B-factory Programme Advisory Committee
Full report for the
Annual Review Meeting

5-7 February 2024, Tsukuba, Japan

A. Andreazza* (Milano), P. Collins* (CERN), G. Corti (CERN), M. Demarteau (ORNL), R. Forty (CERN), B. Gavela (Madrid), S. Gori (UCSC), W. Hulsbergen*& (NIKHEF), M. Ishino (Tokyo), P. Mato* (CERN), F. Meijers* (CERN), N. Neufeld (CERN), B. Ratcliff* (SLAC), A. Petrov* (Univ. South Carolina), M. Sullivan& (SLAC), H. Tajima (Nagoya), M. Titov* (Saclay) and chaired by T. Nakada (EPFL)

& Remote participation

* Expert member.

N. Pastrone (Torino) participated remotely as a designated member

12 May 2024

Contents

1 Short summary

The annual review meeting of the B-factory Programme Advisory Committee (BPAC) took place at KEK from 5th to 7th of February, where the committee heard presentations from the accelerator and Belle II groups on the status and upgrade plan. This section gives the committee's feedback for the six questions asked by the management of the Institute of Particle and Nuclear Studies. A detailed report on the findings of the committee is presented in the following sections.

1. **Reviewing the overall project, is there anything on which the collaboration should focus more to strengthen the project?**
   The committee finds that the overall project, with the accelerator complex and experiment, is well focused on obtaining large statistics of high quality data, processing them rapidly and exploiting them for physics analyses.

2. **Are the objectives of SuperKEKB and Belle II upgrades well defined and achievable?**
   The accelerator group and Belle II collaboration have been developing plans for upgrading the accelerator complex and Belle II detector components during the next long shutdown, LS2, with the aim to reach a peak luminosity of 6×10^{35} cm^{-2}s^{-1} and to collect 50 ab^{-1} of data by around 2035. The Conceptual Design Report (CDR) for the LS2 upgrade plans, presented during this meeting, described various ideas for how to increase the machine luminosities, enhance the detector performance and the tolerance against beam backgrounds and radiation damage. Some of them are well advanced, in particular the new vertex detector system (VTX), which is envisioned to replace the current vertex detector, for which *the committee recommends that the group should advance towards the engineering level of studies, including the aspects of system integration and operation.*

   Some of the other subsystems appear to be still at the exploration stage and further technical studies are required. Although those items are marked as "long term", this prevents drawing a coherent picture of the upgrade plan. The machine group presented an upgrade path with a new set of final focusing magnets. Since they have just started machine operations after the Long Shutdown 1 with a lot of changes and improvements having been made, they need first to understand the behaviour of the machine in order to identify the remaining causes of luminosity limitation and to provide a definite scheme for the upgrade.

   While the CDR provides an excellent initial framework to continue the development of the plans for the upgrade, the BPAC is not yet in a position to judge the validity of the plan described in the CDR. While *the committee recommends further R&D work to continue, it also thinks that a new document providing a coherent and integrated description of the accelerator and detector upgrade plan with a quantitative demonstration of expected physics performance should be submitted for making a decision on the project.* Such a document should also demonstrate the tolerance against background rates at the peak luminosity of 6×10^{35} cm^{-2}s^{-1} and integrated radiation damage up to 50 ab^{-1}.

As for the idea to collect data with a longitudinally polarised electron beam, the proposed test of production of polarised e^- in the injector and measuring the polarisation decay time in the SuperKEKB ring will provide essential information for making further decisions. However, *it should be noted that the highest priority of the coming period must be to achieve stable machine operation* with a luminosity above 10^{35} cm^{-2}s^{-1} for data taking.

3. **Can the updated scheme of the data and run dependent MC production accommodate the current and coming high luminosity data?**
The committee supports the new format for the data recording being adequate for the forthcoming high luminosity run. Production of run-dependent simulation data will be crucial for controlling systematics in the analysis. Presented progress for the planning of large production of data is satisfactory. The committee took note of the increased effort in event generators and *encourages the software group towards further collaboration with existing activities elsewhere. Fast calibration by the subdetector groups remains critical* for smooth data processing.

4. **Have all the concerns about the detector operation been addressed?**
The committee was pleased by the progress being made in the detector hardware, online and detector control, on and offline monitoring framework, software and computing for the start of Run 2 data taking.

5. **Are physics analysis plans for the winter conferences and also toward summer convincing? Are we missing any physics opportunities?**
The BPAC considers the analysis activities presented for the winter conferences and beyond very adequate and convincing. Published results demonstrate that the Belle II collaboration has been exploiting well the advantage of being able to reconstruct final states with neutral particles including neutrinos. Merging of the Belle and Belle II analyses has also been successful. The committee is looking forward to hear the progress in the analysis of dark sector and tau physics, as well as other analyses relevant to the understanding of the muon $g-2$ anomaly. With Run 2, the Belle II experiment should be able to start collecting data with statistics competitive with those of the BaBar and Belle experiments, and *the analysis effort on the core B physics should be prepared.*

6. **Is the computing resource estimate adequate?**
The committee took note of the presented computing resource accounting for 2023 and estimates for 2025 to 2028. A dedicated discussion and recommendation will be given by the expert group consisting of G. Carlino, W. Hulsbergen, P. Mcbride and P. Mato, and chaired by the BPAC chair, after receiving the written accounting report and request.

2 Machine related issues

2.1 SuperKEKB and injection complex

2.1.1 Status

During Long Shutdown one (LS1), the accelerator team has worked on a large number of issues to improve SuperKEKB performance. Even though the linac was down only during the regular summer shutdown a great deal of work was done during those times. Several improvements in the linac include adding pulsed magnets (both quadrupole and corrector) which allows for bunch-by-bunch correction of the lattice and the orbit. This should improve bunch to bunch stability and bunch emittances at the injection. This will become important as clean two-bunch injection becomes necessary to keep the stored beam currents steady. The linac must also supply beam to the Photon Factory and the Accelerator Test Facility. The injection point for the HER has a new beam pipe and an improved local magnetic field. This should improve HER injection efficiency and hopefully improve HER injection backgrounds. In conjunction with the Machine Detector Interface (MDI) team for the detector, the accelerator team has refurbished all collimator heads for both rings and has installed the non-linear collimator in the LER. The non-linear collimator should allow for beam collimation at very close to the core of the beam without introducing physical material close to the beam core. This should allow for close collimation without perturbing the core of the beam bunch. If the non-linear collimator works as designed, settings of the other collimators in the ring can be relaxed a little and this should improve the Transverse Mode Coupled Instability (TMCI) threshold. The MDI and accelerator teams worked together on several other issues that are mentioned in the next section. The accelerator was just turning on when this review was going on so there is as yet no information on the status of these improvements.

2.1.2 Concerns

- There are still some mysteries about the emittance blowup of the HER bunch as it travels from the end of the linac through the transport line to the injection point into the main ring.
- The sudden beam losses have not yet been understood.

2.1.3 Recommendations

- The committee encourages the accelerator group to continue to study the emittance blowup of the HER bunches as they travel through the beam transport line.
- The committee also encourages the continued close collaboration with the MDI team in gathering more information about the sudden beam losses.

2.2 Machine detector interface

2.2.1 Status

The improved central beam pipe at the interaction point (IP) with the full PXD and a more complete coating of Au was installed during LS1. These upgrades should reduce the backgrounds seen in the PXD from back-scattered synchrotron radiation. In addition, the shielding has been improved at the face of the cryostats for the final focusing magnets (QCSs). Injection diagnostic tools developed by the MDI team to understand why some injected bunches have more background than others will continue to be a help in understanding injection backgrounds and in improving injection efficiency. The MDI and accelerator teams have also been working together to improve the detection of an unstable beam by doubling the number of local beam loss monitors and adding Beam Orbit Monitors. Acoustic sound detectors have also been installed in various locations to look for possible arcing inside the beam pipe. In addition, efforts are ongoing to further reduce the time between detecting an unstable beam and aborting the beam. If a beam that is going unstable can be detected and aborted before it causes damage or excessive radiation near the IP, then the rings can be rapidly refilled, and data-taking can quickly resume. This should improve the luminosity integration time and minimise downtime required for replacing damaged parts. Beam tuning assisted by Machine Learning was introduced for the optimised machine operation.

2.2.2 Concern

- The sudden beam losses still need to be understood.

2.2.3 Recommendations

- The committee encourages the MDI and accelerator teams to continue to study reducing injection backgrounds and increasing efficiencies.
- The committee also encourages the continued close collaboration between the MDI and accelerator teams in the study of the sudden beam losses.

