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

21-22 February and 28 February-1 March 2022 Remote Meeting

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),
V. Luth (SLAC), P. Mcbride* (FNAL) P. Mato* (CERN),
F. Meijers* (CERN), N. Neufeld (CERN), K. Oide* (KEK),
B. Ratcliff* (SLAC), A.A. Petrov* (Wayne State University),
M. Sullivan (SLAC), H. Tajima (Nagoya), M. Titov* (Saclay)
and chaired by T. Nakada (EPFL)

* Expert member.

27 May 2022

1 Short summary

The 2022 annual review meeting of the B-factory Programme Advisory Committee (BPAC) was held remotely covering the current status and the future activities for the accelerator and Belle II experiment.

The committee was informed that the recent sharp increase in the electricity cost has forced the machine running time in the Japanese Fiscal Year (JFY) 2022 to be reduced almost by half. As a result, the Long Shutdown 1 (LS1) will start already in summer 2022. This was the original plan, which was changed to start after the autumn run to accommodate the delays in the production of the replacement beam pipe at the interaction point and the new pixel vertex detector (PXD2) as presented during the last BPAC focused review in November last year. The total length of the shutdown will now exceed 15 months. This change will have an impact not only on the LS1 planning, but also on physics analysis and computing resource requirements due to the smaller integrated luminosity collected before the start of LS1 and also the longer shutdown period. The Belle II collaboration has been adjusting well to the new situation.

The performance of the injector and SuperKEKB has shown steady improvements in the operation and hardware stability, resulting in over 50 fb^{-1} of data delivered during Run2021c. It is also satisfactory to learn about the success of the energy scan and the committee is looking forward to hearing the physics results from these data. There is also progress in the beam background studies. Close collaboration between the accelerator and Belle II teams has been crucial for this success and should continue. In this context, the position of Background Fellow should be wisely used. The committee is pleased to note that a plan has been worked out for the upgrades of both the injection chain and SuperKEKB during the LS1. For improving the quality of the beam injection, a feasibility study for constructing straight injection lines using the existing tunnels is recommended. Since the injector also serves the photon facilities, the construction of new injection lines may require a KEK-wide discussion. However, it is worth serious consideration given the potentially high gain.

The hardware status of the Belle II detector has been generally stable. The problems of the efficiency loss and dark current of the Central Drift Chamber (CDC) have so far been stabilised. However, adequate human resources should be secured for the careful monitoring of the CDC behaviour and the long-term ageing tests with a test chamber. Considerable effort has been made to understand the efficiency losses in the photon detector (PMT) of the barrel particle identification system. Many but not all of the ALD type PMTs that are showing unexpected gain losses seem to come from a particular batch, and further investigation is required. It remains unclear if there is a temperature effect in the efficiency losses. The committee understood that no decision has been made yet about when to replace the aged PMTs. The Belle II group plans to take out first some of the PMTs during the LS1 for laboratory testing, so that the results can be compared with the in situ performance measurements. With this comparison, the collaboration will decide which PMTs to change at which moment based on the in situ measurements. Since the time needed to replace the PMTs fits within a normal summer shutdown, it is not mandatory for the replacement to be done during the LS1. The committee considers this a sensible approach and urges the collaboration to define the procedure and performance criteria for selecting the PMTs to be replaced. The committee also recommends further procurement of the ageing resistant PMTs (lifetime extended ALD type).

The two major items to be replaced during the LS1 are the beam pipe at the interaction region and the two-layer pixel vertex detector (PXD2). For the fabrication of the beam pipe, the group is proceeding with all the necessary precaution, but it is delicate and the production schedule remains tight. However, a backup plan with the existing components seems to be well worked out. With the current extended LS1 schedule, there is now adequate contingency in the PXD2 production schedule. On the other hand, the committee is concerned whether there will be enough good detectors to fully equip the two layers of the PXD2. The PXD group should make the utmost effort that there will be enough spares and consider further production of the sensor modules. The group may also need to develop a plan of how to configure PXD2 if the layers cannot be fully populated.

The committee is pleased to see the continuous effort to understand the detector performance, which is crucial to reduce systematic uncertainties in the physics analysis which may soon dominate the errors. While high level performance analyses with sophisticated algorithms have been showing promising results, further improvement of the detector-level understanding of the data would be beneficial.

The data taking and offline data processing have been smooth, where many operation processes have been automated with resulting short error recovery times. There is still a bottleneck in validating the calibration due the the limited number of detector experts, which resulted in a delay in producing the run dependent simulation data and processing the RAW data. Since this also affects physics publication, as well as the computing resource requirement, utmost care should be taken to find a solution.

The physics analysis effort with Belle and Belle II data has been very active with an impressive number of publications. It is now clear that the anticipated amount of data by the start of LS1 will be limited to $\sim$500 fb^{-1} for Belle II due to the change in the run schedule. Yet, by combining with Belle data and applying improved event reconstruction and analysis methods, the collaboration should still be able to produce very competitive results on beauty and charm physics, as well as work on the spectroscopy and search for dark sector particles. The committee acknowledges the effort made by the collaboration to develop a strategy to cope with the LS1 period that now exceeds 15 months. Efforts made to get analysis people more actively involved in the trigger and performance studies is highly appreciated. The committee encourages to extend the effort to other areas such as the detector alignment and calibration, and the software for reconstruction and analysis tools.

In this meeting, further information was presented for the upgrade plan, mainly for the Long Shutdown 2 (LS2) in a timeframe of 2026 to 2027, with an emphasis on a new pixel vertex detector. There are only four to five years before the start of LS2. If this time scale is to be kept, the upgrade effort must be much more focused and a more detailed plan has to be worked out. Further discussion of other possible physics analyses would also be appreciated.

