CERN - European Organization for Nuclear Research

LCD-Note-2012-013

Prospects for the Measurement of the Top Mass in a Threshold Scan at CLIC

Katja Seidel*†, Frank Simon*†, Michal Tesař*

* *Max-Planck-Institut für Physik, Munich, Germany,*
† *Excellence Cluster 'Universe', TU München, Garching, Germany*

August 30, 2012

Abstract

We present a study of the capability of CLIC to measure the top quark mass and the strong coupling constant in a scan of the top threshold. The analysis is based on full detector simulations of the CLIC_ILD detector concept using Geant4, including realistic beam-induced background contributions from two photon processes. Event reconstruction is performed using a particle flow algorithm with stringent cuts to control the influence of background. With these simulations the signal and background selection efficiencies are determined. Signal event yields as a function of energy are obtained using these efficiencies together with NNLO top pair cross-sections corrected for ISR and the CLIC beam energy spectrum. For comparison, the analysis is also performed with the ILC beam energy spectrum. We find that a statistical precision of 21 MeV of the top mass is achieved assuming fixed α_s, and a statistical uncertainty of 33 MeV for m_t and 0.0009 for α_s is achieved in a combined extraction of both observables. At the ILC, the statistical uncertainties are between 10% and 20% smaller for the same integrated luminosity. In addition to the statistical uncertainties, systematic errors from theory and from the precision of α_s, as well as the influence of the precision of the background description and of the understanding of the luminosity spectrum have been studied.

Contents

1 Introduction

As the heaviest Standard Model particle, the top quark is of particular interest since it most strongly couples to the Higgs field and may provide sensitivity to Beyond the Standard Model physics. Experiments at e^+e^- colliders offer the possibility for a wide variety of studies involving top quarks, ranging from the precise measurement of top quark properties to the investigation of asymmetries providing large sensitivity to various New Physics models. Among those is the precise determination of the top quark mass, which is possible with two different techniques: through the direct reconstruction of top quarks from their decay products at energies above the production threshold, and through a scan of the top-pair production threshold. The latter technique has the advantage of providing the mass measurement in a theoretically well-defined scheme, while the former measurement can be performed essentially at arbitrary energies above threshold, however with potentially significant uncertainties due to non-perturbative contributions when transferring the measured invariant mass to a theoretically meaningful value. Progress has been made recently in establishing connections between the top mass parameters used in theory and the experimentally observable invariant mass of the decay products [1, 2], but theoretical uncertainties remain substantial.

In this note, we investigate the potential for the determination of the top quark mass from a measurement of the top-pair production cross-section at several energies around the threshold near 350 GeV, with a total integrated luminosity of up to 100 fb^{-1}. This study complements a previous CLIC study of top mass measurements at 500 GeV by means of a direct reconstruction of the invariant mass of the decay products, which showed that the invariant mass of the top quark can be determined with a precision of better than 100 MeV with 100 fb^{-1} in fully hadronic and semi-leptonic decays of the top pairs [3].

2 Experimental Conditions at CLIC at the Top Threshold

The Compact Linear Collider CLIC is a collider concept based on normal conducting accelerating cavities and two-beam acceleration, which is designed to provide up to 3 TeV collision energy. In a staged approach, a shorter, lower energy version would be operated initially, while construction is under way for the full energy phase.

In the present note, we study the case of a 500 GeV CLIC machine operated at energies close to the top pair production threshold. At 350 GeV, the rate of $\gamma\gamma \to$ hadrons events is relatively small, with only 0.05 events per bunch crossing, down by almost an order of magnitude compared to 500 GeV collisions. The effect from pile-up of this background, in particular after the application of the particle flow object selection cuts, is thus very minor. In addition, the measurement at the top threshold is a measurement of the cross section, which requires the separation of signal and background events, but not the precise reconstruction of the invariant mass, which reduces the impact of the background on the analysis.

