# Cross section limits for $e^+e^- \to Z\,S$ in $S \to \tau\tau$ channel at 550 GeV ILC


Weronika Sobień, **Aleksander Filip Żarnecki**

Faculty of Physics, University of Warsaw


**ILD Software and Analysis meeting**
**August 26, 2026**

# Introduction

Light scalar searches considered in the ECFA study, selected as focus topic arXiv:2401.07564

## Previous study

- Delphes level analysis targeting 95 GeV scalar in $S \rightarrow \tau\tau$ decay channel
  $\Rightarrow$ results presented at LCWS'2023
- more detailed Delphes study for 20 – 120 GeV mass, combining different event categories
  $\Rightarrow$ finalized for LCWS'2024 and ICHEP'2024 by Kamil Zembaczyński (M.Sc. thesis)
  also contributed to ECFA study report and LC Vision submission

# Introduction

Light scalar searches considered in the ECFA study, selected as focus topic arXiv:2401.07564

## Previous study

- Delphes level analysis targeting 95 GeV scalar in $S \to \tau\tau$ decay channel
  $\Rightarrow$ results presented at LCWS'2023
- more detailed Delphes study for 20 – 120 GeV mass, combining different event categories
  $\Rightarrow$ finalized for LCWS'2024 and ICHEP'2024 by Kamil Zembaczyński (M.Sc. thesis)
  also contributed to ECFA study report and LC Vision submission

## New study

- dedicated signal samples with $S \to \tau\tau$ generated for 250 GeV and 550 GeV
  too few events in this decay channel for existing MC samples (assuming SM-like decays)
- reconstruction procedure developed by Weronika Sobień for 550 GeV S&A April 15
- cross section limits extracted for 250 GeV ILC S&A July 15
- today: new results from analysis extended to 550 GeV

## Monte Carlo samples

Standard MC-2020 SM background samples for 250 GeV:

- Generated with Whizard v2.8.5 and ILC 250 “SetA” beam parameters.
- Detector response simulation (mostly ILD_l5_o1_v02) with DDSim.
- Standard event reconstruction chain based on Marlin.

Recent large SM background samples for 550 GeV: thanks to Bryan Bliewert

- Generated with Whizard v3.1.4/5 and ILC 250 “SetA” beam parameters.
- Detector response simulation with SGV.
- Standard event reconstruction chain based on Marlin.

New signal MC samples:

- Generated with Whizard v3.1.5 (thanks to Junping Tian)
- SGV fast detector simulation and reconstruction (thanks to Mikael Berggren)

All events generated in pure helicity states (100% beam polarisation) and properly mixed.

## Reconstruction and pre-selection

same procedure for 250 GeV / 550 GeV

Dedicated high-level event reconstruction in Marlin, based on standard ILCsoft processors.

Reconstruction based on 'PandoraPFOs' collection:

- Isolated photons, leptons and tau jets were identified with corresponding processors ( IsolatedPhotonTagging, IsolatedLeptonTagging, TaJetClustering )
- all PFOs, except for isolated photons, clustered into 4 jets (Durham alg., exclusive mode)
- isolated leptons and tau jet candidates linked to Durham jets

Events with multiple isolated photons, hard colinear photons or missing $p_Z$ removed (Z return)

## Reconstruction and pre-selection

same procedure for 250 GeV / 550 GeV

Dedicated high-level event reconstruction in Marlin, based on standard ILCsoft processors.

