

# Measuring the electron Yukawa coupling via resonant $s$ -channel Higgs production at FCC-ee

David d'Enterria<sup>1a</sup>, Andres Poldaru<sup>2</sup>, and George Wojcik<sup>3</sup>

<sup>1</sup> CERN, EP Department, 1211 Geneva, Switzerland

<sup>2</sup> LMU, 80539 Munich, Germany

<sup>3</sup> SLAC, 2575 Sand Hill Rd., Menlo Park, CA, 94025 USA

Received: February 8, 2022/ Revised version: February 8, 2022

**Abstract.** The Future Circular Collider (FCC-ee) offers the unique opportunity of studying the Higgs Yukawa coupling to the electron,  $y_e$ , via resonant  $s$ -channel production,  $e^+e^- \rightarrow H$ , in a dedicated run at  $\sqrt{s} = m_H$ . The signature for direct Higgs production is a small rise in the cross sections for particular final states, consistent with Higgs decays, over the expectations for their occurrence due to Standard Model (SM) background processes involving  $Z^*$ ,  $\gamma^*$ , or  $t$ -channel exchanges alone. Performing such a measurement is remarkably challenging for four main reasons. First, the low value of the  $e^\pm$  mass leads to a tiny  $y_e$  coupling, and correspondingly small cross section:  $\sigma_{ee \rightarrow H} \propto m_e^2 = 0.57$  fb accounting for initial-state  $\gamma$  radiation. Second, the  $e^+e^-$  beams must be monochromatized such that the spread of their center-of-mass (c.m.) energy is commensurate with the narrow width of the SM Higgs boson,  $\Gamma_H = 4.1$  MeV, while keeping large beam luminosities. Third, the Higgs mass must also be known beforehand with a few-MeV accuracy in order to operate the collider at the resonance peak,  $\sqrt{s} = m_H$ . Last but not least, the cross sections of the background processes are many orders-of-magnitude larger than those of the Higgs decay signals. A preliminary generator-level study of 11 Higgs decay channels using a multivariate analysis, which exploits boosted decision trees to discriminate signal and background events, identifies two final states as the most promising ones in terms of statistical significance:  $H \rightarrow gg$  and  $H \rightarrow WW^* \rightarrow \ell\nu + 2$  jets. For a benchmark monochromatization with 4.1-MeV c.m. energy spread (leading to  $\sigma_{ee \rightarrow H} = 0.28$  fb) and  $10 \text{ ab}^{-1}$  of integrated luminosity, a  $1.3\sigma$  signal significance can be reached, corresponding to an upper limit on the  $e^\pm$  Yukawa coupling at 1.6 times the SM value:  $|y_e| < 1.6|y_e^{\text{SM}}|$  at 95% confidence level, per FCC-ee interaction point per year. Directions for future improvements of the study are outlined.

**PACS.** 14.80.Bn Standard Model Higgs boson – 14.60.Cd electrons and positrons

## 1 Introduction

The usual claim that the Brout–Englert–Higgs mechanism of mass generation for elementary particles [1, 2, 3] has been experimentally confirmed at the Large Hadron Collider (LHC) thanks to the Higgs boson discovery [4, 5], and subsequent studies of its properties [6, 7], is only valid so far for the heaviest Standard Model (SM) particles:  $W$  and  $Z$  weak bosons, and quarks and leptons of the third family ( $t$ ,  $b$ , and  $\tau$ ). Today, not only the generation of all neutrino masses remains a mystery [8], but at the end of the LHC lifetime only a fraction of the Higgs Yukawa couplings to the second-family fermions (the muon and, maybe, the charm quark) will have been probed. On the other hand, due to their low masses and thereby small Yukawa couplings to the Higgs field, the mass generation mechanism for the stable matter of the visible universe, composed of  $u$  and  $d$  quarks plus the electron and neutrinos ( $\nu$ ), will remain experimentally untested. The smallest Yukawa coupling, aside from the Dirac  $\nu$ 's case, is that of the electron given by  $y_e = \sqrt{2}m_e/v = 2.9 \cdot 10^{-6}$  for  $m_e = 0.511 \cdot 10^{-3}$  GeV and Higgs vacuum expectation value  $v = (\sqrt{2}G_F)^{-1/2} = 246.22$  GeV. Measuring the Higgs coupling to the electron is impossible at hadron colliders because the  $H \rightarrow e^+e^-$  decay has a very branching fraction of  $\mathcal{B}(H \rightarrow e^+e^-) = 5.22 \cdot 10^{-9}$  (see Eq. (2) below), and is completely swamped by a Drell–Yan  $e^+e^-$  continuum whose cross section is many orders of magnitude larger. Measurements in p-p collisions at the LHC, assuming the SM Higgs production cross section, lead to an upper bound on the branching fraction of  $\mathcal{B}(H \rightarrow e^+e^-) < 3.6 \cdot 10^{-4}$  at 95% confidence level (CL), corresponding to an upper limit on the Yukawa coupling  $y_e \propto \mathcal{B}(H \rightarrow e^+e^-)^{1/2}$  of 260 times the SM value [9, 10]. Such a constraint can be translated into a lower bound on the energy scale of any physics beyond the SM (BSM) affecting  $y_e$ , of  $\Lambda_{\text{BSM}} \approx v^{3/2}(\sqrt{2}m_e \cdot (y_e/y_e^{\text{SM}}))^{-1/2} \gtrsim 8.8$  TeV [11]. Assuming that the sensitivity to the  $H \rightarrow e^+e^-$  decay scales with the square root of the integrated luminosity, the high-luminosity LHC phase with a  $\mathcal{L}_{\text{int}} = 3 \text{ ab}^{-1}$  data sample [12] will result in  $y_e \lesssim 120y_e^{\text{SM}}$  (i.e.  $\Lambda_{\text{BSM}} \gtrsim 13$  TeV).

<sup>a</sup> Corresponding author: david.d'enterria@cern.ch

The possibility of studying resonant Higgs production at leptons colliders has been considered in the literature so far only for  $\mu^+\mu^-$  annihilation at  $\sqrt{s} = m_H$ , notably as a means to directly and precisely measure  $\Gamma_H$ ,  $m_H$ , and the muon Yukawa coupling, by exploiting a large peak production cross section of  $\sigma_{\mu\mu \rightarrow H} = 70$  pb [13]. The same measurement at an  $e^+e^-$  machine had never been seriously considered given the sub-femtobarn cross section for the  $e^+e^- \rightarrow H$  process, suppressed by at least a factor  $m_e^2/m_\mu^2$  compared to the muon collider case. Notwithstanding this difficulty, when the FCC-ee was first proposed [14], it was noticed that the unparalleled integrated luminosities of about  $\mathcal{L}_{\text{int}} = 10 \text{ ab}^{-1}$ /year available at  $\sqrt{s} = 125 \text{ GeV}$ , would make it possible to attempt an observation of the direct production of the scalar boson [15, 16]. Such a consideration motivated a few subsequent works on various  $e^+e^- \rightarrow H$  theoretical [11, 17, 18, 19] and accelerator [20, 21] aspects.

The Feynman diagram for  $s$ -channel Higgs production (and dominant decays) is shown in Fig. 1 (left). Other  $e^+e^- \rightarrow H$  production processes, through W and Z loops, are suppressed by the electron mass for on-shell external fermions and have negligible cross sections [11]. The resonant Higgs cross section at any given c.m. energy  $\sqrt{s}$  is theoretically given by the relativistic Breit–Wigner (BW) expression:

$$\sigma_{ee \rightarrow H} = \frac{4\pi\Gamma_H\Gamma(H \rightarrow e^+e^-)}{(s - m_H^2)^2 + m_H^2\Gamma_H^2}, \quad (1)$$

where  $\Gamma_H = 4.1 \text{ MeV}$  is the total Higgs width [22],  $m_H = 125 \text{ GeV}$  its mass, and the partial decay width  $\Gamma(H \rightarrow e^+e^-)$ , given by the tree-level relation

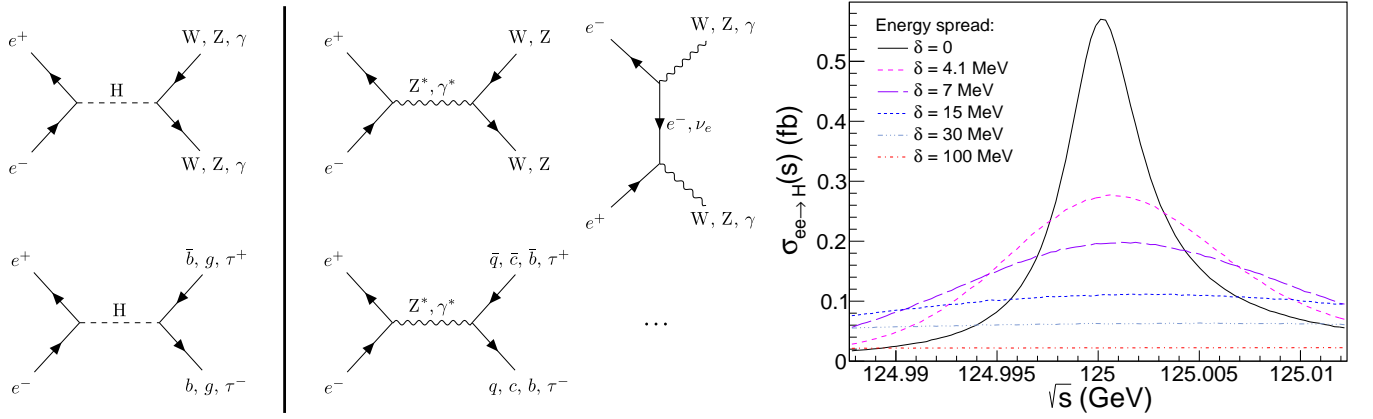
$$\Gamma(H \rightarrow e^+e^-) = \frac{G_F m_H m_e^2}{4\sqrt{2}\pi} \left(1 - \frac{4m_e^2}{m_H^2}\right)^{3/2} = 2.14 \cdot 10^{-11} \text{ GeV}, \quad (2)$$

