CLAS12 analysis note: Measurement of the cross-section of the photoproduction of the J/ψ meson near the production threshold with the CLAS12 detector.

First round of comments (black)

Answer to the first round of comments (blue)

General answer to the first round of review

- We acknowledge the very thorough review of the note provided by the committee. We thank the reviewers for pointing at important aspects which have now been addressed. We have answered all of the comments, providing a single detailed answer in cases where similar issues would be raised several times. In particular, new studies are provided concerning the momentum correction, the calculation of the photon flux and the parametrization of SDMEs. The interpretation of data using out-of-date models is still present as cross-checks, but we aim at including the latest models interpretations in collaboration with theorists once the cross-section is released.
- While reviewing the code and working on the answers of this first round, three errors were found: the ISR corrections were wrongly done (applied to the real flux instead of the virtual), one of the MC samples for the Spring2019 had a wrong initial beam energy, and the flux for the differential cross section was wrongly computed for the spring19 sample. Those issues have been corrected.

General comments discussed among the group:

• Provide more clarity in the way the radiative corrections are applied. You explained how you corrected for the momentum change due to the final state radiation FSR, and the ISR but we did not see any discussion of the external radiative corrections. Your cross section is still dominated by virtual photons flux. The real photon flux is about 20% according to Figure 44.

A figure to illustrate Initial state radiation has been added in section (5.4.2). Section 2.4 now has figures showing which diagrams are included in the radiative effects calculation. Please let us know if this is what you mean by external radiative corrections.

• Concerns about how the flux is calculated (see comments by Derek).

A new subsection (8.4.1) has been written to justify the flux computation. Two points are discussed: 1) A justification for the maximum value of virtuality of the photon is given by analysing the electro-production case, where the scattered electron is detected in the FT. 2) The Frixione Formula is applied to the CLAS12 case, where the FT can be used as a veto, and the maximum virtuality constrained by the FT angular coverage.
Finally, the impact of using a model of flux based on angular consideration on the measurement of the cross section is tested and shown to be small compared to the other systematics.

• The profile of the J/psi fit is certainly not gaussian (see Figure 59). And the fit of the background should be anchored on the side bands and not aborted right after the peak.

We have indeed tested the impact of fitting with a non-gaussian Jpsi peak by varying the fit function in the systematic studies (crystal ball with a left-hand tail). We showed that the effect of changing the fit function is below 15% on the extracted cross-section. The choice of a gaussian is driven by the small statistics of the analysis, which makes the crystal-ball fit difficult to control (the tail parameters have been fixed using MC). Furthermore, the events which have been 'missed' by the fit are taken into account by the acceptance correction.

The fit range now extends to 3.4 GeV (previously 3.3 GeV).

• We agreed on concerns about the 0.16 difference of your normalization factor. This corresponds to a 23% difference which is not small. Somehow you took the absolute difference 0.16 and called it 16% as your generous estimate of the uncertainty. It is not clear why that's the case. See comments on Section 4.

Both methods rely on different assumptions. While the signal particle method uses average efficiency for each particle without correlation of efficiencies, the method based on MC relies on the description of the background using a reweighting technique. While both methods appear imperfect, they do provide estimates for the difference of efficiency between GEMC and the real experiment.
The value for the error of this parameter has been modified according to this remark. This value is in accordance with analysis currently under review involving multi-charge-particle final states.

• Please neglect making plots of ratios like in the bottom of each panel of Figure 17 since these ratios are not used or do not add useful info. Except for the fact that the corrections are sometimes as large as 50% or more.

The ratio plots which did not bring any informations have been removed

• Please make sure that the figures have all the required labels, some are missing and it makes it difficult to guess.

This is corrected now.

In Figure 15 we did not understand why you took the fit for correcting the oscillation as a function of phi. You have precise data and the chi-squared is bad because of those high precision data from -180 to -110 and +180. It looks to us that you need to correct using a functional form that better fits the data.

A new set of functions has been derived using FD protons. This reduces the need to correct for the angle of the proton. We also tested the impact of this corrections wrt the one used in the first version of the note and found no clear difference in the invariant mass spectrum. This is detailed in section 3.1.4 and the new section 3.1.5.

• Having a table of acronyms would be highly appreciated to anyone who would read this note to learn about the analysis.

A glossary has been added at the start of the document

Details Comments/Questions-c/q from Z.-E. M
Reviewer: Z.-E. Meziani

The scheme for identification is page # - line # -c/q for comment or question respectively.
or page #-Fig # -c/q

Section 1: Motivation and previous results:

5-143-c: “maybe accessed in the near threshold region” rather than “can be accessed”

Corrected

5-146-c: May need more references for the comprehensive reviews.

More references will be added to the article. For now, the ‘comprehensive’ has been removed

5-fig.1-c: this diagram is dominant only in the large s regime, near the threshold it is good to have dots indicating that there could be 2 and more gluons exchanged.

A new diagram has been made to represent this

5-153-c: The first GlueX measurement of 2019 is a different reference. Perhaps, it is good to mention first the 2019 reference with one t distribution and then the 2023 reference with 3 t distributions.

Reference has been added

8-171-c: The GPD approach does not use vector-meson-dominance and equations 2 and 3 of line 170 are a first order expansion in ξ.

We will make sure this is clear in the article. The text has been slightly reformatted accordingly.

8-179-c: Equation (4) is an approximation, one can use the exact expression of *F(s),* and instead of the factor of 8 it is (*2t +8m2*)/*4m2* where *m* is the mass of the proton

The formula for Ttilde from PHYSICAL REVIEW D 106, 086004 (2022) is now in the note. We could note find the form that you mentioned. Could you point us to the right reference ?

Section 2: Analysis Code, Data and Monte Carlo Samples

13-258-c: “used to normalized MC..” should be “used to normalize MC..”

Corrected

13-274-c: I do not understand how does 103 x Q [in C] becomes x Q[in mC]. It must be 10-3 rather than 103

Corrected

14-279-c: Perhaps it would be good to mention weighted by what? I presume in the unweighted case the cross section of the physics process is included in the generation. In the other case the weight is assigned at the filling of histograms.

Sentence added to clarify the statement

15-303-c: In equation (26) one momentum of the J/psi event should be that of *e-,* there is a misprint.

Corrected

15-326-c: Maybe define "ECIN/ECOUT" in the SF.

A sentence has been added to clarify the equation

15-fig.9-q: It would be good to evaluate the AUC (Area Under the Curve) to measure the classification performance. What has changed between F18 and S19 that makes S19 data looking much better?

AUCs are now provided for all ROC curves.

The 2019 data are better calibrated, especially the ECAL which is the main subsystem used by the AI PID algorithm.

