Concerning the analysis we have the following questions/issues:

1. Since the phi(1020) is the dominant contribution to the KK final
state, it seems very important that the phi lineshape in the fit is
correctly described: small problems in the phi lineshape may have a
big impact on the extracted contributions of the other
resonances. For example, the relative contribution of phi and S-wave (fig
18,19) must be very sensitive to the exact description of the phi.

Unfortunately, there are no plots shown in the paper from which one
can conclude that the lineshape is fine, so we strongly recommend that
such a plot is added.The only plot that is there, fit 22 top-left,
looks very poor in the phi region. The analysis note reports a chi
value of 73/20. The conclusion in the text that the fit described the
data well (line 554) then does not seem justified. Also, the overall
chi2 of the fit (table 3) is very poor, with a 'probability' smaller
than 10^-7.  This cannot be ignored and must either be solved or
explained in the paper.

Fig 22(a) is the right plot to show the fit to phi region. We had improved the fit with chi2/ndof=90/60 which is relatively reasonable. If we do a relativistic spin-1 BW+linear S-wave with correct convolution in a 1D fit, we got similar chi2/ndof.

As for the chi/ndof in table 3, current 3D binning is using equal bin size to calculate this chi2, there are lots of bins having low statistics, as we know that even Poisson binned likelihood fit with low statistics could have bias. So we change to use an adaptive binning scheme which requires at least 25 events in one bin. We got chi2/ndof = 649/(569-24) corresponding to p-value = 0.14%.

Dalitz-plot analyses are a model of the final state. They may be other resonances below our detection threshold of 3 sigma, some of which have poorly measured masses and widths, so getting “acceptable” fits is a matter of judgment. If you compare with other Dalitz-plot analysis's chi2/ndof with event numbers of about 20,000, ours is quite good.

2. Concerning the phi line shape we are particularly worried about the
method that is used to account for the finite KK mass resolution. For
the phi(1050) one cannot ignore the resolution since it is a good
fraction of the natural width.  In the paper there is nothing written
about this. In the appendix of the analysis note you write that you
artifically increase the natural width of the phi to account for the
resolution. However, while the resolution mostly affects the 'peak',
changing the natural width affects the shape of the phi also in its
tails.

I have attached a plot that illustrates this: in green is the
relativistic breit wigner (with a phase space factor taking the lower
KK threshold into account, but not the upper threshold). In blue is
the same pdf convolved with a Gaussian (using RooFFTConv). In red is
the relativistic breit wigner with the width decribed in the analysis
note. You see that the effects on both peak and tail are different:
the left side of the peak look different; the contribution to the tail
from the 'fudged' breit wigner is too large.

Thanks for your plots. As for the tail, because the two curves’ normalization is arbitrary, you can first scale the one with enlarged width to match with the one with convolved, then compare the peak. Actually we did look at this when we determined how to set the effective width. Here we show a plot where the points are toy MC with 1M events of phi BW convolved with a Gaussian using RooNumConvPdf. Then it’s fitted to pure BW with free width show in histogram. To see how the two are different in terms of current data size, we scale both to data size and recomputed the errors on the points.



We understand that it is not trivial to take the resolution into
account: It is a strong function of the KK invariant mass for instance
and this dependence needs to be modeled. Furthermore, you'll need to
rely on numeric integration which makes the fit notably
slower. However, we think that given the level of detail with which
you report for instance the phi/S-wave fraction, the resolution is
such an important issue that it needs to be dealt with properly.

Our decision to use an effective BW is based on this detailed TOY study for resolution. As the study shows that we can get correct fit fractions including phi/S-wave fraction. We don’t think it’s a problem using current method. In additional, we give physical phi/S-wave fraction vs m(KK) using phi PDG width with zero resolution. So the plot is not affected by the enlarged-width phi lineshape.

3. We are a bit worried about your use of sidebands for background
model, in particular because of the B mass constraint. For signal
events the mass constraint will improve the mKK and mJpsiK
resolution. However, for background it will make it worse and the
effect is bigger the further away you are from the Bs peak. This
potentially affects a real phi contribution in the background. Such a
contribution was observed in the 0.3/fb time-dependent analysis note,
see ANA 2011-036, figure 6. Your projection of the mKK sideband does
not show evidence of a sharp phi peak in the combinatorial
background. It is unclear to us if this is because of binning, because
there none (due to the tighter selection), or because it is smeared
out as a result of the mass constraint. If it is the latter, then the
background shape extracted from the sideband is not representative for
the signal region.