3 Belle II detector

Many improvements have been made during the long shutdown LS1 for better Belle II detector performance. The new PXD2 has been fully integrated with the new beam pipe, with better SR shielding and cooling ability; the CDC gas and HV systems have been improved; MCP-PMTs in the TOP have been replaced with life-extended PMTs; there has been full migration to PCIe40 readout; the HLT units have been reinforced; and additional beam background shielding has been introduced. Improvements have also been made for more efficient data taking: Single Event Upset mitigation in the CDC front-end firmware; a non-zero standby voltage for TOP; a new injection veto for dead-time reduction; the introduction of a partial stop-abort-load-start (SALS) scheme; new data quality monitoring, environmental monitor and alarm system; new HV control and injection enable system. Detailed presentations were made at this meeting for the PXD, CDC, TOP and KLM, which are further discussed below. A visit of the committee to Tsukuba Hall allowed the improvements in monitoring and alarm systems to be demonstrated at the experimental control room, which was appreciated.

A comprehensive Belle II detector paper is being prepared, summarising the technical concepts and designs, and basic performance. The BPAC welcomes this initiative and encourages its timely completion.

3.1 Vertex detector (VXD)

Since the 2023 review, the Belle II collaboration has carefully planned and carried out to completion the full VXD removal and re-installation with the replacement of PXD1 with PXD2. After the PXD2 had been mounted on the beam pipe and its functionality verified, the VXD was extracted in May 2023. The SVD was removed from the extracted VXD, it was then integrated with the PXD2 in June and the upgraded VXD inserted back into Belle II at the end of July. The development of this process, a well-orchestrated effort of the communities involving the beam pipe, PXD and SVD, was also a topic of the June and November 2023 focused reviews. The committee congratulates the VXD community on the successful execution of the upgrade and is eager to see the performance of the detector after the restart of SuperKEKB beam collisions.

3.1.1 Pixel detector (PXD)

3.1.1.1 Status

A detailed update of PXD2 status and follow-up from the November review was presented to the committee, as well as the modified PXD management structure coming from the switch of focus from detector building to operation. The committee would like to express special thanks to Carsten Niebuhr, retiring from his position as PXD Project Leader, for his dedication during the planning and successful implementation of the PXD2 upgrade.

After the installation and services reconnection, all the 40 PXD2 modules (20 ladders, 8 in layer 1 and 12 in layer 2), are communicating and able to take data. The IBBelle CO_2 operational temperature was lowered from the $-20°C$ used with PXD1 to $-25°C$ needed for operation of the full PXD2 with ladders 2.7 and 2.8 powered-off. The performance has been tested with cosmic rays, measuring an efficiency $> 98\%$ in the active regions. Few modules have shown instabilities or noise issues; only one needed to be disabled because of excessive noise rate. Studies are ongoing to understand the origin of the observed behaviour and recover the affected module. Some DAQ instabilities are under investigation.

Focus of the attention for the detector long-term reliability is the bending of ladders due to thermal effects, which results in a sagitta of a few hundreds of μm for most ladders, but there are two consecutive ladders (2.7, 2.8) showing a sagitta deformation up to 1 mm in cosmic rays runs and there is a risk of clashing with neighbouring ladders.

Detailed studies have been performed at MPP and DESY to understand the behaviour of the ladders with particular attention to the glue joint. It has been observed that the Araldite in the glue joint softens above 46–48°C, resulting in a permanent deformation. The joint remains intact and the ceramic stiffeners do not detach from the modules. The FOS temperature monitoring shows a maximum temperature in situ of 38°C, providing some margin before this condition is met by the real detector. It should be noted that FOS does not measure the temperature at the glue joint and the measurements were taken in absence of beam loading. Endurance tests of 100k mechanical cycles with 1 to 2 mm sagitta at room temperature as well as thermal cycles between −30°C and 70°C without deformation stress on modules were performed and no failure has been observed, giving a better idea on the glue joint reliability. Any deformations will be carefully monitored with real beam conditions during future data taking. Ladders 2.7 and 2.8 will be initially taken out of operation, and their deformation will be tracked indirectly, using tomography from reconstructed hadronic interaction vertices, in order to assess the actual risk of clashing.

Documentation about the PXD2 installation and commissioning is gathered in Confluence, including links to database and elog entries, while GitLab issues are used to track progress in the commissioning. Collections of pictures to support the documentation have also been prepared.

For long term support of the PXD2 operation, a small-scale shadow system is being setup at DESY. It will be used to test improvement in firmware and calibration scans, and also to train new shifters. Availability of experts and shifters for the early 2024 run is good, in the long term a change of strategy is planned, relying only on on-call experts.

3.1.1.2 Concerns

- While much progress has been achieved in the understanding of the thermal deformation of ladders and the reliability of the glue joints, there are still open issues, such as the real operational temperature in the presence of beam load.
- Detector documentation is spread between multiple sources, and some may be lost if expert people will leave the PXD2 operation team.
- During the discussion, the worry was raised about the long-term support for PXD2 operation and the need for new commitments from the German or additional institutes.

3.1.1.3 Recommendations

- Continue the effort on the understanding of the risks associated with ladders bowing, in order to reach a decision about the temperature of operation of the detector and the switching on of ladders 2.7 and 2.8. A Finite Element Analysis may be useful to develop a quantitative model of the effect and to extrapolate outside of the regions for which data are available.

- Complete the documentation effort, including also full information about PXD1, which should still be considered a spare in case of accidents with PXD2.
- In order to guarantee the continued long-term operation of the PXD2 detector, providing a clear estimate of the required human resources, at KEK and remote, would be beneficial.

3.1.2 Silicon strip vertex detector (SVD)

3.1.2.1 Status

No specific SVD presentation was given in this review. Initial results after re-installation were provided at the November 2023 focused review.

After re-installation all sensors were operational and initial cosmic ray data showed that performance is well aligned with pre-LS1 results in terms of pulse shape and efficiency. Lowering the operational temperature of the IBBelle plant to -30°C has the positive effect to reduce the SVD L3 leakage currents.

The SVD readout has also performed the transition from COPPER to PCIe40 and the committee looks forward to feedback from the start of beam collisions.

3.1.2.2 Concern

- No particular concern is raised by the SVD situation.

3.1.2.3 Recommendation

- It is recommended to find a solution for the long-term availability of the excellent documentation developed for the VXD removal and re-installation.

3.2 Central drift chamber (CDC)

3.2.1 Status

Many improvements for the CDC gas system have been accomplished during LS1: water and oxygen sensors were re-calibrated, additional sensors for comparison and calibration, and extra gas lines to increase gas flow circulation rate have been installed. The water sensor had not been maintained for five years. As an outcome, the difference between the real water content inside CDC and the readout value was getting larger over the years, reaching ten times higher value ($\sim 10,000$ ppm) than that readout at the end of Run 1. Therefore, the degradation of CDC gain in the end of 2022 was predominantly due to the much higher water content inside the drift chamber. During 2023, the CDC water content was decreased, using silica gel tubes, down to the target value of 1600 ppm and no spurious activity or HV issues have been observed meantime. The CDC performance (median ADC) in 2024 cosmic runs with a water content of 1500 ppm was recovered as expected. This is much better than in 2022 and is comparable to 2017 data. The default oxygen sensor was also re-calibrated in the company and showed consistent output value with another O_2 sensor installed in CDC gas line. The O_2 filter creates small amounts of water inside the drift chamber. Therefore, new gas dryer system using N_2, and replacing silica gel tubes, will be installed for Run 2 data taking.

Significant progress in the implementation of data monitoring and alarm systems has been achieved during LS1. In particular, monitors for relative and absolute gas pressure installed near the detector, water and oxygen sensors, measurement of gas flow rates, and CDC leak current ratios are included into the DCS. Control room shifters are notified if thresholds are crossed. From January 2023, the global cosmic run has started with a goal to operate the CDC during Run 2 with O_2 content of 30 ppm, H_2O of 1500 ppm, circulation gas flow rate of $\sim$ 8 litres per minute, and a fresh gas flow rate of 200 cc/min.

