Measurement of the relative branching fraction of $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-$ and $\Lambda_b^0 \to \Lambda_c^+ D_s^{*-}$ decays – WG review round 1

Based on v1r0, Nov. 23 2021.

Comments from Liupan An

Q Eq.2 It may help to add a clear definition of "f(Pc)".

A Done.

Q l.59-61 Naively, I would expect that a heavier decay product means a larger energy scale for its production, then a better factorization. Could you explain it a bit more?

A Rephrased. The line of arguments is given in Ref. [1] (Sec. C).

Q l.169-184 It is not clear what cuts are used for the misID veto. I suppose they are combination of mass window and PID cuts? Could you add a table for it?

A Added.

Q l.189 How can the cuts be looser than the previous step? Are the cuts in previous iteration used only for determination of misID veto cuts?

A Yes, this refers to "historical" iterations. The cuts are looser, because we added outlier cuts and vetoes. This sentence is removed now.

Q l.206-207 I don't see the mass cuts in Table 9.

A This seems to be leftover. Added another sentence to clarify that the cut is selection-dependent.

Q l.394-397 Is there any quantitative estimation of how larger these contributions are?

A Not that I'm aware of. I would be surprised if the $\Delta S = 2$ decays have branching fractions above 10^{-7}. For the no charm decay the suppression from the CKM matrix should be in the order of 2000 w.r.t. $\Lambda_c^+ D_s^-$; plus the suppression from the fact that $pK^-\pi^+$ and $K^-K^+\pi^-$ have to be within the charm mass windows, which is easily a factor 100.

Q l.408-409 This approach looks not very clear to me. When determining M and mu using GetMean and Landau, did you use the data sample directly? Will it be affected by backgrounds then?

A Rephrased. It indeed sounded like an important procedure that is necessary to follow, where it was instead just a method to get a good initial estimate for the parameter.

Q l.411-412 How is the uncertainty on lambda considered in the final uncertainty?

A Rephrased. See answer above.

Q P.22 I'm trying to understand the whole procedure. The analytic shape is chosen to be Box, Hill and Horns. Then the simulation studies are just to validate this choice? Did you fix anything wrt MC in the data sample fit?

A Yes, eventually all the partially reconstructed quasi-two-body decays could be described by these 3 shapes including convolution. In the final fit, we only fixed the D_s mass and the D_s^* to D_s mass difference to their PDG values, which determine the kinematic endpoints. In a fit-variation to obtain a systematic uncertainty we also fix the D_s^* branching fractions.

Q l.463 Did you check if the efficiency-drop can happen also for other modes? In the following studies of some partial background, only fast simulation is used.

A The other modes extend far below the chosen mass range, and the efficiency drop should be a kinematically induced acceptance effect. In this region (the upper mass tail) the effect should be negligible.

Q l.516-527 Similar question as before, is this to validate the choice of Bukin function? Is any parameter fixed for data fit? If so, as shown in Fig.8e, the AmpGen model looks to be not in perfect agreement with the data. Is it for sure the effect is negligibly small?

A Yes, this shows that the Bukin p.d.f. can fit feeddown shapes of PHSP (a,c) $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{*0} K^-$ and also a the shapes that have used a more realistic amplitude model (b,d). Please note the Fig. 8(e) is not a fit. It shows the amplitude model described in the text that we used to see how much a PHSP model deviates from a more realistic one. Now, the PHSP model is replaced by a model simpler than the AmpGen one, but much more realistic than PHSP. From that model, that was used in full simulation, we get KDE p.d.f.s to describe $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{*0} K^-$. This is a change of fit strategy, as it could be shown that the KDEs perform much better than the Bukin p.d.f.

Q l.628-636 In methods a and b where the pi0/gamma ratio is not fixed, do they give results consistent with the known value? In practice, how do you combined pi0 and gamma mode to the total D* yield?