This was studied during the review procedure that with constraint or without gives consistent distributions. We do have a small phi mass peak from combinatorial background that appears if we zoom in phi region, but the smearing effects are small and within stat. uncertainty





Here the Left plots are for lower sideband and right plots for upper sideband. Here black is with constraint and red is without constraint.

4. You explain that you use toy simulations to compute the
uncertainties on the fit parameters because it is hard to propagate
uncertainties. We are really puzzled by this: if you can compute
"f(x)", then propagating uncertainties is as easy as computing
f(x+sigma(x)). If you need correlations, you can use the same method
to compute the Jacobian and then transport the covariance matrix. Why
did you use toys?

This is not toy MC. We use the full covariant matrix and central value from the fit to generate 500 sample parameter (amplitudes and phases) sets. For each set, the fit fractions are calculated. The distribution of one of fit fractions is used to set the statistical uncertainty.

The fact that the chi2 of your fit comes out so poorly makes us also
worried about the statistical uncertaintiesin general. Do the toy
simulations also have such a poor chi2/dof?

As we don’t do what you think, so we don’t have such comparison. We will also repeat that the Dalitz-plot analysis is model dependent in that components can still be involved but are not 3 sigma significant with our level of statistics.

5. You report the ratio of the lambda=0 and lamba=1 components of the
phi(1020) and claim that it is compatible with the time-dependent
analyis. However, we see a few problems with this comparison.  First,
at face value the reported numbers are incompatible: the uncertainty
for the time-dependent analysis is dominated by the systematic error
on the angular acceptance. Whether that error be right or not, there
is no reason to assume that it is any different for this analysis, so
they are fully correlated. The statistical errors should also be
correlated. So, the numbers are not close at all.

We agree that statistical errors are correlated. While don’t agree that on systematic error. We use different procedure to take into account K momentum difference between MC/data for angular acceptance. Here we think the difference is physics, i.e. Bs Pt&P in MC is not perfect, while time-dependent analysis think it’s due to K efficiency.

Second, we are probably comparing apples to pears here: the
time-dependent analysis reports amplitudes at t=0, while these are
integrated amplitudes. For dG!=0 these are not supposed to be
identical.

Yes. You are right, our time integrated number tend to smaller this value due to dG!=0, which is discussed in the ANA-note.

Finally, the number that you report is not corrected for the finite
decay time acceptance. We do not know what the effect of that is,
since you don't discuss it anywhere, but it may well account for part
of the difference.

Maybe.

In short, in our view this ratio should not be reported in this paper:
without the acceptance correction it is not useful to anybody. The
comparison the our other result is off. Instead, it suffices to say
something like "The value of the ratio of the two phi(1020)
contributions is consistent with that extracted from the
time-dependent analysis." We also think that the effect on the
branching fractions of the time acceptance in case dG!=0 needs to be
discussed somewhere, for example when discussing systematic
uncertainties.

We agree we shouldn’t quote our systematic error on this and just mention it’s RESONANBLY consistent with time-depend analysis.

We have set 1.5% systematic uncertainty due to dG!=0 in the new version.

6. In your discussion of the systematic uncertainties you discuss one
test where you increase the IP chi2 of the kaons. What systematic
effect if this supposed to probe? You explain that it changes the
chi2/dof of the fit marginally, but isn't it more relevant how it
affects the fitted parameters? (You could have a perfect fit with an
incompatible result.)

This is an important test for the acceptance systematics because the IPCHI2 cut is one of two sources to generate the non-flat acceptance. Another source is PT>250 MeV cut on Kaons.

7. For the efficiencies in table 8 an assumption in the MC is made on
the size of ADeltaGamma (or D, or cos(phis), if you wish). This
affects the time-dependent rate and hence the efficiencies. What is
the systematic uncertainty assigned to these assumptions?

That’s has been already incorporated in our Delta Gamma systematic, we vary lifetime by Delta Gamma/4 to calculate the efficiency uncertainty.

We have also a small number of more editorial comments:

- line 310: it looks odd to define the helicity angle as an angle
  between vectors in two different lorentz frames. Instead, define it
  as "the angle between the muplus and minus the direction of the B in
  the J/psi rest frame." (You can also take mu-minus if you don't like
  the 'minus' for the B direction.)

This definition is generally used. The J/psi momentum in B rest frame is
a boost vector to boost muplus to J/psi rest frame. The boost vector's
direction is not changed when boosting along itself.

- line 314: In case you have actually used DecayTreeFitter: I (WH)
  agree that we reference the DTF paper far too often in LHCb, but
  since these Dalitz analyses actually really profit from it, I
  propose to add a reference to DOI: 10.1016/j.nima.2005.06.078.

ok, thank you.

