Some Title Involving $ZH \to llb\bar{b}$

A DISSERTATION PRESENTED
BY
STEPHEN K. CHAN
TO
THE DEPARTMENT OF PHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
PHYSICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
NOVEMBER 2017

Thesis advisor: Professor John Huth Stephen K. Chan

Some Title Involving $ZH \to llb\bar{b}$

Abstract

The Higgs looks more Standard Model by the day. The bulk of this thesis is vomiting up what amount to book reports of the main analysis documents, a technical paper, and a RI quality theory steak haché.

For a "unique intellectual contribution," I made three different BDT's and went shake and bake to what I'm sure will be a set of inconclusive results of dubious actual scientific value.

Contents

This is the dedication.

Acknowledgments

This thesis would not have been possible without large amounts of espresso.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

0

Introduction

Much has been said It has

[31] (henceforth refered to as the "fiducial analysis")

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

1

Theory

Much has been said about the so-called Standard Model of particle physics

Noli turbare circulos meos

Archimedes

2

The Large Hadron Collider and the ATLAS Detector

Look at[26], [19]

The CERN accelerator complex and its experiments stand as a testament to human inge-

nuity and

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

3

Object Definitions and Event Selection

Much has been said

3.1 Object Definitions

Detailed descriptions of object definitions may be found in [21]. Tables 3.2, 3.3, and 3.4 are reproduced from [31] and summarize electron, muon, and jet definitions, respectively.

3.2 Event Selection and Analysis Regions

This analysis focuses specifically on the 2-lepton channel of the fiducial analysis, with the event selection and analysis region definitions being identical. Common to all lepton channels in the fiducial analysis is the set of requirements on the jets in a given event. There must be at least two central jets and exactly two signal jets that have been "b-tagged" according to the MV2c10 algorithm[3], with at least one of these b-jets having a $p_T > 45$. For MVA training and certain background samples, a process known as "truth-tagging" is applied instead of the standard b-tagging to boost sample statistics and stabilize training/fits (cf.[31] Section 4.2 for details). After event selection, the *muon-in-jet* and *PtReco* corrections, described in[21] 6.3.3-4, are applied to the b-jets, and, taking advantage of the fact that the 2-lepton final state is closed, a kinematic fitter is applied (cf.[19]); these objects are only used for MVA training/fit inputs.

In addition to the common selections, there are 2-lepton specific selections. All events are required to pass an un-prescaled single lepton trigger, a full list of which may be found in Tables 5,6 of[21] with the requirement that one of the two selected leptons in the event must have fired the trigger. There must be 2 VH-loose leptons, and at least one of these must be a ZH-signal lepton. This lepton pair must have an invariant mass between 81 and 101 . In order to increase analysis sensitivity, the analysis is split into orthogonal regions based on the number of jets and the transverse momentum of the Z candidate (the vectoral sum of the lepton pair): 2 and $\geq$ 3 jets; p_T^V in $[75, 150)$, $[150, \infty)$. In addition to the signal regions where the leptons are required to be the same flavor (e or μ), there are top $e - \mu$ control regions used to constrain the top backgrounds.

Category	Requirement
Trigger	un-prescaled, single lepton
Jets	$\geq$ 2 central jets; 2 b-tagged signal jets, harder jet with $p_T > 45$
Leptons	2 VH-loose leptons ($\geq$ 1 ZH-signal lepton); same (opp) flavor for SR (CR)
$m_{\ell\ell}$	$m_{\ell\ell} \in (81, 101)$ GeV
$p_{\mathrm{T}}^{\mathrm{V}}$regions ()	$[75, 150)$, $[150, \infty)$

Table 3.1: Event selectrion requirements

Electron Selection		"η	ID	d_0^{sig}	$\lvert\Delta z_0^{BL} \sin\vartheta\rvert$	Isolation
$VH-loose$	>7 GeV	$\lvert\eta\rvert < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	LooseTrackOnly
$ZH-signal$	>27 GeV	$\lvert\eta\rvert < 2.47$	LH Loose + B-layer cut	< 5	< 0.5 mm	LooseTrackOnly
$WH-signal$	>27 GeV	$\lvert\eta\rvert < 2.47$	LH Tight	< 5	< 0.5 mm	FixedCutHighPtCaloOnly

Table 3.2: Electron selection requirements.

Muon Selection		"η	ID	d_0^{sig}	$\lvert\Delta z_0^{BL} \sin\vartheta\rvert$	Isolation
$VH-loose$	>7 GeV	$\lvert\eta\rvert < 2.7$	Loose quality	< 3	< 0.5 mm	LooseTrackOnly
$ZH-signal$	>27 GeV	$\lvert\eta\rvert < 2.5$	Loose quality	< 3	< 0.5 mm	LooseTrackOnly
$WH-signal$	>25 GeV	$\lvert\eta\rvert < 2.5$	Medium quality	< 3	< 0.5 mm	FixedCutHighPtTrackOnly

Table 3.3: Muon selection requirements.

Jet Category	Selection Requirements
Forward Jets	jet cleaning $p_{\mathrm{T}} > 30$ GeV $2.5 \leq \lvert\eta\rvert < 4.5$
Signal Jets	$p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$ jet cleaning JVT ≥ 0.59 if ($p_{\mathrm{T}} < 60$ GeV and $\lvert\eta\rvert < 2.4$)

Table 3.4: `AntiKt4EMTopoJets` selection requirements. The jet cleaning is applied via the JetCleaningTool, that removes events in regions corresponding to hot calorimeter cells.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

4

Signal and Bacgkround Modeling

This chapter summarizes the modeling of the dominant signal and background processes in this analysis, including corrections and systematic uncertainties (set in `this` font) related to each process.

ATL-COM-PHYS-2016-1724 section 4 and Main reference: ATL-COM-PHYS-2016-1747.pdf[33]

(we use the latter as a scaffold for this, basically plucking out stuff section by section)

Most of these studies (unless noted) are truth-level studies (particle level) done in Rivet [20] (F.4 for Rivet/reco comparison....not there)

4.1 Signal Processes

The dominant process is Higgstrahlung; ggF is ∼14%

$q\bar{q}$ Powheg with MiNLO (multiscale improved NLO generator Ref 1) applied as generator, Pythia 8 + AZNLO tune (Ref 3) + NNPDF3.0 PDF (Ref 4) set; $gg \rightarrow ZH$ Powheg + Pythia 8 (Ref 2)

Alternate samples: MadGraph 5(ref 26)_aMC@NLO+Pythia 8

Cross section: this is done NNLO in QCD and NLO in EW except for ggZH NLO+NNL (QCD). WH normalized to values in the table; ZH: total as 0.88, ggZH as 0.12, and then qqZH as total - ggZH (refs 15-18)

Process	σ (pb)
WH	1.37 ± 0.04
W^+H	0.84
W^-H	0.53
ZH	$0.88^{+0.04}_{-0.03}$
$gg \rightarrow ZH$	0.12
$qq \rightarrow ZH$	0.76

Table 4.1: Summary of inclusive cross sections for signal processes.

NLO EWK correction: same as Run 1; they use HAWK to calculate a differential cross section as a function of pTV (take their Figure 3) for a correction factor of $k_{EW}^{NLO}(p_T^V) = 1 + \delta_{EW}$; qqVH only.

N(N)LO EWK systematic: $\Delta_{EW} = \max\{1\%, \delta^2_{EW}, \Delta \quad\}$, δ_{EW} from above correction, $\Delta \quad$ is γ induced cross section uncertainty to the total [WZ]H xsec

Overall signal acceptance uncertainties: cross section and branching ratio (LHC Higgs WG;[28],[32])

- ATLAS_BR_bb (1.7%)
- ATLAS_QCDscale_(VH|ggZH) (0.7%, 27%) vary $\mu_{R,F}$ for renorm/factorization scale by 1/3 to 3 of original value
  - to get ggZH; assume QCD scale σ same for qq[WZ]H,; assume ref 20 inclusive ZH production and take diff in quadrature of inc. and qqZH
- ATLAS_pdf_Higgs_(V[WZ]H|ggZH) (1.9%, 1.6%, 5.0%) (also α_s, 68%CL on PDF4LHC15_nnlo_mc PDF set)
  - qqWH is bigger here than ZH; get ggZH from 19, qqZH from 20 assuming ggZH small, so overall ZH is qqVH

Analysis specific: analysis category acceptances; pTV, mBB shape

- PS/UE (Table 4)
  - MadGraph vs. A14 varied (tunes); nominal Powheg/Minlo/Pythia8 vs Powheg+minlo+Herwig7 (PS)
  - Now vary up and down in each nLep x nJet bin and save as a ratio wrt nominal
  - $\sum_{tunes} \max_{tune} \left(\left|R_{up} - R_{down}\right|\right) \oplus \sigma_{PS}$ (ATLAS_UEPS_VH_hbb)
  - Now add a 2/3 jet ratio systematic (i.e. (2/3 acceptane ratio nominal)/(2/3 ratio alternative)) (ATLAS_UEPS_VH_hbb_32JR); combine in same way
  - pTV (mBB) shape: linear (quadratic); fit up and down for each variation; 2/3jet separate for mBB; use histogran with largest deviation as shape (ATLAS_UEPS_VH_hbb_(VPT|MBB))
    * shape only, except for L2 pTV (shape+norm)

  - ggZH same as qqZH and correlated

- Scale variations (Table 5)
  - Vary μ_R, μ_F (probably the 1/2 to 2 scheme in steps with no more than blah blah)
    * Stewart Tackmann for nJet bins (QCDscale_VH_ANA_hbb_J[23]; both for 2jet)
    * JVeto for L[01] (3 jet exclusive)
  - Same pTV, mBB NP scheme for nLep/nJet as for UEPS
  - ggZH same as qqZH and de-correlated (Run 1 says they're difference)

- PDF+α_s
  - Powheg/Minlo/Pythia8 v. PDF4LHC15_30 PDF set; reco-level distributions (all others use Rivet, which doesn't like a lot of weight variations)
  - PDF: quad sum of variations of PDF uncertainties (go through the set? no probably the same way as above)
  - α_s: average of variations from altering α_s
  - pdf_HIGGS_VH_ANA_hbb, pdf_VH_ANA_hbb_(VPT|MBB)
  - Same pTV, mBB NP scheme for nLep/nJet as for UEPS

4.2 Background

Main backgrounds are V+jet, ttbar, VV, single top (, and multijet in 1lep)

4.2.1 V+jet

cf.[18] for details of MC generation

- Sherpa 2.2.1@NLO[27] for matrix element (ME) and PS tuning (Tables 7–10)
  - ME's for up to 2 (3–4) partons at NLO (LO); for more, use showering (Sherpa's own UEPS)
  - "The merging of different parton multiplicities is achieved through a matching scheme based on the CKKW-L [24] [25] merging technique using a merging scale of Qcut = 20 GeV"[30][29]
  - 5 quark flavors mass(less) quarks in the shower (ME)
  - max (H_T, P_T^V) slices: [0–70, 70–140, 140–280, 280–500, 500–1000, >1000] GeV
  - Slices in [CB](Veto|Filter) for flavors
    - BFilter: at least 1 b-hadron with $|\eta| < 4, p_T > 0$ GeV
    - CFilterBVeto: at least 1 c-hadron with $|\eta| < 3, p_T > 4$ GeV; veto events which pass the BFilter
    - CVetoBVeto veto events which pass the BFilter or the CFilterBVeto
  - Variations of $\mu_{R,F}$ at 0.5, 2; PDF variation for MMHT2014nnlo68cl and CT14nnlo
  - Sherpa 2.1 for resummation scale at 0.5, 2; CKKW 15, 20 GeV
- Alternate samples use MadGraph5+Pythia8 (UEPS)
  - LO QCD ME's, merging parton multiplicities up to 4 (for more, use PS), NNPDF2.3 LO PDFs; A14 tune (ATLAS)
  - CKKW-L scheme with a merging scale of Qcut = 30 GeV.
  - 5 flavor scheme

- Cross section *k*-factors: our generators are NLO, but V production is known to NNLO—add factors to rescale
  - Take total events, average over lepton flavors for filter efficiencies, and compare to NNLO (ref 27)
  - For L2, there's a 40 GeV generator mLL cut, but the NNLO calcu is done in (66,116) GeV, so another scale
  - For L0, take L2 (since NNLO not calc) and correct for $BR(Z \to \nu\nu) / BR(Z \to \ell\ell)$, consider with no mass cuts, remove "Z/γ^* interference"
  - Differences between nominal and alternative MC's can be explained to higher order BR's and EW schemes w.r.t. PDG recommendations

Anyway, V+jet is broken up into V+hf (V+b*, V+cc), V+cl, V+l(ight)

- Relative acceptance between regions
  - Understand correlation between/among regions (you can float these normalizaitons in the fit to fix your understanding of things using more ifnrmation)
  - 2jet vs 3(p)jet for L[01](2), L0 vs. L2 (Z+hf), L0 vs. L1 (W+hf), WCR vs. SR for L1 (W+hf)
  - These norm's are RooGaussian's with priors from MC studies (Rivet, Appendix A[33])
  - Their uncertainties are double ratios between regions and then MC's with components...
    * Envelope of varying μ_R, μ_F in Sherpa
    * 0.5 $\sum_{\oplus}$ (up-down on CKKW, merging scale variation; weird because done with Sherpa 2.1, so no central value comparison)
    * max variation between nominal/alt PDF reweighting
    * diff btw Sherpa/MadGraph

- pTV, mBB shape uncertainties: data driven and MC techniques—you normalize distributions to the same area, compare, then do functional fits, then pick the biggest one and symmetrize
- W+jets
  - Normalization/acceptance systs (Table 13): `Sys(Wcl|Wl)Norm` (one for all regions is fine since $b-$tagging suppresses), a floating `norm_Wbb`, `SysWbbNorm_(J3|DWhfCR_L1|L0)`; J3 is 3-to-2 jet; DWhfCR_L1 is CR-SR; L0 is L0-L1
  - Flavor composition (Tables 14, 15): W+hf breakdown; `Sys(Wbc|Wbl|Wcc)WbbRatio`
  - pTV: a linear `SysWPtV`, which happens to be Serpa 2.2.1 v. MadGraph in all regions (largest variation)
  - mBB: a linear `SysWMbb`, which happens to be Serpa 2.2.1 v. MadGraph in all regions (largest variation) (not a typo; it's the same as pTV)
- Z+jets: L[02] SR only (topemucr is pretty pure; not really in L1)
  - Normalization/acceptance (Table 16): `Sys(Zcl|Zl)Norm` (one for all regions is fine since $b-$tagging suppresses; less than 1% here), a floating `norm_Zbb`, `SysWbbNorm_(L2_J3|J3|0L))`; L2_J3, J3 is 3-to-2 jet (L2 correlates lo/hi pTV; L0 is separate because of selection differences); 0L is 0 to 2 lepton (hi pTV only)
  - Flavor composition (Tables 17): Z+hf breakdown; `Sys(Zbc|Zbl|Zcc)ZbbRatio`—norm uncertainties with diff priors in L0, L2-2jet, L2-3pjet; Sherpa 2.2.1 v MG main diff
  - L2 CR: METHT < 3.5, [012]-tag, 2 and 3pjet, no mJJ in (110,140) GeV for 2tag, pTV regions; subtract off non Z+jet and then scale MC to data
  - pTV: shape+norm, fit to data in L2 CR; $\pm 0.2 \log_{10}\left(p_T^V/500\ \text{GeV}\right)$
  - mBB: shape only, fit to data in L2 CR; $\pm 0.0005 \log_{10}\left(m_{jj} - 100\ \text{GeV}\right)$

4.2.2 Top-Pair Production

MC production—h_{damp} is transverse momentum scale at which Sudakov resummation becomes unimportant: smaller damp means higher suppression (cf. Table 20)

- Powheg+Pythia8
  - Powheg: NNPDF3.0 (NLO) for ME (Powheg); $h_{damp} = 1.5 m_{top}$ (resummation damping factor for ME/PS matching; controls high pT rad)
  - Pythia: PS,UE,had; v 8.210, A14 PDF set, NNPDF2.3 LO for PS; pTdef=2, pThard=0 control Powheg/Pythia8 merging thorugh shower vetoing
  - $\sigma_{t\bar{t}}\left(m_{top} = 172.5\,\text{GeV}\right) = 831.76^{+40}_{-46}$ pb: NNLO QCD; NNLL soft gluon terms;
    - QCD scale variations: $^{+19.77}_{-29.20}$ pb; PDF: ± 35.06 pb: "The e PDF and α_s uncertainties were calculated using the PDF4LHC prescription [8] with the MSTW2008 68
    - 3.3 times higher than 8 TeV

- Powheg+Herwig7: different PS. UE. had, MPI; H7UE tune
- MadGraph 5_aMC@NLO+Pythia 8.2: different hard scatter (i.e. ME)
- Powheg+Pythia8 low radiation sample (double $\mu_{R,F}$; h_{damp}, pTdef, pThard same; A14 tune Var3c Down variation used)
- Powheg+Pythia8 high radiation sample (halve $\mu_{R,F}$; pTdef, pThard same; $h_{damp} = 3 m_{top}$ (doubled) A14 tune Var3c Up variation used)

Systematics—Rivet

- Powheg+Pythia8
  -

4.3 Notes

Notes from Kevin's thesis: Signal:

- *pTV NLOEWK* The signal processes have some pTV dependence at next to leading order (NLO) due to electroweak corrections
- *TheoryQCDScale, TheoryPDF* for renormalization/scale uncertainties, PDF uncertainties
- *TheoryAcc_J[23]* Stewart-Tackmann stuff
- *TheoryAccPDF* do acceptance calculations with different PDF's
- *TheoryVPtQCD* this is one of those functional things—probably different in Run2; linear of pTV

Background

- *ZDPhi* $\Delta\varphi\,(b_1, b_2)$ mismodeling; shape—another linear of dphi; a correction and the correction is a systematic for each event
- *ZPtV* $\Delta\varphi\,(b_1, b_2)$ mismodeling; const+log and half the correction is a systematic for each event
- *Z+jet Normalizations* broken down by flavor region; both Norm's and Ratio between regions
- *ZMbb* const(mbb e-3 -c const); systematic
- *ttbar* pT, (2/3 jet ratio across generators), mBB
- *VV* NLO xsec, *s*/PDF's, mJJ

4.3.1 Stewart-Tackmann

A way to calculate uncertainties on processes in different nJet bins[34]. Generically:

$$\sigma_{\geq N} = \sigma_N + \sigma_{\geq N+1} \tag{4.1}$$

There's some quantity that you make a cutoff in an integral that defines the border between jet regions.

$$\sigma_{\geq N} = \int_0^{p_{cut}} \frac{d\sigma_N}{dp} + \int_{p_{cut}} \frac{d\sigma_{\geq N+1}}{dp} \tag{4.2}$$

So for some fucking reason, inclusive cross sections are easier to calculate, so you can just vary α_S in the usual way for those and treat the two inclusive cross sections. Anywho, we assume the inclusive uncertainties are uncorrelated, for a covariance matrix for $\{\sigma_{\geq N}, \sigma_N, \sigma_{\geq N+1}\}$ of:

$$\Sigma = \begin{pmatrix} \Delta^2_{\geq N} & \Delta^2_{\geq N} & 0 \\ \Delta^2_{\geq N} & \Delta^2_{\geq N} + \Delta^2_{\geq N+1} & -\Delta^2_{\geq N+1} \\ 0 & -\Delta^2_{\geq N+1} & \Delta^2_{\geq N+1} \end{pmatrix} \tag{4.3}$$

The main idea is that you have Sudakov double logs of p/Q, where $Q = m_H$ or whatever scale your hard process occurs at, and p_{cut} is usually something like a p_T cutoff. Now, the $N+1$ term in that matrix is actually some uncertainty associated with your cutoff, but your double logs will dominate your higher order terms...the paper has this reasoning:

"In the limit $\alpha_s L^2 \approx 1$, the fixed-order perturbative expansion breaks down and the logarithmic

terms must be resummed to all orders in α_s to obtain a meaningful result. For typical experimental values of pcut fixed-order perturbation theory can still be considered, but the logarithms cause large corrections at each order and dominate the series. This means varying the scale in α_s in Eq. (9) directly tracks the size of the large logarithms and therefore allows one to get some estimate of the size of missing higher-order terms caused by p_{cut}, that correspond to Δ_{cut}. Therefore, we can approximate $\Delta_{cut} = \Delta_{\geq 1}$, where $\Delta_{\geq 1}$ is obtained from the scale variation for $\sigma_{\geq 1}$."