The detector model used in the present study is a variant of CLIC_ILD [4], a detector concept based on Particle Flow event reconstruction. It consists of a low-mass, high-precision vertex detector and an inner silicon tracker, surrounded by a large-volume time projection chamber, followed by highly granular electromagnetic and hadronic calorimeters contained inside a 4 T solenoidal magnet with instrumented flux return for muon identification. The detector design is based on the ILD detector concept for the ILC, adapted to account for the higher energy (3 TeV) and more severe background conditions at CLIC. This leads to an increased radius of the innermost layer of the vertex detector, which sits at 31 mm compared to 16 mm in ILD at the ILC. At 500 GeV, the background is significantly reduced compared to the 3 TeV case, permitting modifications of the detector to optimize its performance for the lower collision energy. In particular the innermost vertex detector layer for CLIC_ILD can move in by 6 mm to a radius of 25 mm, improving flavor tagging at low momentum. To distinguish the modified detector design from the 3 TeV design, the detector model is referred to as CLIC_ILD_CDR500.

3 Simulation Strategy

For the correct description of the cross-section near threshold, the inclusion of high-order QCD contributions is necessary. Since no appropriate event generator publicly is available at present, the study follows the strategy of earlier studies performed for the TESLA collider [5] by factorising the simulation study into the determination of event selection efficiency and background contamination and the calculation of the top-pair production threshold. In this approach, the signal selection and background rejection is determined using fully simulated top-pair signal events as well as relevant background channels at a nominal center of mass energy of 352 GeV, slightly above the production threshold for the selected top mass of 174 GeV. This energy is chosen to be able to generate the events with PYTHIA, which requires a center-of-mass energy in excess of twice the generator top mass. Data points along the threshold curve are then generated by taking the signal cross section determined using NNLO calculations combined with the selection efficiency, adding background events assuming a constant level over the considered energy range of 10 GeV as determined from the full simulations. In the following, more details are given on the individual steps.

3.1 Top Pair Production Cross Section

The top-pair signal cross-section is determined using full NNLO calculations provided by the code TOPPIK [6, 7]. The top mass input is set to 174 GeV in the 1S mass scheme [6]. The strong coupling constant α_s is taken to be 0.118. Since TOPPIK provides the cross section in units of R, the ratio of $\sigma(e^+e^- \to X)$ to $\sigma(e^+e^- \to \mu^+\mu^-)$, the appropriate conversion factor of the energy-dependent cross section $e^+e^- \to \mu^+\mu^-$ is applied in addition.

Since this cross section is calculated for the energy at the e^+e^- vertex, additional corrections for initial state radiation (ISR) and for the beam energy spectrum of the accelerator have to be applied, as discussed in the following.

3.1.1 Initial State Radiation

ISR reduces the available collision energy E' due to the radiation of photons off the incoming electron and positron prior to the collision. This effect in general lowers the signal cross-section, since events are shifted to lower energies with typically a lower top-pair cross-section. The electron and positron "structure functions" are taken from the approximate YFS (Yennie-Frautschi-

Suura) solution as given in [8], which provides the normalized probability density for a given fraction of the lepton momentum x (ranging from 0 to 1) in the final collision.

For practical reasons, the folding of the ISR distribution with the theoretically calculated cross section is performed numerically. For this, a histogram of the structure function with 0.175 MeV wide bins is built, with the value in each bin taken by evaluating the approximate YFS solution at the bin center. The highest-energy bin is topped off to ensure correct normalization, accounting for the extreme increase in the structure function near 1. The folding is performed by evaluating 100 000 randomly generated energy points with the individual beam energies distributed according to this histogram. The mean value of the cross-section of these 100 000 trials is taken as the ISR-corrected cross section at a given center-of-mass energy.

3.1.2 Beam Energy Spectrum

The beam energy distribution also influences the cross section as a function of collider energy. The beam energy distribution is roughly characterized by the width of the main luminosity peak and by a longer tail to lower energies from beamstrahlung. To be able to compare the impact of the different beam energy distributions of CLIC and ILC, spectra from both colliders, operated at 350 GeV, are used to calculate the final signal cross-section.

