PHYSICAL REVIEW D **84,** 014033 (2011)

Measuring the top Yukawa coupling at the ILC at $\sqrt{s} = 500$ GeV

Ryo Yonamine,[1] Katsumasa Ikematsu,[2] Tomohiko Tanabe,[3] Keisuke Fujii,[4] Yuichiro Kiyo,[4] Yukinari Sumino,[5] and Hiroshi Yokoya[6,7]

[1]*Department of Particle and Nuclear Physics, The Graduate University for Advanced Studies (Sokendai), Tsukuba 305-0801, Japan*
[2]*Department of Physics, Universität Siegen, D-57068 Siegen, Germany*
[3]*International Center for Elementary Particle Physics, The University of Tokyo, Tokyo 113-0033, Japan*
[4]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801, Japan*
[5]*Department of Physics, Tohoku University, Sendai 980-8578, Japan*
[6]*Theory Unit, Physics Department, CERN, CH-1211 Geneva, Switzerland*
[7]*Department of Physics and National Center for Theoretical Sciences, National Taiwan University, Taipei 10617, Taiwan*

(Received 28 April 2011; published 26 July 2011)

We report on the feasibility of the direct measurement of the top Yukawa coupling g_t at the International Linear Collider during its first phase of operation with a center-of-mass energy of 500 GeV. The signal and background models incorporate the nonrelativistic QCD corrections which enhance the production cross section near the $t\bar{t}$ threshold. The $e^+e^- \to t\bar{t}H$ signal is reconstructed in the 6-jet + lepton and the 8-jet modes. The results from the two channels are combined. The background processes considered are $e^+e^- \to t\bar{b}W^-/\bar{t}bW^+$ (which includes $e^+e^- \to t\bar{t}$), $e^+e^- \to t\bar{t}Z$, and $e^+e^- \to t\bar{t}g^* \to t\bar{t}b\bar{b}$. We use a realistic fast Monte Carlo detector simulation. Signal events are selected using event shape variables, through jet clustering, and by identifying heavy flavor jets. Assuming a Higgs mass of 120 GeV, polarized electron and positron beams with $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$, and an integrated luminosity of 1 ab^{-1}, we estimate that the $e^+e^- \to t\bar{t}H$ events can be seen with a statistical significance of 5.2σ, corresponding to the relative top Yukawa coupling measurement accuracy of $|\Delta g_t/g_t| = 10\%$.

DOI: 10.1103/PhysRevD.84.014033 PACS numbers: 13.66.Jn, 14.65.Ha, 14.80.Bn

I. INTRODUCTION

The standard model (SM) of elementary particle physics stands on two pillars. The first pillar is the gauge principle, which has been verified by precision electroweak measurements. The second pillar consists of the electroweak symmetry breaking which is yet to be tested by experiment. The discovery of the Higgs boson will be of particular importance in explaining the mass generation mechanism. Within the SM, the Yukawa interaction of the top quark and the Higgs boson generates the mass term which breaks the electroweak gauge symmetry. The measurement of the strength of the top Yukawa coupling g_t can shed light on the mechanism behind the generation of the top quark mass.

The top-quark Yukawa interaction could be measured indirectly using the production mechanism of the Higgs boson through the top quark loop at the LHC experiments. The indirect measurement unfortunately cannot give a full description of the top-quark Yukawa interaction for the following reason. If we were to observe an anomaly in the Higgs production cross section, it would be difficult to distinguish whether this effect is due to an anomaly in the top Yukawa interaction itself, or there are contributions from unknown particles propagating in the loop connecting the initial state and the Higgs boson. In order to distinguish these two effects, it would be highly desirable to measure the top Yukawa interaction directly. At the LHC, the direct production process $gg \to t\bar{t}H$ in the $H \to b\bar{b}$ channel is marred by jet combinatorial background [1]. While the $H \to \gamma\gamma$ or $H \to \tau^+\tau^-$ channels are expected to yield cleaner signals [2], which could allow for the discovery of the $gg \to t\bar{t}H$ process, the uncertainty in the top Yukawa coupling value would be affected by the potentially large uncertainties in the Higgs branching fraction measurements. We show that a future e^+e^- linear collider, such as the International Linear Collider (ILC), can play a critical role in the determination of the top Yukawa coupling through the direct measurement of $e^+e^- \to t\bar{t}H$ in the $H \to b\bar{b}$ channel.

Feasibility studies of the top-quark Yukawa interaction at a future e^+e^- linear collider have a long history [3,4]. A serious feasibility study of a direct measurement of the top Yukawa coupling using the process $e^+e^- \to t\bar{t}H$ at the center-of-mass (CM) energy $\sqrt{s} = 800$ GeV was performed in [5,6], which incorporated realistic experimental conditions expected at a linear collider experiment. More recently, there has been increased interest in how well the top Yukawa coupling can be measured in the first phase of a linear collider experiment, whose CM energy reaches up to $\sqrt{s} = 500$ GeV. An analysis for $\sqrt{s} = 500$ GeV was carried out in [7]. It was noted that at this energy the bound-state effects between t and $\bar{t}$ enhance the $t\bar{t}H$ production cross section significantly, since the relative momentum of t and $\bar{t}$ is typically small [8–14]. A reanalysis was performed in the Snowmass workshop, incorporating the enhancement effect by $t\bar{t}$ resonance formation as well as an enhancement effect that can be obtained by polarizing the e^+e^- beams [15]. The conclusion was that the top

Yukawa coupling can be measured to roughly 10% accuracy, including statistical errors only, with an integrated luminosity of 1 ab^{-1}.

In this paper, we investigate the feasibility of measuring the top Yukawa coupling at $\sqrt{s} = 500$ GeV using the process $e^+e^- \to t\bar{t}H$ for the Higgs mass of 120 GeV. The new aspects of this study as compared to the previous ones are as follows. We implement the enhancement factor by $t\bar{t}$ bound-state effects into the event generator, both for the $t\bar{t}H$ signal and the $t\bar{t}Z$ background events; the latter is particularly important since the expected measurement accuracy of the top Yukawa coupling is significantly affected by the number of these background events. In addition, we perform a fairly detailed detector simulation which takes into account the realistic energy resolution of the calorimeter components (see Sec. IV).

This paper is organized as follows. In Sec. II, we present our method for including the enhancement by $t\bar{t}$ bound-state effects. In Sec. III, the signatures of the $t\bar{t}H$ process and the possible background processes are outlined. The analysis framework used for the event generation and the detector simulations is discussed in Sec. IV. We discuss the event selection procedure in detail in Sec. V for the 6-jet plus lepton mode analysis, and in Sec. VI for the 8-jet mode. We summarize the accuracy estimate of the top Yukawa coupling measurement in Sec. VII. The measurement is assumed to be dominated by the statistical uncertainty.