They use the example of ggF Higgs production with $\{\sigma_{total}, \sigma_0, \sigma_{\geq 1}\}$ and say this works to all N for all processes, provided one picks $\mu \approx Q$ so you can use perturbative expansions.

Anyway, the upshot is this: we've got 2 and 3 jet bins. For 2 jet TheoryAcc_J2 and TheoryAcc_J3; 3 jet has TheoryAcc_J3, which is anti-correlated with the 2 jet J3 term

4.3.2 CKKW-L

When you're looking to generate MC events, there are two main event generators. There are the parton shower event generators (PSEG), like Pythia, and the matrix element generators (MEG) like MadGraph or Powheg, both of which have nice and not-so-nice features. If we follow [29], section 2, we get a nice illustration. Sherpa does both and stitches things together for you.

So PSEG's have the nice feature that you don't get nasty infinities. You start with some primary hard scatter (say $e^+e^- \rightarrow q\bar{q}$) and then let your incoming and outgoing partons cascade via iterative $1 \rightarrow 2$ branching. You order the emissions by some "evolution scale ϱ," starting at ϱ_0 and descreasing unitl you reach some pre-determined cutoff ϱ_c (usually to match some model) to generate $0, 1 \ldots n$ extra partons, there are exlusive cross sections involving well-ordered, intermediate scales ϱ_i,

some phase space variables (like momentum fractions z_i) denoted Ω_i, probabilities of non-emission between scales in the form of Sudakov form factors $\Delta_{S_i}\left(\varrho_i, \varrho_{i+1}\right)$, coefficients c_{nn}^{PS} associated with splitting functions that depend on ϱ_i, Ω_i and sum over flavors, blah blah.

The Δ's looke like:

$$\Delta_S\left(\varrho_i, \varrho_{i+1}\right) = \exp\left(-\int_{i+1}^{i} \frac{d\varrho}{\varrho}\alpha_s\left(\varrho\right)\int dz\, P\left(z\right)\right) \tag{4.4}$$

and these can be written as a perturbative series in α_s ("duh")

$$\begin{aligned}
\sigma_{+0} &= \sigma_0 \Delta_{S_0}\left(\varrho_0, \varrho_c\right) \qquad (4.5)\\
\sigma_{+n} &= \sigma_0 c_{nn}^{PS} \Delta_{S_n}\left(\varrho_0, \varrho_c\right)\prod_{i=1}^{n}\alpha_s(\varrho_i \Delta_{S_i-1}\left(\varrho_{i-1}, \varrho_i\right) d\varrho_i d\Omega_i
\end{aligned}$$

(4.7)

$$\sigma_{+n} = \sigma_0 c_{nn}^{PS}\left(1 + c_{n,n+1}^{PS}\alpha_s + c_{n,n+2}^{PS}\alpha_s^2 + \ldots\right)\prod_{i=1}^{n} d\varrho_i d\Omega_i \tag{4.8}$$

Now, these c_{ij}^{PS} blow up in the soft/collinear limit of $\varrho_c \to 0$, but a resummation in all order for the Δ's gives a finite result for each cross section. Moreover, $\sum_0^\infty \sigma_{+i} = \sigma_0$. *The problem is that for several hard partons, this description only makes sense for strict ordering (the intermediate states) of hard partons because of the splitting funciton dependent coefficients.*

For MEG's, the picture is simpler because we use tree-level matrix elements for each parton final state. However, the cross-sections are *inclusive* (so each of these is at least n jets), and these all blow up in the soft/collinear regime, where the resummation gets nasty. The authors note that you can make PSEG's look like the MEG for the first emission ($c_{\text{II}}^{PS} \to c_{\text{II}}^{ME}$).

$$\sigma_{+0} = \sigma_0 \tag{4.9}$$

$$\sigma_{+n} = \sigma_0 \alpha_s^n c_{nn}^{ME} \prod_{i=1}^{n} d\Omega_i \tag{4.10}$$

So what to do? "...the solution should be obvious." Just use the MEG to generate your partons over some Q_{cut}, reweight the generated states with the Sudakov form factors, and use the PSEG to make parton showers for these final state objects so that the showers make everything under Q_{cut}. But those Sudakov scales need an ordered set of emission scales since all the diagrams are added together.

How does one set up an ordered set of scales? You can use the $k_\perp-$algorithm (takes pairs based on something like p_T (??)); use those scales as arguments to α_s; use $k_\perp-$algorithm resolution as a cut-off. This approach is good to NLL but has some discontinuities. Anyway, $k_\perp$ is basically the same thing as k_t clustering for jets[22]. Actually, it *is* the same exact thing for lepton colliders, so they use the angle between particles times a minimum square energy instead of ΔR and define beam jets...they also don't have the minimum distance built in, so there's a d_{cut}, which can be the square energy or some other thing; you can define it by the resolution in y you want by $y_{cut} = Q_0/d_{cut}$). Remember, k_t starts with your softest stuff and clusters upwards from there. For the resolution variable, remem-

ber that you have some characteristic distance after which things. Blah blah, so you pre-cluster (their topocluster type stuff for hadronic deposits based on the min $E_{T,i}^2, E_{T,j}^2\ \vartheta_{ij}^2$ metric) until all distances remaining are bigger than d_{cut}. Now define $y_{cut} = Q_o^2/d_{cut}$ and use $y_{kl} = d_{kl}/d_{cut}$ and cluster until all bigger than y_{cut}. *The important thing from this mess is just that y_{cut} is the resolution mentioned above; you don't have this mess with our usual algorithms because distances come in with $\Delta R^2/R_{alg}^2$, so if distances remaining are bigger, the plain "beam distance" keeps things unclustered.*

Dipole Cascade Model

You can also use the dipole cascade model ($2 \to 3$ where the 2 partons are a color dipole). The dipoles mean you don't have to do an angular ordering for the partons? Your ϱ is $p_\perp^2 = \frac{s_{12}s_{23}}{s_{123}}$ where s's are invariant masses of the combinations. There's also a rapidity associated with the $p_\perp$'s: $y = \frac{1}{2}\ln\left(\frac{s_{12}}{s_{23}}\right.$. The emission probability depends on splitting functions, which in turn depend on the parton pair type (parton 2 is the emitted one in this convention), where $x_i = 2E_i/\sqrt{s_{123}}$:

$$D_{q\bar{q}}\left(p_\perp^2, y\right) = \frac{2}{3\pi}\frac{x_1^2 + x_3^2}{(1-x_1)(1-x_3)} \tag{4.11}$$

$$D_{qg}\left(p_\perp^2, y\right) = \frac{3}{4\pi}\frac{x_1^2 + x_3^3}{(1-x_1)(1-x_3)} \tag{4.12}$$

$$D_{gg}\left(p_\perp^2, y\right) = \frac{3}{4\pi}\frac{x_1^3 + x_3^3}{(1-x_1)(1-x_3)} \tag{4.13}$$

$$\tag{4.14}$$

Finally, we get the probability:

$$dP\left(p_\perp^2, y\right) = \alpha_s(p_\perp^2) D_{ij}\left(p_\perp^2, y\right) \exp\left(-\int_{p_\perp^2} \frac{{p'_\perp}^2}{{p'_\perp}^2} \int dy' \alpha_s\left({p'_\perp}^2\right) D_{ij}\left({p'_\perp}^2, y'\right)\right) \frac{dp_\perp^2}{p_\perp^2} dy \quad (4.15)$$

(notice your old exp friend, the Sudakov). Also, hey, look, your intermediate partons are on shell, unlike in a $1 \to 2$ cascade sicne your dipole asorbs recoil, and your inverse cascade is a well-behaved "jet clustering" algorithm. But $g \to q\bar{q}$ has to be done by hand. Basically, you use this cascade/shower on your MEG partons to get scales that you reweight by $\prod_i \alpha_s\left(p_{\perp i}\right) / \alpha_s\left(p_{\perp c}\right)^n$ for some n.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

5

Experimental Systematic Uncertainties

ATL-COM-PHYS-2016-1724 section 6 points to 1674 (section 11)

Much has been said about the so-called Standard Model of particle physics

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

6

Multivariate Analysis

Much has been said

6.1 Training Samples and Variables

A subset of samples described in ?? was used for multivariate analysis (MVA) training, with $qqZH \to \ell\ell b\bar{b}$ and $ggZH \to \ell\ell b\bar{b}$ used as signal samples and Z+jets, $t\bar{t}$, and VV used as background samples.

Truth tagging TRF is used on all samples to improve training statistics and stability.

The standard set of variables taken as a baseline is the same as used in the fiducial analysis. The variables fall into several main categories: energy/momentum scales of composite objects (m_{bb}, m_{bbj}, p_T^V, $m_{\ell\ell}$), angles ($\Delta R\,(b_1, b_2)$, $\Delta\phi\,(\ell_1 + \ell_2, b_1 + b_2)$, $\Delta\eta\,(\ell_1 + \ell_2, b_1 + b_2)$), transverse momenta of the jets in the event ($p_T^{b_1}$, $p_T^{b_2}$, $p_T^{j_3}$), and E_T^{miss}. The kinematic fitter is used to derive the invariant masses of both the b-quark and lepton pairs (m_{bb}, $m_{\ell\ell}$).*

In addition to the standard set of variables used for MVA training, two additional sets of variables were used: the Lorentz Invariants (LI) and RestFrames (RF) inspired variable sets. These will be discussed below.

6.2 Lorentz Invariants

The LI variables, first put forth by S. Hagebock and others, are based upon the fact that the four-vectors of an event are determined, all of the information in an event are encoded into 16 quantities: the ten inner products of the four vectors, the three Euler angles, and the three parameters specifying the boost of the ZH system. The masses of the four final state objects are not considered very useful and so can be removed to leave six meaningful inner products (the ${}_4C_2$ combinations between distinct final state four vectors). Since these inner products can have an ill-defined physical interpretation and in order to help MVA training, each inner product is scaled by:

$$x \to \frac{x}{x + c} \tag{6.1}$$

*All other quantities use their `CxAODFramework` defaults.

where c is the mean of the distribution in the signal MC distribution. These are denoted `xi_yj`, where `x` and `y` are either `j` (for jet) or `l` (for lepton) and the indices are either 0 (1) for the leading (subleading) object by p_T in the event.

The number of useful angles can be reduced by recognizing some symmetries inherent in the final state. The symmetry around the beam axis eliminates one angle, while the boost being mostly in the beam direction (z) marginalizes the utility of the boost angles. This leaves the boost in the z direction, denoted `gamma_ZHz`, and two angles chosen to be the angle between the $b\bar{b}$ system and the beam (`angle_bb_z`) and the angle between $(b_1 + b_2)\times\hat{z}$ and $(b_1 + b_2)\times(l_1 + l_2)$ (`angle_bbz_bbll`).

An important advantage of the LI variable set is that all of the variables are in it are orthogonal in the signal case by construction. A drawback of this framework in a completely closed final state is that there is no way to treat E_T^{miss} in a Lorentz invariant way. In practice, however, it is found that the E_T^{miss} has correlations comparable to other correlations between variables (cf. Figure 6.2 (b) and (e)). There is also no prescription for any additional jets in the event beyond the two b-tagged jets. They are simply ignored in these variable calculations.

6.3 RestFrames Variables

The RestFrames variables[3], calculated using the software package of the same name, is based upon the concept that the most natural frame in which to analyze objects of the signal decay tree is in their individual production (rest) frames. The signal decay tree for Higgs production in association with a Z boson is show in Figure 6.1. Generally, one does not typically have enough information to

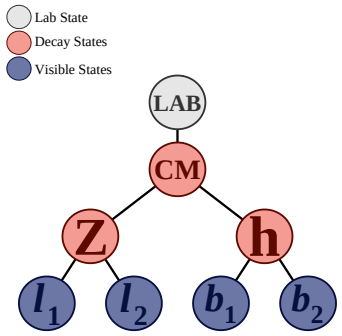


Figure 6.1: The $ZH \to \ell\ell b\bar{b}$ decay tree.

determine exactly each of the object production frames and the boosts between the frames[†], but in a completely closed final state like $ZH \to \ell\ell b\bar{b}$, this can be done in the usual way by adding the four vectors of the final state objects and solving the usual equations.

Each frame has associated with it the boost from its immediate parent and a mass scale; that mass and the angles between the Euclidean three vector associated with boost and the axis of the decay products provide useful variables. In general, the polar angle (typically given as a cosine) is considered more useful than the azimuthal angle (typically just a $\Delta\varphi$), though there are exceptions. The Z frame, for example, has `MZ`, which is just the usual $m_{\ell\ell}$, `cosZ`, the cosine of the polar angle between the lepton momentum direction in their production frame and the boost from the ZH center of mass (CM) frame, and the angle `dphiCMZ`.

In addition to the variables attached to individual object rest frames, energy scales associated with the CM frame can be used to contextualize other event level quantities. In particular, one can use the mass of the CM frame as a natural scale to evaluate the momentum of the CM frame, and the p_T

[†]There are ways to do this for generic decay trees, though, and this is the focus of much of[2] and the functionality of the RestFrames package.

of the CM frame as a natural scale for the event's E_T^{miss}, yielding the variables:

$$R_{p_T} = \frac{p_{T,CM}}{p_{T,CM} + M_{CM}}, \; R_{p_z} = \frac{p_{z,CM}}{p_{z,CM} + M_{CM}}, \; R_{met} = \frac{E_T^{miss}}{E_T^{miss} + p_{T,CM}} \tag{6.2}$$

denoted `Rpt`, `Rpz`, and `Rmet`.

For this analysis, the RF variables were chosen to be `MCM`, `MH`, `MZ`, `cosCM`, `cosH`, `cosZ`, `dphiCMH`, `Rpt`, `Rpz`, and `Rmet`.