Section 3: General Analysis Strategy and Tools

21-357-c: "may radiated photons…" should be "may radiate photons…"

Corrected

21-fig.12-q: Can you clarify what you call azimuthal angle difference, is it $\Delta\phi$ or $\Delta\theta$? Are you working in the lab frame with the z axis along the beam line? The units are not clear in the plot are these degrees or milliradians?

The wording has been corrected, the polar angle difference \Delta \theta is used. We required the leptons and photons to have the same theta at the vertex. All the angles are defined in the lab frame with z axis in the beam direction. And the plots are in degrees.

21-367-c: "… to the electron positron mass.." "…to the electron positron invariant mass…"

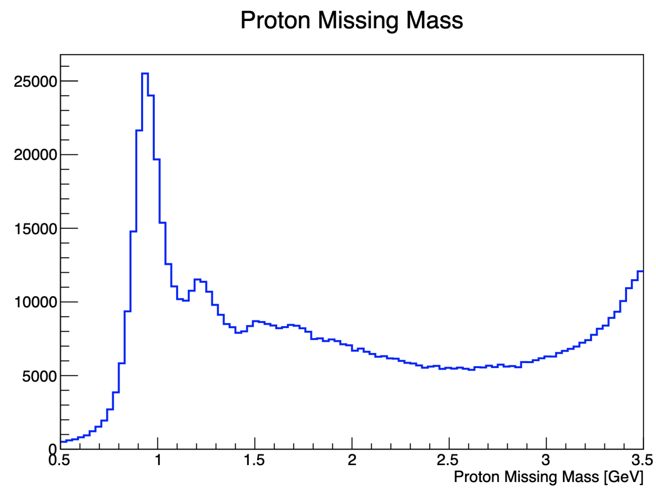
Corrected

22-95-c: "…elastics scattering should be "…elastic scattering…"

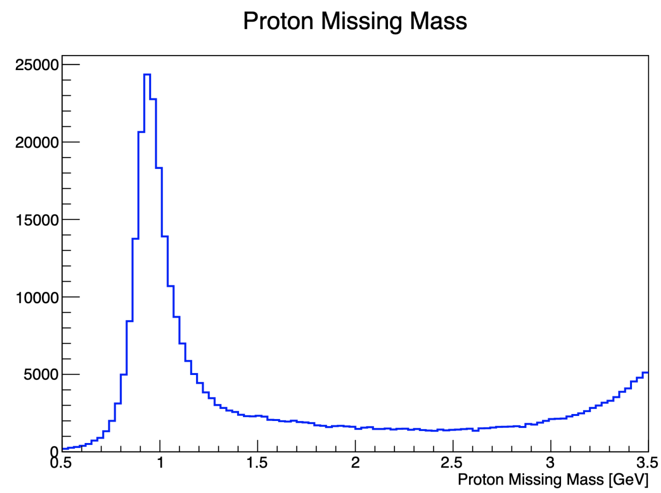
Corrected

23-fig.14 q: Do you understand the "hump" at the proton missing mass of 1.5 GeV in the outbending (bottom of figure) data? Is it an acceptance effect?

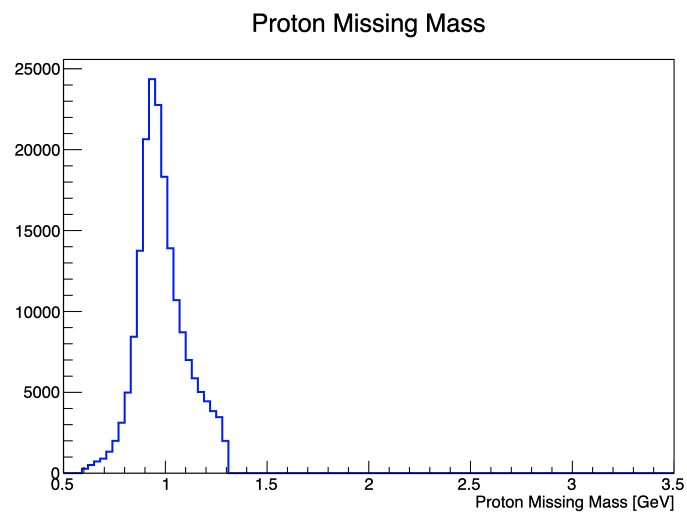
This comes from resonances in the background that are more prominent before the cut on $\Delta\Phi$ (plots 1 & 2 before & after cut). This background is removed with the cut on the proton missing mass (plot 3).



1



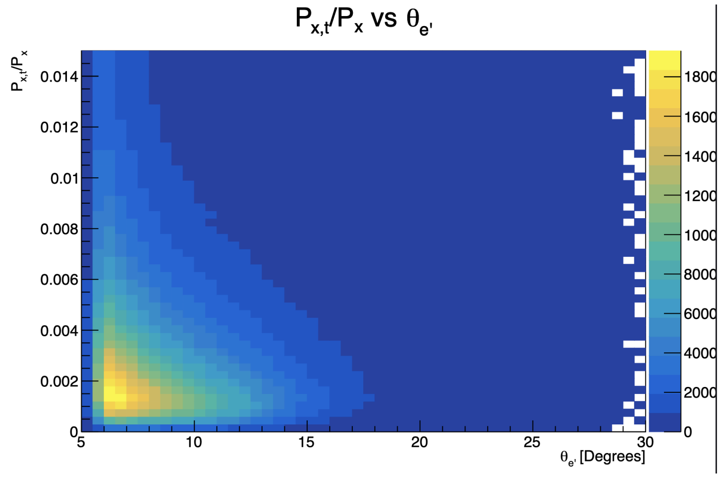
2



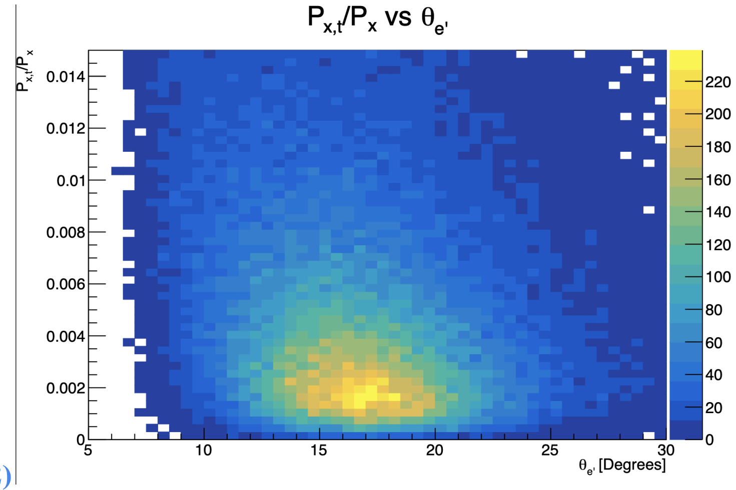
3

23-fig.14 q: Do you understand why the Px,t/Px for the inbending is wider? To have the same relative tolerance between inbending and outbending the cut at 0.015 cannot be the same for both.