- equation (15): where does this expression come from? Please explain
  or add a reference.

It is obtained from Eq. 5, considering non-resonance is S-wave (i.e. LR=0, LB=1) and uniform in phase space (i.e. A\_R=1).

- line 373: the reported chi2 (198/303) is extremely unlikely. Is
  this a typo, or is something wrong with estimated errors? If the
  latter, fix it, or don't report it.

removed line 373

- figure 8: why is the scale arbitrary? cannot you define it?

one can generate the toy at any number of events, so it is arbitrary.

- line 417: "The situation ... is .. confused" Replace 'confused' with
  'confusing' (or 'uncertain' or something equivalent).

ok, thanks.

- line 434: refer back to Eq.20: '... Flatte resonance shape, see Eq.20.'

ok, thanks. (e.g. we used Eq. 12)

- line 441: The parameterization of the acceptance doesn't look
  trivial. Does this really allow you to do an 'analytical' integration
  over the helicity angle?

yes, The cosθJ/ψ dependent efficiency is proportional to (1+α\*cos2θJ/ψ ), and that is for the amplitude are (1- cos2θJ/ψ ) for helicity = 0 and (1+cos2θJ/ψ )/2 for |helicity|=1. So nothing stops the analytical integration over the helicity angle.

- line 496-499: you don't quote a result for Bs->J/psif0(980). Presumably
  because of the strong dependence on the parametrization. Add a comment
  here: 'Therefore we refrain from quoting a branching fraction
  measurement for the decay Bs->J/psif0(980).'

added to the tex, thanks.
- line 533: please specify how large the f2'(1525) mass region is.

it is one full width of f2'(1525) mass(see L531-532), this is 84 MeV what we measured.

- figure 15b, and fig 17 are identical appart from the curves. Please
  choose one.

we prefer to keep as is.

- figure 17: there is either a real problem with the fit, or the
  projection of the blue curve (total pdf) is wrong: it is absent for
  the smallest bin.

in ROOT histogram, smooth curve starts from the centre of the lowest bin.

- figure 18 and 19: From the text it is unclear what is in these
  figures. At first we had not realized that this is not the result of
  a binned fit to the data, but rather just an integral of your fitted
  PDF. We propose to correct the text. It should be stated somewhere
  that the error bars in the figrue are correlated. Finally, we see
  little reason not to make the bin size in figures much smaller, or
  better still, just draw continuous bands. Finally, is there a good
  motivation to show both figures?

We think the figue captions clearly describe what is in the figures. Fig. 18 was in the original draft and Fig. 19 was requested by the PC.

- figure 21: Add it in the caption that the helicity distributions are
  compatible with expectations for spin-1 and spin-2, respectively.

good suggestion. added to the caption. Thank you.

- line 553-555: as we commented above, this conclusion does not seem
  justified by the plots. See e.g. figure 22, top left and fit 23,
  bottom left.

We already mentioned that Fig.22 and 23 should have been updated, (we have updated in the latest version of Analysis note and paper draft). There were some changes on the results and plots after version 4 of the paper draft, where the largest changes was moving the phi mass from 1.0198 to 1.0195 GeV). In the previous study Fig. 22 (a) gives chi2/ndf=150/60 (1 MeV bin, which is 73/20 for 3 MeV bin size), whereas the model of version 5 and later gives 90/60, which is reasonably good to explain the phi region."

For Fig 23(e): As it mentioned in the tex that <Y^0\_4> represents the D-wave if we consider no partial waves higher than the D-wave contribute. We have tested by adding f2(1270), but this does not help to improve left side of the Fig. Possibly something higher than D-wave contributes which we haven't considered in the analysis.

- line 599-602: from the description we do not understand what you
  have done, so it is hard to imagine that somebody from outside the
  collaboration could. Please consider to rephrase this.

We used the standard procedure documented https://twiki.cern.ch/twiki/bin/view/LHCb/TrackingEffRatio and prefer to keep as is. We also think its clear and well defined as is.

- line 603: A sentence is needed to clearify/warn that these
  branching fractions correspond to time-integrated values, and that for
  example, these numbers cannot straighforwardly be used to draw
  conclusions on the magnitude of decay amplitudes;
  "Note that the branching fractions reported here correspond to the
  time-integrated branching ratios, which differ from the branching ratio
  at t=0 [arXiv:1204.1735], and therefore special care needs to be taken
  to draw conclusions on the amplitude level."

added following line "Note that these are the time-integrated branching fractions."

- line 619: remove 'long' in "long tracks"

ok, thanks.