is tiny due to its dependence on the square of the  $e^\pm$  mass. From the BW expression (1), it is clear that an accurate knowledge of the  $m_H$  value is critical to maximize the resonant cross section. Combining three  $e^+e^- \rightarrow HZ$  measurements at FCC-ee (recoil mass, peak cross section, and threshold scan), a  $\mathcal{O}(2 \text{ MeV})$  mass precision is achievable [23] before a dedicated  $e^+e^- \rightarrow H$  run. In addition, the FCC-ee beam energies will be monitored with a relative precision of  $10^{-6}$  [24], warranting a sub-MeV accuracy of the exact point in the Higgs lineshape being probed at any moment. Taking  $m_H = 125 \text{ GeV}$ , Eq. (1) gives  $\sigma_{ee \rightarrow H} = 4\pi\mathcal{B}(H \rightarrow e^+e^-)/m_H^2 = 1.64 \text{ fb}$  as peak cross section. Two effects, however, lead to a significant broadening of the Born-level result: (i) initial-state  $\gamma$  radiation (ISR) reduces the cross section and generates an asymmetry of the Higgs lineshape, and (ii) the actual beams are never perfectly monoenergetic, i.e. the collision  $\sqrt{s}$  has a spread  $\delta\sqrt{s}$  around its center value, further leading to a smearing of the BW peak. The reduction of the BW cross section due to IS photon emission(s) is of factor of 0.35 and leads to  $\sigma_{ee \rightarrow H} = 0.57 \text{ fb}$  [17]. The additional impact of a given c.m. energy spread on the Higgs BW shape can be quantified through the convolution of BW and Gaussian distributions, i.e. a relativistic Voigtian function [25]. Figure 1 (right) shows the Higgs lineshape for various  $\delta\sqrt{s}$  values. The combination of ISR plus  $\delta\sqrt{s} = \Gamma_H = 4.1 \text{ MeV}$  reduces the peak Higgs cross section by a total factor of 0.17, down to  $\sigma_{ee \rightarrow H} = 0.28 \text{ fb}$ . As a baseline study, we will use this latter value as our default expectation for the signal production cross section, and compute the corresponding significance for a 1-year operation with  $10 \text{ ab}^{-1}$  integrated luminosities per FCC-ee interaction point (IP). The computed signal yields and associated significances can then be subsequently rescaled to any other choice of  $(\delta\sqrt{s}, \mathcal{L}_{\text{int}})$  values given by the chosen beam monochromatization scheme [20, 21].

## 2 Analysis strategy. Simulation of signal and background processes

The strategy to observe the resonant production of the Higgs boson is based on identifying final states in  $e^+e^-$  collisions at  $\sqrt{s} = m_H$ , consistent with any of the H decay modes, that lead to a small increase (but, hopefully, statistically significant when combined together) of the measured cross sections with respect to the theoretical expectation for their occurrence via background processes alone, involving  $Z^*$ ,  $\gamma^*$ , or  $t$ -channel exchanges (Fig. 1, center diagrams). The assumption is that, after various years of FCC-ee operation at the Z pole and HZ c.m. energies [26, 27], the theoretical knowledge of the overwhelming background cross sections will be at the  $10^{-5}$  level or better [28], and that experimental systematic uncertainties (detector acceptance, reconstruction efficiencies, luminosity, etc.) will be controlled at the same level of precision [27, 29] and/or will partially cancel out in ratios of number of signal over backgrounds yields. Under such circumstances, the proposed measurement can be considered as a very-rare “counting experiment” that aims at adding up the individual statistical significances for various final states consistent with known Higgs decay channels in the hope to observe an excess above the background counts expectations.

In order to carry out our simulation studies, we generate individual samples of  $10^5$ – $10^7$   $e^+e^-$  annihilation events at  $\sqrt{s} = m_Z = 125.00 \text{ GeV}$  with the PYTHIA 8 Monte Carlo (MC) code [30], for each of the 11 final states for signal and associated backgrounds listed in Table 1. The Higgs decay branching fractions used are those from the HDECAY



**Fig. 1.** Typical diagrams for the direct Higgs channel production (left) decaying into electroweak bosons (top) and fermions or gluons (bottom), and associated backgrounds (center), considered in this work. Right: Resonant Higgs production cross section, including ISR effects, for several values of the  $e^+e^-$  c.m. energy spread  $\delta_{\sqrt{s}} = 0, 4.1, 7, 15, 30$ , and  $100$  MeV [17].

code at NLO accuracy [31]. The PYTHIA 8 signal cross sections are absolutely normalized to match our benchmark  $\sigma_{ee \rightarrow H} = 0.28$  fb value for ISR plus  $\delta_{\sqrt{s}} = 4.1$ -MeV energy spread discussed above (second curve of Fig. 1 right). Higgs decay modes not listed in Table 1 are either completely swamped by background (e.g.  $H \rightarrow ZZ^* \rightarrow 4j$ ) or have too low  $B$ 's (e.g.  $H \rightarrow ZZ^* \rightarrow 4\ell$ ) and thereby have zero expected counts for any realistic integrated luminosity. The generator-level background cross sections in Table 1 are indicatively quoted without ISR to avoid artificial enhancements of their values due to radiative-returns to the Z pole, which can be easily removed experimentally (e.g. tagging the ISR photon and/or imposing requirements on the total energy of the event). The last column lists the indicative signal-over-background ( $S/B$ ) expected for the dominant (irreducible) background of each channel, at the generator level without any analysis cuts. Three broad categories can be identified:

- i) Final states with pairs of jets or tau leptons, with very large backgrounds leading to  $S/B \approx 10^{-7}$ – $10^{-5}$ , except for the  $H \rightarrow gg$  case for which no actual physical background exists ( $Z^*, \gamma^*$  do not couple to gluons), but for an experimental misidentification probability of light-quarks for gluons that we take as 1% (Table 2);
- ii) Final states from intermediate  $WW^*$  decays, with  $S/B \approx 10^{-3}$ ;
- iii) Final states from intermediate  $ZZ^*$  decays with  $S/B \approx 10^{-2}$ , but very small signal cross sections.

In addition, the last row of the table lists the Higgs diphoton decay mode (discovery channel at the LHC) that suffers from both, a tiny signal cross section and 8 orders-of-magnitude larger backgrounds. A swift analysis of this table allows one to identify two channels with some potentiality in terms of statistical significances,  $H \rightarrow gg$  and  $H \rightarrow WW^* \rightarrow \ell\nu 2j$ , which both feature  $\sim 25$ -ab cross sections and  $S/B \approx 10^{-3}$ .

**Table 1.** Cross sections (including ISR and  $\delta_{\sqrt{s}} = 4.1$  MeV) times branching fractions ( $B$ ) for 11 final states in  $e^+e^- \rightarrow H(XX)$  signal processes and associated dominant  $e^+e^- \rightarrow XX$  backgrounds (without ISR), and ratio of signal-over-background for each channel before any analysis cuts (the digluon  $S/B$  quoted assumes a light- $q \rightarrow g$  mistagging rate of 1%).

Higgs decay channel	$B$	$\sigma \times B$	Irreducible background	$\sigma$	$S/B$
$e^+e^- \rightarrow H \rightarrow b\bar{b}$	58.2%	164 ab	$e^+e^- \rightarrow b\bar{b}$	19 pb	$\mathcal{O}(10^{-5})$
$e^+e^- \rightarrow H \rightarrow gg$	8.2%	23 ab	$e^+e^- \rightarrow q\bar{q}$	61 pb	$\mathcal{O}(10^{-3})$
$e^+e^- \rightarrow H \rightarrow \tau\tau$	6.3%	18 ab	$e^+e^- \rightarrow \tau\tau$	10 pb	$\mathcal{O}(10^{-6})$
$e^+e^- \rightarrow H \rightarrow c\bar{c}$	2.9%	8.2 ab	$e^+e^- \rightarrow c\bar{c}$	22 pb	$\mathcal{O}(10^{-7})$
$e^+e^- \rightarrow H \rightarrow WW^* \rightarrow \ell\nu 2j$	$21.4\% \times 67.6\% \times 32.4\% \times 2$	26.5 ab	$e^+e^- \rightarrow WW^* \rightarrow \ell\nu 2j$	23 fb	$\mathcal{O}(10^{-3})$
$e^+e^- \rightarrow H \rightarrow WW^* \rightarrow 2\ell 2\nu$	$21.4\% \times 32.4\% \times 32.4\%$	6.4 ab	$e^+e^- \rightarrow WW^* \rightarrow 2\ell 2\nu$	5.6 fb	$\mathcal{O}(10^{-3})$
$e^+e^- \rightarrow H \rightarrow WW^* \rightarrow 4j$	$21.4\% \times 67.6\% \times 67.6\%$	27.6 ab	$e^+e^- \rightarrow WW^* \rightarrow 4j$	24 fb	$\mathcal{O}(10^{-3})$
$e^+e^- \rightarrow H \rightarrow ZZ^* \rightarrow 2j 2\nu$	$2.6\% \times 70\% \times 20\% \times 2$	2 ab	$e^+e^- \rightarrow ZZ^* \rightarrow 2j 2\nu$	273 ab	$\mathcal{O}(10^{-2})$
$e^+e^- \rightarrow H \rightarrow ZZ^* \rightarrow 2\ell 2j$	$2.6\% \times 70\% \times 10\% \times 2$	1 ab	$e^+e^- \rightarrow ZZ^* \rightarrow 2\ell 2j$	136 ab	$\mathcal{O}(10^{-2})$
$e^+e^- \rightarrow H \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	$2.6\% \times 20\% \times 10\% \times 2$	0.3 ab	$e^+e^- \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	39 ab	$\mathcal{O}(10^{-2})$
$e^+e^- \rightarrow H \rightarrow \gamma\gamma$	0.23%	0.65 ab	$e^+e^- \rightarrow \gamma\gamma$	79 pb	$\mathcal{O}(10^{-8})$