A So far, we calculated back the efficiency correction and got results that agree with the PDG. As the new simulation samples no longer force the π^0/γ into acceptance, we get an better handle on this. Now, we can directly fit the ratio, assuming that there are no residual efficiency-effects (they should be negligibly small). We quote the ratios in the fit tables of the new note.

Q l.806-807 How is this "50 %" obtained? Did I miss anything here?

A This was a conservative (lazy) estimate to take into account that we didn't do our work as careful as it was done in the dedicated measurement. We have now removed the D^0 BDT and argue that the uncertainty introduced by the Λ_c^+ BDT is negligible. We are confident that the calibration is valid, as we have cross-checked it with alternative selection strategies (either using the same *D-from-B* BDT cuts in signal and reference channel, or by removing *D-from-B* BDT cuts alltogether).

Q l.834-840 The procedure to get the final numbers can be explained more clearly.

A Done.

Q l.841-843 Have you considered the trigger efficiency uncertainty?

A No, trigger selections are exactly the same for signal and reference channel. Any residual difference should be small and reasonably well described by simulation. Uncertainties that come with this are expected to be much below our sensitivity.

Comments from Lorenzo Capriotti

General comments

Q I see that some of the selection procedures are not described in detail in the note but rather on Marian's thesis. I think they should be at least summarised in the note, with the most relevant numerical information and plots, just to have most material in the same document. It is fine to put a reference to the thesis but adding information in the note would allow people to read it without having to dig into other documents unless they want to explore specific details.

A We rephrased the selection section, trying to flesh out the individual steps, and add information on the cut scanning procedure.

Q The choice of the cut values for the four different selections is not very well justified.

A We think they are well motivated, but not written up clearly enough. We added more information and moved the section.

Q In general I found the note more descriptive than informative. I think that additional, quantitative information is often required to ensure a high-quality review.

A We have now added many details and hope that we found a better compromise. We still provide several forward pointers to the code, and prepared it such that the entire analysis can be run with a push of the button. Therefore, we left out some of the details that in our eyes were not worth mentioning. We see that the style of the previous iteration is not common in ANA-notes, and therefore switched to passive voice, added details and moved descriptive parts to the appendix.

Detailed comments:

Q l.156 please list also the fiducial cuts in table 6. Furthermore, it is not clear to me if you cut on the D-from-B BDTs or if you use it only in the 3D scan (I suspect the latter but I think it is worth specifying that).

A The fiducial cuts were listed in Table 7, as 6 only lists the cuts applied in DaVinci. The BDT cuts are listed in Table 9.

Q l.165 it would be useful to add details about this cut optimisation procedure in the note. For instance which figure of merit is used, the cut values for the three projections, the signal efficiency and background rejection, etc.

A This part is rephrased. See also our answer to Zan's comment on adding plots: "These steps would have to be re-done by hand, and we never looked at such plots in the first place. By inspecting the stdout of the optimization, we can say that the FoM gradients in the optimal region are relatively flat. This is why we ran the optimization only in the very beginning of the analysis and fixed these cuts early on. We don't think that the information contained in such plots justifies the effort of producing them."

Q l.169 could you please give more details on how the vetoes are applied?

A Done.

Q l.188 is this cut applied on top of the previous optimised cut or is it a whole new cut? How do you optimise the previous cut?

A This was not very clear. Rephrased. There is no additional optimization step.

Q l.199 from Tab.9 it looks like the PID variables are not always tightened. Also, what do you mean by "aligned" and why would that be a good selection strategy? From Tab.9 I see ProbNNk > 0.2 for the K- in LcDK and Prod(ProbNN)>0.02 for the Ds.

A They are tightened w.r.t. selection A. Changed "aligned" to equalized – so that their efficiencies cancel in the ratio (to first order).

Q Tabs.7-8-9 apart from the vague description given in ll.188-199, how are the specific cut values for each selection decided? Are they the result of some sort of optimisation? Could you please provide more details?

A Done; added several statements in the text.