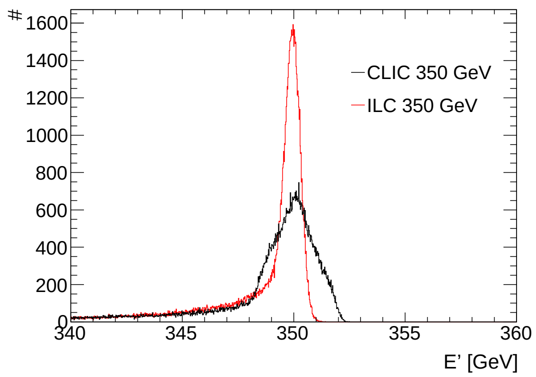

Figure 1: Beam energy spectrum for CLIC and ILC at 350 GeV.

Figure 1 shows the high-energy part of the beam energy spectrum of CLIC and ILC operated at 350 GeV. As for the case of ISR, the folding of the signal cross-section with the beam energy spectrum is performed numerically using 100 000 beam events at each energy point.

3.1.3 Combined Cross-Section

The final signal cross-section is obtained by combining the effects of ISR and of the beam energy spread. Using 100 000 trials per energy point with the collision energy determined from the beam energy distribution and ISR taken into account based on the resulting collision energy, the top pair cross section at both CLIC and ILC is determined based on the TOPPIK calculations.

Figure 2 shows the distribution of the real collision energy E' for CLIC and ILC for beam energy spectrum and ISR separately as well as the resulting combined spectrum. The effect on

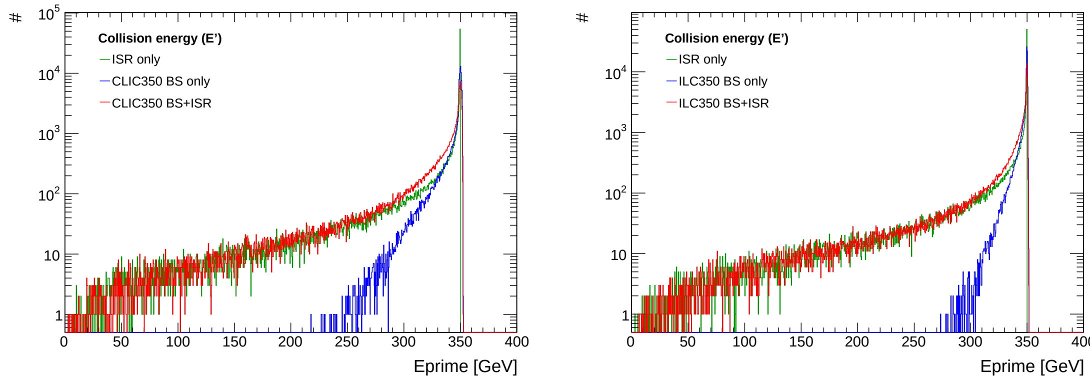

Figure 2: E' distribution taking ISR and beam energy spectrum (CLIC (left) and ILC (right)) into account.

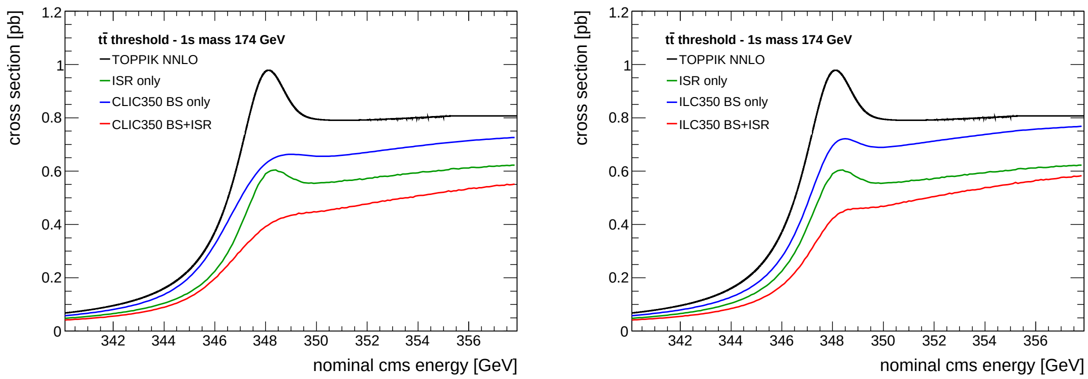

Figure 3: Top pair production cross-section from theory calculations, with beam energy spectrum and ISR as well as for all effects combined for both CLIC (left) and ILC (right).

