Prospects for light exotic scalar measurements in di-tau decay channel at future Higgs factories.

Weronika Sobień, **Aleksander Filip Żarnecki**
Faculty of Physics, University of Warsaw

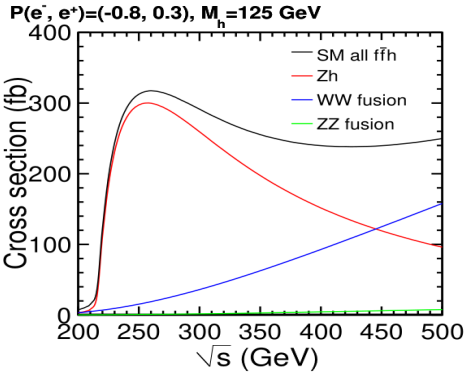


The International Workshop on Future Linear Colliders 2026
Beyond Standard Model session

November 11, 2026

Scalar production at e^+e^- Higgs factory

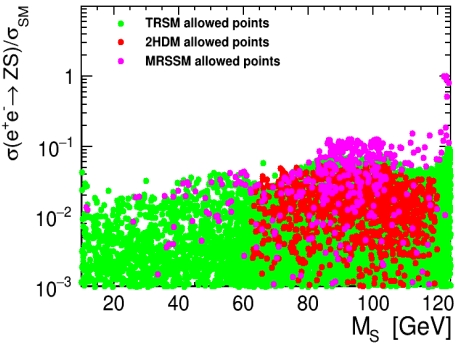
Precision measurements of 125 GeV scalar are clearly the primary target...



In the Zh production channel (higgs-strahlung), dominant below 450 GeV we can use "Z-tagging" for unbiased selection of events.

Scalar production at e^+e^- Higgs factory

Precision measurements of 125 GeV scalar are clearly the primary target...



ECFA'2023 workshop, details in the backup

In the Zh production channel (higgs-strahlung), dominant below 450 GeV we can use "Z-tagging" for unbiased selection of events.

But production of additional, light exotic scalar states is still not excluded by the existing data!

"Z-tagging" should reveal new scalar production in ZS channel (scalar-strahlung)

$\Rightarrow$ decay-mode independent search

But we can also look for particular decays...

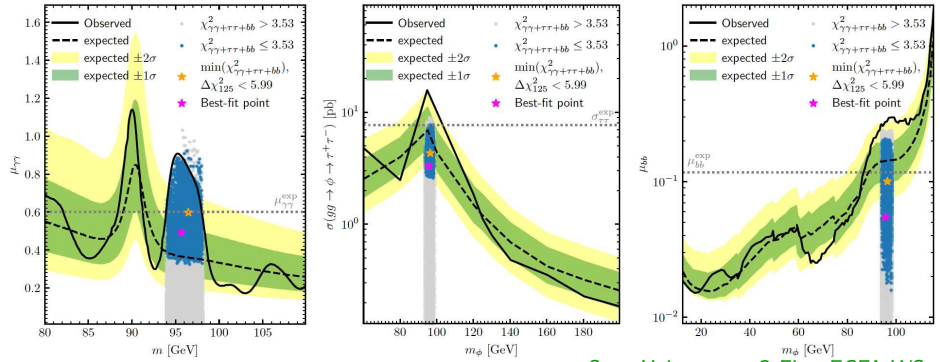
$b\bar{b}$ channel expected to dominate in many models

T. Biekötter, S.Heinemeyer, G. Weiglein
N2HDM analysis, arXiv:2203.13180

Experimental hints for $S \to \tau^+\tau^-$

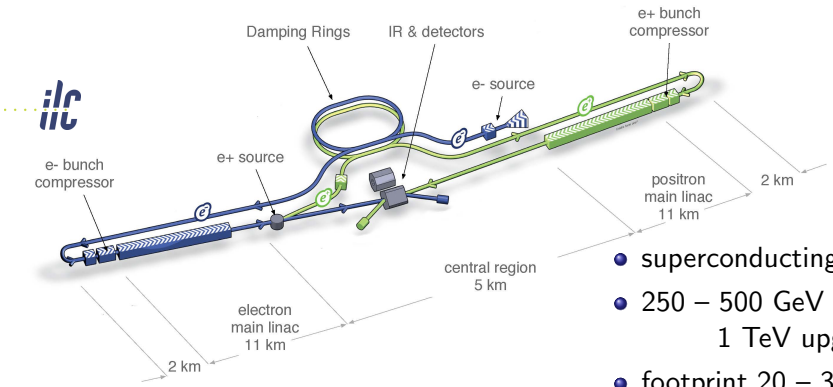
Some discrepancies point to new scalar with mass of ∼95 GeV and dominant decay to $\tau\tau$...

