Draft version 1.0

ATLAS NOTE

February 4, 2015

Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014

D. Adams[1], C. Anastopoulos[2], A. Andreazza[3], M. Aoki[4], L. Asquith[5], M. Begel[1], F. Bernlochner[6], U. Blumenschein[7], A. Bocci[8], S. Cheatham[9], W. Davey[10], P.-A Delsart[11], P.-O. DeViveiros[12], A. Dewhurst[13], D. Duschinger[14], F. Filthaut[15], P. Francavilla[16], F. Garberson[17], S. Head[18], A. Henrichs[17], A. Hoecker[19], M. Kagan[20], B. Kersevan[21], T.J. Khoo[22], B. Lenzi[19], D. Lopez Mateos[23], B. Malaescu[16], Z. Marshall[23], T. Martin[24], C. Meyer[25], A. Morley[26], W. Murray[24], M. zur Nedden[27], R. Nicolaidou[28], S. Pagan Griso[23], G. Pasztor[29], P. Petroff[30], G. Piacquadio[20], C. Pizio[3], R. Polifka[31], X. Poveda[32], R. Reece[33], F. Ruehr[34], F. Salvatore[35], R. Sandstroem[36], T. Scanlon[37], D. Scheirich[38], S. Schramm[31], A. Schwartzman[20], K. Suruliz[2], M. Sutton[35], E. Thompson[40], M. Tripiana[41], A. Tuna[42], S. Viel[43], M. Vincter[29], I. Vivarelli[35], M. Wielers[13], A. Wildauer[36], T. Yamanaka[37], Z. Zinonos[7]

[1]BNL Brookhaven, [2]Sheffield, [3]Milano, [4]KEK, [5]Argonne, [6]Victoria, [7]Goettingen, [8]Duke, [9]McGill, [10]Bonn, [11]LPSC Grenoble, [12]Nikhef, [13]RAL, [14]Dresden, [15]Nijmegen, [17]Yale, [18]Birmingham, [19]CERN, [20]SLAC, [21]Ljubljana, [22]Cambridge, [23]Harvard, [16]LPNHE Paris, [23]LBNL Berkeley, [24]Warwick, [25]Chicago, [26]Stockholm KTH, [27]Berlin HU, [28]CEA Saclay, [29]Carleton, [30]LAL Orsay, [31]Toronto, [32]Indiana, [33]UC Santa Cruz, [34]Freiburg, [35]Sussex, [36]MPI Munich, [37]UC London, [38]CU Prague, [40]Columbia, [41]Barcelona, [42]Pennsylvania, [43]UBC Vancouver, [37]Tokyo

ATL-PHYS-INT-2014-018
04/10/2016

Abstract

Between autumn 2013 and spring 2014 four study groups consisting of experts from the combined performance and physics groups evaluated the potential for physics objects and analysis harmonisation in preparation of the LHC Run-2. The areas covered were tracking and vertexing, particle identification and isolation, jets and missing transverse momentum, as well as overlap removal. Flavour tagging will be covered in a later update of this ongoing and evolving effort. This document summarises the main findings and formulates recommendations for future analyses.

© Copyright 2015 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-3.0 license.

Contents

1 Introduction

This document presents the results of four study groups formed in Nov 2013 with the aim to harmonise the use of combined performance (CP) objects and anlaysis requirements between the physics groups for the LHC Run-2. Harmonisation (or standardisation and diversity reduction) has three central purposes:

1. Reduction of disk and CPU usage;
2. Reduction in the need to understand and calibrate objects;
3. Harmonisation of the analysis flow to ease comparison, combination, review, transparency between physics group, etc.

A first exploratory harmonisation workshop held in Sep 2013 [1] has revealed that part of the diversity among the ATLAS analyses is driven by lack of coordination rather than physics reasons. This is where harmonisation can and should occur. Harmonised analysis will benefit from the more centralised Run-2 analysis model:

- Basic and popular physics objects are computed during the reconstruction and saved as xAOD;
- Refined/customised objects can be built from the xAOD during the derivation step;
- CP object calibration and uncertainties are standardised in the analysis framework and can be applied at any stage;
- xAOD → xAOD reprocessing allows to update stored CP objects.

During a dedicated Physics Coordination (PC) meeting in Nov 2013 the harmonisation effort was subdivided into five study groups:

Group 1: Tracking and Vertexing (led by Simone Pagan Griso and Anthony Morley)

Group 2: Particle identification and isolation (led by Manuella Vincter)

Group 3: Jet and missing E_{T} (led by Michael Begel and David Lopez Mateos)

Group 4: Flavour tagging (led by Frank Filthaut and Tim Scanlon)

Group 5: Overlap removal (led by Ulrike Blumenschein)

alongside which this document is organised. Since the tools for flavour tagging were in the midst of significant developments, the work of Group 4 was postponed to May 2014. The results will be included in an updated edition of this document. The physics and CP groups appointed contacts for these study groups [2] and dedicated group meetings between Jan and May 2014 led to the present harmonisation recommendations document. The focus of this first edition is to outline what can be harmonised for *most* analyses assuming the current state-of-the-art CP objects and analysis tools. It is to formulate a basic template and roadmap anticipating future document updates.