II. INCLUSION OF $t\bar{t}$ BOUND-STATE EFFECTS

Theoretical analyses of the $t\bar{t}$ bound-state effects on the cross section for $e^+e^- \to t\bar{t}H$ with the next-to-leading logarithmic accuracy are given in [14]. To our knowledge, there have been no analyses of $t\bar{t}$ bound-state effects for the background process $e^+e^- \to t\bar{t}Z$ which are included consistently with the $e^+e^- \to t\bar{t}H$ signal. In our event generator, all the QCD corrections are incorporated consistently only with the leading-order accuracy. Hence, for consistency, we incorporate the $t\bar{t}$ bound-state effects on the signal and background cross sections with the leading-order accuracy; we also incorporate some of the important next-to-leading-order corrections in the bound-state effects.

There are a number of (tree-level) Feynman diagrams contributing to each of the processes $e^+e^- \to t\bar{t}H$, $e^+e^- \to t\bar{t}Z$, and $e^+e^- \to t\bar{t}g^*$, where the $t(\bar{t})$ subsequently decays into $bW^+(\bar{b}W^-)$. Let us denote these amplitudes for the process $i \to f$ as $A_{t\bar{t}}(i \to f)$. The tree-level amplitudes are modified as follows:

$$A_{t\bar{t}}(i \to f) = [A_{t\bar{t}}(i \to f)]_{\text{tree}} \sqrt{K_{i\to f}} F(\hat{s}_{t\bar{t}}, \vec{p}; m_t, \Gamma_t, \alpha_s). \tag{1}$$

F represents a process-independent enhancement factor that incorporates $t\bar{t}$ S-wave bound-state effects; $\sqrt{\hat{s}_{t\bar{t}}}$ denotes the CM energy of t and $\bar{t}$ as reconstructed from the final $bW^+\bar{b}W^-$ system; $\vec{p}$ is the three-momentum of t in the CM frame of t and $\bar{t}$; m_t and Γ_t denote the pole mass and width of the top quark, respectively. Close to the threshold of $t\bar{t}$ pair production, this factor F incorporates the bound-state effects according to the nonrelativistic bound-state theory, while for higher values of $\sqrt{\hat{s}_{t\bar{t}}}$, the factor F is smoothly interpolated to unity:

$$F = \begin{cases} \frac{G(E,\vec{p})}{G_0(E,\vec{p})} & E \equiv \sqrt{\hat{s}_{t\bar{t}}} - 2m_t \ll m_t, \\ 1 & E \gtrsim m_t. \end{cases} \tag{2}$$

The nonrelativistic Green function is defined by

$$\left[(E + i\Gamma_t) - \left\{-\frac{\nabla^2}{m_t} + V_{\text{QCD}}(r)\right\}\right] G(E, \vec{x}) = \delta^3(\vec{x}), \tag{3}$$

$$G(E, \vec{p}) = \int d^3\vec{x} e^{-i\vec{p}\cdot\vec{x}} G(E, \vec{x}), \tag{4}$$

where $r = |\vec{x}|$ and $V_{\text{QCD}}(r)$ is the next-to-leading-order QCD potential [14]. $G_0(E, \vec{p})$ is the nonrelativistic Green function of a free $t\bar{t}$ pair, which is defined via Eqs. (3) and (4) after setting $V_{\text{QCD}}(r)$ to zero. The enhancement factor F is explained in more detail in Ref. [16].

In Eq. (1) $K_{i\to f}$ denotes a process-dependent hard-vertex correction factor, which arises as a part of the next-to-leading-order corrections. To a good approximation this factor is independent of kinematical variables for the signal process $e^+e^- \to t\bar{t}H$. In fact with a choice $K_{i\to f} = 0.843$, we reproduce the $e^+e^- \to t\bar{t}H$ differential cross section at $\sqrt{s} = 500$ GeV shown in Fig. 5(a) of Ref. [13]. [The next-to-leading-logarithmic curve in the threshold region of $t\bar{t}$ and the $\mathcal{O}(\alpha_s)$ curve at higher $\sqrt{\hat{s}_{t\bar{t}}}$ or lower E_H.] We adopt this value of $K_{i\to f}$ for the signal process. For the background processes, we choose $K_{i\to f} = 1$, since these factors are unknown and since the deviation of these factors from unity is part of the next-to-leading-order corrections not fully accounted for in our analysis.

III. SIGNAL AND BACKGROUND PROCESSES

The diagrams for $e^+e^- \to t\bar{t}H$ with subsequent top decays $t \to bW^+(\bar{t} \to \bar{b}W^-)$ are shown in Fig. 1.

The first and second diagrams containing the top Yukawa coupling g_t are the targets of this study. The contribution to the cross section coming from the third diagram, where the Higgs radiates off of the intermediate Z boson, is negligible at $\sqrt{s} = 500$ GeV, as shown in Fig. 2. As a result, the number of $t\bar{t}H$ events is proportional to g_t^2 to a very good approximation, which enables us to perform a simple analysis by event counting.

In this study, the Higgs boson is reconstructed in the two b-jet mode $H \to b\bar{b}(68\%)$. Our $t\bar{t}H$ signal can be classified

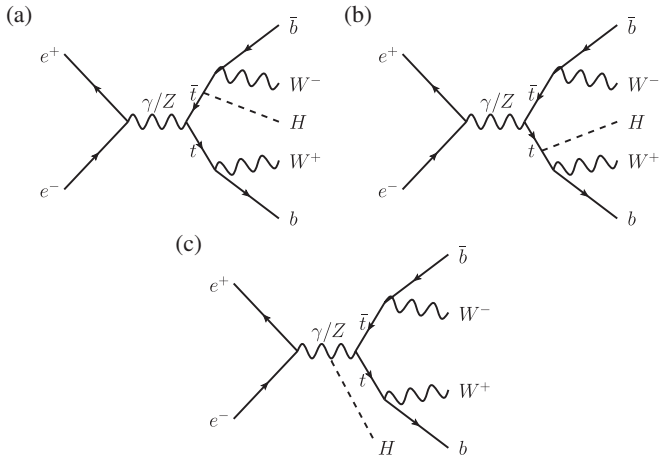

FIG. 1. Feynman diagrams for the $e^+e^- \to t\bar{t}H$ process.