the top pair production cross-section is shown in Figure 3. The cross-section with all effects included is used to determine the signal yield as a function of nominal collision energy in the subsequent analysis steps.

3.2 Signal Selection Efficiency and Background Contamination

The event selection efficiency and the background contributions, mainly from di- and tri-boson production, are determined using events generated with PYTHIA at a collision energy of 352 GeV with a top mass of 174 GeV. These events are fully simulated including pile-up from $\gamma\gamma \to$ hadrons background. For signal identification and background rejection the same technique as for the 500 GeV top mass study [3] is used. The top pair events are identified in the fully hadronic decay mode $t\bar{t} \to W^+bW^-\bar{b} \to q\bar{q}q\bar{q}b\bar{b}$ and in the semi-leptonic mode $t\bar{t} \to W^+bW^-\bar{b} \to q\bar{q}\ell^\pm\nu_\ell b\bar{b}$, $(l = e, \mu)$. The events are clustered into six or four jets depending on the number of identified isolated leptons. A kinematic fit with constraints on overall energy, on the difference of the two top masses and on the mass of the intermediate W bosons is used to form the top candidates. The fit also provides powerful background rejection, since most background events fail the kinematic fit. Additional background reduction is obtained with a binned likelihood using flavor tagging and other event variables to discriminate signal from background.

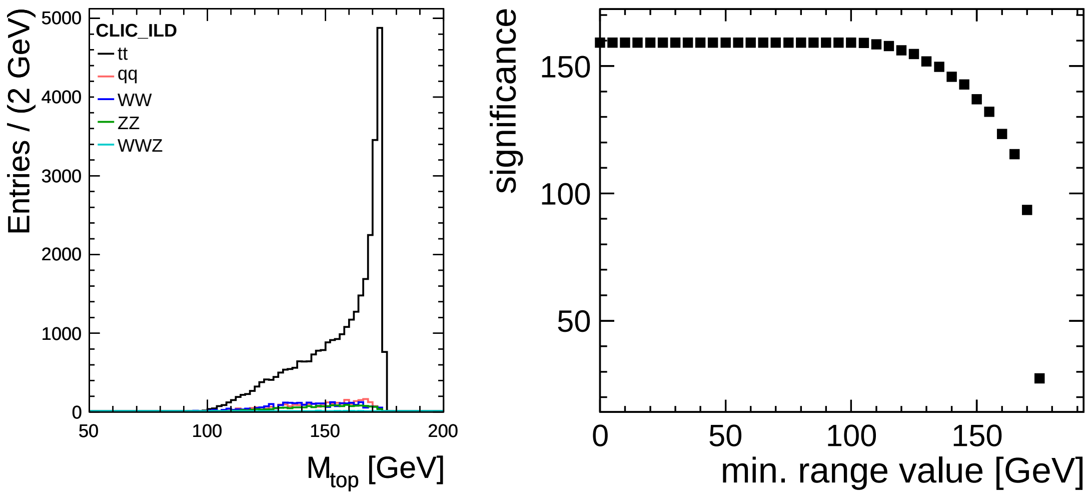

Figure 4: Reconstructed top quark mass for accepted events. Signal as well as each of the backgrounds are shown separately (left). Signal significance as a function of the value of a potential invariant mass cut above threshold for a top pair production cross-section of 450 fb^{-1} (right).