6.4 Correlations

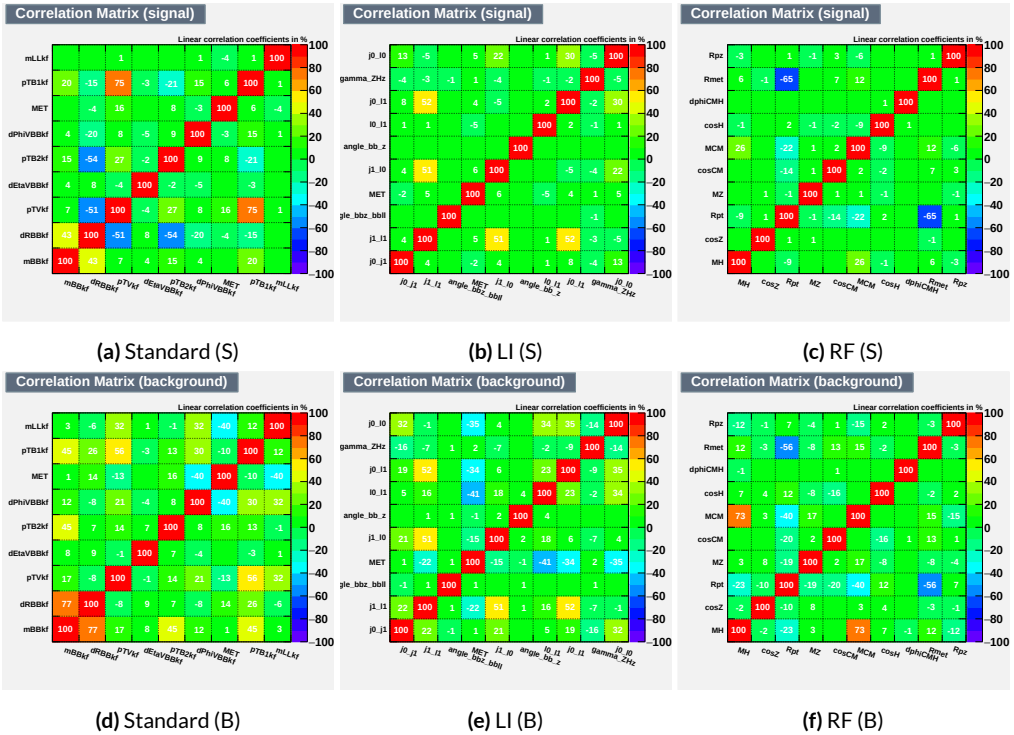
The variable sets used in these studies are summarized in Table ??, while figure 6.2 shows the correlations for the signal and background samples for the standard, LI, and RF variable sets in the most significant analysis signal region, the 2 jet, $p_T^V > 150$ bin (matrices, as well as input distributions, for all regions made be found in Appendix ??). As can be seen in the correlation matrices, variable correlations tend to be much lower, particularly for the signal hypothesis, for the LI and RF variables than for the standard set. Notable exceptions are jet/lepton inner products in the LI set and `Rmet` and `Rpt` in the RF case (not surprising given variable definitions). While this feature is not

6.5 MVA Training

MVA training and hyperparameter optimization (in this case, just the order in which variables are fed into the BDT) is conducted using the "holdout" method LOOK at roger's thesis for a reference. In this scheme, events are divided into three equal portions (in this case using `EventNumber%3`),

Variable Set	Variables
Standard	`mBB, mLL, (mBBJ), pTV, pTB1, pTB2, (pTJ3), dRBB, dPhiVBB, dEtaVBB, MET`
Lorentz Invariants	`j0_j1, j0_l1, l0_l1, j1_l1, j0_l0, j1_l0, gamma_ZHz, angle_bbz_bbll angle_bb_z, MET`
RestFrames	`MH, MCM, MZ, cosH, cosCM, cosZ, Rpz, Rpt, dphiCMH, Rmet`

Table 6.1: Variables used in MVA training. Variables in parentheses are only used in the ≥ 3 jet regions.



(a) Standard (S) **(b)** LI (S) **(c)** RF (S)

(d) Standard (B) **(e)** LI (B) **(f)** RF (B)

Figure 6.2: Signal and background variable correlations for the three variable sets in the 2 jet, $p_T^V > 150$.

with the first third (the "training" set) being used for the initial training, the second third (the "validation" set) being used for hyperparameter optimizaiton, and the final third (the "testing" set) used to evaluate the performanace of the final discriminants in each analysis region.

The MVA discriminant used is a boosted descision tree (BDT). Training is done in TMVA using the training settings of the fiducial analysis[31‡]. For the purposes of hyperparameterization and testing, transformation D with $z_s = z_b = 10$ is applied to the BDT distributions, and the sum in quadrature of the significance $S/\sqrt{S+B}$ in each bin is calculated for each pair of distributions.

Variable ranking is done iteratively in each analysis region. In each set, the validation significance of a BDT using an initial subset of variables is calculated (dRBB and mBB for the standard set; jo_j1 for the LI set; and MH for the RF set). Each of the remaining unranked variables are then added separately, one at a time, to the BDT. The variable yielding the highest validation significance is then added to the set list of ranked variables and removed from the list of unranked variables. This process is repeated until no variables remain. A plot of the ranking for the LI set in the 3+ jet, low p_T^V region can be seen in Figure 6.3.

Once variables have been ranked, the BDT may be used both to evaluate performance in a simplified anlaysis scenario in the absence of systematic uncertainties (described below in Section 6.6) and to create xml files for the production of inputs for an analysis including systematics. Following the approach taken in the fiducial analysis, BDT discriminants using two "k-folds" are produced to prevent overtraining, since the samples used for training are the same as those used to produce in-

‡Namely, `!H:!V:BoostType=AdaBoost:AdaBoostBeta=0.15:SeparationType=GiniIndex:-PruneMethod=NoPruning:NTrees=200:MaxDepth=4:nCuts=100:nEventsMin=5%`

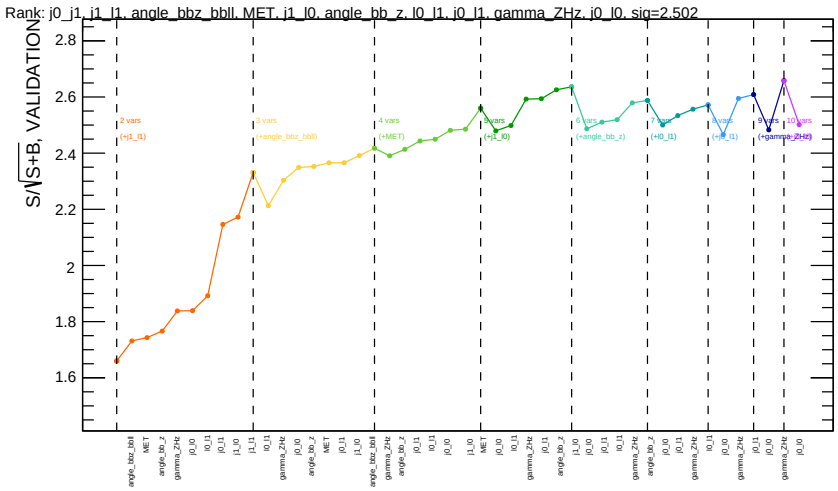


Figure 6.3: An example of a variable ranking to determine variable order in the BDT discriminant. The ranking for the 2 jet, $p_T^V > 150$ region for the LI set is shown.

puts for the full profile likelihood fit. In this scheme, a BDT trained on events with an even (odd) `EventNumber` are used to evaluate events with an odd (even) `EventNumber`.

6.6 Statistics Only BDT Performance

As described above, cumulative significances can be extracted from signal and background BDT output distributions of the testing set of events can be constructed using the discriminant with the optimal variable ranking. An example of such an output is shown in Figure 6.4, with the training BDT distributions in points and the testing distributions after transformation D as block histograms. For better comparison of the distributions, all histograms have been scaled to have the same normalization.

As can be seen in the summary of cumulative significances for each of these analysis regions and variable sets in 6.5, the performance of each of the variable sets is quite similar. The standard set

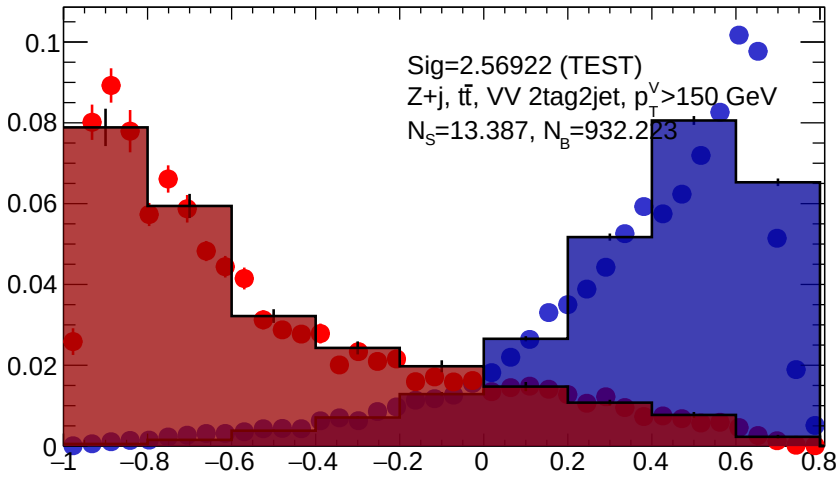


Figure 6.4: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the 2 jet, $p_T^V > 150$ region for the LI set is shown.

performs best, with the LI (RF) set having a cumulative significance that is 7.9% (6.9%) lower. That these figures are all so high ($\sim$ 4.5) relative to numbers one typically sees in an analysis including systematics is in part due to the fact that many of the most significant bins occur at high values of the BDT output, which, as can be seen in Figure 6.4, contain a small fraction of background events.

A full set of ranking and testing plots, as well as correlation matrices and input variable distributions, may be found in Appendix ??.

Correlation, ranking, and input variable plots for the standard, Lorentz Invariant, and Rest-Frames variable sets.

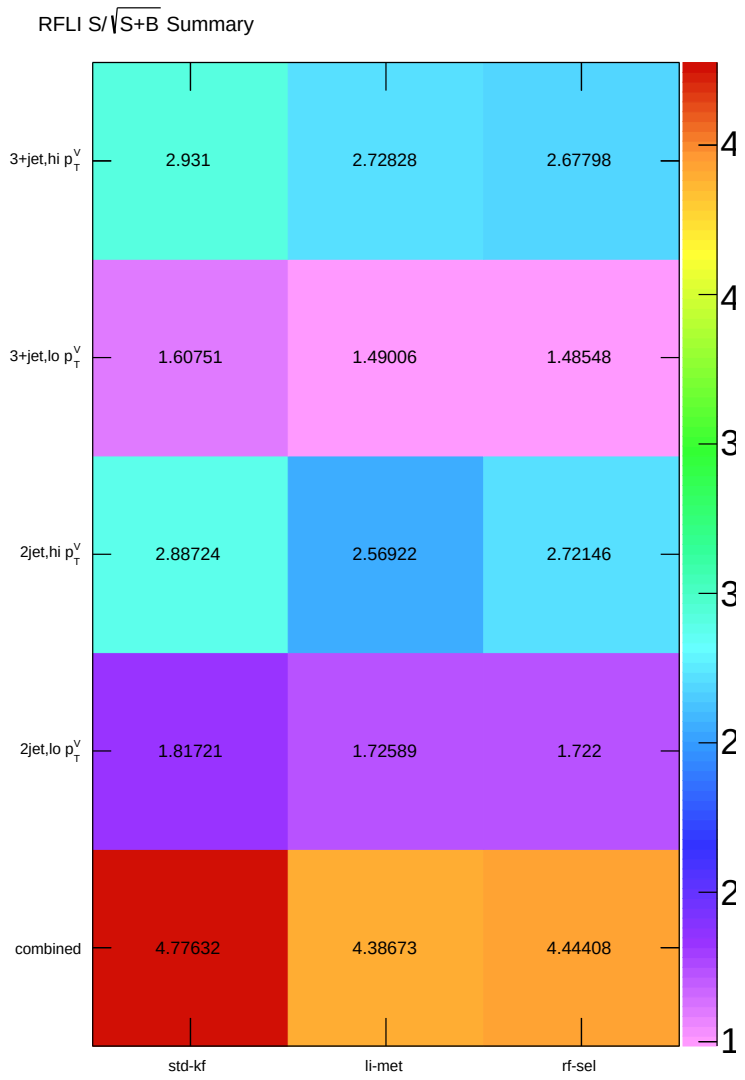
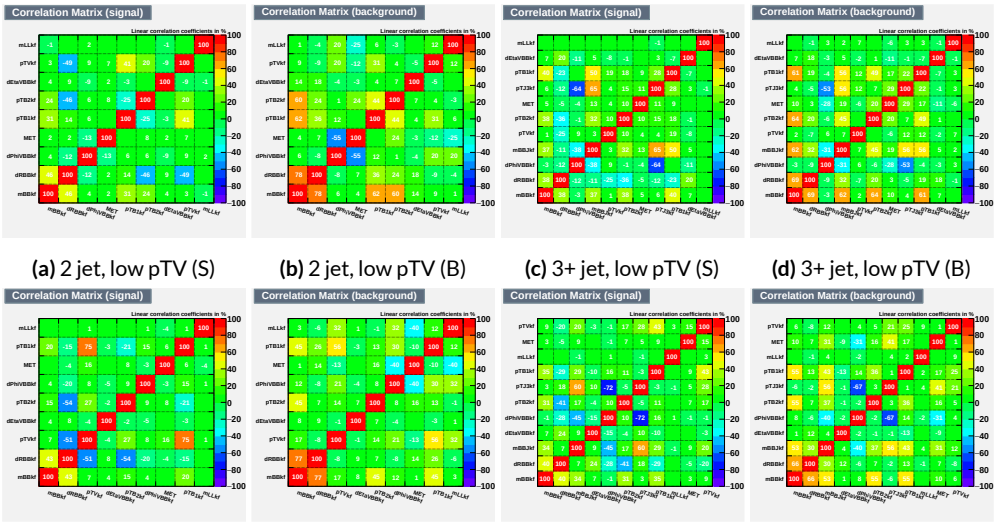


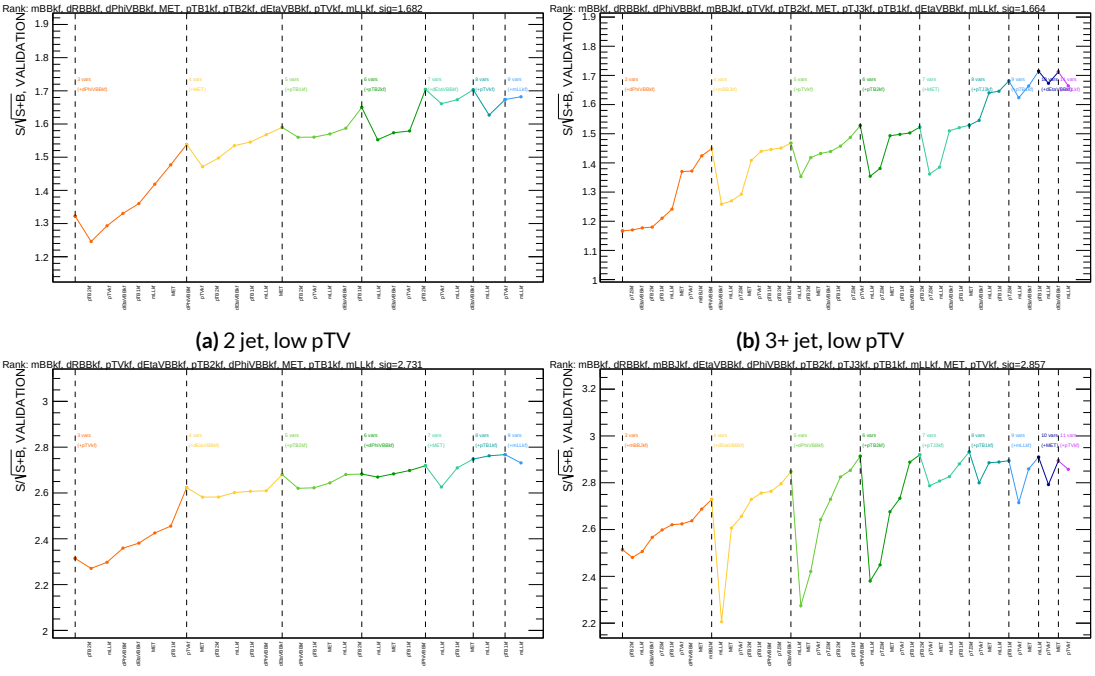
Figure 6.5: Results of testing significances sorted by analysis region and variable set.



(a) 2 jet, low pTV (S) **(b)** 2 jet, low pTV (B) **(c)** 3+ jet, low pTV (S) **(d)** 3+ jet, low pTV (B)

(e) 2 jet, high pTV (S) **(f)** 2 jet, high pTV (B) **(g)** 3+ jet, high pTV (S) **(h)** 3+ jet, high pTV (B)

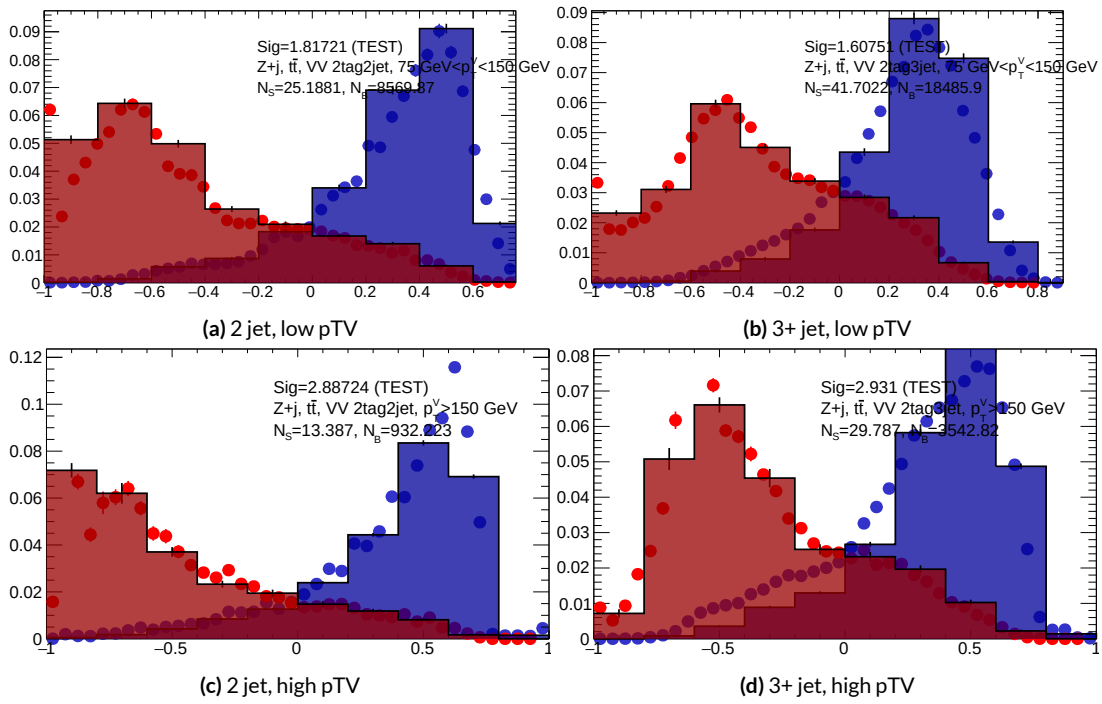
Figure 6.6: Signal and background variable correlations for the standard variable set.



(a) 2 jet, low pTV **(b)** 3+ jet, low pTV

(c) 2 jet, high pTV **(d)** 3+ jet, high pTV

Figure 6.7: Signal and background variable correlations for the standard variable set.



