

General: It is not fully understand why the analysis is split in 7 and 8 TeV. In particular concerning the choice of the binning and multivariate optimisation.

L2: "An observation of ...". That is better than "Any observation".

You actually say later that a branching ration of 10^{-40} is still within the SM (w/ massive ν).

L14: Write "\taum" instead of "tau". To be consistent with the rest of the text.

L13-25: This paragraph contains a lot of "LHCb". By the end the reader surely knows who did the measurement. Suggest to remove the occurrences in L18 and L21.

L18: Add comma after CL.

L22: remove hyphen in cross section. Although the EB guidelines may give another suggestion, this word is not hyphenised in any dictionary (included those suggested by the EB).

L25: Add "at 7 TeV." at the end of this sentence.

L29-30: Include term "blind analysis" : "To avoid potential bias a blind analysis is performed by initially excluding $\mu\text{-}\mu\text{+}\mu\text{-}$ candidates"

L31+94: "expected mass resolution". This sounds a bit like we are still trying to understand our detector. But in reality, we should talk about the mass resolution of our detector (which of course depends on the momenta of the final state particles and the Q -value).

Not about mass resolution of a specific decay. Suggest "intrinsic mass resolution" or similar.

L61: "In THE simulation ..." Which one is that? This is the first mentioning of a simulation. Add a sentence to explain the purpose of the simulation.

L69: there is no such thing as a "b hadron". Either a "b-quark" or a "hadron containing a b-quark". Also in L.71.

L86: "The decay $D_s^- \rightarrow \eta \dots$ ". Start a new paragraph before this sentence.

L93: "The signal region ...". Start a new paragraph before this sentence.

L112: "An ensemble selected ... is used." This is too much

information in one sentence. Better to explain a little bit more. Ensemble selection is new to me. If you keep using it like this, please add a hyphen "ensemble-selected".

L127+134: You say that the PDF is calibrated. But which PDF? This is not mentioned before. Don't you mean that the classifiers are calibrated such that they return a probability or likelihood? In that case the mentioning of likelihood in L136 also makes sense (now it does not).

L139: ".... the signal BF assumption." Explain, lingo.

L136 & Fig. 1: Why the uneven bin widths, rather than equal-width ones? Optimization of the separation power is claimed, but do these binwidth have any significant effect?

L152: why meson. Shouldn't it be heavy hadron instead, to not exclude baryons?

L154: "There IS a large number of ...". Not are. Number is singular!

L160: "... has a mass distribution CONSISTENT with ...": as the process is different, it cannot be "consistent". Perhaps "... has a SIMILAR mass distribution ..."

L165-169: "... for each bin ... using an extended, unbinned" -> unfortunate choice of words (too many different "bins"), consider rephrasing.

L166: "for each bin in ..., ... and MASS" a fit is made to the spectra outside the signal windows. This is contradictory. Isn't the (shape of) the mass-spectrum outside the signal region fitted using an unbinned likelihood for each 2D bin in MPID and M3body?

L169: "the SMALL differences" -> "the differences". BTW, differences in WHAT? So even better "the differences in the estimated number of background events in the signal region are taken as systematic uncertainties." Perhaps even including a statement somewhere that this difference is used to estimate the a priori unknown shape of the background.

L183 "BR(Ds->Phi(K+K-)pi) is taken from the BaBar amplitude analysis [25], which considers only the Phi->K+K- resonant part of the Ds decay."

Is the second half of the sentence indeed needed? We would skip the second half of the sentence, as the first half of the sentence already

implies that this refers to the resonant $\Phi \rightarrow K+K^-$ part.

L184-186: This sounds like $KK\pi$ coming via ϕ is a bad thing, while explicitly normalising to $\phi(\mu\mu)\pi$. So the ϕ is what is needed.
What is precisely meant? Please explain.

L184 "This is motivated by the negligible contribution of non-resonant $D_s \rightarrow \mu\mu\pi$ events seen in the data."

Q1: What does "This" refer to? If you refer to the calculation of $BR(D_s \rightarrow \Phi(\mu\mu)\pi)$, then we propose a rephrase: "The estimate of $BR(D_s \rightarrow \mu\mu\pi)$ using $BR(D_s \rightarrow \Phi(K+K^-)\pi)$ is more precise than the direct measurement of $BR(D_s \rightarrow \mu\mu\pi)$."

L185: Could you explain how the non-resonant $D_s \rightarrow \mu\mu\pi$ events would affect $BR(D_s \rightarrow \Phi(\mu\mu)\pi)$?

L206: is the effect of eliminating the lowest likelihood bins taken into account (a la L200-201)?

L207: "The YIELD of .. candidates in THE data, N_{cal} , ... IS determined ..." (not plural, add the).

L208: "The VARIATION in the YIELD if ... IS considered ..." (singular)

L209: variation of the widths of the Gaussian components: by how much are these widths varied and what is this variation motivated by?

L213-216: The text does not make clear whether the data of the individual bins in M3Body and MPID are used individually or combined (excluding the lowest likelihood bins). If so, some measure to substantiate the claim that "no significant evidence for an excess is observed" could/should be given. E.g. for 7TeV the bin $[0.40-0.45, 0.46-0.54]$ expects 2.89 ± 0.63 and sees 6 and $[0.45-0.54, 0.54-0.65]$ expects 2.83 ± 0.63 and sees 8! Of the 25 measurements, several seems rather high, which in the combination are offset by a fair number of low or very low number of observed counts.
For 8TeV there is indeed little to worry. Please comment on the consistency of "no excess" per bin.

L217: "estimates". Not estimations.

L230: "In summary, the previous LHCb limits ..."

L231: Remove "3.0 fb⁻¹ of". Jargon.

L232: "No evidence for any signal is found" -> "No evidence for the LFV decay $\tau \rightarrow 3\mu$ is found."

L232: "The present limits supersede" -> There is only one, unless the 90% and 95% CL. limits are counted as two ... Replace by "The presented limits"

L233: "B-factories". B in italic.

L233: allow -> give

L233: "... should allow improved ..." -> "... improve the constraints placed on"

L234: Remove "to be placed"

Fig. 6/supplementary material: What is the color coding?