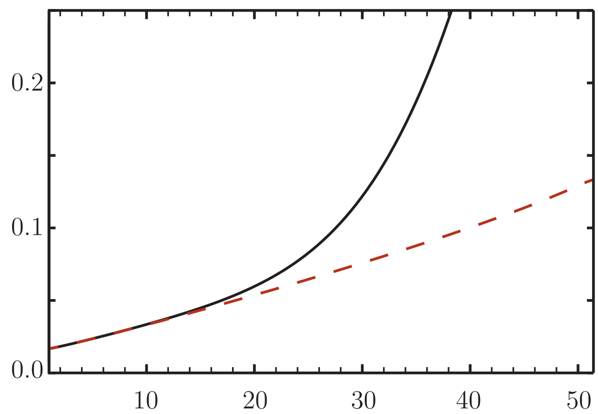

Fig. 2. The ratio $\Delta^{\rm stat}\langle q^n \rangle_{\rm exp}/\langle q^n \rangle_{\rm exp}$ depending on n ($M_t = 175\,{\rm GeV}$, $N = 4000$). The continuous curve represents the results described by formula (10). The dashed curve represents the results obtained by the method of effective moments

To give an idea of the behavior of $\Delta^{\rm stat}\langle q^n \rangle_{\rm exp}$, we present in Fig. 2 (continuous curve) the ratio $\Delta^{\rm stat}\langle q^n \rangle_{\rm exp}/\langle q^n \rangle_{\rm exp}$ calculated at $N = 4000$ (corresponds to the LHC case, see Sect. 5). It is seen from the figure that, with increasing n, the ratio grows. This is explained by the shift (to the right) of the position of the maximum of $q^n F(q)$ from the position of the maximum of $F(q)$, where the statistics are largest. As a result the statistical reliability of $\langle q^n \rangle_{\rm exp}$ decreases. Another important property of the ratio is the change of the mode of the growth beginning with $n \approx 15$. This is explained by the emergence of a noticeable contribution from the large-q tail in $q^n F(q)$. The latter property is illustrated by the set of the curves represented by Fig. 3.

In fact the emergence of a noticeable contribution from the tail is an undesirable effect since in the tail the uncertainty from the background is comparable with the signal process. To avoid this difficulty one can correct the definition of the moments by introducing a cut-off in the integral in (4). The position of the cut-off should be determined so as to isolate the second (unphysical) peak in the tail of $q^n F(q)$ but simultaneously to keep as much statistical significance as possible of the sample events. It is clear that the optimal

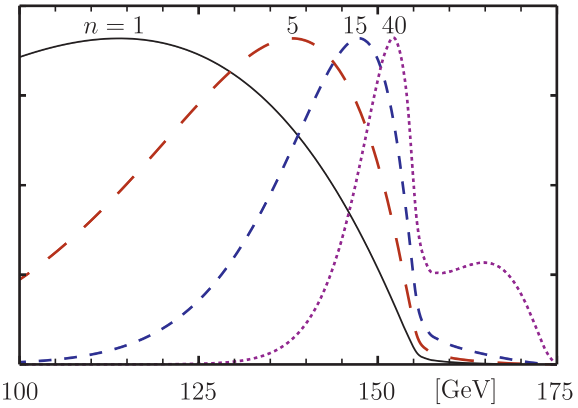

Fig. 3. The shape of the function $q^n F(q)$ at $M_t = 175\,\text{GeV}$, $n = 1, 5, 15, 40$ (in arbitrary normalization)

cut-off should be placed in the neighborhood of a local minimum between the two peaks of $q^n F(q)$ (if the second peak appears). From Fig. 3 it is seen that at $n \approx 40$ the sought point is about two half-widths to the right of the maximum of $q^n F(q)$. So a simplified algorithm for the cut-off may be determined by setting $\Lambda_n = \min\{q_{n\,\text{extr}} + 2\Gamma_{n\,\text{right}}, M\}$, where $q_{n\,\text{extr}}$ is the position of the maximum of $q^n F(q)$ and $\Gamma_{n\,\text{right}}$ is the half-width from the right. Thus we come to the following definition of the effective moments:

$$\langle q^n \rangle^{\text{eff}} = \int_0^{\Lambda_n} \mathrm{d}q\; q^n F(q) \bigg/ \int_0^{\Lambda_n} \mathrm{d}q\; F(q)\,. \qquad (11)$$

In the experimentally determined effective moments the cut-off must be the same. Ultimately $\Delta^{\text{stat}} \langle q^n \rangle^{\text{eff}}_{\text{exp}}$ is defined by formula (10) with $\langle q^{2n} \rangle$ replaced by $\langle q^{2n} \rangle^{\text{eff}}$ with the introduction of the cut-off Λ_n instead of Λ_{2n}. The latter anomalous prescription follows immediately from the derivation of the formula (10).