No further selection based on the reconstructed top quark mass is performed since this does not provide a substantial additional background rejection, while it would add potential systematic uncertainties from the additional cut. Figure 4 shows the reconstructed invariant mass distribution for top quark candidates after all selections for accepted signal and background events, as well as the signal significance above threshold (assuming a signal cross-section of 450 fb) as a function of a possible invariant mass cut. Overall, a signal selection efficiency of 70.2% is achieved, with an efficiency in excess of 90% for the selected fully-hadronic and semi-leptonic decay modes. For the major background channels, the cross-section is reduced by two to three orders of magnitude. Table 1 summarizes the signal and background cross-sections before and after selection.

Even though the study is performed using the CLIC_ILD detector model and CLIC background conditions, the conclusions drawn about the signal selection efficiency and background contamination also apply to ILC and the ILD detector. In terms of detector model, the most relevant difference is the radius of the innermost vertex detector layer, which is larger at CLIC due to the higher background level of incoherent e^+e^- pairs. For the identification of $t\bar{t}$ events, b-tagging is crucial, but not the separation of charm and bottom. Thus, the differences in performance of the two detector models are expected to be negligible for this analysis. The same

Table 1: Signal and considered physics background processes, with their cross section calculated for CLIC at 352 GeV before and after event selection. The combined background cross-section after selection is 78 fb.

type	$e^+e^- \to$	σ at 352 GeV	selected σ
Signal (m_{top} =174 GeV)	$t\bar{t}$	450 fb	316 fb
Background	$q\bar{q}$	25.2 pb	28 fb
Background	WW	11.5 pb	28 fb
Background	ZZ	865 fb	19 fb
Background	WWZ	10 fb	3 fb

also applies for the background rejection. Thus, the selection efficiencies and background levels determined for CLIC are also used for a study of a threshold scan at ILC.

3.3 Generation of Data Points

Simulated data points are generated by taking the ISR and beam spectrum corrected top pair cross-section at the desired energy to calculate the nominal number of events expected. The simulated number of signal events is determined on a random basis following a gaussian distribution with the mean set to the nominal number of events and the standard deviation given by the square root of that number. With the same method, background events are added, using a constant cross-section of 78 fb as discussed above. It is assumed that the nominal background contribution is well known both from theory and from measurements below threshold, so the nominal number of background events is subtracted from the signal, leaving just the statistical variations on top of the signal data with its own statistical uncertainty.

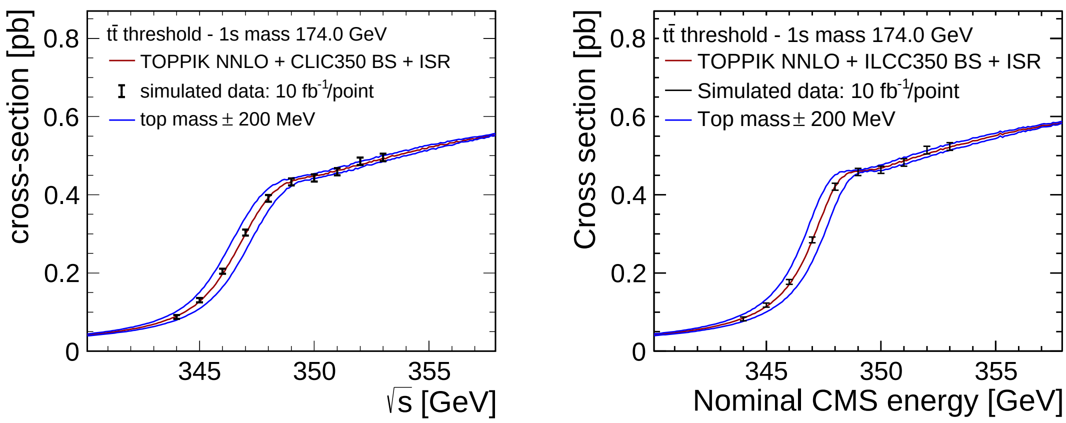

Figure 5: Background-subtracted simulated cross-section measurements for 10 fb^{-1} per data point, together with the cross-section for the generator mass of 174 GeV as well as for a shift in mass of $\pm$200 MeV for both CLIC (left) and ILC (right)).

