Recovering events with Soft Secondary Vertices in the diHiggs to 4b analysis

Lukas Roscher
DESY

Introduction and motivation

The Large Hadron Collider (LHC) collides protons at a centre-of-mass energy of 13.6 TeV. ATLAS is one of the large experiments recording these collisions. Its inner detector measures charged-particle tracks, while the calorimeters measure the energy deposits of traversing particles. At a hadron collider, quarks and gluons (partons) play an important role in a lot of processes of interest. Partons can not be observed as free particles as they hadronize and produce collimated sprays of particles, which are reconstructed as hadronic jets. Jets containing a b-hadron are of big interest in a lot of physics processes at the LHC andcan be identified through dedicated b-tagging algorithms.

Higgs-pair production is an important target of the ATLAS physics programme and has not been observed yet. The coupling parameter λ in the Higgs potential is not measured yet. The Standard Model predicts its value, while physics beyond the Standard Model can modify it. The value of this coupling changes both the production rate and the shape of the m_{HH} distribution, with the low-mass region being especially important. Measuring or constraining this coupling is a main motivation of measuring Higgs-pair production. The $HH \to b\bar{b}b\bar{b}$ decay has the highest branching ratio and therefore a large signal yield. At the same time, it has a huge QCD multijet background. The resolved $HH \to b\bar{b}b\bar{b}$ analysis uses at least four reconstructed jets and four b-tags to select this final state.

This requirement creates a simple problem. Jets in ATLAS are reconstructed and calibrated only above a transverse-momentum threshold of 20 GeV. Additionally, b-taggers are not perfect. If one Higgs boson produces a soft b-quark, the corresponding b-hadron may not form a jet. The b-tagger then has no jet to analyse, and an otherwise useful $HH \to b\bar{b}b\bar{b}$ event fails the four-b selection. Here, "soft" refers to the low momentum of the b-hadron. My analysis studies whether *soft secondary vertices* (SSVs) can identify the missing b-hadrons and recover such events.

Recovering the missing b-hadron

A b-hadron has a lifetime of about 1.5 ps. It travels a measurable distance through the detector before it decays, so its tracks do not point back exactly to the primary proton–proton collision. Instead, they form a displaced secondary vertex. Standard b-tagging searches for this signature inside a reconstructed jet.

The NewVrtSecInclusive (NVSI) algorithm starts from the tracks instead. It combines them into displaced vertices and applies several cleaning and vertex-merging steps. Most importantly, it is not restricted to tracks inside jets. It can therefore reconstruct the decay of a soft b-hadron even if no reconstructed jet is present. The SSV is not a replacement for the full jet energy (since it doesn't contain energy information from the calorimeters). It is first of all a tag and a direction: evidence that the three-b event may actually contain the fourth b-hadron from $HH \to b\bar{b}b\bar{b}$.

The new selection requires at least three jets, exactly three b-tags (instead of 4) and at least one "good" SSV outside all b-tagged jets. It is orthogonal to the standard $\geq 4j$, $\geq 4b$ region, so it adds new events without double counting those already used in the standard analysis. About 90% of all reconstructed SSVs are already inside b-jets and do not add new information. The remaining SSVs are the interesting ones: Fig. 1 shows a clear population outside the b-jet radius of $\Delta R = 0.4$. Of these useful SSVs, 48.9% are inside a non-b-tagged jet and 51.1% are isolated from all selected jets.

Signal gain and background

The SSV selection recovers a sizeable signal contribution. Table 1 shows the yields of the inclusive $3b1$SSV category relative to the standard $\geq 4b$ category.

Event category	Signal yield	QCD $bbbb$ yield
Standard $\geq 4b$	100.0%	100.0%
Inclusive $3b1$SSV	42.5%	66.9%

Table 1: Weighted yields relative to the standard $\geq 4b$ category. Since the two selections are orthogonal, the $3b1$SSV signal is an additional contribution.