The behavior of $\Delta^{\text{stat}} \langle q^n \rangle^{\text{eff}}_{\text{exp}} / \langle q^n \rangle^{\text{eff}}_{\text{exp}}$ is shown by the dashed curve in Fig. 2. It is seen from the figure that the transition to the effective moments implies no noticeable modification up to $n \approx 15$, while at larger n the growth of the ratio stabilizes. A similar behavior is observed in the basic formalism (without the transition to the effective moments) in the limit $\Gamma_W \to 0$, when the large-q tail identically vanishes.

4 Systematic errors

Proceeding to the systematic errors it is necessary first to clarify their origin. For this purpose we use the analysis of [6] of the errors of the $M_{J/\psi\ell}$ distribution simulated with the PYTHIA and/or HERVIG event generators. For the main sources of the systematic errors [6] found the uncertainties in the b-quark fragmentation (including the final-state radiation) and uncertainties in the background processes. It is clear that the same sources should be the main ones when solving the inverse problem, the determination of the $M_{b\ell}$ distribution from the $M_{J/\psi\ell}$ distribution, which is considered virtually to be the data. In the LC case we expect the same sources of systematic errors.

On this basis we consider first the error resulting from the uncertainty in the b-quark fragmentation. For brevity we call this the type I error. At the level of the $M_{b\ell}$ distribution it appears as the uncertainty in the bin number in which the number of events, N_i, is measured. In the continuous case this error becomes the uncertainty Δq in the determination of the q variable.

Suppose that Δq is sufficiently small. Then, neglecting the nonlinear effects, we have

$$\Delta^{\text{sys I}} \langle q^n \rangle_{\text{exp}} = \int_0^M \mathrm{d}q\; \left[q^n F(q)\right]'\, \Delta q\,. \qquad (12)$$

Here the prime means a derivative with respect to q. The systematic error I of the effective moment $\langle q^n \rangle^{\text{eff}}_{\text{exp}}$ is estimated similarly, by replacing the upper bound M by Λ_n and then dividing the result by the normalization factor, as in formula (11). The normalization factor itself should be the same, as it controls the total number of events that are taken into consideration when determining the effective moment.

We carry out the determination of Δq with the aid of the following reasoning. First we note that the invariant mass q^2 is actually the doubled scalar product of the 4-momenta of the b quark and the lepton ℓ. So in the laboratory frame it can be represented as $q^2 = E_b\, K$, where E_b is the energy of the b quark, and K is a factor proportional to the energy of the lepton ℓ. (Additionally K includes a dependence on angular variables which, however, is relatively weak.) Furthermore, by calculating the differential we obtain $\Delta q = \frac{1}{2}\,(\Delta E_b/E_b + \Delta K/K)\, q$, where ΔE_b and ΔK are the corresponding errors. A more precise estimation is determined by the sum of the quadratures. Thus we arrive at a linear dependence with a certain coefficient r,

$$\Delta q = r\, q\,, \qquad r = \frac{1}{2}\left[\left(\frac{\Delta E_b}{E_b}\right)^2 + \left(\frac{\Delta K}{K}\right)^2\right]^{1/2}. \qquad (13)$$

The systematic error arising after subtraction of the background we call the type II error. This appears in the absolute value of the distribution function. So it should be described as an additive contribution δF to the function F. Correspondingly, we obtain the following formula for the type II error of the moments:

$$\Delta^{\text{sys II}} \langle q^n \rangle_{\text{exp}} = \int_0^M \mathrm{d}q\; q^n\, \delta F(q)\,. \qquad (14)$$

The type II error of the effective moments $\langle q^n \rangle^{\text{eff}}_{\text{exp}}$ is defined by a similar formula with the modifications listed below (12).

It is reasonable to determine $\delta F(q)$ under the supposition that it vanishes at the boundaries of phase space, and that when passing from small q to large q it changes sign only once. The simplest form of a function satisfying these requirements is a third-degree polynomial,

$$\delta F = h\, q\, (q - M/2)\, (q - M). \qquad (15)$$

The parameter h in (15) describes the amplitude of the error and is subject to further determination.