$pp \to h_{95} \to \gamma\gamma$ $\quad gg \to h_{95} \to \tau^+\tau^-$ $\quad e^+e^- \to Zh_{95} \to Zb\bar{b}$



Sven Heinemeyer @ First ECFA WS, October 2022

International Linear Collider Technical Design (TDR) presented in 2013 arXiv:1306.6328

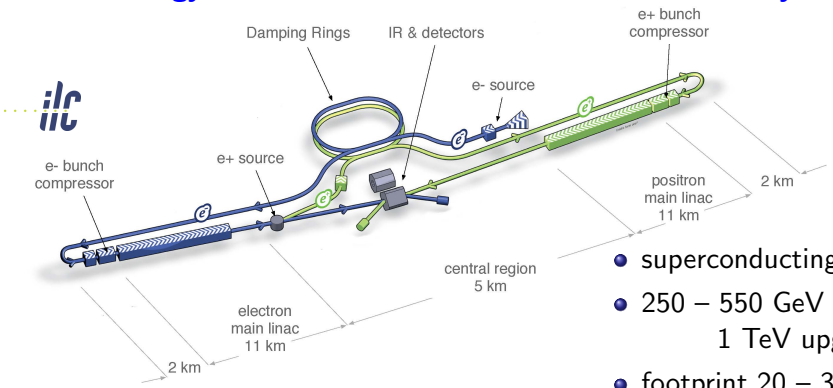


- superconducting accelerating cavities
- 250 – 500 GeV c.m.s. energy (baseline), 1 TeV upgrade possible
- footprint 20 – 31 km

- polarisation for both e^- and e^+ (80%/30%)
- staged construction, starting as 250 GeV Higgs factory arXiv:1903.01629

International Linear Collider Technical Design (TDR) presented in 2013

Technology also considered for a Linear Collider Facility at CERN arXiv:2503.24049



- superconducting accelerating cavities
- 250 – 550 GeV c.m.s. energy (baseline), 1 TeV upgrade possible
- footprint 20 – 33 km

- polarisation for both e^- and e^+ (80%/60%)
- staged construction, starting as 250 GeV Higgs factory

Event simulation

Study based on existing ILD Monte Carlo samples for SM processes at 250 GeV and 550 GeV:

- generated with Whizard v.2.8.5/v.3.1,
- using default (SetA) ILC beam-spectrum,
- simulated and reconstructed using
  a full (GEANT 4) ILD detector simulation/SGV Fast simulation
  with ILD_l5_o2_v02 model and ILCSoft v02-02-01,
- processed using MARLIN modular framework

Signal samples, for scalar particle mass from 10 GeV up to 160 GeV for the 250 GeV scenario and from 20 GeV up to 460 GeV for the 550 GeV energy, were generated and processed with the same tools, but the detector simulation was done with SGV.

Total luminosity of $2\,\mathrm{ab}^{-1}/4\,\mathrm{ab}^{-1}$, with $\pm 80\%/\pm 30\%$ polarisation for e^-/e^+ (H-20 scenario).

Event reconstruction

same procedure for 250 GeV / 550 GeV

Candidate event reconstruction, based on Particle Flow Objects reconstructed with Pandora:
('PandoraPFOs' collection standard ILD reconstruction chain)

- 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)

Event reconstruction

same procedure for 250 GeV / 550 GeV

Candidate event reconstruction, based on Particle Flow Objects reconstructed with Pandora:
('PandoraPFOs' collection standard ILD reconstruction chain)

- 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" (!)

