Boosting Strong Higgs Pair Production at the LHC

Juan Rojo
STFC Rutherford Fellow
Rudolf Peierls Center for Theoretical Physics, University of Oxford

BOOST2014 Workshop
UCL, 21/08/2014

Apologies .....

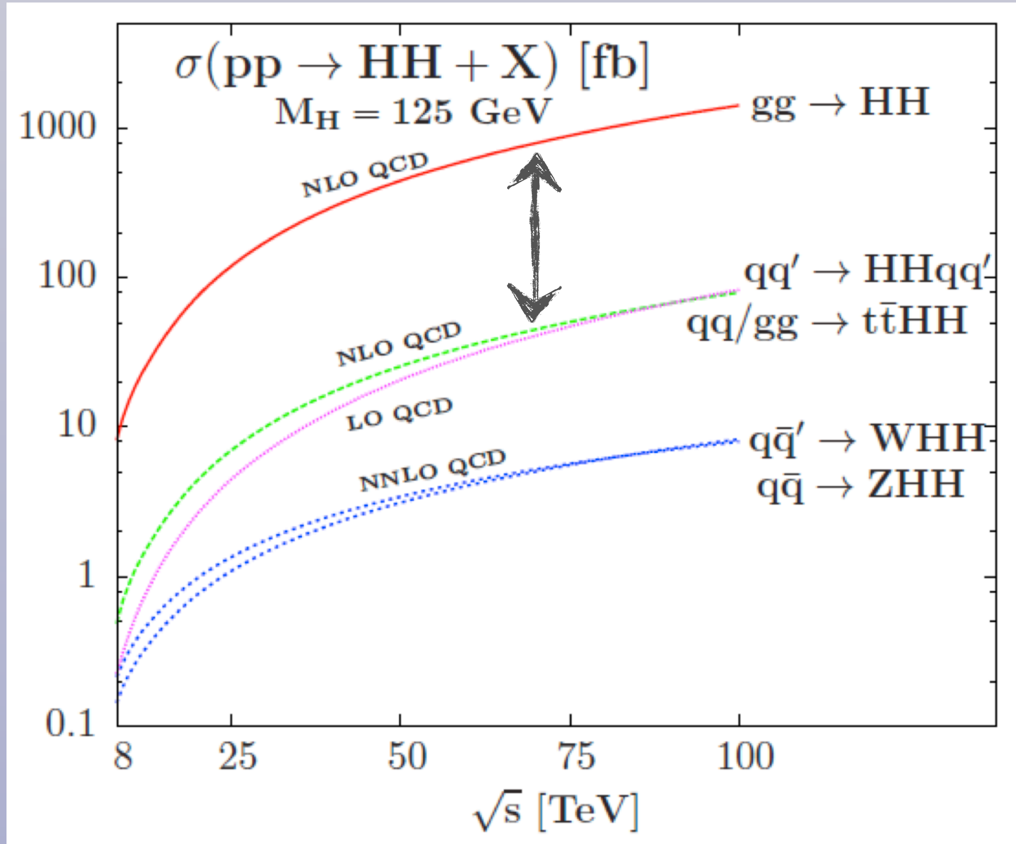
- For being able to participate only today in this fascinating Workshop!
- But I am in the right middle of the family move to Oxford ....

Motivation

Why Higgs Pair Production in Vector-Boson-Fusion relevant?

Motivation

- **Higgs pair production** is one of the most crucial processes for future LHC runs, since it allows to perform stringent tests of our understanding of electroweak symmetry breaking
- In the SM, the dominant process is gluon fusion, with **~33 fb (~1.5 pb)** at 14 TeV (100 TeV): direct sensitivity to the **Higgs trilinear coupling λ_3**
- Higgs pair production in **Vector-Boson Fusion** is small in the SM: **2 fb (80 fb)** at 14 TeV (100 TeV), yet provides unique information on the **hhVV coupling**
- VBF Higgs pair production can be **substantially enhanced** in scenarios where electroweak symmetry breaking is broken by **new strong dynamics** (like in composite Higgs models)

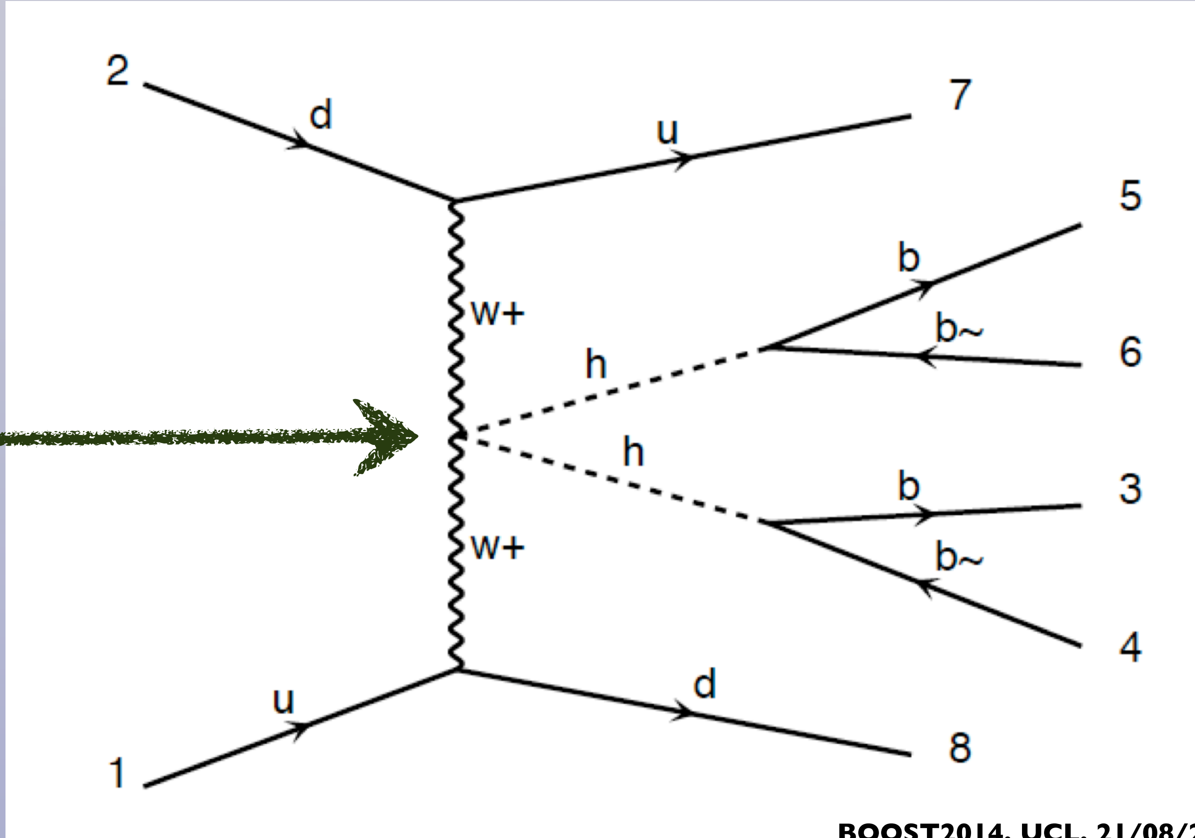


J. Baglio et al
arxiv:1212.5581

Motivation

- **Higgs pair production** is one of the most crucial processes for future LHC runs, since it allows to perform stringent tests of our understanding of electroweak symmetry breaking
- In the SM, the dominant process is gluon fusion, with ~**33 fb (~1.5 pb)** at 14 TeV (100 TeV): direct sensitivity to the **Higgs trilinear coupling λ_3**
- Higgs pair production in **Vector-Boson Fusion** is small in the SM: **2 fb (80 fb)** at 14 TeV (100 TeV), yet provides unique information on the **hhVV coupling**
- VBF Higgs pair production can be **substantially enhanced** in scenarios where electroweak symmetry breaking is broken by new **strong dynamics** (like in composite Higgs models)
- Closely related process to **WW scattering**

Unique direct sensitivity to the hhVV coupling and the the SM unitarization mechanism



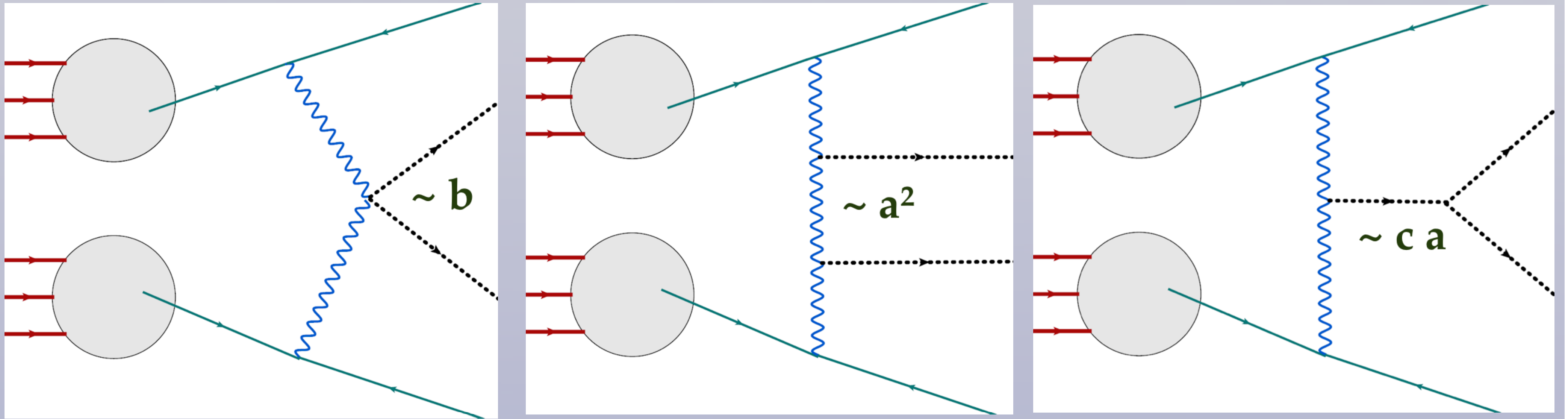
Strong Double Higgs Pair production

- In **composite Higgs models** with **new strong dynamics**, the predictions for VBF Higgs pair production at the hadron colliders can be substantially enhanced as compared to their SM values