5 Numerical results

We assign the following values for the parameters with global meaning:

$$M_W = 80.4 \text{ GeV}, \ \Gamma_W = 2.1 \text{ GeV}, \ M_t = M = 175 \text{ GeV}. \quad (16)$$

The remaining parameters N, r, and h depend on the conditions under consideration. Recall that N means the volume of the representative sample of events, the parameter r characterizes the error in the invariant mass of the $b\ell$ system, and h describes the error arising after the subtraction of the background processes.

With reference to the LHC case, we fix the parameters N, r and h on the basis of the results of [6]. Since in that work the $M_{J/\psi\ell}$ distribution was determined at $N = 4000$ (with kinematic cuts and for four years of LHC operation), in our investigation we set this value for N, as well. We fix the parameters r and h based on the properties of the $M_{J/\psi\ell}$ distribution and the direct results derived in [6] from these properties. First we use the estimation $\Delta^{\text{sys}}\langle M_{J/\psi\ell}\rangle = +0.3/-0.4\,\text{GeV}$ and the result derived from this $\Delta M_t = +0.6/-0.8\,\text{GeV}$. Considering the framework of our investigation, and the latter quantity as the uncertainty of the input parameter M_t, we get by direct calculation $\Delta^{\text{sys}}\langle q\rangle_{\text{exp}} = +0.47/-0.62\,\text{GeV}$. By comparing this with $\Delta^{\text{sys}}\langle M_{J/\psi\ell}\rangle$ we obtain an energy scale factor of 1.6, which describes the spreading of the $M_{J/\psi\ell}$ distribution when converting it to the $M_{b\ell}$ distribution. Using the mentioned factor, from $\Delta^{\text{sys II}}\langle M_{J/\psi\ell}\rangle_{\text{exp}} \lesssim 0.15\,\text{GeV}$ [6] we further derive an estimation $\Delta^{\text{sys II}}\langle q\rangle_{\text{exp}} \lesssim 0.24\,\text{GeV}$. From this result and the relations (14) and (15) we get $h \simeq 1.7 \times 10^{-10}\,\text{GeV}^{-4}$. (Hereinafter we take the upper bounds as the estimations.)

Knowing $\Delta^{\text{sys}}\langle q\rangle_{\text{exp}}$ and $\Delta^{\text{sys II}}\langle q\rangle_{\text{exp}}$, we immediately get $\Delta^{\text{sys I}}\langle q\rangle_{\text{exp}} \simeq +0.41/-0.57$. From here and the formula (12) it follows that $r \simeq 0.004$–0.006; we use the average value $r = 0.005$. It is worth noticing that the same estimation for r follows from formula (13) when taking into consideration the 1% precision of the determination of the energy of the b jets expected at LHC [1], and additionally neglecting $\Delta K/K$ compared to $\Delta E_b/E_b$.

Now, as we know N, r, h, we can calculate $\Delta^{\text{stat}}\langle q^n\rangle_{\text{exp}}$ and $\Delta^{\text{sys I,II}}\langle q^n\rangle_{\text{exp}}$ at any n. So we calculate $\Delta^{\text{stat}} M_{t(n)}$ and $\Delta^{\text{sys I,II}} M_{t(n)}$. The dependence on n of these errors is shown by the solid lines on Figs. 4–6. The dashed lines show the errors obtained by the method of effective moments. (The break of the slope in the dashed line in Fig. 5 is explained by the change of the sign in the integral in formula (12) appearing after introducing the cut-off Λ_n.) In Table 1 we present the numerical results at some n and the summed-quadrature errors $\Delta^{\text{sys}} M_{t(n)}$ and $\Delta M_{t(n)}$. It

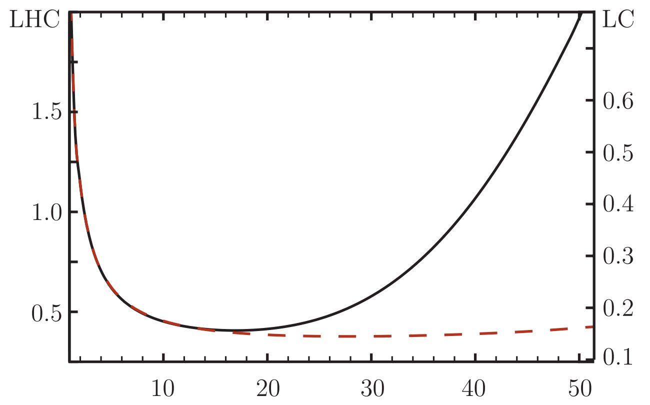

Fig. 4. The statistical error $\Delta^{\text{stat}} M_{t(n)}$ as a function of n. The dashed curve represents the results obtained by the method of effective moments. The left and right vertical axes are in GeV for the results from the LHC and LC, respectively

