

Hi Buddhini,

Thanks again for all your work. Also thanks for a draft with line numbers, double spacing and a larger font. I really appreciate it!

I'm disappointed that you're not following the same practice I used on the Wien 0 PRL and the general NIM paper though. On those, I responded directly and explicitly to each and every comment from each and every person who wrote comments, explaining whether their change got made, and if not, why not. Well, at the 95% level anyway. At a minimum, those comments which don't get included deserve an explanation. It's a tremendous burden to write replies like that, but it's also a burden to make the comments. Mostly it's because everyone whose name is on the paper has to be on board with what's written in the paper, and if their objections get ignored, they need to be persuaded that was the right thing to do. In my case it sometimes happened that I was wrong to reject some changes that were asked for, and this only came out because I gave people the opportunity to argue with me about it. Sometimes it went the other way. A small subset of comments took a lot of research to check out and a lot of back-and-forth to straighten out. Ultimately I felt that writing and editing the papers I agreed to write and edit meant that I had to achieve a consensus from all the authors on the content of the paper. That is a huge job when we have so many co-authors. It can get heated and frustrating sometimes. But I felt strongly that it was an important part of my job as editor to listen carefully to my collaborators, even when I initially disagreed with what they were saying.

On the NIM paper, I even posted (almost) all those replies to the mailing list, because people in the collaboration wanted to see the discourse. It was a real pain, but this kind of transparency is something I think we should strive for. People certainly asked for it. They certainly deserve it as stakeholders in the paper. I also think it's important that if you reject a comment someone makes, that they get the opportunity to see your argument, so they can make a counter-argument if they feel so inclined. For example, you rejected many of my comments on the 1st draft, and since I don't know why, I don't know what to do about it. I also don't know if anyone else made the same comments, which might (or might not!) add weight to my own. Anyway I think you need to do this for the 2nd and subsequent drafts. I know it's a lot of work, but I don't feel too bad asking you to do this since I did it for the previous 2 papers. I think it's work that has to be done.

I thought I'd confine my comments to this file, where I have room to supply references and so on, and explain my comments more fully. So I'm just sending this, and not also an annotated scan of red ink on a paper copy like I did for the 1st draft. A lot of my comments are the same as for the 1st draft, because you didn't include them, and haven't explained why...

***The big picture:** De-emphasize the proton radius connection in the motivation section of the paper- it's too thin and too has too small an impact on the proton radius problem (unless I've misunderstood what you're saying there). It's only interesting as a side-comment, or perhaps we need to discuss this more thoroughly to see if a stronger connection can actually be made. Instead emphasize that for the 1st time, we present an asymmetry precision small enough that it's comparable with the theoretical uncertainties, and much smaller than the differences between the various calculations, which was not generally the case in the past. In the past the*

experimental errors were usually similar or greater than the differences between the calculations. So this new level of precision you present allows us for the first time to unambiguously distinguish between various models, and thus learn some TPE physics by looking at which model assumptions and ingredients are used in each calculation. Finally, compare to more experiments than you do. Our collaborators who were involved in those will insist on it anyway. But in any case, it's just good physics to compare to what has been measured previously. Finally, clearly define what was floated and fix the equations so they are consistent and transparent.

OK, now for the (very) detailed picture:

Lines 54-55: I think that greatly exaggerates and even misrepresents the (practically non-existent) TPE effect in the charge radius puzzle.

Lines 81-84: I'm not really objecting to your statement here, or the language you used here, which seems to me to be carefully constructed to avoid objections that might otherwise arise to this statement. But I think you definitely need to reference whatever source you have that supports this claim. Otherwise you probably do need to change the words "... seems to indicate TPE effects ..." to something more speculative like "... may indicate..." or "... might indicate ...". Now, to be clear, I'm not objecting to you pointing out that the widely known discrepancy in higher Q^2 GE/GM between Rosenbluth and pol. xfer is due in large part to TPE. Yes- no problem with that. I'm questioning the connection you make in that same sentence to ρ_M . I don't know where that connection comes from, even though it seems like a reasonable connection. Can you firm it up? I can't find any evidence for it.