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)^2 - V(h) + \frac{v^2}{4}\mathrm{Tr}\left(D_\mu\Sigma^\dagger D^\mu\Sigma\right)\left[1 + 2a\,\frac{h}{v} + b\,\frac{h^2}{v^2} + \dots\right]$$

$$- m_i\,\bar{\psi}_{Li}\,\Sigma\left(1 + c\,\frac{h}{v} + \dots\right)\psi_{Ri} + \text{h.c.}\,,$$

in SM, a=b



- In the SM, **a=b=c=1**
- The **hVV coupling** is constrained from single Higgs production up to **O(10-20%)**
- No model independent direct constraints available on the **hhVV and hhh couplings** yet

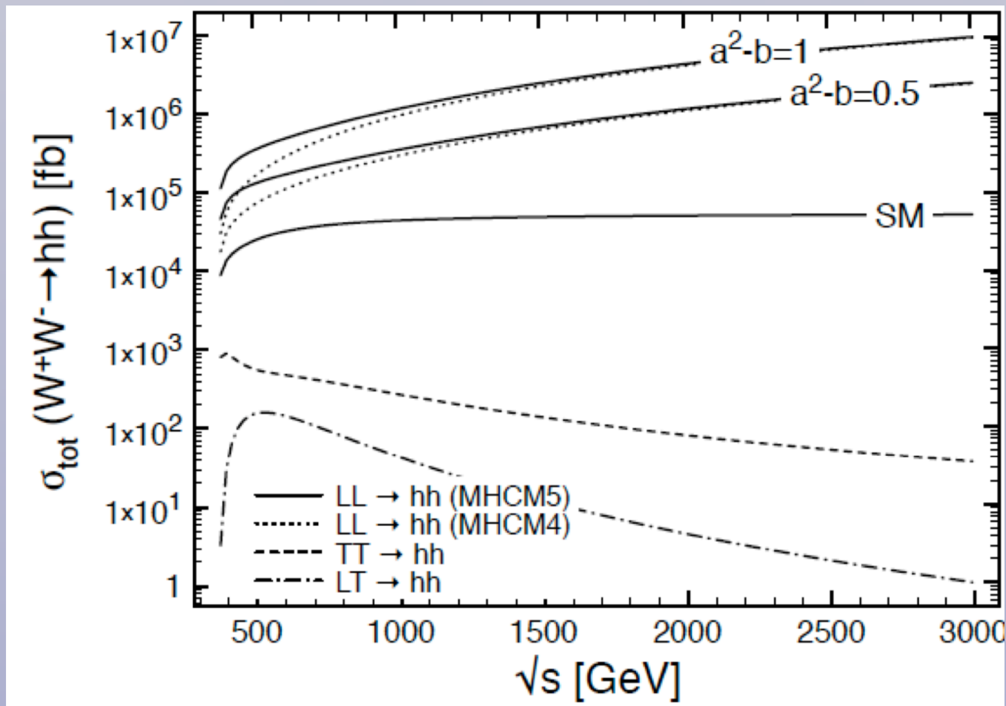
Strong Double Higgs Pair production

In **composite Higgs models** with **new strong dynamics**, the predictions for **VBF Higgs pair production** at the hadron colliders can be substantially enhanced as compared to their SM values

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)^2 - V(h) + \frac{v^2}{4}\mathrm{Tr}\left(D_\mu \Sigma^\dagger D^\mu \Sigma\right)\left[1 + 2a\frac{h}{v} + b\frac{h^2}{v^2} + \dots\right]$$
$$- m_i \bar{\psi}_{Li}\, \Sigma \left(1 + c\frac{h}{v} + \dots\right)\psi_{Ri} + \mathrm{h.c.}\,,$$

in SM, a=b

$$\frac{d\sigma_{LL\to hh}}{dt} \simeq \frac{(b-a^2)^2}{32\pi v^4}\,, \qquad \frac{d\sigma_{TT\to hh}}{dt} \simeq \frac{g^4(a^4 + (b-a^2)^2)}{64\pi s^2}\,,$$



- Most **striking signature** is that the **new strong dynamics** lead to a much harder distribution in **M_{HH}** as compared to the SM
- Original feasibility study by **Contino, Grojean, Moretti, Piccinini and Rattazi in arxiv:1002.1011** assumed a **180 GeV Higgs** and focused on the dominant **WW final state**
- Our goal is to revisit the analysis for **m_H=125 GeV** for the final states with larger BR.
- For **b ≠ a**, Higgs pairs produced with **large boosts: jet substructure techniques** needed

Scale-Invariant Resonance Tagging

M. Gouzevich, A. Oliveira, J. Rojo, R. Rosenfeld, G. Salam, V. Sanz
arXiv:1303.6636, JHEP 07 (2013) 148

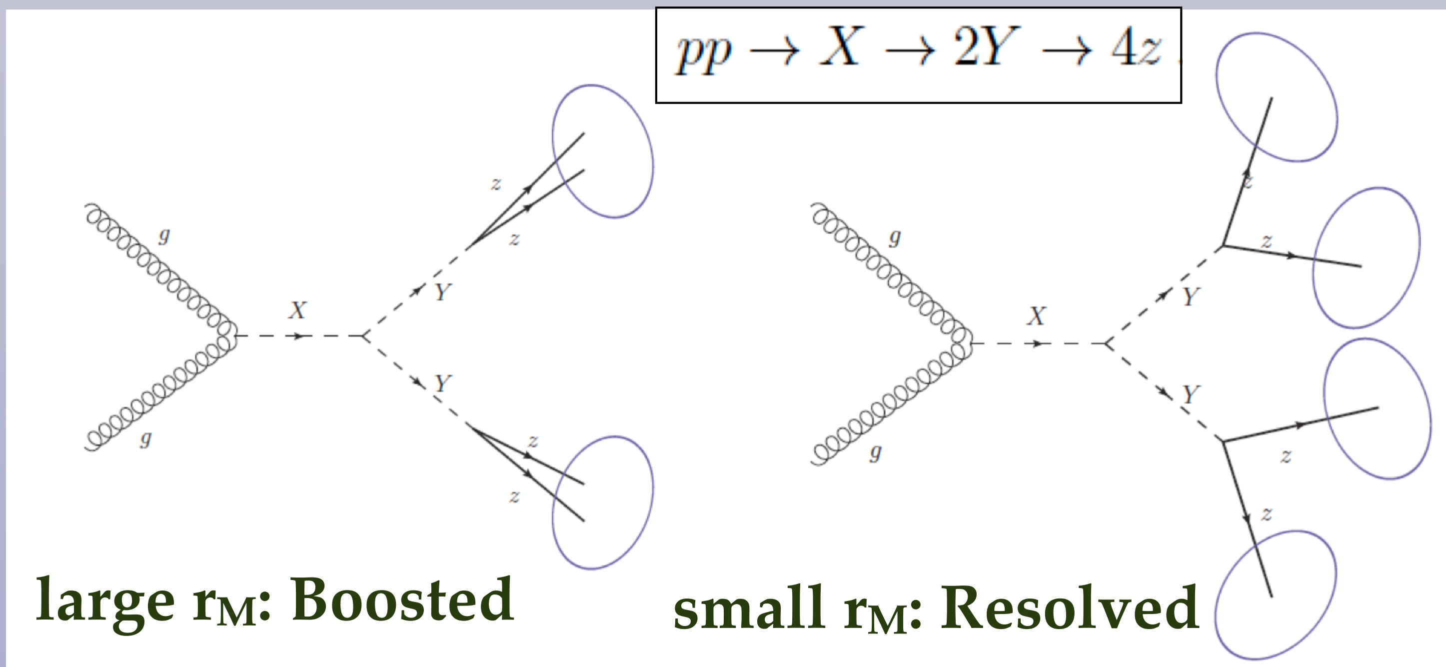
See also my talk at **BOOST2013** (Thanks Gavin!)

Motivation for scale-invariant tagging

- Many BSM scenarios involve **resonant pair production** of heavy (SM and BSM) particles
- For example, let's consider a generic production kinematics of the type

