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

1–2 and 8-9 March 2021 Remote Meeting

A. Andreazza* (Milano), P. Collins* (CERN), G. Corti (CERN),
M. Demarteau (ORNL), R. Forty (CERN), B. Gavela (Madrid),
G. Carlino* (Napoli), S. Gori (UCSC), W. Hulsbergen* (NIKHEF),
M. Ishino (Tokyo), V. Lüth (SLAC), P. Mcbride* (FNAL)
P. Mato* (CERN), F. Meijers* (CERN), N. Neufeld (CERN),
K. Oide* (KEK), B. Ratcliff* (SLAC), A. Petrov* (Wayne State University),
M. Sullivan (SLAC), H. Tajima (Nagoya), M. Titov* (Saclay)
and chaired by T. Nakada (EPFL)
* Expert members

5 July 2021

1 Short summary

Due to the COVID restrictions, this year's annual review meeting of the B-factory Programme Advisory Committee (BPAC) took place remotely on the 1st, 2nd, 8th and 9th of March 2021. Since the meeting duration of the each day was limited owing to the time differences among the locations of the participants, slides and video recordings for the presentations were made available prior to the meeting so that the reviewers could pose their questions remotely in advance. The committee thanks the accelerator group and Belle II collaboration for their effort to provide slides in a timely way covering the status of the SuperKEKB machine and the Belle II experiment, as well as the physics analysis with the Belle data. This section provides the most important findings and recommendations of the committee and a full report is provided in the sections to follow.

The SuperKEKB machine has been operating at higher luminosities with lower beam currents than those of the KEKB machine and the committee appreciates the continuous effort and improvements by the machine group for stable running of the machine. It appears that there are persistent obstacles, such as the increase of emittance for the high energy beam in the injection line, which prevent the machine from increasing luminosities without raising the background above the acceptable level for the experiment. The machine group is invited to make a plan for analysing the issues currently limiting the machine to increase luminosities keeping the machine background under control and devising actions together with the resource requirement. Increased efforts in simulation would be a part of a such plan. In this context, the proposed "background fellow", who will be financed by the Belle II common fund for maintenance and operation to work in the background study group, is an excellent idea. The committee fully supports this. For gaining more human resources in the simulation effort, collaboration with other accelerator laboratories worldwide should be strengthened. Interest in SuperKEKB operating with low emittance beams is increasing and there could be also accelerator experts elsewhere wishing to participate in the machine operation. The committee understands that the current COVID restrictions make such collaborative work difficult, but it should be explored with the long term gain in mind. Another important issue on a relatively short time scale is the consolidation of the ageing machine components, in particular for the injector linac. The committee urges the machine group and the KEK management to address deferred maintenance in a timely manner by making an inventory of critical spare parts and making a replacement plan.

For the longer term, the committee thinks that the characteristics of the machine must be fully understood and exploited first to increase luminosities, before attempting a major hardware upgrade such as the new superconducting focusing quadrupole magnets (QCS) in the interaction region. Understanding of the hardware limitation of the machine would allow a prioritised and optimal upgrade path to be made.

The committee is pleased to see that the analysis with Belle data is still very active and fruitful, resulting in many publications, and this effort should be supported until a comparable statistics will become available for the Belle II data. The analysis of Belle II data is advancing well. Although the priority of data taking until the long shutdown in 2022 (LS1) must be given to collecting high statistics $\Upsilon(4S)$ data, the committee supports a short run at higher energies since this will provide a physics opportunity unique to the Belle II experiment. Closer interaction between the Belle and the Belle II collaboration would be very beneficial. The experience in assessing systematic uncertainties gained by the Belle analysis will be very useful for the future Belle II analysis, although the detectors and beam conditions of the two experiments are different.

The committee did not find major concerns with the Belle II operation for data taking and processing. With continuous effort in improving the detector performance monitoring and automation of the operation, for both online and offline tasks, the efficiency for running the experiment during this difficult period with COVID has been impressive. The idea of opening "operation fellow" positions with financial support to Master and PhD students located at KEK for their shift work is very interesting and a good idea during the COVID epidemic when the number of people available at KEK is limited. The outcome of this programme must be carefully analysed before deciding whether it should be continued after the COVID restrictions will be lifted.

The plan to operate the experiment with an increased beam background rate, i.e. up to 3 MHz per photomultiplier tube (PMT) of the barrel particle identification system (TOP), until the start of LS1 was presented. This would ease the machine operation at high luminosities for delivering $\sim$0.7 ab^{-1} of data before LS1. There has been no indication of ageing in the PMT performance so far. Given projected PMT rates, such degradation is not expected to be observable in the data until about the summer of 2021. During LS1, all of the conventional PMTs, and eventually the normal ALD type PMTs if required, will be replaced with PMTs that have extended lifetime with ALD treatment. The committee, therefore, thinks that running with this level of background to obtain high luminosities is, in principle, a sensible plan. In practice, the running condition should be continuously optimised to maximise the "useful" integrated luminosity by carefully monitoring the performance of the detector in particular the effect of ageing. The high level trigger (HLT) becomes important for running with increased background rates, in order to keep the collected data within the pledged computing capacity. There is some room for further improvement in the HLT performance. Work for the 2021 summer shutdown should be carefully planned for the repairs necessary to keep the detector operating at its best until the LS1.

Although the overall status of the Belle II experiment is good, the collaboration should further enhance efforts to understand and improve the detector performance in a systematic way. In particular, TOP needs increased efforts by detector and analysis experts working more closely together. Observed discrepancies between the data and simulation must be resolved for all sub-detectors. Further studies should be made to see if the detector performance can be boosted with improved use of the detector information, e.g. by studying the performance in full two dimension, momentum and polar angle, rather than in their projections separately, or analysing the impact of photons with long path lengths and/or steep polar angles in the quartz.

The committee noted the revised schedule for the LS1, which is now foreseen to start in July 2022 instead of January 2022 in the original plan. The duration of the shutdown is also extended to ten months. Those changes will give welcome margin for the experiment running under the COVID conditions, including the preparation and installation of several detector components. Care should be taken so that the budget for the 2023 running would not be reduced. The committee is pleased with the general progress made for the LS1 preparation. Recently developed ideas to gain space needed for the cables from the vertex detector system, which includes modifying the front cap of the QCS, are very encouraging. Although further work is needed to examine their feasibility, the anticipated gain in space is such that the increased number of cables from the fully equipped pixel detector can be accommodated.

For the long-term future upgrade, the committee recommends that the collaboration make a more systematic approach to study how the overall performance of the Belle II detector can be improved for physics goals, rather than focusing too fast on some specific detector upgrades. The timeline for possible major upgrades appears to be more relaxed now. Those ideas should be separated from the consolidation effort for ensuring the running of the current Belle II detector, which should be planned carefully.

The resource request for the computing in 2022 was also presented and will be evaluated by the computing resource scrutiny group later. The committee learned that 2.5 PB more tape space has been used in 2020 compared to the pledge since more data were collected than anticipated. Since this shortage should not affect the 2021 data taking, it would be highly appreciated by the committee if the computing centres could provide the additional tape storage.

The committee is very pleased to learn that a total of seven months is granted for the running in the Japanese fiscal year 2021. It is important to collect as much data as possible before the LS1 and the committee is looking forward to following the progress.

2 Physics of Belle and Belle II

2.1 Belle

2.1.1 Status

Over a period of 10 years, the Belle detector recorded the largest data samples worldwide of bottomonium states, 772M $B\bar{B}$ events at the $\Upsilon(4S)$ and smaller unique samples at the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(5S)$ resonances. About a decade after the last data were recorded, the Belle collaboration remains scientifically very active. In fact, three international groups recently joined the collaboration. Last year, the Belle collaboration submitted 24 papers to journals and presented 21 talks at 10 conferences. Their citation record remains very strong and together with Belle II and LHCb, Belle is a key contributor to the Heavy Flavour Averaging Group.

With the experience in the broad physics program, in-depth understanding of common detector components, a tested software system, and largely unique data sets, Belle scientists are playing an important role in the early Belle II program. On the other hand, the Belle II software with its newly developed analysis tools will enhance some of the Belle analyses, especially as some publications are being planned as joint analyses based on the two data samples. Joint meetings on a variety of physics analyses and on future operation at different energies are becoming more frequent. Both Collaborations realise that these joint efforts will have a great impact now, as Belle II is expected to significantly enlarge the total $\Upsilon(4S)$ data sample and is developing new machine learning techniques for complex analyses.

Currently Belle's data storage and computing relies on KEKCC, using about 7% of CPU, 9% of disk space and 3% of tapes. For the future, there is no dedicated budget (or other resource) to keep the Belle data at KEK. Porting Belle software to the new operating system will be a major task. The next replacement of the KEKCC is planned for 2024, at a time when the Belle II data sample is expected to exceed the Belle sample.

Of the many recently published results only a few can be addressed here, with focus on B mesons decays, but including bottomonium, charm and tau lepton studies.