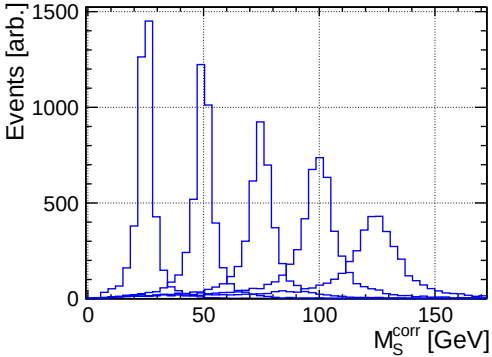
Scalar mass reconstruction

neutrino energies from collinear approximation

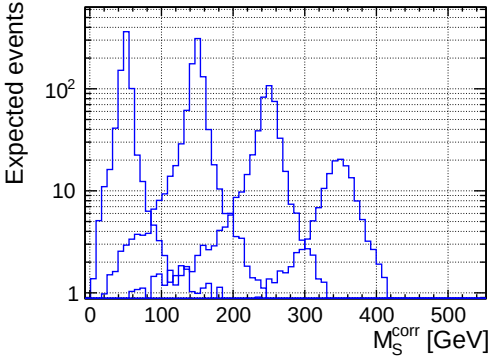
see backup slides

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

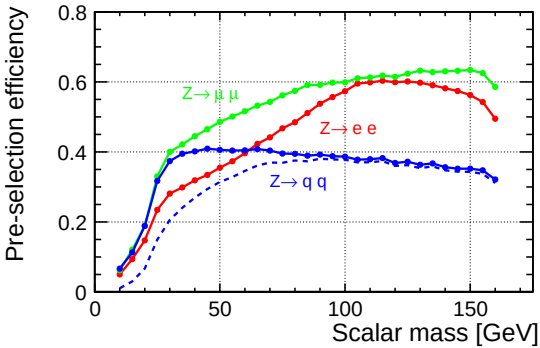


All tau decay channels combined

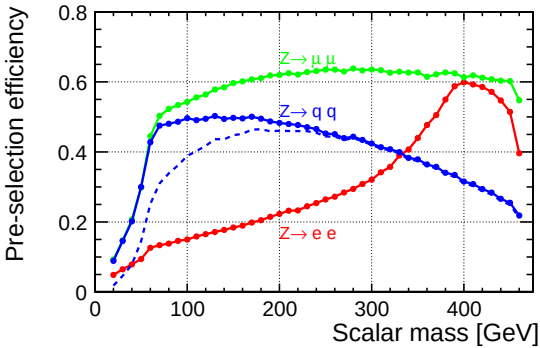
Signal reconstruction efficiency

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

Expected background

Expected SM background contributions after signal reconstruction 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

BDT input variables

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

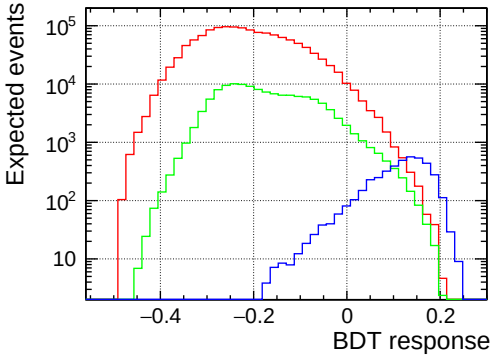
- reconstructed mass of Z candidate (di-lepton or di-jet) and recoil mass
- measured invariant mass of S candidate (di-tau) and corrected S mass
- polar angle of Z candidate, asimuthal angle distance between tau candidates
- scalar decay angles in its CMS
- minimum and maximum distances for PFO clustering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