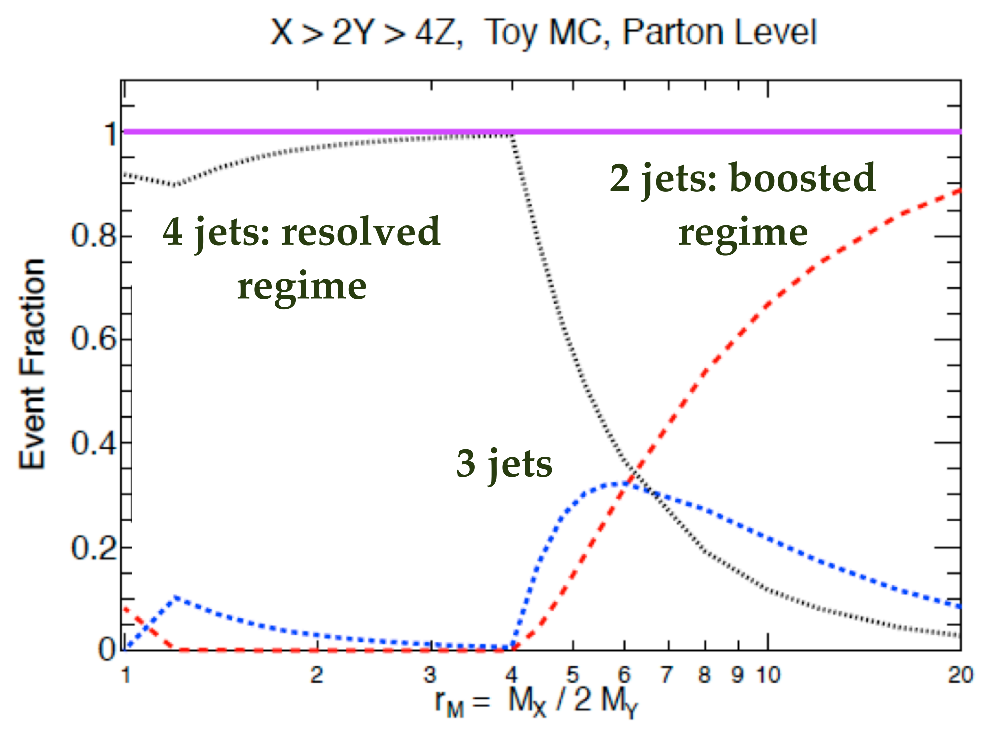
$$pp \to X \to 2Y \to 4z$$

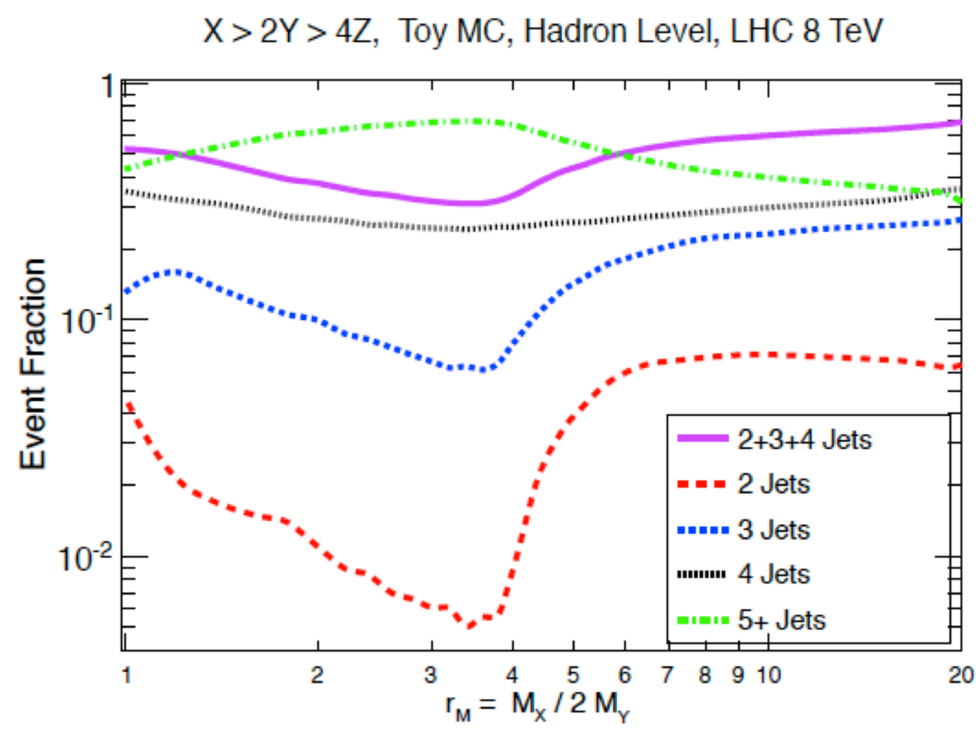
- Depending on the value of the mass ratio $\mathbf{r_M = M_X/2M_Y}$ different final state topologies arise
- For large $\mathbf{r_M}$ the intermediate heavy particles *Y* will be **highly boosted,**
- For small $\mathbf{r_M}$ the *Y* particles are produced close to rest, and the four decay particles **z** are well separated in the detector: **resolved regime**
- Design a **search strategy** that efficiently explores simultaneously **the whole $\mathbf{r_M}$ range,** and improves the overall efficiency by including the **intermediate mass** regime



Event Classification

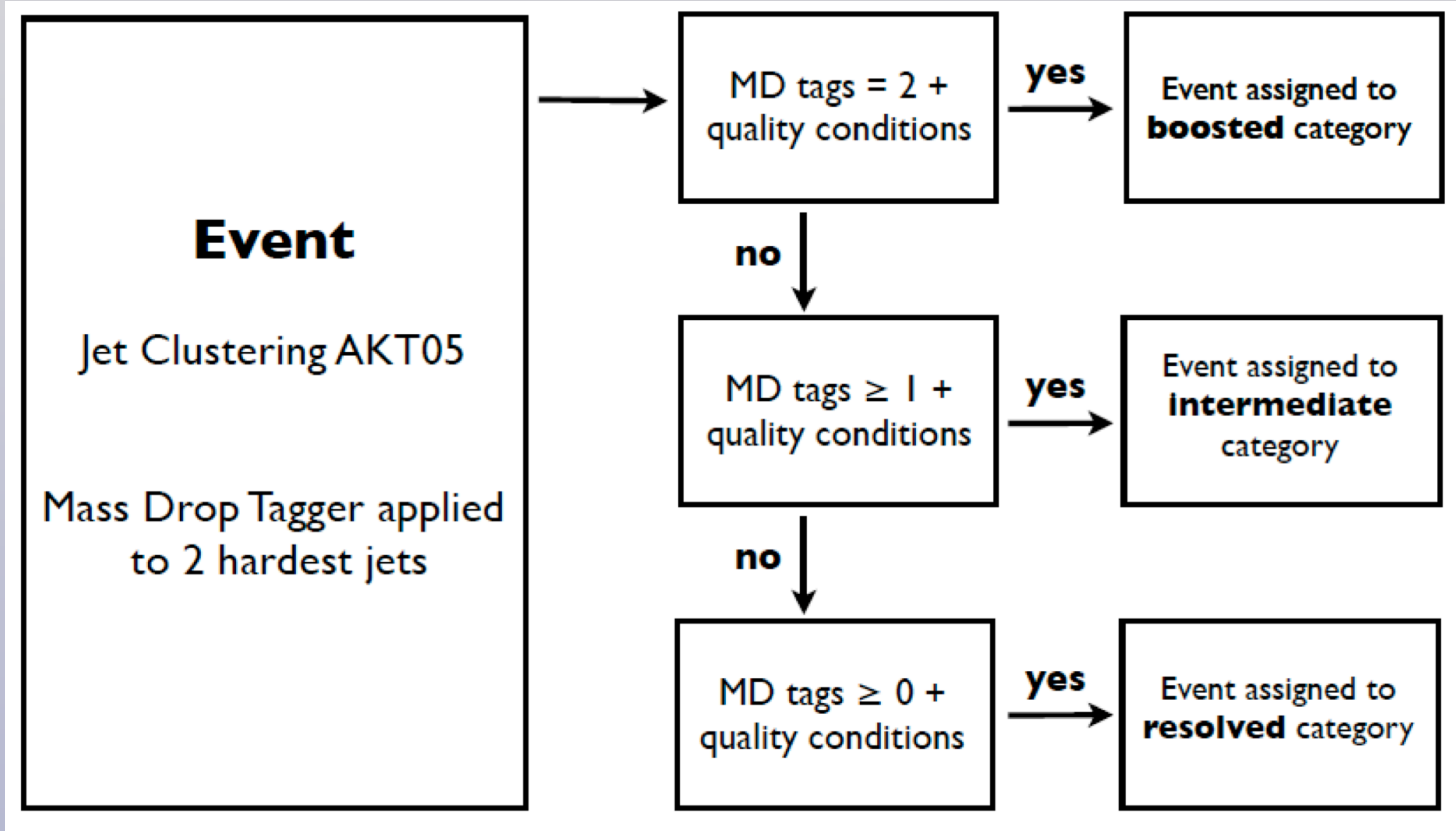
- Generated events for **X > 2Y > 4z** with an **in-house toy MC**, interfaced to **Pythia8** for showering and hadronization
- At **parton level**, without cuts, the classification of the event topology, *boosted*, *resolved* or *intermediate*, can be trivially obtained **based on the number of jets**
- But at **hadron level** with **realistic cuts** such naive classification is not feasible





Event Classification

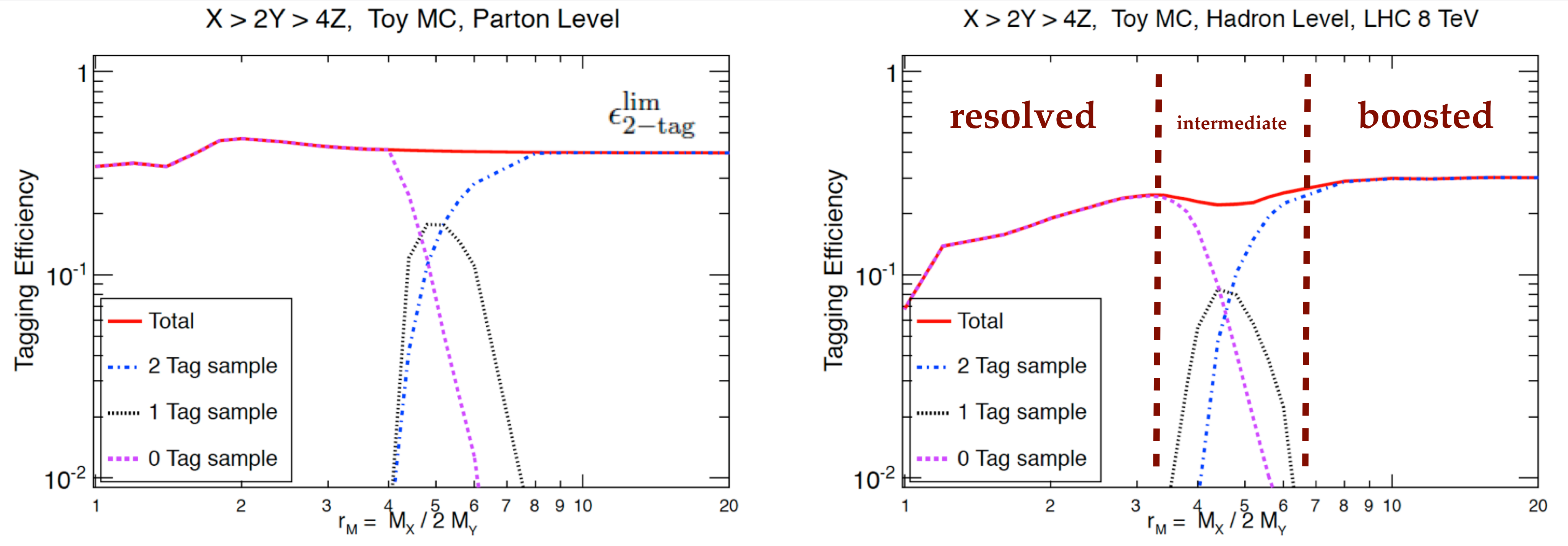
- Use **event classification** based on the number of **mass-drop substructure tags** of leading two jets