It is worth noting that the background cross sections computed with PYTHIA 8 for two-particle final states ( $e^+e^- \rightarrow q\bar{q}, c\bar{c}, b\bar{b}, \tau\tau, \gamma\gamma$ ) are found consistent with those obtained running alternative calculators, such as MADGRAPH 5 [32],

[33], but that those for 4-fermion processes with intermediate  $WW^*$  and  $ZZ^*$  are prone to ambiguities in the internal definition of the contributing diagrams, and the ISR treatment, and are not always numerically compatible among them. We trust that such differences will not significantly alter our final results, given that the applied multivariate analysis will remove most non-signal-like topologies, but a dedicated study of 4-fermion backgrounds with an alternative MC generator (such as WHIZARD [34] or KKMC [35]) is left for a forthcoming work. In this context, a few of the quoted background diboson cross sections in Table 1 should be just taken as indicative of the order-of-magnitude irreducible contributions expected for the corresponding Higgs decay.

### 3 Event reconstruction and preselection

Signal and background events are generated, showered, and decayed with PYTHIA 8 (v2.26). Initial state radiation is activated for all backgrounds, and the signal cross section samples are scaled to the ISR-plus-energy-spread benchmark point discussed in Section 1. A detector polar angle acceptance of  $5^\circ \lesssim \theta \lesssim 175^\circ$  is assumed for all final-state particles (defined as those with lifetime  $c\tau_0 > 10$  mm). The FASTJET package [36] is used to reconstruct all jets using the  $k_T$  algorithm [37, 38] in its exclusive variant that clusterizes all hadrons in the event into a prefixed number  $N_j = 2, 4$  of jets (the  $N_j$  choice depends on the particular final state aimed at, e.g.  $H \rightarrow q\bar{q} \rightarrow 2j$ ,  $H \rightarrow WW^*, ZZ^* \rightarrow 2j + \ell/\nu$ , or  $H \rightarrow WW^* \rightarrow 4j$ ). Whenever photons or charged leptons are required to be isolated, standard criteria are applied: the sum of all particles energies must be below 1 GeV within a radius  $\Delta R = 0.25$  around the  $\gamma$  or  $\ell^\pm$  direction. Neutrinos and particles beyond the angular acceptance are added to the missing energy ( $E_{\text{miss}}$ ) 4-vector. The impact of detector (in)efficiencies on the reconstruction of relevant final states is implemented in a simplified manner, according to the performances listed in Table 2. The (mis)tagging jet-flavour performances are beyond the current state-of-the-art

**Table 2.** Bottom ( $b$ ), charm ( $c$ ), and light ( $uds$ ) quarks, gluon ( $g$ ), tau lepton ( $\tau_{\text{had}}$ , hadronically decaying), and photon/electron (mis)reconstruction performances assumed in this study.

	$b$ jets	$c$ jets	gluon jets	$\tau_{\text{had}}$	$\gamma, e^\pm$
reco/tagging efficiency ( $\varepsilon_i$ )	80%	70%	70%	80%	100%
mistagging rates ( $\varepsilon_{j \rightarrow i}^{\text{mistag}}$ )	1% (for $c$ jet)	5% (for $b$ jet)	1% (for $uds$ jets)	$\sim 0\%$ (for $b, c$ -jets)	0.01% ( $e^\pm$ for $\gamma$ )
	0.01% (for $udsg$ jets)	0.1% (for $udsg$ jets)	0.001, 0.01% (for $b, c$ -jets)	$\sim 0\%$ (for $udsg$ jets)	

reached at the LHC today, but reasonably achievable in the “clean” environment of  $e^+e^-$  collisions with dedicated high-precision FCC-ee detectors after various years of operation at the Z pole and HZ energies. More details on the various jet working points assumed are provided in the next section. We note that since the analysis boils down to basically just counting the number of events sharing a given predefined final state, any detector resolution/smearing effects on kinematic properties of the reconstructed objects (jets,  $\ell^\pm$ ,  $\gamma, \dots$ ) impact identically signals and backgrounds, will be very well controlled comparing real data and simulations, and can be accounted for here just through a (small) assigned systematic uncertainty on the final yields when computing the final statistical significance of each channel.

In Table 3, we list the criteria applied to all signal and backgrounds events aiming at a first preselection of final-state topologies consistent with each considered Higgs decay channel. The goal of this first set of cuts is to remove reducible backgrounds as much as possible, while keeping the largest possible signal cross section. For the  $H \rightarrow \tau\tau$  channel, we consider only the fully hadronic ( $\tau_{\text{had}}\tau_{\text{had}}$ ) decay, which is  $0.65 \cdot 0.65 / (0.35 \cdot 0.35) \approx 3.5$  times more probable than the fully leptonic one  $H \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$  (that has thereby a negligible number of signal counts expected after cuts). The last column quotes the approximate percentage of cross section signal retained by the chosen criteria.

### 4 Multivariate analysis (MVA) per channel

For each reconstructed event of all generated MC samples passing the aforementioned preselection criteria per target Higgs channel, we define  $\mathcal{O}(50)$  variables for single and combined ( $n$ -wise) physics objects (jets, charged leptons, photons, neutrinos), as well as for global event properties, in order to provide as much information as possible to a subsequent MVA used to discriminate signal and the remaining backgrounds. The defined variables include kinematic components ( $p_T, \eta, \phi, E$ ), charge, mass (invariant and transverse),... for each single object —as well the same quantities for sums and differences of 4-momenta of selected  $n$ -wise objects combinations—, the maximum and minimum values of  $p_T^{i(ij)}, \eta^{i(ij)}, \phi^{i(ij)}, m_{ij}, \dots$  in the event for all (pairs of) objects  $i(ij)$ , as well as quantities associated with global event topologies (sphericity, linearity, aplanarity, thrust max/min,...). Angular information is particularly useful in diboson channels with decay leptons in order to separate final states coming through the spin-0 Higgs resonance or proceeding through  $t$ -channel processes or via the spin-1  $s$ -channel continuum and/or  $Z^*, \gamma^*, W^\pm$  decays. For such cases, angular discrimination variables based on the Matrix Element Likelihood Analysis (MELA) [39] are also defined

**Table 3.** Minimal event final-state definition for each considered Higgs decay channel, and associated preselection efficiency (after acceptance, and reconstruction (in)efficiencies of Table 2). The  $\ell^\pm$  symbol indicates  $e^\pm, \mu^\pm, \tau_{\text{lep}}^\pm$  charged leptons.

Target Higgs decay	Final state definition	Signal presel. efficiency
$H \rightarrow b\bar{b}$	2 (excl.) jets, 1 $b$ -tagged jet, no $\tau_{\text{had}}$	80%
$H \rightarrow gg$	2 (excl.) gluon-tagged jets, 0 isolated $\ell^\pm$	50%
$H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$	Exactly 2 $\tau_{\text{had}}$ , 0 isolated $\ell^\pm$	65%
$H \rightarrow c\bar{c}$	2 (excl.) jets, 1 $c$ -tagged jet, no $\tau_{\text{had}}$	70%
$H \rightarrow WW^* \rightarrow \ell\nu 2j$	1 isolated $\ell^\pm$ , $E_{\text{miss}} > 2$ GeV, 2 (excl.) jets	$\sim 100\%$
$H \rightarrow WW^* \rightarrow 2\ell 2\nu$	2 isolated opp.-charge $\ell^\pm$ , $E_{\text{miss}} > 2$ GeV, 0 non-isol. $\ell^\pm$ , 0 charged hadrons	$\sim 100\%$
$H \rightarrow WW^* \rightarrow 4j$	4 (excl.) jets, $\geq 1$ $c$ -tag jets, 0 $b$ -, $g$ -tag jets; jets with $m_{j1j2} \approx m_W$ not both $c$ -tagged, 0 $\tau_{\text{had}}$ , 0 isolated $\ell^\pm$	70%
$H \rightarrow ZZ^* \rightarrow 2j 2\nu$	2 (excl.) jets, $E_{\text{miss}} > 30$ GeV, 0 isolated $\ell^\pm$ , 0 $\tau_{\text{had}}$	$\sim 100\%$
$H \rightarrow ZZ^* \rightarrow 2\ell 2j$	2 isolated opposite-charge $\ell^\pm$ , 2 (excl.) jets, 0 $\tau_{\text{had}}$	$\sim 100\%$
$H \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	2 isolated opp.-charge $\ell^\pm$ , $E_{\text{miss}} > 2$ GeV, 0 non-isol. $\ell^\pm$ , 0 charged hadrons	$\sim 100\%$
$H \rightarrow \gamma\gamma$	2 (excl.) isolated photons	$\sim 100\%$

and incorporated into the MVA. We used the TMVA framework [40] to train and test boosted decision-tree (BDT) classifiers in order to provide statistical discrimination between each Higgs decay channel and all relevant background final states, and maximize the signal significance. Examples of the BDT variables used for a particular channel ( $H \rightarrow WW^* \rightarrow \ell\nu 2j$ ) are shown in Fig. 2 (right) later, as well as listed with their individual relative weights in the final signal significance in Table 5.