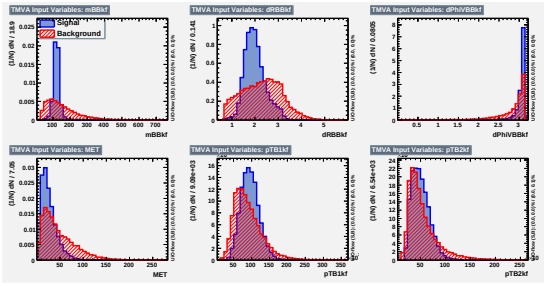
(a) 2 jet, low pTV

(b) 3+ jet, low pTV

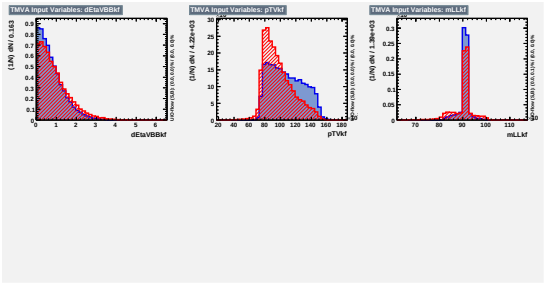
(c) 2 jet, high pTV

(d) 3+ jet, high pTV

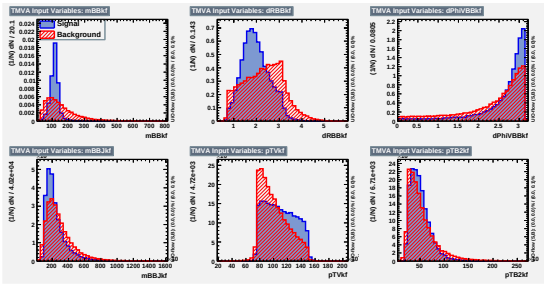
Figure 6.8: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the standard variable set.



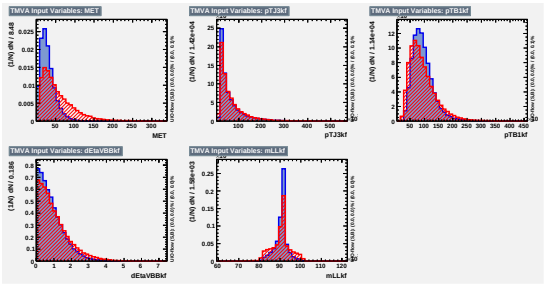
(a) 2 jet, low pTV (1/2)



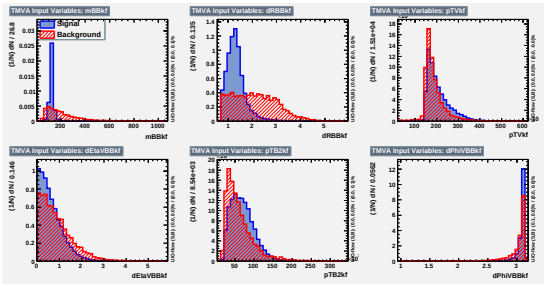
(b) 2 jet, low pTV (2/2)



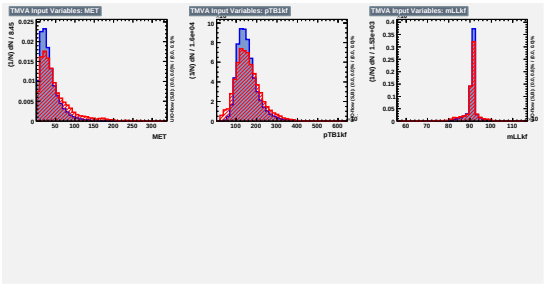
(c) 3+ jet, low pTV (1/2)



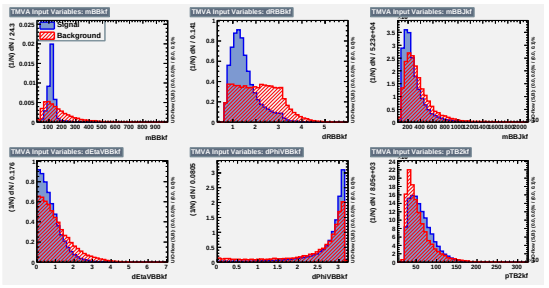
(d) 3+ jet, low pTV (2/2)



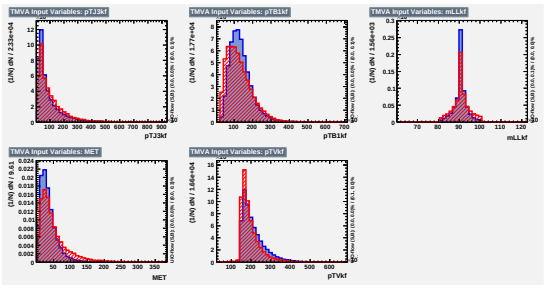
(e) 2 jet, high pTV (1/2)



(f) 2 jet, high pTV (2/2)

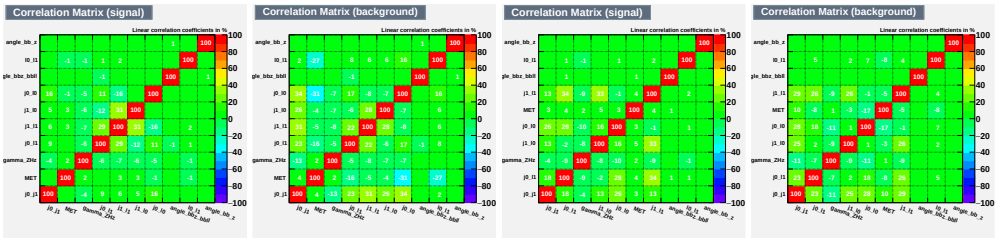


(g) 3+ jet, high pTV (1/2)



(h) 3+ jet, high pTV (2/2)

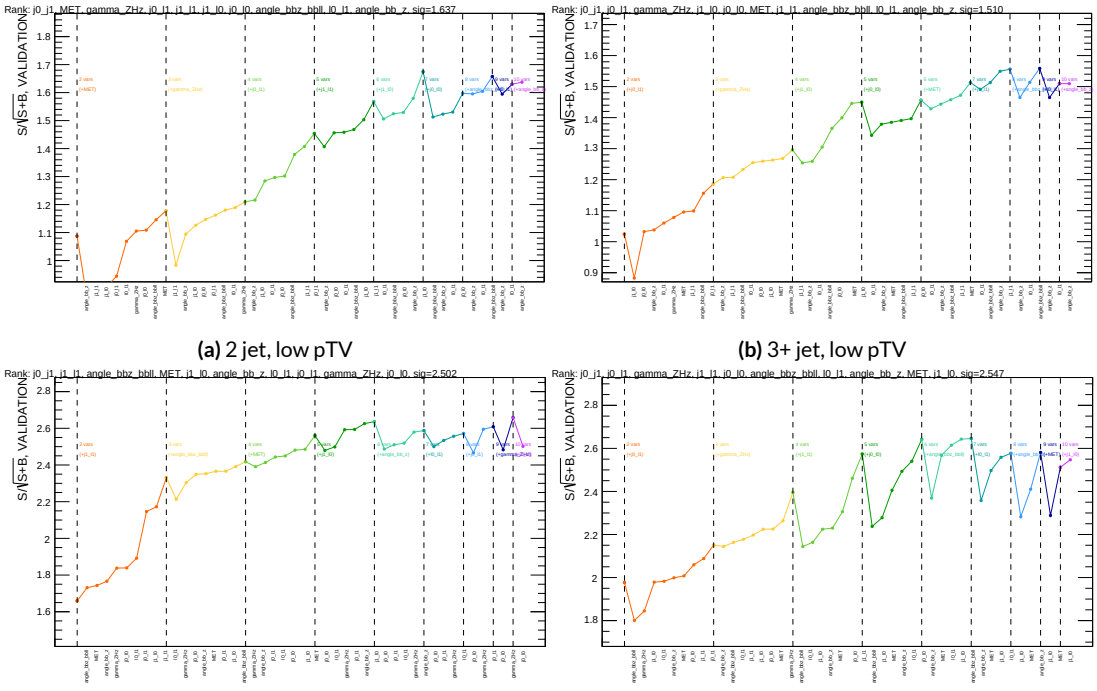
Figure 6.9: Input variables for the standard variable set.



(a) 2 jet, low pTV (S) **(b)** 2 jet, low pTV (B) **(c)** 3+ jet, low pTV (S) **(d)** 3+ jet, low pTV (B)

(e) 2 jet, high pTV (S) **(f)** 2 jet, high pTV (B) **(g)** 3+ jet, high pTV (S) **(h)** 3+ jet, high pTV (B)

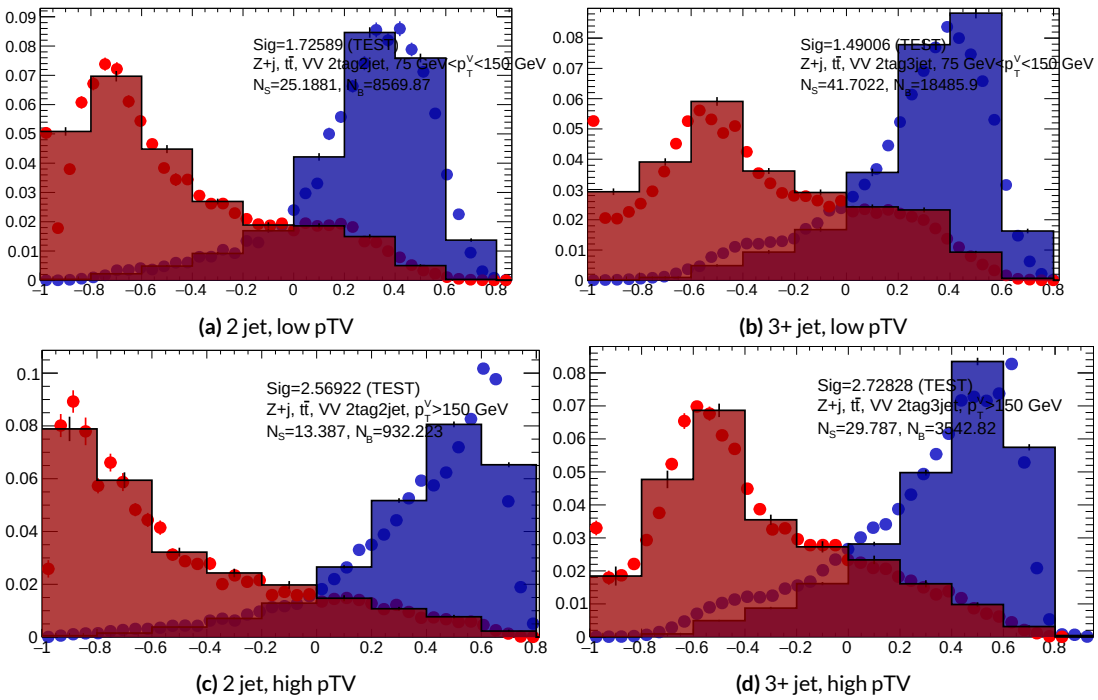
Figure 6.10: Signal and background variable correlations for the LI variable set.



(a) 2 jet, low pTV **(b)** 3+ jet, low pTV

(c) 2 jet, high pTV **(d)** 3+ jet, high pTV

Figure 6.11: Signal and background variable correlations for the LI variable set.



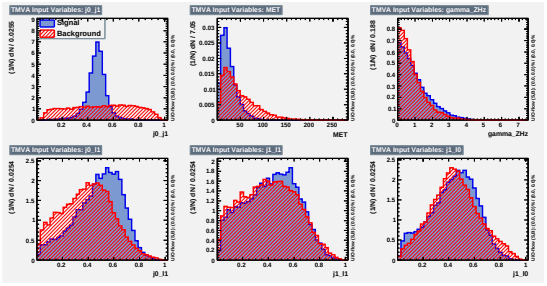
(a) 2 jet, low pTV

(b) 3+ jet, low pTV

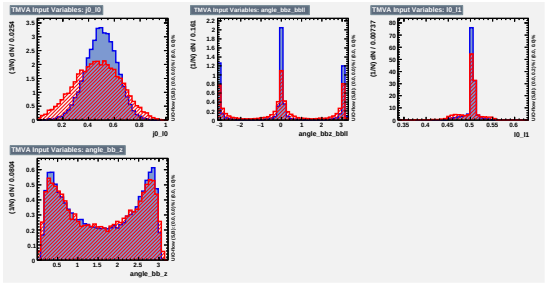
(c) 2 jet, high pTV

(d) 3+ jet, high pTV

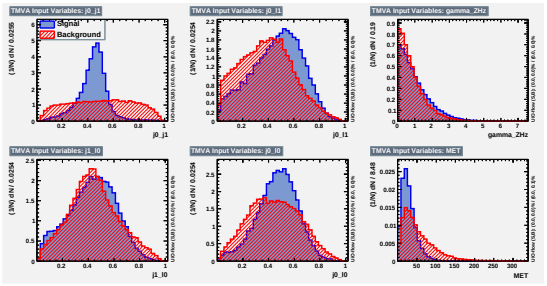
Figure 6.12: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the LI variable set.



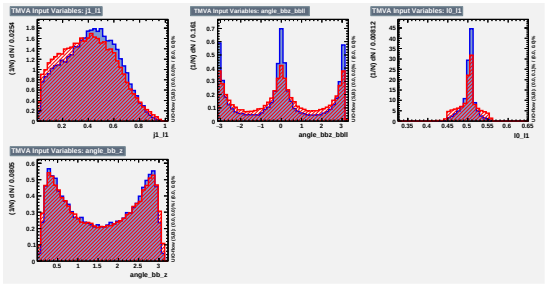
(a) 2 jet, low pTV (1/2)



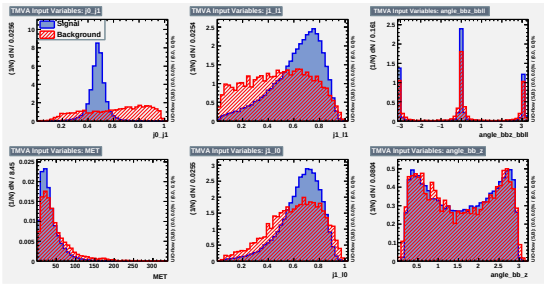
(b) 2 jet, low pTV (2/2)



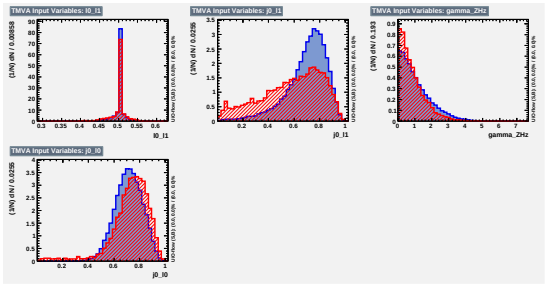
(c) 3+ jet, low pTV (1/2)



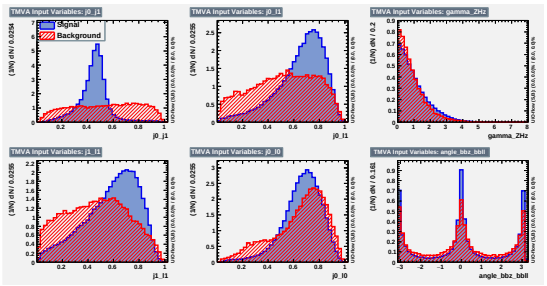
(d) 3+ jet, low pTV (2/2)



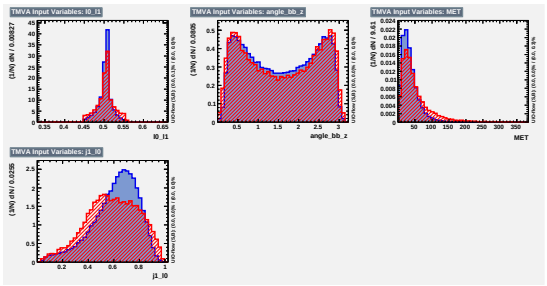
(e) 2 jet, high pTV (1/2)



(f) 2 jet, high pTV (2/2)

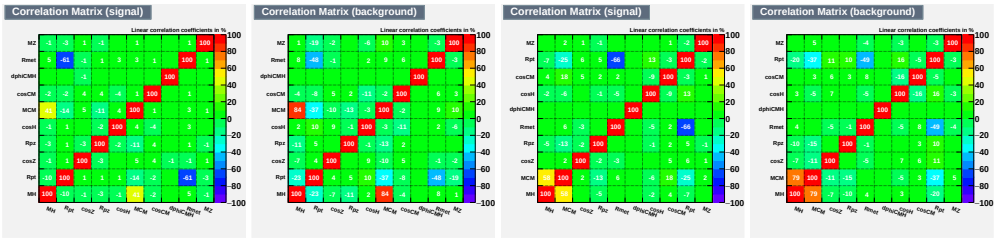


(g) 3+ jet, high pTV (1/2)

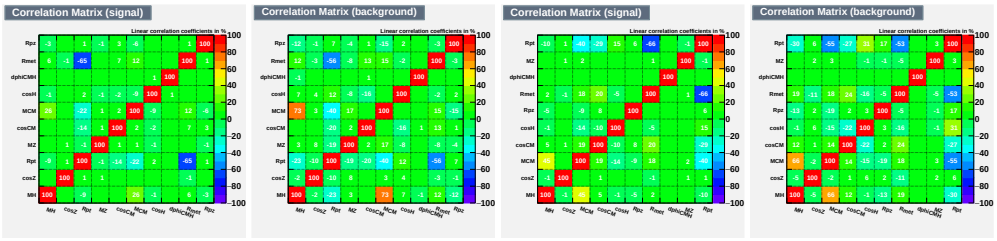


(h) 3+ jet, high pTV (2/2)

Figure 6.13: Input variables for the LI variable set.