Q Sec.2.3 can you please add LcDK, LcDs, Lc, Ds and D0 mass plots for the four selection strategies? Better if you could use the same frame for each channel. So, for instance, a plot of the Lc mass with 4 histograms corresponding to the selections, and so on.

A Good idea. Added.

Q Fig.5 please add Dalitz plots of signal MC and selected data. This will allow to check the Dalitz boundaries having included smearing from resolution and efficiencies.

A We use square Dalitz plots, and show them in Sec. 4.2.

Q Sec.3.1-3.3 you make a lot of (reasonable) assumptions in these sections. You could validate at least some of them by quickly generating some million events with RapidSim for each assumption. This will not take into account the resonant structure but will give an idea on where the mass boundaries are after resolution.

A Does this still apply after the update? Could you please be more specific?

Q l.369 do you take into account Lb->Lc+K-pi+pi- with pi->K misID? In general there are no misID studies. Why?

A We have done extensive misID studies, see the (amount of) vetos we have applied. The background you mention has been studied, but is not observed, as the $\overline{D}^0$ or D_s^- selection, including mass window, is very efficienct at suppressing it.

Q l.413 I think there might be a logical issue here. Eq. 6 is obtained in the zeta->0 limit. You first constrain lambda using Eq. 6 and then verify that zeta is compatible with zero; then, you fix zeta = 0 in the fit. But you cannot fix zeta taking as justification the fact that you measure it to be zero, as that measurement might be biased by your choice of lambda. Did I understand correctly what you did?

A It is not as bad as it sounds. All studies have first been done with $\Lambda_b^0 \to \Lambda_c^+ D_s^-$, where it is possible to leave more parameters floating, and still get a fast converging fit. Leaving both λ and ζ free there, we recover the case we describe. Fixing $\zeta = 0$ gives more numerical stability (see description in RooFit).

Q Sec.3.5 despite the exhaustive digression, this section lacks the most important information, i.e. how do you exactly obtain the combination of hills, boxes and horns for each channel and if you use or not your MC as a signal proxy. Is the limited MC size taken into account? Please add more quantitative details.

A We have moved the physics discussion to an Appendix, and hope the rest gives all information needed.

Q l.572 what about the partially-reconstructed components? Do you mean that they are included in what you call here "signal"?

A We mean the $\Lambda_c^+ \to pK^-\pi^+$ signal here. The goal is to model $\Lambda_b^0 \to \Lambda_c^+ K^+\pi^-K^-$ in 3D; so we need the $\Lambda_c^+ \to pK^-\pi^+$ signal, a phase-space model in $M_{\rm inv}(K^-K^+\pi^-)$ and a Λ_b^0 "peak" that takes the kinematic constraint from $\overline{D}^0$ or D_s^- into account.

Q l.637 so, just to be clear, you are saying that selection A with fit c will be the baseline measurement and all the other combinations are used to estimate the systematic uncertainties?

A Yes and no. Selection **A** was used to get the final result; Fit **c** is the one we showed, but in the end fits **a** and **b** contributed with the same weight to the final result as fit **c**. In the new iteration, we have one baseline fit and selection, and calculate systematic uncertainties from other fits, while varying the selection is taken as robustness study.

Q Eq.8 please add also the individual efficiencies (i.e. before the ratio). I find it curious that they are all above unity. Do you understand why?

A With the new MC they are probably closer to what you would expect, as they are all using the same generator level selection strategy.

Q Eq.9 please add also the individual efficiencies.

A Can you give more information on what one can learn from individual efficiencies? Quoting rough numbers for $\Lambda_b^0 \to \Lambda_c^+ D_s^-$ could be a compromise.

Q l.843 how small?

A In the order of several permille relative uncertainty (added to the text).

Q Sec.5 the treatment of systematic uncertainties needs to be expanded a bit. It would be nice to have a summary table of all sources of systematics with their value. Also, it is not clear to me how you take into account the various non-baseline selections and strategies in the evaluation of systematics, how you combine the different sources and how you ckeck for correlations.