- **Boosted category:** each MD tagged jet assumed to be a Y resonance candidate
- **Intermediate category:** MD tagged jet first Y resonance, then pair the other two leading jets in event
- **Resolved category:** The two Y resonance candidates determined from dijet pairing that minimizes M_Y difference

Scale-invariant tagging

- Tagging efficiency **independent of the value of the mass ratio** (except hadron level small $\mathbf{r_M}$)
- **Smooth interpolation** between the boosted and resolved regimes



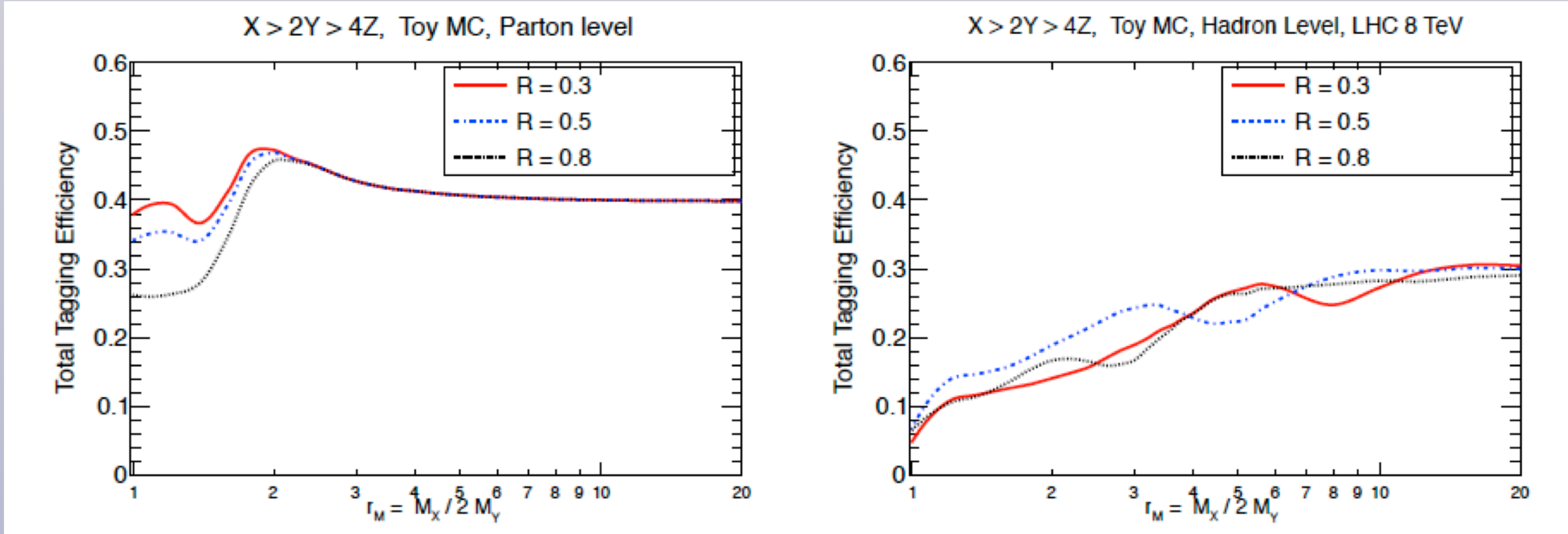
- At parton level the **tagging efficiency in the boosted limit** can be computed analytically

$$\epsilon^{\lim}_{2-\rm tag} \equiv \epsilon_{2-\rm tag}\left(r_M \gg 1\right) = \left(1 - \frac{2y_{\rm cut}}{1+y_{\rm cut}}\right)^2 \cdot \frac{\exp(\Delta y_{\max})-1}{\exp(\Delta y_{\max})+1} \sim 0.40$$

Scale-invariant tagging: with a **single analysis**, explore **simultaneously both the boosted and resolved regimes**, with a smooth interpolation for intermediate masses

Scale-invariant tagging

- Tagging efficiency is also **independent of the value of jet radius**
- **The relative classification of the events** in the **resolved**, **boosted** and **intermediate** categories depends on **R,** but the total tagging efficiency is reasonably **R-independent**

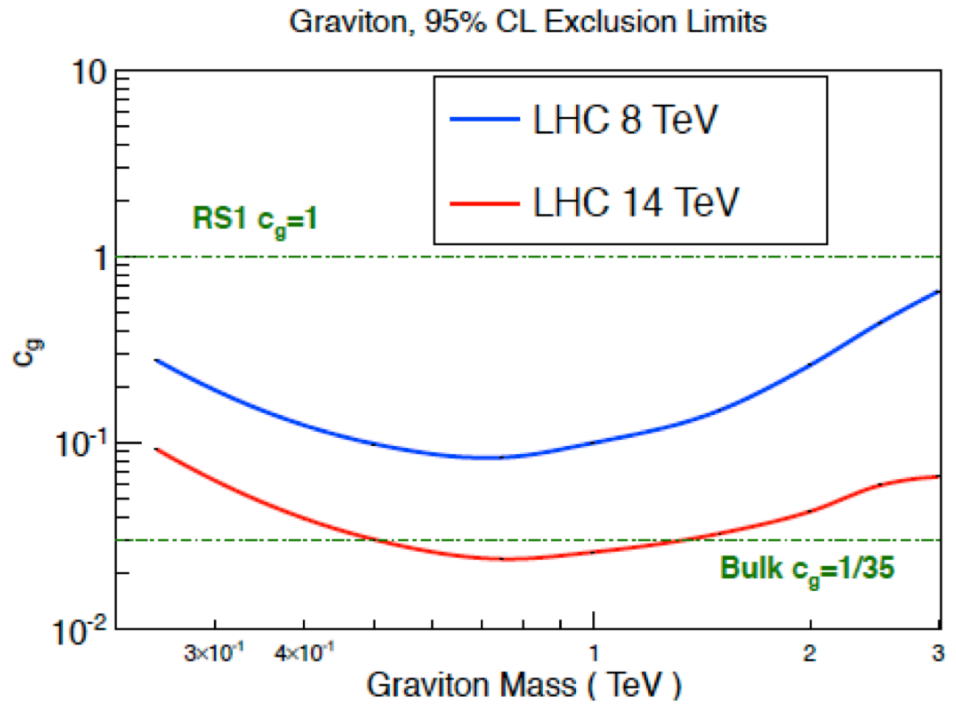


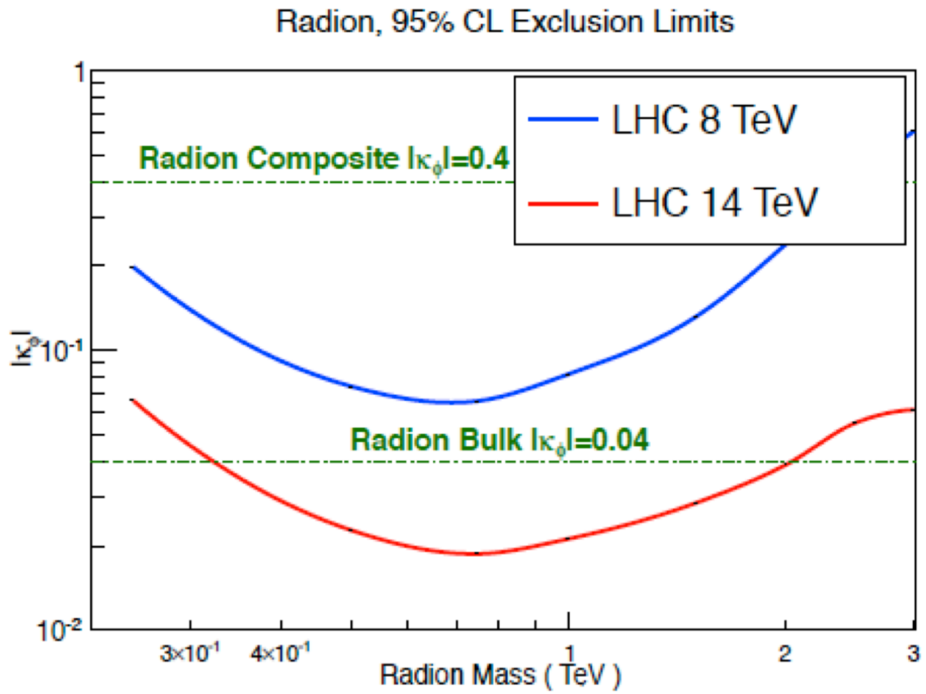
Scale-invariant tagging: with a **single analysis**, explore **simultaneously both the boosted and resolved regimes**, with a smooth interpolation for intermediate masses

Radius-independent tagging: Results are resilient against choice of **R**

Application to HH -> 4b

- Using the scale-invariant tagging, determined **model-independent bounds** on generic resonances (for all masses) decaying into Higgs pairs
- Also obtained explicit bounds in the parameter space of **Warped Extra Dimensions** models
- Good sensitivity for **all resonance masses,** for **resolved, boosted and semiboosted regimes**





Resolved **Boosted** **Resolved** **Boosted**

Boosting Strong Higgs Pair Production at the LHC

O. Bondu, R. Contino, A. Massironi and J. Rojo
in preparation

preliminary results in the **Les Houches 2013 BSM report**

Event Generation

- The low rates for Higgs pair production in VBF emphasize the need for final states with large branching fractions: we use here the **4b2j** and **2b2τ2j** final states.
- **Signal events** have been generated with **MadGraph5_aMC@NLO**, with the **hhV, hhVV and hhh couplings** rescaled in a way that $c_V=c_{2V}=c_3=1$ are the SM values

$$C_V \kappa_{hhV} hhV \,, \qquad C_{2V} \kappa_{hhVV'} VV' \,, \qquad C_3 \kappa_{hhh} hhh$$