Table 4 lists the number of signal and background(s) events expected after preselection and BDT output cuts, for 9 different final states. We omit the  $H \rightarrow \gamma\gamma, c\bar{c}$  channels from the table given that they are fully swamped by backgrounds and have a negligible statistical significance. The first observation is that except for the  $H \rightarrow b\bar{b}$  decay, which is anyway overwhelmed by the continuum background, the final number of signal events is (well) below 100 counts for each individual channel and that the remaining backgrounds counts are orders-of-magnitude larger. Therefore, the leading uncertainty of the signal will be of statistical nature, and evidence of any excess will rely on an accurate control of the background systematic uncertainties (which must be well below the statistical ones). Among the listed channels, we observe that  $H \rightarrow gg$  and  $H(WW^*) \rightarrow \ell\nu 2j$  feature the largest  $S/\sqrt{B}$  significances<sup>1</sup>, and are discussed in more detail in dedicated subsections below. The  $H \rightarrow b\bar{b}$  channel suffers from a very large irreducible background, the MVA is unable to improve the rejection of the  $e^+e^- \rightarrow b\bar{b}$  continuum much beyond the preselection result, and the final statistical significance remains very low ( $S/\sqrt{B} \approx 0.12$ ). Although orders-of-magnitude smaller, we also quote the number of misidentified  $e^+e^- \rightarrow c\bar{c}, q\bar{q}$  background events expected for this channel, so as to assess the potential contamination from such processes if the mistagging points assumed in Table 2 are changed. The  $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$  decay mode (as well as, similarly, the  $H \rightarrow c\bar{c}$  one not listed) suffers from very low signal counts and a daunting continuum background that yields a negligible statistical significance ( $S/\sqrt{B} \approx 0.02$ ). Among  $H \rightarrow WW^*$  final states, the fully leptonic one ( $2\ell 2\nu$ ) features the smallest branching fraction, and thereby very low final signal counts. For the two others, lepton+jets ( $\ell\nu 2j$ ) and fully hadronic ( $4j$ ) decays, although they have the same branching fraction, only the former can take advantage of background removal by exploiting the different  $W^\pm \rightarrow \ell^\pm$  decay lepton polarizations for signal and background processes, as explained below. Finally, Table 4 shows that the  $H \rightarrow ZZ^*$  final states will have less than  $\sim 10$  signal events expected after cuts, over much larger backgrounds, and appear statistically marginal in terms of signal significance.

### Analysis of $e^+e^- \rightarrow H(gg) \rightarrow jj$

At face value, the digluon decay is a very promising signal channel as it has the third most abundant Higgs branching fraction ( $\mathcal{B}(H \rightarrow gg) = 8.2\%$ ), and has no irreducible physical background because  $Z$  and  $\gamma$  bosons do not couple to gluons. However, the production of light quark ( $uds$ ) pairs in the much more abundant  $e^+e^- \rightarrow Z^*, \gamma^* \rightarrow q\bar{q}$  process (with cross sections million times larger than that of the signal, Table 1), jeopardizes the observation of  $H \rightarrow gg$  because experimentally separating jets issuing from the showering and hadronization of light-quarks and from gluons is not perfect<sup>2</sup>. An illustrative case would be the emission of a very hard gluon from each one of the  $Z^* \rightarrow q\bar{q}$  quarks, that could mimic the Higgs digluon final state. Fortunately, in the last years there has been tremendous progress on quark-gluon

<sup>1</sup> The actual significance per channel is computed for a single-bin counting experiment using a profile likelihood approach, as explained in Section 5, but it is numerically consistent with this standard naive estimation.

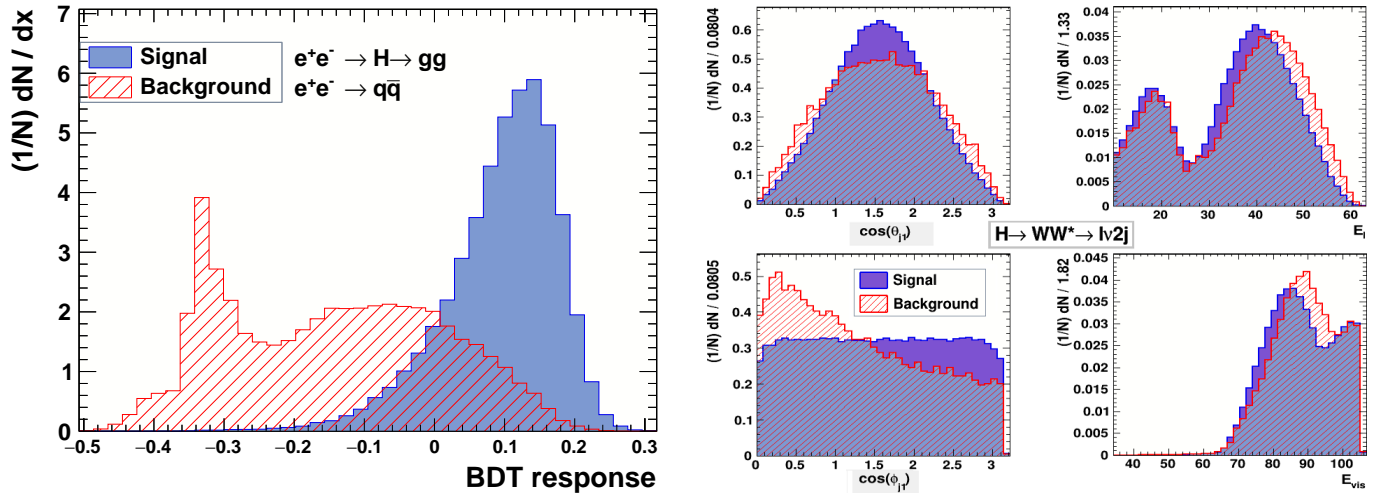
<sup>2</sup> Separating heavy-quark ( $c, b$ ) from gluon jets is easier given the presence of *single* bottom/charm hadrons in the former, which in the gluon case only appear in pairs, through gluon splitting, with suppressed probabilities  $g_{g \rightarrow c\bar{c}, b\bar{b}} \approx 3\%, 0.3\%$  at the



**Table 4.** Number of reconstructed events expected after preselection N(presel.) and BDT output N(MVA) cuts, for  $s$ -channel Higgs decay modes and associated dominant backgrounds in  $e^+e^-$  collisions at  $\sqrt{s} = m_H$  ( $\delta_{\sqrt{s}} = 4.1$  MeV and  $\mathcal{L}_{\text{int}} = 10 \text{ ab}^{-1}$ ).

Channel	N(presel.)	N(MVA)	Channel	N(presel.)	N(MVA)	Channel	N(presel.)	N(MVA)
$H \rightarrow b\bar{b}$	1320	1220	$H \rightarrow gg$	110	55	$H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$	48	13
$e^+e^- \rightarrow b\bar{b}$	$1.5 \cdot 10^8$	$1.1 \cdot 10^8$	$e^+e^- \rightarrow q\bar{q}$	61 000	2400	$e^+e^- \rightarrow \tau_{\text{had}}\tau_{\text{had}}$	$2.7 \cdot 10^7$	$3.8 \cdot 10^5$
$e^+e^- \rightarrow c\bar{c}$	$1.4 \cdot 10^6$	$9.4 \cdot 10^5$	$e^+e^- \rightarrow c\bar{c}$	220	$\sim 10$			
$e^+e^- \rightarrow q\bar{q}$	$3.0 \cdot 10^4$	4800	$e^+e^- \rightarrow b\bar{b}$	20	$\sim 1$			
$H \rightarrow WW^* \rightarrow \ell\nu 2j$	265	55	$H \rightarrow WW^* \rightarrow 2\ell 2\nu$	64	25	$H \rightarrow WW^* \rightarrow 4j$	180	27
$e^+e^- \rightarrow WW^* \rightarrow \ell\nu 2j$	$2.3 \cdot 10^5$	11 000	$e^+e^- \rightarrow WW^* \rightarrow 2\ell 2\nu$	$5.6 \cdot 10^4$	7600	$e^+e^- \rightarrow WW^* \rightarrow 4j$	$1.3 \cdot 10^5$	14 000
$e^+e^- \rightarrow b\bar{b}$	1100	–	$e^+e^- \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	1360	$\sim 5$	$e^+e^- \rightarrow ZZ^* \rightarrow 4j$	$4.7 \cdot 10^3$	20
$e^+e^- \rightarrow c\bar{c}, q\bar{q}$	150	–	$e^+e^- \rightarrow \tau\tau$	$1.2 \cdot 10^7$	–	$e^+e^- \rightarrow b\bar{b}, c\bar{c}$	$5 \cdot 10^5$	7 000
$H \rightarrow ZZ^* \rightarrow 2j 2\nu$	21	11	$H \rightarrow ZZ^* \rightarrow 2\ell 2j$	10	4	$H \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	3	0.8
$e^+e^- \rightarrow ZZ^* \rightarrow 2j 2\nu$	2700	1000	$e^+e^- \rightarrow ZZ^* \rightarrow 2\ell 2j$	1000	500	$e^+e^- \rightarrow ZZ^* \rightarrow 2\ell 2\nu$	270	70
$e^+e^- \rightarrow WW^* \rightarrow 2j 2\nu$	6100	400	$e^+e^- \rightarrow WW^* \rightarrow 2\ell 2j$	$3.3 \cdot 10^4$	$\sim 1$	$e^+e^- \rightarrow WW^* \rightarrow 2\ell 2\nu$	$3.3 \cdot 10^4$	260
$e^+e^- \rightarrow b\bar{b}, c\bar{c}, q\bar{q}$	7000	–	$e^+e^- \rightarrow b\bar{b}, c\bar{c}, q\bar{q}$	400	–	$e^+e^- \rightarrow b\bar{b}, c\bar{c}, q\bar{q}$	390	–
$e^+e^- \rightarrow \tau\tau$	1700	$\sim 2$				$e^+e^- \rightarrow \tau\tau$	$3 \cdot 10^4$	–