- A new extraction of the CKM matrix element $|V_{ub}|$ from inclusive $B \to X_u \ell^+ \nu_\ell$ decays has resulted in values that are compatible with measurements based on exclusive decays at the level of 1.3σ. This analysis employed machine learning techniques to enhance the tagging of $B\bar{B}$ events and the separation of the signal from the dominant $B \to X_c \ell^+ \nu_\ell$ background. This also resulted in a substantially enlarged kinematic phase space for the signal.

- Standard Model predictions of the rate of B^0 decays into invisible final states with or without a single photon are extremely small (10^{-9} or 10^{-16}, respectively), due to the chiral suppression. The Belle search for such states did not indicate a significant signal and resulted in limits on possible new physics contributions to these decays or to the presence of light dark sector particles.

- Another search for B^0 decays involving dark sector particles was based on samples of 4 charged particles, either two lepton pairs, or a lepton pair and a pion pair. This resulted in a range of upper limits of about 10^{-7} for the branching fraction of the decay mode $B^0 \to A'A'$ with the A' decaying to either muons, electrons or charged pions.

- The studies of rare decays of the type $B^+ \to p\bar{p}\pi^+\pi^0$ and $B^0 \to p\bar{p}\pi^+\pi^-$ are new. They are expected to provide insight to factorisation.

- A number of studies of CP violation in rare B decays involving multiple kaons were presented. The three-body decay $B \to \phi\phi K$ proceeds via a $b \to ss\bar{s}$ loop, but the same final state can also originate from the tree-level process $B \to \eta_c(\to \phi\phi)K$. Due to interference of these processes, no CP violation is expected, and none was observed. Studies of the decay $B \to K_S K_S K_S$ revealed CP asymmetry in mixing at a level of 2.5 standard deviations, consistent with expectations.

- A recently reported measurement of the cross sections for $e^+e^- \to \Upsilon(nS)\pi^+\pi^-$ ($n = 1, 2, 3$) at c.m. energies from 10.52 to 11.02 GeV revealed a new structure, at a mass of 10.75 GeV and with a width of 35.5 MeV. The corresponding global significance (including systematic uncertainties) is of 5.2 standard deviations. The nature of the new structure is not obvious: it could be a resonance, may be a signal for the not yet observed $Y(3D)$ state or an exotic state, e.g., a compact tetra-quark. This is just one of several puzzling results from the relatively small, but unique, higher energy data samples.

- Results of other studies of strong and electromagnetic decays off the Υ states, in addition to exotic states, such as X(3872), and η_{c2} production in B decays, are of great interest.

- The first measurement of the $D^0 \to K^0\omega$ decay, a CP odd final state, is important for studies of the charm mixing parameter y_{CP}.

- A study of radiative decays of orbitally excited charm baryons $\Xi_c(2790)$ and $\Xi_c(2815)$ resulted in measurements of partial widths which are to be compared to theoretical models that probe the inner structure of these heavy baryons.

- Searches for lepton-number and baryon-number violations in B meson and τ lepton decays as well as studies of ordinary and exotic hadrons have resulted in some of the most stringent tests of the Standard Model.

It is expected that Belle data will continue to serve as a most valuable resource for many future analyses, including lepton-flavour violating decays of the B mesons and τ leptons, studies of ordinary and exotic hadrons, and dark sector particles. The development of new analysis techniques applied to the large set of different processes is likely to lead to broader in-depth understanding and hopefully some valuable insight into the unknown.

2.1.2 Concerns

- More intense interactions of the Belle and Belle II scientists are very important now, while the first large Belle II data set at the $\Upsilon(4S)$ is being recorded. These interactions need to become more frequent to result in joint efforts to enhance the detector performance and improve the data analyses, with special attention to systematic uncertainties, given that statistical errors will shrink.

2.1.3 Recommendations

- Plans for the future of the Belle Collaboration should be developed to assure that there is appropriate support, in particular for data storage and preservation, access to computing, and sharing of other resources and expertise.
- Plans for the future support and potential unification of the Belle and Belle II collaborations need to be developed to assure the necessary resources.
- Measurements of the radiative penguin transition $B \to K^*\gamma$ which is sensitive to the CKM matrix element V_{ts} have not been updated recently. The potential for studies of the $B \to \rho\gamma$ transition and the extraction of the CKM ratio V_{td}/V_{ts} should be explored.

2.2 Belle II

2.2.1 Status

Extensive preparations for the very broad Belle II physics program continue to be impressive. In preparation of the data set of 0.5 ab^{-1} to be recorded before LS1, 29 analyses were assigned high priority based on their scientific importance, their novelty and feasibility, and the strength and expertise of the working group. The participation by many analysis groups and close interactions with detector performance groups will be critical, as will be close attention to details of the MC simulations of the detector responses and the dynamics of the particle decays. Preliminary results submitted to conferences have demonstrated the progress in the development of these complex analyses. Many of them confirmed earlier measurements.

Given the current data set of close to 100 fb^{-1}, early publications have focused on topics that had not been studied in the past, or have been performed to exploit new analysis techniques.

- Precision measurements of fundamental particle properties like particle masses and lifetimes represent critical tests of the detector performance and MC simulations, and are likely to exceed the precision of existing measurements.
- Similarly, time-dependent CP violation studies involving many different decay modes of B and D mesons remain to be further explored.
- In many cases, blinding of the results should be used to avoid bias from earlier measurements or theoretical predictions.
- In recent years, the dark sector has drawn a lot of attention. This is an area where Belle II results are likely to have a strong impact; two unique papers based on very small data sets were published in PRL last year, others are in preparation.

Certain measurements could benefit from combining Belle and Belle II data, taking advantage of larger data samples and newly developed analysis techniques.

2.2.1.1 Leptonic and Semileptonic B decays

Measurements of the CKM matrix elements have been at the core of the research program at B factories. With future, much larger data sets and the recent development of powerful multi-variable event selections, Belle II has a great potential to significantly upgrade our current knowledge, test theoretical predictions, and hopefully uncover some unexpected features.

The development of a hadronic tag reconstruction for the study of B decays with an unobserved neutrino is very challenging, but critical for the next generation of measurements of semileptonic and leptonic decays of B mesons produced at the $\Upsilon(4S)$. A detailed description of the calibration algorithms for tag-side reconstruction and preliminary results on the inclusive $B \to X\ell\nu$ decays - including first estimates of various uncertainties - were presented at ICHEP last year, based on a $\Upsilon(4S)$ sample of 34.6 fb^{-1} of data and 100 fb^{-1} of MC simulations.

The tag-side reconstruction algorithm employs a hierarchical reconstruction of exclusive hadronic B decay chains, in which each unique particle decay has its own multivariate classifier. The new algorithm is based on 10,000 decay chains and more than 200 BDTs and resulted in an increase of the tagging efficiency by about factor two, and a very low purity of 10%.

A major task will be to assess and account for differences in the performance of the algorithm in data and simulation. To correct for these effects for the signal $B \to X\ell\nu$ decays, first calibration factors have been derived. For both B^0 and B^+ decays a variety of selections are used to determine the tag-side B multi-variate classifier. For a very loose selection the calibration factors are 0.653 ± 0.020 and 0.830 ± 0.029 for tag-side B^+ and B^0 mesons, respectively. The calibration factors increase with a tighter selection of the signal and especially of the tag decays and they increase the purity of the samples, and thereby reduce the systematic uncertainties. The purity of the tags and thereby the contamination of the B flavour needs to be addressed, especially for inclusive signal decay modes. The calibration factor provides a normalisation, but does not deal with the accuracy of the simulation of the huge number of decay chains.

The first analysis of a $B \to D^*\ell\nu$ sample tagged by hadronic B decay resulted in a signal sample with very modest background as indicated by the missing mass distribution, thus confirming the tag calibration procedure. First fits to the inclusive charged lepton spectrum are very encouraging, they are dominated by D, D^*, $D^{**}\ell\nu$ decays and large continuum background.

Considerable efforts have been made to assess the uncertainties. The dominant systematic uncertainty is associated with the shapes of the various contributions to fitted data samples. The next largest sources of uncertainty are those associated with the branching fractions for various $B \to X\ell\nu$ decays and also the tracking of low momentum charged particles and the PID performance, which are still under study. Of high priority are measurements of the CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$ using inclusive $B \to X_c\ell\nu$ and $B \to X_u\ell\nu$ decays and also exclusive decays, like $B \to D^*\ell\nu$ and $B \to D\ell\nu$, as well as $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays.

Tests of lepton flavour universality in the form of the ratios of branching fractions, $R(D)$ and $R(D^*)$, comparing $B \to D^{(*)}\ell\nu$ decays involving the τ heavy lepton with those involving electrons or muons, have resulted in an intriguing puzzle. Future tests will require a very large data set, and will benefit from the hadronic tag selections tuned to individual signal decay modes.

In parallel, semi-leptonic tagging processes are being explored. They rely on a very small set of tag decay modes. The first application of this method was a preliminary measurement of the B meson decay with the largest exclusive branching fraction, $B \to D^*\ell\nu$, resulting in a sample of very low backgrounds. Measurements of purely leptonic decays, $B^+ \to \tau^+\nu$ and also $B^+ \to \mu^+\nu$ are also very challenging and will also require a very large data set. They will benefit from both hadronic and semileptonic tags. Initial studies are quite encouraging.