- Analytical dependence of the cross-section in these parameters given by (R. Contino et al, arxiv:1309.7038)

$$\sigma = c_V^4 \sigma_{\rm SM} \left(1 + A\delta_{c_{2V}} + B\delta_{c_3} + C\delta_{c_{2V}}\delta_{c_3} + D\delta_{c_{2V}}^2 + E\delta_{c_3}^2\right) \qquad \delta_{c_{2V}} \equiv 1 - \frac{c_{2V}}{c_V^2} \quad , \delta_{c_3} \equiv 1 - \frac{c_3}{c_V}$$

- Background events have been generated with **Alpgen and MadGraph5**
  - ☑ **4b2j final state** QCD multijet production of **4b2j and 2b4j**
  - ☑ **2b2τ2j final state: ttjj** dominant background, **2b4j** also relevant when light jets fake **τ** leptons
- **Signal and background** parton level events are then showered and hadronized with **Pythia8.** Jet clustering is performed using **FastJet** with the anti-kT algorithm with **R=0.4**
- **Realistic b-tagging and τ-tagging,** including mistag rates, along the lines of ATLAS/CMS, has been implemented.
- We have studied the results both at the **LHC 14 TeV with 300 fb^{-1} and 3000 fb^{-1}** and at an **FCC at 100 TeV with 3000 fb^{-1}**

Selection and analysis cuts

Our basic selection cuts, including the **vector-boson fusion cuts** to suppress background, are

Basic acceptance cuts

$$p_{Tj} \geq 25 \text{ GeV}\,, \quad p_{Tb} \geq 25 \text{ GeV}\,, \quad p_{T\tau} \geq 25 \text{ GeV}$$

$$|\eta_j| \leq 4.5\,, \quad |\eta_b| \leq 2.5\,, \quad |\eta_\tau| \leq 2.5$$

$$\Delta R_{j\tau} \geq 0.4\,, \quad \Delta R_{b\tau} \geq 0.4\,, \quad \Delta R_{\tau\tau} \geq 0.2\,,$$

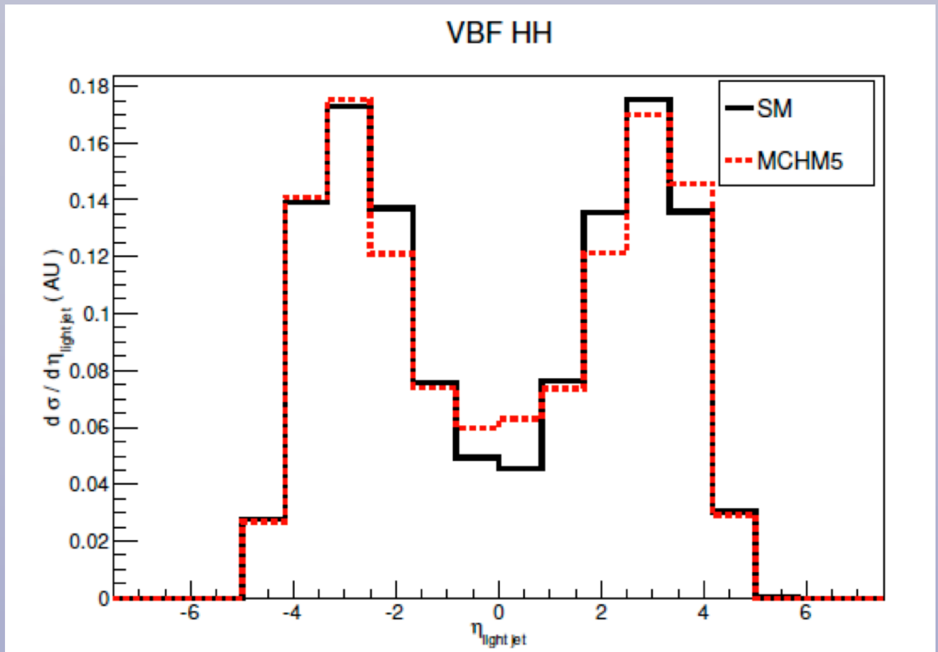
VBF cuts

$$m_{jj} \geq 1000 \text{ GeV}\,, \qquad \Delta y_{jj} \geq 5.0\,.$$

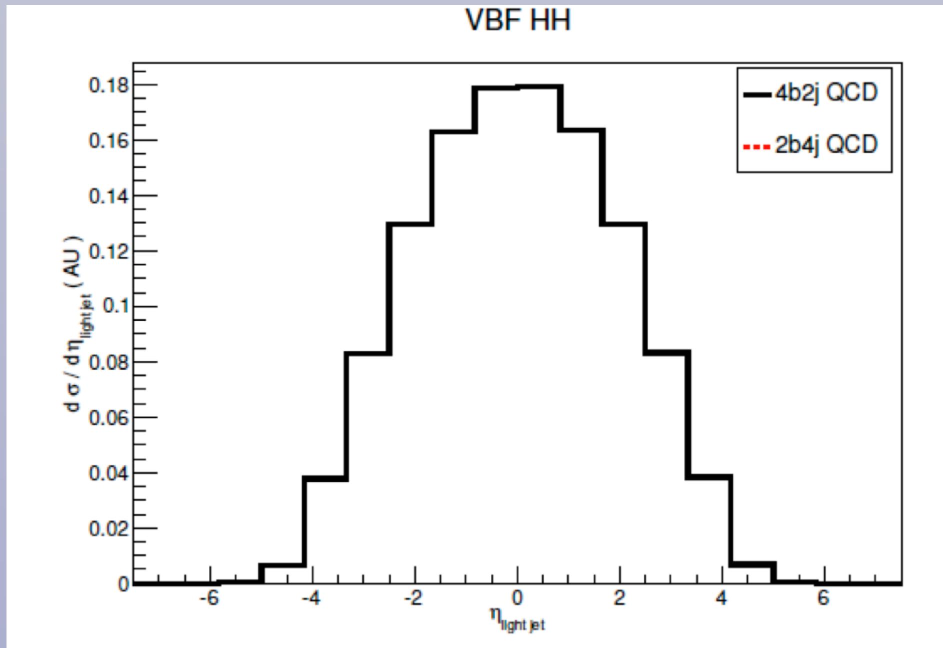
Central jet veto

$$p_{Tj,3} \leq 30 \text{ GeV} \quad \text{if} \quad \eta_{j,\text{min}} \leq \eta_{j,3} \leq \eta_{j,\text{max}}$$

VBF tagging jets rapidity - Signal



VBF tagging jets rapidity - Background



Event rates at the LHC and beyond

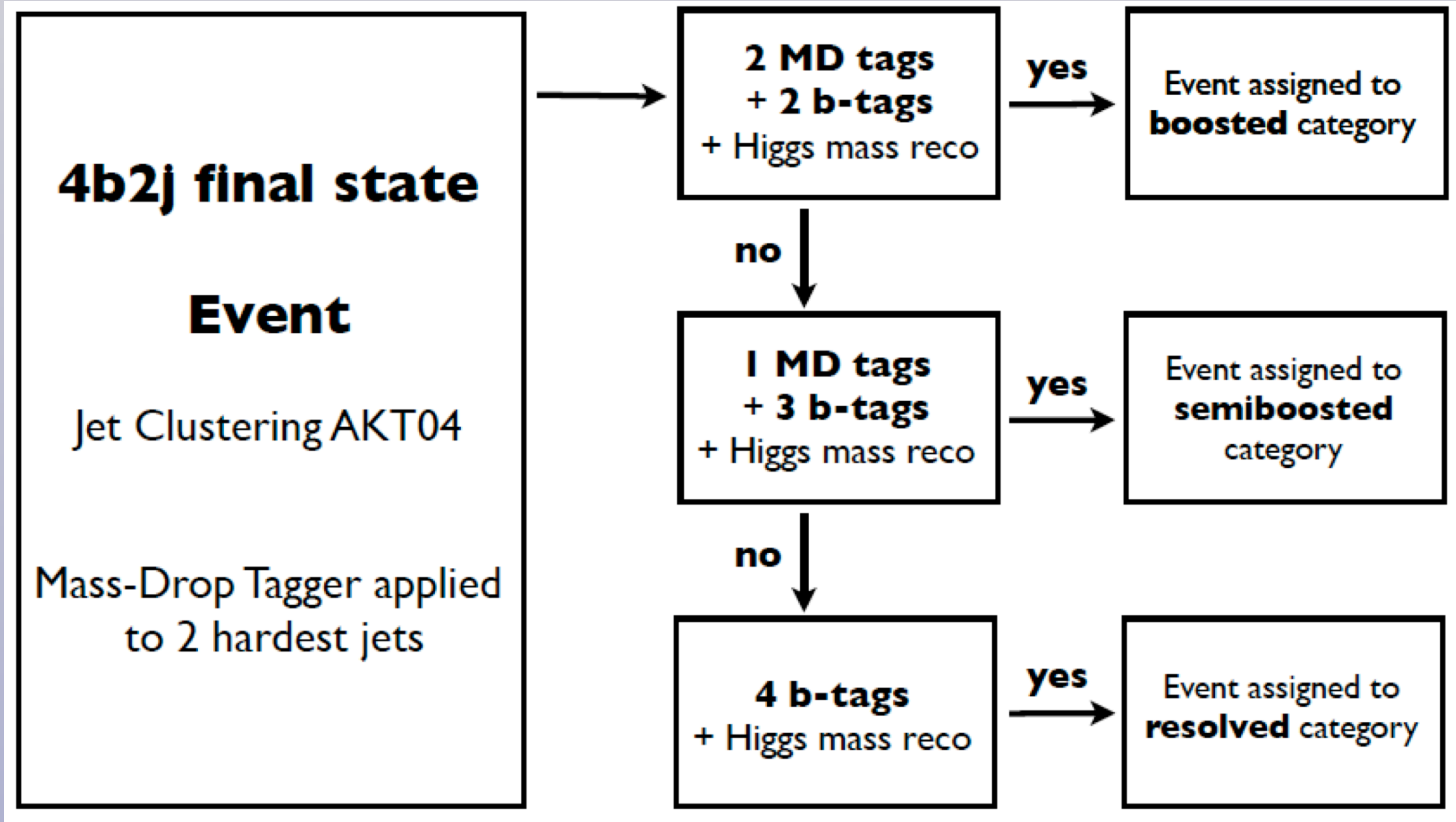
LHC 14 TeV	$\sigma(4b)$ (fb)	$N_{\rm ev}(3\,{\rm ab}^{-1})$
Standard Model	0.10	300
$c_V, c_{2V}, c_3 = 1.0, 0.0, 1.0$	2.45	7380
$c_V, c_{2V}, c_3 = 1.0, 2.0, 1.0$	1.59	4770
$c_V, c_{2V}, c_3 = 1.0, 1.0, 0.0$	0.29	870
$c_V, c_{2V}, c_3 = 1.0, 1.0, 3.0$	0.27	810
MCHM5 with $\xi = 0.3$ $c_V, c_{2V}, c_3 = 0.84, 0.40, 0.48$	0.41	1230