A new test chamber for ageing studies was built using the same wires, feedthrough and gas sealing used for the Belle II CDC. The final target is to irradiate the test chamber up to $\mathcal{O}(1)$ C/cm with beta-ray ^{90}Sr source. Note that $\sim$ 6% gain reduction was measured at 1 C/cm in the earlier studies with Belle CDC chamber. In case future ageing studies will reveal severe ageing effects. such as larger gain drop and/or Malter effect, the same test chamber can be used to study different remedies and/or operation with hydrocarbon-free mixture (e.g. $He/CO_2/CF_4$). Chamber current, as well as water and oxygen content, will have to be monitored during the entire campaign. Regular pulse-height measurements with an Fe^{55} source, at irradiated and reference spots of the wire to compare gas gain, are envisaged. Meantime, simulation of the spatial distribution of beta-ray source intensity as well measurement of wire-hit rates will be performed.

3.2.2 Concerns

- Slow start of ageing study with a test chamber.
- There is a significant uncertainty in background projection for the post-LS1 CDC operating conditions, which can affect CDC triggering, tracking and PID performance. Assuming more favourable background scenarios, the CDC current could be of the order of $\sim$ 200 (500) μA per layer at a luminosity of $2(6)\times10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ that would result in a total accumulated charge in the innermost CDC layers of $\sim 2(5)$ C/cm after ten years. Such an accumulated charge exceeds the normal survival condition for drift chamber operation with a hydrocarbon-based gas mixture.

3.2.3 Recommendations

- Oxygen and water content, any sporadic activity in the chamber, e.g. increased sparking rate, leakage current or noise hits is of a paramount importance and should be carefully online monitored, since the luminosity and background rates will increase during the course of Run 2.
- The optimal set of operating conditions during the test chamber irradiation should be identified and finalised. The final target should be to irradiate the test chamber up to $\mathcal{O}(1)$ C/cm. Careful monitoring of parameters during the ageing test and regular Fe^{55} pulse-height measurements at irradiated and reference wire spots have to be done.

3.3 Particle identification system

3.3.1 Barrel system (TOP)

3.3.1.1 Status

In the period following the BPAC meeting in February 2023 and preceding the June meeting, the TOP group successfully performed the planned LS1 replacement program for the PMTs, and exchanged and repaired the front-end electronics, all within the scheduled time frame. This excellent work was discussed in the June 2023 report.

Additional projects for TOP operations were undertaken to improve the data taking efficiency during the upcoming runs at much higher expected luminosities. These included TOP boardstack recovery while running, and bringing the TOP to a non-zero HV during standby so that the transition back to full HV is faster than during the Run 1 data taking period. TOP monitor and alarm systems have also been improved as part of the new overall system discussed at the meeting.

The PMT quantum efficiency (QE) measurement program at Nagoya University has continued. A clear correlation was observed between the bench measurements at the university and those measured by the detector monitoring. However, the slope in the ratios between the two measurements was different from unity. After recalibration of the reference point assuming a quadratic dependence of the degradation with integrated charge, this appears to be much better understood, and the slope is now close to one, although there is still a small systematic shift.

Further lifetime tests are underway using PMTs removed from the detector. So far the tubes are not following the expected degradation curves, and, indeed, none show clear degradation to date at 20 °C. The tests will continue first collecting more data at room temperature and then at 40 °C to try to understand possible temperature effects on lifetime. There is some preliminary indication that high temperatures may play a role in the faster degradation for the conventional PMTs. Production and testing continue in order to have 224 new tubes from HPK available well before LS2 in case there is significant degradation in the remaining ALD and conventional PMTs during the high luminosity runs of the next few years.

Substantial work by the TOP group is underway to further understand and improve the TOP performance. Several improvements have been made in the analysis software between release 6 and 8, but the disagreement between data and MC persists and seems especially large in certain regions of dip angle at higher momenta, as shown especially by data plots distributed after the meeting. This clearly needs further analysis.

The group is also focused on preparing for the new running period, following the extensive changes implemented during LS1. Calibration is nearly complete. Cosmic ray data show the expected improvements in photon number in each slot based on the exchange of PMTs. The detector is expected to be ready to take data when the run commences.

3.3.1.2 Concerns

- It is unclear if non-operating PMTs sitting at a high temperature will age at the same rate as active tubes at the same temperature.
- The overall performance of TOP is not yet fully understood or well modelled.

3.3.1.3 Recommendations

- Efforts to better understand the loss of PMT QE, and especially to understand and mitigate the effects of temperature on the PMT efficiencies, should continue.
- Contributions from the TOP group to characterise and model the performance at the detector level are very useful and should continue as much as possible.

3.3.2 Endcap system (ARICH)

3.3.2.1 Status

ARICH is a stable detector that has generally run well during all data runs to date. There was no specific ARICH detector presentation during this review.

No hardware intervention inside the ARICH detector was scheduled during LS1 as it remained attached to the ECL. A low voltage cable was repaired, bringing five HAPDs back. However, six HAPDs had problems after endcap insertion, and had to be disabled. The ARICH group is still working to recover them. The number of dead channels is thus quite stable at 6.6%.

There were 17 noisy HAPDs in a localised area of the ARICH during the recent cosmic ray runs. The noise persisted even when the bias cables were disconnected. The noise has now disappeared, without being understood.

3.3.2.2 Concern

- The origin of the noise in the localised area of the ARICH is undetermined, and the effect on data taking if it were to return remains uncertain.

3.3.2.3 Recommendation

- Monitor the performance of the HAPDs carefully during the current run to be able to react quickly if the noise returns, and consider possible countermeasures in case it does.

3.4 K-Long Muon Detector (KLM)

3.4.1 Status

During LS1, the monitoring system of the gas flow was fully installed and the system commissioning is progressing in situ. The readout system of the environment data and the alarm system are integrated into the Belle II central system. Control room shifters should notice and report any problems, and the experts can take proper actions. It is expected a similar incident that caused low efficiency of the RPC BB2 layers, due to the interruption of gas flow and continuous operation under water vapour, should never be repeated.

As preparation for recovering the efficiency of the damaged BB2 RPCs by flushing the chambers with a gas mixture containing ammonia, the effect was tested using five RPCs, specially made for the test (A03, A08, A09, A10, and A11). The result is summarised that their efficiency dropped from e.g. 60% to less than 10% after injecting water vapour, and then following the treatment of gas with ammonia for two cycles, the efficiency was recovered to the original level for three cases out of five samples. In parallel to the test, a conceptual design of the gas treatment in the Belle II KLM system was presented.

3.4.2 Concern

- The benefit of the treatment with ammonia gas on the test RPCs is observed but it is limited to three cases out of five samples.
- The original efficiency of test RPCs used for the test was only 50 to 60%.
- A small amount of residual ammonia after treatment could damage the other good chambers.

3.4.3 Recommendations

- A review including external experts should be held before starting the ammonia gas treatment, to be sure any intervention will not result in damaging the other chambers with good performance.
- Ideally, the test of the effect for recovering the efficiency of RPCs should be made using chambers that have normal efficiency at the start, to draw reliable conclusions.

3.5 Trigger and data acquisition system (DAQ)

3.5.1 Status

A series of upgrades has been applied to the DAQ and High Level Trigger to improve both robustness and performance. The transition to the PCIe40 based readout is completed, additional CPU resources for the HLT will allow operation with at least 20 kHz rate. The raw-file writing and transfer to Offline has been simplified by using now the same ROOT format, as used subsequently. The performance and stability of the system in cosmic ray tests are very encouraging.

The L1 trigger has also been significantly upgraded. New, more efficient and complex firmwares are supported by more capable hardware, the UT4 board in several subsystems.

A strategy has been prepared to progressively tighten the trigger conditions if the limit of the DAQ should be reached.

New, improved triggers have been tested with cosmic rays where possible. Others will be commissioned with beam, particularly important among them the improved injection veto, which should significantly improve the overall data-taking efficiency of the experiment.

The committee is pleased to see that there is steady progress in attracting more institutes and people to the many important tasks in the trigger project. Significant progress has been achieved in many parts of the trigger and DAQ.

3.5.2 Concern

- There is still a known weakness in the *b2tt* link when running over some of the existing copper cables.

3.5.3 Recommendations

- If the workload of the team during running allows, priority should be given to finalise the tests on the optical version of the *b2tt* so that it can replace the copper links during the shutdown if needed.

- The efforts to secure and increase the person-effort and institute commitments for the trigger should continue.

3.6 Detector control and monitoring

3.6.1 Status

The Detector Control System (DCS) is the principal system for experiment control and provides monitoring and alarms to the shift crews in the control room as well as remote. The review and consolidation process of the DCS has advanced well during LS1. The HV control and injection enable mechanism have been given the highest priority and the new version is operational.