Reconstruction based on 'PandoraPFOs' collection:

- Isolated photons, leptons and tau jets were identified with corresponding processors ( IsolatedPhotonTagging, IsolatedLeptonTagging, TaJetClustering )
- all PFOs, except for isolated photons, clustered into 4 jets (Durham alg., exclusive mode)
- isolated leptons and tau jet candidates linked to Durham jets

Events with multiple isolated photons, hard colinear photons or missing $p_Z$ removed (Z return)

Z decay channel is identified first:

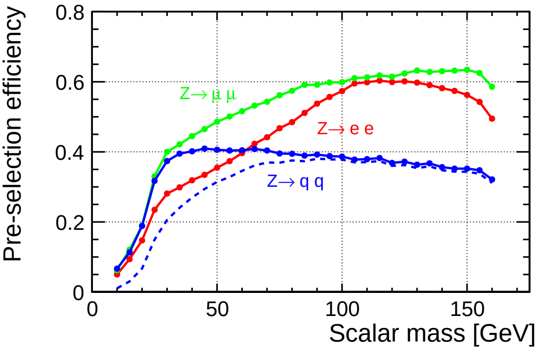
- events with two massive jets ($M_{j_1} > 5/8$ GeV, $M_{j_2} > 3$ GeV) considered hadronic Z decays
- otherwise pair of leptons matching $M_Z$ (60–140 GeV) required for leptonic Z decay

Scalar decay: two tau candidates (isolated leptons or tau jets) required – "tight selection" (!)

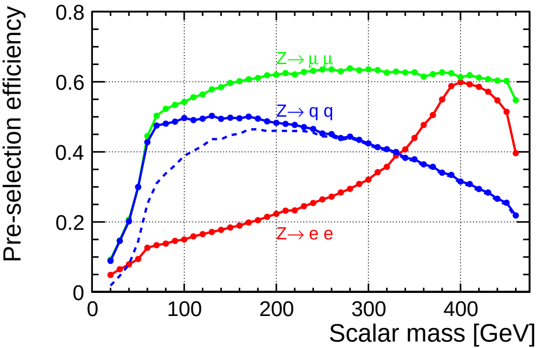
## Signal pre-selection efficiency

All tau decay channels included: (dashed: without merged tau jets)

250 GeV




550 GeV




Pre-selection efficiency for $Z \to e^+e^-$ biased by the contribution from ZZ fusion channel

## Background pre-selection

Expected SM background contributions for the combined runs:

| Background category | Expected events | |
|---|---|---|
| | 2 $ab^{-1}$ @ 250 GeV | 4 $ab^{-1}$ @ 550 GeV |
| 2f | 589'000 | 2'630 |
| 4f | 736'000 | 771'000 |
| 6f | 80 | 2'180 |
| SM Higgs | 25'800 | 36'500 |
| Total BG | 1'350'000 | 812'000 |

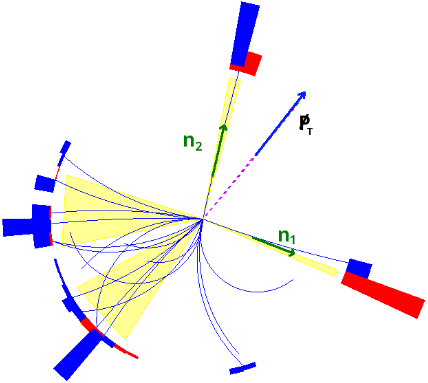
| | | |
|---|---|---|
| 50 GeV/250 GeV signal* | 4'200 / – | 770 / 480 |

* for 1% of SM cross section at this mass

## Scalar mass correction

arXiv:1509.01885

Example signal event with hadronic tau decays




Tau leptons are very boosted $\Rightarrow$ collinear approximation

Assume tau neutrinos are emitted in the tau jet direction.

Their energies can be found from transverse momentum balance:

$$\vec{\not{p}}_T = E_{\nu_1} \cdot \vec{n_1} + E_{\nu_2} \cdot \vec{n_2}$$

where $\vec{n_1}$ and $\vec{n_2}$ are directions of the two tau jets.

Unique solution !