- Cross-sections for the **4b final state** after basic selection cuts
- Event rates very substantially enhanced when the **hhVV** coupling departs from its SM value: high sensitivity to **new strong BSM dynamics**
- Large increase in events rates when going up to **100 TeV**: greatly improved sensitivity

FCC 100 TeV	$\sigma(4b)$ (fb)	$N_{\rm ev}(10\,{\rm ab}^{-1})$
Standard Model	4.53	45.3K
$c_V, c_{2V}, c_3 = 1.0, 0.0, 1.0$	327	3.3M
$c_V, c_{2V}, c_3 = 1.0, 2.0, 1.0$	280	2.8M
$c_V, c_{2V}, c_3 = 1.0, 1.0, 0.0$	11.0	110K
$c_V, c_{2V}, c_3 = 1.0, 1.0, 3.0$	9.2	92K
MCHM5 with $\xi = 0.3$ $c_V, c_{2V}, c_3 = 0.84, 0.40, 0.48$	39	390K

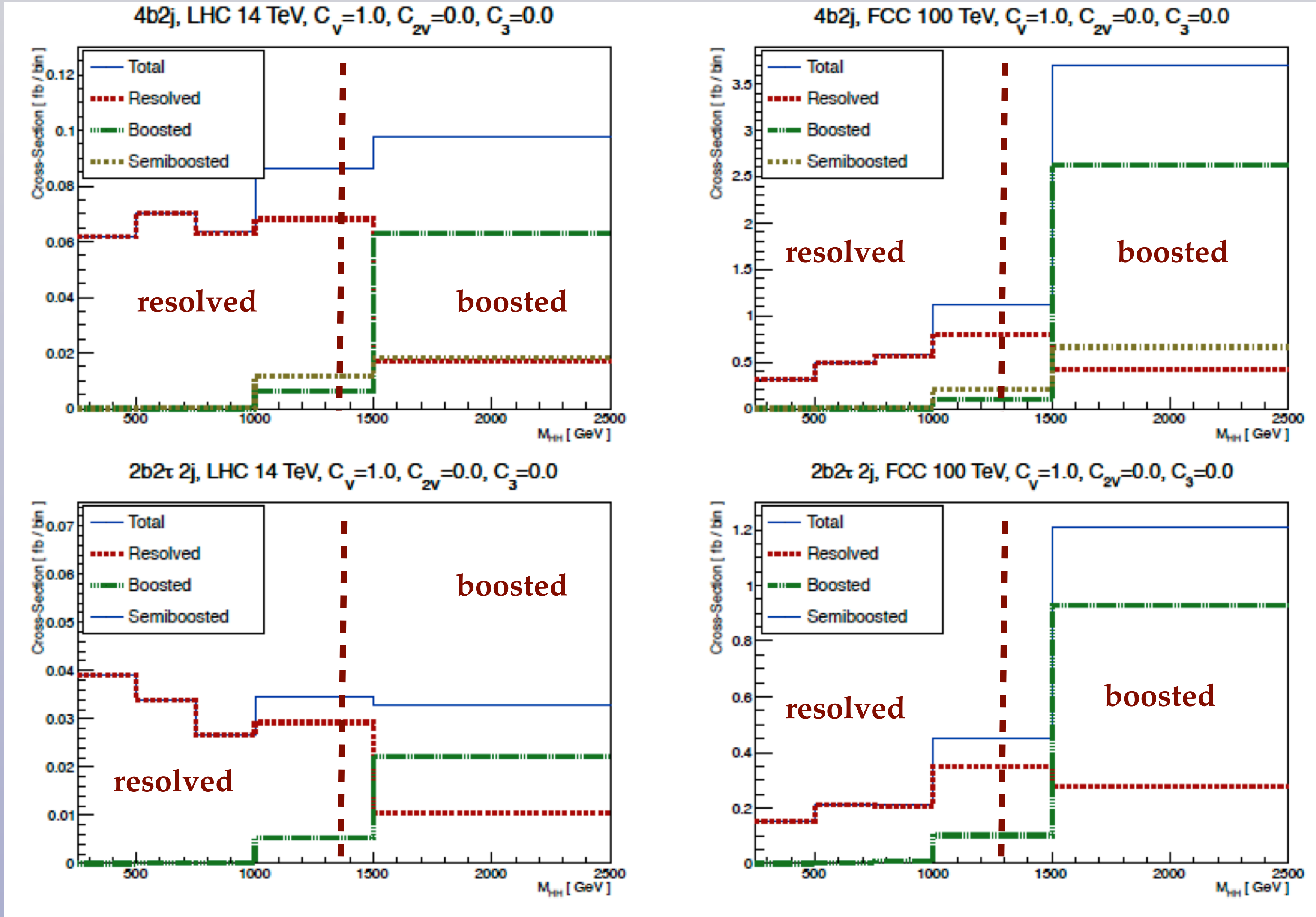
Analysis strategy

- To explore the complete range of SM and BSM scenarios we need to classify, on a event-by-event basis, **all possible signal topologies: boosted, semiboosted and resolved**
- This can be achieved using **scale-invariant resonance tagging**



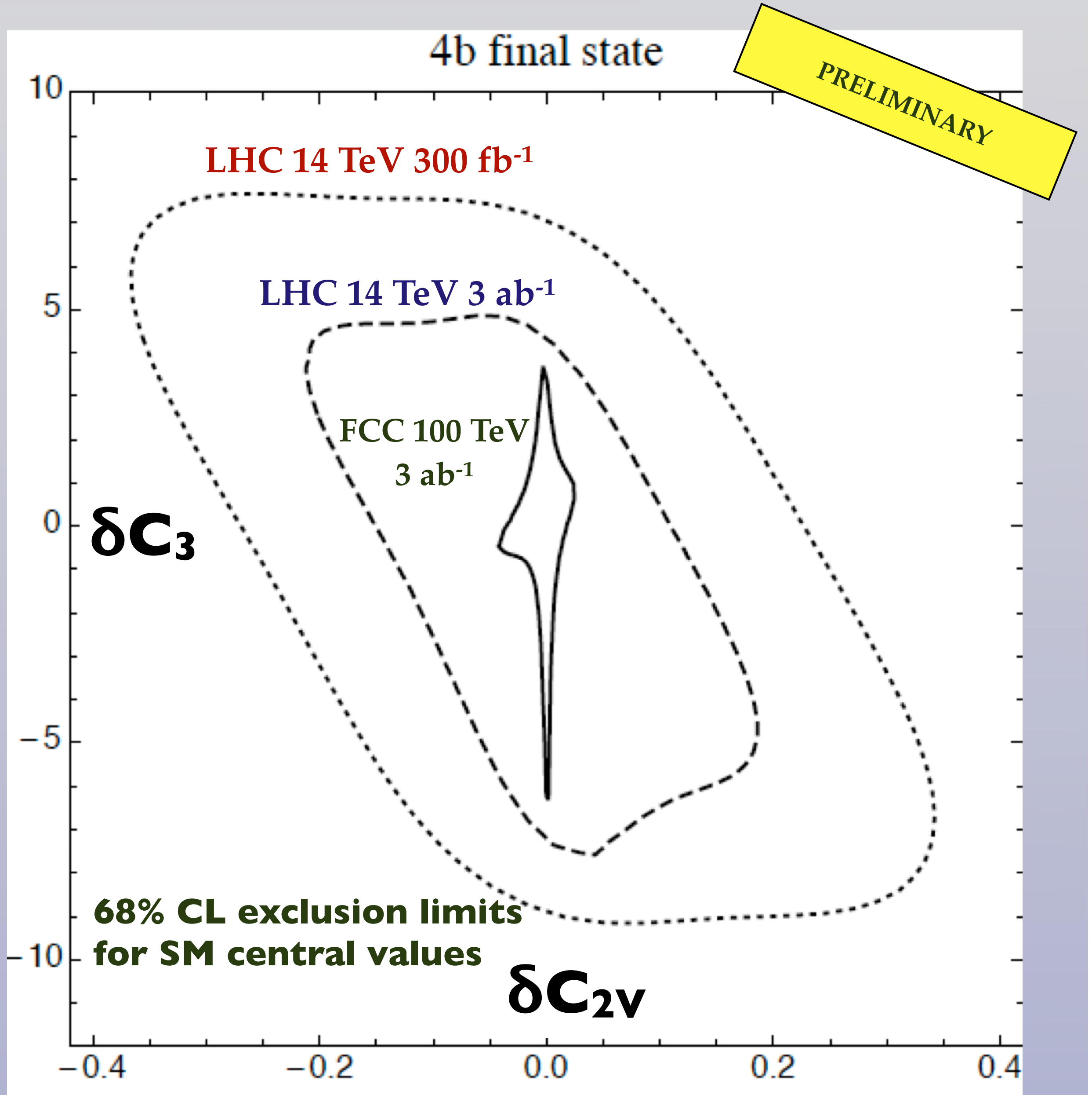
Analysis strategy

- For m_{hh} close to threshold, the **resolved contribution** dominates, while large m_{hh} is the **boosted regime**
- At the LHC, resolved and boosted configurations **similar,** while at the FCC the **boosted regime dominates**
- Boosted techniques crucial since **large m_{hh}** is the region more sensitive to **new strong BSM dynamics**