tagging studies exploiting jet substructure properties with machine learning techniques [42]. The latest LHC results reach  $\varepsilon_g \approx 60\%$  gluon efficiencies with  $\varepsilon_{q \rightarrow g}^{\text{mistag}} \approx 10\%$  false positive rates using advanced multivariate analyses [43, 44], or  $\varepsilon_{q \rightarrow g}^{\text{mistag}} \approx 7\%$  [45] further exploiting Lund jet plane information [46]. Reaching mistagging rates down to  $\varepsilon_{q \rightarrow g}^{\text{mistag}} \approx 1\%$ , while keeping large gluon reconstruction efficiencies, appears feasible in the clean and kinematically constrained QCD environment of future  $e^+e^-$  machines, in particular taking advantage of the very large samples of  $Z \rightarrow q\bar{q}(g)$  events at the Z pole, and the  $\mathcal{O}(10^5)$   $H \rightarrow gg$  events collected during the  $e^+e^- \rightarrow ZH$  runs, available for dedicated studies of the different colour, radiation, spin, charge, hadronization properties of quark and gluon jets [47, 48, 49]. The addition of advanced hadron identification capabilities to the FCC-ee detectors for dedicated flavour [50] (and QCD) studies, will further reduce the parton-to-hadron fragmentation uncertainties [51]. Our assigned (mis)reconstruction jet working point for this channel is  $(\varepsilon_g, \varepsilon_{q \rightarrow g}^{\text{mistag}}) = (70\%, 1\%)$ , which leads to a  $10^{-4}$  background rejection factor when requiring two gluon-tagged exclusive jets in the event. The corresponding number of events expected in  $10 \text{ ab}^{-1}$  for signal and background, after acceptance and efficiency preselections, are 110 and  $\sim 61\,000$  respectively (Table 4). The subsequent MVA is performed removing beforehand any jet variable that may have been potentially used to define the light- $q$ /gluon separation, and which is therefore de facto already accounted for by the chosen preselection (mis)tagging efficiency. An analysis of the BDT response (Fig. 2, left) indicates a maximum significance reached for a BDT output cut that further reduces the background by a factor  $\times 0.06$  while only losing 50% of the signal. The final statistical significance reached for this channel is approximately given by  $S/\sqrt{B} = 55/\sqrt{2400} = 1.1\sigma$  per FCC-ee IP per year.

**Fig. 2.** Left: Example of normalized BDT response distributions for signal and backgrounds in the  $H \rightarrow gg$  channel. Right: Examples of a few of the most discriminating (normalized) BDT variables of the  $H \rightarrow WW^* \rightarrow \ell\nu 2j$  analysis.

Z mass [41]. Although somewhat arbitrary, we have considered tiny but finite mistagging rates  $\varepsilon_{b,c \rightarrow g}^{\text{mistag}} = 0.001, 0.01\%$  to be able to quantify the impact from such sources (rescaling them, if needed) in the number of background events identified as digluons (Table 4).

### Analysis of $e^+e^- \rightarrow H(WW^*) \rightarrow \ell\nu 2j$

The event signature of the  $H(WW^*) \rightarrow \ell\nu 2j$  signal is one isolated charged lepton, missing energy from the neutrino, and two exclusive jets. In principle, such a final state can be present in multiple reducible backgrounds (Table 4), but the MVA study allows to remove basically all of them, leaving just a fraction of the  $e^+e^- \rightarrow WW^* \rightarrow \ell\nu 2j$  continuum. Table 5 lists the BDT variables used in the analysis, together with their relative weight in the final signal significance for this channel. Apart from blindly running the MVA, it is instructive to show the impact of different kinematic cuts to get rid of reducible backgrounds. Thus, for example, a significant fraction of  $e^+e^- \rightarrow q\bar{q}, c\bar{c}, b\bar{b}$  events can be eliminated by requiring e.g.:  $E_{j1,j2} < 52, 45$  GeV,  $m_{W(\ell\nu)} > 12$  GeV,  $E_\ell > 10$  GeV,  $E_{\text{miss}} > 20$  GeV. The additional requirement on the mass of the missing 4-momentum vector  $m_{\text{miss}} < 3$  GeV further discards many  $e^+e^- \rightarrow \tau\tau$  events.

**Table 5.** Indicative list of BDT variables used in the  $H \rightarrow WW^* \rightarrow \ell\nu 2j$  analysis, with their relative weight in the statistical significance for this channel.

$\cos\theta_{j1}$	$E_\ell$	$p_T(jj)$	$\cos\phi_{j1}$	$m_{\text{miss}}$	$E_{\text{visible}}$	$p_T^\ell$	$E_{\text{miss}}$	$p_T(jj\ell)$	$\cos\theta^*$
0.0446	0.0417	0.0409	0.0398	0.0341	0.0328	0.0308	0.03015	0.02726	0.02626
$\eta_{\text{miss}}$	$\eta_{j1}$	$\cos\theta_{j2}$	$\Delta\phi_{jj}$	$m_{T,\text{miss}}$	$m_{W\text{ offsh.}}$	$E_{j,\text{min}}$	$\Delta R_{\text{min},j\ell}$	$\min\Delta\eta_{j\ell}$	$p_T^{j1}$
0.0255	0.0238	0.0220	0.0215	0.0212	0.0212	0.0205	0.0204	0.0192	0.0189
$\max\cos(\ell j)$	$\eta_\ell$	$m(\ell\nu)$	$\min\cos(\ell j)$	$\max\Delta\eta_{jj}$	$m_{W\text{ shell}}$	$m_T(\ell j_1)$	$m_T(jj\ell)$	$m(\ell j_1)$	$m_{j2}$
0.0189	0.0182	0.0179	0.0176	0.0165	0.0160	0.0160	0.0160	0.0156	0.0147
$\cos\phi_{j1,j2}$	$p_T^{j2}$	$\Delta R_{\text{max},j\ell}$	$\eta_{j2}$	lin.spher.	$m_{j1}$	$p_T(\ell j_2)$	$\Delta\theta_{jj}$	$m_T(jj)$	$\Delta R_{jj}$
0.0140	0.0136	0.0136	0.0136	0.0136	0.0134	0.0134	0.0132	0.0131	0.0127
$E_{j,\text{max}}$	$m_T(\ell j_2)$	sphericity	$p_T(\ell j_1)$	$\min\Delta\phi_{j\ell}$	$E_{\text{isol}}$	aplanarity	$\max\Delta\phi_{j\ell}$	$\phi_{j1}$	$m(jj\ell)$
0.0125	0.0121	0.0116	0.0103	0.0102	0.00998	0.00927	0.00914	0.00894	0.00764
$m(\ell j_2)$	$m_{jj}$	$\phi_{j2}$	lin.aplan.	$\phi_\ell$	$\cos\phi^*$	others ( $R_{\text{min}}, \eta_\ell, \dots$ )			
0.00680	0.00641	0.00565	0.00514	0.00512	0.00471	< 0.001			

The remaining background is dominated by the  $WW^*$  continuum that can then be reduced by exploiting, among others, the different  $W^\pm$  polarizations for signal and background processes. The signal decay  $H \rightarrow W^+W^-$  is that of a scalar to a pair of distinguishable spin-1 bosons. The subsequent  $W$  bosons decays maximally violate chirality: a  $W^+$  ( $W^-$ ) boson preferentially emits a  $\ell^+$  ( $\ell^-$ ) along (against) its spin direction. The anticorrelation between the  $W^\pm$  polarizations expected in spin-zero Higgs decays is transferred into a correlation between the momenta of the charged leptons in their decays that manifests itself in the distributions of relative  $\ell^\pm$  polar angles, and a preference for a small azimuthal angle ( $\phi$ ) between the  $\ell^+\ell^-$  pair. Such angular correlations of the emitted charged leptons are encoded into the MELA variables exploited by the ATLAS and CMS collaborations to separate Higgs decays from  $W^+W^-$  backgrounds in their original searches [4, 5]. Examples of discriminating BDT variables distributions for signal and backgrounds are shown in Fig. 2 (right). Applying an appropriate cut on the BDT response output, keeps a 58% efficiency on signal, while removing 80% of the continuum background. The final statistical significance of this final state is of the order of  $S/\sqrt{B} = 55/\sqrt{11\,000} \approx 0.5\sigma$  per FCC-ee IP per year.