In conclusion, the committee is generally pleased with the status of the machine and experiment and encouraged by the progress. The biggest concern is clearly the significantly reduced running time with very uncertain future due to the electricity cost. The effect of the war in Ukraine (that broke out after this meeting) may be temporary, but the pressure to reduce energy consumption will continue. While the effort to secure more funding is important, KEK may also need to develop a plan for a more sustainable operation of the laboratory.

2 Accelerators

2.1 Injectior and SuperKEKB

2.1.1 Status

The SuperKEKB accelerator presentation displayed a much improved and more steady machine performance. Many fixes and improvements were made during the summer shutdown and these have improved stability and reliability. In addition, particular studies to understand the dynamic aperture using rotated sextupoles has yielded improved performance. It was found that the very old thyratrons used for the Low Energy Ring (LER) injection kickers were unreliable and they have been replaced. This has greatly increased the LER injection reliability. The High Energy Ring (HER) injection efficiency has also improved through more careful modelling of the septum magnetic fields and through careful tuning. Procedures for increasing the stored beam currents have been developed and they have allowed the LER beam current to be raised to above 1 A. Startup procedures have also been developed and have resulted in a smoother and faster machine startup after shutdowns. There have been no rapid beam loss events during this last run. The vacuum leak in the cryostat of the final focusing magnet (QCS) has so far remained stable. The scrubbing of the LER beam pipe has slowly lowered the dynamic pressure allowing more beam to be stored.

2.1.2 Concerns

- Although there were no rapid beam loss events, the reason for this is not fully understood.
- The emittance through the HER beam transport line, though much improved, is still high.

2.1.3 Recommendations

- Continue to improve the fast abort system. If the unstable beam can be removed as soon as possible, the number of events with high backgrounds and/or QCS quenches will be minimised,
- Studies of the optics in the HER beam transport line should be continued with the intention of further improving the HER injection efficiency. In addition, the possibility of upgrading the HER transport line by using the straight ahead tunnel now used for the photon factory should also be studied.

2.2 Machine background

2.2.1 Status

The backgrounds in the detector have generally improved over the last run. Initially, the backgrounds started out high, which is generally expected after a long shutdown. After some maintenance work on the machine on November 24$^{\text{th}}$, the backgrounds were very significantly reduced. The cause of this reduction is not yet understood and one machine development period was used for attempting to increase the vacuum pressure locally to see if the old background rates would return. The result of this test was not conclusive. However, it is still a possibility that there was a very local high pressure area that was not detected by the vacuum gauges. Investigations are continuing. After this significant decrease, the background rates are very much closer to the expected rates from simulation and the detector background limits are well above the current machine running conditions. The lower background rates can now allow for the possibility of opening the collimator settings in order to raise the Transverse Mode Coupled Instability (TMCI) threshold. The LER beam-gas background is no longer the dominating background with

Touschek scattering now becoming more important. This is very good news as the final dominant backgrounds for the machine at design running conditions should be Touschek and luminosity related. Collimator optimisation has also been improved and the collimator settings now in use also improve the detector backgrounds. New injector background studies have also improved with the addition of machine learning codes. There were no "catastrophic" beam loss events during the last run. There is no known reason for this, so vigilance is still important and fast signal monitors like CLAWS are being included in the beam abort system. Procedures have been developed for startup after a long down time and for raising the beam currents both of which should help to minimise sudden beam loss events. Earlier detection of collimator damage has also improved and this should also help to understand the background rate versus various collimator settings. Damaged collimators produce more background than expected and in addition they can lower instability thresholds by introducing a larger than expected local impedence.

2.2.2 Concerns

- It was mentioned that the beam pipe in the final focus cryostats may be introducing more beam impedance than is expected.
- The HER Touschek background rate is still higher than expected.

2.2.3 Recommendations

- Whether or not there is indeed more impedance than expected in the cyrostat beam pipes should be verified.
- The committee encourages the background team to continue studying the cause of the higher than expected HER Touschek background.

3 Belle II experiment

3.1 Detector operation

3.1.1 Status

The Belle-II detector has recorded about 54 fb^{-1} of integrated luminosity in run 2021c, bringing the total to 268 fb^{-1}. The overall data taking efficiency was 86.8%, close to the target of 90%, but still with room for improvement. A significant fraction of the inefficiency came from teething trouble due to the migration of the readout from the COPPER to PCIe40 system. Compared to earlier running periods, the background level decreased after a maintenance day in November. There is concern that this change in background conditions is not understood. Closer monitoring of the beam background changes will be necessary. There is a potential lack of personnel, in particular local experts (e.g. for the PXD and KLM), so more remote control is being implemented.

The PXD and SVD detector systems are performing well and have contributed negligibly to the overall downtime of the data taking. The backgrounds in the PXD and

SVD from the various beam sources are steadily improving and have in some cases been reduced by more than a factor of two; for example the background from synchrotron radiation from the high energy ring for the PXD. The Touschek background from the high energy ring has seen a slight increase. Forecasts to design luminosity indicate that the background in the PXD can be kept below the 1% level and in the SVD below 3% occupancy.

During the 2021 run the rate at which the PXD triggered the "over-voltage protection" safety feature, when the power supply unit is shutdown to protect the ASICs, increased suddenly to an average of 0.5 trips per day. An auto-recovery system has been implemented. The source is believed to be single-event upsets in an ASIC. Studies are ongoing to identify its source. An increase in the high voltage current is also observed in the PXD when the beam is off, which is being studied. The PXD was intended to run in "gated mode", to protect it from noisy injection bunches, however, this has not yet been implemented. The gated mode itself, however, causes pedestal fluctuations and the merits of implementing this mode are being studied. Currently, there is no need to run in gated mode.