An existing alarm system with a Phoebus plugin was installed and deployed with DCS, DAQ and Data Quality Monitoring (DQM) and has already helped to identify several detector issues.

The development of the new DQM system has been mostly completed during LS1. The main control room DQM panels have been reorganised and a major revision of the DQM content of all sub-detectors has been done. The DQM is running in the basf2 framework with an interface to EPICS, including the alarm system. A flexible scheme for sub-detector configuration and reference histograms has been devised. The system is ready for the start of the upcoming physics run, while further consolidation can be done during the run.

The committee witnessed convincing live demonstrations of shifter consoles of these systems during a visit to the control room.

3.6.2 Concern

There is no particular concern.

3.6.3 Recommendations

- Documentation and consolidation of DCS and DQM components should continue during Run 2.
- A more powerful alarm system can be an opportunity for more automated actions to reduce the load on the shifters and experts. When the DCS team has some time once the system runs stably, this can be a worthwhile investigation.
- Investigate the usefulness of DQM plots with correlations between sub-detectors.

4 Computing, Data production, Software and Performance

4.1 Computing

4.1.1 Status

Small changes of responsibilities in the coordination of the computing project were presented at the last BPAC meeting.

The new raw data flow from online to offline deployed in LS1 has been successfully put in operation. Transfer of data between HLT storage to KEKCC local, then from KEKCC local to KEKCC grid and later to Raw Data Centres (RDC), has been further exercised in the recent cosmic-ray data taking. The committee appreciates that all possible issues on this critical data transfer between online and offline were successfully investigated. An additional safety margin of having enough storage for 40 days in case data-transfer would stop due to unforeseen troubles is in place.

Work has continued in re-designing the production system to produce smaller number of adequately-sized multi-run files for smoother analysis. Further integration of Rucio client tools with DIRAC to overcome some inefficiencies and to reduce possible delays has already been done in the BelleRawDIRAC instance and is being propagated to the BelleDIRAC instance. The collaboration informed the committee of their plan to replace the Metadata File Catalog (AMGA) with the equivalent Rucio functionality. The KISTI team assured them of AMGA support as long as needed. The committee sees this plan positively as it will result in a reduction in the number of components and in an overall simplification.

The collaboration intends to participate in the development of DiracX, the evolution of DIRAC with the goal to address some of its issues and bottlenecks. They are currently working on a prototype of the BelleDiracX extension. The plan is to deploy it in production only after 2025.

The move of the re-calibration site from DESY to KEKCC is ongoing. All the cDST data needed for the calibration hosted on TAPE has been replicated to KEK in a new Storage Element (SE). Replication policies have been updated so that new data will automatically be copied to the new SE. At the same time, the data staging for reprocessing has been further automated, but some manual operations are still necessary. It would be nice to pursue a full automation.

Hardware renewal of KEKCC will take place between June and August 2024. Migration of data from the current to the new system is scheduled for July-August. The new computing centre will probably have a similar CPU capacity as the current one, but the disk and tape capacities will increase by about 20%. Based on the experience of the last renewal in 2020, the collaboration does not expect any problem.

4.1.2 Concerns

- Production of output datasets without run boundaries is already partially implemented to reduce the impact of data fragmentation. Data fragmentation is further reduced with the use of skims for user analysis. However, the production of run dependent MC is still done run by run, and continues to suffer from this problem.
- Data productions often struggle to achieve 100% completion due to long tails caused by a number of issues. Several causes are responsible for few jobs getting stuck in DIRAC and/or in the DDM system. They need to be investigated with an accurate monitoring to be able to react quickly.

4.1.3 Recommendations

- Once the causes of jobs getting stuck are diagnosed, the changes necessary to increase the robustness of the data production system should be introduced.
- The committee would strongly encourage the full automation of data staging, including un-staging, and its direct drive by the production system.
- The committee strongly supports for the collaboration to directly participate in the development of the new DiracX, in the spirit of contributing to a community development, even if there is an obvious focus on the BelleDiracX extension.

4.2 Data processing

4.2.1 Status

The committee was given a report on *proc16* to reprocess all Run 1 data. *Release-08* of the *basf2* software with new calibrations for better physics performance, as well as several other improvements will be used. The full re-processing is divided in chunks to be processed sequentially, with the last chunk expected to be ready in October. The expected timeline is based on past experience and takes into account concurrent prompt processing of new data. Skimmed samples will be produced after each chunk. Production of generic run dependent Monte Carlo (MCrd) will proceed in parallel with the reprocessing of the data.

The migration of the Airflow job scheduler from DESY-NAF to KEKCC was not completed in time for *proc16*. For the replacement of the Airflow system the main complexity comes from its website interface. The plan for replacing it is to reduce complexity and keep only the essential features. A prototype is envisioned to use *b2luigi* as job scheduler and use a GitLab repository to collect expert feedback instead of a custom website. Personpower is needed for the implementation of such a replacement.

Prompt processing is expected to re-start in March using the same *release-08* software as the reprocessing. Prompt calibration will be done at BNL as in the past and generic MCrd will be produced in parallel with the data as for the re-processing. Run independent Monte Carlo (MCri) will be done in two halves: one with Run 1 conditions for easy comparison with previous campaigns, the other half with Run 2 conditions for first studies.

The last production of skims for data and MCrd samples included a better merge of files without run boundaries to reduce significantly the number of output files. This has a very positive impact for analysers reducing drastically the number of user jobs to run. A re-skimming campaign has started with 28 new skims and 21 improved ones, which has received enthusiastic response from Physics WGs. In addition, a new skimming strategy to group together skims with large overlaps and adding a flag has been proposed to streamline the skim production.

4.2.2 Concerns

- The calibration of the first *proc16* reprocessing chunk1 was run by hand due to the delay in the migration of Airflow to KEKCC. This is worrisome combined with the reduced maintenance for Airflow.

- Production of signal MCrd suffers from very long processing time due to the large number of signals (530) and only about seven signals that can be produced per week. The main reason for this is the inefficiency of running many short productions on the Grid distributed environment, due to the combination of number of signals and background data fragmentation for very short runs. Certainly running multiple signals per production will mitigate the problem, as well as skipping short runs with little impact on the physics. An alternative solution could be merging files of short runs into a larger multi-run file.

4.2.3 Recommendations

- The committee urges the collaboration to provide a prototype for replacement based on *b2luigi* as soon as possible as it allows for an easier support of calibration workflow.

- It is recommended to pursue the possibility to merge short runs with the same or very similar conditions into a single multi-run. This would also help in all subsequent productions and avoid, for every re-processing, MCrd production and skimming production having to cope with individual short runs and the overhead this implies.

4.3 Core Software and infrastructure

4.3.1 Status

The validation of *release-08* of the Belle II Analysis Software Framework (*basf2*) was almost completed at the time of the BPAC meeting, after the feature freeze in June 2023. The last patch release is already in use for data taking, however remaining issues delayed the start of reprocessing. The freeze of the next major release is foreseen for summer 2024, according to the regular yearly release schedule. The committee was informed that the collaboration is considering whether this may be delayed by one year, since no reprocessing is planned for 2025. While a validation campaign would allow in-depth checks of new features and focus development work, it would divert effort from other activities. Many new features are being developed in reconstruction and physics analysis software.

The use of Machine Learning (ML) algorithms is becoming more and more wide spread in all parts of the software. The full complement of available ML algorithms is exploited from classical Deep Neural Network and Boosted Decision Trees to state-of-the-art algorithms to deal with the increasing complexity of the analysis and reconstruction. Convolutional Neural Networks (CNN) are used in charged particle ID. Graph Neural Networks (GNN) algorithms are being used in Flavour Tagging and Tracking algorithms. Development of a GNN clustering algorithm for the ECL trigger to be implemented on FPGAs is in progress. ML techniques are also being used in simulation of detector response, as with Generative Adversarial Networks (GANs) for the PXD. Many algorithms are already integrated the *basf2* software, with a python/C++ interface provided in addition. The use of ML algorithms in the Belle II software is expected to rise due to the increasing amount of data and growing complexity of the analyses.

More and more automation of regular tasks is being introduced in the software, from monitoring of tracking performance to retraining and calibration of physics analysis ML-based algorithms. Verification of software releases is carried out exploiting the Validation Interface for Belle II (VIBE) tool, reducing manual workload. Good synergy is in place between the Performance and Software groups, as evident in the recent validation campaign. Analysts have become more involved in the development of new functional features, including the reconstruction software. The systematics framework is also key to a timely and complete validation of software updates and new features.