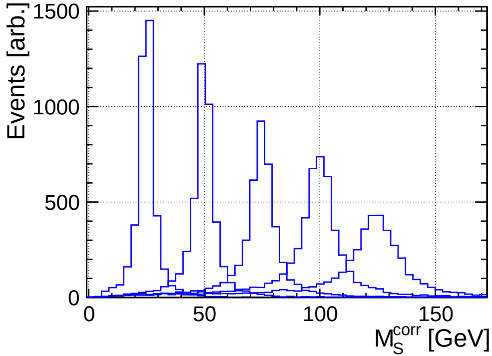
Works also for semi-leptonic and leptonic events!

Because of small tau mass $\Rightarrow$ small invariant mass of neutrino pair

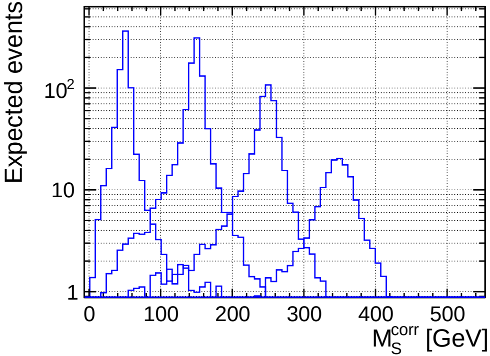
## Mass correction

Example distributions of corrected mass:

Scalar masses of 25 to 125 GeV at 250 GeV




Scalar masses of 50 to 350 GeV at 550 GeV




All tau decay channels combined

## BDT input variables

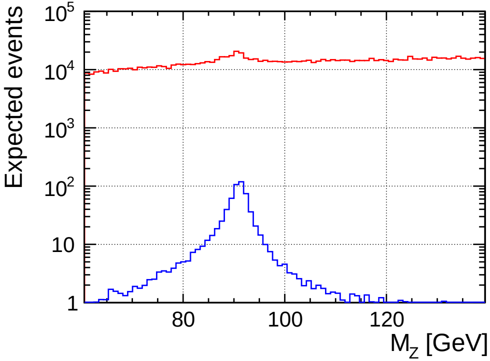
Event classification using BDT (from root TMVA) with following input variables:

- reconstructed mass of Z candidate (di-lepton or di-jet) and recoil mass
- measured invariant mass of S candidate (di-tau) and corrected S mass
- polar angle of Z candidate, asimuthal angle distance between tau candidates
- scalar decay angles in its CMS
- minimum and maximum distances for PFO clustering into four jets
- total reconstructed energy and energy of ISR photon
- flags indicating decay channels (for Z and two taus) and event class