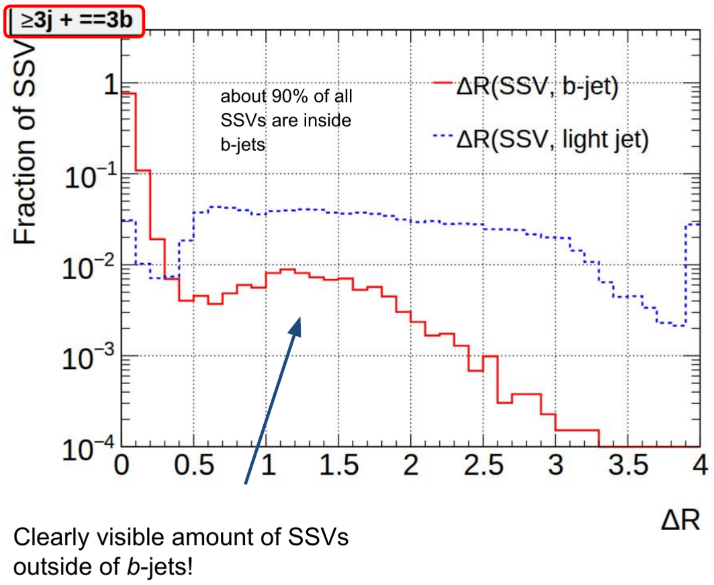

Figure 1: Distance of reconstructed SSVs to the closest b-tagged jet (red) and non-b-tagged jet (blue). Most SSVs lie inside b-jets, but the red tail above $\Delta R = 0.4$ demonstrates that a usable population exists outside them.

The table shows that the SSV category adds about 40% signal. Below $m_{HH} = 400$ GeV, where we are most sensitive to the Higgs self-coupling, the gain reaches about 49%. At the same time, the QCD background increases by 66.9%, more strongly than the signal. Since we want to keep a high signal to background ratio for the final fit on the coupling parameter and signal strength, this 3b1SSV region has to be treated individually, which is statistically possible because the events are orthogonal to the 4b region.

Higgs candidate pairing

A remaining challenge of the current resolved $HH \to b\bar{b}b\bar{b}$ analysis is to assign the reconstructed objects to the two Higgs candidates. With 4 jets (or 3 jets and 1 SSV), we have 3 different combinations of pairing jet pairs to the two Higgs. There are many options on how to pair these jets. The current strategy is taking the minimum angular distance between 2 b-jets, which works well for high m_{HH}, since the decay products are highly collimated. But this pairing startegy loses efficiency at low m_{HH}, exactly where the sensitivity to the Higgs self-coupling is largest, which means we are blind to the most important events (for determining the self-coupling). Future studies must improve this assignment while also checking how well the chosen strategy rejects the QCD background and whether it shapes the background into a signal-like mass distribution. One thing we are studying is to combine different pairing methods, like taking the maximum angular seperation for low m_{HH} and the minimum angular seperation for high m_{HH}.

Summary and outlook

SSVs provide a way to identify a b-hadron even when ATLAS software fails to reconstruct them as a jet. This happens simply because the b-hadron is too soft. Without SSV reconstruction, events where this happens would otherwise be lost, since usual b-tagging requires a reconstructed jet. An inclusive $3b1$SSV region could add about 40% signal overall and about 49% below 400 GeV. This is exactly the region which carries important information on the Higgs self-coupling.

There are a lot of things to consider still. Isolated SSVs need an energy correction. The QCD background must be constrained in an individual fit region. Pairing the 4 higgs candidates to the actual Higgs is a major challenge. The current pairing methods are blind to events at low m_{HH}, wherewe are most sensitive to the Higgs self-coupling. Studies are done to see if the Higgs pairing can combine different pairing strategies to successfully cover the whole range of m_{HH}. With these steps, SSVs can turn events which are currently lost into useful signal candidates.