## 5 Beam monochromatization, expected signal significance and $y_e$ constraints

Table 6 lists the statistical significances, in units of std. deviations  $\sigma$ , for each individual  $s$ -channel Higgs decay channel studied here, for our baseline  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) = (4.1 \text{ MeV}, 10 \text{ ab}^{-1})$  monochromatization assumption. The combined final significance, and associated 95% CL upper limit, are calculated considering a multibin counting experiment with a profile likelihood for hypothesis test and confidence interval, using the ROOSTATS statistical package [52]. We have considered  $10^{-4}$  fractional systematic uncertainties<sup>3</sup> for the backgrounds, consistent with the expected experimental precision aimed at FCC-ee [24]. The final combined significance is  $1.3\sigma$ , which is also very close to the naive quadratic sum of individual  $S/\sqrt{B}$  values per channel. Such a result is equivalent to setting a 95% CL upper limit of 2.6 times the SM Higgs  $s$ -channel cross section, per FCC-ee IP and per year. Since the cross section depends on the square of the electron Yukawa,  $\sigma_{e^+e^- \rightarrow H} \propto y_e^2$ , this corresponds to placing an upper bound on the coupling at  $\sqrt{2.6} = 1.6$  times the SM value, i.e.  $|y_e| < 1.6|y_e^{\text{SM}}|$  (95% CL).

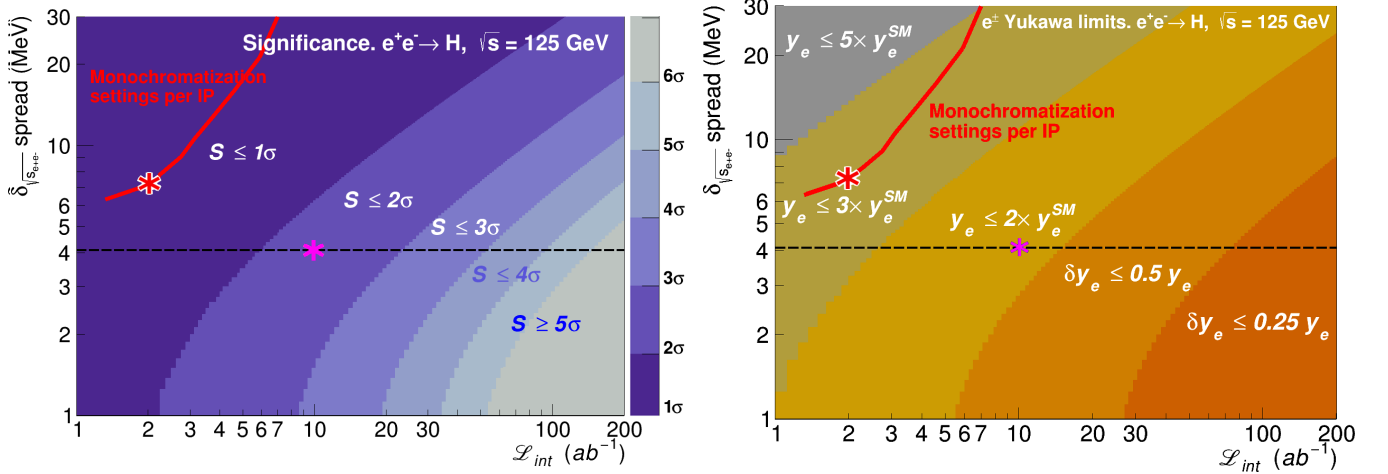
The expected final significance of the  $\sigma_{e^+e^- \rightarrow H}$  measurement, and associated 95% CL limits on  $|y_e|$ , derived for a benchmark  $\delta_{\sqrt{s}} = 4.1 \text{ MeV}$  collision-energy spread and  $\mathcal{L}_{\text{int}} = 10 \text{ ab}^{-1}$  integrated luminosities, can be easily derived

<sup>3</sup> A detailed description of the systematic studies and detector requirements needed to achieve such uncertainties for each of the Higgs final states is beyond the scope of this essay, and will be part of the outcome of the forthcoming FCC feasibility study.

**Table 6.** Individual significances (in std. deviations  $\sigma$ ) expected per decay channel for  $s$ -channel Higgs boson production in  $e^+e^-$  collisions at FCC-ee for  $\mathcal{L}_{\text{int}} = 10 \text{ ab}^{-1}$  and  $\delta_{\sqrt{s}} = 4.1 \text{ MeV}$ . The last column quotes the combined significance.

$H \rightarrow gg$	$H \rightarrow WW^* \rightarrow \ell\nu\ 2j; 2\ell\ 2\nu; 4j$	$H \rightarrow ZZ^* \rightarrow 2j\ 2\nu; 2\ell\ 2j; 2\ell\ 2\nu$	$H \rightarrow b\bar{b}$	$H \rightarrow \tau_{\text{had}}\tau_{\text{had}}; c\bar{c}; \gamma\gamma$	Combined
$1.1\sigma$	$(0.53 \otimes 0.34 \otimes 0.13)\sigma$	$(0.32 \otimes 0.18 \otimes 0.05)\sigma$	$0.13\sigma$	$< 0.02\sigma$	$1.3\sigma$

for any other combination of  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}})$  values achievable through beam monochromatization. Figure 3 shows the bidimensional maps for the significance of  $s$ -channel Higgs production (left) and the corresponding 95% CL upper limits on the electron Yukawa (right), as a function of both parameters. The signal significance, and associated upper limits, improve with the square-root of the integrated luminosity (along the  $x$  axes of both plots), and diminish for larger values  $\delta_{\sqrt{s}}$  (along the  $y$  axes of the maps) following the relativistic Voigtian dependence of the signal yield on the energy spread shown in Fig. 1 (right).



**Fig. 3.** Left: Significance contours (in std. dev. units  $\sigma$ ) in the c.m. energy spread vs. integrated luminosity plane for the resonant  $\sigma_{e^+e^- \rightarrow H}$  cross section at  $\sqrt{s} = m_H$ . Right: Associated upper limits contours (95% CL) on the electron Yukawa  $y_e$ . The red curves show the range of parameters presently reached in FCC-ee monochromatization studies [20,21]. The red star indicates the best signal strength monochromatization point in the plane (the pink star over the  $\delta_{\sqrt{s}} = \Gamma_H = 4.1 \text{ MeV}$  dashed line, indicates the ideal baseline point assumed in our default analysis). All results are given per IP and per year.

The red curves in Fig. 3 show the current expectations for the range of  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}})$  values achievable at FCC-ee with the investigated monochromatization schemes [20,21]. Without monochromatization, the FCC-ee natural collision-energy spread at  $\sqrt{s} = 125 \text{ GeV}$  is about  $\delta_{\sqrt{s}} = 46 \text{ MeV}$  due to synchrotron radiation. Its reduction to the few-MeV level desired for the  $s$ -channel Higgs run can be accomplished by means of monochromatization, e.g. by introducing nonzero horizontal dispersions at the IP ( $D_x^*$ ) of *opposite sign* for the two beams in collisions without a crossing angle. The beam energy spread reduction factor is given by  $\lambda = \sqrt{(D_x^{*2} \sigma_\delta^2)/(\varepsilon_x \beta_x^*) + 1}$ , where  $\beta_{x(y)}^*$  denotes the horizontal (vertical) beta function at the IP and  $\varepsilon_{x(y)}$  the corresponding emittance. The need to generate a significant IP dispersion implies a change of beamline geometry in the interaction region and the use of crab cavities to compensate for the existing, or remaining, crossing angle. A nonzero IP dispersion leads to an increase of the transverse horizontal emittance from beamstrahlung, thereby impacting the beam luminosity. Optimization of the IP optics parameters ( $D_x^*$ ,  $\beta_{x,y}^*$ , ...) yields the corresponding red curves of Fig. 3. For the lowest collision-energy spread achieved of  $\delta_{\sqrt{s}} = 6 \text{ MeV}$ , the anticipated monochromatized luminosity per IP exceeds  $10^{35} \text{ cm}^{-2} \text{ s}^{-1}$  [21]. This translates into an integrated luminosity<sup>4</sup> of at least  $1.2 \text{ ab}^{-1}$  per IP per year. One can reach larger integrated luminosities at the expense of a worse beam energy spread. The point (red star) over the red curves that has the highest signal strength today corresponds to  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) \approx (7 \text{ MeV}, 2 \text{ ab}^{-1})$ , to be compared to our original baseline point (pink star) over the  $\delta_{\sqrt{s}} = \Gamma_H = 4.1 \text{ MeV}$  dashed line. For such a 7-MeV c.m. energy spread, the peak of the relativistic Voigtian distribution describing the  $s$ -channel cross section is located at about 1 MeV above the mass of the Higgs boson (Fig. 1, right). Therefore, the optimal c.m. energy of the dedicated  $e^+e^-$  run needs also to be carefully chosen to maximize the resonant cross section for any given monochromatization point.

<sup>4</sup> Conversion from luminosity ( $\mathcal{L} = 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ ) to integrated luminosity ( $\mathcal{L}_{\text{int}} = 1.2 \text{ ab}^{-1}/\text{year/IP}$ ) assumes 185 physics days per run with a 75% physics efficiency [27].