The CDC operation has been stable, after the gain drop that had been observed in 2021 was corrected by adjusting the level of water added to the detector gas, which is now about 1.3 per mille. This will continue to require careful monitoring. Ageing studies are important to ensure the long-term health of the detector, given the expected increase in luminosity and background. These studies are starting in a test setup, with results expected by the late fall.

TOP operation was essentially stable during runs 2021a and b. The major upgrade was the introduction of the PCIe40 readout system during run 2021c. There were several significant teething issues, but TOP was running fairly smoothly by the end of the run. Masking remains unreliable and some external tools have yet to be migrated. The TOP group continues to work on other recurring operational concerns, especially aiming to improve the speed of the injection veto masking. The plan is to automate all routine recovery procedures including run restarts, which should reduce the load on system experts. In the short term before the start of run 2022a, there are plans to update and test TOP Power Cycling and adapt the GUI and continue to sort out issues with the DAQ.

Since November 2021, further analysis of the TOP MCP-PMTs has provided clearer evidence of deterioration in quantum efficiency (QE), although with large fluctuations both for tubes of different types (conventional, ALD, and lifetime-extended ALD), and between different tubes of the same type. The charge integrated at the PMT photocathodes is now of order 0.12 C/cm^2 for the conventional tubes, and typically about twice that for the ALD and lifetime-extended ALD tubes. A similar additional accumulated charge is expected before the start of LS1. There is significant scatter in the size of the observed QE drop with about 14% of the tubes (of both the conventional and ALD types) showing a QE drop of more than 20%. There appears to be a production date (batch) dependence in the tubes showing the biggest QE drop, but there was no known difference in the production process at Hamamatsu. There are also shifts from the measurement methodology which lead to systematic shifts in the measured QE of 2-

3%. There are plans to remove some PMTs during LS1 to measure QE with the bench system to provide further cross-checks. No further studies were reported regarding the possible influence of environmental factors (especially temperature) on ageing rates, but such studies are being developed. The TOP L1 trigger is being studied actively, but not yet ready to be implemented. The LS1 access and progress on PID performance will be discussed in other sections of this report.

The ARICH ran reliably with stable ring parameters during 2021c and is ready for 2022 data taking. There was no downtime greater than 30 minutes, and no new masked channels during the run. About 94% of the channels are operational and there were no additional tripped or noisy APDs during the run. Only two bias channels failed during 2021, and none at all in 2021c. A total of 101 APDs are disabled at this time, the same as at the start of run 2021c. There were small changes to the HV definition leading to faster ramping times and perhaps to enhanced stability.

The recovery script worked well. Most errors could be fixed by the local shift crew leading to significantly reduced ARICH down-times. The DAQ has moved from the copper to the PCIe40 system for 2022a. The changeover was well prepared and tested and went rather smoothly. At the time of the meeting it was running well during cosmic and DAQ test runs. Substantial software developments are continuing to improve ARICH PID performance, centring on alignment of the tiles, gaps between them, and calibration at this time. There have been minor incremental improvements seen in the pion fake rates. Work is continuing.

Apart from the problem of PCIe40, the largest contribution to the DAQ downtime in period 2021c came from the KLM sub-system. The situation seems to be worse than in period 2021ab. The main cause appears to be the loss of the link between the front-end and back-end electronics. A shifter-friendly tool for smooth recovery is being developed. To mitigate the frequency of link losses, LAN cables are being replaced with optical fibres. The effect should be evaluated carefully.

3.1.2 Concerns

- Since some of the ALD coated tubes appear to be suffering more rapid degradation than expected, the original plan of just replacing the conventional PMTs with lifetime-extended ALD coated PMTs may need to be reconsidered.

- The TOP PID performance is affected by background PMT hit rates, with threshold-like effects beginning at about 2 MHz/tube. This particularly effects the low momentum region. Present software studies indicate that a rather tight correlation is to be expected between kaon efficiency and bunch finder efficiency, with a loss of several percent at the projected nominal background rate of 11 MHz/tube at full luminosity.

3.2 Recommendations

- The effect of increasing rates on the performance of all sub-detectors should be carefully monitored. The detector performance should be monitored online so failures and degradations due to high backgrounds can be detected promptly.

- The cause of the efficiency loss in the TOP PMTs must continue to be studied and understood with high priority: environmental factors (such as tube temperature or neutron radiation) for the in-situ tubes may play a role in the QE deterioration and the ability to extrapolate the bench measurements to the situation in the detector.
- The MCP-PMT QE changes must be carefully monitored during the next run to further improve the understanding of systematic uncertainties. Reliable extrapolations of the long-term impact of the QE loss will be essential for the scheduling of the PMT replacement, as well as considering possible changes to the operating conditions for the PMTs in the ongoing runs to minimise the impact.

3.3 Trigger and DAQ

3.3.1 Status

The data-taking efficiency (luminosity recorded/luminosity delivered) during the 2021c run reached 86.8%. The DAQ downtime was about 4%. About half of it was related to the switch-over from the old COPPER back-end to the newer PCIe40 based one for the TOP and KLM detectors. Most of the PCIe40 issues were initially unexpected, but are now believed to be fixed and should not impact the system anymore. Together with increased automation, this should make it possible to achieve or surpass the target data-taking efficiency of 90% for the 2022a run. The DAQ upgrade is prepared for the ARICH for the 2022a run and will be implemented for the remaining four sub-detectors during LS1 starting in summer 2022.

The Level-1 trigger and HLT (High Level Trigger) selection were performing reliably in run2021c and operating well below the present DAQ limit of about 15 kHz. No significant change in configuration is anticipated before LS1.