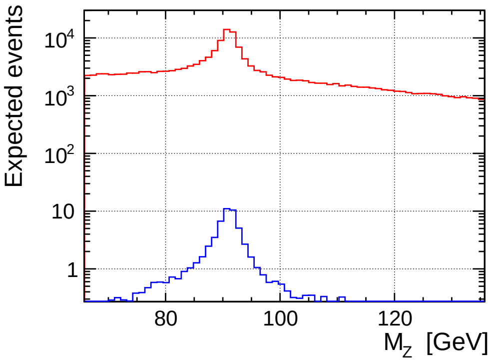
## BDT input variables

**Signal** and **SM background** distributions: Z candidate mass in leptonic Z decays.

250 GeV:




550 GeV:

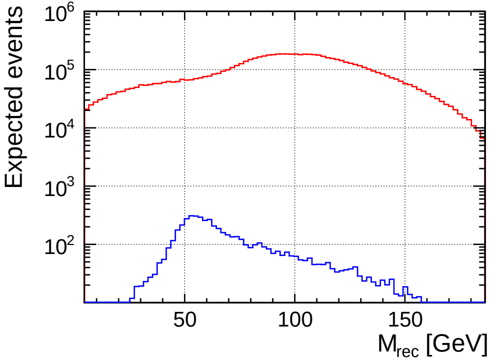



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

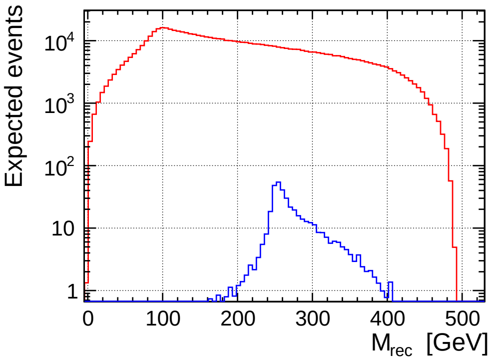
## BDT input variables

**Signal** and **SM background** distributions: recoil mass for hadronic Z decays.

250 GeV:




550 GeV:

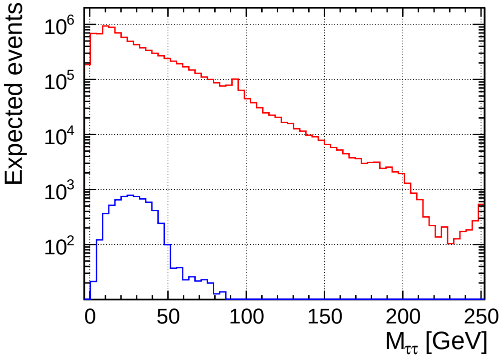



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

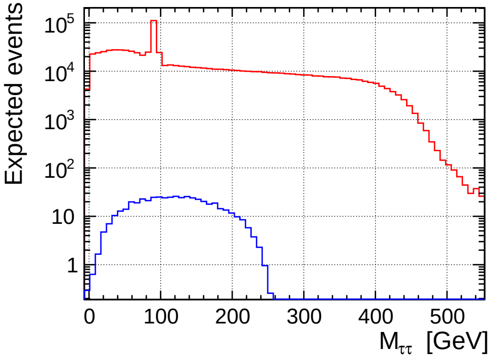
## BDT input variables

**Signal** and **SM background** distributions: raw scalar mass, all Z decays.

250 GeV:




550 GeV:

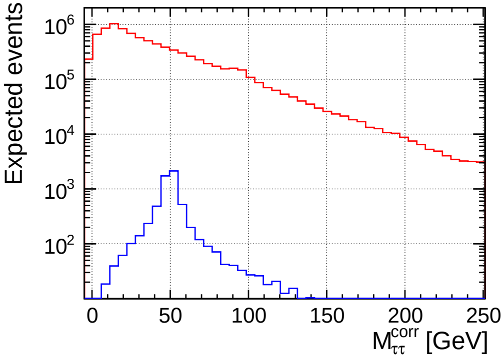



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

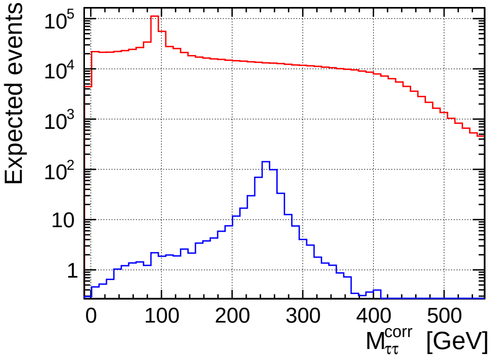
## BDT input variables

**Signal** and **SM background** distributions: corrected scalar mass, all Z decays.

250 GeV:




