

GENERAL COMMENTS

The manuscript is in excellent typographic shape. Make sure to use WH , ZH , VH , etc., rather than WH , ZH , ..., to preserve proper latex spacing.

On the other hand, there is a serious problem of credibility that should be addressed. One can easily predict that this new analysis of the same data as in a previous publication, claiming to use an improved b-tagging code (line 38, page 2), will generate careful scrutiny because the new limits are surprisingly tighter than the previous ones. Despite the detailed effort made for describing the different event selection and the event migrations among tag categories in the two analyses, the effect will remain rather obscure to the general reader. None of the information given has the capability to quantitatively convince the reader that all is good in our taggers. For example, mentioning that no mismodeling is observed in the NN distributions (line 155) does not tell much. The NN gets its discriminating power from multidimensional correlations among quantities. In those, large mismodelings can happen with no immediately visible effect on the one-dimensional distributions. Event migrations are mentioned, but without any quantitative convincing test that one sees what one expects.

To make the case of a statistical fluke credible, the authors should quantify, using many samples of Monte Carlo simulation, what is the chance for the two analyses of observing such discrepant limits. If this number is $< 3\sigma$, then the reader will have a quantitative proof that your claim of a statistical fluctuation is well-founded. According to standard practice, if it's $> 3\sigma$, the reliability of one or both the b-taggers will have to be discussed.

If one does not feel like revisiting the analysis to this extent, although we propose several line-by-line corrections in these paragraphs, we suggest that one should reduce to a minimum the text between lines 154 and 192, by simply declaring that accurate studies have convinced us that the change in the result is primarily due to the different event selection and splitting into tag categories obtained with Hobit with respect to SecVtx. Even if this short-cut to the problem is adopted, the observed limit from the previous analysis should be reported in the plot of fig 3.

LINE BY LINE

Abstract

Line 3. Missing space between "fb⁻¹" and "at"

Line 5. Suggest making this sentence shorter: "The sensitivity is improved by an average of 14%...". Suggest adding a reference to the "previous analysis" (ref.5)

Lines 8 to 10. Suggest deleting the sentence "The significant reductions... algorithm", which does not belong to the abstract

Page 2

- Line 9.** Suggest "...symmetry breaking generates a massive scalar boson..."
- Line 16.** Suggest "...important to understand its SM or non-SM nature."
- Line 17.** The ATLAS and CMS observations make use of the W+W- channel as well.
- Line 19.** "driven by"? Suggest a more appropriate expression.
- Line 20.** No hyphen after "125". "Higgs boson decaying" (delete "of" after "boson").
- Lines 21-22.** Suggest dropping this insignificant comment.
- Line 25.** Suggest "...with a rate much larger than the rate predicted for Higgs-boson production followed by $H \rightarrow b\bar{b}$ decay."
- Line 26.** Suggest "inclusive" rather than "direct"
- Line 27.** Suggest "...in processes where the SN Higgs..."
- Line 28.** Suggest "...boson V (V = W or Z).", so that one can skip "...where V represents..." on line 35.
- Line 30.** Suggest "larger" rather than "better".
- Lines 31-33.** Suggest advancing this comment to line 24 after "...b-bbar final state.", and to suitably adjust it like "Selecting events...improves...However, at the Tevatron..."
- Line 37.** Suggest "...previous CDF analysis in the $E_{t,miss} \bar{b}+b\bar{b}$ search channel..."

Page 3

- Line 42.** Since this is a PRD-RC and not a PRL, space constraints are less stringent. Consider adding a short description of the main detector components here.
- Line 49.** Suggest: "...data taking if $E_{t,miss}(cal) > 45$ GeV, or $E_{t,miss}(cal) > 35$ GeV and two jets are detected." Suggest mentioning that data are collected in CDF on-line by a multi-channel trigger system [include a ref], of which the two mentioned channels are exploited for this sample.
- Line 50.** One is switching here to off-line selection. This should be mentioned.
- Line 51.** Suggest: "...in jet reconstruction are applied [8]."
- Line 57.** Suggest "and on momenta of jet tracks."
- Line 60.** Suggest "... $|\eta| < 2$ and at least one of them should satisfy $|\eta| < 0.9$ to be covered by the vertex detector acceptance."
- Line 64.** Defining this as the "preselection sample" rather than "preselection region" (here and several times later in the paper) would have been more appropriate.
- Line 65.** Delete "now"
- Line 67.** Suggest "...based on the probability that the jet is originated..."
- Line 68.** Suggest "...tend to cluster at values close to 1, whereas ...likely to populate the -1 region."

Page 4

- Table 1, first line of caption.** "Tag category" in the heading, and move to line start. Suggest "...compared to the previous one [5]"
- Line 72.** "tightly" should be defined as appropriate in a physics paper
- Line 77.** Typo: \rightarrow non-negligible
- Line 80.** Suggest ("electroweak mistags")
- Line 92.** "jet direction". "...tends to take larger values for jets..."
- Line 93.** "b quark".

Page 5

Line 104. Sphericity is defined here as an overall event shape. In the caption to fig. 1 (page7) it is “sphericity of the jets in the event”. Neither of these definitions is clear, and they sound inconsistent with each other. Please give a single and more precise definition.

Line 107. “two-tired”? Consider “staged”.

Line 120. Suggest to delete "It was found that".

Line 125. Suggest “...is validated with...”

Page 6

Tab II caption: "amongst" → "among"

Line 128. "As in the previous analysis" is unnecessary

Line 133. Suggest “...with an r.m.s. width equal...”

Lines 134-135. Delete colon after "from". Suggest: “...normalization of the contributions from V and heavy flavor backgrounds,...”, “..difference in b-tagging efficiencies...”. Also, check spacing in percentage ranges → "simulation (8—16%), uncertainty on the top (6.5—10%)..." [until line 137]

Page 8

Caption to fig. 2 “...factor of ten...”

Line 141. Please add more details on the statistical procedure used to set Bayesian limits. In particular one should specify which priors are used for the signal rate and for all nuisance parameters (flat or else? truncated at zero or else?)

Page 9

Line 144, 145. Suggest → “...given in Fig. 3 and Table III, showing a 14% median improvement in expected...”.

Line 154. "we note that" is not necessary.

Line 156. “...substantial event migrations...”

Lines 158 to 160. Suggest inverting some words: “...coming from events that were not present in the signal region of the previous analysis, that are now accepted by using the updated b-tagging algorithm.”

Line 168 and 169. Suggest “the previous one.”

Line 169. “...significant event migrations...”

Line 170. Suggest removing the first “individual”.

Line 174. “one standard deviation”

Page 10

Line 177. “...largest signal-to-background ratios, thus...”

Line 187. Suggest “...the most signal-rich bins in the previous analysis....”

Line 189. “...roughly one-standard deviation..”

Line 192. “...at other assumed values of the Higgs boson mass as well.”

Line 195. The need of toys to support this statement with quantitative evidence is particularly evident at this point.

Line 197. “...migrations from the SecVtx and JetProb tag categories of the previous analysis to...”

Line 203. "on average" (hyphen not needed)

Line 204. "has been investigated" is unnecessary

Page 11

Fig. 3 inset. "Expected "→"expected "

Line 206. "supersede"

Page 12

Line 225. Please use journal references for these two papers.

Line 227. Please italicize "*et al.*"

Lines 230, 231. Include Collaborations.

Line 236. Please italicize "*et al.*"

Page 13.

Line 251. Use journal reference.

Line 252. Missing "arXiv:1205.1812."

Line 258. Delete "(2012)"