3.3.2 Concerns

- Some of the issues uncovered by the PCIe40 upgrade seem to be related to insufficiently tested software and to be depending on the sub-detector.
- The rollback procedure from the PCIe40 to the COPPER readout is currently quite slow.
- Reliability issues in the *Bell2Link* and TTD (Trigger and Timing) distribution chain have been present since the start of Belle II and not yet completely solved.
- In order to recover from data-corruption on the links, it is necessary to actually *reprogram* in some cases instead of simply reset PCIe40 FPGA.
- The PCIe40 related issues, which are hopefully mostly infancy-problems, are not the only contributors to data-taking inefficiency.

- The upgrade to PCIe40 in ECL, TRG, SVD and CDC will only happen during LS1 and hence will not be able to be validated with beam for quite some time.

3.3.3 Recommendations

- Full-chain testing of the detector-setup with PCIe40 based read-out should be made as realistic as possible, e.g. by running at the highest possible rates. It should also include robustness tests, such as a reboot of a card and the interruption of a link or similar hardware failures. Slow control applications should be a part of the chain.
- The roll-back procedure, even though it will hopefully not needed anymore, should be well planned and documented for each sub-detector, to have a "playbook" in an emergency.
- The initiative to replace the copper *b2tt* link with an end-to-end optical fibre solution is a very reasonable approach and is fully supported by the committee.
- The central firmware experts for the PCIe40 must be kept for a while to make sure that all detectors have a firmware which is optimally made for them. In particular the firmware engineers must be available during the recommissioning period after LS1.
- In addition to the PCIe40 induced issues, other DAQ problems should not be forgotten. The automation of actions should be even more promoted and the monitoring regularly reviewed and improved to optimally serve the needs of the shift-crew.
- There seems to be room for run-time performance improvement in several HLT channels. This is a topic worthy of further discussion in a future review.

3.4 Software

3.4.1 Status

The Belle II software group is making significant efforts to maintain stable and reliable software while at the same time enhancing its already feature-rich base.

The committee is pleased to note the coverage of the software validation program is being extended. The well established procedure followed for software developed within the *basf2* framework is being complemented by validation of the external software used. Regular updates of ROOT are monitored through CI/CD (Continuous integration/Continuous delivery) regression and performance tests of the Belle II core and analysis software. The move to a new version of Geant4 is carried out about one year before its use in Monte Carlo productions: an extensive validation of basic and complex physics quantities using large simulated samples is performed and discussed with relevant experts. Following the issue observed in summer 2021 during the validation of *release-06* a new set of validation tools for basic generator quantities has been put in place to identify as early as possible the effect of adopting new versions of EvtGen and Pythia. The change in version will be made well in advance of the final release for production, following the same procedures as for Geant4.

The next major release (*release-07*) of the Belle II software will provide several improvements in tracking and particle identification. Due to the shortening of the 2022 data taking, a five month delay, with respect to the initial date of July 1$^{\text{st}}$, has been considered for freezing the implementation of the new features for the release. The new date would allow for additional features to be included as well as more time for preliminary validations by the software developers. On the other hand, the pre-LS1 data set processed with *release-07* would be available for physics analysis only after summer 2023. A final decision between keeping the original deadline of July 1$^{\text{st}}$ or delaying it to December 1$^{\text{st}}$ was expected to be taken by mid-March. Belle II has a rich list of long-term potential developments to improve the precision of the physics measurements. They range from refinement of the current reconstruction to exploring the use of multi-variate analysis techniques at reconstruction level, to investigating how to reduce the software execution time. In order to carry them out an appropriate level of human resources drawn from software developers and physics performance analysts is crucial.

The software group has set up a dedicated branch in the software repository to carry out developments needed for Belle II upgrade studies. A synchronisation strategy with the main development branch has been established and dedicate releases are foreseen to produce Monte Carlo samples and carry out relevant performance studies.

Due to the increase in the license fee for the Atlassian collaborative tools used by Belle II, the DESY IT group has started to provide a free GitLab instance to replace them. The migration to the new tools is scheduled in LS1. Preliminary work to move all repositories, to identify replacements, e.g. for JIRA, to implement requested features for specific functionality, e.g. cvmfs access for CI/CD pipelines, is in progress. No major showstoppers have been identified by any of the Belle II groups using the tools.

3.4.2 Concerns

- Validation is improving with emphasis given to preliminary validation and reporting by developers. Nevertheless how changes on one aspect of the software could impact others, e.g. tracking and PID performance, should also not be underestimated in the integration phase.

- The diversity of the software needs of the collaboration are increasing, while human resources to support all aspects of the software, i.e. operation, developments, upgrade, is limited.

- The software developers and those making physics performance studies have different set of programming skills. This can hinder communication and make mobility of people difficult.

3.4.3 Recommendations

- Ways to ease communication and facilitate collaboration between software developers and performance analysts should be encouraged in order to ensure efficient and high quality software. Hackathon type events where they work together on a chosen topic could foster sharing of knowledge and bringing the community closer.

3.5 Data processing

3.5.1 Status

Offline data processing has been in general smooth. Many operation processes have been automated with a noticeable and appreciated reduction of latency. In particular, data readiness latency for calibration and prompt processing of *2021c* data was achieved within a 15-30 days goal. However, there is still very little contingency and any problem may cause long delays. There are some calibrations still not automated (SVD timing, CDC dE/dx and beam energy) that are done manually. Due to the limited number of detector experts, this results in delays in the production of run dependent simulation samples and processing the RAW data.

Two serious problems in MC production affecting both run independent and run dependent data were detected several months after the production was finished. In both cases, wrong wide resonance mass and beam energy spread were used and it was due to payload mis-configuration. The production needed to be re-done for the correct energy spread, while only a partial new production was done for the resonance mass issue.