To measure non-perturbative QCD parameters and the CKM matrix element $|V_{cb}|$, mass moments M_X^n of the hadronic system in inclusive semileptonic $B \to X_c\ell\nu$ decays can be used. This state-of-the-art procedure combines information from mass moments with measured moments of the lepton energy spectrum and of $B \to X_s\gamma$ decays to perform a combined fit using theory predictions based on the Heavy Quark Expansions of QCD to determine $|V_{cb}|$ and the b quark mass m_b. The X-system is defined by the rest of the event, i.e. the sum of all unassigned charged particles and neutral clusters in the ECL. In the future, the systematic uncertainties can be reduced by addressing the bias in the reconstructed M_X distribution as well as more extensive studies of the composition of unmeasured parts of the $B \to X_c\ell\nu$ spectrum. A recent theoretical proposal to include higher order terms promises higher sensitivity to derive $|V_{cb}|$.

Two branching fraction measurements, based on untagged samples, are expected to be ready for publication this summer. A measurement of $B \to D^*\ell\nu$ decays, identified by kinematics, resulted in values consistent with earlier measurements, with errors of 1.1% statistical, 4% systematic, and 10% due to slow pions. The data agree with the Boyd, Grinstein, Lebed prediction and will be used to extract $|V_{cb}|$. A similar analysis is in progress for $B \to D\ell\nu$ decays. Following a previous analysis, $B \to X_u\ell\nu$ decays are identified by their high momentum leptons above the dominant $B \to X_c \ell \nu$ decays and various backgrounds. The expected significance of the signal for this small data sample is $> 3\sigma$.

2.2.1.2 Rare B decays

Studies of rare decays of B-mesons provide interesting opportunities for New Physics searches at B-factories. Recent reports of "anomalies" in $B \to K^{(*)}\ell^+\ell^-$ transitions test our understanding of the Standard Model. This makes the studies of the rare decays $B \to K^{(*)}\nu\bar{\nu}$, which are significantly less affected by long-distance QCD effects, both timely and relevant. Such modes explore unique features of Belle II detector that are not available at LHCb. The first encouraging results of the analysis of $B \to K^{(*)}\nu\bar{\nu}$ with inclusive tagging were reported at the Moriond 2021 meeting. Further studies of this and related decays into invisible final states are eagerly awaited.

2.2.1.3 Dark Sector and Low Multiplicity Events

Belle II has unique capabilities to search for dark sector particles up to masses of a few GeV. Two papers based on 2018 data were published in PRL:

- a search for axion-like particles (ALPs) with a mass above the pion mass decaying into two photons;
- a search for an invisibly decaying Z', produced in association with a pair of muons or a muon and an electron.

These searches resulted in the world's most stringent bounds in large regions of parameter space of dark sector models. There are plans to extend the early searches to much larger data samples to be recorded this year.

Dark matter searches will become a broad and largely unique part of the Belle II research program. A search for dark scalar accompanied by a dark photon ("dark Higgstrahlung") decaying to $A' \to \mu^+\mu^-$ is well advanced. Another analysis currently underway is the search for an invisible dark photon, resulting in a mono-photon signature $e^+e^- \to \gamma A'$. The analysis relies on detailed understanding of the ECL and KLM performance. Based on a sample of 80 fb^{-1}, the measurement is expected to reach much higher sensitivity than an earlier search by BaBar. The collaboration is also planning additional searches like the search for inelastic dark matter, the search for visibly decaying dark photons, and the search for di-photon resonances produced from b decays.

The collaboration has put together a list of high priority dark sector analyses that commits to perform in the coming two years:

1. the search for an invisible dark photon;
2. an updated search for the invisibly decaying Z' produced in association with μ's or e's with a larger luminosity and the search for a Z' decaying to muons or taus;
3. the measurement of the $\gamma\pi^+\pi^-$ cross section.

The first two searches will be able to probe uncharted parameter space. The latter search is particularly relevant to improve the theoretical uncertainty in the Standard Model determination of $(g-2)_\mu$.

Precision measurements involving the τ lepton are being performed, including precision measurements of the mass and lifetime. A number of searches for lepton flavour violating interactions, including $\tau \to \mu^+\mu^+\mu^-$ decays, remain of great interest. Measurements of the ratio of the τ branching fractions to $\pi\nu$ and $K\nu$ will be needed to extract the value of $|V_{us}|$.

2.2.1.4 Special Run at 10.75 GeV

By extending their data beyond the $\Upsilon(4S)$, BaBar operating at the $\Upsilon(3S)$ observed the $\eta_b(1S)$ and $h_b(1P)$ and Belle at the $\Upsilon(5S)$ discovered the $h_b(1P, 2P)$, $\eta_b(2S)$ and the $Z_b(10610, 10650)$, thereby establishing a spectroscopy of new bosons containing a $b\bar{b}$ pair or a single b quark. There are now several bosonic candidates and their spectrum of masses, quantum numbers, and transition amplitudes have shed light on the strong interactions.

In 2019, Belle reported measurement of the cross sections for $e^+e^- \to \Upsilon(nS)\pi^+\pi^-$ ($n = 1$, 2, 3) at c.m. energies from 10.52 to 11.02 GeV. The data revealed three narrow structures, Y(10860), Y(11020) and Y(10750). The first two are consistent with the recently discovered bottomonium resonances, $\Upsilon(5S)$ and $\Upsilon(6S)$, while the nature of the third structure with a mass of 10.75 GeV and a width of 35.5 MeV is not obvious. The peak is out of the range expected for a conventional bottomonium state, although the current uncertainties of the theoretical predictions are large and the data statistics are low. It also does not coincide with any threshold for molecular state. However, it could be a candidate for a tetraquark ($[cq][\bar{c}\bar{q}]$), potentially mixed with the $\Upsilon(4S)$, the closest bottomonium state.

To determine the nature of the three peaks detailed studies of their production and variety of decay modes will be required. As demonstrated in earlier measurements, Belle II is well equipped to compare the decay modes of the three resonances with various predictions for different types of states. As stated above, for a bottomonium states the decays to $\Upsilon(nS)\pi^+\pi^-$ have been observed. For tetraquark states, one expects different final states, for instance $Y(10750) \to \Upsilon(1S)$ + hadrons with higher mass dipion resonances, including $f_0(980)$ and $f_2(1270)$. The preliminary Belle results indicate an enhancement at 10.75 GeV and maybe also at 11 GeV for the final states $\omega\chi_b(1P)$ and $\omega\eta_b(1S)$ consistent with predictions for decays of tetraquarks.

The recently formed the "Beyond $\Upsilon(4S)$ Task Force" has proposed a short special run with a total luminosity of 16 fb^{-1}, with 10 fb^{-1} recorded at 10.75 GeV and additional 6 fb^{-1} recorded above and below the 10.75 GeV peak to establish the decomposition to $B^{(*)}\bar{B}^{(*)}$ final states. The plan is to combine these new Belle II data with existing Belle data and publish a series of results on the various decay modes. Even with this very modest additional data sample complementing the Belle data, a significant advancement of our knowledge is expected. The hope is to record these data before the start of the LS1. In preparation, further analyses of the Belle data, and various Monte Carlo and trigger studies are being pursued.

There is no doubt that the goal to ascertain the line shape of the enhancement at 10.75 GeV and the determination of its decay modes are of great scientific interest. Previous Belle results on this topic have been well received by the experimental and theoretical community. Above all, Belle II is the only experiment worldwide that can provide the data to further study these kind of exotic bottomonium states. The preliminary proposal is still under study and discussions with the accelerator experts are underway. The suggestion to evaluate the Run 2021b data before making this decision seems to be generally supported by the collaboration.

2.2.2 Concerns

- The application of the Full Event Interpretation system with very low purity leads to higher tagging efficiency but large systematic uncertainties, thus fine tuning and detailed studies will be required to optimise the use of this very extensive process.

- Dark matter searches and other analyses of events with low multiplicity of charged particle tracks and photons critically depend on special triggers with high efficiency and effective suppression of large background rates. Future changes in the trigger design need to support these requirements.

- The highest priority of this year's operation is to record a data sample at the $\Upsilon(4S)$ comparable in size to the BaBar sample at this energy. There are four months left for Runs 2021b and 2021c, and there is hope that the current integrated luminosity of less than 1 fb^{-1} per day can be doubled. Experience has shown that the highest integrated luminosities per unit time are recorded in the last few weeks of a long run. Thus there is concern that the recording of this small data sample at higher energies in 2021c is likely to result in a significant reduction of the $\Upsilon(4S)$ data sample.

2.2.3 Recommendations

- Measurements of leptonic and semileptonic B decays are of fundamental importance to test the Standard Model of weak interactions. The assessment of experimental and theoretical uncertainties of the CKM elements and other parameters should be given utmost attention.

- The results of some dark sector searches may relate to the parameter space of more than one dark sector model; thus it would be useful to present results such that different interpretations can be possible. For instance, the results of a search for $B \to K\nu\nu$ decays could be interpreted in terms of a limit on a dark scalar or an ALP produced in B decays, but not detected.