A Done.

Comments from Zan Ren

General comments

Q My general layout comments is that please follow the standard LHCb template for document writing, this will be a great helpful to the increasement of readability. For example, Figure captions go below the figure, table captions go above the table; Figures and tables should normally be placed so that they appear on the same page as their first reference, but at the top or bottom of the page, ... Lots of figures in this version of ANA-note are placed far from where they are first time mentioned, and there are also some figures forgot to be mentioned in the main text. "Table" and "Appendix" are not abbreviated.

A Mostly done. Final layout (moving plots and tables) will be done as we converge to a more final version.

Detailed comments:

Q Abstract: Since these decay modes has been studied (especially, been observed) in a previous published Ph.D thesis (also written by Marian Stahl in year 2018, or Ref [42] in your document) using LHCb Run1 data, I am not sure here statement like "... ... are observed for the first time" is still suitable or not since this work focused on Run2 data, of which is collected later than Run1 period. If the results in Ref [42] has been published, you should not use some phrases like "first observation" again; if not, I suggest you can publish both Run1 and Run2 results in the future and then mentioned that this is the first observation of corresponding decay modes. Then, could you please expand this section, and make it look like a formal "abstract"? For example, add several sentences to the dataset used in this analysis and summarise the final results of your measurements.

A Added new abstract.

Q L2: add "decay" after the 2nd decay chain.

Q L4: It there any reason the word "pentaquark" shown in italic style?

Q L17-18: "QCD" is not an abbr for "strong interaction" but quantum chromodynamics. So, please rephrase here or just remove the text in bracket.

Q L19: 1970s and 1980s

Q L22: cf. Ref. [1]

Q L26: please add a hyphen between "model" and "independent"

Q L30: please add a comma before the second "and"

Q L31: remove extra dot after "cf."

Q L40: "This analysis" -> "The amplitude analysis"

Q L46: The word "Table" should not be abbreviated. Please check and correct similar situations in the whole document.

Q L49: add the corresponding symbol between "events" and "in"

Q L54: "quasi-two-body decay", you define this jargon at the footnote of line 428, please move the relevant footnote here.

Table 1,2,3,4: please located titles and captions above corresponding tables.

A All done

Q L92: Does cleanliness means purity?

A Yes, rephrased.

QL114: collected by LHCb detector

QL122: The first level software

QL124-126: could you please add a reference to this HLT1 line, or make sure that you have list all the selection criteria in this line

A Done

Q Table 6: The description of BPVVDZ seems incomplete (in z direction?)

A should be fine as it describes to what the cut is applied, not what it actually means

Q L157-115: Could you please add a short paragraph to briefly introduce the BDT selection part (instead of just leave a Ref.[C21] here), including the type of the classifier, training samples, target function for optimisation, working point, and so on. This can provide minimal information to the readers in the main text without jumping to the reference.
"background subtracted data " -> background-subtracted data

A Done

Q L180-182: Could you please provide some figures here to show the angle distribution between two final charged tracks?

A Added to appendix

Q Fig.2 : sub-figure (l) is missed, but it is mentioned in the caption

A Done

ig.2 caption : you say these plots are “after all other selection steps”, it that means the selection in Table 7,8,9 has been required? If yes, could you please change the narrative order of the relevant cuts in the text; if not, maybe you need to rephrase here

A Tried to clarify.

Q L183-184: I am confused by this sentence. If my guess is right, referring to my previous question, it seems you first perform the mis-ID and clone-track vetoing and later add the cuts in Table 7&8, right? Furthermore, why these vetoes/cuts can help to “facilitate the efficiency correction procedure”?

A Rephrased.

73, and L179: I think you set some window according to the width of relevant mass peak after swapping the mass hypothesis of some particles. If my understanding is right, could you please provide the width and centre-value of corresponding mass windows? If not, you also need to mentioned how do you remove the mis-ID backgrounds in the text.

A Added a table.