Skimming is currently suffering from some scalability limitations of the production system. There are in general too many jobs running at once that need to be babysat. The production system was designed with the assumption of eight-hour runs being dominant. This is not the case now since any problem that requires a stop/start of the DAQ generates a new run. Most of the data processing jobs are made for a single run, thus the proliferation of jobs. Merging runs may help in reducing the number of jobs but the consequences needs to be studied in detail.

Ideas for improving the efficiency and scalability of the data processing operation for post-LS1 were presented. The committee is pleased to see that many of the proposed changes will certainly go in the right direction: Reduce the re-calibration needs to the minimum, increase the number of re-calibration centres, use a common interface, divide the human workload between long-term commitment managers and short-term service people, and provide them with more powerful automation and monitoring tools. The collaboration should put enough resources now in the realisation of these ideas in order to be ready by end of LS1.

3.5.2 Concerns

- Validation of MC productions is a concern. In addition to the problems found during *release-06* validation, serious problems were detected several months after the production, invalidating the whole production for use in physics analysis. This affected both run-independent and run-dependent samples.

- The scalability issues of the production system for short runs needs to addressed.
- Data processing activities continue to suffer from human resources issues. Operation and development are on the shoulders of the same people, which is good in general in order to develop effective tools, but if the operations load is too large it hinders any progress.

3.5.3 Recommendations

- Thorough MC validation process, from the software itself to the production, need to be established, eventually asking the collaboration of physics groups to provide feedback on the quality of new MC data.
- The use of run-dependent MC in physics analysis is not yet well spread and needs to be strengthened. This will encourage people to provide timely feedback and to engage in a virtuous circle for its use.
- The committee recommends to study and eventually implement merging runs that share similar running conditions to mitigate the scalability issues of the production system. This may imply changes in the data processing jobs because the single run assumption might be deeply embedded in the code.
- Dividing the human workload between long-term commitment managers and short-term service people (i.e. data production shifts) should be implemented as soon as possible to liberate human resources for the development effort for all the needed improvements.

3.6 Detector performance

3.6.1 Status

The Belle II physics analyses rely on measurements of the produced particles, charged and neutral, mostly using several sub-detectors to cover the full range of momenta, energies, and angles and different types of particles, both hadrons and leptons. There has been some progress in those measurements providing fundamental input to various analyses under relatively stable conditions. In many cases the current precision of the results has led to efficiencies and systematic uncertainties, comparable to Belle data. However, this level of performance does not meet the requirements of many future flavor physics studies at high statistics.

The fact that run-dependent Monte Carlo samples are not available for all buckets has led to significantly larger systematic uncertainties. For instance, the uncertainty for the $B\bar{B}$ normalisation of 1.1% for run dependent simulations compared to 2.6% for run independent simulations.

Charged Particle Tracking
For last year's data the charged particle detection efficiencies for the PXD, SVD, and CDC sub-detectors was close to 99%. For the higher momenta the data/MC ratio for the track reconstruction was consistent with 1.0, but low by about 10% for the lowest momenta. A separate study indicated a charge asymmetry of up to 10% for low momentum tracks in the transverse plane for both data and simulations. This is a significant contribution to the overall systematic uncertainty of 0.7%, which is a factor two larger than for Belle data.

The data/MC efficiency ratios for the detection of K_S decays decrease dramatically with increasing distances of the decay vertex from the IP. This is largely due to a reduced number of SVD signal hits and increases in beam backgrounds. The current estimate of the systematic uncertainty of 3% exceeds the Belle estimates by a factor of two.

There have been several indications that the magnetic field map is not accurate in some areas. This leads to errors in the momentum measurements for the charged particles. So-called global track momentum scale factors have been introduced to arrive at values for various particle masses in agreement with those given in Particle Data Book. These scale factors, averaging 0.99971, depend on the track polar angle and charge.

Neutral Particle Detection
Building on the experience at Belle and also BaBar, the current uncertainties of the photon detection are at the level of 1% in the barrel for $E_\gamma > 250\,\mathrm{MeV}$. The introduction of BDTs to differentiate signal photons from beam backgrounds and fake photons has produced encouraging preliminary results and is being applied for missing mass measurements, for instance in semileptonic decays. A considerable effort has resulted in energy dependent corrections of the raw data to address the photon energy bias. These corrections, derived from $\pi^0 \to \gamma\gamma$ decays, are typically smaller than 1%.

The data/MC ratio of the π^0 detection efficiency is largely independent of momentum, but shows some differences for the two data samples under study, specifically evident for the polar angle dependence. Measurements of absolute efficiencies as a function of the kinematics are underway.

The detection of K_L decays is important for the determination of missing energy and serves primarily as a veto. It has been studied in $e^+e^- \to \gamma\phi(K_L K_S)$ production or $D^{*+} \to (D^0 \to K_L\pi^0)\pi^+$ decays, but this will require much larger data samples.

Charged Particle Identification
Charged particle identification is a very complex task relying on six different sub-detectors (SVD, CDC, TOP, ARICH, EMC, and KLM) with their different responses to particle type and momentum, and different geometric acceptance. Charged hadron ID relies primarily on TOP on the barrel and ARICH in the endcap at higher momenta, with substantial contribution from dE/dx of the CDC and SVD, especially at lower momenta. The ARICH reconstruction has been updated to include individual tile properties, alignment and calibrations. Unfortunately,there have been only modest improvements in the

TOP performance, with kaon efficiencies ranging from 70% to 90% and pion mis-ID from 5% at low energy to 20% above 3 GeV and still showing substantial differences between data and simulations, including track momenta and angles, and a clear charge asymmetry. About 20% of the TOP modules are hit not by one, but two tracks and therefore the association of the emitted Cerenkov photons with the individual tracks may be incorrect. Recent studies indicated a potentially tight correlation between the kaon and the bunch finder efficiencies. This is just one of many features that require further study.