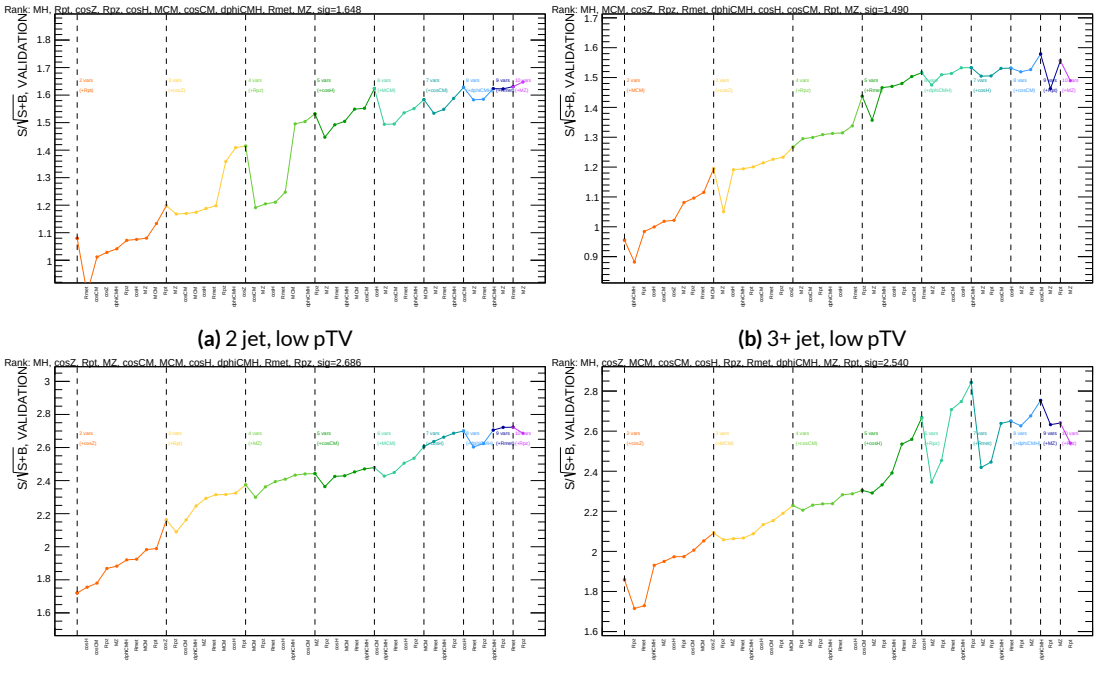


(a) 2 jet, low pTV (S) **(b)** 2 jet, low pTV (B) **(c)** 3+ jet, low pTV (S) **(d)** 3+ jet, low pTV (B)



(e) 2 jet, high pTV (S) **(f)** 2 jet, high pTV (B) **(g)** 3+ jet, high pTV (S) **(h)** 3+ jet, high pTV (B)

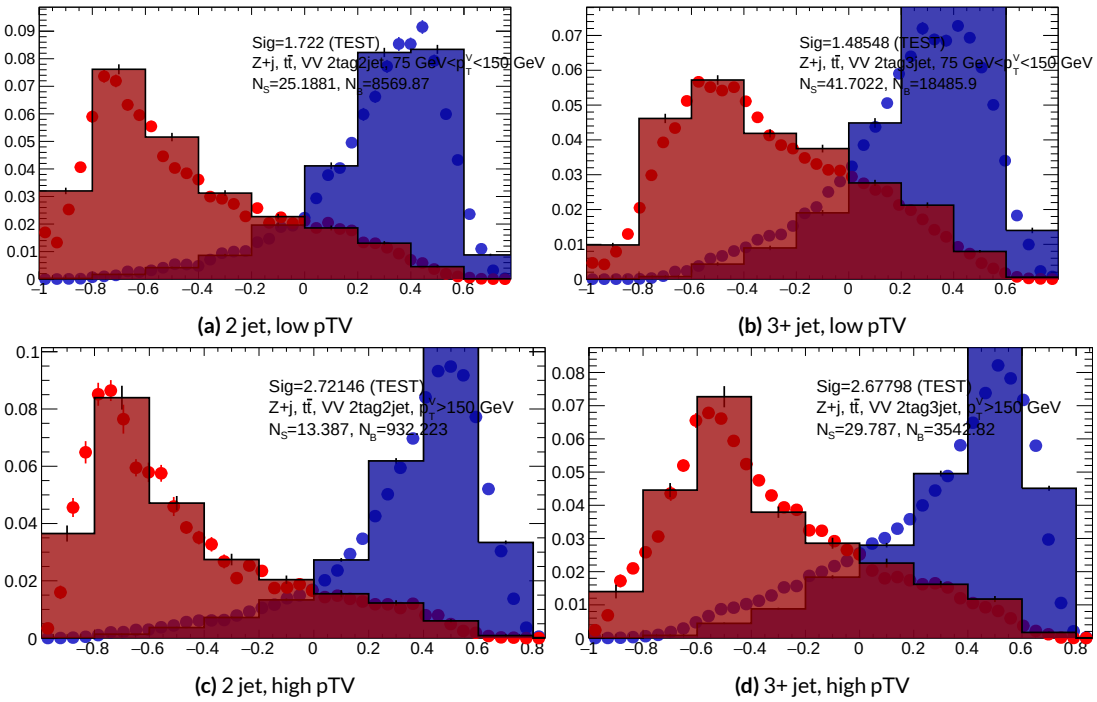
Figure 6.14: Signal and background variable correlations for the RF variable set.



(a) 2 jet, low pTV **(b)** 3+ jet, low pTV

(c) 2 jet, high pTV **(d)** 3+ jet, high pTV

Figure 6.15: Signal and background variable correlations for the RF variable set.



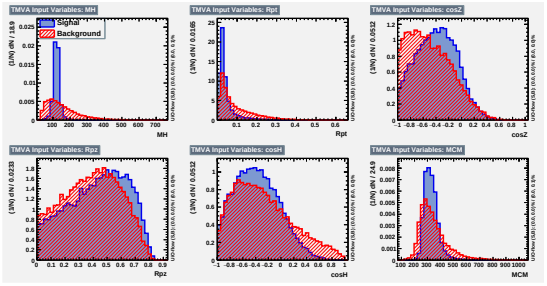
(a) 2 jet, low pTV

(b) 3+ jet, low pTV

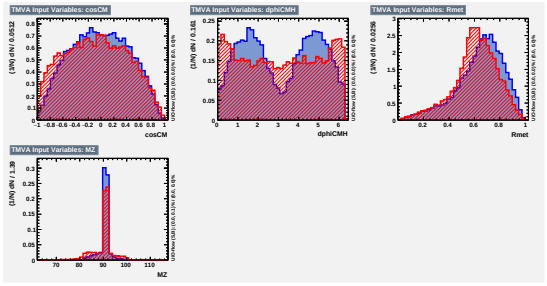
(c) 2 jet, high pTV

(d) 3+ jet, high pTV

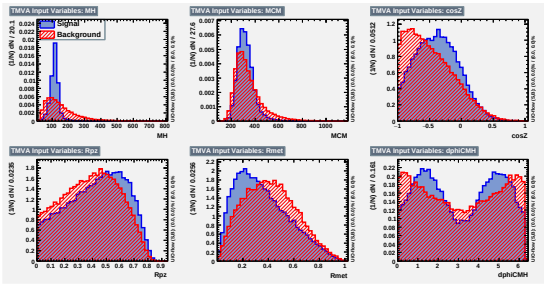
Figure 6.16: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the RF variable set.



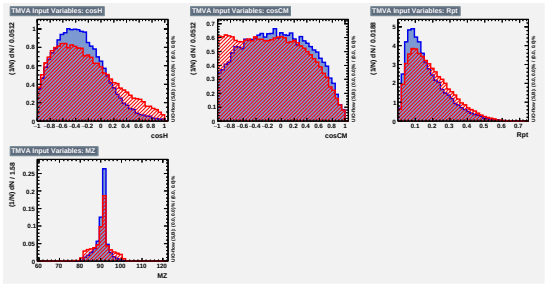
(a) 2 jet, low pTV (1/2)



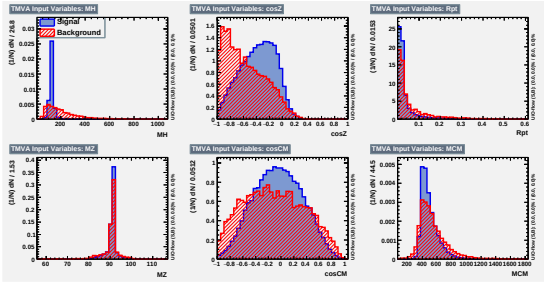
(b) 2 jet, low pTV (2/2)



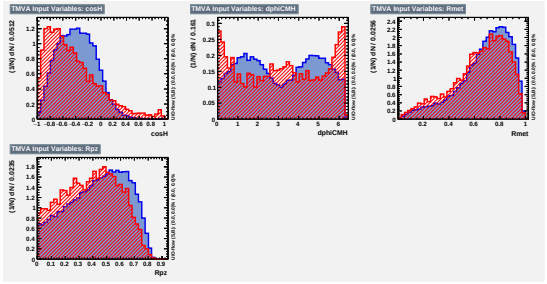
(c) 3+ jet, low pTV (1/2)



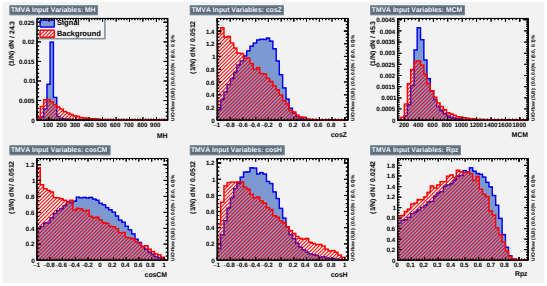
(d) 3+ jet, low pTV (2/2)



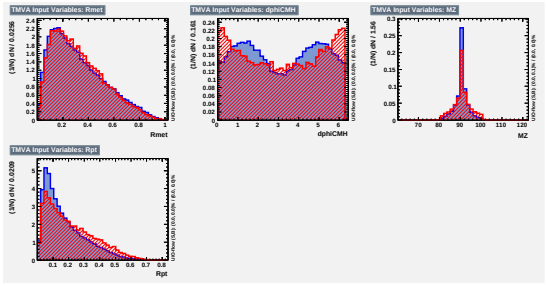
(e) 2 jet, high pTV (1/2)



(f) 2 jet, high pTV (2/2)



(g) 3+ jet, high pTV (1/2)



(h) 3+ jet, high pTV (2/2)

Figure 6.17: Input variables for the RF variable set.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

7

Statistical Fit Model and Validation

Much has been said

7.1 The Fit Model

In order to derive the strength of the signal process $ZH \to \ell\ell b\bar{b}$, denoted μ, and other quantities of interest while taking into account systematic uncertainties or nuisance parameters (NP's, collectively

denoted ϑ), a binned likelihood function is constructed as the product of Poisson distributions over bins:

$$\mathcal{L}(\mu, \vartheta) = \text{Pois}(n \,|\mu S + B) \left[\prod_{i \in \text{bins}} \frac{\mu s_i + b_i}{\mu S + B} \right] \tag{7.1}$$

where n is the total number of events observed, the s_i and b_i are the number of expected signal and background events in each bin, and S and B are the expected total signal and background events. The signal and background expectations depend on the NP's ϑ. NP's related to the normalization of signal and background processes fall into two categories. The first set is left to float freely like μ while the second set are parametrized as log-normal to prevent negative predicted values. All other NP's are parametrized with Gaussian priors.*

One can maximize the likelihood in Equation 7.1 for a fixed value of μ to derive estimators for the NP's ϑ; values of ϑ so derived are denoted $\hat{\hat{\vartheta}}$ to emphasize that these are likelihood maximizing for a given ϑ. The profile likelihood technique finds the likelihood function's maximum by comparing the values of the likelihood over all possible values of μ using these "profiles" and picking the one with the greatest $\mathcal{L}\left(\mu, \hat{\hat{\vartheta}}\right)$ value; these values of μ, ϑ are denoted $\hat{\mu}$ and $\hat{\vartheta}$. The profile likelihood can further be used to construct a test statistic†

$$q = -2 \frac{\mathcal{L}\left(\mu, \hat{\hat{\vartheta}}\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\vartheta}\right)} \tag{7.2}$$

This statistic can be used to derive the usual significance (p value), by setting $\mu = 0$ to find the

*This results in a "penalty" on the NLL discussed below of $\left(\hat{\alpha} - \mu\right)^2 / \sigma^2$, for NP α, normally parametrized with mean μ and variance σ^2 for an MLE of $\hat{\alpha}$.

†The factor of -2 is added so that this statistic gives, in the asymptotic limit of large N, a χ^2 distribution.

compatability with the background-only hypothesis[?] . If there is insufficient evidence for the signal hypothesis, the CL_s method can be used to set limits[?] .

In order to both validate the fit model and study the behavior of fits independent of a given data set, a so-called "Asimov" data set is constructed for a given fit model; this dataset has each bin equal to its expectation value for assumed values of the NP's and a given μ value (in this case, $\mu = 1$).

7.2 Fit Inputs

Inputs to the binned likelihood are distributions of the BDT outputs described above for the signal regions and of m_{bb} for the top $e - \mu$ control regions. These regions split events according to their $p_{\mathrm{T}}^{\mathrm{V}}$ and number of jets in an event. All events are required to have two b-tags, as well as pass the other event selection requirements summarized in Table ??; the only difference between the signal and control region selection is that the same flavor (i.e. leptons both be electrons or muons) is flipped so that events in the control region have exactly one electron and one muon. The BDT outputs are binned using transformation D, while the m_{bb} distributions have 50 bins, with the exception of the 2 jet, high $p_{\mathrm{T}}^{\mathrm{V}}$ region, where a single bin is used due to low statistics.

Input distributions in MC are further divided according to their physics process. The signal processes are divided based on both the identity of associated V and the number of leptons in the final state; $ZH \to \ell\ell b\bar{b}$ events are further separated into distributions for qq and gg initiated processes. The V+jet events are similarly split based on the identity of the V and then further subdivided according to the flavor of the two leading jets in an event, b, c, or l, for a total of six categories. Due to the effectiveness of the 2 b-tag requirement suppressing the presence of both c and l jets, truth-

tagging is used to boost MC statistics in the *cc*, *cl*, and *l* distributions.[‡] For top backgrounds, single top production is split according to production mode (*s*,*t*, and *Wt* production), with $t\bar{t}$ as single category. Diboson background distributions are also split according to the identity of the *V*'s (*ZZ*, *WZ*, and *WW*). Fit input segmentation is summarized in Table ??.

Category	Bins
# of Jets	2, 3+
$p_{\mathrm{T}}^{\mathrm{V}}$ Regions ()	$[75, 150)$, $[150, \infty)$
Sample	data, signal $[(W, qqZ, ggZ) \times n_{lep}]$, V+jet $[(W, Z) \times (bb, bc, bl, cc, cl, l)]$, $t\bar{t}$, diboson (ZZ, WW, WZ), single top (s, t, Wt)

Table 7.1: Fit input segmentation.

7.3 Systematic Uncertainties

A full discussion of systematic uncertainties can be found in [31]. A brief summary of the NP's considered in these studies is provided below.

7.3.1 Modeling and Theoretical Uncertainties

The signal and background physics processes considered in the final statistical fit and their nominal samples are described in Section ??. In addition to the nominal samples, alternate samples, described in [33], are also used to derive systematic uncertainties, also described there—these are summarized in Table 7.2 below. p_T^V systematics are generally shape and normalization, whereas m_{bb} systematics are shape only; these shape systematics are usually parametrized as linear functions.

[‡]Since *WW* is not an important contriubtion to the already small total diboson background, no truth tagging was applied here, in contrast to the fiducial analysis.

Process	Systematics
Signal	$H \to bb$ decay, QCD scale, PDF+α_S scale, UE+PS (acc., p_T^V, m_{bb}, 3/2 jet ratio)
Z+jets	Acc, flavor composition, p_T^V+m_{bb} shape
$t\bar{t}$	Acc, p_T^V+m_{bb} shape
Single top	Acc., p_T^V+m_{bb} shape
Diboson	Overall acc., UE+PS (acc, p_T^V, m_{bb}, 3/2 jet ratio), QCD scale (acc (2, 3 jet, jet veto), p_T^V, m_{bb})

Table 7.2: Summary of modeling systematic uncertainties.

7.3.2 Experimental Systematics

A full discussion may be found in [21], and a full summary table may be found at Table 33 of [31].

Process	Systematics
Jets	21 NP scheme for JES, JER as single NP
E_T^{miss}	trigger efficiency, track-based soft terms, scale uncertainty due to jet tracks
Flavor Tagging	Eigen parameter scheme (CDI File: `2016-20_7-13TeV-MC15-CDI-2017-06-07_v2`)
Electrons	trigger eff, reco/ID eff, isolation eff, energy scale/resoltuion
Muons	trigger eff, reco/ID eff, isolation eff, track to vertex association, momentum resolution/scale
Event	total luminosity, pileup reweighting

Table 7.3: Summary of experimental systematic uncertainties.

For fit model validation, cf. Appendix ??.

7.4 Full Breakdown of Errors

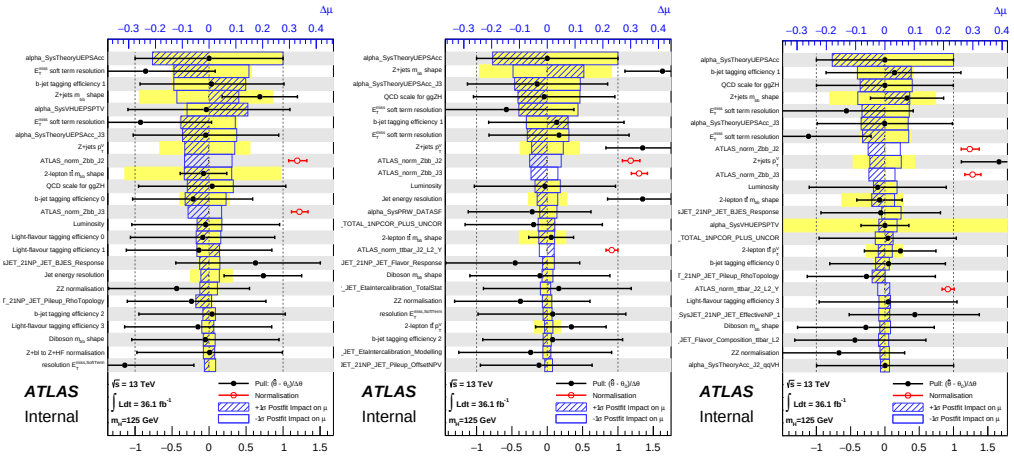
A postfit ranking of nuisance parameters according to their impact on $\hat{\mu}$ for the different variable sets may be found in Figure 7.1.