4.3.2 Concern

- The lack of a validation campaign following a major release until mid-2025 may accumulate potential issues and prevent finding solutions promptly. Focus on development of new features may also decrease.

- ML algorithms are becoming the preferred choice in HEP to cope with high data volume and complex problems. The inherent risk is to accept their successful outcome without deep understanding of the reason.

4.3.3 Recommendations

- A concerted validation campaign, even if reduced in scope and after a minor release, should be pursued, even if some effort may be diverted from other activities.
- Development of new functionality in reconstruction and physics analysis software is progressing steadily. Although resources are at an appropriate level, a suitable level of expert software developers should be foreseen for more in depth technical changes that may be needed, e.g. to cope with the emerging heterogeneous computing landscape.

4.4 Reconstruction software and performance

4.4.1 Status

The committee received reports on the performance of the charged and neutral particle reconstruction and identification. The committee is impressed by the progress in algorithms, monitoring and simulation, which has led to a improvement in the understanding of the data. The committee was pleased to learn that the agreement between data and simulation for the tracking performance has been improved, although the simulation tends to be more optimistic than the data. In particular, tracking efficiency and fake rates give very good agreement. It is also encouraging that the run-dependent simulation gives better agreement than the run-independent simulation.

The hit resolution and efficiency in the simulation of the SVD and CDC are essentially parameterised to follow the data. This has the benefit that analyses can use a well-calibrated simulation, in particular when they rely on the run-dependent samples. In parallel, efforts are ongoing to improve the actual CDC hit simulation by using a realistic computation of the field. This is expected to reduce the parameterised corrections. In the PXD simulation such a parametrisation procedure is not yet applied: as a result the resolution is poorly modelled in the region near the glue joint and in the very forward region.

At the track-level there is still a $\sim$10% difference in track parameter resolution between data and simulation in $e^+e^- \to \mu^+\mu^-$ events, with position parameters (d_0, z_0) more affected than $p_\perp$. This suggests that the high momentum tracks are still affected by data-MC differences in either SVD or PXD. The $K_S \to \pi^+\pi^-$ tracking efficiency in data shows a significant deviation from the simulation for larger decay lengths in the CDC, indicating issues with CDC tracking for off-IP tracks. The effect is parameterised to provide corrections for physics analysis but it is not yet understood. Finally, the large charge asymmetry in charged particle yields observed in *release-06* has been much reduced in *release-08*, but is still larger than 5% in parts of the phase space. This effect seems only partially reproduced in the simulation. Plans for the automation of tracking performance monitoring are in place and implementation is still in progress.

The understanding of the neutral particle reconstruction (photons, π^0s and K_L^0) has made impressive progress as well. A bug fix in *release-08* has lead to a large improvement in the K_L^0 reconstruction. The π^0 energy resolution has been improved in both data and simulation by tuning the clustering algorithm. The reconstructed energy of low-energy photons is calibrated using $\pi^0 \to \gamma\gamma$ decays. The bias and resolution observed in the data differ from that in the simulation. The effect is parameterised and corrections are provided. Several methods have been developed to calibrate the neutral reconstruction efficiency, including the use of $D^{*+} \to D^0\pi^+$ ($D^0 \to K^+\pi^-$, $D^0 \to K^-\pi^+\pi^0$), $\tau \to 3\pi\pi^0\nu_\tau$, and $e^+e^- \to \gamma_{\mathrm{ISR}}\omega(\to \pi^+\pi^-\pi^0)$ modes where different modes have different statistical and systematic errors. There remain discrepancies between different modes, and the Neutrals group plans to converge by mid-Spring. These corrections have been made available to analysts as well, reducing the systematic uncertainty on efficiencies extracted from the simulation to the percent level. The data-simulation agreement for the photon efficiency is good in general except for momentum less than 0.5 GeV/c in the endcap region due to material modelling in front of the ECL.

The particle identification (PID) calibration is in good shape. The PID group provides recipes for data-driven corrections to analysts. There is close interaction with the physics analysis groups to streamline the tools. Lepton identification is also significantly improved by utilising the ECL cluster shape. Unfortunately, a bug was introduced in the dE/dx simulation of the CDC in *release-08*, which led to a large drop in PID performance in the simulation. This bug was quickly spotted by the PID group and corrected in the latest patch release. The actual PID performance is still being improved as well, in particular by exploiting machine learning techniques, which has led to a large gain in K/π separation. The next step is to extend this approach to proton and lepton ID. The integration with the Systematics Framework is ongoing which will make analysis easier.

4.4.2 Concerns

- The variation of the performance of the CDC with time is a challenge for the simulation. It is expected that the variations will be reduced by the improvements in the gas system.
- The misalignment caused by the imperfect glue joints in the PXD may potentially have short-scale time-variations. It is unclear how this will be addressed in the detector alignment.
- For the SVD and CDC, tracking performance differences between data and simulation are reduced by parameterising resolutions. The risk of this is that contributions to the resolution that are the result of poor calibration (in particular alignment) will not be further understood.
- The data-MC differences in the impact parameters of high-momentum tracks seem not yet fully understood. The same holds for the K^0_S reconstruction efficiency at large decay lengths, and for the PXD hit resolution at forward angles.
- The observed charge asymmetry is potentially a limiting factor in CP violation measurements.
- The difference in the energy response (bias and resolution) for low-energy photons in data and simulation is not yet understood.

4.4.3 Recommendations

- The committee realises the importance of data-driven corrections to simulated events for physics analysis. However, the collaboration is advised to keep the development of low-level improvements to the detector simulation going. Such improvement of the simulation can be expected to reduce the dependence on extrapolation and hence systematic uncertainties.
- The committee recommends a detailed study of the PXD alignment near the glue joint, and, should this be necessary, the development of an alignment procedure to follow variations at the relevant time-scale.
- The committee encourages to continue on the efforts to automate tracking performance and neutral reconstruction monitoring, and effort should continue to resolve the remaining charge asymmetry issue.
- The committee encourages continued efforts to find causes of data/simulation discrepancies in particle identification performance.

4.5 Analysis software

4.5.1 Status

The committee was impressed by the report on calibration and improvements on main software tools for physics analysis provided by the Analysis Tools and Trigger & Event Properties groups.

Full Event Interpretation (FEI) is carried out with a Multi Variate Analysis (MVA) algorithm for two predefined types of decay modes, hadronic and semileptonic. Calibration factors, $\epsilon_{data}/\epsilon_{MC}$, are obtained from specific signals and provided for physics analysis. Hadronic-FEI calibration factors to be used in physics analysis for winter 2024 conferences were already provided for *MC15ri* while updated values for *MC15rd* were in preparation. An update was reported to be already underway for analysis targeting the summer conferences to take advantage from the changes in the decay tables and resulting in a $\sim 20\%$ improved yield. Calibration factors for the new Semileptonic-FEI were reported to be under finalisation with the aim to provide them for 2024 summer physics analysis. Different calibration factors are provided to chose on depending on the specific analysis. A new flavour tagger based on GNN, GFlaT, was shown to have better performance with respect to the existing category-based Flavour Tagger with an increase in efficiency of 18%. CP-violation measurements for $B \to J/\psi K_S^0$ with the new GFlaT are in agreement with world average values but with smaller statistical uncertainty with respect to previous measurements. GFLaT can also be applied to Belle data allowing to take full advantage of the higher statistics of combined analysis.

Overall a good L1 trigger performance during Run 1 was reported, with very high efficiency and good data-MC agreement. Three modes with large and well known cross sections are used for luminosity measurement. Better understanding of systematic uncertainties was reported from the use of MCrd samples; they are below 1% for the Bhabha and $\gamma\gamma$ methods. Luminosity measurements over data chunks are an essential input for B-counting and the main source of systematic uncertainties in the conventional method based on determining $N_{B\bar{B}}$ comparing on-resonance data with continuum. An alternative method independent of luminosity and off-resonance counting is under investigation.

The committee was pleased to hear of renewed effort on Event Generators, with a dedicated sub-group established to federate activities on the many generators used by Belle II. It closely collaborates with Physics Working Groups to identify areas of improvement and validate changes before deployment in the software, with the goal to improve agreement between data and physics models.