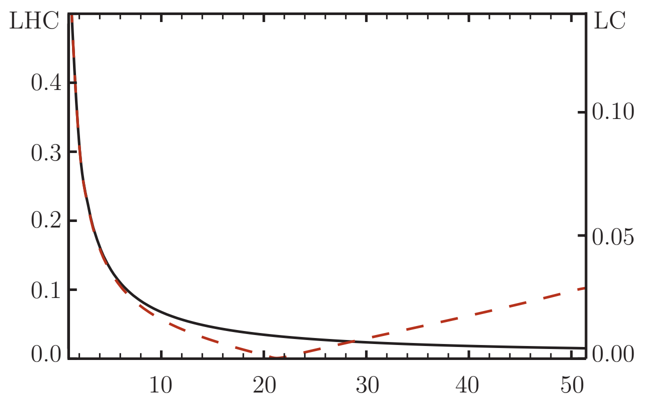

Fig. 5. A repeat of Fig. 4 for $\Delta^{\text{sys I}} M_{t(n)}$

Table 1. Statistical, systematic I and II, and the systematic summed-quadrature errors, presented in GeV for the LHC case. The last column represents the sum of the statistical and the systematic errors. In the brackets we show the results calculated by the method of effective moments (if they are different with the results calculated by the basic method)

n	$\Delta^{\text{stat}} M_{t(n)}$	$\Delta^{\text{sys I}} M_{t(n)}$	$\Delta^{\text{sys II}} M_{t(n)}$	$\Delta^{\text{sys}} M_{t(n)}$	$\Delta M_{t(n)}$
1	2.07	0.62	0.30	0.69	2.18
5	0.62	0.13	0.20 (0.17)	0.24 (0.21)	0.66
10	0.45	0.07 (0.06)	0.20 (0.13)	0.21 (0.14)	0.50 (0.48)
15	0.41 (0.40)	0.05 (0.03)	0.24 (0.12)	0.24 (0.12)	0.48 (0.42)
20	0.42 (0.39)	0.03 (0.00)	0.32 (0.11)	0.32 (0.11)	0.52 (0.40)
30	0.59 (0.38)	0.02 (0.03)	0.63 (0.11)	0.63 (0.11)	0.86 (0.39)

Table 2. The same as in Table 1 in the LC case

n	$\Delta^{\rm stat} M_{t(n)}$	$\Delta^{\rm sys\ I} M_{t(n)}$	$\Delta^{\rm sys\ II} M_{t(n)}$	$\Delta^{\rm sys} M_{t(n)}$	$\Delta M_{t(n)}$
1	0.80	0.17	0.19	0.26	0.84
5	0.24	0.04	0.13 (0.11)	0.13 (0.12)	0.27
10	0.17	0.02	0.13 (0.09)	0.13 (0.09)	0.27 (0.20)
15	0.16	0.01	0.15 (0.08)	0.15 (0.08)	0.22 (0.17)
20	0.16 (0.15)	0.01 (0.00)	0.21 (0.07)	0.21 (0.07)	0.26 (0.16)
30	0.23 (0.15)	0.01	0.41 (0.07)	0.41 (0.07)	0.46 (0.16)

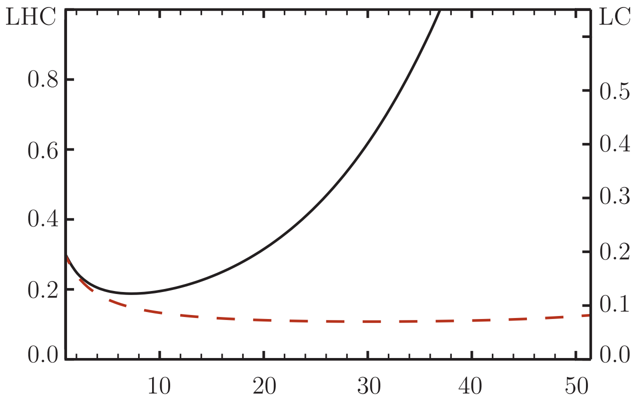

Fig. 6. A repeat of Fig. 4 for $\Delta^{\rm sys\ II} M_{t(n)}$

should be noted that at $n = 1$ the systematic errors in Table 1 practically coincide with those in [6]. The reason is that we have fixed the parameters of the current model by matching the errors of the first moments.