Q Fig 2: How do you obtain the black (labelled as “without veto”), blue (labelled as “with veto”), purple(lablled as “full selection”), and orange (labelled as “vetoed”) histogram? Since the x-axis title contains notation like $M_{\text{inv}}([H_1 \to H_2]H_3)$, are all these four histograms are plotted using the dataset after the $[H_1 \to H_2]$ mass-swap ? And have you applied sWeight (derived from the 1D fit) on the events?

A Rephrased in the caption and text. The definition of the variable is always the same. There is no sWeighting involved. The full sample used for 1D fit without the multiple candidate rejection is shown.

Q L188-191: Could you provide the $\frac{S}{\sqrt{S+B}} \times \frac{S}{S+B}$ distributions w.r.t. corresponding variables? I think this is an optimisation in 4-dim. space (i.e. LcBDT , D0BDT or ProdProbNN , ProbNNk , BPVIPCH2), maybe you can show the projection on some dimensions.

A This part is rephrased. These steps would have to be re-done by hand, and we never looked at such plots in the first place. By inspecting the stdout of the optimization, we can say that the FoM gradients in the optimal region are relatively flat. This is why we ran the optimization only in the very beginning of the analysis and fixed these cuts early on.
We don't think that the information contained in such plots justifies the effort of producing them.

Q L196: You say "without using PID variables", but from Table 7 seems you still required PID cuts, and its even more tighter than strategy A!

A Vetoes are done without PID variables. We need to use a proton `ProbNN` cut (that was missing in Table 9) and a bachelor K `ProbNN` cut, otherwise the bkg is too large

Q L198-199: The BDT cuts are also tightened in Selection D, please mention it in the text.

Q L204 : remaning -> remaining

Q Table 7: Please avoid left the title-row alone on the bottom of one page. Same suggestion for all other tables in this note.

A All done.

Q L207: Table 9 does not include any mass cuts of charm candidates! So what's the exactly mass region of charm particles for these two decay modes? And, is the mass cut then included in all A to D strategies?

A Added a sentence to the text. The reference to Table 9 was a leftover.

Q L211-212 : The sentence "where a score ... are removed" is ungrammatical. Maybe you mean that candidates which have a score less than the maximum one among all the multiple candidate partners in one single event are removed.

Q Fig.3 : Does "removed by PID" contains the removal from BDT cuts in Table 9? If yes, please comment on that in the caption; If not, could you please add the histogram for "removed by BDT" ?

Q Fig.3 : This figure is not mentioned in this section, please mentioned it at a suitable position.

A All done.

Q Table 10 : It is reported that Sim09d/e/f has a bug on multiplicity due to handling of wide resonances (D^{*0} is also affected), please see https://twiki.cern.ch/twiki/bin/view/LHCb/Sim09Differences, I promise that is not of great problematic, but I still suggest to have a look and check.

A We have new MC now.

Q L229: Hesse -> HESSE or MIGRAD -> Migrad

Q L253: The footnote should be : $|j_1 - j_2| \leq J \leq j_1 + j_2$ and can you change the coupling form to relative orbital angular momentum number L? This will be more convenient and intuitive to see it is decay via S-,P-, or D- waves.

Q L282,288 : eq. -> Eq.

Q L317: $D^*_{s0}(2317)^-$, $D_{s1}(2460)^-$

Q L374: Same to my previous questions, Table 9 does not contain any mass window cuts...

A All done.

Q L375-376: Your statement here seems to be inconsistent. From line 238-240 you say that the mass windows is determined after 3D fit, but here you say, for example, m_Kpi should be in [1822, 1912] MeV. And a further question based on line 238-240 is, do you mean you will choose some tighter mass windows (on m_Kpi, mKKpi, mp_Kpi) according to the 3D fit results than the mass windows you mistakenly thought was mentioned in Table 9?

A Apart from referring to Table 9, all should be consistent. For the 3D fit, we need some initial mass window (or call it fit range). Then, the 3D fit is carried out, and the cumulative distribution function of the signal allows us to determine the region that contains 95 % of signal events. This part is re-written to make this more clear.