See backup slides for selected variable distributions

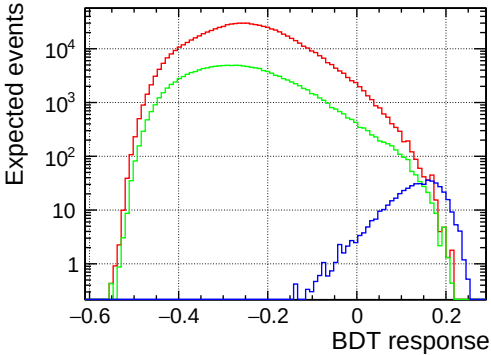
BDT response distribution

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

50 GeV scalar at 250 GeV



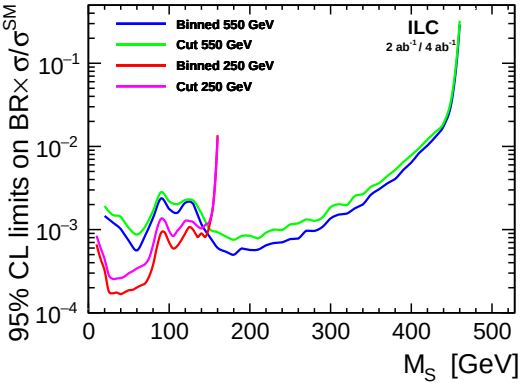
250 GeV scalar at 550 GeV



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

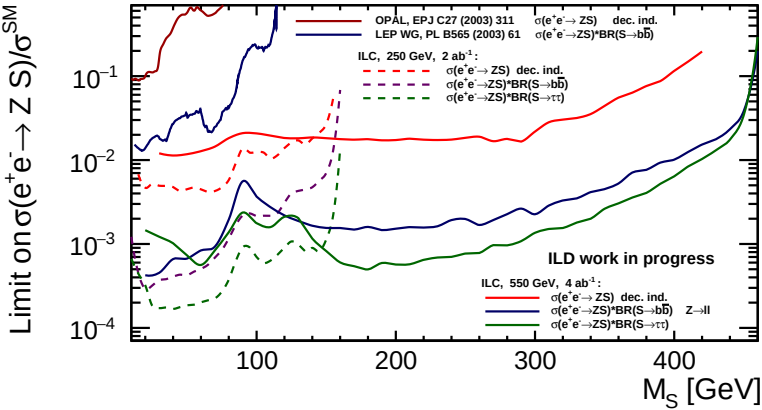
Cross section limits

Final limits calculated from the BDT response histogram for signal and background:
Data from all decay channels and polarisations combined



Cross section limits

Comparison of all recent ILD results from full simulation / SGV



Limits around $m_S = 125\,\text{GeV}$

Our searches are focused on experimental signatures, with minimal model assumptions.
Still, background was calculated as predicted by the SM, including SM Higgs boson.

125 GeV Higgs boson gives significant background contribution for $m_S \sim 125\,\text{GeV}$!

Limits were calculated assuming negligible uncertainty on SM predictions for h_{SM} production!

But what about so called "sum rule"?

In models with extended scalar sector, scalar couplings to Z boson need to fulfill condition:

$$\sum_i g^2_{h_i VV} = g^2_{SM}$$

Limits around $m_S = 125\,\text{GeV}$

Our searches are focused on experimental signatures, with minimal model assumptions. Still, background was calculated as predicted by the SM, including SM Higgs boson.

125 GeV Higgs boson gives significant background contribution for $m_S \sim 125\,\text{GeV}$!

Limits were calculated assuming negligible uncertainty on SM predictions for h_{SM} production!

But what about so called "sum rule"?

If the two scalar states are degenerate in mass, sum of their production cross sections:

$$\sigma(e^+e^- \to Z\,S) + \sigma(e^+e^- \to Z\,h_{125}) \;\approx\; \sigma^{SM}$$

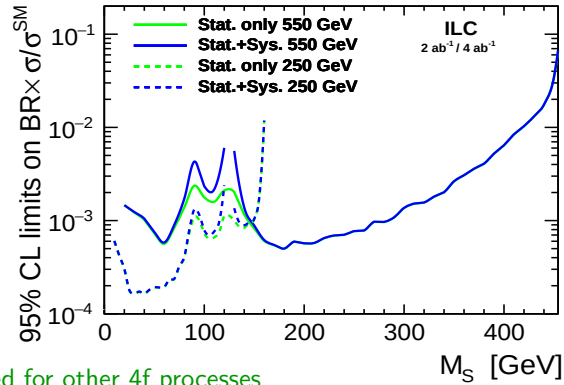
so no excess is expected in production, could only be due to different decay channels.

This has to be taken into account when interpreting limits for scalar mass close to 125 GeV...

Impact of 125 GeV scalar uncertainty

We can try to quantify the impact of model uncertainty by attributing 10% systematic uncertainty to 125 GeV Higgs background predictions

ATLAS Legacy Run 2
arXiv:2608.07332



1% uncertainties assumed for other 4f processes

Conclusions

BSM scenarios with light scalars still not excluded by existing data
Sizable production cross sections for new scalars can coincide with non-standard decay...

Light scalar decays to tau pairs seem a challenging scenario
But also results in most stringent limit on the cross section times branching ratio
promising discovery channel for models with enhanced $\tau^+\tau^-$ decays

Leading background contribution from ZZ events ($qq\tau\tau$ and $\ell\ell\tau\tau$)
and SM Higgs boson production with di-tau decays - irreducible backgrounds