This is due to the difference in polar angular range of scattered electrons (plots below show 1) outbending and 2) inbending).



1)



2)

24-fig.15-c: The $\chi 2$/n.d.f. of the fits in Figure 15 are not good due to the -180o data point, in either inbending or outbending case. The difference with the fit is 2o in ϕ. Why not correct for it explicitly?

See the answers about the momentum correction and the preliminary angular correction in the 'general comment' section

25-16-q: Do we understand why the corrections are the largest at low momentum starting at 1GeV?

Low momentum electrons scatter at large angles where the field is not as well known. In this analysis the leptons momenta are required to be above 1.7 so the region with large corrections is not used.

25-411-c: "…is applied to the both leptons…" should be "…is applied to both leptons…"

Corrected

26-fig.17: What do you do with these ratios? While the average seems to be about 1 the correction in each bin is as large as 50%. Please remove the ratio if it is not used.

The ratios have been removed

Section 4: Background modelization and Normalization Factor

32-468-c: The BH background can be compared to the Monte Carlo on the side bands of the J/ψ peak, not under the peak.

The phrasing has been changed. The intended meaning was indeed what you suggested.

32-469-c: "…can be compared in data and Monte Carlo." Should be : "…can be compared between data and Monte Carlo."

Corrected

38-fig 25-c: The dark blue points are the result of doing what? this is the same of many figures that contain blue points.

Blue points are background-subtracted data. This is mostly used to extract the efficiency normalizing factor. For clarity, it is removed everywhere it is not used.

37-544-c: "…factor the Fall 2018 inbending…" should read " …factor for the Fall inbending data set"

Corrected

38-fig.28-c: The figure is not clear. The blue points are not described. It looks like the J/psi events are nowhere to be seen.

See answer above about the blue points.

In this figure, only events with reconstructed virtuality of the photon above 0.5 GeV are shown. Most of the events in this region are background and we expect no jpsi to be reconstructed.

39-fig.29-c: Data/MC can be as large as 1.5, this is concerning about the power of prediction of the algorithm. Looking at Appendix K for the comparison in the signal region, I wonder what uncertainty is estimated for the correction factor. The evaluated spectra as a function of positron and electron momentum don't line up in the low momentum or positrons and high momentum for electrons.

This method relies on extrapolating a reweighting method to a region where it was not trained. This yields some discrepancies in some parts of the phase space. The momentum spectra are the least well modeled.
To estimate an error on this estimation, we used the comparison with both methods to derive the correction factor.

42-fig. 32: shows a poor description of the Eγ spectrum.

For the efficiency factor, the outbending era is not used due to the very large contribution of the background. We apply the factor obtained for the inbending case.

We have left the plots in the note for completeness.

51-628-c This method results in ω = 0.849 for the outbending case and 0.867 for the inbending case while the other method results in a 0.69 irrespective of the run period. Do we know which one is correct?

In the case of this method, we broke down the efficiency for each era. For the efficiency factor obtained using the complete description of the data (signal and background), we only used fall inbending and spring 2019 data. In the case of outbending the ratio signal/background become extremely large and difficult to interpret. We interpret the difference as a systematic error on the knowledge of the normalization of our cross-section.

Section 5: Integrated cross section

1- Decay of J/psi is not isotropic, this should be taken into account. See publication of GlueX as an example. J/psi-007 in Hall C also considered a non-isotropic decay of the J/psi in the simulations.

See answer to a similar comment by Derek

53-651-c: Why is the blue curve (exponential background) not anchored in the side band.

The fit range now extends to 3.4 GeV (previously 3.3 GeV)

53-651-c: The J/psi fit has no reason to be a gaussian since we know that in the e+e- decay channel radiative corrections change the shape and make it asymmetric.

See the answer to this comment in the ‘general comment’ section at the start of this document

59-fig.49-c The radiative corrections factor is pretty large at low photon energies, what uncertainty is put on it. Does it include the internal and external corrections.

We have tested the radiative effect implementation (ad-hoc and photos). The difference of the two has been studied in detail in section 7.1.9 and 7.2.9. In the lowest energy bin the error is 16%. It does include all corrections (internal and external) as described in *F. Ehlotzky et al.* and *E. Barberio et al.* These two refs are in the note. For clarity the diagrams have been added to the note (see section 2.4).

Section 6: Differential Cross Sections

62-750-c: “in in E_{gamma}…” one “in” suffice.

Corrected

Section 7: Systematic Error Study

No comments

N/A

Section 8: Additional checks

83-fig.76-c: I don't understand why the integrated exponential gives a factor almost 2 smaller uncertainty than an integrated dipole? Why is the integrated fit outside the data in the third bin? Does it mean the dipole fits give large contributions at large t where you do not have data compared to the exp. fits?

The error is mostly driven by the value of the cross section at 0 and the dipole fit yields a larger error for this parameter than the exponential fit.

The fit has been redone in all bins, following all the changes which have been done to the analysis, including the corrections on the flux computation. The integrated differential CS is now within error bars of the integrated CS (except for one bin), see Figure 81 of the note.

Section 9: Physical interpretation of the measured cross sections

• General comment: The GPD model of GJLY was the first attempt with a two gluon exchange leading term calculation in an expansion in ξ. This theoretical approach has been carried to NLO and a Bayesian method was used to analyze the data with $\xi > 0.5$ (Arxiv:2501.10532[hep-ph]

See answer to the last comment of this referee

• The extraction of the mass radius using Dima Kharzeev method is not complete. His description of the cross section uses VMD and you have to find the right place in $E\gamma$ where it is valid which is not obvious. He emphasizes the dilaton rather than the graviton. Therefore, the graph on the right to extract the anomaly is not really useful. Note that it was published as a supplemental material in the J/psi-007 Nature paper but we did not take it seriously. Determining one radius from the 2D data is more appropriate.

Although the model is simple, we believe that it is important to produce this plot for sanity checks and also as a baseline for the same analysis using neutron target which is currently under review.

• I honestly don't see the need to use the perturbative formalism of the HERA region where we know the cross section is dominated by the imaginary part of the amplitude and the factorization was proven and, compare it to a very non-perturbative region where the formalism breaks down. A two gluon exchange picture makes no sense in the threshold region, that is why I mentioned the new reference of Feng Yuan , Guo et al. I am not sure your x means much, even though it seems to indicate that at large x we have a small transverse radius of gluons and at small x it is growing similar to the quarks profile extraction in DVCS. My recommendation is to drop section 9.3.2. and all the projections made by the first version of the GPD model.

This part has been moved to the appendix. Your comment was added to the corresponding text and we will follow your recommendation not to include it in the future publication.