Off-line Analyses

Methods are being developed to combine the information of the sub-detectors using combinations of multi-variable analyses, specifically likelihoods, boosted decision trees (BDT), and convolutional neural networks (CNN). The introduction of BDTs has reduced the electron fake rate by more than a factor two, while on average the systematic uncertainty of the detection efficiency is of order 1-2%. For lepton ID (LID), results based solely on J/ψ dilepton decays, the targeted systematic uncertainty of 0.5% per track has been reached. However, the tracks in this event sample do not extend over the full phase space in momentum and angles, or higher multiplicities. The recent inclusion of SVD information is expected to significantly improve the separation of leptons and kaons for momenta below 1 GeV. Surprisingly, the inclusion of the TOP information resulted in a lower LID efficiency and higher mis-ID. Also, sizeable run dependencies on data/MC ratios have been observed.

Combinations of neural networks and BDT appear to be more promising, though they also will need further optimisation. The sensitivity of machine learning might be very powerful, especially for TOP data, but it would require a variety of training samples. The formation of the physics performance group overseeing the preparations and standardisation of performance studies with the goal to reduce systematic uncertainties are of great importance!

3.6.2 Concerns

- While the stated higher level analyses are very important, the best performance can only be reached with very detailed understanding of sub-detector data under various beam conditions. The amount of data now available should allow detector studies with reasonable statistics while making multi-dimensional cuts in appropriate detector variables, including particle angles, positions, charge, and momenta. The TOP performance also needs further study as a function of background rates and the polar angle of photons in the bar, which essentially requires cuts on the relative photon timings for each track, and will allow PID performance to be studied with different levels of signal to background photons.

- The sub-detector groups will be challenged by the overlapping demands of the operation and calibrations, ongoing preparations for the upcoming long shutdown, and the need to increase analysis efforts to better understand performance at the detector level for a large variety of physics processes under study.

3.6.3 Recommendations

- The detector performance should be scrutinised at all levels to obtain a detailed understanding of the raw data and their calibrations as well as the complex multivariate analyses to optimise ways to combine information from different sub-detectors and thereby enhance the overall performance.
- Off-line performance studies, for both a single or multiple sub-detectors, should be based on a variety of particles produced in different interactions, to record efficiencies and other features for different momenta, angular distributions, event multiplicities, and the impact of background.
- Close collaboration between the sub-detector groups with the physics performance group will be critical, in particular for the TOP and the CDC.
- Given the challenge to address and solve the many detector performance issues, the sub-detector groups are encouraged to engage more scientists to examine the variety of unexplained features and the impact of the beam conditions by carefully studying the performance in the relevant sub-detector phase space and comparing the data with simulations.

3.7 Belle and Belle II physics

3.7.1 Status

The physics analyses based on combined Belle and Belle II data sets have resulted in an impressive number of publications. Belle completed its data taking in 2010, collecting 1.02 ab^{-1} of data, the world's largest data sets at the $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(4S)$, and $\Upsilon(5S)$ resonances, which the collaboration continues to analyse. Five new institutions were admitted to the Belle collaboration[1] under the new fast-track policy.

The Belle collaboration retains excellent visibility by publishing papers, giving conference presentations, and playing an active role in the Heavy Flavor Averaging Group (HFLAV). In 2021-22 newly published analyses include studies of ϕ_3 in $B^+ \to D^0 h^+$ with $D^0 \to K_S h^+ h^-$, Ω_c^0 baryon decays, and measurements of $B^{(*)}\bar{B}^{(*)}$ cross-sections above the $\Upsilon(4S)$ resonance.

The overall progress and plans for physics analyses are excellent. Since the last BPAC review, the Belle collaboration has published 18 papers and submitted six more papers for publication. The research program is broad, including studies of B, D, τ decays, and dark sectors. Analyses of branching fractions and CP-violating asymmetries in D-decays, searches for lepton-flavor violations in the decay of the tau lepton ($\tau \to \ell\gamma$), and B-mesons ($B \to \ell^\pm\tau^\mp$), as well as analyses of inclusive semileptonic B-decays, have been published.

Several new analyses were completed and submitted to the arXiv. This includes bounds on the Electric Dipole Moment (EDM) of the τ-lepton, a unique measurement at

[1] It is desirable to have Belle and Belle II collaboration membership lists (both in terms of the members and the institutions) available on their websites.

an e^+e^- collider. Due to the extremely small SM prediction, this measurement provides a practically background-free probe of New Physics. Also, a search for the leptophilic decay of a new light Z'-boson to $\mu^+\mu^-$ was performed. An important study of B-decays to Λ involving missing energy was performed, motivated by a recently discussed low-scale baryogenesis scenario.

It is now clear that the total data sample to be collected by Belle II by the start of LS1 will be limited to ~ 500 fb^{-1}. Yet, the collaboration has recognised and proven the value of combining Belle and Belle II data sets for sensitive analyses by applying improved event reconstruction and analysis methods.

Belle II performed a number of exciting analyses that are either submitted for publication or already published. The improved management of physics analyses are to be commended, for instance, the reorganisation and streamlining of the analysis approval and the merging of groups working on similar studies to reach the critical mass of researchers. As a result, several exciting analyses were completed in time for the Moriond Meetings. These analyses include studies of the CKM matrix elements V_{cb} and V_{ub} based on semileptonic B-decays, tau-decays, branching ratios, and CP-violating asymmetries in hadronic B-decays, including decay time dependent CP-violating asymmetries. Important studies of rare decays including neutrinos, like $B \to K^{(*)}\nu\bar{\nu}$, are complementary to other studies of electroweak anomalies in $B \to K^{(*)}\ell^+\ell^-$ and also could be used to place constraints on the properties of light dark-sector particles. Other analyses of dark sector particles have also been performed. Particularly, the collaboration completed an update on the search for a new light Z' decaying invisibly and the first search for dark Higgsstrahlung.