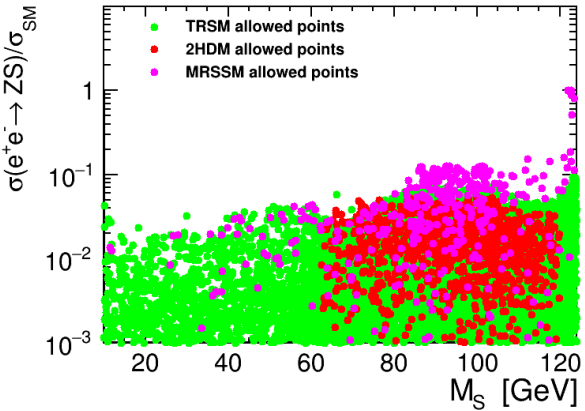
With 550 GeV data limits at 10^{-3} level extended up to masses of about 300 GeV
Limits on $\sigma\times$BR factor 2-3 stronger than for $S \to b\bar{b}$

Significant impact of the possible uncertainty on the 125 GeV Higgs boson contribution.

Thank you!

Possible scenarios (as presented at ECFA'2023 workshop)

Benchmark points consistent with current experimental and theoretical bounds



Two-Real-Singlet Model
thanks to Tania Robens
see arXiv:2209.10996 arXiv:2305.08595

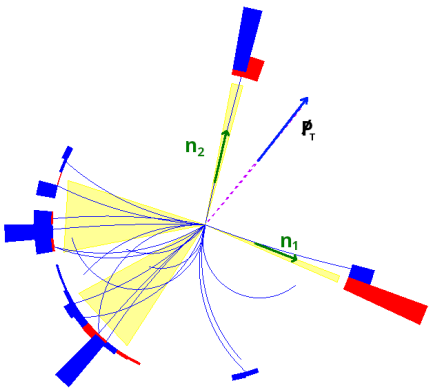
Two Higgs-Doublet Model
thanks to Kateryna Radchenko
thdmTool package, see arXiv:2309.17431

Minimal R-symmetric Supersymmetric SM
thanks to Wojciech Kotlarski arXiv:1511.09334

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:

$$\vec{p}_T = E_{\nu_1} \cdot \vec{n_1} + E_{\nu_2} \cdot \vec{n_2}$$

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

Unique solution !

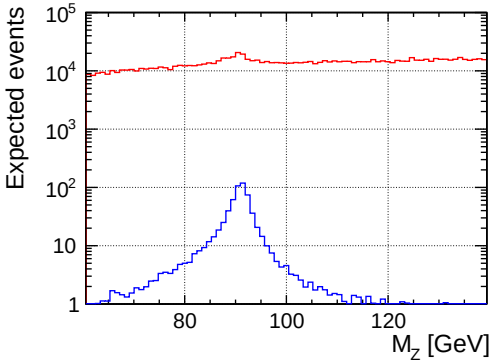
Works also for semi-leptonic and leptonic events!