Line 84-85: I just don't see anything in the Bernauer reference you provide [15] to support the statement you make here. It's possible I missed it, but I looked for it carefully. Bernauer doesn't apply TPE in his fit in that paper, so I don't see how he could possibly make such a statement. However, there are references you might use which do at least weakly support the assertion you are trying to make here that TPE effects might matter: examples include the FF fit from Arrington & Sick in PRC76, 035201 (2007). The most direct connection to the statement you are making about r_M that I could find is Bernauer's reply to the scathing criticism he got from Arrington in response to Bernauer's fit. In his reply, Bernauer explicitly reports the difference he gets in the magnetic radius when using TPE or not. That's in PRL 107, 119102 (2011). Although he reports a 1.5 sigma effect (0.026 fm) on r_M due to TPE, he says that means it's unimportant. You are trying to make the opposite conclusion from the same number. It's 1.5 sigma so I guess you could make a case either way... but it's pretty weak. There are other studies of TPE in the literature, like PRC76, 035205 (2007) or PRC76, 057601 (2005). But they don't report much of an effect either. So I am in favor of watering this down, unless you have some more solid evidence I didn't run across.

Line 88-90: Again we're back to this 7 sigma number. In the meantime I learned from Katherine why I think it's 3.4 sigma but you & her think it's 7 sigma. Anyway I guess the way you wrote it is technically OK, even though the increase from 3.4 to 7 sigma comes from atomic measurements, not ep scattering, so I'll back off.

Line 91: The 11% connection. I read ref [21]. It's tough going, but I don't see where the 11% comes from in that paper. Can you clarify that?

Lines 95-106: I still feel, as before in my comments on the 1st draft, that the connection to MUSE and to this whole proton radius business is way oversold. I would prefer to see the entire discussion on lines 75-106 watered down, because it doesn't seem to me that it is a convincing (or necessary) motivation for the experiment. Now, maybe it is, and the evidence supporting that just needs to be presented in a way that comes through better than it does. I suspect though, after looking at this essentially unchanged section from the 1st draft, that it just needs to be cut way back. I still think, after spending lots of time chasing down references and not finding the connections that are supposed to be in them, that we're hyping a practically non-existent connection, and that we don't need to motivate our paper that way. We can motivate it with stuff that's real, with the stuff you talk about later in the paper. I think the p-rad connection is mostly bogus, and that it should morph into just an interesting side note instead of being highlighted, as it is now, as the main motivation for the measurement, which it isn't. I would love to be wrong about this, because p-radius and MUSE are topical and interesting problems. But then I think you have to make a more convincing argument that it's relevant to our BNSSA measurement. If I'm not wrong about this, then it really should be watered down, and presented as a slightly interesting, potential additional reason to study TPE through BNSSA, even though the connection to the p-rad puzzle seems tenuous at best.

Can you point to the place in ref [23 or 24] where they say the TPE corrections can reach to 1.0-1.3%? I see in [23] where they say "*The estimates for the TPE correction of the muon-proton scattering cross section vary between 0.25% and 0.5%. These estimates are up to a factor three smaller, as compared with TPE corrections for the case of electron-proton elastic scattering in the same lepton kinematical region.*" Is that where you are getting your 1.3%? Shouldn't it be $3 \times 0.5\% = 1.5\%$ then? But, r_M depends on the SLOPE of the xsecs near $Q=0$, not the xsecs themselves, so I am still left wondering how much (if at all) this might affect r_M . And there in [23] the comparison is not to TPE/no TPE, but to TPE in mu-p vs TPE in e-p. So I am very confused about what this all means. Can you tell?.....

One problem I have, which I guess others also have, is that this subject is so unfamiliar to me I feel like we need to check the formulas really well, and that means tracking down each reference and each assertion to make sure it's right. That's hard! Takes time! But someone needs to do it.

Line after 106: Just a convention remark: Decide how to hyphenate beam-normal single-spin asymmetry. Then decide how to abbreviate it: BNSSA or B_n . I think you prefer B_n , fine. But then I think you should call it B_n and not keep on using beam-normal single-spin asymmetry. Consistency!