In the present analysis, we assume a threshold scan with 10 data points with an integrated luminosity of 10 fb^{-1} each. The measurement points are spaced by 1 GeV. Figure 5 shows simulated data points for CLIC and for ILC.

4 Results

Two extractions of the top mass are being considered here: A one-dimensional template fit performed by comparing the simulated data with theory curves calculated in 50 MeV steps in top mass assuming α_s is known, and a two-dimensional template fit in top mass and α_s for a simultaneous determination of the top mass and the strong coupling constant. The measured top mass, and α_s in the case of the 2D fit, is given by the minimum of a parabolic fit to the χ^2 distribution of the different templates. The statistical uncertainty is taken from the standard deviation of the measured mass in 5000 trials with different simulated data points.

In the 1D fit, two main sources of systematic uncertainties are considered: A theory uncertainty taken as an overall normalization uncertainty of the calculated cross section, and an uncertainty from the knowledge of α_s, which is assumed to be known in the 1D fit. For the theory uncertainty, two levels are considered: A normalization uncertainty of 3%, assumed as a reasonably conservative estimate of current theory uncertainties [9], and an uncertainty of 1% optimistically assumed to be achievable with additional theoretical work in time for experiments at linear colliders. To determine the systematic error due to α_s, the current uncertainty of the world average of 0.0007 is assumed. The interpretation of the data points above threshold is particularly sensitive to the overall theory normalization uncertainty and to the strong coupling constant. In the 1D fit, uncertainties can thus be somewhat reduced by just considering the first six data points from 344 GeV to 349 GeV, without a reduction of the statistical sensitivity to the top mass. Table 2 summarizes the results.

Table 2: Results summary for the 1D top mass determination with a threshold scan at CLIC. For the systematic uncertainty originating from α_s, the current error on the world average of 0.0007 is assumed.

1S top mass 1D fit

measurement	stat. error	theory syst. (1%/3%)	α_s syst.
six point scan	21 MeV	15 MeV / 45 MeV	20 MeV
ten point scan	21 MeV	18 MeV / 54 MeV	21 MeV

Since the shape of the cross-section as a function of energy depends both on the top mass and on the strong coupling constant, a simultaneous determination of both is possible with a two-dimensional template fit. Figure 6 shows the resulting precision, and shows the clear correlation of the two variables. Since sensitivity to α_s also comes in through the higher-energy scan points, a reduced scan with six points along the strongly rising region of the cross-section lead to significantly increased uncertainties. In the case of the 2D fit, only the theory uncertainty is considered as a source for systematic uncertainties in the fit. The results are summarized in Table 3.

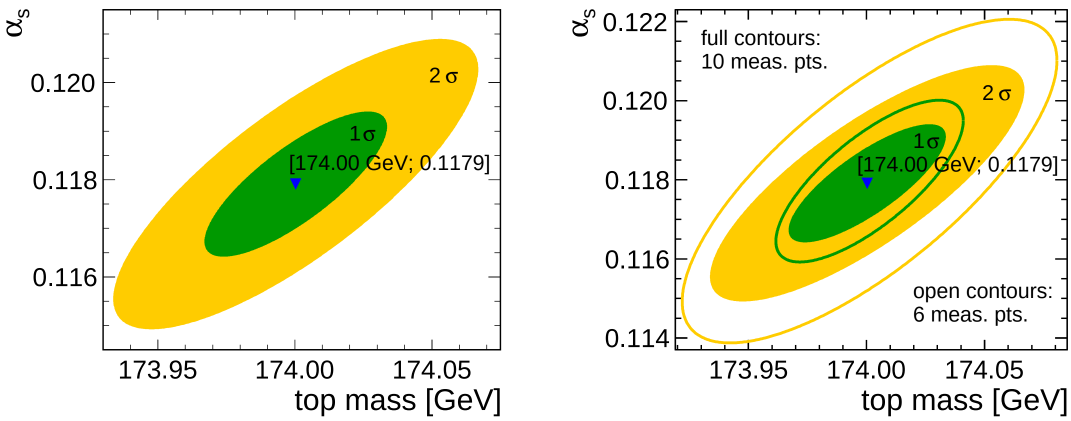

Figure 6: Simultaneous fit of the top mass and the strong coupling constant, showing the correlation of the two variables and the achieved precision (left). Difference in precision of top mass and α_s fit using just the first 6 points in the threshold scan or all 10 points (right).