	Std-KF	LI+MET	RF
Total	+0.608 / -0.511	+0.632 / -0.539	+0.600 / -0.494
DataStat	+0.420 / -0.401	+0.453 / -0.434	+0.424 / -0.404
FullSyst	+0.440 / -0.318	+0.441 / -0.319	+0.425 / -0.284
Floating normalizations	+0.122 / -0.125	+0.110 / -0.111	+0.093 / -0.089
All normalizations	+0.128 / -0.129	+0.112 / -0.112	+0.099 / -0.092
All but normalizations	+0.403 / -0.274	+0.387 / -0.250	+0.382 / -0.227
Jets, MET	+0.180 / -0.097	+0.146 / -0.079	+0.122 / -0.083
Jets	+0.051 / -0.030	+0.044 / -0.035	+0.025 / -0.042
MET	+0.173 / -0.091	+0.140 / -0.074	+0.117 / -0.063
BTag	+0.138 / -0.136	+0.069 / -0.071	+0.076 / -0.078
BTag b	+0.125 / -0.125	+0.067 / -0.070	+0.073 / -0.075
BTag c	+0.018 / -0.016	+0.004 / -0.004	+0.005 / -0.005
BTag light	+0.057 / -0.051	+0.020 / -0.014	+0.009 / -0.018
Leptons	+0.013 / -0.012	+0.029 / -0.026	+0.012 / -0.023
Luminosity	+0.052 / -0.020	+0.050 / -0.016	+0.050 / -0.019
Diboson	+0.043 / -0.039	+0.035 / -0.031	+0.038 / -0.029
Model Zjets	+0.119 / -0.117	+0.124 / -0.127	+0.095 / -0.086
Zjets flt. norm.	+0.080 / -0.106	+0.052 / -0.092	+0.026 / -0.072
Model Wjets	+0.001 / -0.001	+0.001 / -0.001	+0.000 / -0.001
Wjets flt. norm.	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Model ttbar	+0.076 / -0.080	+0.025 / -0.035	+0.025 / -0.040
Model Single Top	+0.015 / -0.015	+0.002 / -0.004	+0.021 / -0.007
Model Multi Jet	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Signal Systematics	+0.262 / -0.087	+0.272 / -0.082	+0.290 / -0.088
MC stat	+0.149 / -0.136	+0.168 / -0.154	+0.153 / -0.136

Table 7.4: Expected error breakdowns for the standard, LI, and RF variable sets

	Std-KF	LI+MET	RF
$\hat{\mu}$	1.7458	1.6467	1.5019
Total	+0.811 / -0.662	+0.778 / -0.641	+0.731 / -0.612
DataStat	+0.502 / -0.484	+0.507 / -0.489	+0.500 / -0.481
FullSyst	+0.637 / -0.451	+0.591 / -0.415	+0.533 / -0.378
Floating normalizations	+0.153 / -0.143	+0.128 / -0.118	+0.110 / -0.109
All normalizations	+0.158 / -0.147	+0.130 / -0.119	+0.112 / -0.110
All but normalizations	+0.599 / -0.402	+0.544 / -0.354	+0.486 / -0.318
Jets, MET	+0.218 / -0.145	+0.198 / -0.113	+0.167 / -0.106
Jets	+0.071 / -0.059	+0.065 / -0.047	+0.036 / -0.051
MET	+0.209 / -0.130	+0.190 / -0.102	+0.152 / -0.077
BTag	+0.162 / -0.166	+0.093 / -0.070	+0.115 / -0.099
BTag b	+0.142 / -0.147	+0.090 / -0.066	+0.110 / -0.094
BTag c	+0.022 / -0.021	+0.006 / -0.006	+0.007 / -0.007
BTag light	+0.074 / -0.072	+0.025 / -0.022	+0.031 / -0.029
Leptons	+0.039 / -0.029	+0.035 / -0.031	+0.034 / -0.030
Luminosity	+0.079 / -0.039	+0.073 / -0.034	+0.069 / -0.032
Diboson	+0.047 / -0.043	+0.031 / -0.028	+0.029 / -0.028
Model Zjets	+0.164 / -0.152	+0.141 / -0.143	+0.101 / -0.105
Zjets flt. norm.	+0.070 / -0.109	+0.041 / -0.086	+0.033 / -0.083
Model Wjets	+0.001 / -0.001	+0.001 / -0.000	+0.001 / -0.001
Wjets flt. norm.	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Model ttbar	+0.067 / -0.102	+0.029 / -0.040	+0.040 / -0.048
Model Single Top	+0.015 / -0.020	+0.001 / -0.005	+0.004 / -0.006
Model Multi Jet	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Signal Systematics	+0.434 / -0.183	+0.418 / -0.190	+0.364 / -0.152
MC stat	+0.226 / -0.201	+0.221 / -0.200	+0.212 / -0.189

Table 7.5: Observed signal strengths, and error breakdowns for the standard, LI, and RF variable sets

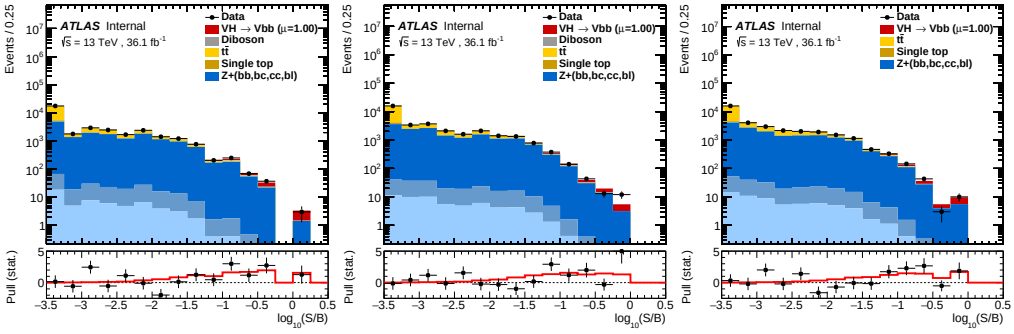


(a) Standard **(b)** LI **(c)** RF

Figure 7.1: Plots for the top 25 nuisance parameters according to their postfit impact on $\hat{\mu}$ for the standard (a), LI (b), and RF (c) variable sets.

7.5 S/B Plot

Plots for the binned S/B in signal region distributions may be found in Figure 7.2.



(a) Standard **(b)** LI **(c)** RF

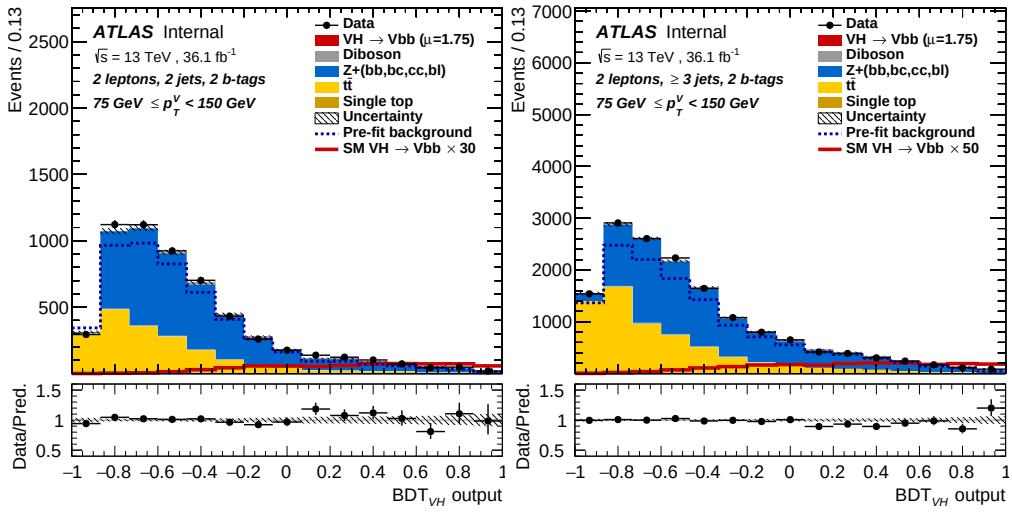
Figure 7.2: Binned S/B plots for the standard (a), LI (b), and RF (c) variable sets.

7.6 Postfit Distributions

Postfit distributions for the MVA discriminant (m_{bb}) distribution in the signal (top $e-\mu$ control) region for the standard, Lorentz Invariant, and RestFrames variable sets.

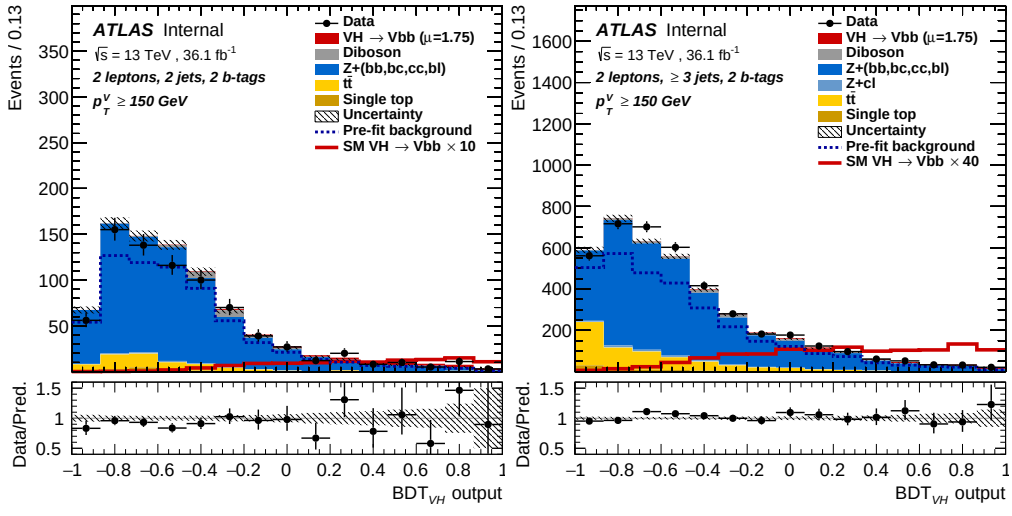
7.7 Nuisance Parameter Pulls

As can be seen in Figures 7.9–7.13, the fits for the three different variable sets are fairly similar from a NP pull perspective. Black is the standard variable set, red is the LI set, and blue is the RF set.



(a) 2 jet, low pTV

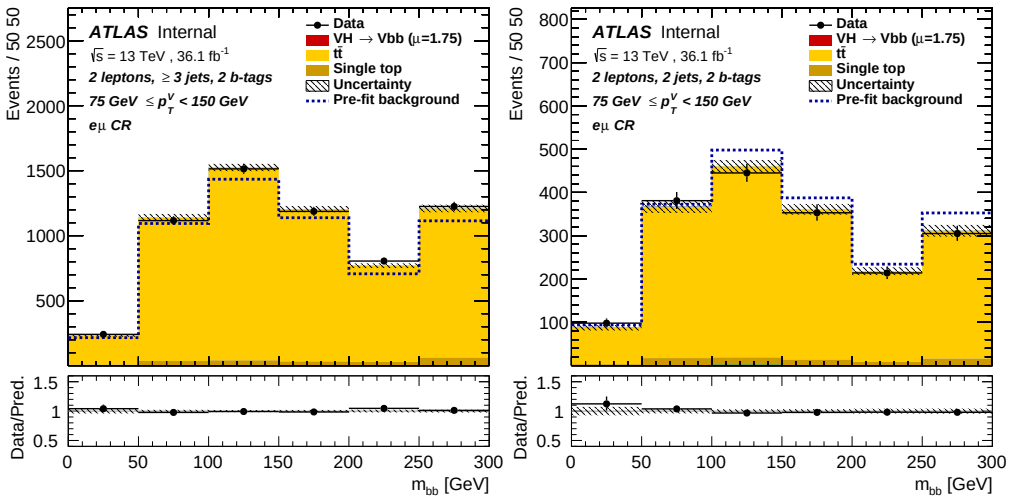
(b) 3+ jet, low pTV



(c) 2 jet, high pTV

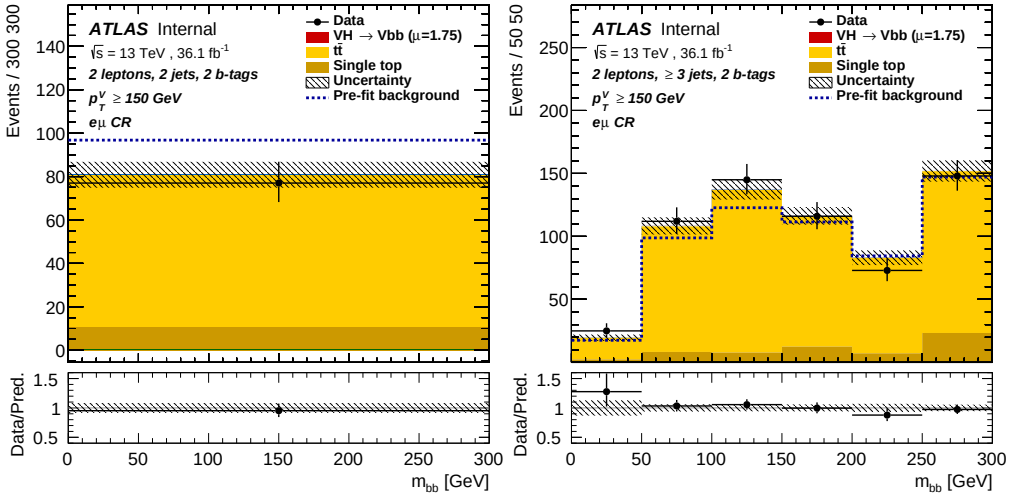
(d) 3+ jet, high pTV

Figure 7.3: Postfit BDT_{VH} plots in the signal region for the standard variable set.



(a) 3+ jet, low pTV

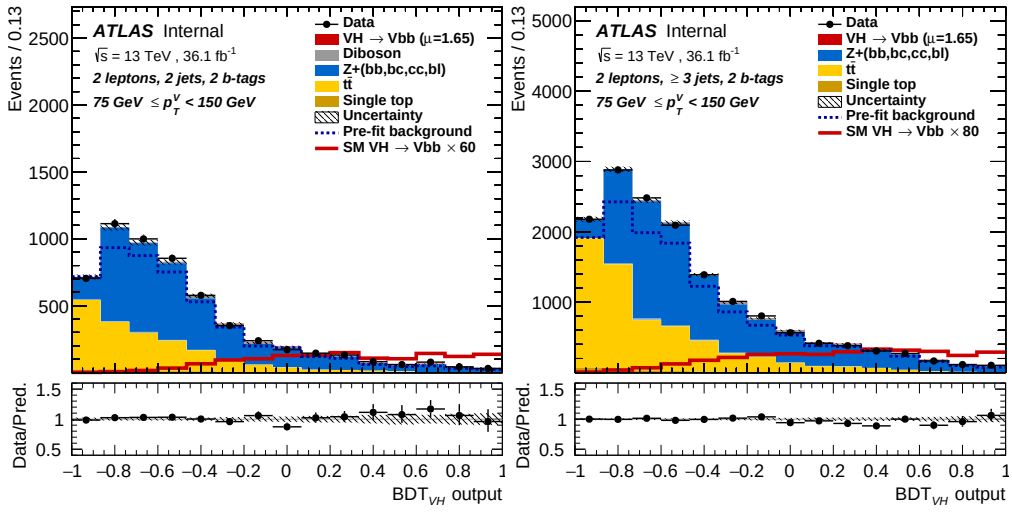
(b) 2 jet, low pTV



(c) 2 jet, high pTV

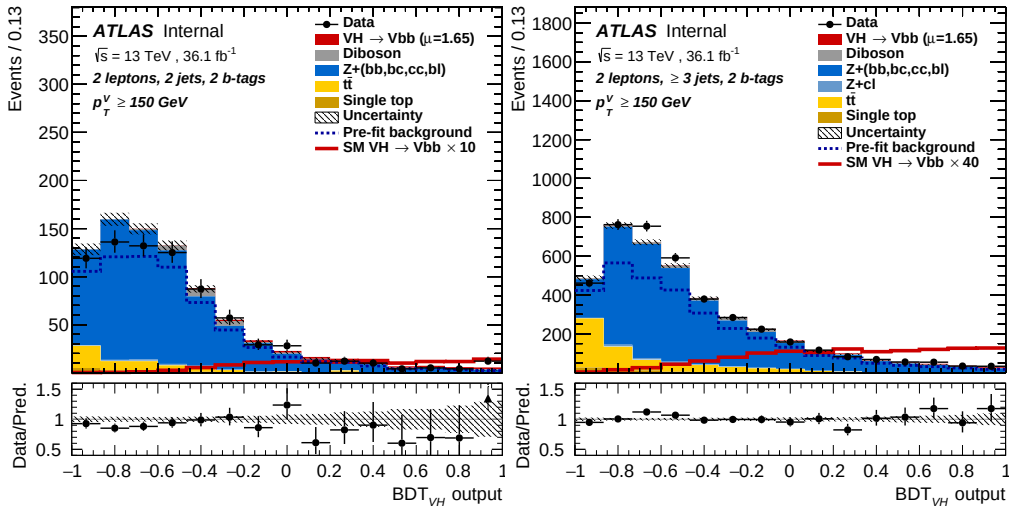
(d) 3+ jet, high pTV

Figure 7.4: Postfit m_{bb} plots in the top $e - \mu$ CR for the standard variable set.



