Full list of questions for CMD-3 arxiv:2302.08834

Separation e/μ/π

1. Energy deposition in LXe calorimeter: is it energy deposition for each of 7 layers or just information about energy in the whole cluster?
   Is energy deposition considered to be 0 when the track has no clusters attached? What is the probability of cluster loss for the data events (and the full simulation)?
2. Fig.3-4 show 2D-plots for the momentum and energy deposition methods at 2 CM energies, one where each method work best (0.5 GeV for momentum and 0.956 GeV for energy) and the other at their limit where they do not perform well but are still used (0.9 GeV for momentum and 0.548 GeV for energy). In the comparison with other experiments the problematic region is 0.6 - 0.8 GeV. Need to see the corresponding plots at these energies, i.e. 0.6, 0.7, 0.8 GeV.
3. Could you show the residuals for Fig. 6 and 7?
4. Can we see similar plots for CM energies 0.5 and 0.9 GeV?
5. Still in Fig. 6-7: the ππ and ee distributions show shoulders in the tails. What are the reasons for these structures?
6. The 2D reference distributions contain 36 and 57 parameters treated as nuisance parameters in the likelihood fit. Provide more information on the nature of these parameters, their time dependence, the checks with data and how they impact the systematic uncertainty on the cross section. Is it possible to show a data-MC comparison for individual PDFs, e.g. by applying strong cuts for one of the tracks?
7. How are you sure that there is no double counting, i.e. the corrections applied on the PDF don't include already some of the corrections mentioned after (like for example the ones in Section 4.2)?
8. Fig.8: the double ratio $N_{\pi\pi}/N_{ee}$ for the 2 methods is fitted between 0.6 and 0.9 GeV and found to be consistent with 1 within 0.2%. The fit is dominated by the large statistics at the ρ peak while uncertainties are much larger in the tails. Is it reasonable to quote a constant systematic uncertainty on this ratio of 0.2% throughout the range 0.381-1 GeV?
9. Show a blow-up of Fig. 8 in the range 0.7 – 0.82 GeV
10. Please show systematic uncertainty for the two separation methods (track, cluster) as a function of CM energy
11. The results of the 2 methods are averaged with weights given by their systematic uncertainties inverse squared. Show the plot of the weights as function of CM energy.

Efficiencies

12. e+e- →γγ (with one photon converted to e+e-, where one of the e+- takes most of the energy) can give a tiny contribution to the test sample. I guess it's pretty easy for CMD-3 to suppress such events, but this contribution cannot be exactly 0.
13. Fig. 11-12: what is the CM energy?
14. Fig.12: Are the differences between ππ and ee at the edges reproduced by Monte Carlo?
15. Fig.15: Is it possible to see the corresponding plot for the ππ efficiency? Why is there a large scattering of efficiencies on a scale beyond the error bars?

16. In Figs. 15 - 18 there seem to be very significant structures in the efficiency corrections (i.e. very significant differences between neighbouring points). What are these caused by? Maybe some fast time dependence?
17. In Fig. 17 a few points seem to have efficiencies larger than 1. What is causing this?
18. Tracking: clarify the separation made between 'base efficiency' (track selection cuts) and inefficiency from sources specific to particle type (decay, multiple scattering, bremsstrahlung, nuclear interactions).
19. Tracking plots are given for MC simulation only. Need to see data/MC tests.
20. No systematic uncertainty quoted for tracking efficiency resulting from data/MC tests, however overall efficiency inside fiducial volume is 94% (2018), 87% (2013)
21. "*Not all of such lost events were included in the test samples and were accounted for in the efficiency determination described in the previous section".* How are obtained these test samples?
22. "*The already accounted part of this losses for $\pi\pi$ events is about 30% at ρ resonance energies, ee 5% and $\mu\mu$ less than 10%.".* What the "accounted losses" do refer to?
23. How the sample with "subset of events with no any listed above processes" are obtained
24. Pion interaction losses: assumed nuclear cross sections in GEANT known to 20%, leading to quoted 0.2% systematic. Do you have a check comparing distributions of data with MC?
25. Pion decay losses: large effect even at ρ peak (1%) limited test from momentum distribution, but affects tracking, momentum, energy deposition, 0.1% quoted systematic. How well is the "decay" correction known?