Q L382 : Maybe you can add ",to a certain degree," between the word "are" and "counterbalanced". The word "counterbalanced" is too strong, and might contradicts the next sentence.

A Rephrased.

Q L384-385 : Just to confirm, is the sentence "the suppression from xxx fractions" a description for single-charm/charmless decay?

A This only concerns $\Lambda_b^0 \to \Lambda_c^+ K^+ \pi^- K^-$

Q L399-412 : Could you please give a formal formula of 2-side Hypatia p.d.f.? And then tell the reader the corresponding symbol of a parameter (i.e. σ) when it is mentioned in the text.

A Done.

Q L408-409: Could you please provide the results (as an appendix) for getting the mode value $\mu(p_X^2)$ by fitting a Landau p.d.f.?

A This is a rule of thumb calculation that is not worth documenting in greater detail. The description is somewhat outdated and contains some confusing statements. Eventually, it only allowed us to get a first estimate in which region we can expect the lambda parameter to be. When doing 1D fits of the LcDs or LcD0K signals later, we let this parameter floating and saw that the fit does not depend much on whether we vary lambda within a couple of sigma. We will write this up a bit more comprehensive.

Q L424: the lower script should be italic Λ_c^{*+}

Q L426: remove the extra space between Eq. and (5)

Q L429: "Hill and Horns" -> "Hill, and Horns"

Q L469: "e.g. [51]" -> "e.g. in Ref.[51]"

A All done.

Q Fig. 7: For the Fig7©, if the dataset is from a full simulation, could you add more details about it, since in Sec 2.4 you only provide the Event number is 15496200 ? Are the ratio of $\frac{D_s\gamma}{D_s\pi^0}$ event number fixed to PDG branching fractions? Maybe you need to attached the decfile (especially for 15496200), this will be more clear what process (and their fraction) is exactly simulated.

A Cited dec file. Updated description in Sec. 2.4

Fig. 8 ,Fig.9 : For the inv. mass for fitting, do you use reconstructed mass or M(XX) computed from MC-truth information (i.e. X_TRUE_PX,x_TRUEPY,...X_TRUE_PZ,X_TRUE_E)? Especially for those marked as "LHCb full simulation".

A We use the reconstructed mass.

Q L544: Remove the Ref. [47] (this has been reference before)

A Prefer to keep as is.

Q L570: Please correct $M_{\text{inv}}(K^-)K^+\pi^-$.

A Done.

Q L571-L572: The equation number is missed, and seems that the is a pair of strange usage of bracket $\mu(...)$ at the numerator. If you want to say that the parameter μ in the Gaussian is a function of $M_{\text{inv}}(KK\pi)$, you can just type $\mu = \mu_{\Lambda_b^0} + \rho(M_{\text{inv}}(KK\pi) - \mu_{D_s})$.

A This is on purpose as we don't reference to the equation in the text. The rest is re-written.

Q L571-575: Maybe this is a misunderstanding, but it is deeply incorrect for me to model the single-charm decay in 3D fit like this, since you say (in line 563) while calculating the $M(\Lambda_b^0)$, a mass constrain to m(charm) is required. This is finished by setting the m(charm)=constant in a DecayTreeFitter I guess. And in Figs 10(b) and 10(f), if the clear mass-correlation between m(charm) and $M(\Lambda_b^0)$ is seen, the must derived without any mass constraint to m_charm (I denoted this as $M_{\rm reco}(\Lambda_b^0)$), only in this case (i.e. without mass constraint), you can describe the single-charm (i.e., only Lc, no D) decay by a D mass-dependent mean in signal shape for $M_{\rm reco}(\Lambda_b^0)$, a signal shape in $M_{\rm reco}(\Lambda_c^+)$, and an exponential bkg shape in $M_{\rm reco}(D)$. If you use a 3D space of $M_{\rm Dmassconst}(\Lambda_b^0) \otimes M_{\rm reco}(D) \otimes M_{\rm reco}(\Lambda_c^+)$, the line shape for signal charm decay in $M_{\rm Dmassconst}(\Lambda_b^0)$ should never be described as a mean-dependent Gaussian, and the exactly shape need to be derived from corresponding MC with D-mass constraint just like what you have done for the 1D fit.