• If you want your GPD approach to be useful you need to follow the procedure of the paper Arxiv:2501.10532[hep-ph] where the method is NLO with a controlled approximation. The method used in your analysis is leading order and we know it is not valid per the authors but rather a first attempt. Of course in all cases this method has controllable approximations if $\xi > 0.5$. It is worth to see the impact of your data with $\xi >$ 0.5 using this approach. I am sure the authors will be happy to have your data and do it themselves.

See answer to the comment below

• 96-fig.89-c: I would drop the GPD model as presented and recommend the new method.

We will contact the authors, once the cross-sections values have been approved.

Review : Derek Glazier, University of Glasgow.

1 Motivations and previous work

Good.

N/A

2 Analysis Code, Data and MC Samples

2.3 Monte-Carlo samples and processing – for clarification, JpsiGen generates both real and virtual photoproduction Jpsi events ?

It does, see answers below

- Note I checked the code and it seems it is only real photoproduction. It does use N_EPA for flux normalisation.

Both flux are computed and kept in the event weights

- This means the virtual photo-production process, which is the main contribution is not fully modelled.

The post-processing of the MC samples later in the analysis chain has both contributions (EPA and real photons)

3 General Analysis strategy and tools

Line 348, Fig 9 : "It can be seen that in the inbending case, the two curves are very close to each other, while in the outbending case, the discrepancy is larger."
- in my impression the two curves are certainly not very close to each other. For the outbending case the agreement is just terrible. S19 looks much better and can be classed as close.

The wording is now changed. While the agreement is not great, the ratio of the data and MC efficiencies is then corrected when needed in the rest of the analysis (in the computation of the total efficiencies)

- The effect of this discrepancy is estimated through Line 350, ", the ratio of signal efficiency is computed as a function of the cut applied on the BDT output."
It is not entirely clear how this is done, even from the referenced analysis note. I presume the signal comes from radiative events ? And you fit this with no BDT cut, then refit with varying cuts. You do this for the simulation and data and take the ratio. Is this correct ? Can you please clarify this in the text. In addition you should include some of the signal fits for data and simulation so we can be convinced everything looks reasonable (I.e fit with no cut, fit with cut =0 say).

The signal is extracted directly from experimental data, using radiative events (e-> e' \gamma). We first perform a fit to the relevant distribution (e.g., \Delta\theta) without any BDT cut to establish the baseline signal yield. Then, we apply a range of cuts on the BDT output and repeat the fit at each cut value to extract the remaining signal. The signal efficiency is computed as the ratio of the signal yield after each BDT cut to the yield with no cut. This procedure is applied independently to both data and simulation, and the ratio of efficiencies (data/simulation) is used to quantify the agreement and performance of the lepton ID model.

The required plots are provided below and have been added to the note.

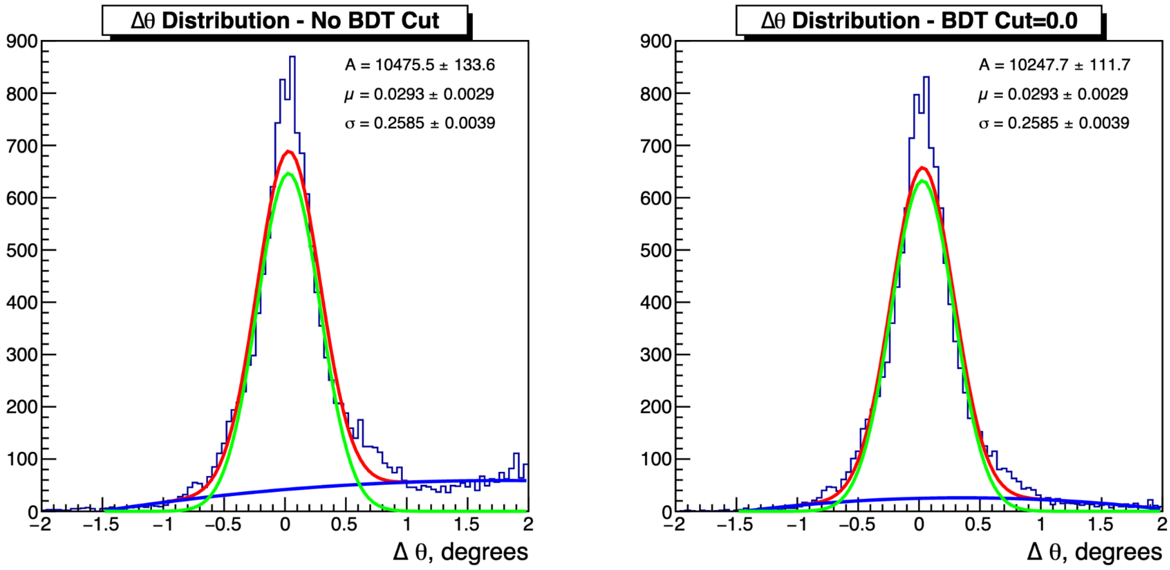


Fig 10, There are some fits to the lines in the plots. Can you clarify if these are used in sect 4.4 or you just use the measured value at the cut value you used ?

The fits are not used. They have been removed from the figures and the note is actualized accordingly

Plots in 10-11 are only for F18in, what about F18out where the situation looks less favourable.

Appendix O has these figures.

3.1.4 Lepton momentum correction

Fig 15 right plots. Apply the correction from the magenta line is going to make the angle reconstruction worse for the most left sector. Why do this ?

See the answer provided to this point in the ‘general comment’ section at the start of this document

4 Background Modelling and Normalisation factor

“Normalisation factor” is very generic. What specifically is being normalised ? Can we show it mathematically here ? Is it simulated BH to data BH ? I think it is not correct to call this a normalisation factor. This is a correction factor for the normalisation.

The name has been changed to efficiency correction factor, to reflect better what is done in this section of the analysis

Line 471 : “Normalisation factor” - “the ratio of efficiency in data and Monte-Carlo”Is this factor unrelated to any luminosity normalisation factors ?

See answer to the previous comment

Figure 23,24 – what are the blue points exactly ? Please add definition to caption as probably too much hassle to add to legend.

See later comment about these points. They have now been removed for clarity.

Line 512 : “Once the re-weighting procedure has been validated, we can use the same weights to model our background in this region of interest.” What does it mean, the same weight ? The training weights are defined for events in different Q2 ranges, how can you use these weights for signal region ? Do you mean you use the BDT to produce new weights in the signal region ?