In the LC case, unfortunately, there are no published results that would allow us in a similar way to fix the parameters of the model. Therefore we mainly make use of indirect methods. We fix the parameter N according to the following reasoning. First we note that $\sigma(e^+e^- \to t\bar{t}) \approx 0.6\,\text{pb}$ at $\sqrt{s} = 500\,\text{GeV}$. So, at an integrated luminosity of $300\,\text{fb}^{-1}$, corresponding to 1–2 years of running, approximately 180 000 $t\bar{t}$ pairs must be generated. Since the branching of the process (1) is close to 30%, only 54 000 $t\bar{t}$ events relate to our investigation. The efficiency of their detection we estimate as follows. Suppose that at LC the detection efficiency of W-pairs decaying in a semi-leptonic channel will be the same as at LEP2, that is $\sim 80\%$ [15]. In addition, following [3], we suppose that the b-jet tagging efficiency at LC will be about 80%. In summary this gives an acceptance of 50%, which implies $N = 27\,000$.

We fix the parameter r based on the systematic error of M_t obtained in [11] in the direct reconstruction approach of $t\bar{t}$ events in the semi-leptonic channel. Additionally we use the note that, in the kinematic range near the upper $M_{b\ell}$ endpoint, the neutrino practically does not contribute to the total invariant mass of the decay products of the top quark. Therefore the determination of the M_t in the mentioned range is practically the same as the determination of the $M_{b\ell}$ invariant mass. [11] obtained $\Delta^{\rm sys} M_t = 250\,\text{MeV}$. So we set $\Delta^{\rm sys} M_{b\ell} = 250\,\text{MeV}$. Assuming that this is the type I error, we equate Δq to this value. Finally, by setting $\Delta q = rq$, $q \simeq M_t$ we get $r = 0.0014$.

We fix the parameter h by proceeding according to [6], based on the decreasing systematic type II error for the average $\langle M_{{\rm J}/\psi\ell}\rangle$ below 0.1 GeV for increasing statistics up to $N \sim 10^4$. Applying the method above to this, we obtain a rough estimate $h \simeq 1.1 \times 10^{-10}\,\text{GeV}^{-4}$.

Knowing N, r, h, we find $\Delta^{\rm stat}\langle q^n\rangle_{\rm exp}$, $\Delta^{\rm sys\ I,II}\langle q^n\rangle_{\rm exp}$ and then $\Delta^{\rm stat} M_{t(n)}$, $\Delta^{\rm sys\ I,II} M_{t(n)}$. Since in our model the pattern of the dependence on n is the common one in the LC and LHC cases, the difference between these cases appears in the scales of the errors only. This allows us to present the results on Figs. 4–6 by adding new scales. The numerical results are presented in Table 2. It is interesting to note that $\Delta^{\rm sys\ I} M_{t(1)}$ turns out to be smaller than the $\Delta^{\rm sys} M_t$ obtained in [11] in the framework of the direct reconstruction of events. Nevertheless this does not lead to an inconsistency. In fact, we equate the $\Delta^{\rm sys} M_t$ of [11] to Δq at $q \simeq M_t$, but the dominant contributions to $\Delta^{\rm sys\ I}\langle q\rangle_{\rm exp}$ are formed at strictly smaller q than M_t, which is obvious from formulas (12) and (13). This effect decreases $\Delta^{\rm sys\ I} M_{t(1)}$ compared to the $\Delta^{\rm sys} M_t$ of [11].

6 Theoretical uncertainty

The analysis of the previous sections shows that the experimental accuracy of the M_t determination can be considerably improved by moving to high degrees of the moments. So, to achieve an eventual high accuracy the theoretical uncertainty becomes increasingly crucial. In this regard it is important to understand whether the theoretical error in the determination of M_t can be made smaller than the experimental error for high degrees of the moments. If this will be possible then the theoretical error will not spoil the expected accuracy. Below we discuss this question in a somewhat qualitative manner since the possibility of a solution is of initial importance.

Let us begin with the observation that the origin of the theoretical uncertainty in the M_t determination is connected with the uncertainty of the calculation of the theoretical moment $\langle q^n\rangle$ in (5). Furthermore, in the determination of $\Delta M_{t(n)}$ the corresponding error $\Delta^{\rm th}\langle q^n\rangle$ is to be added (in quadratures) to $\Delta\langle q^n\rangle_{\rm exp}$ in formula (6). So the problem is reduced to the question of the possibility of carrying out the calculations precisely enough that the error $\Delta^{\rm th}\langle q^n\rangle$ is smaller than $\Delta\langle q^n\rangle_{\rm exp}$.