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

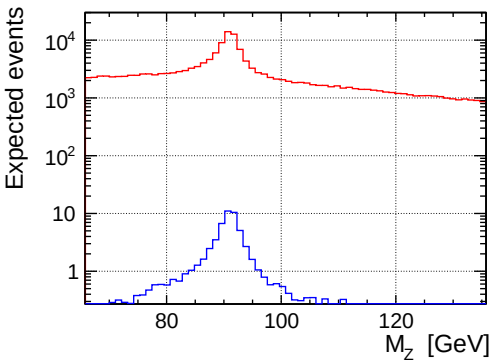
BDT input variables

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

250 GeV:



550 GeV:

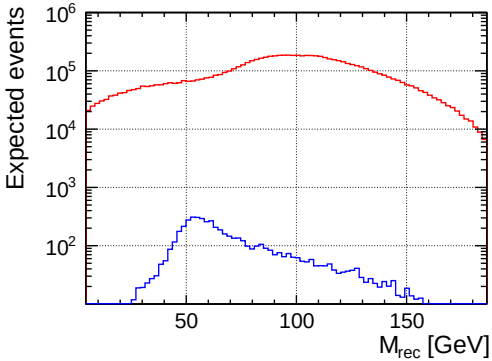


Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

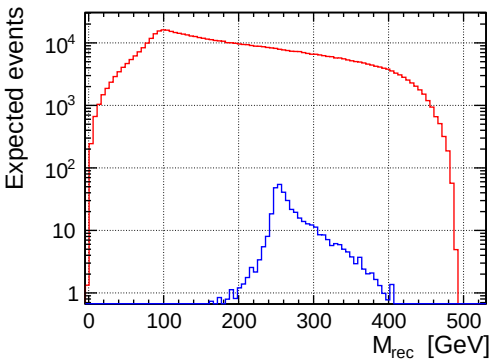
BDT input variables

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

250 GeV:



550 GeV:

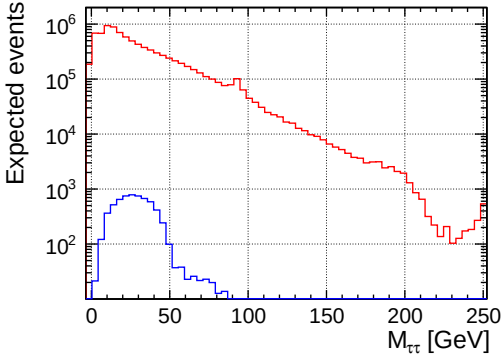


Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

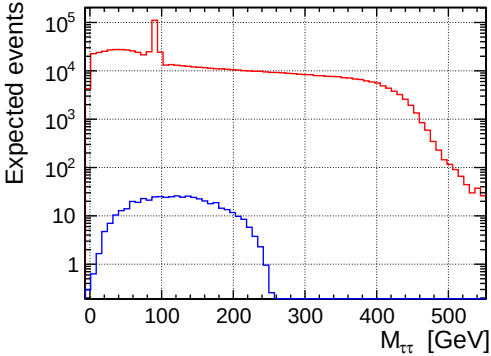
BDT input variables

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

250 GeV:



550 GeV:

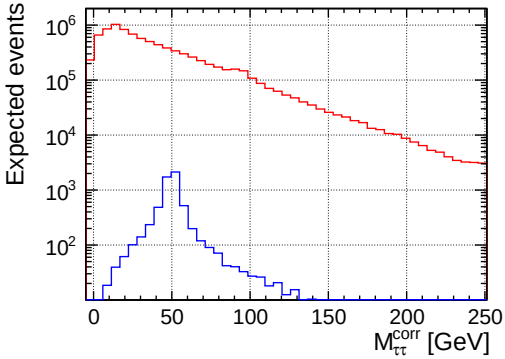


Signal of 50 GeV / 250 GeV scalar production normalized to 1% of SM cross section

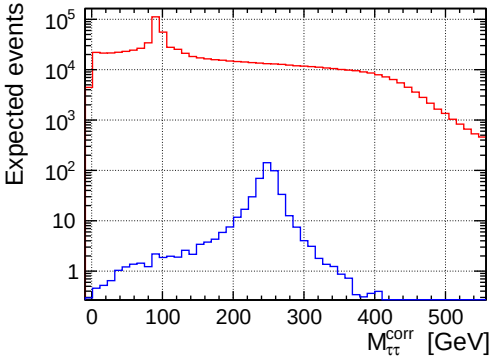
BDT input variables

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

Cross section limits from BDT histogram

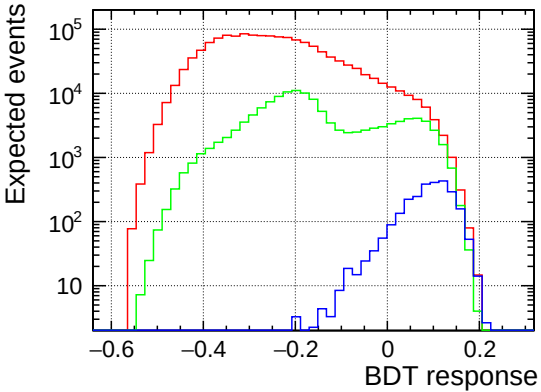
Limits are extracted from the "template fit" to the BDT response histogram:

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

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

where b_i is the number of expected SM background events in each bin and s_i is the number of expected signal events assuming production cross section

$$\sigma_{ref} = 0.01 \cdot \sigma^{SM}\Big|_{m_h = m_S}$$



$M_S = 95$ GeV at 250 GeV

Cross section limits from BDT histogram

Limits are extracted 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

where b_i is the number of expected SM background events in each bin and s_i is the number of expected signal events assuming production cross section

$$\sigma_{ref} = 0.01 \cdot \sigma^{SM}\Big|_{m_h = m_S}$$

Impact of systematic uncertainties can be taken into account as additional nuisance parameters in the fit:

$$\mu_i = b_i + \frac{\sigma}{\sigma_{ref}} \cdot s_i + \sum_{k=1}^{N_{syst}} \Delta_{ki} \cdot \delta_k$$

where Δ_{ki} describes change of SM background in bin i under unit variation of systematic uncertainty k