Recent improvements and current efforts in MC modelling were presented. An update of Branching Fractions (BF) in the EvtGen inclusive decay table was carried out with matching of $BF(B^+/B^0 \to D^{(*)}n\pi)$ to recent measurements and PDG values. Activities for EvtGen are now focusing on the review of charm decay models to identify necessary additions and improvements. This affect systematic contributions rising from the mismodelling of D decays in relevant analysis, e.g. $B \to K\nu\nu$. The committee was also informed of ongoing efforts to tune Pythia hadronisation parameters. The new tune was not deployed for use, due to inconclusive results from an extensive validation campaign and worse data-MC agreement in some distributions. A new strategy is now in place to cope with the complexity of the task, upon recommendation from the author of Pythia's reference tuning. Improvements were also reported across event generators used for low multiplicity and dark sector analysis, with particular attention to modelling of radiative corrections. New generators are being introduce to complement those already in use. Deployment of changes in the Belle II generator software suite is coordinated with the Software and Data Processing groups; documentation and tracking of changes is provided.

4.5.1.1 Concern

- No particular concern.

4.5.2 Recommendations

- It was reported that interaction with generators' authors is done through individual efforts. The committee encourages the collaboration to interact closely with relevant generators authors and provide them with contributions. It also suggests to be involved in HEP-wide initiative like the HEP Software Foundation (HSF) Generators Working Group to more widely share experiences and needs.

- Generator tuning is a complex task with many constrains. The committee encourages the collaboration to share results relevant for tuning in the form of HepData and possibly RIVET plugins as it would allow generators authors to include them in their own tuning. This would be valuable to both the Belle II collaboration and the HEP community at large.

5 Physics

5.1 Status

The committee congratulates the Belle collaboration for continuing to produce world leading results: in the past year, 18 new papers were completed, covering a broad range of physics topics. Among the most important ones are: the first observation of the $B \to \omega\omega$ decay; the first simultaneous inclusive and exclusive determination of V_{ub}; the most stringent upper bound on the branching ratio for $\Upsilon(2S) \to \ell\tau$ with $\ell = e, \mu$.

The merging of the Belle and Belle II analyses that was approved in June 2023 is very well organised and successful. As of now, all ongoing Belle (or Belle + Belle II) analyses, except for those in the final stage, are transferred to Belle II. These analyses will follow the Belle II procedure. With a few exceptions based on scientific merit, it will not be allowed anymore to start a new analysis with Belle only dataset. There are 36 new analyses that are following this merging. Also future talks at conferences will be organised centrally by the Belle II speakers committee, with a few exceptions for the Belle-specific analyses.

Many new interesting results have been completed by the Belle II collaboration. In particular, there are five new results since the last B2GM meeting in November. Among the most important ones are: the world leading bound on the lepton flavour violating decay $\tau \to 3\mu$; the world leading measurement of lepton-flavour universality in leptonic tau decays; the measurement of the cross section for $e^+e^- \to \pi^+\pi^-\pi^0$. This cross-section is important for the determination of hadronic vacuum polarisation (HVP) contribution to $(g-2)_\mu$ and shows a slight tension compared to the corresponding BaBar measurement (at the $2.6\,\sigma$ level).

Eleven additional analyses will be ready for Moriond or shortly thereafter. In particular, the first result on CP violation in charm (Belle II shows a much better resolution than Belle for the reconstruction of D mesons). In the summer, 40 new analyses are expected for ICHEP, utilising Run 1 data. In the future, the collaboration will want to leverage Run 1 data as much as possible, and, at the same time, prepare for the Run 2 sample. Belle II has the ambitious goal to run stably at above $10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in 2024, reaching a luminosity of $150\,\mathrm{fb}^{-1}$ per month, so as to exceed $1\,\mathrm{ab}^{-1}$ over the year.

The collaboration is pursuing the analyses of inclusive and exclusive B-decays generated by the electroweak penguin operators. Such transitions, mediated by the $b \to s(d)\gamma(\gamma)$ and $b \to s(d)\ell^+\ell^-$ quark currents are sensitive probes of new physics. The collaboration reported the first evidence for $B^+ \to K^+\nu\bar{\nu}$ decays by $b \to s\nu\bar{\nu}$ quark current. Sensitivity of the recently reported result for $B \to \gamma\gamma$ is approaching the SM level.

Last year, the collaboration started a new format for the Physics Week meeting. The topic of the meeting was V_{cb}. This is a very important topic for many reasons, including the anomaly in its measurements and the fact that it enters in many very rare processes that are sensitive to New Physics at high energy scales, like $B \to X\ell\ell$ and $K \to \pi\nu\nu$. The meeting saw a good balance of talks and discussion sessions. The goal will be to have a report by the Summer of 2024, that contains a prioritised list of issues that can be resolved on a 3-4 year time scale, and a proposed approach.

5.2 Concern

- The progress on the analysis of the dark sector (dark photon, ALPs etc.) seems to have slowed down, while these are prime and quite unique physics goals of Belle II.

5.3 Recommendations

- The committee encourages the collaboration to stick to the aforementioned luminosity goals.
- The committee recommends to concentrate the collaboration's efforts in analysing the largest contributor to the HVP contribution to $g-2$, the $\pi^+\pi^-$ channel.
- Since with the Run 2 data, the Belle II experiment should be able to start collecting statistics competitive with those of the BaBar and Belle experiments, the analysis effort on the core B physics should be prepared, including the analyses exploiting quantum coherence of the initial $B\bar{B}$ state.
- The collaboration should not slow down in the completion of low multiplicity analyses (tau and dark sector).
- The new format of the Physics Week was a success, and the committee encourages the collaboration to continue this type of meeting in the coming years. The committee is looking forward to a successful meeting on tau physics and dark sectors in the coming Autumn.
- The committee appreciates the collaboration's decision to release the full likelihood on HEPdata for the $B \to K\nu\nu$ decay analyses, and encourages to do this for other future analyses as well.

6 Upgrade plans

6.1 Overall scope and schedule

Despite the remarkable efforts made by the machine experts, the luminosity that was achieved at SuperKEKB in the runs prior to the first long shutdown, LS1, did not match the initial expectations, with significant backgrounds observed in the Belle II detector when the luminosity was increased. The maximum instantaneous luminosity achieved so far is $4.65 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$, while a value of at least $1 \times 10^{35}\,\mathrm{cm^{-2}s^{-1}}$ is required to allow $1\,\mathrm{ab^{-1}}$ to be integrated per year. During LS1 a large amount of work has been done on the accelerator complex to address the known issues, and the coming run will demonstrate whether those changes have been successful and the luminosity can be pushed to $10^{35}\,\mathrm{cm^{-2}s^{-1}}$ or beyond. Nevertheless, to reach the original goal of $50\,\mathrm{ab^{-1}}$ will require a further increase in luminosity to the design value of $8 \times 10^{35}\,\mathrm{cm^{-2}s^{-1}}$, or at least to the intermediate target of $6 \times 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, if that goal is to be achieved in a reasonable timescale during the next decade. This is expected to require a more radical reconfiguration of the machine, in particular for the final focus close to the interaction point. A second long shutdown (LS2) has been foreseen for those changes to be implemented, currently scheduled in 2028-29.

The Belle II experiment is also considering modifications to its detectors, to handle the higher background than foreseen and ensure the physics performance is maintained over the full lifetime of the experiment. The collaboration submitted a Conceptual Design Report (CDR) for the Belle II Detector Upgrades to this meeting of the BPAC [BELLE2-REPORT-2023-001]. As noted in the introduction to that document, it presents a number of possible detector improvements that could be implemented in the short or medium term, along with other ideas for longer term upgrades, that go beyond the currently planned programme. The various upgrades are largely independent of each other and can be selected individually on the basis of actual need, physics performance, technical readiness, and funding availability. In this sense they do not form a single coherent picture of the future Belle II, but rather a menu of possibilities. A different kind of opportunity is explored by the option of beam polarisation, which requires a SuperKEKB upgrade but no detector modification, opening a new and unique programme of precision physics, for which only a brief summary was presented. The upgrade plans of SuperKEKB to reach higher luminosity are also discussed.

The BPAC has reviewed the CDR, and recognises the large amount of excellent work that it contains. However, given the structure of the document explained above, it does not represent a coherent proposal for a coordinated upgrade of the experiment and collider on a common timescale, and as such the committee cannot make a recommendation concerning the document as a whole. Instead, comments from the BPAC review are provided here to the individual sections of the CDR, following the order of their presentation in the document.