I could not find Eq. 1 in Ref [25]. I can find it in other places though, for example in PRL107, 022501, which is the G0 bkwd BNSSA paper you need to reference anyway.

What about Eq. 2? You need to reference where that came from also, unless it came from your own derivation. For epsilon (line 107), since that is a commonly used kinematic variable, you need to make sure to point to its definition (like in [23], for example) since you're using it as a polarization variable, if it's not defined in whatever currently missing reference you provide for Eq. 2.

line 110: need a space in Eq. 1

line 113: 3 words in, you need a reference.

Everywhere (multiple times, like on line 115: depend~~ance~~ → dependence).

Line 117: 0.1%. No, the expectation is a few ppm, not 0.1%.

Lines 121-123: Well, certainly PVES expt's are not motivated by leakage. So you might try replacing that sentence with something like "In fact, companion measurements of B_n are necessary in PVES expt's in order to properly account for the effects of residual xverse pol. in the beam."

Line 125: Could stick in a ref [25] but you do do it a few lines later, so it's not 100% necessary.

Line 128: ~~general performance of the~~

Line 130-131: Again, I think the jargon of run 1 and run 2, and the dates the runs took place, do not belong in the paper.

Line 135: (spin pointing up **at the target**). I like that you made this distinction before, since there is precession. Let's stick to it and make it here too. Also here, and many other places, you use "-" instead of "\$-\$". The former is a hyphen. The latter is a minus sign. They look very different in latex output.

Line 136: setup → apparatus

Line 137-138 This is pretty awkward. I'd suggest putting a period after Hall C on line 137. Then a new sentence put together from parts of your old one could read "The 34.4 cm long unpolarized liquid hydrogen tgt was contained in an aluminum cell with ~0.1 mm thick windows." Before it kinda sounded like the tgt was an aluminum tgt instead of a LH2 tgt.

Line 139: ~~the~~ Moller polarimeter → **a** Moller polarimeter. Add a reference to the Hall C Moller pol. paper (you'll find it in our NIM paper somewhere).

Line 140: extraneous space before the "%". Delete ~~for the full measurement~~ .

Line 141 & 143: is → was

Line 144: **I wish!** You quote our angle as 7.9 deg +/- 0.03 deg. That would be very nice. Where did the 0.03 error come from? I don't believe it, because if you differentiate our Q^2 equation with respect to theta, and then apply this error at 7.9 degrees, you find that our Q^2 resolution would be 0.7%! That's close to what we hope it will eventually be, but right now the tracking group is quoting 2.4%.

Line 144: ~~then~~ and perhaps also change elastic electrons to elastically scattered electrons. This latter change is maybe getting too picky so you can leave it if you want.

Line 148: use minus sign, not hyphen

Line 149: ~~and represents a 180° rotation in the plane.~~

Line 157: "in the scattering plane" seems awkward to me since we had 8 of them. I want to get rid of those words, how about: "...the measured asymmetry in a given detector has an azimuthal..."

Eq. 3: No, this is still messed up. First of all, it's inconsistent with Eq. 4, where you define B_n to have a different relationship to A_{exp} as you do in Eq. 3. I think you added the middle part of the equation trying to address an earlier question of mine, but this solution doesn't work I think. Sorry.

It's also inconsistent with definitions in the literature, like in PRL107, 022501. Worst of all though is that you still haven't made crystal clear and explicit where the floating constants appear in this equation, and say exactly what is and what is not floated in your fits. That HAS to be made clear. Reading it, I can't even tell whether some of this is floated in a fit to the xverse data and some to the longitudinal data or what. It's too confusing. I think the fit: what data are fit, what parameters are in the fit, what is floated and what not, has to be made 100% clear to the reader.

Line 168-170: "The experimental asymmetry...". Which one? The xverse or the longitudinal one? In the previous sentence you were talking about both. So now you have to specify which one. Also, isn't another way to interpret this sentence that the 2 additional fit parameters we use to correct the main PVES msrmnt were poorly determined in this fit? That would certainly be bad for Qweak. So I think this sentence needs work to be sure to explain clearly what you're trying to say here.

Tab. 1? Should that not always be Table x?