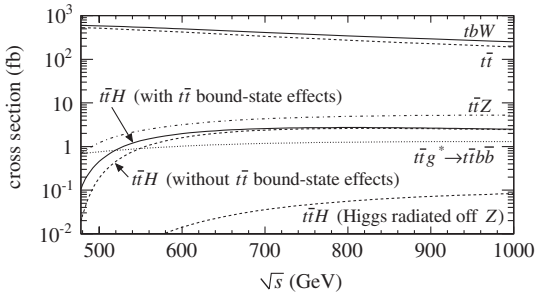

FIG. 2. Production cross section of the $e^+e^- \to t\bar{t}H$ signal (shown with and without $t\bar{t}$ bound-state effects), together with those of the main background processes, $t\bar{t}H$ (Higgs radiated off the Z boson), $t\bar{t}Z$, $t\bar{t}$, $t\bar{b}W^-/\bar{t}bW^+$ (denoted as tbW), and $t\bar{t}g^* \to t\bar{t}b\bar{b}$, as a function of the CM energy without beam polarizations. The initial state radiation and beamstrahlung effects are included.

into three groups, depending on the decay mode of the W bosons. Their branching fractions are

(i) 8-jet mode: 45%,
(ii) 6-jet + lepton mode (e or μ): 29%,
(iii) 4-jet + 2-lepton mode (ee, $e\mu$, or $\mu\mu$): 5%,

where we have omitted the contribution of the top decays to tau ($t \to b\tau^+\nu_\tau$ and $\bar{t} \to \bar{b}\tau^-\bar{\nu}_\tau$), since we only reconstruct electrons and muons from the top in this study. The 8-jet mode and the 6-jet + lepton mode are chosen for reconstruction.

The following processes are identified as possible background sources which can mimic the $t\bar{t}H$ signatures:

(i) $e^+e^- \to t\bar{b}W^-/\bar{t}bW^+ \to bW^+\bar{b}W^-$,
(ii) $e^+e^- \to t\bar{t}Z \to bW^+\bar{b}W^-b\bar{b}$,
(iii) $e^+e^- \to t\bar{t}g^* \to bW^+\bar{b}W^-b\bar{b}$.

The cross sections for these processes are shown as a function of $\sqrt{s}$ in Fig. 2.

We will refer to the $e^+e^- \to t\bar{b}W^-/\bar{t}bW^+$ process as $e^+e^- \to tbW$. The $e^+e^- \to tbW$ process includes the $e^+e^- \to t\bar{t}$ process. The $e^+e^- \to tbW$ final state consists of up to two b jets, as opposed to four b jets for our $t\bar{t}H$ signal. The tbW channel can be therefore reduced to a small fraction by identifying the flavor of the b quarks in the final state (b tagging) and by counting the number of b jets. Because of the large tbW cross section, a significant amount of tbW background remains even if there is a small rate of event misreconstruction, which occurs equally likely for events in and away from the top pair resonance, thus making it important to include the nonresonant contributions.

In the event generation, the top quark decays explicitly as $t \to bW^+$ ($\bar{t} \to \bar{b}W^-$), before the hadronization step. Thus, in order to take into the background due to hard gluon emissions from the top quark, we separately include the independent contribution from the $e^+e^- \to t\bar{t}g^*$ background.

In contrast to the tbW process, the processes $t\bar{t}Z$ and $t\bar{t}g^*$ can have identical final states as those of the $t\bar{t}H$ process if the Z boson or the hard gluon g^* decays into a $b\bar{b}$ pair. In this case, the signal extraction will depend strongly on the resolution of the Higgs mass reconstructed from the two b jets. The unpolarized cross section for $t\bar{t}Z$ is 1.3 fb, including the $t\bar{t}$ bound-state effects similar to that expected for the signal process; without including this correction, the cross section becomes 0.7 fb. For $t\bar{t}g^* \to t\bar{t}b\bar{b}$, the unpolarized cross section is 0.7 fb. We note that there is no $t\bar{t}$ bound-state enhancement in the $t\bar{t}g^*$ process because the $t\bar{t}$ system is not a color singlet in this case. The cross sections at $\sqrt{s} = 500$ GeV for our signal and background processes are summarized in Table I.

The signal and background samples have been produced with pure beam polarizations. Unless otherwise noted, our results weight these samples to match the beam polarizations of $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$ [17].

TABLE I. Cross sections at $\sqrt{s} = 500$ GeV for the signal and background processes are shown for the different beam polarizations. The $e^+e^- \to t\bar{t}H$ and $e^+e^- \to t\bar{t}Z$ processes include the $t\bar{t}$ bound-state effects. The $t\bar{t}H$, $t\bar{t}Z$, and $t\bar{t}g^*$ processes all decay as $bW^+\bar{b}W^-b\bar{b}$ while the $\bar{t}bW^+/t\bar{b}W^-$ process (denoted as tbW) decays as $bW^+\bar{b}W^-$. The number of events N used in this study is shown for each sample, along with its equivalent luminosity $\mathcal{L}$.

Process	σ (fb)	N	$\mathcal{L}$ (ab^{-1})
$e^-_L e^+_R \to t\bar{t}H$	1.07	5.00×10^4	47.8
$e^-_L e^+_R \to t\bar{t}Z$	4.04	5.00×10^4	12.4
$e^-_L e^+_R \to t\bar{t}g^*$	1.93	5.00×10^4	25.9
$e^-_L e^+_R \to tbW$	1633	1.00×10^7	6.1
$e^-_R e^+_L \to t\bar{t}H$	0.45	5.00×10^4	92.6
$e^-_R e^+_L \to t\bar{t}Z$	1.32	5.00×10^4	37.8
$e^-_R e^+_L \to t\bar{t}g^*$	0.86	5.00×10^4	58.2
$e^-_R e^+_L \to tbW$	700	1.00×10^7	14.3

IV. ANALYSIS FRAMEWORK

Signal and background events are generated using the PHYSSIM [18] event generator, based on the full helicity amplitudes including gauge boson decays, calculated using HELAS [19] and BASES [20], which properly takes into account the angular distributions of the decay products. For the event generation, the following values are used: $\alpha(M_Z) = 1/128$, $\sin^2\theta_W = 0.230$, $\alpha_s(M_Z) = 0.120$, $M_W = 80.0$ GeV, $M_Z = 91.18$ GeV, $M_t = 175$ GeV, and $M_H = 120$ GeV. The effects of initial state radiation and beamstrahlung are included. The $t\bar{t}$ bound-state effects results in a roughly twofold increase in the $t\bar{t}H$ signal cross section at $\sqrt{s} = 500$ GeV, as shown in Fig. 3.