550 GeV:

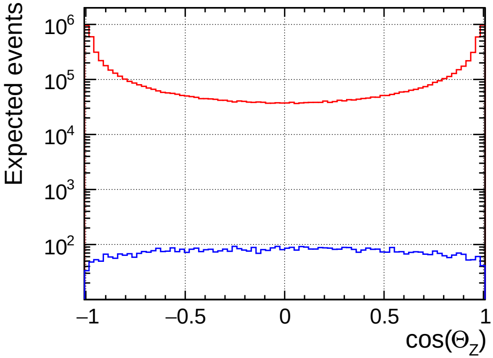



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

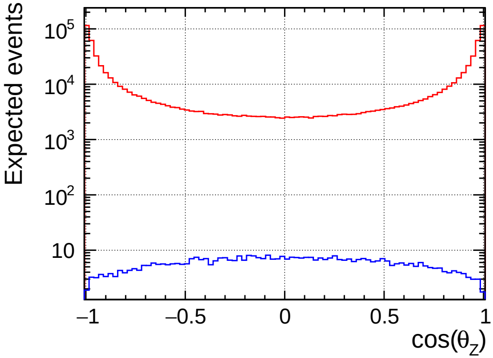
## BDT input variables

**Signal** and **SM background** distributions: Z candidate polar angle

250 GeV:




550 GeV:

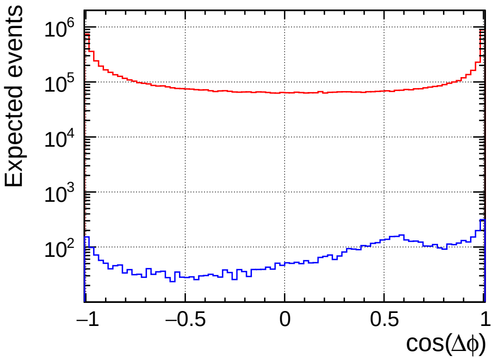



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

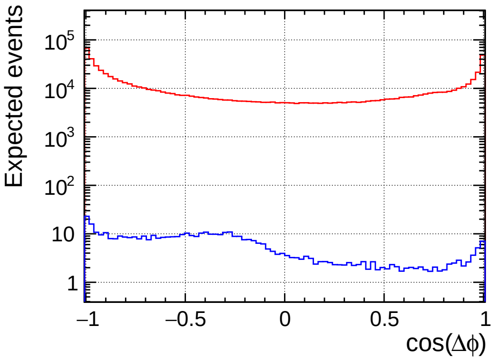
## BDT input variables

**Signal** and **SM background** distributions: azimuthal angle between tau candidates

250 GeV:




550 GeV:

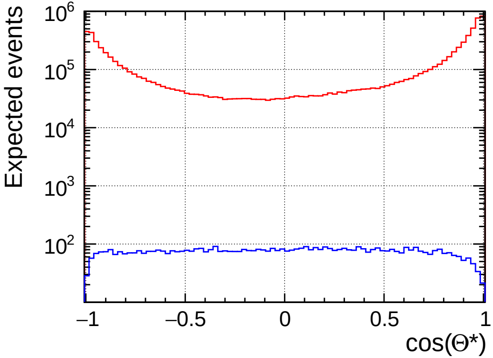



Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

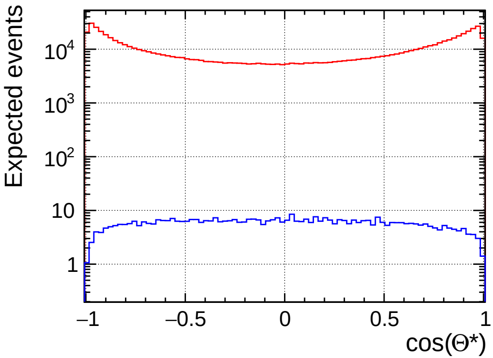
## BDT input variables

**Signal** and **SM background** distributions: scalar decay angle (in CMS)

250 GeV:




550 GeV:

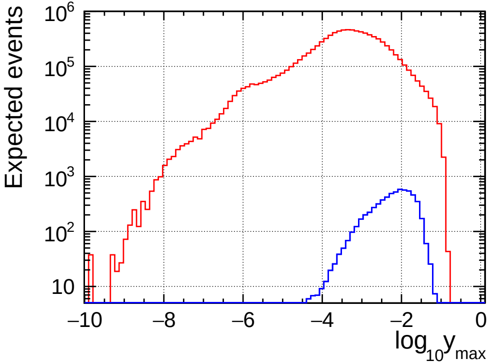