(a) 2 jet, low pTV

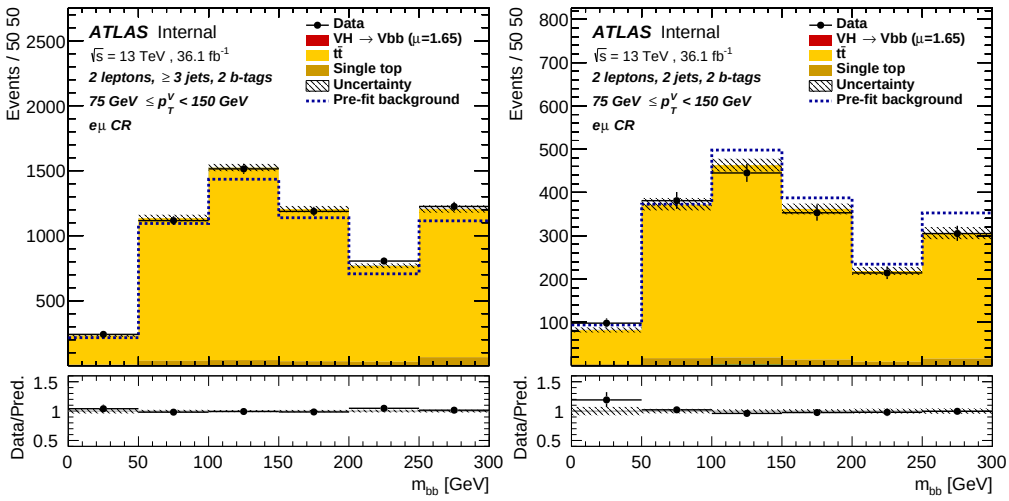
(b) 3+ jet, low pTV



(c) 2 jet, high pTV

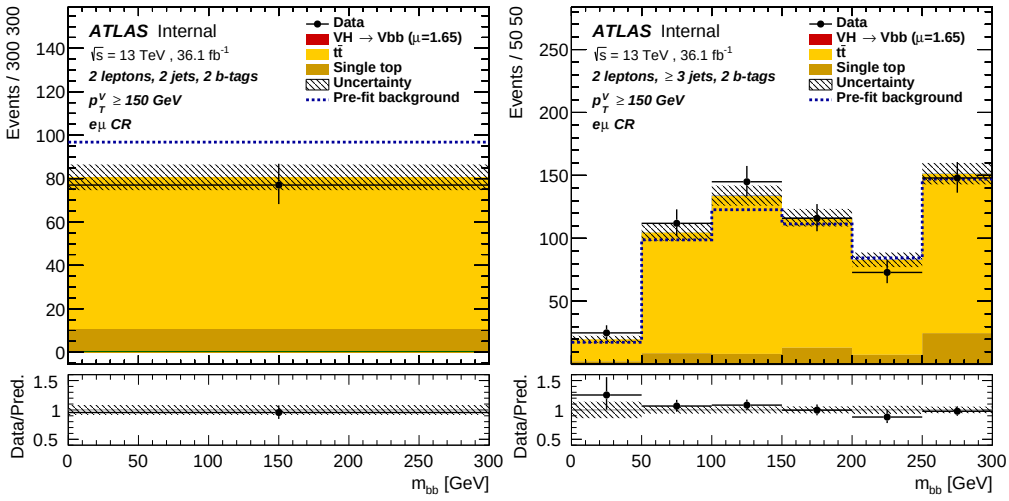
(d) 3+ jet, high pTV

Figure 7.5: Postfit BDT_{VH} plots in the signal region for the LI variable set.



(a) 3+ jet, low pTV

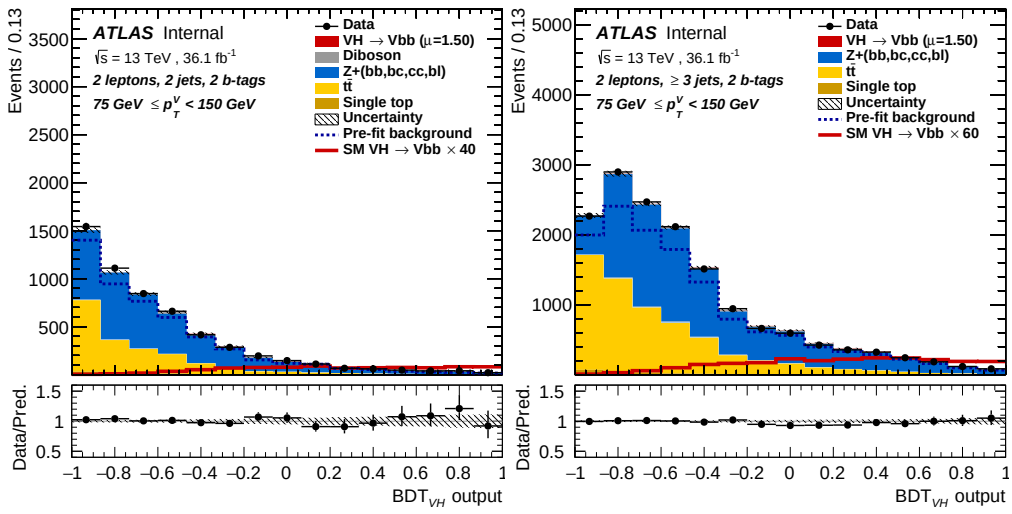
(b) 2 jet, low pTV



(c) 2 jet, high pTV

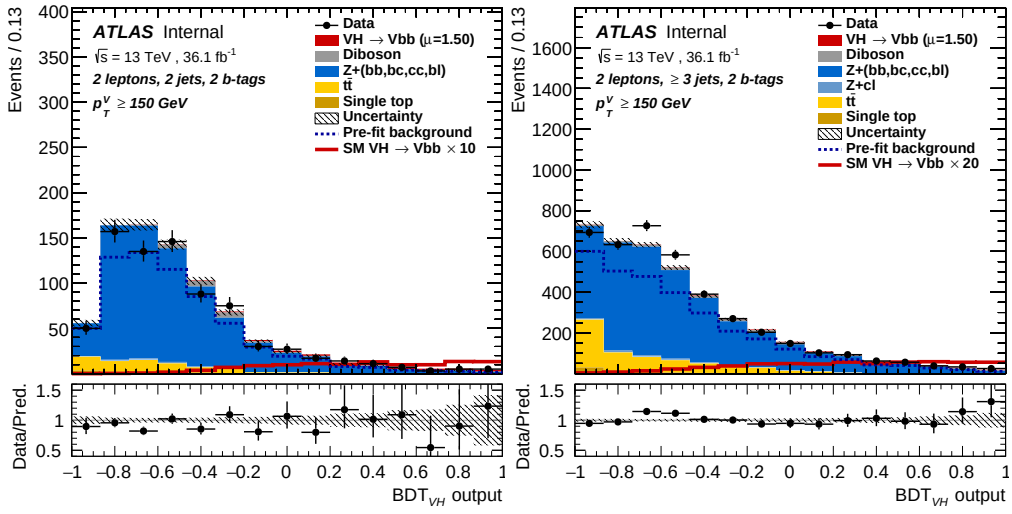
(d) 3+ jet, high pTV

Figure 7.6: Postfit m_{bb} plots in the top $e-\mu$ CR for the LI variable set.



(a) 2 jet, low pTV

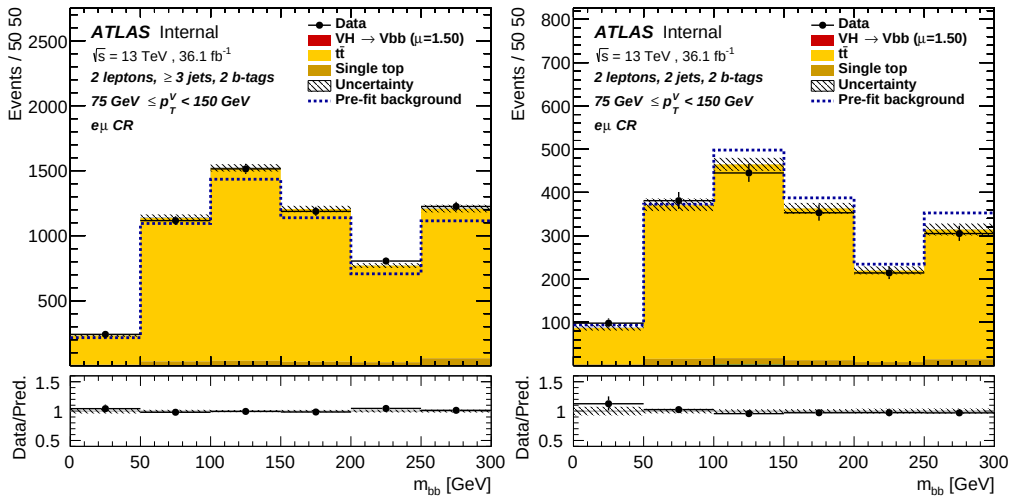
(b) 3+ jet, low pTV



(c) 2 jet, high pTV

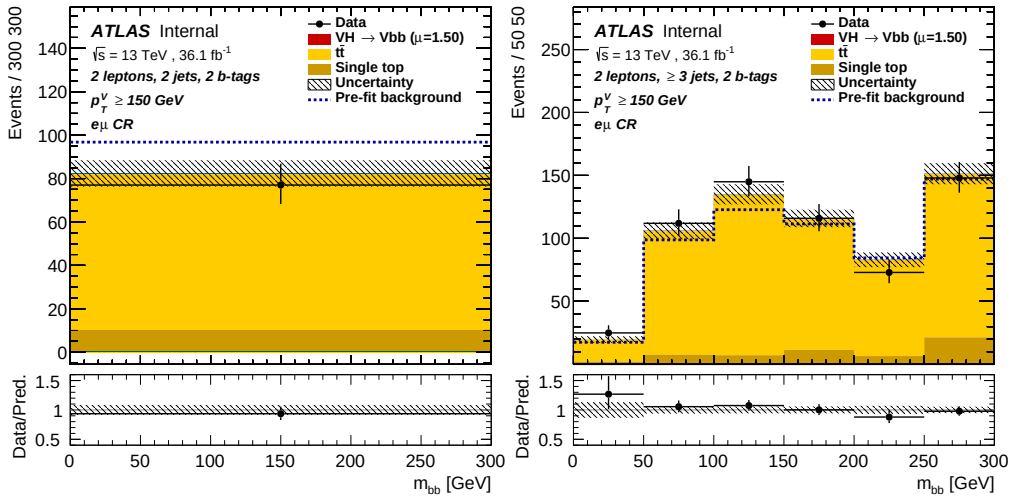
(d) 3+ jet, high pTV

Figure 7.7: Postfit BDT_{VH} plots in the signal region for the RF variable set.



(a) 3+ jet, low pTV

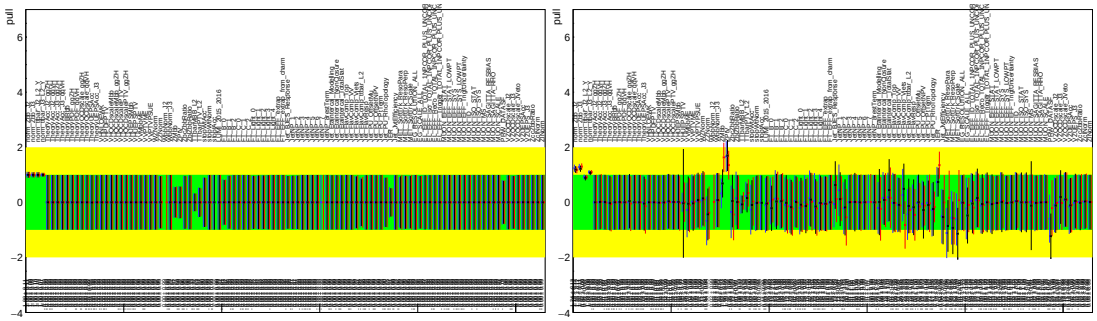
(b) 2 jet, low pTV



(c) 2 jet, high pTV

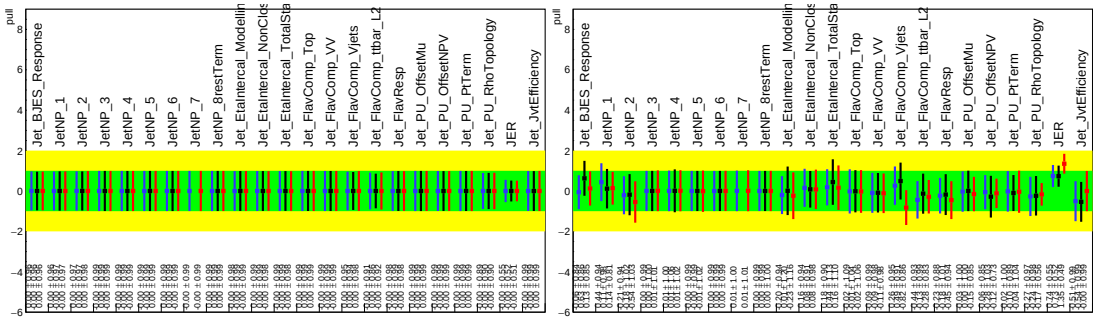
(d) 3+ jet, high pTV

Figure 7.8: Postfit m_{bb} plots in the top $e-\mu$ CR for the RF variable set.



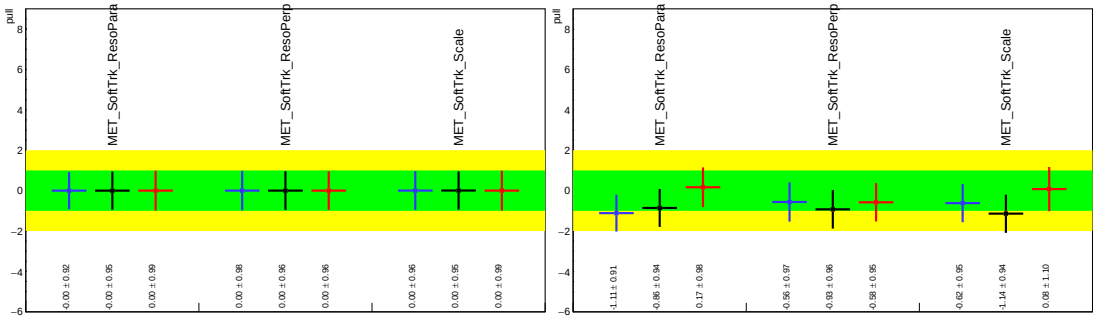
(a) Asimov **(b)** Observed

Figure 7.9: Pull comparison for all NP's but MC stats.



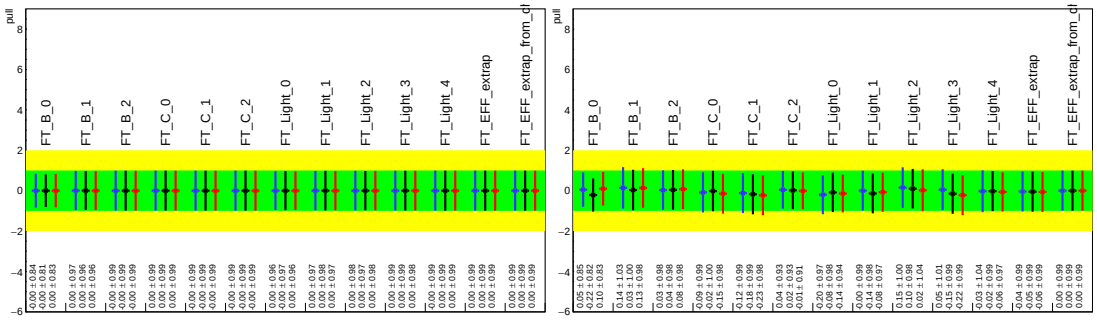
(a) Asimov **(b)** Observed

Figure 7.10: Pull comparison for jet NP's.



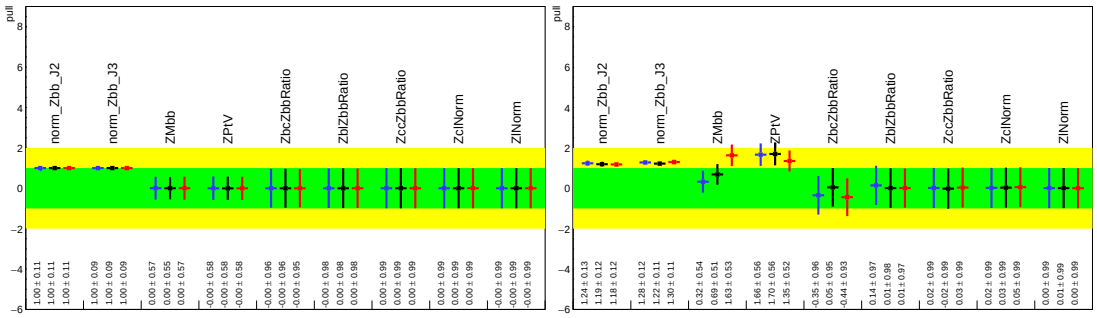
(a) Asimov **(b)** Observed

Figure 7.11: Pull comparison for MET NP's.



(a) Asimov **(b)** Observed

Figure 7.12: Pull comparison for Flavour Tagging NP's.