The four-momenta of the final-state quarks and leptons are passed as input to PYTHIA 6.4 [21] for parton showering and hadronization. The detector response is simulated using the QUICKSIM [22] fast Monte Carlo detector simulator.

The detector consists of the beam pipe, a vertex detector, a drift chamber, an electromagnetic calorimeter (ECAL), and a hadronic calorimeter (HCAL). The crossing angle of the beams is also taken into account. Each hit in the tracking detector is smeared according to the detector resolution specified in Table II. For each charged particle, the parameters describing its helical trajectory are smeared according to the full covariance matrix of the parameters. Calorimeters are simulated down to the level of individual cells with possible overlaps of energy deposits from nearby particles. Each hit in the calorimeter cell is smeared according to Table II. The calorimeter cell hits are clustered and then matched to the tracks of charged particles. ECAL clusters which are consistent with a charged track are subtracted based on the particle flow approach [23]. HCAL clusters whose energy is consistent with charged hadrons are removed, while for clusters with inconsistent energy matching, as is the case when neutral hadrons are present, HCAL energy deposits are statistically subtracted by an amount weighted by the geometrical overlap between the charged hadrons and the HCAL clusters. The resulting jet energy resolution for $Z \to q\bar{q}$ events is found to be similar to the case of studies with the detailed detector simulations [24,25]. For events with high jet multiplicities, the jet energy resolution is degraded due to the confusion in the jet clustering procedure.

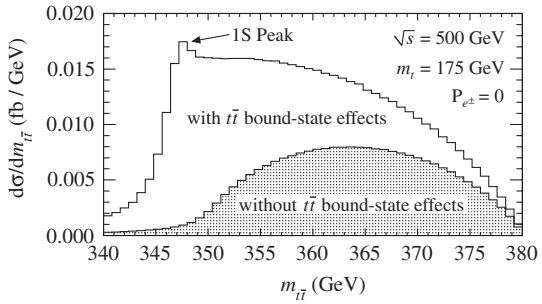

FIG. 3. Differential production cross section as a function of the invariant mass of the $t\bar{t}$ system for unpolarized beams $(P_{e^-}, P_{e^+}) = (0.0, 0.0)$ with $\sqrt{s} = 500$ GeV. The shaded histogram represents the leading-order values. The white histogram includes the $t\bar{t}$ bound-state effects.

TABLE II. Detector parameters. p, p_T, and E are measured in units of GeV. The angle θ is measured from the beam axis.

Detector	Resolution	Coverage
Vertex detector	$\sigma_b = 7.0 \oplus (20.0/p\sin^{3/2}\theta)\ \mu$m	$\lvert\cos\theta\rvert \le 0.90$
Drift chamber	$\sigma_{P_T}/P_T = 1.1 \times 10^{-4} p_T \oplus 0.1\%$	$\lvert\cos\theta\rvert \le 0.95$
ECAL	$\sigma_E/E = 15\%/\sqrt{E} \oplus 1\%$	$\lvert\cos\theta\rvert \le 0.90$
HCAL	$\sigma_E/E = 40\%/\sqrt{E} \oplus 2\%$	$\lvert\cos\theta\rvert \le 0.90$

V. ANALYSIS OF THE 6-JET + LEPTON MODE

We describe the event selection for the analysis of the 6-jet + lepton mode first, followed by the 8-jet mode. Similar techniques are used in both modes. The main differences between the two analyses are the presence of a lepton in the 6-jet + lepton mode and the number of jets in the final state.

A. Identification of an isolated lepton

In the 6-jet + lepton mode, the lepton from the $W \to \ell\nu$ tends to be energetic and isolated from the rest of the event. To identify such a lepton (e or μ), a cone with a half-opening angle $\theta_{\rm cone}$ is constructed around each track (lepton candidate). The cone energy $E_{\rm cone}$ is defined to be the sum of the energy of all the tracks inside the cone, excluding the lepton candidate. The value of $\theta_{\rm cone}$ which gives $\cos\theta_{\rm cone} = 0.98$ is found to be optimal for our event selection.

Figure 4 shows the distribution of the cone energy versus the lepton candidate energy. The energetic isolated leptons

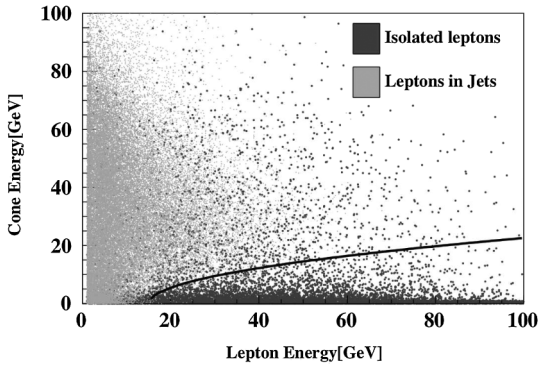

FIG. 4. Distribution of the cone energy and the lepton energy, shown for leptons from W decays (black dots) and leptons originating in heavy flavor jets (gray dots). The black curve shows the cut boundary for the lepton selection; leptons below the curve are identified as isolated leptons.

from the W decay have a high lepton energy and a low cone energy, thus populating the lower right region, shown as black dots in Fig. 4. Leptons from heavy flavor jets are likely to be less energetic and have a higher cone energy, shown as gray dots in Fig. 4. The selection of isolated leptons is performed by applying a cut on the cone energy which varies as the lepton energy and is given by the equation $E_{\text{cone}} < \sqrt{6(E_\ell - 15)}$, where E_{cone} and E_ℓ are given in units of GeV. For the 6-jet + lepton event selection, we require one and only one isolated lepton.

B. Event shape

To exploit the differences in the event topology between $e^+e^- \to t\bar{t}H$ and $e^+e^- \to tbW$ events, we use the *thrust* T, defined as [26,27]

$$T = \max_{|\hat{n}|=1} \frac{\sum_i |\hat{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}, \tag{5}$$

where the index i runs over each reconstructed particle with $\vec{p}_i$ corresponding to its momentum, and T is to be maximized with respect to the unit vector $\hat{n}$ corresponding to the axis in which the overall event shape is stretched. The thrust T tends to unity for 2-jet-like events while it tends to $1/2$ for isotropic events. Because the $t\bar{t}$ events tend to have fewer jets, with a higher average jet energy, T tends to be higher for tbW events compared to $t\bar{t}H$ events. The requirement of $T < 0.77$ is found to be optimal in the 6-jet + lepton analysis.