Table 3: Results summary for the 2D simultaneous top mass and α_s determination with a threshold scan at CLIC.

1S top mass and α_s combined 2D fit

measurement	m_t stat. error	m_t th. syst. (1%/3%)	α_s stat. error	α_s th. syst. (1%/3%)
six point scan	39 MeV	2 MeV / 4 MeV	0.0014	0.0008 / 0.0020
ten point scan	33 MeV	6 MeV / 13 MeV	0.0009	0.0009 / 0.0022

5 Results for ILC beam conditions

The influence of the beam energy spectrum of the accelerator is studied by repeating the analysis using the ILC beam energy spectrum, as discussed in Section 3.1.2. The faster rise of the cross section due to the sharper main luminosity peak is expected to lead to somewhat reduced statistical uncertainties on the top mass for a given integrated luminosity due to increased differences between different mass hypotheses in the threshold region. As for the CLIC analysis, an integrated luminosity of 10 fb^{-1} per point is assumed. The same one- and two-dimensional fits of m_t and m_t and α_s combined are also performed for data points generated with the ILC beam spectrum. For simplicity, only the ten point fits are performed, and the systematic errors taking into account the theory normalization uncertainty and the uncertainty of α_s are not determined. It is however expected that these uncertainties are quantitatively very similar to the CLIC case.

Table 4 summarizes the results of both 1D and 2D fits, while Figure 7 shows the results of the combined extraction of the top mass and the strong coupling constant, illustrating the statistical uncertainty and the correlation of the two variables. In comparison to the statistical precision achieved assuming the CLIC beam energy spectrum, in the ILC case a 15% smaller uncertainty is observed in the 1D top mass fit, and a 20% smaller uncertainty on the top mass and a 10%

Table 4: Results summary for the 1D top mass fit and the 2D simultaneous top mass and α_s determination with a threshold at ILC.

ILC 1D 1S top mass and 2D 1S top mass and α_s combined fit

measurement	m_t stat. error	α_s stat. error
ten point scan 1D fit	18 MeV	-
ten point scan 2D fit	27 MeV	0.0008

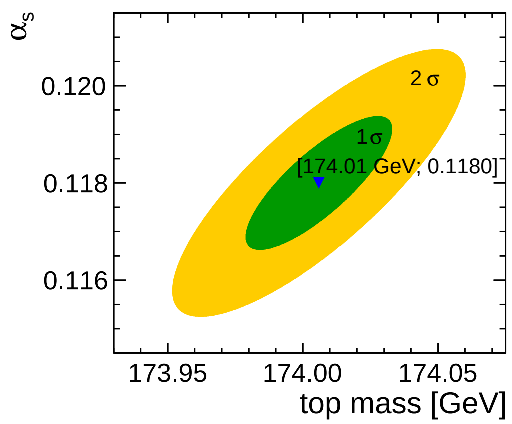

Figure 7: Simultaneous fit of the top mass and the strong coupling constant for data points simulated with the ILC beam energy spectrum, showing the correlation of the two variables and the achieved precision.

smaller uncertainty on α_s is obtained in the combined extraction. Compared to the systematic uncertainties originating from theory and from the precision of the strong coupling constant these differences are negligible.

6 Additional Systematic Studies

In addition to the theory uncertainties and the uncertainty of α_s in the case of the 1D fit, additional potential sources for systematic errors were studied.

A potential dependence of the result on the precise choice of energy values for the scan in relation to the top mass was excluded by shifting the measurement points to higher energies by 0.5 GeV without a significant change in the determined mass and α_s values.