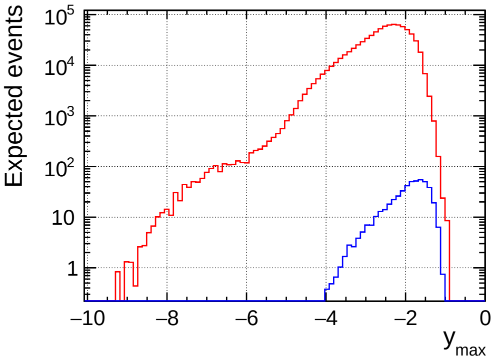

Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

## BDT input variables

**Signal** and **SM background** distributions: clustering parameter

250 GeV: 550 GeV:



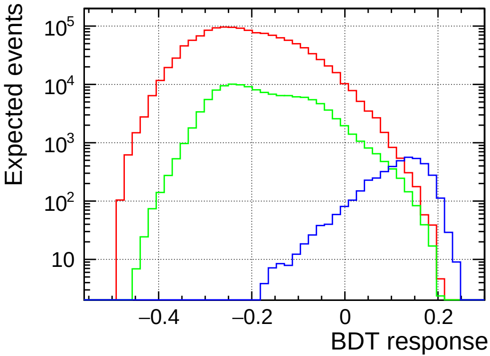




Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

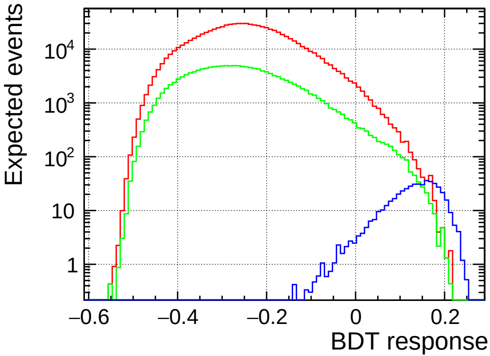
## BDT response distribution

Example of BST response distribution: total SM bg., semi-leptonic ZZ, signal.

50 GeV scalar at 250 GeV




250 GeV scalar at 550 GeV

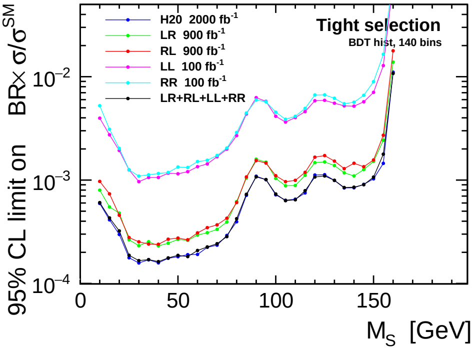



Signal of normalized to 1% of SM cross section for given mass

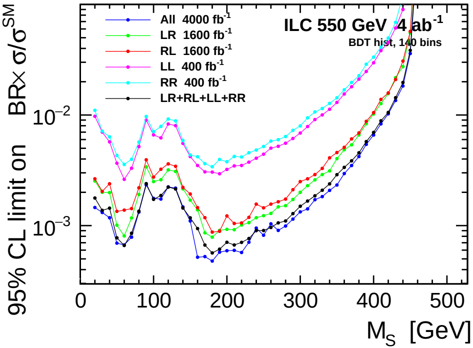
## Beam polarisation

Results presented so far were obtained when combining data with different beam polarisation

250 GeV running




550 GeV running

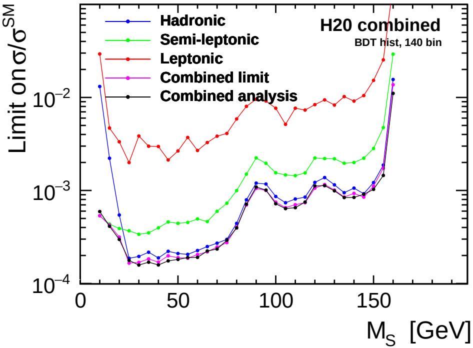



No improvement from combining limits calculated separately for each polarization!