C. Jet clustering

We employ the Durham jet clustering algorithm [28] to separate the event into 6 jets, after taking out the isolated lepton. In the Durham algorithm, each particle is regarded as a jet on its own to begin with; a jet pair i and j gets combined if the pair has the lowest Y_{ij} value which is defined as

$$Y_{ij} = \frac{\min\{E_i^2, E_j^2\}(1 - \cos\theta_{ij})}{E_{\text{cm}}^2}, \tag{6}$$

where θ_{ij} is the angle between the momentum vectors of the two jets, and $E_{\text{cm}} = \sqrt{s}$. The jet clustering is allowed to continue until there are two jets remaining, with the value of Y_{ij} being recorded at each transition from n jets to $n-1$ jets, which we call $Y_{n\to n-1}$. It is found that $Y_{5\to4}$ is useful for discriminating $t\bar{t}H$ 6-jet + lepton events from tbW events. This is due to the fact that, after identifying the isolated lepton ℓ from the semileptonic decay $t \to bW \to b\ell\nu$, tbW events cannot have more than four jets without a gluon emission. Taking all the backgrounds into account, we require $Y_{5\to4} > 0.005$. The jet configuration for $n = 6$ is used for the rest of the analysis of the 6-jet + lepton mode.

D. Identification of heavy flavor jets

The identification of jet flavor is critical for the suppression of tbW background due to the differences in the number of b jets in the final state between tbW events, which produce two b jets, and our $t\bar{t}H$ signal, which results in four b jets. The jet flavor is identified by looking at the number of secondary tracks belonging to the jet. We count the number of tracks with an impact parameter significance (in three dimensions) greater than a certain threshold value Q; if the count is equal to or greater than a certain number N_Q, the jet is identified as a b jet. The two numbers (Q, N_Q) are optimized in our b-tagging selection.

In the 6-jet + lepton analysis, we define two criteria for the identification of b jets. We define *loose* b jets as jets passing the b-tagging requirement of $(Q, N_Q) = (2.0, 2)$; *tight* b jets are defined by those passing the b-tagging requirement of $(Q, N_Q) = (2.5, 4)$. Note that the set of tight b jets is a subset of loose b jets. We require at least four loose b jets in the event; two out of the four are also required to pass the tight b-jet criteria.

The b-tagging efficiency is estimated using a sample of $Z \to q\bar{q}$ events at $\sqrt{s} = 91.2$ GeV and is found to be 81% (47%) for the loose (tight) selection. The rate of incorrectly identifying a jet originating from a lighter quark as a fake b jet is estimated to be 40% (3.2%) in a sample of c jets, and 0.5% (0.08%) for s, d, and u-jet samples for the loose (tight) selection. In a multijet environment, the b-tagging efficiencies decrease due to overlapping jets. For 6-jet + lepton and 8-jet events, this effect typically reduces the b-tagging efficiencies by roughly 10%.

Because of the large fake rate of b jets coming from c jets, final states involving charm quarks, such as $t\bar{t}(Z \to c\bar{c})$ and $t\bar{t}(g^* \to c\bar{c})$, can be an additional source of background. We do not include these backgrounds in this study since we expect their impact to be reduced with the use of more sophisticated b-tagging methods [29]. According to the simulation studies by the International Large Detector Concept Group [24], which employ the methods of Ref. [29], the purity of b jets is found to be 80% (98%) in a sample of $Z \to q\bar{q}$ events at $\sqrt{s} = 91.2$ GeV, corresponding to the fake rate of at most 25% (1.3%) due to c jets, for b-tagging efficiencies equivalent to our loose (tight) selection criteria. Similar performances are found in studies by the Silicon Detector Concept Group [25]. While these numbers can be compared with our b-tagging performance, a realistic estimate of background due to c jets requires a future study using full detector simulation.

E. Top and Higgs reconstruction

The Higgs candidate ($H \to b\bar{b}$) is formed by requiring one tight b jet and one loose b jet. The hadronic top candidate is formed by combining three jets, one of which

TABLE III. Summary of cuts in the analysis of the 6-jet + lepton mode, denoted as $6j$. We denote the 4-jet + 2-lepton mode as $4j$, and the 8-jet mode as $8j$. Estimated yields are given assuming an integrated luminosity of 1 ab^{-1} with beam polarizations $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$. Refer to the text for the details of the b-tagging requirement and the mass cuts.

	$t\bar{t}H$ $(6j)$	$t\bar{t}H$ $(8j)$	$t\bar{t}H$ $(4j)$	tbW	$t\bar{t}Z$	$t\bar{t}g^*$ $(b\bar{b})$
No cuts	282.3	289.5	68.3	980 738.5	2406.9	1159.6
Single isolated lepton	179.6	20.7	28.3	340 069.0	790.6	397.7
Thrust <0.77	145.7	18.5	19.2	144 999.0	616.7	266.0
$Y_{5\to4} > 0.005$	125.5	16.6	9.2	12 297.7	416.2	113.7
b tagging	49.0	1.3	2.9	172.9	53.3	37.8
Mass cuts	39.5	1.2	0.4	23.0	33.9	13.2

must be (at least) a loose b jet. Because there are multiple possible ways to combine the six jets in this way, we define the quantity χ^2 as

$$\chi^2 = \left(\frac{m_{j_1 j_2} - M_H}{\sigma_H}\right)^2 + \left(\frac{m_{j_3 j_4 j_5} - M_t}{\sigma_t}\right)^2 + \left(\frac{m_{j_3 j_4} - M_W}{\sigma_W}\right)^2, \tag{7}$$

and choose the jet combination which minimizes the χ^2 value. Here, m_{jj} (m_{jjj}) is the invariant mass of the two-jet (three-jet) system; the two jets j_1 and j_2 are used to form the Higgs candidate, while j_3, j_4, and j_5 are the three jets used to reconstruct the top candidate which decays hadronically. The masses M_t, M_W, and M_H are taken to be the same values used in the event generation. The widths σ_t, σ_W, and σ_H correspond to the mass resolutions in the case of perfect jet clustering and jet combinations. These values are determined to be $\sigma_t = 14.3$ GeV, $\sigma_W = 9.3$ GeV, and $\sigma_H = 17.7$ GeV by combining the reconstructed four-momenta of final particles in the $e^+e^- \to t\bar{t}H$ sample using Monte Carlo information.

Final cuts are applied to the resulting invariant mass distributions. For the 6-jet + lepton mode, we require the top mass to be in the range of $140 < m_t < 205$ GeV and the Higgs mass to be in the range of $95 < m_H < 150$ GeV, where the range has been optimized in steps of 5 GeV.