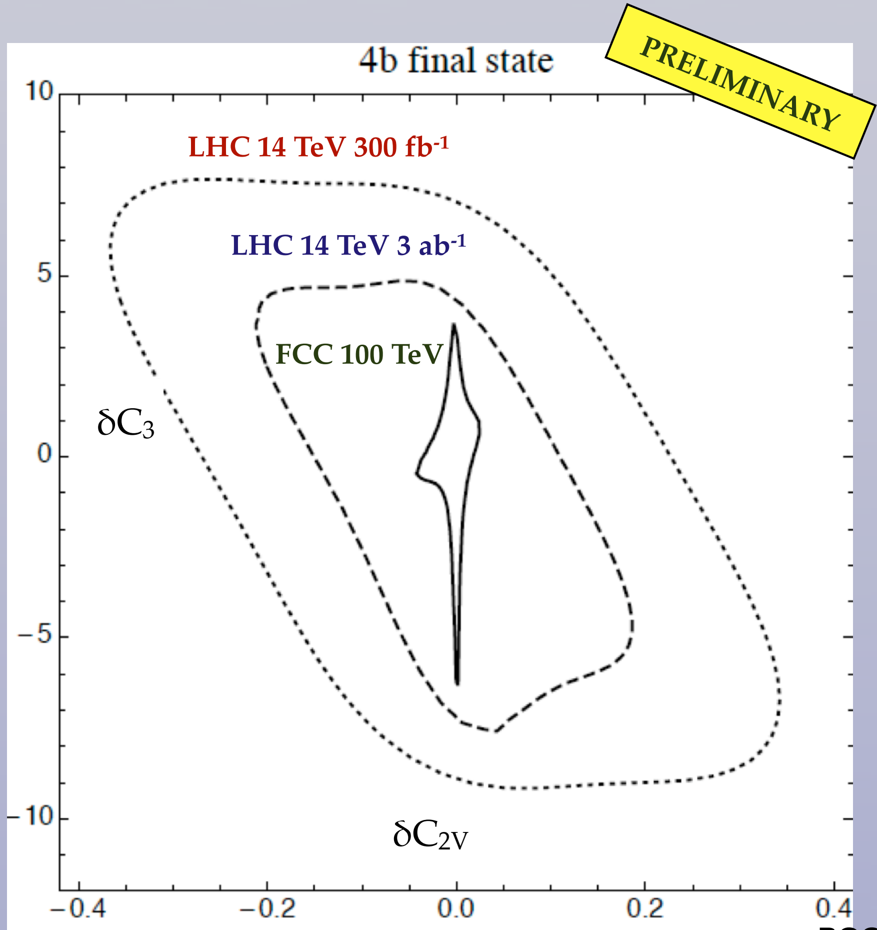
Preliminary Results

- In the **4b final state,** 14 TeV with 300 fb^{-1} (3000 fb^{-1}) the **hhVV** coupling can be measured with **good precision: ~25-30% (10-15%)**
- As expected, the precision on the **Higgs trilinear coupling** is worse than in **gg->hh** (since backgrounds dominate *hh* threshold region)
- At the **FCC,** the **hhVV** coupling can be pinned down with **very high, few percent precision**
- We have included a 50% error in the backgrounds, to account for theory and experimental uncertainties
- Encouraging to begin to explore **Higgs pair-production in VBF** already at the **LHC Run II!**



Summary and outlook: HH VBF

- **Higgs pair production** in the **vector-boson fusion channel** provides unique information of the mechanism underlying **electroweak symmetry breaking**
- Deviations from the SM value for the **hhVV** coupling induce large differences in event rates that grow strongly with $\mathbf{m_{hh}}$, where the di-Higgs system is **boosted** and **jet substructure techniques** are required
- Preliminary results indicate that in the **4b2j** final state we can probe at the **LHC deviations in C_{2V} as small as 25-30% (10-15%)** with **300 (3000) fb^{-1}**, while at the FCC we find **few-percent accuracy**
- Complementary constraints from the **2b2τ2j** final state, with smaller rates but with reduced backgrounds

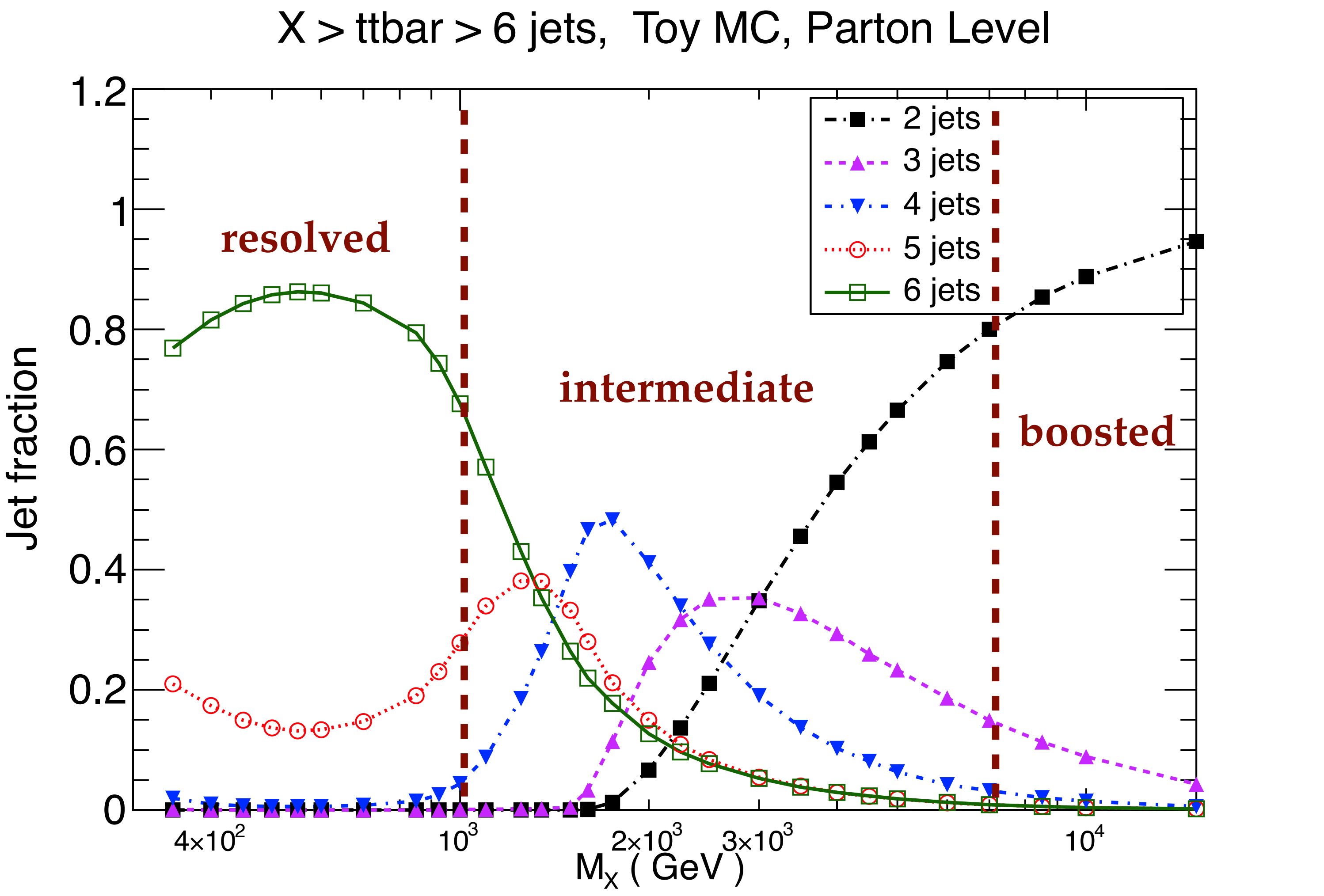


Summary and outlook: scale-invariant tagging

- **Top quark pair production** is widely used in BSM searches
- Typically searches are separated into the **boosted** and **fully resolved** regimes
- It would be desirable to **merge the two regimes** into a common analysis, while improving the overall efficiency by including the **intermediate regime** as well - **Work in progress**