Line 171: similar kinematics? Similar? Weren't they essentially identical? I would just delete that phrase "~~Due to similar kinematics,~~" and start the sentence "The error-weighted average..."

Line 173: How was the 0.9938 determined and with what uncertainty?

Line 176-177: Why not either provide these corrections and their errors, or better, refer to our Wien 0 PRL for that info.

Table 1: Again, I think we should consider lumping the 2 vertical msrmnts together and getting rid of the run 1/run 2 jargon. But I guess that could be debated.

Line 188-189: What does "no correction was applied" and "the dilutions were taken into consideration" mean? How can both those statements be true?

Line 194: You say B_t is estimated to be 10^{-11} and you provide a reference to a private communication. That's good- thank you for the reference which had to be there. But, since it's just a private communication, we can't see what was used to generate the reference. Is there anything you can add to fill this out, anything you can say about how they made this estimate? Just in broad strokes?

Line 195: We usually try to avoid words like "me" "us" and so on to the extent possible. Why not say "in this expt." Instead of "by us"?

Eq. 4: Missing some kind of period or whatever.

Lines 198-209: Why not just refer to the Wien 0 PRL for some of these details? Or at least say if it's the same as in that paper.

Line 209: Add the reference to the Wien 0 PRL. Your thesis should also stay as a reference but use the PRL also. While we're talking about your thesis reference: it's almost useless without a link to it. You need to add a link in the reference itself at the end of the paper so people can actually find it. We did that for our previous 2 papers, and in fact didn't add thesis references unless they had a link.

Line 211- rest of the paper: Let's figure out a shorthand way to refer to the 3 calculations without having to provide first and middle initials, and last names for all authors, every time you refer to a given calculation. It's very awkward otherwise.

Line 217: similar behavior → similar angular behavior

Line 218: in our acceptance → at our energy, central values → normalizations

Line 220: TPE → TPE diagram?

Line 227: kinematics are → kinematics are the

Line 231: has → have

Line 232: "... and the continuum". What do you mean here? "...to the continuum."?

Line 233: "The possible under prediction..." → The over-prediction...

Line 234: Good, I'm glad you stuck in a reference to another experiment like I asked for. It's a good step in the right direction, but you need to do more. What about the many other msrmnts, like G0 bkwrđ (PRL107, 022501) or Happex/Prex (PRL109, 192501) or Sample (see Steve Wells). What about describing the other calculations even, in the G0 forward paper you did mention? There were 4 curves, you only refer to one of them. I think we can learn more physics from a more thorough comparison of your result with these others. And it's somewhere around here that you should hammer in the message I suggested in red at the beginning of my comments, regarding how our small error bar can finally help us distinguish between different models and the importance (or not) of the mechanisms they include.

Line 236-239: This sentence seems to come out of nowhere, there were several things about it I didn't like but thought could be mostly fixed by starting that sentence with something like the following: "The present expt., which has uncertainties comparable to those of the calculation of [29], clearly suggests that ..." Then later in that same sentence, I found the reference to the 2pi threshold in the cm frame to be weird. Why not either specify it in the lab frame (321 MeV), or even just drop it all together? Saying the 2 pi threshold is probably good enough, without having to provide a number.

Line 239-242: Another sentence that seems to come out of nowhere. It caught me off balance and made me wonder where all this business about photo-production and Compton amplitudes came from. It needs to be filled out and clarified.

Line 242-244: I think this sentence is too obvious, and should be deleted.

Line 244: If you delete the previous sentence, then start this next one without the leading modifiers "~~Therefore, as already shown,~~ this measurement..."

Line 248: Isn't it really only a tool to test DRs if you have precise real parts measured to test them against? Why not provide that info, via references, to make your argument air tight?

OK, this is obviously a lot of comments. I definitely would like to see at least another draft to see how this paper evolves. I would like to have a discussion with you on those points I've made you don't agree with. I think the elements of a great paper are here already. But we need to trim some stuff that doesn't belong, add some stuff that does, and clarify a few things. We'll get there together, no worries. I still think it's a great start and you should be very proud of getting us all here.

Thanks, Buddhini!

Greg