- The collaboration is encouraged to continue a broad program for dark sector searches in the coming years, exploiting the much larger data samples and thereby exploring larger regions of the parameter space. Particularly important will be the development of a longer term program for dark sector searches.

- The decision on the special run near 10.75 GeV should be deferred until after the end of Run 2021b to assess the overall performance of the machine and the size of the recorded $\Upsilon(4S)$ sample, and to estimate the time required to raise the energy (of one beam or both) and to fine tune the operation to achieve stable conditions to collect data at these higher energies.

3 Machine and background status

3.1 Status

Progress is steadily being made in machine performance and in background understanding and remediation. The accelerator is becoming better understood by the machine group. The 2020c run was only 2 months long and there were several unexpected difficulties in getting the accelerator performance back to what it had been last June. However, the machine group succeeded in continuing to scrub the vacuum in both rings resulting in lower steady state background levels in the detector. The recent understanding of the Transverse Mode Coupled Instability (TMCI) threshold is a strong step toward understanding the specific luminosity decrease as the bunch current increases. The accelerator and detector teams have decided to optimise the accelerator at a β_y^* of 1 mm and gradually increase beam currents. Injection emittance is still high, especially in the HER and injection backgrounds, in general, remain high. The background team has made very significant progress in matching measured backgrounds against MC predictions. The ratio of MC to measured value is close to one for a large majority of background signals and this has generated increased confidence in extrapolating the present background levels to the design luminosity performance. The committee congratulates both the accelerator and the background teams for their combined effort in achieving this very impressive goal. A collimator jaw was damaged in the 2020c run, and the background and accelerator teams worked together to measure the collimator jaw positions using the beam. This revealed some positional discrepancies in some of the collimator jaws and with this information they were able to explain how the jaw came to be damaged. Future runs will include occasionally remeasuring the jaw positions and updating the jaw position database.

3.2 Concerns

- Injection backgrounds continue to remain high.
- Injection efficiency continues to be low.
- The TMCI threshold induced by the collimators sets another constraint on the achievable luminosity.

3.3 Recommendations

- As has been mentioned in several presentations, the committee strongly encourages improving the injection performance. This needs to include a plan to upgrade the injector hardware and sufficient running time to come to an understanding concerning some of the injection issues like the emittance blowup in the HER beam transport line.
- The decision to maximise luminosity delivery during the next run does not preclude further background studies and the committee encourages the continuation of this effort to further understand the backgrounds. New background sources may appear as beam currents slowly increase and as the vacuum slowly improves.
- Explore possibilities of low impedance collimation including optimised shapes and nonlinear effects and specialised materials.
- The committee thinks that the decision to optimise the accelerator at a β_y^* of 1 mm is sound and that it will go a long way toward getting a better understanding of the machine.

4 Belle II detector hardware status

4.1 Status

The committee commends the Belle II detector group for their successful data taking during the challenging COVID year of 2020. All detectors delivered high quality data and remained rather stable during the run. There were incremental improvements in stability and improved monitoring as the run progressed. Nearly 90% data taking efficiency was attained. Further improvements to factory mode operation are anticipated. All detectors are considered fully ready for steady data taking in 2021, and capable of coping with the higher expected background levels with the 1 mm beam optics as the currents and luminosity are increased. Monitoring of detector performance and background conditions are continuous and have also been improved. It is important to continue to monitor radiation dosage for the SVD and refine estimates for the inner detectors and for neutrons. The TOP PMT limits will be relaxed to 3 MHz/tube to allow easier conditions for machine tuning and data taking operations.

4.1.1 PXD

At the individual detector level, the PXD operated stably using additional monitoring and optimisation tools. Increased HV and bulk currents were observed, eventually reaching Power Supply (PS) limits that resulted in performance degradation at the end of the run. Dedicated X-ray studies showed that the current saturated with increasing HV at around 5 to 7 mA, so the PS current limits could be increased.

4.1.2 SVD

Both the hardware and software for the SVD performed well in 2020 and are ready for the 2021ab runs. The online systems, including calibration, operated stably. The radiation damage effects are at the expected level and do not affect performance. It is crucial to continue careful studies of radiation effects and to accurately model the dose estimates as the luminosity increases. SVD reconstruction time for the HLT was improved during the run, and this work should also continue. Successful tests were carried out with a new DAQ mode. Rejection cuts based on hit time are ready in case backgrounds are higher than anticipated in 2021ab, and the DAQ, DQM and HLT should be able to cope with the expected rates.

4.1.3 CDC

The committee was pleased to learn that CDC demonstrated stable operation in run 2020c with good performance. The injection background, clearly seen by the fast current (DESY) monitor, dominates operational currents in CDC. Typical storage currents are of the order of 40 μA per layer, while, during injection, CDC current spikes can be two to four times higher. Due to this effect and on the longer term, hit rate limitations and higher CDC occupancy might compromise tracking performance; more studies are anticipated. Thanks to repairs during the last summer shutdown, the number of masked front-end boards has decreased substantially (from ten to two). In particular, it was found that the cross talk between the clock and trigger signals had caused DAQ errors, such as lost links. After dedicated studies, layer 54 was included back into physics data taking (at nominal HV) since the middle of run 2020c. First check of raw data from layer 54 revealed similar performance as in the other CDC layers, i.e. no degradation in gas gain. No dark ("Malter") currents have been observed in the chamber throughout run 2020c. Ratios of CDC tracking efficiencies per layer, divided into earlier and later running periods, are consistent with 1.0 within large statistical uncertainties. More precise wire efficiency studies are expected to be done with new calibration constants.

4.1.4 TOP

TOP operations were stable in 2020, with 95.6% of channels operating in run 2020c. Eight boardstacks need to be replaced during the 2022 shutdown. A few high hit rate PMTs, increasing current draws, hit rate spikes, etc., have been observed and not been understood yet. They are being closely followed as they occur in all three PMT types and may impact PMT lifetime. The DAQ firmware has generally been well-behaved, but several problems still occur that create issues while running. Understanding the board-stack crashes is especially urgent. The reconstruction code is being ported from FORTRAN to C++, including documentation. This should broaden the personpower base of those who can work on improving the code, and there is some early evidence that it may improve performance. TOP is the subdetetor with the most elaborate simulation tasks, so substantial studies are also underway to improve speed, and PID and simulation performance via machine learning.

4.1.5 ARICH

The ARICH ran reliably with stable ring parameters during 2020 and is ready for 2021ab data taking. One merger board is still excluded due to a disconnection from the -2 V power line. and can only be fixed with both yokes open, probably in 2022. There has been a slight increase in the number of masked HAPDs due to bias and HV problems. Channels fail occasionally and need to be disabled. The failure rate is about 1.5% per year. It must be carefully watched as it could become a significant concern for long term operation, although simulation studies show a modest impact over the long term if the rate does not increase. Updates have been made to the firmware to provide recovery from single event upsets. Substantial work is ongoing to try to improve PID performance, centred mainly on using a more accurate PDF, and on dealing better with tracks that scatter or decay before reaching the ARICH ($\sim$10% of tracks).

4.1.6 ECL

All ECL channels are alive, and no new noisy channels have been found in the system. The firmware of the front-end electronics, as well as the DAQ software, have been working fine, and the system operated stably during run 2020c. The alignment of crystal positions is in progress using events that contain muon pairs. The displacement of crystals is visible, especially in the endcap region, with misalignments up to 0.15 degrees in θ and 0.4 degrees in ϕ.

4.1.7 KLM

Major efforts to consolidate the KLM system and reinforce the grounding scheme were made before the start of the 2020c run. Reinforcing the grounding scheme and shortening the CAT7 cables brought stability to the system. After threshold and gain calibrations were performed, the layer efficiency has improved with most layer efficiencies now exceed 80%.

4.2 Concerns

- The challenge for 2021 will be to keep the detector operating well for data taking as luminosities and backgrounds increase while improving factory mode operations so that data taking efficiency exceeds 90% routinely. This requires continued attention to improving the monitoring, and to automate the DAQ recovery procedures, to decrease the level of effort that must be expended on operations. In particular, SCROD processing abort/reset needs work, and PID performance still needs to be better understood and improved.
- Even if no persistent dark currents have been observed during run 2020c, careful online monitoring of the CDC chamber performance (using online tools and DQM histograms, fast current monitors and offline wire efficiency studies) is mandatory.

- Shortage of personnel for the CDC tasks (such as those necessary for ageing tests, new gas mixture studies) remains critical;
- Although KLM layer efficiencies have improved, they are still substantially lower than expected.
- The detector group's person power will be challenged by the overlapping demands of operations and calibrations, ongoing preparations for the 2022 Long Shutdown, and the need to increase analysis efforts to better understand performance at the detector level to support the highest possible quality physics analyses of the 2021 data set.