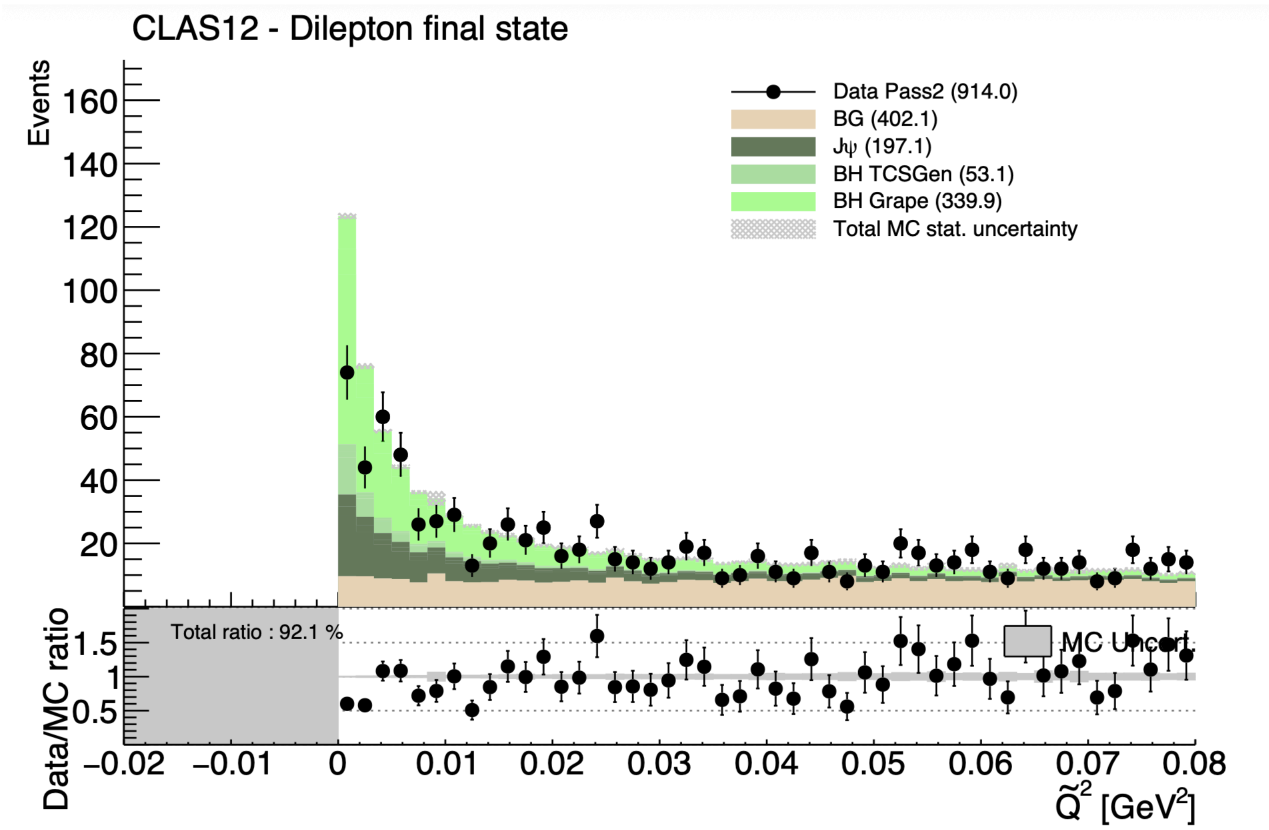
We assume that the weight can be used for the signal region. We essentially extrapolate the model to the signal region

There seems to be an implicit assumption that all 9 training variables are independent of Q2 (so you can use the same reweighter at all Q2) This could be demonstrated by plotting the 9 variables versus Q2. It is a similar issue to use of sPlot right ? There Q2 would be the discriminatory variable.

Instead, we look at the behaviour of the prediction in the high Q2 region (in Appendix L). The extrapolation should work on both sides of the Q2 range. The integrated discrepancy is 10%.

Fig. 25, It would be good to see the BG distribution as Q2→0 . It looks like (if I take the difference black points – blue points it falls off towards Q2→0). This is the key part of the BG distribution though. Is its behaviour understood, it just comes from the mixed events right ?

The Q2 dependence of the background is not constrained explicitly by the method we used to model it. Below is the plot required, zoomed in the peak region. The Q2 dependence of the background is flat.



So now if I understand right, the total BG + simulations has a larger signal peak than the data by about 20-30% (at Q2=0), and this gives the correction factor.

The corrections factor is obtained using figure 42 and 43 (in the new note), using all events below Q2=0.5. A fairly large fraction of the events have a reconstructed Q2 (\tildeQ2) above 0

4.2.4 Overall background normalization factor – But we are assuming we know the Jspi cross- section, which is what we are trying to measure. This we cannot do, right ? Why not do this up to invariant mass 2.8 where there is no Jpsi cross section info required ?

This is done for a Qtilde2 region where there is no signal Qtilde2>0.5 GeV2. And indeed the expected number of JPsi is small in this region (around 3), so we can safely consider that all events in the data are background. (Figure 25 intend to explain this)

Line 565 - "The blue data points correspond to the data histograms where the background spectra was subtracted bin per bin." Now we know what it is ! The terminology is a bit confusing. I tend to think of background as everything but Jpsi, but here it is everything but exclusive dilepton right ?

Yes. It is now removed from all plots.

Line 571 : now we restrict the mass range to not include the Jpsi. Why not through-out ?

We only use the mass range where BH is dominant (just below the peak), so that the efficiency correction factor does not depend on the JPsi simulation. The rest of the procedure (the background modelization) also does not use the jpsi events at all.

Line 579 : The correction factor is calculated for BH events, which have very different angular distribution (in dilepton CM), at photon energy and invariant mass largely below Jpsi. So how good is this factor to be used for Jpsi ? The overall given systematic uncertainty on this procedure is large so I guess that accounts for it.

By providing two estimates of the correction factor, we indeed expect to account for these kind of effects.

4.4 Estimation of the normalization factor from single particle efficiencies

This section is a bit hard to follow. But the point seems to be that the correction factor is 0.849 compared to 0.69 previously, or a difference of 23 %. No explanation is given for this large discrepancy, or which one we should prefer . Note on line 644 we are told it is 0.70.

Line 644 statement has been modified for consistency.
Explanation for the large difference is given in the main comment at the start of the document

5 Integrated Cross Sections

5.3 J/ψ peak fitting procedure - why are we using a Gaussian line shape for the signal? The reconstructed Jspi mass peak will not be accurately reproduced with a Gaussian, better to use simulated templates. In the appendix the width seems to vary randomly by quite a bit bin-to-bin.

See the answer given to this point in the 'general comment' section at the start of this document

5.4 Photon flux Equation 49 is not the correct virtual photon flux for Jspi lowQ2-photoprodution. It ignores the phase-space for virtual photons tends to 0 at Jpsi threshold. i.e 0<Q2max<0.2 due to kinematic limits. This will only have a significant effect in the lowest energy bin I think, but is straightforward to apply.

A new section has been added concerning the flux specifically. This comment is addressed there. See the general comment at the start of these comments.

Line 672 – I do not see any explanation for the Q2 max in the given appendix Q.2

This appendix was missing. The explanation has been moved to the section regarding the flue and the Q2max value is now motivated.