The precise knowledge of the non-top background after event selection is crucial for the measurement of the signal cross section. The effect of an imperfect background description is studied by subtracting 5% and 10% too little or too much background before the fit. The 5% variation results in a 20 MeV shift in the top mass and 0.0005 in α_s, in both cases approximately two thirds of the statistical uncertainty. Subtracting only 90% of the background leads to a shift of twice the size, but also significantly reduces the stability of the template fit. Subtracting 110% of the background leads to a 30 MeV shift of the top mass and a shift of 0.0014 in α_s. This shows that an understanding of the background contamination on the level of 5% or better is important to keep systematic effects substantially below the statistical uncertainties.

The knowledge of the beam energy spectrum is very important for the correct description of the signal cross section, and thus also for the precision of the template fit. A full study has not yet been performed, but a very preliminary first study indicates that already a 20% uncertainty of the RMS width of the main luminosity peak results in top mass uncertainties of approximately 75 MeV, far in excess of the statistical uncertainties. Further studies to quantify the effects of realistic uncertainties of the beam energy spectrum are needed.

7 Conclusions

In this study, we have investigated the achievable precision of the top quark mass with a threshold scan at CLIC. Compared to the direct measurement of the invariant mass of the top quark decay products the threshold scan has the advantage that the mass is directly determined in a theoretically well-defined mass definition. The study uses event selection efficiencies and background contaminations from fully simulated events including the effects of the CLIC beam spectrum and $\gamma\gamma \to$ hadrons backgrounds and top pair signal cross-sections from NNLO calculations corrected for ISR and the beam energy spectrum. With an integrated luminosity of 100 fb^{-1} divided across ten data points spaced by 1 GeV, a statistical precision of the top quark mass in the 1S scheme of 33 MeV is obtained in a combined fit together with the strong coupling constant, which is determined with a precision of 0.0009. A one-dimensional fit with fixed α_s yields a precision of 21 MeV. Using the ILC beam energy spectrum instead results in 15% to 20% smaller uncertainties on the mass and in a 10% smaller uncertainty of the strong coupling constant. Combined systematic uncertainties from theory and background control are expected to be of similar order as the statistical errors. Together with a previous study of top quark mass measurements from direct reconstruction of the decay products this study demonstrates that precision top measurements are possible at CLIC both at and above threshold.

References

[1] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart. Jets from massive unstable particles: Top-mass determination. *Phys.Rev.*, vol. D77 p. 074010, 2008.

[2] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart. Top Jets in the Peak Region: Factorization Analysis with NLL Resummation. *Phys.Rev.*, vol. D77 p. 114003, 2008.

[3] K. Seidel, S. Poss, and F. Simon. Top quark pair production at a 500 GeV CLIC collider, 2011. CERN LCD-2011-026.

[4] A. Münnich and A. Sailer. The CLIC ILD CDR Geometry for the CDR Monte Carlo Mass Production. *LCD-Note-2011-002*, 2011.

[5] M. Martinez and R. Miquel. Multiparameter fits to the $t\bar{t}$ threshold observables at a future e^+e^- linear collider. *Eur. J. Phys.*, vol. C27 pp. 49–55, 2003. hep-ph/0207315.

[6] A. H. Hoang and T. Teubner. Top quark pair production close to threshold: Top mass, width and momentum distribution. *Phys. Rev.*, vol. D60 p. 114027, 1999. CERN-TH-99-59, DESY-99-047, hep-ph/9904468.

[7] A. H. Hoang and T. Teubner. Top quark pair production at threshold: Complete next-to-next-to-leading order relativistic corrections. *Phys. Rev.*, vol. D58 p. 114023, 1998. UCSD-PTH-98-01, DESY-98-008, hep-ph/9801397.

[8] M. Skrzypek and S. Jadach. Exact and approximate solutions for the electron nonsinglet structure function in QED. *Z.Phys.*, vol. C49 pp. 577–584, 1991.

[9] A. Hoang and M. Stahlhofen. NNLL Top-Antitop Production at Threshold. 2011. arXiv:1111.4486 [hep-ph].