Compared to our baseline values (pink stars on the plots), the signal significance for the currently best monochromatization settings,  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) \approx (7 \text{ MeV}, 2 \text{ ab}^{-1})$ , drops to  $\mathcal{S} \approx 0.4\sigma/\text{year/IP}$ , and the corresponding upper bound on the  $e^\pm$  Yukawa becomes  $y_e \lesssim 2.5y_e^{\text{SM}}$  (95% CL) per year and per IP. Assuming 2 years of FCC-ee operation at the Higgs pole and combining four detectors/IPs, this would translate into a  $1.2\sigma$  significance and a  $y_e \lesssim 1.6y_e^{\text{SM}}$  limit. Such a result, although clearly short of an evidence for  $s$ -channel Higgs production, is still about 100 (30) times better [54] than that reachable at HL-LHC (FCC-hh [53]), and would imply setting a constraint on new physics affecting the electron-Higgs coupling above  $\Lambda_{\text{BSM}} \gtrsim 110 \text{ TeV}$ .

Given that any improved analysis of the Higgs decay channels is unlikely to increase much more the final signal significance, alternative paths need to be considered in order to measure more precisely the electron Yukawa coupling at FCC-ee. The possibility of introducing beam longitudinal polarizations ( $P_L$ ) would enhance the signal by  $(1 + P_L^2)$  and suppress backgrounds by  $(1 - P_L^2)$ , i.e. running with  $P_L = 68\%$  (90%) would increase by a factor of two (four) the statistical significance of the signal. However, for realistic longitudinal polarizations reachable at FCC-ee ( $P_L = 20\text{--}30\%$ ) the gain would be insufficient and higher polarizations would significantly reduce the luminosity. The only approach seemingly left to carry out an  $e^+e^- \rightarrow H$  measurement with a sensitivity reaching the SM electron Yukawa level requires improving the beam monochromatization beyond the current state-of-the-art [20, 21]. Alternative or modified monochromatization scenarios [55, 56, 57] are being explored that however, for now, do not improve the results of the red curves shown in Fig. 3.

## 6 Summary and outlook

The prospects for a potential FCC-ee measurement of the direct  $s$ -channel Higgs boson production in  $e^+e^-$  collisions at  $\sqrt{s} = m_H$  have been studied as a means to determine the Higgs Yukawa coupling of the electron ( $y_e$ ). The three main challenges of such a measurement have been discussed: (i) the need to accurately know (within MeV’s) beforehand the value of the Higgs boson mass where to operate the collider, (ii) the smallness of the resonant Higgs boson cross section (few hundred ab) due to ISR and beam energy spread ( $\delta_{\sqrt{s}}$ ) that requires to monochromatize the beams, i.e. reduce  $\delta_{\sqrt{s}}$  at the few MeV scale, while still delivering large (few  $\text{ab}^{-1}$ ) integrated luminosities  $\mathcal{L}_{\text{int}}$ , and (iii) the existence of multiple backgrounds with orders-of-magnitude larger cross sections than the Higgs signal decay channels themselves. The knowledge of  $m_H$  with a few MeV accuracy seems feasible at FCC-ee as per dedicated studies reported in Ref. [23]. This present work has focused on the points (ii) and (iii) above, by performing a generator-level study that has chosen as benchmark point a baseline monochromatization scheme leading to  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) = (4.1 \text{ MeV}, 10 \text{ ab}^{-1})$ , corresponding to a peak  $s$ -channel cross section of  $\sigma_{e^+e^- \rightarrow H} = 280 \text{ ab}$ .

Large simulated event samples of signal and associated backgrounds have been generated with the PYTHIA 8 Monte Carlo (MC) code for 11 Higgs boson decay channels. A simplified description of the expected experimental performances has been assumed for the reconstruction and (mis)tagging of heavy-quark ( $c$ ,  $b$ ) and light-quark and gluons ( $udsg$ ) jets, photons, electrons, and hadronically decaying tau leptons. Generic preselection criteria have been defined that target the 11 Higgs boson channels, suppressing reducible backgrounds while keeping the largest fraction of the signal events. A subsequent multivariate analysis of  $\mathcal{O}(50)$  kinematic and global topological variables, defined for each event, has been carried out. Boosted-Decision-Trees (BDT) classifiers have been trained on signal and background events, to maximize the signal significances for each individual channel. The most significant Higgs decay channels are found to be  $H \rightarrow gg$  (for a gluon efficiency of 70% and a  $uds$ -for- $g$  jet mistagging rate of 1%), and  $H \rightarrow WW^* \rightarrow \ell\nu 2j$ . Combining all results, a  $1.3\sigma$  signal significance can be achieved, corresponding to an upper limit on the  $e^\pm$  Yukawa coupling at 1.6 times the SM value:  $|y_e| < 1.6|y_e^{\text{SM}}|$  at 95% confidence level (CL), per FCC-ee interaction point (IP) and per year. Such a bound is about  $\times 100$  ( $\times 30$ ) times better than that reachable at HL-LHC (FCC-hh), and can be translated into a lower limit on the energy scale of any physics beyond the SM (BSM) affecting the electron Yukawa coupling, of  $\Lambda_{\text{BSM}} \approx v^{3/2}(\sqrt{2}m_e \cdot (y_e/y_e^{\text{SM}}))^{-1/2} \gtrsim 110 \text{ TeV}$ .

Details on the status of ongoing FCC-ee monochromatization studies have been provided. The current monochromatization settings with largest Higgs signal strength correspond to  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) \approx (7 \text{ MeV}, 2 \text{ ab}^{-1})$ , and translate into a  $0.4\sigma$  significance on the Higgs boson cross section, or correspondingly a  $|y_e| < 2.6|y_e^{\text{SM}}|$  (95% CL) upper bound, per IP and per year. Forthcoming extension and consolidation of this work, in the context of the anticipated FCC feasibility study, require at least the following activities:

- (i) Confirming the current signal significances with alternative MC event generators for the Higgs diboson backgrounds, in particular for the promising  $H \rightarrow WW^* \rightarrow \ell\nu 2j$  channel.
- (ii) Studying the improvements of the FCC-ee detectors design needed in order to achieve the required accuracy and precision in key aspects of the analysis, such as the small light-quark-for-gluon mistagging efficiency of 1% assumed in the key  $H \rightarrow gg$  channel.
- (iii) Redoing the analysis using a more realistic (parametrized or full simulation) description of the detector response to more accurately assess the impact on the final signal significances of the reconstruction and selection efficiencies expected at FCC-ee.

- (iv) Continuing and extending the accelerator monochromatization studies to improve the currently best FCC-ee working point of  $(\delta_{\sqrt{s}}, \mathcal{L}_{\text{int}}) \approx (7 \text{ MeV}, 2 \text{ ab}^{-1})$ , aiming at further reducing  $\delta_{\sqrt{s}}$  while increasing  $\mathcal{L}$ , and developing the corresponding optical lattices for the required beam optics parameters at the IP.

It is worth noting that running FCC-ee at  $\sqrt{s} = m_H$  for a couple (or more) years can provide many more scientific outputs than the direct  $s$ -channel measurement considered here. Indeed, integrating tens of  $\text{ab}^{-1}$  in  $e^+e^-$  collisions at 125 GeV provides useful means to accurately determine the number of light neutrino families via  $Z(\nu\nu)\gamma$  radiative return [58], search for weakly-coupled BSM physics between the  $Z$  and Higgs mass poles [59], and carry out other luminosity-demanding SM studies not accessible at the  $Z$  pole.

In summary, the results presented in this essay demonstrate that FCC-ee is the most well-suited (if not, arguably, the unique) collider that can aim at a measurement of the electron Yukawa coupling via direct  $s$ -channel Higgs boson production. Such a measurement has many fundamental physics motivations and implications, among which: (i) it will explore the (so far hypothetical) Higgs mass generation mechanism for elementary particles of the first family of fermions that form the stable matter of the visible universe, (ii) it will scrutinize the electron's Yukawa coupling that, through its impact on the electron mass, sets the size of atoms and their energy levels (the Bohr radius is proportional to  $1/m_e$ ), (iii) it can access BSM scalar physics connected to the electron above the  $\sim 100 \text{ TeV}$  scale, and (iv) it can directly probe the potential presence of any new particle that is quasi-degenerate (at the MeV level) with the Higgs boson mass.

**Acknowledgements.** We thank Roy Aleksan, Alain Blondel, Patrick Janot, and Frank Zimmermann for valuable discussions and feedback on various aspects of this study.