In practice it is convenient to compare relative errors such as $\Delta\langle q^n\rangle/\langle q^n\rangle$ instead of the proper errors $\Delta\langle q^n\rangle$. Fortunately the experimental relative error $\Delta\langle q^n\rangle_{\rm exp}/\langle q^n\rangle_{\rm exp}$

grows as we move to higher degrees of the moments; its behavior is similar to that represented in Fig. 2. In particular, when moving from $n = 1$ to $n = 15$, $\Delta\langle q^n\rangle_{\rm exp}/\langle q^n\rangle_{\rm exp}$ increases from 1.8% to 5.2%(4.5%) in the LHC case, and from 0.7% to 2.4%(1.9%) in the LC case. So with increasing n the requirement for the theoretical relative error $\Delta^{\rm th}\langle q^n\rangle/\langle q^n\rangle$ weakens.

Generally a theoretical error arises from a parametric uncertainty and an intrinsic uncertainty in the calculation itself. The parametric uncertainty originates mainly from the parameters that are least accurately known. In the given case these are the widths of the W boson and of the top quark. The analysis of [6] shows that the uncertainties in these parameters practically do not affect the first moment. Moreover, even switching-off the widths gives a negligible effect. In the 15-th moment, varying Γ_W within the experimental error $\Delta\Gamma_W = 0.04\,\text{GeV}$ results in $\Delta\langle q^{15}\rangle/\langle q^{15}\rangle = 0.09\%(0.06\%)$, which is also insignificant. Unfortunately we cannot estimate the variance of the moments by varying the width of the top quark, since from the very beginning we use the narrow-width approximation for the top quarks. Nevertheless, based on the results of [6] we expect a negligible variance of the moments in this case too. This is corroborated by the lack of reasons leading to appreciably greater sensitivity of the moments to the width of the top quarks than to the width of the W boson.

It should be mentioned, however, that the complete switching-off of the widths can vary the high-degree moments noticeably. Thus, setting $\Gamma_W = 0$ implies a shift of $\langle q^{15}\rangle$ by 4.6%(3.2%), which can be compared with the above estimations for the experimental relative errors. This means that the calculation of the high-degree moments must be carried out while taking into consideration realistic values for the widths. The latter requirement, of course, is unnecessary for the estimation of the errors only, which we are investigating herein.

Now let us consider the errors of the calculation itself. First we note that all the processes in (1) go far above the thresholds of the production of unstable particles, the W bosons and the top quarks. Therefore their production and decay can be described by standard methods [16], namely with the Dyson resummation in the leading-order calculation and the pole approximation when calculating the perturbation-theory corrections.[3] Thus, the problem is reduced to the estimation of the order of the perturbation theory that is necessary to satisfy the required precision of the calculation.

Further we note that the corrections to the moments will be calculated by calculating the corrections to the distribution $F(q)$. For kinematic reasons the basic features of the behavior of the latter corrections should follow the behavior of the distribution. Namely, $\Delta^{\rm th}F(q)$ must vanish at the ends of the kinematic region because of the vanishing phase volume. Furthermore, $\Delta^{\rm th}F(q)$ must almost vanish on the tail at large q due to the small size of the width of the W boson. (Recall that, in the limit $\Gamma_W = 0$, the distribution is completely suppressed on the tail for kinematic reasons.) So $\Delta^{\rm th}F(q)$ must be precisely known mainly in the middle of the kinematic region but not near its ends, including the tail. When going to the high-degree moments this condition is maintained. Moreover, in some sense it even becomes stronger. Really, at low q the contributions to the moments are additionally suppressed by the factor q^n. At large q, in the case of the effective moments, the contributions are completely suppressed by the cut-off Λ_n. In addition, the larger n is, the larger the distance between the cut-off and M_t, the right boundary of the actual range of the kinematic variable (since $\Lambda_n \to \Lambda$ from the right as $n \to \infty$). In particular, $\Lambda_1 = 171\,\text{GeV}$ but $\Lambda_{15} = 160\,\text{GeV}$, which is 15 GeV away from M_t. This feature is valuable for our investigation as the cut-off of the ends of the region of the kinematic variable implies a suppression of the large logarithms that can arise near the ends when calculating perturbation-theory corrections. In the final analysis this allows us to use the naive counting method to estimate corrections to the moments.[4]