A Indeed, there seems to be a misunderstanding. We don't use DTF. The passage is rephrased to make clear how $M_{\rm inv}(\Lambda_c^+ D_s^-)$ is defined and how this is connected to ρ.

Q Fig 10: For sub-figures ©(d)(g)(h), what exactly fit functions are used when you derive the blue curve marked as "Approximation"? Could you show some details for this fit as an appendix?

A We forgot to describe this p.d.f. in the text and added it.

Q Fig 10: Correlated to my previous question, is sub-figures (a)(b)(e)(f) derived from real data or full simulation or fast simulation? And, for the x-axis variable, Lambda_b mass, it there a mass constraint on m(charm)? And could you explain that why the mass distribution for the single-charm decay looks so strange (a too-wide, asymmetrical plateau), since it do have a real Λ_b^0, so the line shape I expect still to be a peaklike (e.g. Gaussian, CB, or Hypatia). Sorry for bring too much questions on this part, but it is really so different from my previous practical experience.

A The shape is a consequence of our definition of the Lb mass. With a DTF, you would not be able to "rotate" the single-charm component as we did. In hindsight, this rotation gives us a better handle on the single charm component. (initially it was a workaround for a bug in `Phys` that didn't allow particle substitutions in the DTF).

Q L603 : D -> D

A Done.

Q L606-610 : It is OK to take shapes for partially reco components from MC and fixed some parameters that has large correlations. However, from previous sections, there are several types of MC are used, some of them are officially generated while the others are produced by fast-simulation or gen-level simulation. This will cause the mass resolution effect not be well considered, especially for the fast-simulation/gen-level simulation. If both the MC are gen-level, the detector resolution must be take into consideration, furthermore, the partially reco. components usually show a wider distribution in mass, the resolution might also varies as the Q-value or the mass difference $M_{\rm reco}(\Lambda_b^0) - \sum_{\rm finaltrackX} m_{PDG}(X)$. If both of them are from full-simulation, its OK for me.

A The resolution is a floating parameter in our fits and shared between the partially reconstructed components. For generator-level simulation, we use a numerical convolution with a Gaussian resolution function. The uncertainty that comes with fixing parameters is covered by the systematic uncertainties coming from different fit models.

Q L622: So in both scenario b&c, the exp bkg is change to linear, right?

A Made it explicit in the text.

Q L624-625: From this sentence and fit results in Table 9, I guess you only fix the branching fraction of $\mathcal{B}(D_s^{*-} \to D_s^- \gamma)$ to PDG value in scenario c, right? If yes, please rephrase here.

A Correct. But this is what the text describes already.

Q Table 9 : The stat. uncertainty for $\mathcal{B}(D_s^{*-} \to D_s^- \pi^0)$ is 0.0, really?

A Yes, the parameter has been fixed. I built in an if else statement to not print the error.

Q Tabel 9 : Could you please provide the entire fitting results, including all free parameters and their uncertainties?

A We are doing several hundred fits in the analysis. This would certainly be too verbose. Can you specify what you would like to see explicitly? If you would really like to see them all, the analysis is fully reproducible with a single command after following the setup instruction in the repository. We invite you to run it yourself and take a closer look. I'm also happy to help in case there are issues.

Q Fig 11 : This figure is not mentioned in the text.

A Fixed.

Q Fig 11: For the "Comb. Lc" and "Comb. D" components, do they share same parameter (e.g, the exponential coeff) in the background line shape of $M(\Lambda_b^0)$? And, similar question, for the red filled component (i.e. single charm), in the figure © $M(KK\pi)$ or (f), do they share same parameter with "Comb. D" or "Comb. bkg components?