4.3 Recommendations

- The detector groups are encouraged to strongly enhance their efforts to help understand and improve detector performance, especially by carefully studying the performance of detectors in the relevant detector phase space (as discussed in Section 1), comparing with MC, and improving reconstruction software. In particular, the TOP seems most in need of substantially increased effort.
- Every effort should be made to ensure that the level of personnel engaged in CDC hardware and software studies is adequate to address critical concerns for the CDC long-term performance.
- Efforts to optimise the KLM layer efficiencies should continue, focusing on the layers with lower efficiencies.

5 Belle II detector performance

5.1 Status

The latest run 2020c in October–December 2020 was characterised by challenges in running the accelerator at high luminosity, with the luminosity lower than in the previous run. However, the Belle II detector performance was relatively stable and the data taking efficiency was about 89% within the physics running time, approaching the target. The sub-detectors all worked well, with no major problems.

One of the limitations that has been applied to the accelerator operation up to now has been that the background rate in the MCP-PMT photodetectors of the TOP should be kept below 1.2 MHz per PMT, to limit the ageing of the detectors due to the integrated charge at their photocathodes. The decision has been taken to replace the MCP-PMTs most vulnerable to that ageing (those without ALD treatment) during the shutdown LS1 currently scheduled to start in Summer 2022. As a result, the background rate limitation will be raised to 3 MHz per PMT until LS1. The committee agrees that this is a pragmatic approach, to help in maximising the integrated luminosity. It notes that the allowed background rate has been calculated under the conservative assumption of continuous operation at high rate, whereas in practice it can be expected that the integrated rate will be lower by a factor of two or more, due to interruptions in the accelerator operation. Until now there has been no measurable decrease in quantum efficiency (QE) of the PMTs due to ageing. With the higher rate maintained after LS1 it is expected that the next class of PMTs (with ALD treatment but not lifetime-extended) will need to be replaced by about the time of the following long shutdown, and this replacement campaign is also foreseen. When the background limitation from the PMTs is relaxed, other sub-detectors may see undesirable effects as the rate increases. At present the simulation studies indicate that the detectors should all be able to handle the increased rate without significantly degraded performance, in particular as they were designed to cope with the higher luminosity that is expected in the long term from the accelerator. Care will need to be taken, for example in monitoring the CDC currents.

Good progress was reported on the detector performance, with focused reviews presented on charged particle tracking and neutral reconstruction, and a more comprehensive discussion of the charged hadron identification. In general the committee would welcome presentations that give an overview of the performance, rather than only focusing on a few topics, and show the performance separately by sub-detector as well as combined. However, this wider perspective was largely covered in the discussion around each presentation. The performance of lepton reconstruction was not reviewed at this meeting, which was unfortunate for such an annual review. The committee commends the collaboration on achieving performance for many aspects that already reaches, or even supersedes, that of the original Belle detector.

The tracking report focused on two issues, the slow pion efficiency and the performance of V^0 reconstruction (K^0_S, Λ). While the tracking efficiency is well reproduced in the simulation for momenta above 200 MeV, as determined using a tag-and-probe method on tau pair events, at the last review a discrepancy had been reported at lower momenta, where a 20% higher efficiency was seen for data compared to MC for the slow pion from $B \to D^*$ decays. This has now been understood as due to peaking backgrounds in the M_{bc} distribution, and has been corrected by instead fitting the ΔE distribution when extracting the efficiency. For the V^0 reconstruction detailed data-MC comparisons have been performed using large inclusive samples of K^0_S and Λ. Some deviations have been seen, up to a 20% enhancement of the data over MC ratio of the yield of reconstructed V^0s, as a function of the x-y distance of the decay vertex from the IP. Interestingly the discrepancy appears to be a smooth function of the distance, in the region between VXD layers 4 and 5, indicating that this may be caused by a bias in the vertex reconstruction, as that is the transition region between tracks which are CDC-only or have CDC+SVD hits. V^0 reconstruction may be the place where the effect of increased background is first seen, particularly from extra hits in the CDC.

The report on the performance for neutral reconstruction focused on the issue reported by physics analysis groups raised at a previous review, of the energy bias seen in the distribution of π^0s from channels such as $B^+ \to \overline{D}^0\pi^+$, with $\overline{D}^0 \to K^+\pi^-\pi^0$. This has been traced to an energy bias of up to 0.7% seen in data for low energy photons (below 1 GeV). A correction for this can be applied at analysis level and is in preparation. Photon timing cuts were also presented, where loose cuts can help to reduce out-of-time background while maintaining high signal efficiency. Such timing selection is expected to become more important with the increase of out-of-time background in the future. The photon timing resolution is a factor two worse for data than MC: the accumulation of crystal-by-crystal mis-calibrations in data is suspected of causing this, and the ECL group is now re-evaluating the current status of the crystal timing calibrations. Progress is being made in unfolding the tracking resolution from the determination of the photon energy resolution. Finally some encouraging work was presented on the identification of neutral hadrons, for both K_L^0 and neutrons, combining pulse shape discrimination with shower-shapes, with a note in preparation.

A comprehensive review was presented of charged hadron identification, from the combination of the sub-detectors that contribute. The ARICH is performing well for the tracks which fall within its acceptance, and is well described by the simulation. The CDC dE/dx gives good separation at low momenta below the cross-over point (up to around 1 GeV for K-π separation) and approaches 2 sigma separation above 2 GeV, a region where it is currently providing better separation than the TOP. The efficiency based on the CDC dE/dx shows a dip for kaon identification in data at central polar angles. This drop in efficiency has been interpreted as a space-charge effect for ionisation electrons produced near the sense wire, and a more detailed simulation might reproduce the observed affect. On the other hand, a similar effect was observed in the BaBar drift chamber and was resolved by slightly increasing the operating voltage of the chamber. The loss can be explained by the short paths in the CDC cells for central tracks. dE/dx from the SVD has not yet been added to the combined performance, and should give significant improvement for soft tracks (below 1 GeV). The performance of TOP is not yet achieving expectations, and there may be room for further optimisation. It would be good to keep the pion mis-identification rate below 5% over a wide as possible region of the overall phase-space, given the large number of pions in many decays of interest. There remains discrepancy with the simulation, in particular at high momenta and forward polar angles. Photons with long path lengths in the TOP radiator may not be well simulated, as the detailed description of their absorption will become more important, and it may be instructive to separate them (e.g. from a 2D binning of performance vs momentum and angle, or applying timing cuts).

From the combined performance, the kaon efficiency at 10% pion miss-identification rate is 84% in data and 90% in MC, and it was noted that a similar data-MC discrepancy persisted in the Belle particle ID, although both figures were a few percent higher there. The focus is now on improving the TOP performance, e.g. correcting the PDFs used in the likelihood calculation for tracks that are extrapolated to the TOP but actually scatter or decay-in-flight before reaching the detector. Comparison of analytic and MC-based calculations of PDFs are seen to give large differences, which may be a key to understanding discrepancies. Moving from FORTAN to C++ has provided an opportunity for reviewing the reconstruction software and should help broaden the access to the younger physicists, who can now help scrutinise the performance and pursue alternative approaches e.g. for the PDFs.

5.2 Concerns

- The increased background rate is expected to start showing noticeable ageing for a subset of the MCP-PMTs of the TOP in the coming runs. The effect is subject to significant uncertainty, given that the tolerance of the PMTs has been determined by accelerated ageing for a small number of subsample. Some PMTs are also showing high rates, that might be a sign of an alternative damage mechanism e.g. due to increasing ion feedback.

- The analysis of the TOP is complex, and its performance is not yet fully understood. There are persistent discrepancies between data and simulation.

5.3 Recommendations

- The QE and high rate behaviour of the TOP PMTs should be carefully monitored. There should be a flexible approach to the background rate limit, so that it can be adapted if necessary during the run.

- Close connection should be encouraged between the detector experts and those studying overall performance for physics. There may be room for more people to get involved in bringing the details of the detector knowledge to the performance studies, to enhance the combined performance. Strengthening the engagement of experts with Belle experience may be helpful here.

- The committee recommends to improve the CDC simulation, where track level dE/dx simulation may not be adequate to describe the details of the likelihood distribution in the central region, and to investigate with the detector experts if different operating conditions may partly recover the loss of performance in this region. Remaining sources of disagreement between simulation and data at the detector level for the TOP should be identified with high priority, as they may be key to understanding and improving its performance: e.g. the behaviour of long path-length photons and the analysis of probability distribution functions over the plane of position and time for the detector.

- Detector performance should continue to be scrutinised and maintained at a high level, ready for the large data-set that is foreseen to be taken in the coming runs before the shutdown, paying close attention to systematic effects given that the statistical uncertainties on physics measurements will shrink with the increasing data.

6 Trigger, DAQ, detector control and operation

6.1 Status

The overall operation of the experiment in the 2020c period was quite satisfactory. Almost 90% of data-taking efficiency (DAQ live time) has been achieved. The remaining dead-time is due to the injection veto (about 5%), and various detector problems or calibrations.

The trigger had a stable L1 rate of about 3.5 kHz, well below the DAQ limit of 12 kHz. The new trigger board (UT4) has started production without problems and will continue to replace its predecessor during 2021. The two 3D-track triggers (conventional and neural-network based) are now ready to go into production, and comparison studies will take place during 2021. The trigger group will study now the maximum allowable rate the trigger system can cope with. The trigger is now available in the simulation.