One factor I do not see, the flux should be calculated using generated Q2, not reconstructed. This is an asymmetric exponential distribution, so resolution effects will shift reconstructed Q2 mean to higher values. So what value of Q2max should be used in flux calculation if the reconstructed max is 0.2 ? – it probably should be lower.

The flux is computed with the generated Q2

How do you account for the finite virtual photon Q2 dependence when calculating acceptance? If you ignore it your acceptance will be too large. You can check if this is a significant effect by comparing the number of accepted (Q2<0.2) BH events for photo and electro-produced (Grape vs TCSGen) given the same number of generated events for each.

The effect of the finiteness of Q2 is taken into account in the EPA flux which is used in the acceptance calculation.

Equation 50 should not use Eqn 49, but the 2D virtual photon flux formula, then by sampling in this way it might work. - Correction, previous sentence won't work as there is no Q2 dependence of generated events.

N/A

Alternatively varying Q2max as function of photon energy in Eqn 49 might work (I.e. close to threshold Q2max<0.2) . Note the effect should be smaller closer to threshold, so cross-section should be a bit larger.

See previous comment. Also, by implementing the Frixione formula with an angular cut, we actually indirectly vary the Q2max has a function of energy. The Frixione flux is also used in the acceptance calculation of the newly added comparison dealing with flux calculation. The effect is shown to be small

3 5.5 Detection efficiency

Line 706 : how is the uncertainty in the acceptance computed ?

No uncertainty on the acceptance is taken into account

Figure 46 - it is not addressed in the text but in the caption there is the comment : "using the full MC dataset. The latest has been ignored as it does not take into account the small statistics available in data and in particular was yielding large χ 2 for the fits."

As mentioned a Gaussian will not well represent the reconstructed Jspi shape, so it is no surprise the chi2 blows up if you have large statistics. I do not see a convincing argument why the large statistics derivation is not better. Is the point that due to low stats you are misfitting the peak and the way to correct for that is to misfit the simulation in the same way so the effect cancels ?

> In data, using a gaussian or a crystal ball do not yield very different results. In addition, we have to fix most of the crystal ball tail parameters to be able to obtain reasonable parameters. If using the full MC statistics one would then also need to fix these parameters and there would be only small difference with the current method.

Why not perform event-by-event maximum likelihood which is more reliable with low statistics ?

> Future analysis of JPsi photoproduction with few events (ie on RGE/D data) will use this method.

Appendix G first plot – you see the Gaussian model is only letting you count ¼ of your actual signal events (16/62) , and about half in the second.

> See comments above on the fit procedure

It is not explicitly mentioned, but I assume JPsiGen is generating flat e+ e- helicity angles. This can lead to a difference in acceptance than for the real data, which will have different distributions based on the SDMEs of the Jpsi. Note, GlueX used a 1+cos^2(theta) distribution (model dependent, assumes Jspi take all of photon polarisation), you may also fit SDMEs directly or use a model that goes beyond GlueX assumption, This may have O(10%) effect for CLAS12. The simplest way to correct this is just to sum 1+cos^2(theta) over your reconstructed simulated events in each bin, then divide by the number of reconstructed simulated events in that bin. If it ==1 then no correction is needed. However, the GlueX assumption may not be great.
Note BH has a (1+cos^2(theta)) /sin^2(theta) distribution.

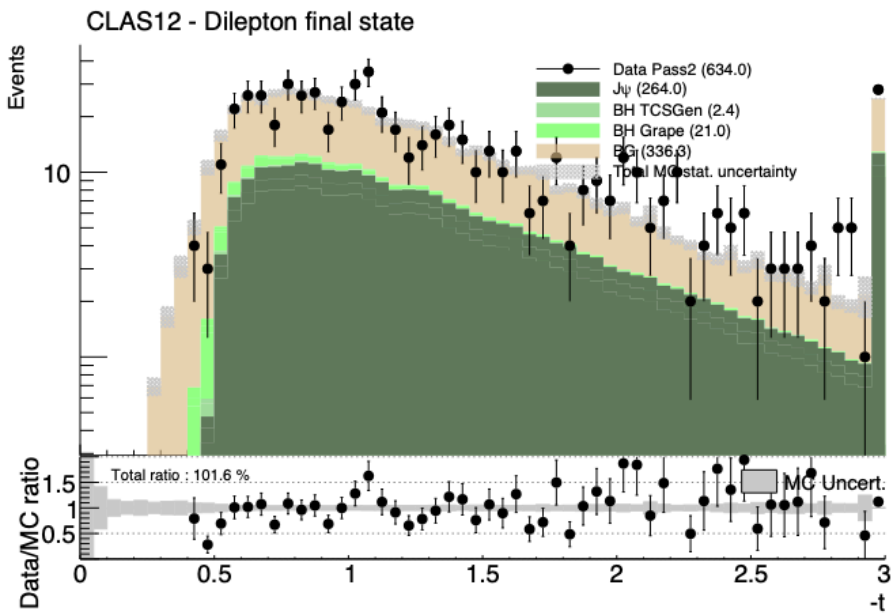
> To address the SDME dependence, the Gotfried-Jackson angles are computed and the effect on the cross section is added in the systematic section of the note. The effect is 10%. See comment at the start of this document.

What value of t-slope parameter was used in JpsiGen ? Is it consistent with the results extracted in this paper ? B = 1 or so ? This is used to integrate the acceptance for the cross section so good agreement is required.

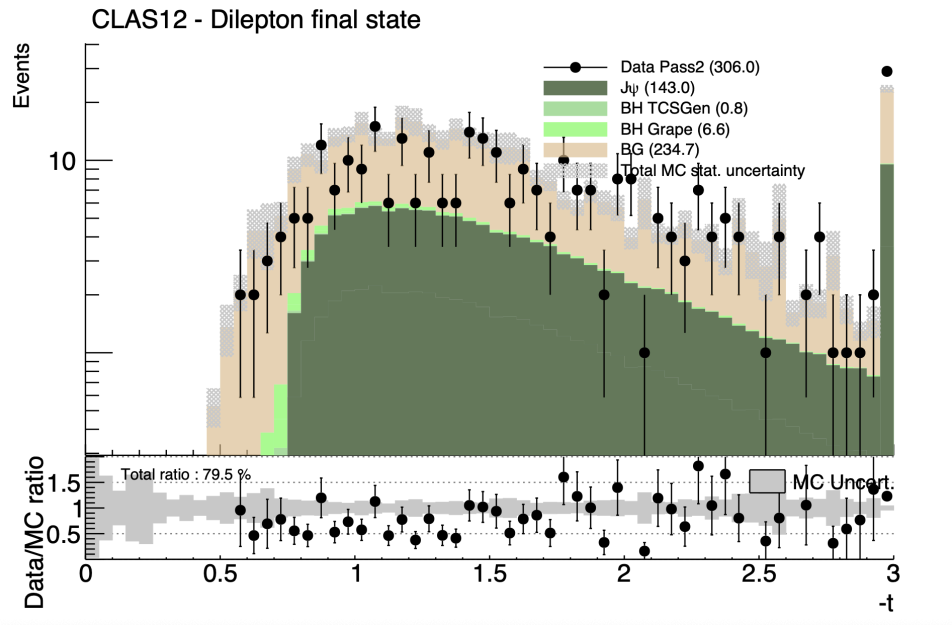
> The value used in JPsiGen is the one from the the fairly old Brodsky paper on 2 gluon exchange https://doi.org/10.1016/S0370-2693(00)01373-3. In this paper, the t slope of the exponential is 1.13. We find slopes from 1.02\pm0.27 to 1.31\pm0.14, which is in agreement with the slope used in JPsigen

Can you show Figure 41 projected onto the t-axis and compared to JpsiGen ?