A Good point. Added more info to the text.

Q Fig 12: This figure is also not mentioned in the text.

A It is in line 637.

Q L669: Do you mean you add some weights for each year's gen-level efficiency according the corresponding integrated lumi.? If yes, could you please add a table to show what the exactly number of the weight is?

A Table added.

Q L659-700: Could you please provide the plots for before-corr/after-corr comparison on the PID variables?

A Added to Appendix.

Q L680-681: In Section 2, you listed the L0 triggers used in this analysis, so do you use both L0_TIS and L0_TOS , or just required one of them? If L0_TIS is used, even though the name of selected L0 lines are totally same for both signal and reference decay, there is still some unignorable difference between them, which can not be eliminate by just a MC-correction. The MC-correction only reduced the disparity between real-data and simulation for samples we can offline obtained, but the L0 trigger decision is made while data-taking.

A We don't expect different L0 rates in signal and reference channel as we reconstruct them with the same final state particles. Also the kinematics is much alike. We cross-checked our assumptions in Apendix A.

QLine 705, Fig 13,15,18, and so on. Could you please give a definition on the "Square Dalitz variables" (in range [0,1]) and ?

A Done.

Q L709: remove "Λ_b^0 and"?

A This should be fine. In Lc Ds these are pT, eta and the Lc SDP; in LcD0K there are 2 mode dimensions from the Lb SDP

Q L758 and Figs.18,19,20,21: It seems that for both $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$ decay, you required the same simple mass window (5400-5540 MeV) in signal samples, I recognised that separate these kinds of decays in not so easy, but from the fitting results these two components occurs in the same mass region, do you have an idea how many bias will bring to this correction by the same mass window?

A Indeed, the same mass window is used. This is not a big issue as the variables that we are weighting the MC in do not depend much on the $\overline{D}^{*0}$ decay mode. I.e. the Λ_c^+ SDP, the Λ_b^0 SDP and pT, η very much look the same for both partially reconstructed modes. With the new MC, all $\overline{D}^{*0}$ decay modes are treated equally.

Q L764-L766: Please define the $N_{\text{eff}}^{(X)}$ and make the superscript (truth or genvl) consistent.

A Done.

Q L765: Is this formula correct? Shouldn't it be $N_{\text{eff}}^{k} = \sum_{i \in \text{passedselection}k} w_i$. If w_i is factorised to "w_i^{kine}", "w_i^{Dalitz}", just take the direct product is OK, i.e. $w_i = w_i^{\text{kine}}(p_{\text{T}}, \eta) \times w_i^{\text{kine}}(m_{12}, \theta_{12})$. Why there is squared in the fraction?

A By weighting, the effective sample size is decreased. If the weights are not normalized, uncertainties will be underestimated.

Q L777: I think the first time you mentioned sPlot is in line 731, so please remove this reference.

A Done.

Q L778: FD$\hat{1}$0
chi$\hat{2}$ -> FD
chi$\hat{2}$

A The 10 is a footnote. Prefer to keep as is.

Q L797: please move the reference of CB function to where the first time you mention it, maybe line 417.

A It is cited the first time here.

Q L809: Are these BDT efficiency values for selection strategy A? What's the value for other selection B,C,D?

A 1; or extremely close to 1 (due to tiny differences in the kinematics of signal and reference channel). This is by construction.

Q Eq.(11): Could you please add the label such as stat., syst, ... to tell the reader where the uncertainty is come from?

A Sure, I must have forgotten this.

Q Eq.(11): How do you figure out the final stat. uncertainty from the 18 measurements?

A This is described in line 837-839 and 841-843 (v1r0). In the new version, we follow a more traditional approach.

References

[1] J. Zhu, Z.-T. Wei, and H.-W. Ke, *Semileptonic and nonleptonic weak decays of* Λ_b^0, Phys. Rev. D **99** (2019) 054020, arXiv:1803.01297.