The DAQ has worked well. All detectors except the PXD are still using the original readout-board COPPER, the upgrade to the PCIe40 will be made for KLM and TOP in summer 2021. The COPPER boards are quite old by now and have been one of the main contributors to DAQ down-time. Hardware replacements will improve the situation.

The data-reduction mechanism in the ONSEN system was tested successfully but not yet applied. The data-flow in the HLT has been newly based on the industry-standard ZeroMQ protocol. The HLT is now at about 75% of the nominal capacity and capable of processing about 14 kHz of L1-rate. Several issues were fixed to improved the robustness of the HLT. There is still a remaining sensitivity to “big” events, which can cause out-of-memory conditions.

Significant progress has been made in the monitoring system which is now based on a modern ELK (Elasticsearch, Logstash, Kibana) stack.

The personnel situation in the DAQ group has improved both by a new group and by adding people paid by the collaboration maintenance and operation fund.

In the control system a major reason for instability in the timing distribution has been identified. Some of the CAT7 cables were found to have insufficient signal transmission qualities and were replaced by better alternatives, which improved the situation considerably.

The slow-control system was very stable during the 2020c period and did not cause any significant down-time. Consolidation is progressing and the slow-control team has put quite some effort into providing remote access tools to facilitate running during the COVID pandemic.

6.2 Concerns

- The shortage of personnel in the trigger group remains a concern, in particular for sub-system triggers, e.g. the ECL. This may delay important studies and improvements.

- Currently, the memory use by the HLT is quite high and it may require a work-around which reduces the overall computing power of the HLT cluster.

- The fast control TTS links have shown robustness issues and caused downtime since the beginning of Belle II operations.

6.3 Recommendations

- Slow-control software consolidation and computer and network administration would profit from dedicated effort (system administrator). This could offload the core slow-control team from important but routine tasks.
- The LHC experiments have also faced the problem of large memory footprint of their HLT applications. There is a priori no reason why the techniques that they developed could not be used also for the benefit of the Belle II HLT, so when some effort is available it could be a good investment to study this work.
- As the TTS cables will remain after the replacement of the COPPER by the PCIe40 boards in the context of the DAQ upgrade, they will require continued attention to avoid downtime.

7 Computing, software and data preservation

7.1 Status

7.1.1 Computing

The committee congratulates the Computing group on the successful transition to operation of Rucio for distributed data management (DDM). The migration took place in January 2021 and involved developers in four different times zones over a span of four days through the week-end ending with the system ready for final tests on Monday morning. The certification process that started in June 2020 was key to the success of the migration allowing them to establish a clear list of actions and related responsibility. Minor issues were found and quickly fixed in the two days testing period, longer than the eight hours initially assumed. The transition was performed smoothly thanks to the deep commitment of everyone involved, coupled with very effective communication channels. Many benefits in computing operation have already been observed from the new Rucio DDM. The transfer backlog accumulated before the Rucio migration was quickly finished with 100k files/hour throughput. Discussions are in progress within the Computing group to enable more of Rucio features. Automatic deletion of files, popularity of files and datasets to optimise disk space are expected soon. The committee commend the group for the debriefing exercise planned at BNL.

The production system was reported to be more robust compared to last year. Various updates targeting performance improvement have been introduced, e.g. parallel processing. A scheduling mechanism to replace the manual one-by-one submission of productions is planned for deployment in 2021. Automatic staging of raw data via Rucio is under discussion. Scout jobs, where the bulk of the jobs are submitted only after a smaller number have been successfully completed, has been introduced in the production environment to cope with failures of users' productions and will soon become the default. To cope with the long standing lack of person power for production operation, a new booking workflow has been put in place for experts shifts, with a manager taking care of the assignment to each institute in geographical regions. Rucio dedicated training for expert shifters will be setup.

The online-offline raw data copy and conversion is handled by the Core Computing group and has so far been stable. Nevertheless it has been estimated that the system would barely cope with four times the 2020c data taking rate; as a result it will be necessary to introduce HLT filtering before reaching peak luminosity in 2021. A second copy of the whole raw data is currently hosted at BNL. Starting from the 2021 Japanese fiscal year, six Multiple Raw Data Centers (RAWDC) will share the copy, each with a fraction ranging from 10% to 30%. The distribution of the data to multi RAWDC by Rucio will be tested with the 2021a data, and the *BelleRawDiracData* backend for Grid registration is in preparation.

The status of the current reprocessing, *proc12* to produce the baseline data set for the 2021 summer conferences was presented. It encompasses all data collected in 2019 (*exp7, 8, 10*) and 2020 (*exp12, 14*) corresponding to an overall integrated luminosity of 78.7 fb^{-1} at the $\Upsilon(4S)$ and 9.2 fb^{-1} off-resonance. Reprocessing started at the end of December 2020 and the re-calibration was expected to be completed by the end of March 2021 when production of re-processed mDST would start. The data pre-processing was being done mostly at BNL with the re-calibration split in three stages with sequential RAW data staging/purging. All events from *exp7-12* ($\sim$ 70 fb^{-1}) also required a one-off pre-processing to adapt to the different HLT skims used in later processing. Some problems were observed with the merging of the *RawProcessing* output files where $\sim$ 1-2% of problematic tasks caused long tails in processing time due to data transfer. The deployment of Rucio appears to have mitigated the problem with less movement away from the *RawProcessing* sites for the merging. Triggering of the staging and purging of files is still done by hand by the distributed computing team, but automation of both is planned to be put in place in 2021. Preparation of prompt processing of 2021a-b data with *proc12* was also mentioned with the goal to add at least four processed prompt *buckets* to the data set for the 2021 summer conferences, potentially reaching 150 fb^{-1} of *proc12*+prompt processed data.

The experience with the Belle II calibration in 2020 was reported. The calibration is stable and shows overall good quality with many improvements introduced since 2020c data taking, e.g. on CDC dE/dx. The calibration is performed locally at the calibration centres: BNL for prompt calibration (since autumn 2020) and DESY for re-calibration for major processing. The four step calibration flow has been automated and is managed by an *Airflow*-based job scheduler hosted at DESY. Prompt calibration is performed bi-weekly during data taking. Local calibrations are followed by raw calibration and alignment and finally post-tracking quantities are determined. The resulting *cDST* file are transferred to DESY for detector studies and re-calibration. Despite the automation, the calibration is still slower than what will be necessary for efficient future turn-around. Automatic production of calibration plots has been implemented in the Calibration Alignment Framework (CAF) software and integrated in *AirFlow*. The complete sequence of automated production, validation and sign off will be exercised in 2021a data taking in order to reach the calibration target of one *bucket*/week. For long term sustainability the calibration is planned to be moved to the Grid for 2023 data taking.

A status report on Monte Carlo (MC) production was also presented. The 2020 *MC13* campaign using the older software release 4 was being concluded and would provide samples for analysis to be presented at Moriond 2021. The 2021 MC14 campaign had just started: the run independent *MC14a* was progressing well with half of the planned equivalent 1 ab^{-1} generic sample already produced. The run dependent *MC14b* equivalent to *proc12* was planned to start after calibrations would be ready. It will be followed by *MC14b* corresponding to 2021a-b data. Storage needs are to be carefully checked with inputs from physics and performance groups for signal samples.

Further updates in computing infrastructure are forthcoming following the KEKCC replacement. The move from *RHEL6* to *CentOS7* for *VOMS* and *StoRM* will result in a downtime of one week to be accommodated during the summer shutdown. Python2 to Python3 upgrade is is planned over the summer. The move to non-SRM protocol will also requires an update from *DIRAC* v6 to v7.

The request for the 2022 computing resources were presented and the computing resource scrutiny group will evaluate them in June. A shortage of tape space was reported due to the use of more than what expected in 2020 since more data was collected than anticipated. Computing centres were asked to provide additional tape storage in order not to affect 2021 data taking. The committee supports the request.

7.1.2 Software

The committee congratulates the software group on the impressive work reported at the annual review meeting.

The committee was reminded of the category of releases Belle II supports ranging from major releases with changes requiring in depth careful validation to light releases only for physics analysis. A change in strategy has been introduced with only one major release for production planned per year. The latest Release 5 is currently being used for data taking in *proc12* and MC production, *MC14*. Automated deployment and validation of patches is in place while major releases are still manually validated. The major effort for release 5 was the Geant4 upgrade from 10.1.2 to 10.6.1 with VecGeom solids enabled and optimised Belle II custom physics lists. The extensive validation carried out has shown no effect on physics quantities.

Considerable improvement on processing time was reported with 30-35% decrease observed both in simulation and in reconstruction. The improvement on simulation is a result of the new Geant4 version and physics list, while the faster offline reconstruction is due to optimisations.

The target for the next major release 6 is summer 2021, with the main focus on the implementation of timing effects and further improvements for Data/MC agreement. Additional improvements on tracking and high level quantities, e.g. skim flagging, are also envisioned. The introduction of timing effects will require a coordinated validation effort on reconstruction for all sub-detectors as well as modification in the simulation chain to be performed well ahead of the release. The software group intends to automate it as much as possible.