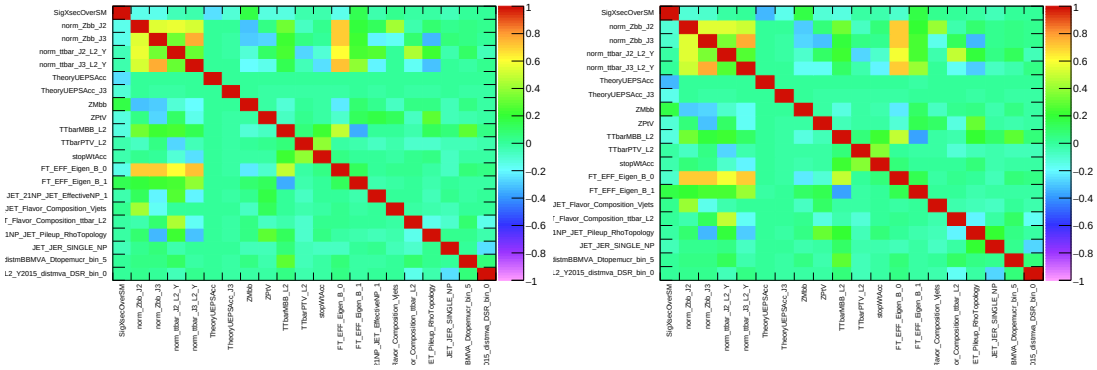


(a) Asimov **(b)** Observed

Figure 7.13: Pull comparison for Z+jets NP's.

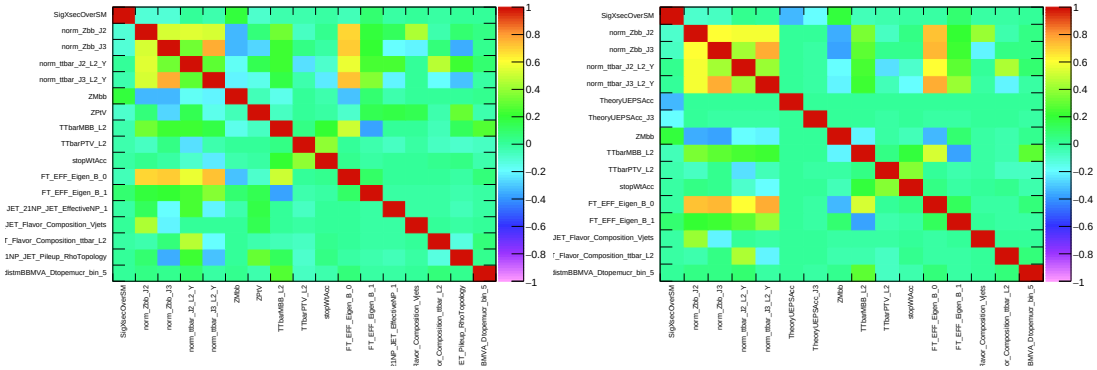
7.8 Nuisance Parameter Correlations

Nuisance parameter correlation matrices (for correlations with magnitude at least 0.25) for all three variable set fits can be found in Figure ?? for Asimov fits and Figure ?? for observed fits.



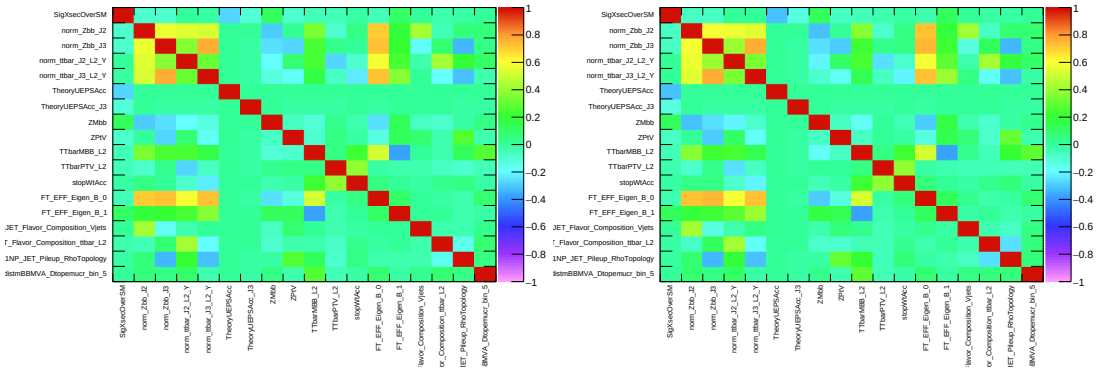
(a) Asimov **(b)** Observed

Figure 7.14: NP correlations for standard variable fits.



(a) Asimov **(b)** Observed

Figure 7.15: NP correlations for standard variable fits.



(a) Asimov **(b)** Observed

Figure 7.16: NP correlations for standard variable fits.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

8

Fit Results

Much has been said Expected and observed sensitivities for the different variable sets may be found in Table 8.1. The RF fits feature the highest expected sensitivities, outperforming the standard set by 3.5% and 3.4% for fits to Asimov and observed datasets, respectively. The LI variable has a lower set than both for expected fits to both Asimov and data with a 6.7% (1.7%) significane than

the standard set for the Asimov (observed) dataset. While the fit using standard variables does have a higher observed significance than both the LI and RF fits, by 2.8% and 8.6%, respectively, these numbers should be viewed in the context of the best fit $\hat{\mu}$ values, discussed below. That is, the standard set may yield the highest sensitivity for this particular dataset, but this is not necessarily the case for any dataset.

	Standard	LI	RF
Expected (Asimov)	2.057	1.919	2.129
Expected (data)	1.76	1.73	1.796
Observed (data)	2.865	2.786	2.62

Table 8.1: Expected (for both data and Asimov) and observed sensitivties for the standard, LI, and RF variable sets.

A summary of fitted signal strengths end errors for both the Asimov and observed datasets are shown in Figure 8.1. For reference, the standalone two lepton fit from the fiducial analysis is given .

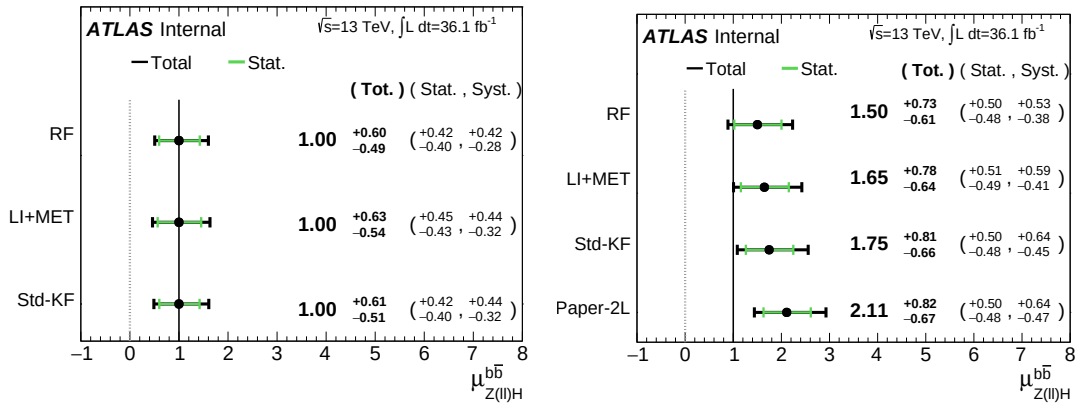


Figure 8.1: $\hat{\mu}$ best fit values and error summary for the standard, LI, and RF variable sets; for reference, the values for the fiducial analysis 2 lepton standalone fit is included under the heading ``Paper-2L.''

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

9

Conclusions

Much has been said

A

Telescoping Jets

Some teljet[23]

B

Micromegas Trigger Misalignment

Regurgitate[24] and also mention[25]

References

[18] (2016). *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*. Technical Report ATL-PHYS-PUB-2016-003, CERN, Geneva.

[19] Ahmadov, F., Alio, L., Allbrooke, B., Bristow, T., Buescher, D., Buzatu, A., Coadou, Y., Debenedetti, C., Enari, Y., Facini, G., Fisher, W., Francavilla, P., Gaycken, G., Gentil, J., Goncalo, R., Gonzalez Parra, G., Grivaz, J., Gwilliam, C., Hageboeck, S., Halladjian, G., Jackson, M., Jamin, D., Jansky, R., Kiuchi, K., Kostyukhin, V., Lohwasser, K., & Lopez Mateos, D, e. a. (2014). *Supporting Document for the Search for the bb decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector*. Technical Report ATL-COM-PHYS-2014-051, CERN, Geneva.

[20] Buckley, A., Butterworth, J., Grellscheid, D., Hoeth, H., Lonnblad, L., Monk, J., Schulz, H., & Siegert, F. (2010). Rivet user manual.

[21] Buzatu, A. & Wang, W. (2016). *Object selections for SM Higgs boson produced in association with a vector boson in which H-gt;bb and V decays leptonically with Run-2 data: Object support note for VH(bb) 2015+2016 dataset publication*. Technical Report ATL-COM-PHYS-2016-1674, CERN, Geneva. This is a support note for the VH(bb) SM publication using the 2015+2016 datasets.

[22] Catani, S., Dokshitzer, Y. L., Seymour, M. H., & Webber, B. R. (1993). Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions. *Nucl. Phys. B*, 406(CERN-TH-6775-93. LU-TP-93-2), 187–224. 38 p.

[23] Chan, S., Huth, J., Lopez Mateos, D., & Mercurio, K. (2015a). *$ZH \rightarrow llb\bar{b}$ Analysis with Telescoping Jets*. Technical Report ATL-PHYS-INT-2015-002, CERN, Geneva.

[24] Chan, S. K.-w., Lopez Mateos, D., & Huth, J. (2015b). *Micromegas Trigger Processor Algorithm Performance in Nominal, Misaligned, and Corrected Misalignment Conditions*. Technical Report ATL-COM-UPGRADE-2015-033, CERN, Geneva.

[25] Clark, B., Lopez Mateos, D., Felt, N., Huth, J., & Oliver, J. (2014). *An Algorithm for Micromegas Segment Reconstruction in the Level-1 Trigger of the New Small Wheel*. Technical Report ATL-UPGRADE-INT-2014-001, CERN, Geneva.

[26] Collaboration, T. A., Aad, G., Abat, E., Abdallah, J., & A A Abdelalim, e. a. (2008). The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08), S08003.

[27] Gleisberg, T., Höche, S., Krauss, F., Schönherr, M., Schumann, S., Siegert, F., & Winter, J. (2009). Event generation with sherpa 1.1. *Journal of High Energy Physics*, 2009(02), 007.

[28] Heinemeyer, S. & Mariotti, e. a. (2013). *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. Technical Report CERN-2013-004. CERN-2013-004, Geneva. Comments: 404 pages, 139 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[29] Lavesson, N. & Lonnblad, L. (2005). W+jets matrix elements and the dipole cascade.

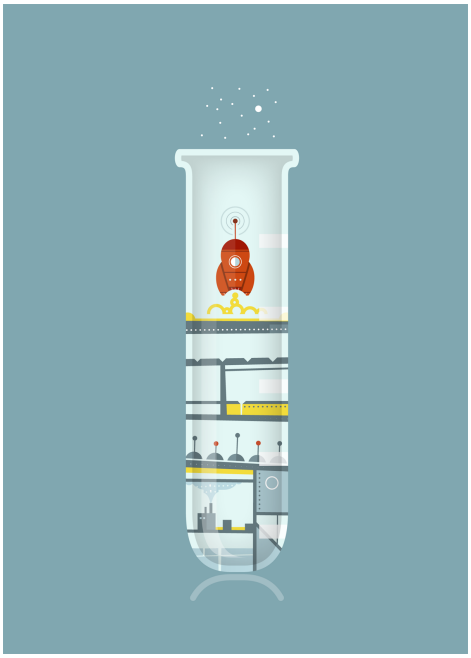
[30] Lonnblad, L. (2001). Correcting the colour-dipole cascade model with fixed order matrix elements.

[31] Masubuchi, T., Benitez, J., Bell, A. S., Argyropoulos, S., Arnold, H., Amaral Coutinho, Y., Sanchez Pineda, A. R., Buzatu, A., Calderini, G., & Chan, Stephen Kam-wah, e. a. (2016). *Search for a Standard Model Higgs boson produced in association with a vector boson and decaying to a pair of b-quarks*. Technical Report ATL-COM-PHYS-2016-1724, CERN, Geneva.

[32] Mellado Garcia, B., Musella, P., Grazzini, M., & Harlander, R. (2016). CERN Report 4: Part I Standard Model Predictions.

[33] Robson, A., Piacquadio, G., & Schopf, E. (2016). *Signal and Background Modelling Studies for the Standard Model VH, $H \to b\bar{b}$ and Related Searches: Modelling support note for VH(bb) 2015+2016 dataset publication*. Technical Report ATL-COM-PHYS-2016-1747, CERN, Geneva. This is a support note for the VH(bb) SM publication using the 2015+2016 datasets.

[34] Stewart, I. W. & Tackmann, F. J. (2011). Theory uncertainties for higgs and other searches using jet bins.



This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, *Science Experiment 02*, was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at github.com/suchow/Dissertate or from its lead author, Jordan Suchow, at suchow@post.harvard.edu.

References

[18] (2016). *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*. Technical Report ATL-PHYS-PUB-2016-003, CERN, Geneva.

[19] Ahmadov, F., Alio, L., Allbrooke, B., Bristow, T., Buescher, D., Buzatu, A., Coadou, Y., Debenedetti, C., Enari, Y., Facini, G., Fisher, W., Francavilla, P., Gaycken, G., Gentil, J., Goncalo, R., Gonzalez Parra, G., Grivaz, J., Gwilliam, C., Hageboeck, S., Halladjian, G., Jackson, M., Jamin, D., Jansky, R., Kiuchi, K., Kostyukhin, V., Lohwasser, K., & Lopez Mateos, D, e. a. (2014). *Supporting Document for the Search for the bb decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector*. Technical Report ATL-COM-PHYS-2014-051, CERN, Geneva.

[20] Buckley, A., Butterworth, J., Grellscheid, D., Hoeth, H., Lonnblad, L., Monk, J., Schulz, H., & Siegert, F. (2010). Rivet user manual.

[21] Buzatu, A. & Wang, W. (2016). *Object selections for SM Higgs boson produced in association with a vector boson in which H-gt;bb and V decays leptonically with Run-2 data: Object support note for VH(bb) 2015+2016 dataset publication*. Technical Report ATL-COM-PHYS-

2016-1674, CERN, Geneva. This is a support note for the VH(bb) SM publication using the 2015+2016 datasets.

[22] Catani, S., Dokshitzer, Y. L., Seymour, M. H., & Webber, B. R. (1993). Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions. *Nucl. Phys. B*, 406(CERN-TH-6775-93. LU-TP-93-2), 187–224. 38 p.

[23] Chan, S., Huth, J., Lopez Mateos, D., & Mercurio, K. (2015a). *$ZH \to llb\bar{b}$ Analysis with Telescoping Jets*. Technical Report ATL-PHYS-INT-2015-002, CERN, Geneva.

[24] Chan, S. K.-w., Lopez Mateos, D., & Huth, J. (2015b). *Micromegas Trigger Processor Algorithm Performance in Nominal, Misaligned, and Corrected Misalignment Conditions*. Technical Report ATL-COM-UPGRADE-2015-033, CERN, Geneva.

[25] Clark, B., Lopez Mateos, D., Felt, N., Huth, J., & Oliver, J. (2014). *An Algorithm for Micromegas Segment Reconstruction in the Level-1 Trigger of the New Small Wheel*. Technical Report ATL-UPGRADE-INT-2014-001, CERN, Geneva.

[26] Collaboration, T. A., Aad, G., Abat, E., Abdallah, J., & A A Abdelalim, e. a. (2008). The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08), S08003.

[27] Gleisberg, T., Höche, S., Krauss, F., Schönherr, M., Schumann, S., Siegert, F., & Winter, J. (2009). Event generation with sherpa 1.1. *Journal of High Energy Physics*, 2009(02), 007.

[28] Heinemeyer, S. & Mariotti, e. a. (2013). *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*.

Technical Report CERN-2013-004. CERN-2013-004, Geneva. Comments: 404 pages, 139 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[29] Lavesson, N. & Lonnblad, L. (2005). W+jets matrix elements and the dipole cascade.

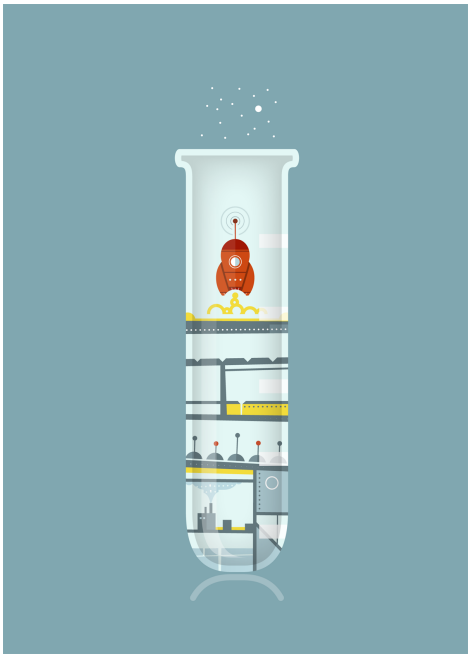
[30] Lonnblad, L. (2001). Correcting the colour-dipole cascade model with fixed order matrix elements.

[31] Masubuchi, T., Benitez, J., Bell, A. S., Argyropoulos, S., Arnold, H., Amaral Coutinho, Y., Sanchez Pineda, A. R., Buzatu, A., Calderini, G., & Chan, Stephen Kam-wah, e. a. (2016). *Search for a Standard Model Higgs boson produced in association with a vector boson and decaying to a pair of b-quarks*. Technical Report ATL-COM-PHYS-2016-1724, CERN, Geneva.

[32] Mellado Garcia, B., Musella, P., Grazzini, M., & Harlander, R. (2016). CERN Report 4: Part I Standard Model Predictions.

[33] Robson, A., Piacquadio, G., & Schopf, E. (2016). *Signal and Background Modelling Studies for the Standard Model VH, $H \to b\bar{b}$ and Related Searches: Modelling support note for VH(bb) 2015+2016 dataset publication*. Technical Report ATL-COM-PHYS-2016-1747, CERN, Geneva. This is a support note for the VH(bb) SM publication using the 2015+2016 datasets.

[34] Stewart, I. W. & Tackmann, F. J. (2011). Theory uncertainties for higgs and other searches using jet bins.



This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface.

The above illustration, *Science Experiment 02*, was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at github.com/suchow/Dissertate or from its lead author, Jordan Suchow, at suchow@post.harvard.edu.