F. Results

We summarize the yields after applying each cut for the case of polarized beams $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$ in Table III, where the yields are normalized assuming an integrated luminosity of 1 ab^{-1}. The resulting distributions

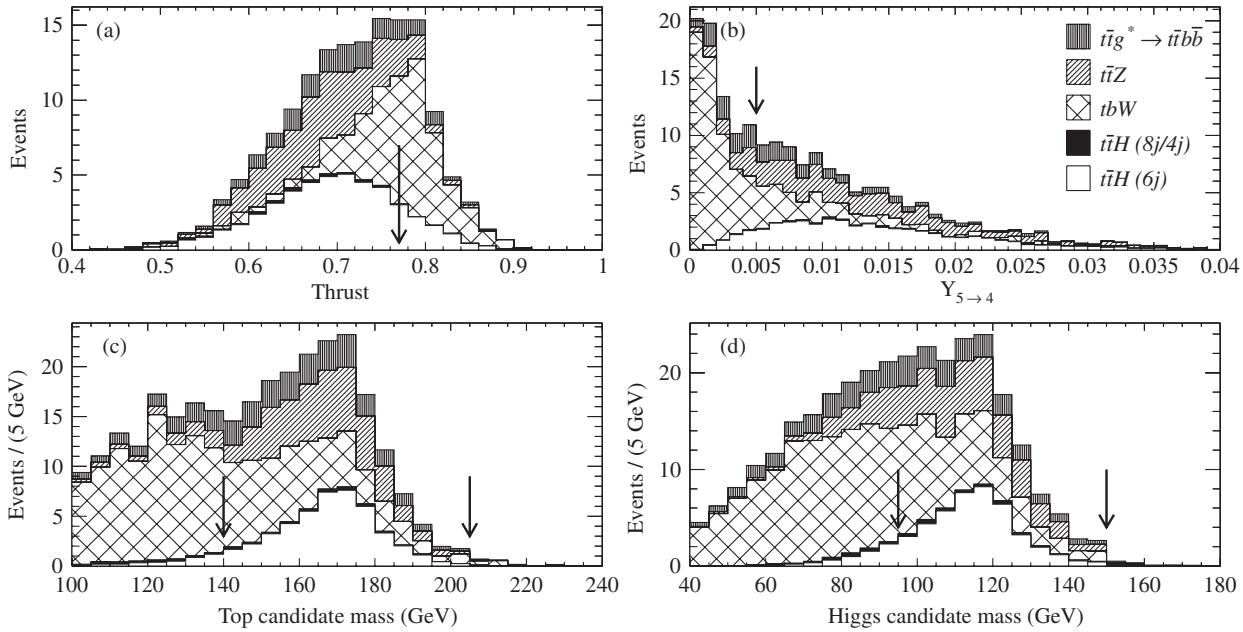

FIG. 5. The distributions of the cut variables in the 6-jet + lepton analysis are shown: (a) thrust, (b) $Y_{5\to4}$, (c) mass of the top candidate, and (d) mass of the Higgs candidate. Each sample is weighted assuming an integrated luminosity of 1 ab^{-1} with beam polarizations $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$. In each of these four plots, all the event selection criteria are applied except for the cut on the variable shown. The arrows indicate the optimized cut values.

for the thrust, $Y_{5\to4}$, the top candidate mass, and the Higgs candidate mass, after applying all the other cuts, are shown in Fig. 5. The signal significance is estimated to be 3.7, corresponding to the measurement accuracy of the top Yukawa coupling of $|\Delta g_t/g_t| = 14\%$. With unpolarized beams $(P_{e^-}, P_{e^+}) = (0.0, 0.0)$, the significance becomes 2.9, corresponding to $|\Delta g_t/g_t| = 17\%$.

VI. ANALYSIS OF THE 8-JET MODE

A. Isolated lepton rejection

In the 8-jet analysis, the 6-jet + lepton mode can become a source of background as a result of splitting the jets by the jet clustering procedure. To reduce this kind of background, we look for isolated leptons using the same prescription used in the 6-jet + lepton analysis. Events containing one or more isolated leptons are discarded in the 8-jet analysis. This procedure ensures that the 6-jet + lepton and 8-jet samples are statistically independent from each other, allowing for a straightforward combination of the two results.

B. Event shape

Similarly to the 6-jet + lepton analysis, the thrust variable T is used to reduce the $t\bar{t}$ background. It is found that $T < 0.7$ is optimal for the 8-jet mode.

C. Jet clustering

In the 8-jet mode analysis, the jet clustering is performed over all particles in the event to form eight jets. We keep the jet transition values $Y_{n\to n-1}$. The value for $Y_{8\to7}$ is found to be useful in discriminating $t\bar{t}H$ events from $t\bar{t}$ events. We require $Y_{8\to7} > 0.0009$ in the event selection.

D. Identification of heavy flavor jets

We follow a similar procedure as in the 6-jet + lepton mode for the identification of b jets. We use a different optimization for the tight b jet, which is modified to be $(Q, N_Q) = (3.0, 2)$. The definition of the loose b jet remains the same.

E. Top and Higgs reconstruction

The Higgs candidate ($H \to b\bar{b}$) is formed by requiring one tight b jet and one loose b jet. One of the top candidates is required to contain a tight b jet, while the other top is required to have (at least) a loose b jet. Because there are multiple possible combinations of jets, we define the quantity χ^2 similarly to the 6-jet + lepton mode as

$$\chi^2 = \left(\frac{m_{j_1 j_2} - M_H}{\sigma_H}\right)^2 + \left(\frac{m_{j_3 j_4 j_5} - M_t}{\sigma_t}\right)^2 + \left(\frac{m_{j_6 j_7 j_8} - M_t}{\sigma_t}\right)^2 + \left(\frac{m_{j_3 j_4} - M_W}{\sigma_W}\right)^2 + \left(\frac{m_{j_6 j_7} - M_W}{\sigma_W}\right)^2, \quad (8)$$

and choose the combination of jets which minimizes the χ^2 value. Here, j_1 and j_2 are used to form the Higgs candidate. The three jets j_3, j_4, and j_5 are used to reconstruct the first top candidate, while j_6, j_7, and j_8 are used to reconstruct the second top candidate. The same values for σ_t, σ_W, and σ_H are used as in the 6-jet + lepton analysis.

Final cuts are applied on the invariant mass of the top and Higgs candidate. For both top candidates, the mass is required to be in the range of $140\ \text{GeV} < m_{jjj} < 215\ \text{GeV}$. The Higgs candidate mass is required to be in the range of $80\ \text{GeV} < m_{jj} < 150\ \text{GeV}$.