Strong effort in is place to speed up the HLT reconstruction to optimise throughput and for which a 30% improvement has already been achieved. A strategy has been defined for further optimisation with short, mid and long term achievable targets with different level of potential gains and efforts required. The current main strategy of the software group is to avoid any impact on physics performance by deferring some computational parts to later stages in the decision process and making sure the results are identical on an event-by-event basis.

It was reported that the *B2BII* software tool to be able to perform the same analysis on Belle and Belle II data has now a fully automated configuration. The committee was very impressed that Belle data high level objects, with no new reconstruction or simulation, can be used seamlessly with Belle II high level tools, e.g. FEI,.

The Collaboration is starting to address long term data preservation for analysis and introduced compatibility tests for data objects in *uDST* to support it.

A new systematic framework to automate and unify the calculation of systematic uncertainties or Data/MC weights has been setup, inspired by a method developed by LHCb. The framework uses a new *b2luigi* tool based on the Spotify Luigi workflow management, that was felt to fit best the requirement and for which the connection to the Belle II Grid analysis job submission, *gbasf2*, was already in place.

The Belle II Institutional Board ratified the final decision to make the software open source, aiming for an LGPL licence. The well established procedure in software products like Apache and Google of Contributor License Agreement (CLA) will be followed with the copyright remaining with the authors. A restructuring of the repository and clean up of the code to remove any sensitive information will take place before the new repository is made public. Taking advantage of the complete change in repository a more inclusive renaming of the principal branch from *master* to *main* will take place at the same time. The move is planned for July 2021 after the Release 6 feature freeze to minimise the impact of a two weeks period when it will not be possible to publish new code.

The COVID epidemic forced a shift in the software training approach: providing self training material has been favoured following the model of Software Carpentry and HSF with references to their training material for *software prerequisites* topics. A first online StarterKit training event in the new style took place in October 2020, where the new self training *Online Book* was complemented with a KickOff event and daily Q&A zoom sessions. The Belle II software group is evaluating to keep this online format for initial training and provide in person lectures for advanced training after the COVID restrictions will be lifted.

7.1.3 Data Preservation

The committee was provided with the draft of a *Recommended Belle II Data Preservation Plan* document prepared by the Data Preservation Task Force. Two different periods where physics publications would continue after the completion of SuperKEKB running and Belle II data taking have been defined based on the analysis of publication and citation information of completed collider experiments. The first period, immediately after operation is completed, would involve final calibration and processing of data and

Monte Carlo as well as Analysis and generation of signal Monte Carlo. Experience from past experiments indicates that this period would extend over about 5 years when around 20% to 30% of the publications takes place. Afterward the fraction of publications is rather small requiring minimal resources: it would coincide to when direct funding of the Belle II collaboration would end. The resource needs could in principle be met via funding agencies' ongoing commitments to make research data available to the public. In order to support data preservation a new *final* DST (*fDST*) processed data product with all necessary information for all stages of Belle II lifetime should be developed. Analysis infrastructure would also need to be preserved. Documentation of the analysis workflow will need to be provided to ensure reproducibility. Formalisation of analysis workflows would help for that. Additional effort would be needed to support data access beyond former Belle II collaborators.

7.2 Concerns

- Will need to monitor carefully massive user job submission.
- The time needed to fully automate the validation of major releases should not be underestimated.

7.3 Recommendations

- Continue with exploitation of Rucio functionalities, e.g. automatic deletion of files, to ease the operation of distributed computing. The committee supports the team effort on spreading knowledge of the system.
- The Computing group is encouraged to work closely with the physics and performance groups for setting priorities for production jobs.
- If in the longer term any changes become necessary for optimisation of HLT that affect the physics performance these should be fully evaluated and discussed with physics and performance groups before proceeding.
- A new workflow management tool is being used for the newly established framework. While no additional development for python module used is required it will require some dedicated knowledge and maintenance. How to harmonise it with the calibration and distributed computing workflow managements, *AirFlow* and *DIRAC*, should be investigated.
- The committee supports the plans for Data Preservation outlined in the document they were provided. This will enable post-operation analysis and publication at a level comparable to that of similar completed experiments. The collaboration should make a cost benefit analysis of resources needed, including support by experts, to enable access to data and maintenance of computing resources to the post-operation period.

8 Long shutdown 2022

8.1 Status

The collaboration developed a detailed work-plan for the long shutdown in 2022. The shutdown is now scheduled from mid-July 2022 to mid-May 2023. The later start date provides more time for collecting a sizeable data sample for physics analyses and for preparatory work, like the delivery of the new beam pipe. Key points of this shutdown are the installation of the fully equipped PXD2 on a new beam pipe and the replacement of conventional PMTs with life-extended ALD MCP-PMTs. Integration of all the activities, including interaction with SuperKEKB, is organised by the mechanical structure design group leader and the VXD schedule management. A detailed schedule was provided as support material to this meeting.

Concerning the PXD production, there are still two sensor lots, PXD9-20 and PXD9-21 being processed. During the flip-chip of PXD9-20, a vacuum pump failure caused a wrong temperature cycle. After the vacuum pump repair, an interlock on the temperature cycle has been added and the process re-optimised splitting the SWB and DHP+DCD bump-bonding steps. An additional six weeks delay has been incurred due to the replacement of the silicon plate of the furnace. Of the 27 modules affected by the flip-chip issue, 23 were recovered. To ensure their reliability, a thermal cycling step is now introduced before kapton attachment. While proceeding with module and ladder assembly using the PXD9-20 sensors, the PXD9-21 batch is going through the phase 3 of wafer processing (thinning and copper deposition). There are currently eight Layer 1 (L1) ladders fully assembled and tested and 12 Layer 2 (L2) ladders. To reach the required number of spares, seven more L1 ladders and four more L2 ladders are required. To achieve this goal only four modules from the PXD9-21 batch are needed. A few of these modules may have about 1% efficiency loss due to shorts between the gate-lines and the *cleargate* signal that are fixed by cutting the affected lines. Remaining sensors will be kept for assembly with the new generation of switcher chips, up to 25 additional modules can be assembled with the available parts. The schedule had been already updated with the additional delays incurred. Availability of PXD2 is not on the critical path. The assembly of the first half shell can start in April with the best modules, the second half shell will wait for the availability of the remaining ladders. Shipment to KEK is foreseen in October 2021. The committee welcomes the recruitment of four PhD students and post-docs that will be involved in the installation and operation of PXD2.

The PXD2 will be mounted on a new beam pipe. It consists of two crotch parts and an IP pipe section with a shorter straight region than the current one. The gold plated inner wall of the IP section will reduce the flux of synchrotron radiation photons. The production of the crotch parts is on schedule. The IP section was found to have a deformation of 0.8 mm for the inner tube and 0.3 mm for the outer tube after the brazing process. This was due to the omission of installing spacers during the process. A second IP section will be built and the delay has been absorbed in the schedule contingency, keeping fixed the beam pipe delivery date in September. The beam pipe replacement will also address an observed hotspot in the forward IP section by improving the cooling system on the crotch.

Replacing the PXD and the beam pipe requires the removal of the whole VXD and disassembly of the SVD half shells. There is additional work related to the VXD reinstallation, including additional shielding on the bellows, modification of the radiation monitor system with new diamond detectors and non-halogen connectors, the replacement of the CDC inner-support ring and QCS heads. The latter action will increase the space available between the CDC and the QCS, necessary to route the full set of PXD2 cables. It will also allow some margin for the relative movements of CDC and QCS, that showed drifts up to 250 μm during Belle II operation. At the time of the meeting the design of these components was not completed and alternative options for cable routing, including modification of the PXD Infiniband cables were presented as fall-back solutions.

The committee appreciates the decision to create a VXD2022 installation group and a VXD2022 review committee, involving members from all groups of the VXD community. The VXD2022 installation group has developed and documented the first version of the disassembly procedure and it is taking care of the knowledge transfer of the beam pipe/VXD assembly. In order to optimise the schedule, VXD disassembly will make use of the ARICH room, while the PXD attachment can start at an earlier date in the B4 clean room. The disassembly procedure was reviewed for the first time in February, suggesting the need to develop also a PXD detachment procedure, so that the present L1 can be preserved as a spare. The review process will continue to cover all aspects of VXD reinstallation.

The 2022 long shutdown also provides the opportunity for replacing the conventional MCP-PMTs with life-extended ALD MCP-PMTs. The parts have all been procured and 257 PMTs in hand, corresponding to the 224 required and additional spares. Replacing the conventional PMTs in the 2022 shutdown and the ALD PMTs in the 2026 shutdown allows TOP to be operated at a much higher background rate of 5 MHz/PMT. Bad boardstacks will be exchanged together with the PMTs. Most components are already available, but some of the front boards are being checked before shipping to KEK, and the optical cookies will be produced in the second half of 2021. The total TOP intervention is compressed into two weeks, with additional time for CDC electronics disassembly and reinstallation. A training plan for people involved in this replacement has been laid out, using a spare module and building a mock-up of the CDC ring. The overall schedule and personpower needs seem to be reasonable and attainable at this time.