Toy MC for heavy resonance X decaying to ttbar

Fully hadronic decays

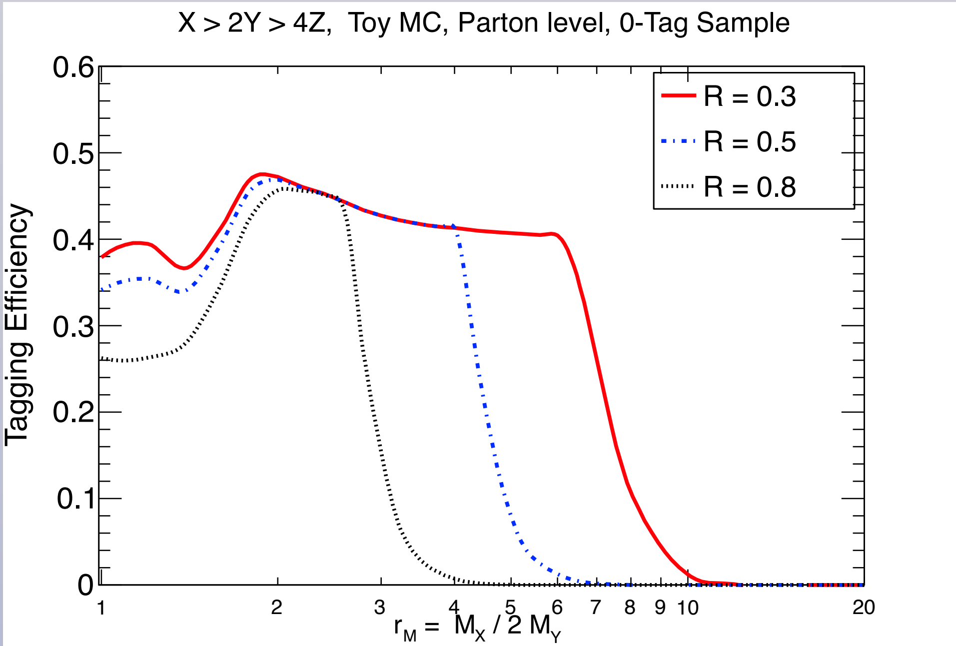


Extra Material

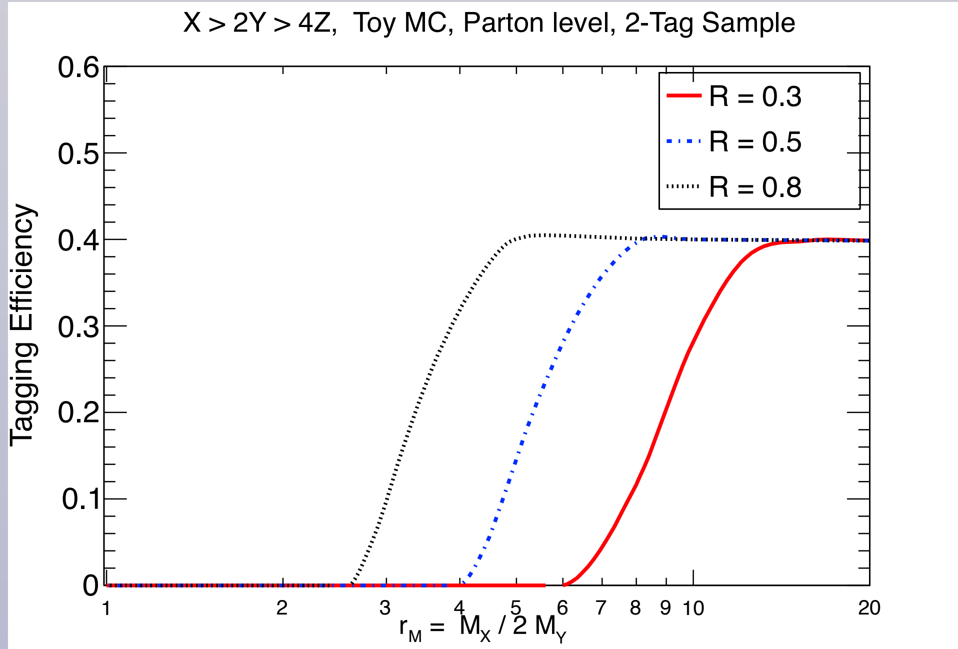
Scale-invariant tagging

- Tagging efficiency is also **independent of the value of jet radius**
- **The relative classification of the events** in the **resolved**, **boosted** and **intermediate** categories depends on **R,** but the total tagging efficiency is reasonably **R-independent**

Resolved



Boosted

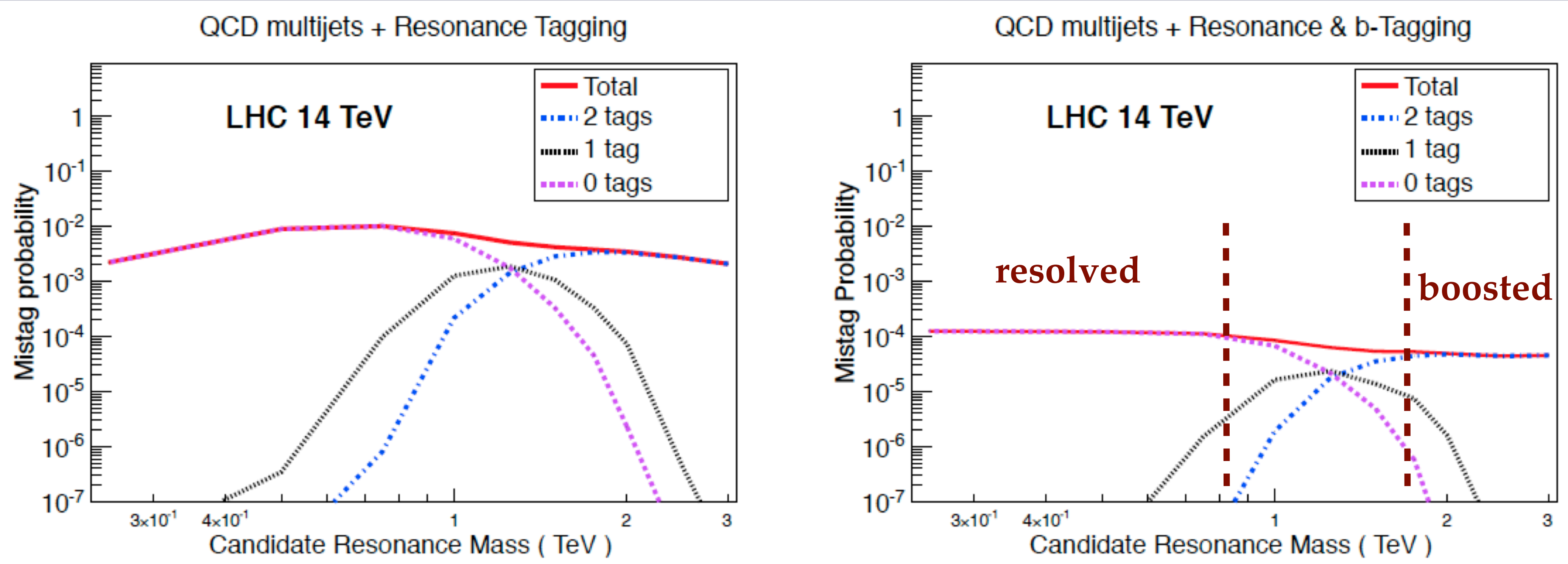


Scale-invariant tagging: with a **single analysis**, explore **simultaneously both the boosted and resolved regimes**, with a smooth interpolation for intermediate masses

Radius-independent tagging: Results are resilient against choice of **R**

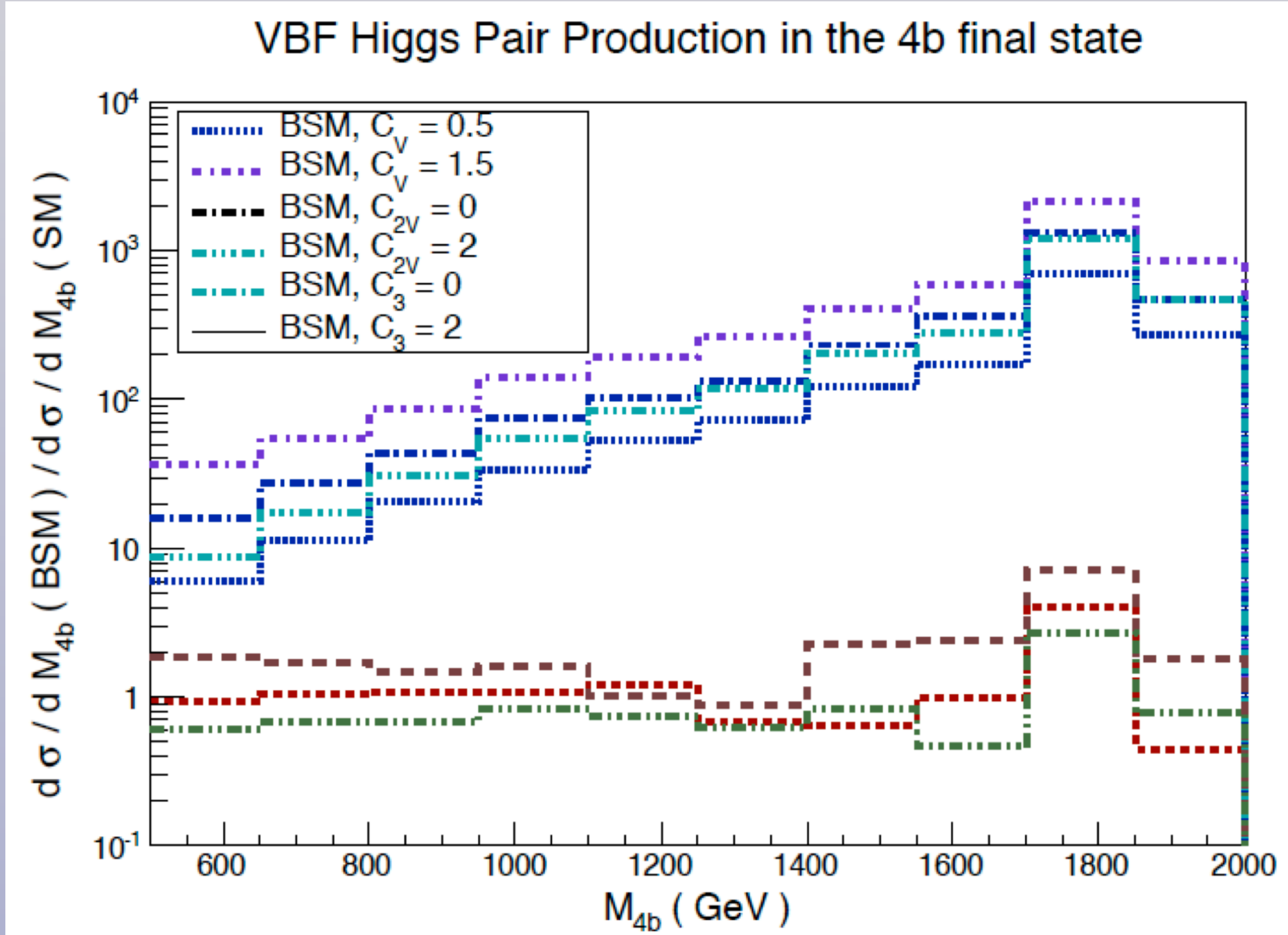
Background rejection rates

- As the tagging efficiency, the **background rejection rate is scale invariant**: 10^{-4} for all masses
- The flat background rejection rate arises from a **non-trivial combination** of the contributions from the *boosted, resolved* and *intermediate* **jet tagging categories**



Boosting the diHiggs final state

- The ratio of BSM/SM cross-sections **grows strongly** as a function of $\mathbf{M_{HH}}$
- Exploiting the **hardness of the $\mathbf{M_{HH}}$ distribution** in composite Higgs models is the key to tame the overwhelming QCD backgrounds, and requires the use of **boosted jet techniques**



Scale-invariant tagging