Despite record-setting peak luminosities, several interesting physics analyses, including studies related to anomalies, e.g., violation of lepton-flavor universality in rare B-decays, are not yet ready to compete with published analyses. The collaboration understands the challenges associated with some of the analyses and makes the good use of the available resources. They managed to produce several results in areas where the Belle II detector enables unique contributions with limited statistics, such as studies of charmed baryon lifetimes.

Data recently recorded above the $\Upsilon(4S)$ resonance at 10.75 GeV, are expected to reveal the nature of the observed peak structure in $e^+e^- \to \pi\pi\Upsilon(nS)$. The analysis should be given high priority. The use of machine learning methods, similar to those employed by the LHCb for pentaquark searches, should be encouraged.

Further information on the physics motivation was presented for the upgrade plan for the Long Shutdown 2 (LS2) from 2026 to 2027. Further discussion of the future physics program would also be appreciated. Since the output of the Snowmass process will be highly influential worldwide, the participation of the Belle II collaboration is essential and highly recommended.

3.7.2 Concerns

- The future physics results will be seriously compromised without a substantially better understanding of the detector performance.

- While the combined Belle and Belle II data should be exploited for now, the future goal is the exploitation of the much larger Belle II data sample. .

3.7.3 Recommendations

- The committee encourages the continuation of the interactions of the Belle and Belle II scientists to perform combined analyses.
- The committee notes the superior performance of Belle II in vertex reconstruction and excellent control of systematic uncertainties in measurements of the lifetimes of D-mesons and Λ_c baryons. The collaboration should exploit the possibility of precision measurements of the lifetimes of other charmed baryons.
- The analysis of the available data above Υ resonance should be given priority to reveal the nature of the unexpected observed peak. New analysis methods, possibly including machine learning methods, should be developed.
- Analyses of $e^+e^-(\gamma) \to \pi\pi$ events at low q^2 (with the radiative return) should be pursued to understand inconsistencies in current experimental data and to assess the contribution of hadronic vacuum polarization to muon $g-2$.
- Additional new searches for light dark-sector particles should be pursued in a relatively short time scale (e.g., the search for an invisible dark photon). Belle II will have a unique opportunity to set the most stringent constraint (or discover) on the parameter space of many of these dark sector models.
- The committee acknowledges efforts by the collaboration to cope with the LS1 shutdown, which now exceeds 15 months. For instance, the active involvement of the physics analysis teams in trigger and performance studies would be highly appreciated. The committee encourages plans to extend similar engagements to other areas, such as the detector alignment and calibration and further developments of the reconstruction and analysis software tools.
- Since the output of the Snowmass process will be highly influential worldwide, the participation of the Belle II collaboration in the process is highly valued and should be continued.

3.8 Long Shutdown 1 Preparation

3.8.1 Status

Beam pipe:
To mitigate the effects of synchrotron radiation on the PXD, a new beam pipe is being produced for the PXD2. Unfortunately, delamination of the gold plating of the new beam pipe could not be repaired and a new inner section had to be procured. To mitigate further risk, the team adopted a two-pronged approach to identify the root cause of the delamination: study of the material condition of the beam pipe surface around the

delaminated area and a review of the sputtering process and equipment itself. The team is to be commended for the very thorough studies that have been carried out. X-ray Photoelectron Spectroscopy has revealed that the delamination was caused by the presence of an unwanted titanium oxide layer on the solid titanium surface. Good vacuum control will be essential for a successful sputtering process and the setup needs to be updated. The gold electrode for the sputtering will also be replaced. There are also indications that oxygen, carbon and fluorine contamination is caused by the use of Teflon in the old setup. Almost all materials used in the sputtering chamber will be replaced with new stainless steel or ceramic parts. The schedule calls for sputtering tests between now and April, with the gold spattering on the new beam pipe taking place in May 2022. After further integration with the heavy metal part the new beam pipe should be ready by October 2022. The availability of the new beam pipe drives the overall schedule.

A spare cylinder set exists in case the sputtering process fails and it has been verified that this pipe set is usable.

PXD:

The PXD2 has made several improvements to the assembly and transportation procedures based on the recent experience. The adhesive foils are now cut by hand to avoid excessive burrs; the CO_2 pipes are pre-bent to avoid stress on the Support Cooling Block; and more rigid cases are being used for the transport. At the time of the review layer 1 of the half-shell was fully populated and two ladders had been mounted on the layer 2 half-shell.

The full complement of ladders needed for layer 1 and layer 2 is 8 and 12, respectively. For layer 1, two more ladders are needed. Given the current ladder production, this is achievable with two spare ladders. For layer 2, one more ladder is needed. The production indicated that it was possible to meet the production goal plus three spare ladders. During the review, however, it was learned that two layer 2 ladders were broken. This puts the project in a precarious position with possibly only one spare layer 2 ladder.

The rework of problematic modules is continuing at HLL, albeit at a slower pace than expected due to a variety of reasons. The restrictions due to the pandemic limit progress and assembly of new modules is delayed due to the lack of good switcher chips. The testing of the switcher chips is halted due to a defective needle card and limited availability of personnel and the new switcher chips are compromised due to strongly misaligned bumps.

The current schedule calls for the PXD2 to be completed by July 2022, which gives the project about three months of schedule contingency. This schedule includes ten weeks of testing at DESY. The committee supports the LS1 strategy to first ensure safe arrival of the PXD2 at KEK before removal and disassembly of the existing VXD. The schedule and procedures for extraction and re-installation of the VXD with the new PXD2 are being developed and documented, and are being reviewed, however, without several key experts present. Several scenarios have been developed including scenarios where repairs will have to be made to the SVD. This level of contingency planning and risk mitigation is very valuable and the committee was very impressed with the proactive approach.