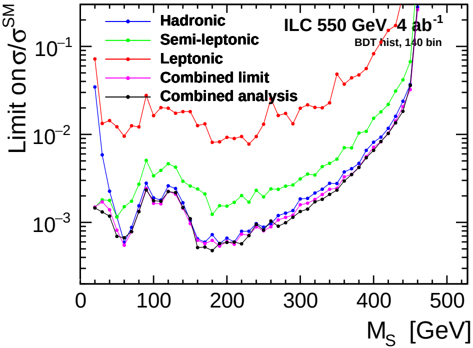
## Tau decays

Results presented so far were obtained from all tau decay channels combined.

250 GeV running




550 GeV running



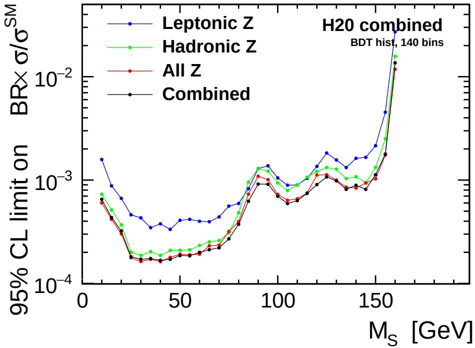

No improvement from combining limits calculated separately for each event type!

# Comparison of limits

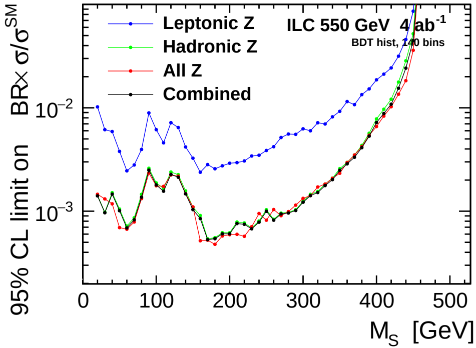
## Z decays

Results presented so far were obtained from all Z decay channels combined.

250 GeV running




550 GeV running

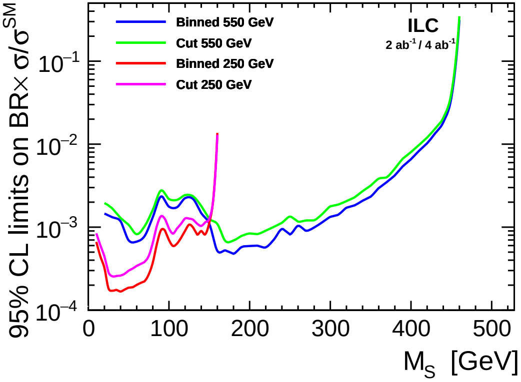



Small improvement from combining limits from hadronic and leptonic Z decay only at 250 GeV.

## Final result

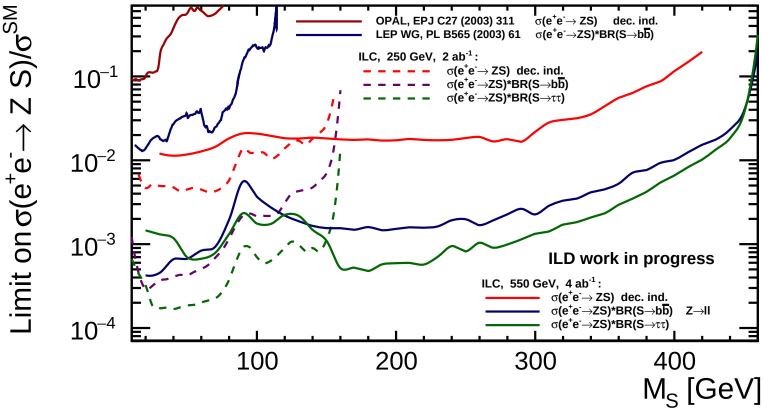
Final limits calculated from the BDT response histogram for signal and background

Limits improve by 20–40% compared to BDT cut method, see backup sliders for details



## Final result

Comparison of all recent ILD results from full simulation / SGV



## Conclusions

Full simulation/SGV analysis of exotic light scalar searches in $S \rightarrow \tau\tau$ channel completed.