F. Results

The estimated signal yields are summarized in Table IV for the case of polarized beams $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$, assuming an integrated luminosity of $1\ \text{ab}^{-1}$. The resulting distributions for the thrust, $Y_{8\to7}$, the top mass, and the Higgs mass are shown in Fig. 6. The signal significance in the 8-jet mode is 3.7, corresponding to the measurement accuracy of the top Yukawa coupling of $|\Delta g_t/g_t| = 14\%$. With unpolarized beams $(P_{e^-}, P_{e^+}) = (0.0, 0.0)$, the significance becomes 2.8, corresponding to $|\Delta g_t/g_t| = 18\%$.

VII. CONCLUSIONS

We have evaluated the accuracy of the top Yukawa coupling at $\sqrt{s} = 500\ \text{GeV}$, taking into account the $t\bar{t}$ bound-state effects for the $e^+e^- \to t\bar{t}H$ signal sample as well as the $e^+e^- \to t\bar{t}Z$ background sample. Other

TABLE IV. Summary of cuts in the analysis of the 8-jet mode, denoted as $8j$. We denote the 6-jet + lepton mode as $6j$, and the 4-jet + 2-lepton mode as $4j$. Estimated yields are given assuming an integrated luminosity of $1\ \text{ab}^{-1}$ with beam polarizations $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$. Refer to the text for the details of the b-tagging requirement and the mass cuts.

	$t\bar{t}H$ $(8j)$	$t\bar{t}H$ $(6j)$	$t\bar{t}H$ $(4j)$	tbW	$t\bar{t}Z$	$t\bar{t}g^*$ $(b\bar{b})$
No cuts	289.5	282.3	68.3	980 738.5	2406.9	1159.6
Reject isolated leptons	266.3	85.6	6.6	589 716.0	1351.4	701.2
Thrust <0.7	167.7	44.0	2.7	107 227.0	818.0	311.5
$Y_{8\to7} > 0.0009$	113.8	13.0	0.3	4048.1	349.6	67.1
b tagging	66.6	6.8	0.1	442.6	77.6	39.8
Mass cuts	50.1	0.4	0.0	75.6	47.6	14.1

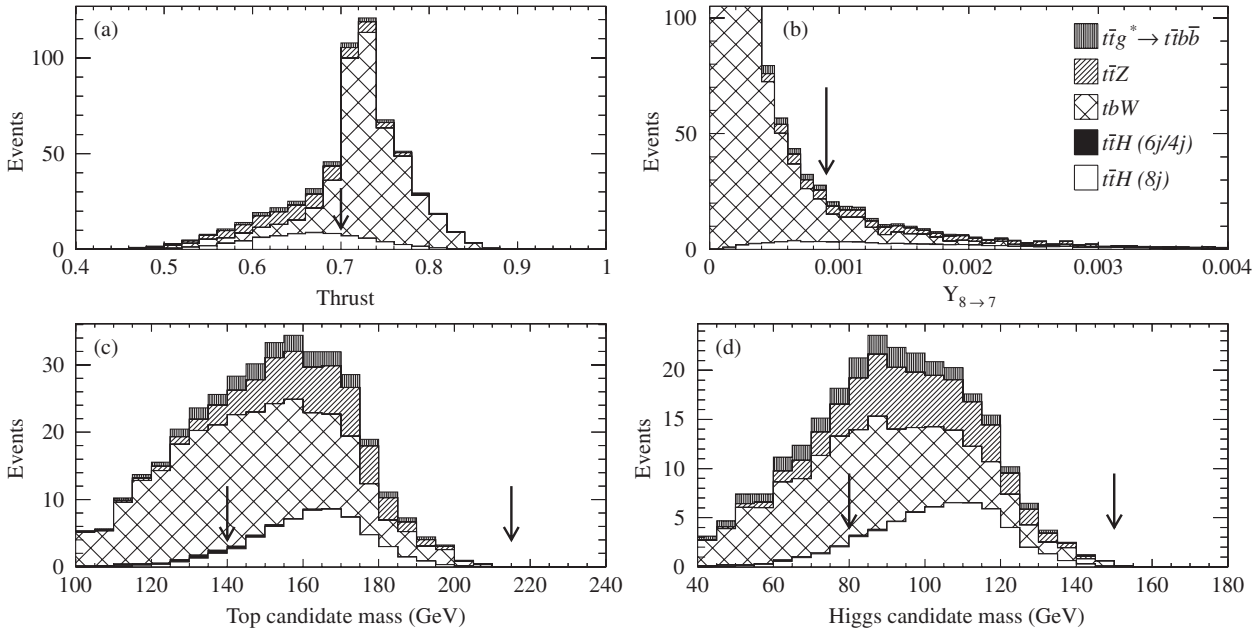

FIG. 6. The distributions of the cut variables in the 8-jet analysis are shown: (a) thrust, (b) $Y_{8\to7}$, (c) mass of the top candidate, and (d) mass of the Higgs candidate. Each sample is weighted assuming an integrated luminosity of 1 ab^{-1} with beam polarizations $(P_{e^-}, P_{e^+}) = (-0.8, +0.3)$. In each of these four plots, all the event selection criteria are applied except for the cut on the variable shown. The arrows indicate the optimized cut values.

backgrounds considered were $e^+e^- \to t\bar{b}W^-/\bar{t}bW^+ \to bW^+\bar{b}W^-$ and $e^+e^- \to t\bar{t}g^* \to bW^+\bar{b}W^-b\bar{b}$. A simple cut-and-count analysis was performed for the 6-jet + lepton and 8-jet signal decay modes. We assume an integrated luminosity of 1 ab^{-1}. Because the 6-jet + lepton sample and the 8-jet sample are statistically independent, the combined significance can be computed simply by summing the significances of the two modes in quadrature, assuming Gaussian statistics.

With polarized beams $(P_{e^-}, P_{e^+}) = (-0.8, 0.3)$, the combined significance is 5.2, corresponding to the measurement accuracy of the top Yukawa coupling of $|\Delta g_t/g_t| = 10\%$. With unpolarized beams $(P_{e^-}, P_{e^+}) = (0.0, 0.0)$, the combined significance becomes 4.0, corresponding to $|\Delta g_t/g_t| = 13\%$. Note that these numbers only take into account the statistical uncertainty.