## References

1. F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.
2. P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132–133.
3. P. W. Higgs, *Broken symmetry and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
4. ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1–29, [arXiv:1207.7214 \[hep-ex\]](#).
5. CMS, S. Chatrchyan *et al.*, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30–61, [arXiv:1207.7235 \[hep-ex\]](#).
6. H. Gray and P. Janot, *Higgs Physics*, *Comptes Rendus Physique* **21** no. 1, (2020) 23–43.
7. S. D. Bass, A. De Roeck, and M. Kado, *The Higgs boson - its implications and prospects for future discoveries*, [arXiv:2104.06821 \[hep-ph\]](#).
8. F. F. Deppisch, P. S. Bhupal Dev, and A. Pilaftsis, *Neutrinos and collider physics*, *New J. Phys.* **17** no. 7, (2015) 075019, [arXiv:1502.06541 \[hep-ph\]](#).
9. CMS, V. Khachatryan *et al.*, *Search for a standard model-like Higgs boson in the  $\mu^+\mu^-$  and  $e^+e^-$  decay channels at the LHC*, *Phys. Lett. B* **744** (2015) 184–207, [arXiv:1410.6679 \[hep-ex\]](#).
10. ATLAS, G. Aad *et al.*, *Search for the Higgs boson decays  $H \rightarrow ee$  and  $H \rightarrow e\mu$  in  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector*, *Phys. Lett. B* **801** (2020) 135148, [arXiv:1909.10235 \[hep-ex\]](#).
11. W. Altmannshofer, J. Brod, and M. Schmaltz, *Experimental constraints on the coupling of the Higgs boson to electrons*, *JHEP* **05** (2015) 125, [arXiv:1503.04830 \[hep-ph\]](#).
12. M. Cepeda *et al.*, *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 221–584, [arXiv:1902.00134 \[hep-ph\]](#).
13. V. D. Barger, M. S. Berger, J. F. Gunion, and T. Han, *Higgs boson physics in the  $s$  channel at  $\mu^+\mu^-$  colliders*, *Phys. Rept.* **286** (1997) 1–51, [arXiv:hep-ph/9602415](#).
14. TLEP Design Study Working Group, M. Bicer *et al.*, *First Look at the Physics Case of TLEP*, *JHEP* **01** (2014) 164, [arXiv:1308.6176 \[hep-ex\]](#).
15. D. d'Enterria, *Search for resonant  $s$ -channel Higgs production at the FCC-ee*. Fcc-ee/tlep physics workshop (tlep7), cern, <https://indico.cern.ch/event/313708/contributions/1687826/>, 19–21 June 2014. FCC-ee/TLEP physics workshop (TLEP7) 19–21 June 2014, CERN.
16. D. d'Enterria, *Higgs physics at the Future Circular Collider*, *PoS ICHEP2016* (2017) 434, [arXiv:1701.02663 \[hep-ex\]](#).
17. S. Jadach and R. A. Kycia, *Lineshape of the Higgs boson in future lepton colliders*, *Phys. Lett. B* **755** (2016) 58–63, [arXiv:1509.02406 \[hep-ph\]](#).
18. M. Greco, T. Han, and Z. Liu, *ISR effects for resonant Higgs production at future lepton colliders*, *Phys. Lett. B* **763** (2016) 409–415, [arXiv:1607.03210 \[hep-ph\]](#).
19. A. Dery, C. Frugiuele, and Y. Nir, *Large Higgs-electron Yukawa coupling in 2HDM*, *JHEP* **04** (2018) 044, [arXiv:1712.04514 \[hep-ph\]](#).
20. F. Zimmermann and M. Valdivia García, *Optimized monochromatization for direct Higgs production in future circular  $e^+e^-$  colliders*, in *8th Intl. Particle Accel. Conf.*, [10.18429/JACoW-IPAC2017-WEPIK015](#). 2017.

21. M. Valdivia García and F. Zimmermann, *Effect of emittance constraints on monochromatization at the Future Circular  $e^+e^-$  Collider*, in *10th Intl. Particle Accel. Conf.*, [10.18429/JACoW-IPAC2019-MOPMP035](#), 6, 2019; and private comm.
22. D. de Florian *et al.*, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922 \[hep-ph\]](#).
23. P. Azzurri, G. Bernardi, S. Braibant, D. d'Enterria, J. Eysermans, P. Janot, A. Li, and E. Perez, *A special Higgs challenge: Measuring the mass and production cross section with ultimate precision at FCC-ee*, [arXiv:2106.15438 \[hep-ex\]](#).
24. A. Blondel *et al.*, *Polarization and centre-of-mass energy calibration at FCC-ee*, [arXiv:1909.12245 \[physics.acc-ph\]](#).
25. R. A. Kycia and S. Jadach, *Relativistic Voigt profile for unstable particles in high energy physics*, *J. Math. Anal. Appl.* **463** no. 2, (2018) 1040–1051, [arXiv:1711.09304 \[math-ph\]](#).
26. FCC, A. Abada *et al.*, *FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1*, *Eur. Phys. J. C* **79** no. 6, (2019) 474.
27. FCC, A. Abada *et al.*, *FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2*, *Eur. Phys. J. ST* **228** no. 2, (2019) 261–623.
28. A. Blondel, J. Gluza, S. Jadach, P. Janot, and T. Riemann, eds., *Theory for the FCC-ee: Report on the 11th FCC-ee Workshop Theory and Experiments*, vol. 3/2020 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 5, 2019. [arXiv:1905.05078 \[hep-ph\]](#).
29. S. Jadach, W. Placzek, M. Skrzypek, B. F. L. Ward, and S. A. Yost, *The path to 0.01% theoretical luminosity precision for the FCC-ee*, *Phys. Lett. B* **790** (2019) 314–321, [arXiv:1812.01004 \[hep-ph\]](#).
30. T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](#).
31. A. Djouadi, J. Kalinowski, M. Muehlleitner, and M. Spira, *HDECAY: Twenty++ years after*, *Comput. Phys. Commun.* **238** (2019) 214–231, [arXiv:1801.09506 \[hep-ph\]](#).
32. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#).
33. V. Hirschi and O. Mattelaer, *Automated event generation for loop-induced processes*, *JHEP* **10** (2015) 146, [arXiv:1507.00020 \[hep-ph\]](#).
34. W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating multi-particle processes at LHC and ILC*, *Eur. Phys. J. C* **71** (2011) 1742, [arXiv:0708.4233 \[hep-ph\]](#).
35. A. Arbuzov, S. Jadach, Z. Was, B. F. L. Ward, and S. A. Yost, *The Monte Carlo Program KKMC, for lepton or quark pair production at LEP/SLC energies—Updates of electroweak calculations*, *Comput. Phys. Commun.* **260** (2021) 107734, [arXiv:2007.07964 \[hep-ph\]](#).
36. M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097 \[hep-ph\]](#).
37. S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant  $k_T$  clustering algorithms for hadron hadron collisions*, *Nucl. Phys. B* **406** (1993) 187–224.
38. S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160–3166, [arXiv:hep-ph/9305266](#).
39. CMS, S. Chatrchyan *et al.*, *Study of the mass and spin-parity of the Higgs boson candidate via its Decays to Z boson pairs*, *Phys. Rev. Lett.* **110** no. 8, (2013) 081803, [arXiv:1212.6639 \[hep-ex\]](#).
40. A. Hocker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, [arXiv:physics/0703039](#).
41. Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, *PTEP* **2020** no. 8, (2020) 083C01.
42. A. J. Larkoski, I. Moult, and B. Nachman, *Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning*, *Phys. Rept.* **841** (2020) 1–63, [arXiv:1709.04464 \[hep-ph\]](#).
43. G. Kasieczka, N. Kiefer, T. Plehn, and J. M. Thompson, *Quark-gluon tagging: Machine learning vs detector*, *SciPost Phys.* **6** no. 6, (2019) 069, [arXiv:1812.09223 \[hep-ph\]](#).
44. J. Filipek, S.-C. Hsu, J. Kruper, K. Mohan, and B. Nachman, *Identifying the quantum properties of hadronic resonances using machine learning*, [arXiv:2105.04582 \[hep-ph\]](#).
45. C. K. Khosa and S. Marzani, *Higgs tagging with the Lund jet plane*, [arXiv:2105.03989 \[hep-ph\]](#).
46. F. A. Dreyer, G. P. Salam, and G. Soyez, *The Lund Jet Plane*, *JHEP* **12** (2018) 064, [arXiv:1807.04758 \[hep-ph\]](#).
47. P. Skands and D. d'Enterria, *QCD and  $\gamma\gamma$  studies at FCC-ee*, *PoS ICHEP2016* (2017) 1156, [arXiv:1610.06254 \[hep-ph\]](#).
48. D. d'Enterria and P. Skands, eds., *Proceedings, Parton Radiation and Fragmentation from LHC to FCC-ee: CERN, Geneva, Switzerland, Nov. 22-23, 2016*, 1, 2017. [arXiv:1702.01329 \[hep-ph\]](#).
49. D. d'Enterria, *Precision Quantum Chromodynamics*, *CERN Yellow Reports: Monographs* **3** (2020) 38–50.
50. G. Wilkinson, *Particle identification at FCC-ee*, [arXiv:2106.01253 \[physics.ins-det\]](#).
51. D. d'Enterria, *Ultimate light-quark vs gluon jet discrimination at FCC-ee*. In preparation, (2021).
52. L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, *PoS ACAT2010* (2010) 057, [arXiv:1009.1003 \[physics.data-an\]](#).
53. FCC, A. Abada *et al.*, *FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3*, *Eur. Phys. J. ST* **228** no. 4, (2019) 755–1107.
54. N. Alipour Tehrani *et al.*, *FCC-ee: Your Questions Answered*, in *CERN Council Open Symposium on the Update of European Strategy for Particle Physics*, A. Blondel and P. Janot, eds., 6, 2019, [arXiv:1906.02693 \[hep-ph\]](#).

- 55. A. A. Zholents, *Sophisticated accelerator techniques for colliding beam experiments*, *Nucl. Instrum. Meth. A* **265** (1988) 179–185.
- 56. A. Bogomyagkov and E. Levichev, *Collision monochromatization in  $e^+e^-$  colliders*, *Phys. Rev. Accel. Beams* **20** no. 5, (2017) 051001, [arXiv:1702.03634 \[physics.acc-ph\]](#), [Erratum: Phys.Rev.Accel.Beams 21, 029902 (2018)].
- 57. V. I. Telnov, *Monochromatization of  $e^+e^-$  colliders with a large crossing angle*, [arXiv:2008.13668 \[physics.acc-ph\]](#).
- 58. K. J. F. Gaemers, R. Gastmans, and F. M. Renard, *Neutrino Counting in  $e^+e^-$  collisions*, *Phys. Rev. D* **19** (1979) 1605.
- 59. P. Agrawal *et al.*, *Feebly-Interacting Particles:FIPs 2020 Workshop Report*, [arXiv:2102.12143 \[hep-ph\]](#).