Leading background contribution from ZZ events ($qq\tau\tau$ and $\ell\ell\tau\tau$) - irreducible background

Limits dominated by fully hadronic channel (hadronic Z decays, hadronic $\tau$ decays)

No impact of beam polarisation observed (except for effective luminosity)

With 550 GeV data limits at $10^{-3}$ level extended up to masses of about 300 GeV

Limits on $\sigma \times$BR factor 2-3 stronger than for $S \rightarrow b\bar{b}$

250 GeV results already presented at ICHEP'2026 and Corfu'2026

550 GeV results to be presented at HDays'2026

## Cross section limits from BDT histogram

BDT cut rejects 40-60% of signal events. Information on the response distribution is not used.

Better limits can be obtained from the "template fit" to the BDT response histogram:

$$\frac{\delta_\sigma}{\sigma_{ref}} = \left( \sum_{i \in bins} \frac{s_i^2}{b_i} \right)^{-\frac{1}{2}}$$

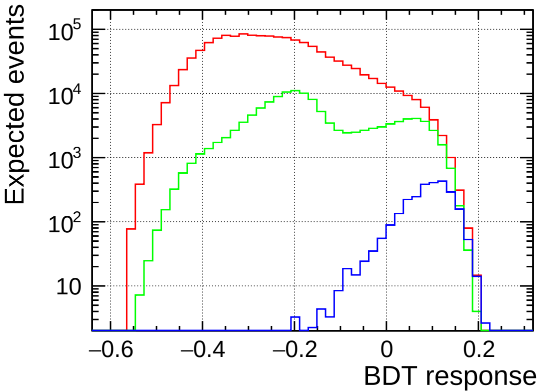
corresponds to combined signal significance in $S \ll B$ limit

Bins with few background events combined to obtain

$$b_i = \sum w_i \geq N_{min} \quad \text{and} \quad \frac{\left(\sum w_i\right)^2}{\sum w_i^2} \geq N_{min}$$

expected events — effective MC events

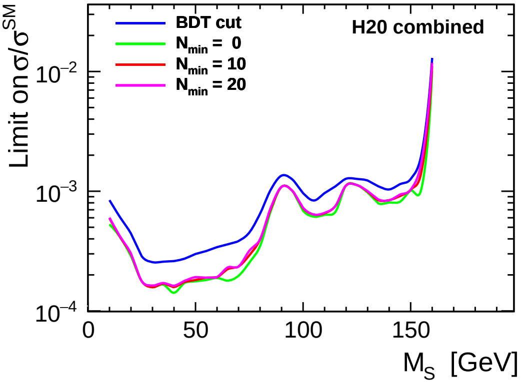
where $w_i$ is the MC event weight.

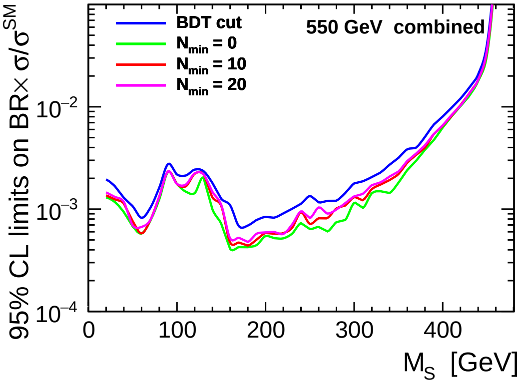



$M_S = 95$ GeV at 250 GeV

## Cross section limits from BDT histogram

Compared to BDT cut method, limits improve by 20–40%:




Stable results of re-binning procedure, largely suppressing fluctuations visible for $N_{min} = 0$