Radiative corrections

26. Two generators used (MCGPJ, BabaYaga) NLO+NNLO approximative with some differences found for ee: give more information. Does it affect also the μμ and ππ samples?
27. A problem is mentioned for the momentum distributions with MCGPJ. Please show Fig. 6-7 using MCGPJ.
28. The problem is claimed to be partly cured by introducing an angular distribution for the photon jets. Is this physical? Wouldn't you expect different angles for each extra photon?
29. It seems to affect measurement only above 0.75 GeV for pions, but above 0.4 GeV for muons (Fig. 20, also 1.3% difference quoted p. 36, 10 x larger than the statistical accuracy). In Fig. 30 the agreement is with BabaYaga. Yet MCGPJ is used for pions. Please clarify.
30. How can you justify a 0.2% error for the ππ mode in MCGP given the large uncertainties seen for the Bhabha mode?
31. Why quote a systematic uncertainty on the RC only from form factor parametrizations in other experiments, since the iterative procedure uses the CMD-3 data and so should be self-consistent?
32. The RC are large +8% at 0.9 GeV and -9% at 0.7 GeV. What is the uncertainty specific to this analysis, from the used generators. The number 0.2% quoted is for the

integrated cross sections ('declared' by MCGPJ authors) , but apparently not listed in Table 2. Also what about NLO+HO differential cross sections? Need to be clarified.

33. In Fig 21 would be possible to distinguish the different sources of RC (ISR, FSR and VP) for the three sample (ee, μμ, ππ) also when Babayaga@NLO and MCJPG are used (for ee, μμ)?

General procedure

34. In Fig. 2 the BaBar statistical uncertainties within the enhanced bin sizes are probably computed using the published statistical covariance matrix, accounting for the (anti-)correlations between the original bins. Can you please confirm?
    If that's the case, please note that this matrix includes both data and MC stat uncertainties (e.g. from the unfolding), including also the uncertainty from MuMu luminosity, from background subtraction and from data/MC corrections.
    Do the values of the statistical uncertainties for the other experiments include the statistical uncertainties from the background subtraction and from the other corrections that are being applied?
35. The fact that the value for $N_{\pi\pi}/N_{ee}$ is consistent within 0.2% between angular fit, momentum and energy separation does not automatically imply that the systematic uncertainty is <0.2% if there are correlations between the methods (for ex. angular distribution and momentum separation, both using only tracks). Please justify.
36. Since 0.2% is quoted for the range 0.381 - 1 GeV can the test be repeated below/above the restricted 0.7 - 0.82 GeV range used in the test?
37. In Fig. 28 right there seem to be some structures (i.e. oscillations) in the shape of sigma(mu mu) / QED, while the chi2 fit is also not sensitive to these structures (i.e. one can obtain a good chi2/ndof even in presence of significant oscillations on sub-ranges). Has the significance of these oscillations been quantified and has their source been studied? It seems the fit in this figure is also not performed on the full sqrt(s) range.
38. What is the difference between the sigma(mu mu) / QED plot in Fig. 30 and the one in Fig. 28 right ? The above-mentioned oscillations seem to be absent in Fig. 30.
39. Would it be possible to have a prescription for the treatment of the systematic uncertainties in phenomenological studies? Are all the sources to be treated as independent between each-other, while each of them is fully correlated across the phase-space?
40. Section 6.2: What would be the agreement if one uses only MC samples (without additional data/MC corrections)?
41. Is there any way to check the correctness of the corrections obtained from the fit to the PDF vs the ones obtained by the fit in the angle? (It's unclear to me what the corrections from PDF refer to, and therefore having a dependence on the angle could be useful to clarify that)
42. DM2 results above 1.35 GeV are used to constrain the high mass part with ρ(1420) and ρ(1700). DM2 results strongly disagree with BABAR in this region with a large negative interference not seen in DM2 because of μμ background. Does it affect the fit in the ρ region?
43. Since it is only mentioned without any detail in the conclusion, can you clarify how the blinding of the results was achieved?

Comparison to other experiments

44. There is the same trend for CMD-3/BABAR or /CMD2 or /SND: CMD-3 excess of up to 5% around 0.7 ± 1 GeV (left side of ρ peak), excess extending to the highest energies for CMD-3/KLOE. But hard to distinguish the different contributions in Figs. 34-35. Plot separately CMD-3 fit/all other experiments to better assess the discrepancies.
45. Comment on the fact that accuracy for a_μ[0.6-0.88] is similar for CMD-2 and CMD-3 despite much larger data set for the latter.
46. The paper cannot avoid a study and a discussion concerning the CMD-2/CMD-3 strong discrepancy which are absent at the moment, despite similar detectors, analysis and group: outline the major differences in the detector and the analysis procedure, compare distributions, dig out where the problem occurs.
47. More generally can an effort be made between CMD-3 and SND groups (the 2 scan experiments running at BINP) to understand their discrepancy? Maybe the institute can help to straighten out the embarrassing situation?
48. The central values of the K+K-, π+π-, ancillary 3π measurements all tend to be higher than other experiments at a similar level of 4%, which of course for the 2π channel looks most spectacular. Have possible common systematic effects across channels been investigated?
49. What are the plans for publishing this analysis: short/long papers? Do you intend to perform additional checks before submitting to a journal?