The cancellation of the 2022c run implies that there is no need to transport the MARCO cooling system to KEK and that the IBBelle system and present phase 3 infrastructure at KEK can be used for pre-commissioning of the half-shells.

VXD:
The team has to be congratulated with their proactive approach to address single point failures. Experts for the SVD and the CO_2 piping have been identified which should significantly mitigate the risks. Vulnerabilities in the areas of slow control and VXD extraction remain. Good progress has been made on studying interference issues at the VXD endcap and the shielding of the bellows. Agreement has been reached on the required clearances for the shields, bellows and RVC.

TOP:
The basic plan is to replace all 224 conventional TOP MCP-PMTs during LS1 if the measured QE performance deterioration requires. The decision must be made by Jan. 2023 or earlier so that the work can be completed before the beginning of the 2023 run. Enough required tubes, 257, are fully tested and available. All other required parts (like cookies) are being produced now with PMT module assembly scheduled starting in March 2022. Installation work is being practised with the spare module. The required spare boardstacks have been tested multiple times in Hawaii and are staged and ready at KEK.

ARICH:
For LS1 2022, the ARICH will not be detached from the ECL, so there are no plans to do any work inside the detector. One failed LV cable will be repaired so the related five HAPDs should recover.

3.8.2 Concerns

- The schedule for the new beam pipe is very tight and the availability of the new beam pipe is a critical component to the success of LS1.
- It appears that the collaboration has a very limited number of switcher chips to complete the assembly of the required ladders. The misaligned bumps are worrisome.
- The continued misfortune regarding the availability of ladders is of great concern. There now is a non-zero probability that the PXD2 may not be fully populated.
- The loss of expertise for the pixel detector is a key concern. A well-trained installation team is needed to carry out the delicate work.
- Although there has been excellent effort to document the extraction and installation procedure of the VXD, to mitigate the impact of disappearing expertise and

the possibility of continuing travel restrictions, there is no substitute for hands-on presence.

- Timely decision will be needed whether to replace the TOP PMTs to be well prepared for the delicate work.
- The TOP PMT and electronics replacement will be challenging and require experienced crews.

3.8.3 Recommendations

- It is recommended to start making contingency plans for the unfortunate situation that the PXD2 detector may not be complete and optimise it for physics performance.
- Practice the installation of the VXD often to minimise the risk of accidents, mimicking the real detector as closely as possible including a realistic modelling of the cable plant. It is strongly recommended to practice the beam pipe assembly and the VXD support assembly with the key experts.
- Effort should be given to early and advanced planning to allow for the possibility of former experts to participate in the extraction and installation process.
- Continue practising the installation of the TOP MCP-PMTs to ensure that there are enough people adequately trained to ensure timely installation.

3.9 Future upgrade

3.9.1 Status

The Upgrade Working Group (UWG), which was formed in October 2018, gave an update on their activities. Motivations for the upgrade studies are to improve the detector robustness against background, to increase the long-term radiation resistance, to develop the technologies to cope with possible future SuperKEKB and Belle II upgrade paths, and to improve further the physics performance. A set of Expressions of Interest (EoIs) were collected at the beginning of 2021 and have been analysed by the Upgrade Advisory Committee (UAC) formed at the end of 2020. The upgrades have a wide time span with options that can be implemented independently; the 2026 SuperKEKB upgrades planned as part of the KEK road map (CDC electronics and radiation monitors), those targeting the LS2 in 2026 (replacement of the VXD and upgrades of TOP and KLM) and some long-term developments (ECL upgrade and STOPGAP to cover the gaps in TOP coverage). Other systems were invited to verify their capabilities at high luminosities. The BPAC was pleased to hear also the plans for the upgrade of the trigger system, which have been presented at this annual review.

As for the VXD systems, there is the option to replace the current detectors with improved DEPFET and fine pitch double-sided strip detectors, keeping the Belle II geometry, or to implement an all-pixel solution, either with SoI or CMOS technology.

The TOP upgrade concerns both the photosensors, moving to life-extended ALD-MCPs or to SiPMs for the whole detector, and an upgrade of the front-end electronics. For KLM, options include the replacement of RPC with scintillators and an upgrade of the readout electronics.

A proper evaluation of different options depends on the future SuperKEKB performance evolution, in terms of delivered luminosity and extrapolation of machine backgrounds, and the upgrade plans, especially if a change in the interaction region will be required. Two white papers, one on the detector upgrades with physics justification and one on the feasibility for running with a polarised electron beam, will be submitted shortly to the Snowmass process. A Conceptual Design Report for the upgrade is foreseen by the end of the year.

The committee appreciates the effort by the UWG and the UAC to develop the physics case for a detector upgrade and to steer the development of an upgrade path. For the upgrade with the time frame of 2026, when there could be a possible replacement of the QCS, the Belle II Collaboration will need to make the transition from the R&D phase to a construction project in a short time. Some R&D will proceed in view of possible extension of the SuperKEKB and Belle II projects after 2032.

3.9.2 Concerns

- Physics and performance studies are still not at the level needed to address and compare quantitatively different upgrade options.
- The task of UWG is quite broad covering a large time span with different scenarios. For the upgrade considered during LS2, not much time is left for the R&D and making technical choices among different options.
- There are large uncertainties in the exact scope and timeline of the accelerator upgrade. This will have a significant impact on the schedule of the detector upgrade and choice of technologies.

3.9.3 Recommendations

- A decision-taking process for the different upgrade options should be included in the CDR preparation.
- Evaluation of the physics case for the upgrade should continue and be sharpened.
- Expression of Interest for the trigger upgrades is welcome for further discussion by the BPAC.