With this in mind we use a rather rough approach that is based on a comparison between the corrections to the moments and the width of the top quark. (Notice that the width is actually the zero moment, accurate to the normalization.) The key reason for the approach is the observation that the integrals for the moments and the width, and for the corrections to the moments and the width, accumulate contributions mainly from the middle region of the kinematic variable. So, supposing that in this region the correction to the distribution $\Delta^{\rm th}F(q)$ varies weakly in the units of $F(q)$, one can expect that the corrections to the moments and to the width should be close to each other in relative units. By closeness here we mean on the order of several units. It is worth mentioning that, even in the case of the experimental errors, which depend strongly on the shape of the distribution, the relative errors at $n = 1$ and $n = 15$ only differ from each other by a factor of 2.5–3.5.

As we know, the electroweak one-loop correction to the top-quark width amounts to approximately 2%. The QCD one-loop correction is near 10%, while the two-loop correction is near to 2%. (See [1] and the references therein.) The comparison of these values with the estimates above for the experimental relative errors demonstrates that the one-loop electroweak and two-loop QCD corrections are enough to remain within the required limits (for both the LHC and LC cases). It should be noted that these corrections to the distribution can certainly be calculated since the corrections to the width have been calculated. Finally we also note that only the direct calculations of these corrections can explicitly solve the problem of the theoretical errors to the moments.

[3] As variants, one can exploit the method of effective field theory for calculating the resonant processes [17] or the modified perturbation theory based on distribution theory [18, 19].

[4] It should be emphasized that we discuss here the corrections to the $t \to b$ transition but not to the b-quark fragmentation, including the perturbative fragmentation. The latter process is described by the convolution of the cross section with the fragmentation function, and this operation is to be fulfilled in the framework of MC event generators.

The mentioned calculation, however, would not yet entirely close the problem of the theoretical error of the determination of M_t because of the problem of nonperturbative nature caused by the renormalon contribution. Below, for completeness, we only briefly consider this problem as its solution is known, at least conceptually. The problem is actually connected with the kind of the mass that is to be determined through an experimental measurement. In fact there are different masses, but only a Lagrangian mass is ultimately valuable since only this mass can be constrained by other fundamental parameters of the theory. The important representatives of the Lagrangian mass are the pole and $\overline{\text{MS}}$ masses. The directly measurable mass is the pole mass, which is determined by the kinematics. Correspondingly, the algorithms currently used to extract M_t from the data are tuned to the pole mass. However, because of the renormalon contribution pole-mass determination faces an extra uncertainty of order $O(\Lambda_{QCD})$ [20]. Numerically this can amount to hundreds of MeVs.

This difficulty can be bypassed in the framework of the following algorithm (below we state one of its possible variants) [20]. First, all theoretical calculations are to be fulfilled in terms of the pole mass. Then the value of the pole mass is to be determined from matching with the data. Remember, at this stage the result includes the renormalon contribution. Further, by means of the well-known formula relating the pole mass to the $\overline{\text{MS}}$ mass (see [1], for example), the $\overline{\text{MS}}$ mass is determined. At this step the result again contains the renormalon contribution but, as is declared, this cancels the previous one. (So the inaccuracy in the relation between the pole and $\overline{\text{MS}}$ masses is charged to the pole mass.) Direct calculations in certain examples [8, 9] demonstrate the effectiveness of this algorithm.

So, the problem is initially stated as though for the pole-mass determination, but at the final stage the $\overline{\text{MS}}$ mass is determined. This allows one to avoid a theoretical systematic uncertainty of order of $O(\Lambda_{QCD})$ caused by the renormalon contribution. Returning to our results, we see that the theoretical error of the top-mass determination can really be made smaller than the experimental error.

7 Discussion

The major result of this article is the detection of the effect of decreasing the statistical and systematic errors of the top-quark mass measured from the $M_{b\ell}$ distribution, when applying the technique of moments and proceeding to moments of high degree. The optimal value of the degree that minimizes the errors is found to be near $n = 15$.

To determinate the errors we have used a simple model. Its parameters in the LHC case have been fixed on the basis of the results obtained earlier [6] by the MC modeling method. As applied to LC the parameters have mainly been fixed by indirect methods. Knowing the parameters and their dependence on the degree n of the moments, we have estimated the errors as a function of n and have found the optimal value of n that minimizes the errors. The optimal value $n = 15$ is clearly visible in the framework of the basic method of calculating the moments. The application of the technique of effective moments decreases the errors at $n = 15$ by 10–20%, but further increases of n practically do not vary the results (see Figs. 4–6 and Tables 1–2).