The $e^+e^- \to t\bar{t}Z$ and $e^+e^- \to t\bar{t}g^*$ backgrounds survive the event selection procedure primarily because of the overlapping of the dijet mass for the Higgs candidate. This can be reduced by improving the jet energy resolution and the jet clustering procedures, which in turn improves the mass resolution of the Higgs candidate. The production rate of the $e^+e^- \to tbW$ process has a large systematic uncertainty, and therefore must be modeled accurately, particularly in the tails of its kinematically allowed region.

Our results indicate that the measurement of the top Yukawa coupling is possible down to the 10% level of statistical precision at the ILC with $\sqrt{s} = 500$ GeV after taking into account the $t\bar{t}$ bound-state effects, which agrees with previous predictions [15]. It will be critical to reduce the systematic effects down to the level comparable to the statistical uncertainties. We expect the systematic uncertainties coming from the determination of the background rates to be the dominant effect. The amount of tbW background can be estimated by measuring the tbW cross section at $\sqrt{s} = 500$ GeV. The $t\bar{t}$ bound-state effects must also be verified by measuring the $t\bar{t}$ cross section at its production threshold ($\sqrt{s} \approx 350$ GeV) which will be used to estimate the rate of the $e^+e^- \to t\bar{t}H$ signal and the $e^+e^- \to t\bar{t}Z$ background. For this, it will be necessary to measure the differential cross section of $e^+e^- \to t\bar{t}$ in order to separate the Higgs-exchange contribution via the t channel which itself contains the top Yukawa coupling.

On the theoretical front, it will be desirable to reduce the uncertainties in the production cross section coming from loop corrections, which will be critical for precise background estimation. For the $e^+e^- \to t\bar{t}$ process, the electroweak corrections are known at the one-loop level [30], with further improvements expected in the coming years. QCD corrections are already known at the three-loop level [31–35]. For the $e^+e^- \to t\bar{t}Z$ process, the known QCD corrections at the one-loop level [36] include the $t\bar{t}$ bound-state effects. Since our study also incorporates the $t\bar{t}$ bound-state effects, it will be necessary to calculate the higher order corrections in order to properly estimate the theoretical uncertainties in the $e^+e^- \to t\bar{t}Z$ cross section.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to all members of the ILC physics subgroup [37] for useful discussions. Among them, A. Ishikawa and S. Uozumi deserve

special mention for their work during the initial stages of this work. The authors would like to thank J. Kanzaki for discussions on the discovery potential of the $t\bar{t}H$ process at the LHC. This work is supported in part by the Creative Scientific Research Grant No. 18GS0202 of the Japan Society for Promotion of Science (JSPS), the JSPS Core University Program, and the JSPS Grant-in-Aid for Scientific Research No. 22244031.

[1] D. Benedetti *et al.*, J. Phys. G **34**, N221 (2007).
[2] E. Gross and L. Živković, Eur. Phys. J. C **59**, 731 (2009).
[3] K. Hagiwara, H. Murayama, and I. Watanabe, Nucl. Phys. **B367**, 257 (1991).
[4] A. Djouadi, J. Kalinowski, and P. M. Zerwas, Mod. Phys. Lett. A **7**, 1765 (1992).
[5] A. Juste and G. Merino, arXiv:hep-ph/9910301.
[6] A. Gay, Eur. Phys. J. C **49**, 489 (2007).
[7] H. Baer, S. Dawson, and L. Reina, Phys. Rev. D **61**, 013002 (1999).
[8] S. Dittmaier, M. Krämer, Y. Liao, M. Spira, and P. M. Zerwas, Phys. Lett. B **441**, 383 (1998).
[9] S. Dawson and L. Reina, Phys. Rev. D **59**, 054012 (1999).
[10] G. Bélanger *et al.*, Phys. Lett. B **571**, 163 (2003).
[11] A. Denner, S. Dittmaier, M. Roth, and M. M. Weber, Nucl. Phys. **B680**, 85 (2004).
[12] Y. You *et al.*, Phys. Lett. B **571**, 85 (2003).
[13] C. Farrell and A. H. Hoang, Phys. Rev. D **72**, 014007 (2005).
[14] C. Farrell and A. H. Hoang, Phys. Rev. D **74**, 014008 (2006).
[15] A. Juste *et al.*, *Proceedings of the 2005 International Linear Collider Physics and Detector Workshop and 2nd ILC Accelerator Workshop (Snowmass 2005)*, econf C0508141, PLEN0043 (2005).
[16] Y. Sumino and H. Yokoya, J. High Energy Phys. 09 (2010) 034.
[17] J. Brau *et al.*, Report No. ILC-REPORT-2007-001.
[18] http://www-jlc.kek.jp/subg/offl/physsim/.
[19] H. Murayama, I. Watanabe, and K. Hagiwara, KEK Report No. 91-11, 1992.
[20] S. Kawabata, Comput. Phys. Commun. **41**, 127 (1986).
[21] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.
[22] http://www-jlc.kek.jp/subg/offl/jsf/.
[23] J.-C. Brient and H. Videau, in *Proceedings of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)*, eConf C010630, E3047 (2001).
[24] T. Abe *et al.* (ILD Concept Group-Linear Collider), arXiv:1006.3396.
[25] H. Aihara *et al.* (SiD Concept Group-Linear Collider), arXiv:0911.0006.
[26] S. Brandt, Ch. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).
[27] E. Farhi, Phys. Rev. Lett. **39**, 1587 (1977).
[28] S. Catani, Yu. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. B **269**, 432 (1991).
[29] D. Bailey *et al.* (LCFI Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **610**, 573 (2009).
[30] W. Beenakker, S. C. van der Marck, and W. Hollik, Nucl. Phys. **B365**, 24 (1991).
[31] G. Grunberg, Y. J. Ng, and S.-H. H. Tye, Phys. Rev. D **21**, 62 (1980).
[32] J. Jersák, E. Laermann, and P. M. Zerwas, Phys. Rev. D **25**, 1218 (1982).
[33] K. G. Chetyrkin, J. H. Kühn, and M. Steinhauser, Nucl. Phys. **B482**, 213 (1996).
[34] K. G. Chetyrkin, J. H. Kühn, and M. Steinhauser, Nucl. Phys. **B505**, 40 (1997).
[35] K. G. Chetyrkin, A. H. Hoang, J. H. Kühn, M. Steinhauser, and T. Teubner, Eur. Phys. J. C **2**, 137 (1998).
[36] L. Dai, W.-G. Ma, R.-Y. Zhang, L. Guo, and S.-M. Wang, Phys. Rev. D **78**, 094010 (2008).
[37] http://www-jlc.kek.jp/subg/physics/ilcphys/.