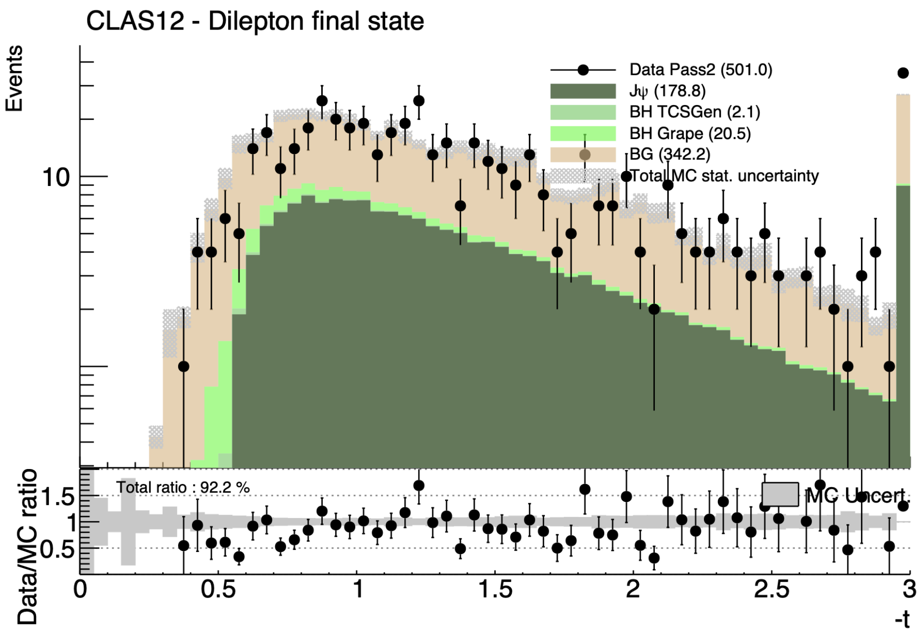
> Below are the t spectra including all contributions (BH and background) for each dataset and in the mass region of the JPsi. Note: the scale of the jpsi peak is roughly adjusted to match MC and data in this region. The shape agreement is a justification that JPsiGen modelization is under control and reproduces data well.



Inbending Fall 2018



Outbending Fall 2018



Spring 2019

6 Differential Cross Sections

It is not clear to me that when you bin over a large Egamma range you can reliably extract the slope from fits to t due to the varying tmin (actually also the t-slope varies). Can it be demonstrated that this works OK ? Via simulation perhaps ?

> The varying t_min is taken care by the bin-volume-correction. While the energy bins are large, the slope shown here will correspond to the average over the bin. As mentioned in a comment by Zein-Eddine, we mostly do this extraction for direct comparison with what has been done before.

7 Systematic Errors Study

General comment. Changing cuts by some arbitrary value then assigning the deviation as a systematic uncertainty is not a rigorous procedure (but unfortunately everyone does it). In particular where you end up with almost the same event samples (say 90% the same) then you will not see any significant deviation. This sort of thing should be banned! There is a nice work by Barlow on this https://arxiv.org/abs/hep-ex/0207026

> We agree that the method we used to assess systematics in this analysis can be improved. However, we disagree that this analysis should be the only one to need this. Suggesting better methods for the systematic evaluation should be done at a higher level (the collaboration level), maybe as guidelines.

7.1.4 AI lepton PID score "We assessed the effect of the cut applied on the score provided by the BDT-based Lepton PID by" - it is not clear, when doing this did you also change the value for simulation ?

> Yes for both Data and Simulation

The deviation between this response for simulation and data is likely to give some uncertainty. I.e try with different cuts for simulation and data (as simulation does not agree precisely with data)I did not see a plot for the BDT response in data compared to simulation. This would help understand if a 0.05 shift is sufficient.

The deviation between MC and data in the response, is taken into account in the overall correction factor. The cut is now varied by +-0.1. This corresponds to a range where the signal efficiency in simulation is above 80%. The same variation and justification was used in the TCS analysis which uses a very similar BDT. For comparison between MC and data see the dedicated note.

7.1.7 Normalization factor – The difference between the 2 factors is 0.16, hence as a % it should be 0.16/0.7 or 22% (Assuming this is a reasonable way to estimate the uncertainty)

This is corrected

8 Additional Checks

8.4 Consistency between measured differential and integrated cross-sections

“One can verify that, for each model, the integrated data are compatible with the integrated differential cross-sections.” – this does not quite seem to be the case for the high energy bin Fig 76

See answer given to Zein-Eddine on similar issue

F. Benmokhtar

I congratulate the authors for the thorough analysis note and their results, extracting the J/psi cross sections and the gluons gravitational FF. Here is what I learned from it:

The authors are measuring proton-photoproduction of J/psi of a proton target from an electron beam into a real photon, with production of electron-positron. The main motivation is to access the gravitational form factors through the cross section. There are also possible contributions from box diagrams and pentaquarks that need to be disentangled from the main physics goal.

Previous measurements existed from Jlab from GlueX and Hall C, where the gravitational FF were extracted from t-dependance analysis.

Data fall 2018 (inbending/outbending) and Spring 2019(inbending) runs, followed by common tools for these data sets.

Detecting the final state particle and reconstructing the missing scattered electron and ???, we can also present the invariant mass of the pair-electron (invariant mass of the electron pair as a function of Q2 and missing Mass of X which is the scattered electron.

Very long list of analysis: proton momentum corrections, radiative corrections, lepton momentum corrections, background subtractions, etc… Detection efficiencies, systematic analysis was done.

- General remarks:

a) The GlueX results are in red and the actual result in a cyan color. For the publication, it will be better to switch the colors, as the eye is first attracted to dark colors.

> Thank you for the suggestion. The color have been changed and I took the advantage of the comment to use the IBM colorblind safe palette instead

Detailed comments/suggestions/questions/concerns:

. Page 34- Figure 23: I personally have an issue with these figures. The data is much higher than the initial simulations, but this is not the problem, the problem is that the shapes are way off and the location of the J/Psi peak in the data (in M_ee for example) does not match to the simulations. Is there some kinematical shift that is not taken into account in the analysis?