8.2 Concerns

- The long shutdown activities involve not only the upgrade of the PXD2 and TOP, but also the SVD disassembly and work on the CDC.

- Even if the right emphasis is given to the inclusion of new personpower and transfer of competencies, it seems there are still tasks depending on the availability of a single key person. These may pose risks to the schedule.

- The PXD2 schedules appears to be credible and the detector assembly does not seem to be on the critical path for the 2022 shutdown. Nevertheless, additional failures in the final steps of PXD9-21 processing, ladder production or half-shell assembly may still have an impact on the schedule or on the quality of the final detector.

- The VXD disassembly and installation procedure is still not completely defined and there are open issues, like the need to remove the cover of Infiniband cables that poses a reliability concern.
- Simulations of the performance degradation of the ALD MCP-PMTs indicate that efficiencies of some of these PMTs could drop to 80% QE by sometime in 2026. As the run proceeds, careful monitoring will be required to assess the problem to understand this better, and plan a response. Replacements may be needed which will require appropriate financial investment.
- Other issues with the PMTs, such as hit rate spikes, high hit rate areas for tubes and current fluctuations need careful study. These issues may be a concern for the longer term.

8.3 Recommendations

- It is advised to evaluate the risks and implications of unforeseen incidents during the repairs and upgrades of the VXD and TOP during the major shutdown activities, and to develop contingency plans to be able to quickly react in such situations.
- The committee also advises to evaluate which activities in the shutdown schedule are sensitive to the availability of critical personpower and verify if the schedule has enough flexibility to accommodate for unforeseen absences, like it was the case of travel restriction due the pandemic. Following this verification some mitigating actions might be considered.
- The committee recommends to carry on the PXD2 assembly according to the updated schedule presented at the meeting. Emphasis should be given to avoid modification to the PXD2 assembly and QC procedures, which could have unexpected side effects, and may require the (re)training of the people involved in the half shell assembly.
- The committee recommends to plan the assembly of the remaining sensors from the PXD2 production into modules with the new generation of SWB.
- The documentation and review process of the VXD disassembly and installation should be continued until all procedures and design issues, also affecting other systems, as in the case of the QCS cover replacement, are completely solved.
- The committee strongly supports the outlined path for the preparation of the PMT replacement, in particular the preparation of the missing optical cookies and of the installation mock-up for training.

- The QE degradation model for the MCP-PMTs has yet to be tested in the experiment, since no degradation is expected as yet. With the rate increase to 5 MHz/PMT, it will be extremely useful to monitor the actual QE degradation in order to validate the model and the projections until 2026.

9 Future upgrade

9.1 Status

An upgrade plan for SuperKEKB has been presented to MEXT as part of the "Roadmap 2020" process with the aim to increase the peak luminosity by at least a factor of two and reach an integrated luminosity of 50 ab^{-1} around 2031. The plan calls for a partial RF-power upgrade and an upgrade of the full interaction region in 2026, called "intermediate upgrade". It was selected as one of the 15 highest priority projects by MEXT, although a final decision is deferred and will depend on the progress improving the performance of the accelerator complex in its current configuration. Based on the feedback from the BPAC as well as the recommendations from the Accelerator Review Committee, regular long-term planning meetings have been scheduled where the upgrade plans are being discussed. For example, the need and feasibility of the IR upgrade is currently being reviewed. Studies to improve the narrow dynamic aperture have to-date not resulted in a higher luminosity. The addition of rotatable sextupoles, studied in simulations, is expected to improve the luminosity and will be tried in the coming months. Increasing the physics aperture is also being studied, requiring the installation of a new beam pipe and the relocation of several magnets. Simulations also indicate that the Touschek beam lifetime could be increased significantly. This would require modifications to the machine, so no results are expected to be available soon. A long-term plan is being developed to address deferred maintenance of accelerator components, especially for the linac and the transfer lines. The team is to be congratulated on vigorously pursuing the upgrade of the accelerator to reap the maximum physics benefit and is encouraged to continue to develop a coherent plan to improve the performance of the machine.

The Upgrade Working Group, which was formed in October 2018, gave an update on their activities. The current mandate of this working group is to coordinate activities in response to four considerations: an evaluation at what integrated luminosity a major detector upgrade would be required; a review of technical solutions for subdetectors that cannot withstand the expected background levels; the physics impact of luminosities five times the design luminosity and how the Belle II detector would perform at those luminosity levels; and an evaluation of options in response to potential forthcoming upgrades of the accelerator. The focus of the group is the timeframe of 2026, the current date for the major accelerator upgrade, when there will be a possible replacement of the QCS. This timeline does not leave a lot of room for extensive R&D for the intermediate upgrade. An upgrade after 2032 is also being considered for those components and systems which require extensive R&D and prototyping. The group has called for expressions of interest (EOIs) for the high-luminosity running. Those subsystems not planning to upgrade their detector had to submit a report demonstrating the robustness of their system for running at high luminosity. This process was concluded in March 2021. Ten EOIs were received, with two being able to complete before 2026 and two extending beyond the 2026 timeframe. The EOIs covered many subsystems, ranging from improved radiation monitoring with the diamond detectors, to several proposals addressing the vertexing and tracking system. EOIs were also submitted for the KLM, CDC, ECL and TOP detectors.

The writing of a white paper on the feasibility and physics case for running with a polarised electron beam has been delayed and is now expected to be submitted in the summer or fall of 2021. A polarised beam would bring unique capabilities to the experiment and allow for a precision measurement of $\sin^2 \vartheta_W$ and the electric dipole moment and anomalous magnetic moment of the τ-lepton. The plan is to produce white papers for Snowmass in early 2022 and conceptual design reports for an upgraded Belle II detector by the summer of 2022. An upgrade advisory committee has been created to review the proposals and direct the process. For this effort to be successful it will be critical to have reliable background extrapolations available.

The committee is pleased to see serious efforts exploring the physics case and technical feasibility to maximise the scientific return on the investment in SuperKEKB and Belle II. Many interesting physics channels, deviations from the Standard Model, and dark sector portals could be studied with an upgraded facility. Given the ongoing analyses of the current data, now is a good time to fully explore the accessibility and sensitivity to the broader physics opportunities of a luminosity upgrade, to quantify the scientific return and to define the detector specifications.

9.2 Concerns

- SuperKEKB is a complex machine and some of the initial attempts to improve the performance of the machine have not been as successful as expected. There are many options to improve the performance of the current machine; in addition there is a very large parameter space to be explored on what path to follow to upgrade the machine. This upgrade will have to be considered in the light of the financial and resource impact, the difficulty in its implementation, and the expected performance gain. The committee is concerned that at the moment, with the limited resources available, the process is relatively diffuse and will not converge on the time scale proposed.

- The collaboration is to be congratulated on initiating the EOI process to identify limitations of the current detector and explore upgrades. There is a serious concern, however, that the various boundary conditions, such as funding, timescale for accelerator upgrade, interface with the accelerator, are all not well defined. Furthermore, the scope of the detector upgrade seems rather large.

- The approach to the detector upgrade seems to be mainly driven by the interest of individual groups and upgrade options rather than physics performance. This may result in an imbalance in overall detector performance, where some subdetectors may be the limiting factor for reaching the physics goals, whereas the performance of other upgraded detectors may exceed the performance needs. Few examples were given motivated by a physics objective. Continuing that line of thought all the way to final physics results is encouraged.

- The presentations at the BPAC were very informative but the exact scope of the accelerator upgrade was not clear. Also, the process for reaching a decision was not discussed and there is some uncertainty when the accelerator upgrade will take place. It is noted that some accelerator parameters can have a significant influence on the proposed upgrades for the detector. This strong interplay between the machine upgrade and the detector upgrade and the current limited coordination is worrisome. For example, the upgrade of the IR would dictate aspects of the layout of the pixel detector and tracker.

9.3 Recommendations

- Address the deferred maintenance of the accelerator complex immediately and develop a long-term plan for replacement of obsolete or ageing components under the normal operating scenario of the machine to increase its reliability.
- Set up a regular meeting between the detector scientists and the accelerator physicists working on their respective upgrade, e.g. a few times per year, to inform each other and to ensure that possible incompatibilities between the upgrades are caught at an early stage.
- Develop a complete list of accelerator improvement projects, for the current machine as well as for its future upgrade, and evaluate the anticipated impact in terms of improvement in performance. Prioritise the task list according to impact and feasibility.
- Renew the effort to reach out to the international accelerator community to engage in the upgrade when the pandemic eases.
- Complete the review process of the EOIs as quickly as possible, identify the highest impact projects and develop those into conceptual designs that can lead to approval for construction. For each detector upgrade considered, the requirements on detector performance should be set in relation to well-defined physics objectives to ensure that the overall performance of the upgraded detectors is balanced to avoid a situation where one subdetector is overdesigned whereas another subdetector lags in performance to meet the desired physics goal.
- Decide on the decision-taking process for the different upgrade options, which should help stay on the aggressive schedule.
- Continue to evaluate and sharpen the physics case for the upgrade, including electron beam polarisation.