At the optimal value $n = 15$ the total error ΔM_t is found to be close to 500 MeV in the LHC case, and close to 200 MeV in the LC case. In the LHC case this accuracy is more than twice that obtained by the other methods [1], including the original method in [6]. In the LC case the estimated accuracy of the M_t determination is close to that expected by scanning the $t\bar{t}$ production threshold [8, 9].

In conclusion it should be mentioned, once again, that at the intermediate stage of the analysis we have introduced simplifications that allow us to minimize the calculations. However at the final stage all the estimations have been made on the basis of realistic values of the parameters. This peculiarity should not reduce the legitimacy of the detected behavior of the errors and, moreover, of their estimations. Nevertheless the quantitative outcomes could be improved by further calculations based on the direct application of a proper MC event generator.

References

1. M. Beneke et al. (conveners), A. Ahmadov et al., Top quark physics, CERN 2000-004 [hep-ph/0003033]
2. J.A. Aguilar-Saavedra et al. (ECFA/DESY LC physics working group), TESLA technical design report Part III, DESY-01-011 [hep-ph/0106315]
3. T. Abe et al. (American Linear Collider Working Group), Linear Collider physics resource book for Snowmass 2001, in: Proceedings of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001), edited by N. Graf, SLAC-R-570 [Parts 1–4: hep-ex/0106055–hep-ex/0106058]
4. K. Abe et al. (ACFA linear collider working group) (ed.), Particle physics experiments at JLC, KEK-REPORT-2001-11 [hep-ph/0109166]
5. E. Accomando et al. (Report of the CLIC Physics Working Group), Physics at the CLIC multi-TeV linear collider, CERN-2004-005[hep-ph/0412251]
6. A. Kharchilava, Phys. Lett. B **476**, 73 (2000)
7. V.S. Fadin, V.A. Khose, JETP Lett. **46**, 525 (1987); Sov. Nucl. Phys. **48**, 309 (1988); Physics of elementary particles. Proc. XXIV LNPI Winter School, Leningrad, 1989, p. 3
8. A.H. Hoang, M. Beneke, K. Melnikov, T. Nagano, A. Ota, A.A. Penin, A.A. Pivovarov, A. Signer, V.A. Smirnov, Y. Sumino, T. Teubner, O. Yakovlev, A. Yelkhovsky, Jur. Phys. J. Direct C **3**, 1 (2002)
9. A.H. Hoang, A.V. Manohar, I.V. Stewart, T. Teubner, Phys. Rev. D **65**, 014014 (2002)
10. M. Martinez, R. Miquel, Eur. Phys. J. C **27**, 49–55 (2003)
11. S.V. Chekanov, Eur. Phys. J. C **26**, 173 (2002)
12. S.V. Chekanov, V.L. Morgunov, Phys. Rev. D **67**, 074011 (2003)
13. G. Corcella, E.K. Irish, M.H. Seymour, HERWIG for top physics at the linear collider, hep-ph/0012319
14. G. Corcella, M.L. Mangano, M.H. Seymour, JHEP **0007**, 004 (2000)
15. M. Campanelli, Int. Mod. Phys. A **14**, 3277 (1999); R. Strohmer, Int. Mod. Phys. A **18** 5127 (2003)

16. M.W. Grunewald et al., Four-fermion production in electron–positron collisions, The LEP2-MC workshop 1999/2000, hep-ph/0005309
17. W. Beenakker, F.A. Berends, A.P. Chapovsky, Nucl. Phys. B **573**, 503 (2000); A.P. Chapovsky, V.A. Khoze, A. Signer, W.J. Stirling, Nucl.Phys. B **621**, 257 (2002); M. Beneke, A.P. Chapovsky, A. Signer, G. Zanderighi, Nucl. Phys. B **686**, 205 (2004)
18. F. Tkachov, in: Proc. of the 32nd PNPI Winter School on Nuclear and Particle Physics, Ya.I. Azimov et al. (ed.) St.Petersburg, PNPI, 1999, p. 166 [hep-ph/9802307]
19. M.L. Nekrasov, Eur. Phys. J. C **19**, 441 (2001); Phys. Lett. B **545**, 119 (2002)
20. I.I. Bigi, M.A. Shifman, N.G. Uraltsev, A.I. Vainstein, Phys. Rev. D **50**, 2234 (1994); M. Beneke, V.M. Braun, Nucl. Phys. B **426**, 301 (1994)