> These figures (in section 4.1) only have an illustrative purpose. Indeed we show that the data are not well described by the monte-carlo when the latter only included JPsi and Bethe-Heitler events. The discrepancy is solved by including an additional background described in the following section (4.2). Once this background is included, we plot the same spectra and show that the agreement is restored. This method serves two purpose: first we show that we understand all the contributions of our data samples, second, by subtracting the background, we are able to determine how well the Bethe-Heitler signal is reproduced by the Monte-Carlo, giving us an estimate for the efficiency normalizing factor (section 4.3)

Was the bin migration done for this analysis?

The bin migration is part of the acceptance correction. For this, we used a simple ratio Recontructed/Generated obtained from JPsi Monte-Carlo simulations

. **Page 35**-Figure 24: Same remarks here. The problem is not the scale, but the shapes. Is there some Physical phenomena that was not taken into account?

See previous comment for Figure 23/ page 34

Page 35-Fig 25: The Xtitle should be Q2Tilda, to be consistent with the text. Please add a key for the purple points and the gray filling in the legend.

The blue points have been removed for clarity. The grey bands (the statistical error on the Monte-Carlo) have been added to the legend. The X label has been corrected to match the text definition.

Page 35, section 4.2.3: With reweighting, do you mean just normalizing the simulation bin by bin for each variable to get the same shape? There is a link to the source paper, but it will be good to explain in this note. Did you have to perform iterations to get to this match? There are some bin that still do not match. Do you have an explanation for this? Maybe I am just missing the definition of what is called reweighting. If it is collective re-weighting, then can you please explain how were you able to flip (left to right) the max of the simulation for theta electron for example, to fit the max of the data.

The reweighting technique uses a boosted decision tree to try to distinguish a source sample (our mixed event) and a target sample (the background events in the training region). Once this is done, one can use the ratio of the output of the BDT for both samples to reweight each event in the source sample. By propagating the weight hence obtained in histograms made from the reweighted source sample, one obtains exactly the target sample histograms. This explains for example the flip in the maximum of the theta spectrum.
Because we use this BDT in a region where it was not trained (we essentially extrapolate our model to the signal region), there can be some discrepancies.

Page 38, section 4.2.5, Validation: is directing us to Appendix K to see the agreement after the re-weighting in the signal region. I still see big discrepancies in some variables: like Pe' , theta_e', etc… Is there a way to re-weight? Maybe go through iterations (or more iterations?) reproduce the data?

The method is based on reproducing the spectra of variables in a different Q2 region as the one of the signal. As seen in appendix J, in the training region, the model is reproducing the data extremely well and thus there seems to be no advantage to do iteration.
The discrepancies arise from the fact that we use the same weights to reweight the mixed events in the signal region. Doing so relies on the fact that the background events kinematic do not depend much on Q2. While this is validated by looking at reweighted events in a high Q2 region (appendix L), there are possibilities that the background model does not capture all the variation of the real background. Nevertheless, the correction factor is derived using only the invariant mass spectrum, which is fairly well reproduced (Figure 124.a)

Page 39, Fig 29. We do not see full good agreement between data and simulation especially around Q2 of 2.2 and 2.7. Is there an explanation for this?

We found that the number of events is reproduced at the 10% level in the validation region. However we do not expect full agreement as the reweighting has not been trained on this Q_tilde^2 region.

Page 41 Figure 31, (b) missing 5 purple points in the range -0.1 to 0 of Mx^2 . (also the key for the purple points and the gray band)

The blue points represent the background-subtracted data. Sometimes in a given bin, the estimated background is larger than the data, and the blue points could be negative, and not showing on the plots. These points are not used, so we have removed them.

Page 42: Why aren't there purple points in the 2.6 to 3 GeV in the M_ee plot? This might indicate and support what I see that the data is much lower than the simulation in that region.

This is the same explanation as above. Additionally this plots is for outbending, where the dataset is dominated by background, making the background subtraction difficult to control.

Skip to page 53 (didn't understand much all what is being done in the pages in between).

See the answer of the comment for page 34 for a general overview of this section of the manuscript

Page 53: Figure 42. Why is the blue fit (the exponential stopping at 3.3? it should go further, so it will be probably lower, and the green Gaussian will have a higher amplitude.

The fit now extends to 3.4 GeV.

Page 55: Please explain what goes to GRAPE (why don't you use RADGEN + Pythia: calculates radiative corrections by simulating the emission of additional photons from charged particles involved in a scattering process, effectively accounting for the quantum fluctuations of the electromagnetic field and adjusting the kinematics of the primary event.) What does GRAPE do?

RADGEN+Pythia is not well tested in the kinematic regime of Q2~0. GRAPE has been extensively used for HERA analysis, and for the processes involved in this analysis. It also includes initial state radiative effects.

Page 58: What is the difference between Radiative corrections and ISR? Can you add some Feynman diagrams so we understand what you are correcting for?

A figure to illustrate ISR has been added in section 5.4.2.

Section 2.4 now has figures showing which diagrams are included in the radiative effects calculation

Page 62-64: the three figures can be gathered in one with a legend, so we will be able to see the evolution with Egamma.

Changed to the suggested layout

Page 68: I do not understand the AI lepton PID score. Please explain.

In order to assess the impact of using our Machine Learning particle ID, we computed the cross-section for various values of the cut applied on the output of the algorithm. The difference of these computations is assigned as the systematic induced by our Machine Learning algorithm.

Page 71: section 7.1.7 a difference in the normalization factors you've derived with two different methods is 16%, looks to large to validate the results.

See the discussion about this point in the general comment answer at the start of the document

Page 75: I do not understand section 7.2.4

See answer given for comment about Page 68

Page 76: section 7.2.5 Please explain how was the systematic variation associate with the proton PID measured?

The event builder of CLAS12 provide a variable call xi_PID for the proton. It corresponds to the difference between the expected time-of-flight (from tracking) and the measured one, divided by the resolution of the detector the tof has been measured in (in this analysis the FTOF). In this analysis, we did not apply any cut on this variable. To assess the effect of this choice, we recomputed the cross-section with cuts applied on the xi_pid of the proton (at 2 and 3 sigma). However, because of calibration issues, the width and the mean of this variable are not the same for data (for each dataset respectively) and for monte-carlo. We thus applied this cut based on the measured width and average of the xi_pid PID for each dataset and for each monte carlo sample.

Page 78. Normalization factor is 16% is this the same one as in section 7.1.7? if not, please explain.

It is the same. Added a link to the relevant section for clarity

Section 9 and up: please use a dark strong color for this data and lighter colors for GlueX. Agree with Zein-Eddine, should compare to more recent models.

See previous answer for the color palette. See Answer to Zein-Eddine for the model comparison