6.2 SuperKEKB upgrade plan

Two options to improve the SuperKEKB performance for significantly higher luminosities were presented. They involve modifying the final focus cryostats and superconducting magnets for the larger dynamic aperture. The first option does not change the outer dimensions of the cryostat and moves the QC1P quadrupoles as close to the IP as possible inside the present cryostat envelope and, in addition, the inner radius of the quadrupole is increased through the use of thinner superconducting wire, which requires further investigation.

The second option modifies the outer dimensions of the cryostat and moves the final focus magnets closer to the IP and also includes an increase in the quadrupole radius. This option also extends the compensation solenoid toward the IP which will significantly improve the decoupling of the QC1 field from the Belle II solenoid field. These changes would have a significant impact on the available space for the detector, and the beam pipe components such as the IP beam pipe, BPM and vacuum sealing. On the other hand, it has a much better chance for making a large improvement in machine performance and further studies are needed to clarify how much expected improvement can be gained. Significant R&D work is needed before the new cryostat design can be finalised and a close contact with the vertex detector upgrade group is required.

6.3 Physics performance for the upgrade

The committee appreciates the efforts towards ensuring improvements on overall physics performance, and in particular to improve robustness against backgrounds and detector performance. The importance of progressing on the slow pion reconstruction efficiency from B decays has been convincingly argued. The foreseen improvement on K_L reconstruction allows for significant improvements in golden channels (e.g. on vetoing K_L or photons), providing an additional handle to suppress backgrounds for instance in channels with neutrinos or dark sector particles. This will affect key measurements such as the inclusive determination of $|V_{ub}|$, crucial τ-channel analyses such as those in $B \to \tau\nu$, $B \to D^{(*)}\tau\nu$, and also $B \to K\nu\bar{\nu}$, $D \to \pi\nu\bar{\nu}$ and $B \to \tau^+\tau^-$. The assessment of this impact remains at present rather qualitative, except for a few examples.

6.4 Vertex detector

The CDR outlined a proposal to completely replace the current VXD detector with a new all-pixel detector using a single technology. This replacement is motivated by the desire to have a replacement of the current VXD in case of severe accidents with the following advantages: being a more performant detector that is robust against machine-induced backgrounds operating at a luminosity of above 10^{35} cm^{-2} s^{-1}, and able to keep the community aligned with the main strategic R&D lines worldwide. Physics performance requirements lead to the detector having a range in radius of 14 ? 135 mm, a single point resolution better than 15 μm and a material budget of 0.2%/0.7% of X_0 for the inner/outer layers. To be able to withstand the expected running conditions, the detector should be able to handle a hit rate of 120 MHz/cm^2 and a total ionising dose and a neutron fluence of 100 Mrad and $5 \times 10^{14}\, n_{\mathrm{eq}}/\mathrm{cm}^2$, respectively, during 10 years of operations. The technology of choice is the depleted CMOS monolithic active pixel sensor technology (DMAPS) by TowerJazz (TJ). This would enable the construction of an all-pixel detector. With the constraint on the power consumption of less than 200 mW/cm^2 and an integration time between 50-100 ns, air-cooling seems possible for the two innermost layers.

The current layout has two air-cooled inner layers (0.2% of X_0) followed by three water-cooled layer (one 0.5% and two 0.8% of X_0) of straight ladders with pixel sensors with 33 μm pitch. The pixel sensor for the new VXD, dubbed Obelix, is based on the TJ-Monopix2 sensor developed within the framework of the LHC upgrades. This sensor has been tested extensively and meets the Belle II requirements in terms of efficiency, power consumption and radiation hardness. The design is currently being adapted to fit the Belle II geometry and trigger. The first prototype will have an 896×464 pixel matrix with 33×33 μm^2 pixels and is scheduled to be submitted in Q2 of 2024. The current tracker design calls for 117 straight ladders populated with 2552 sensors. Prototyping of inner and outer layer ladders has started. The inner ladders use blank silicon with a redistribution layer (RDL) to electrically connect the sensors and heaters to mimic the heat load. A notional schedule and cost estimate for a new VXD was provided.

An alternative to the DMAPS technology was described based on the Silicon-On-Insulator (SOI) technology, the Dual Timer Pixel (DuTiP) detector, but this option was far less advanced than the DMAPS option.

The performance of two VTX geometries was studied, which differ only in the location of the third layer. The nominal geometry of the new VTX detector has the third layer placed at a distance of 3.9 cm from the interaction point (IP), and the alternative geometry has the third layer placed at 6.9 cm from the IP. The performance of these two configurations are compared to the current VXD detector. The studies indicate that the new VTX in both configurations has a marked improvement in track reconstruction efficiency with comparable purity. Also a significant improvement in vertex resolution is obtained. Since the momentum resolution is driven by the CDC, no noticeable change is seen for the new geometry. The situation is different for specific processes. For the K_S^0 reconstruction efficiency, the simulation studies show that the current VXD has still better performance in terms of overall reconstruction efficiency. The new VTX, that has one layer less than the current VXD, shows a 5% to 10% falloff of the K_S^0 efficiency compared to the current VXD depending on the options for the VTX geometry.

The new VXD based on the DMAPS technology could now proceed towards engineering level studies, including aspects of system integration and operation. Special attention is needed for the air-cooling option and the electro-optical interface. Refined simulation studies are also required to define the metric for the design of the detector based on overall physics impact and to explore the option for adding an additional layer to the current layout to have the same number of space points as the current tracker.

Although the design of the new vertex detector is well advanced, it is very tightly integrated with the design of the new interaction region. Ensuring a coherent upgrade plan, well coordinated between the machine and the vertex detector, with commensurate timelines, will be essential.

6.5 CDC

The CDR describes medium-term CDC upgrade with an emphasis on the radiation hardness of components of the CDC readout. The main critical items on the boards are the optical transceivers that are very sensitive to the total ionising dose. Therefore, the current proposal is to replace these components with higher radiation tolerant optical transceivers. The increase in beam currents and luminosities also demands the CDC-based trigger system to effectively select physics events under high background conditions.

In addition, the CDC is one of sub-detectors that exhibited highest sensitivities to the machine related backgrounds during Run 1. Even without possible performance degradation due to high injection background, the CDC background level is estimated to be at the detector limit, both in terms of particle rates and the total accumulated charge on wires. In particular, the estimation of the total accumulated charge in the

CDC innermost layers after ten years of running varies between 2 and 5 C/cm per wire. This represents a significant risk for the long-term CDC operation and requires future studies for a back-up solution, even if a replacement of inner layers of CDC, or the whole chamber, is very challenging.

Given the limitations for the CDC mentioned above, the overall optimisation of the future tracking system consisting of gaseous and solid state detectors will be of paramount importance.

6.6 TOP

TOP has been working in an environment with higher backgrounds than originally anticipated at the operating luminosities achieved to date, leading to concerns about degradation in the quantum efficiency (QE) of the improved lifetime MCP-PMTs as luminosities continue to increase. This degradation is under intensive study as reported during this meeting, but it is not yet well understood. The default plan is to replace the lower life MCP-PMTs in the detector (at appropriate intervals) as they lose QE with newly produced life-extended tubes, while continuing to work with HPK to produce even longer extended-life MCP-PMTs, if feasible.

The CDR briefly describes a possible change of the photodetectors from the present MCP-PMTs to SiPMs run in single-photon detection mode at low temperature. This would be a major development project, requiring substantial engineering for detector cooling, and DAQ, detailed understanding of detector performance with these devices, and a clear understanding of their resistance to backgrounds. The BPAC appreciates the potential of this approach, but suggests that any such a major change is most appropriate to consider for a longer term than the five year time horizon being described, after substantial R&D and better understanding of the loss of performance of the present MCP-PMTs.

The TOP proposal also discussed a baseline upgrade to the readout electronics, both to reduce their size and power consumed. This would facilitate the implementation of the cooling needed to operate the SiPMs, if such detectors were to be installed. However, it is less clear of its importance under the baseline assumption which retains MCP-PMTs as the photon detectors. A sharper justification for this large scale project in the baseline scenario should be made as the costs and engineering development required are substantial.

6.7 KLM

Two options for the KLM upgrade are described in the CDR: (1) replace RPC layers with scintillators read by SiPM, or (2) change the operation mode of RPC to the avalanche mode and change the pre-amplifier electronics. The benefit of the upgrade options with fast timing determination will be studied with simulation for each case.

The RPCs have been operated since the beginning of the Belle experiment, thus it may be safer to replace RPCs with a new device. On the other hand, signs of performance degradation have not yet been seen. The replacement of KLM layers interferes with the services of the other detectors. The timing of the replacement should be studied carefully for the first option, while there is no interference in the case of the second option.

6.8 Trigger and DAQ

This part of the CDR begins with a general appreciation of the impact of machine learning (and more generally AI) techniques on many areas of modern computing and the related rise of more or less specialised co-processors to efficiently run these workloads, most notably GPGPUs but also FPGAs.

The potential of ML techniques, in particular for pattern-recognition and triggering tasks, is undeniable. On the other hand, the effort should not be underestimated to consistently and efficiently graft these algorithms and the heterogeneous hardware, where these algorithms run most efficiently, onto existing computing frameworks such as BASF. The LHC experiments have already been making similar efforts and Belle II collaboration can benefit from their experiences.

The trigger subsection describes a natural continuation of ongoing efforts to cope with high backgrounds. It is assumed that advanced algorithms used in the future will require significantly more hardware resources and it is proposed to upgrade the UT5 board based on a modern data-centre class FPGA. Given the complexity and cost of such future FPGA boards, the development effort should not be underestimated also for the firmware. The effort will require a robust team with a long-term view and continuous engineering support. When features are used or needed that are not in the focus of the mainstream users, good relations with industry, in particular the FPGA manufacturer, are important. Progress in the trigger will not only be vital for a major upgrade of Belle II, but also likely to benefit the current detector.

The upgrade of the timing distribution system described follows very similar lines as the discussion of the trigger. It is a natural evolution of the current system, which will benefit any future Belle II DAQ, regardless of whether the rest of the detector or the accelerator are significantly upgraded or not.

Upgrade of back-end DAQ boards is the subject of natural upgrade path of developing a board with much more powerful components available in the future. The greater processing power of these boards will also enable sophisticated event feature extraction that can offload subsequent processing in the HLT.

In addition to these natural evolutionary update ideas, it could be the right moment for the collaboration to think about more radical changes. For instance, the front-end electronics could be redesigned to use directly Ethernet as output protocol (on optical links). Then, the whole back-end design could be radically simplified and one could also use commercial FPGA cards for receiving data rather than producing expensive custom made back-end cards.[1]

A trigger-less readout is particularly attractive for the non-B physics program for Belle II, while the L1-trigger overall has a very high efficiency for B-events. The data rates are not discussed in the CDR. Higher data rates have an impact on cost, power

[1] For the timing-distribution and triggering it is much less obvious that commercial boards will be available, which can do the required tasks in a cost-efficient way.

consumption and cooling needs of the front-end electronics. On the other hand, electronics being developed for the HL-LHC could show that really large rates of data can be managed with modern technology. The Belle II collaboration should keep an open mind in exploring this option and to weigh the cost also with the decreased complexity in the system design.

6.9 Beam polarisation

Collecting 10 to 40 ab^{-1} of data with 70% longitudinally polarised electrons provides a prospect for new physics that could include:

- Unprecedented precision in the measurement of certain electroweak parameters, and specifically on $\sin^2\theta_W$ at c.m.s. energy of 10.58 GeV with a level of precision similar to that attained at measurements at the Z pole, that is, in a very different energy range.

- To measure the $g-2$ of the τ lepton with 10^{-5} precision and with 10^{-6} in the longer term, after a very high luminosity upgrade. This would enable meaningful tests of BSM theories such as MFV (minimal flavour violation) and others. The Pauli factor F_2 would be extracted from observable asymmetries, selecting τ-pair events in which both τ^+ and τ^- decay semileptonically, which allows the reconstruction of the production plane and direction of flight. The asymmetries between data taken with a left-polarised and right-polarised beam benefit from cancellations of systematic uncertainties associated with the detector asymmetries. However, those cancellations for the very high precision measurement require to improve the current uncertainties on m_τ and $M_{\Upsilon(1S)}$ (10^{-5} level) to the 10^{-6} level. Furthermore, a precision on the τ $g-2$ at the 10^{-6} level necessitates on the theory side the computation of two-loop effects presently not available.

- Measurement of the τ electric dipole moment at the 10^{-20} e·cm level would be an impressive two orders of magnitude improvement over present bounds. The improvement using polarised beams results because they enable measurement of the polarisation of a single τ, rather than measurement of correlations between two τ leptons produced in the same event.

Other interesting τ physics goals are also considered, including lepton flavour violation (LFV) in that sector.

The proposal described in the CDR to polarise the electron beam is in its early stages and further studies are needed to verify the anticipated performance and with minimal impact on the main goal of achieving maximum peak luminosity with a significant increase in the integrated luminosity.

6.10 Longer term upgrades

Longer term upgrades are discussed in the CDR for the ARICH and the ECL. For the ARICH there is the concern that HPK has stopped the production line for their HAPD detectors. Without an upgrade, there are no fallback options other than losing the phase space for particle ID, if significant numbers should begin to misbehave. SiPMs and LAPPD MCP-MCPs are discussed as possible alternative photodetector technologies for the upgrade, but either option will require significant further R&D before a concrete solution can be proposed.

For the ECL, long term upgrade options have been discussed to cope with the high occupancy and background rate while maintaining good particle reconstruction performance. They include replacing the CsI(Tl) crystals with pure CsI scintillator counters, that would be faster, or adding a preshower detector in front of the ECL. Complementing the PiN diode photosensors with avalanche photodiodes or SiPM are also discussed. Both the replacement of the crystals and the addition of a preshower are costly options. Therefore, further studies will be required to justify such upgrades.

Longer term upgrade of the CDC with a TPC-based tracking system coupled to Micro Pattern Gaseous Detector readout and dE/dx capability is briefly described. However, this change would also require major modification of the Belle II trigger strategy, which is currently largely based on the CDC.

6.11 Concerns

- As the machine group has just started machine operations after LS1, with a lot of changes and improvements having been made, they need first to understand the behaviour of the machine in order to identify the remaining causes of luminosity limitation and to provide a definite scheme for the upgrade. In addition, the R&D presented as necessary for preparing the upgrade of the final focus quadrupoles would last at least three years, and the final construction could take a further three years. This implies that the current schedule foreseen for LS2 will be difficult to maintain.

- The current CDR does not provide a coherent overall picture for the detector upgrades. The detector upgrade plan may not be compatible with the timescale needed for the machine upgrades. The final design of the upgraded detector is likely to depend on the choices made for the machine upgrade, as those may impact the geometry at the interaction region. The metric for selecting among the proposed detector upgrades was not well-defined in the CDR. Although some performance studies are included it lacks a complete study and a clear definition of the parameters that will drive decisions, concerning surviving the high background conditions and maintaining or improving the physics performance.

- The addition of longitudinal polarisation would provide an interesting extension of the physics reach of Belle II, but would imply significant interventions on the machine which risk to perturb the current priority of increasing the luminosity and integrating a large dataset. For the longer term with very high statistics, the τ $g-2$ goal of the polarisation programme requires external pre-conditions on two fronts: 1) to improve the current uncertainties on m_τ and $M_{\Upsilon(1S)}$ (10^{-5} level) to the 10^{-6}

level; 2) the theoretical estimation of two-loop effects, whose implementation lies beyond the collaboration efforts.

6.12 Recommendations

- The R&D required for defining the machine upgrade should proceed in a timely manner, to avoid any unnecessary delay to the schedule for the upgrade, and the timescale for LS2 should be adapted as required.
- Studies are well advanced for the new vertex detector system (VTX) that is envisioned to replace the current vertex detector. Work on this detector should advance towards the engineering level studies, including the aspects of system integration and operation. This will enable progress to be made quickly on the final design once the machine upgrade has been defined.
- Further R&D work should also continue for the other proposed detector upgrades, but the BPAC considers that a new document providing a coherent and integrated description of the accelerator and detector upgrade plan, with a quantitative demonstration of expected physics performance, will be required once the overall scope has been defined, for a decision to be made on the project. Such a document should also demonstrate the tolerance against background rates at a luminosity of $6 \times 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and radiation damage corresponding to an integrated luminosity of $50\,\mathrm{ab}^{-1}$.
- It is important to make the progression from qualitative to quantitative statements on the expected physics impact of the upgrade (for instance the KLM upgrade) and to demonstrate the overall performance of the upgraded detector for the main physics channels.
- The committee appreciates that the proposed test of production of polarised electrons in the injector and measuring the polarisation decay time in the SuperKEKB ring will provide essential information for making further decisions on the idea to collect data with a longitudinally polarised beam. However, it should be noted that the highest priority of the coming period must remain to achieve stable machine operation with a luminosity above $10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for data taking.