The research and development in the CP and detetor reconstruction groups does not stop after this harmonisation exercise. However, any new physics object or procedure should be discussed in the context of harmonisation before it is introduced for physics use. Harmonisation is a continuing intellectual and technical effort that any large collaboration needs to invest in if it wants to work efficiently.

Finally, we emphasise that analysis harmonisation goes beyond objects. It includes data-driven background and calibration methods, experimental systematic uncertainties, theoretical uncertainties, generators, PDFs, etc. Although these topics are beyond the scope of this document, they must not be forgotten.

2 Harmonisation Study Group 1: Tracking and Vertexing

The main harmonisation topics to be addressed by the tracking and vertex reconstruction harmonisation group can be classified in three areas.

- Track quality selections and default parameters:
  - Definition of quality requirements for reconstructed tracks;
  - Impact parameter reference frame choice;
  - Impact parameter selection for prompt electrons and muons.
- Primary vertex:
  - Quality requirements for hard-scattering and pileup primary vertices;
  - Selection of hard-scattering primary vertex.
- Track-vertex association:
  - Selections for classifying tracks compatible with primary vertices.

The Inner Detector track reconstruction was optimised during Run-1 to provide the best possible reconstruction efficiency for charged particles originating from the LHC beam interaction point and reasonable efficiency from secondary interactions. A variety of track reconstruction setups were used during Run-1 but for the vast majority of the proton–proton collisions tracks with transverse momentum (p_{T}) greater than 400 MeV[1] and 7 hits in the silicon tracker were attempted to be reconstructed by the primary track reconstruction algorithm.

The physics requirements for Run-2 are very similar to Run-1 and therefore the track reconstruction setup will be very similar. However the addition of the IBL [3] and expected higher luminosity of the LHC will require changes.

2.1 Track selection and track parameter usage

2.1.1 Track selection

During Run-1 two main working points were defined [4] for tracks that originating from the proton-proton collision point. The first is tuned to provide a very good efficiency at the price of some fake tracks being present. They are applied during reconstruction and all tracks produced by the primary track reconstruction algorithm satisfy these criteria. We will refer to this baseline set of requirements as "loose". The fake tracks that are especially prevalent at larger $|\eta|$ and can be suppressed by further quality requirements or track-object matching (object being a muon segment, a calorimeter object, a primary vertex, etc..).

A more robust set of requirements, tightening the number of silicon hits requirement to 9 and requiring no pixel holes, was developed to reduce the fake tracks contribution to a negligible level at the cost of a drop in tracking efficiency by a few percent. This selection will be referred to as "tight". The tight selection has been used for a variety of applications, including primary vertex reconstruction.

It has been found that several different track selections have been used in physics analysis and CP object reconstruction [5]. Such selections are sometimes motivated but often just a custom derivation of an early reference such as the minimum-bias paper [6], and are not optimal for the higher pileup environment experienced in the bulk of Run-1 data or expected for Run-2. The majority of these custom

[1] The minimum p_{T} threshold was defined to allow for the reconstruction of all charged particles that originate from proton–proton collision point and reach the entrance of the barrel calorimeter

derivations was geared towards a class of tracks not covered by the two above working points, that is, tracks with well measured impact parameters. This is something we hope to address in these recommendations.

The development of the tracking setup for Run-2 is still in progress and many improvements in the core tracking software are implemented to achieve a better tracking performance under high pileup conditions. It is therefore too early to provide final recommendations on track quality selections. However, **we expect to provide for Run-2 three sets of track quality selection recommendations**:

- **"Loose"**. Analogously to the Run-1 working point, this will correspond to all the tracks saved as output by the reconstruction and will focus on obtaining highly efficient charged particle reconstruction at the cost of a non-negligible rate of fake tracks. The target is to have greater than 99% efficiency for all track with transverse momentum greater than 1 GeVand leave 8 silicon hits in the detector.

- **"Loose-primary"**. A track selection targeting to improve the track quality, especially impact parameter measurements, but still be as efficient as possible for particles produced before the first measurement plane. A reduction in fake track rate with respect to the loose selection is expected but is not the focus of this track selection class. It is expected to represent a good baseline for selecting tracks originating from b-hadron or tau decays.

- **"Tight-primary"**. Analogously to the Run-1 working point, this extra quality selection will be optimised to reduce fake tracks to a negligible level (below 1% at $\mu = 41$) at the price of a reduced tracking efficiency. The selection will be optimised on primary charged particles and may not be ideal for particles produced by decays significantly displaced from the proton–proton interaction point (radii greater than the first measurement plane).

Note that no explicit recommendation will be made for quality requirements of tracks originating from long-lived particle decays. As the use cases are limited and quite varied depending on physics of the process analysis specific optimisation will be required.

It is expected that for both reconstruction and analysis one of the three above track classes can be used. Additional selection criteria will be used to determine the origin of a track as discussed in Sections 2.1.2 and 2.3.

2.1.2 Default track parameter representation

The default track parameter representation in Run-1 has been the perigee representation with respect to the nominal ATLAS centre (0,0,0). Although this is a convenient representation in terms of consistency it is not the most useful one for physics analysis as it is not possible (without extrapolating the track to its true origin) to determine the correct kinematics of the charged particle and to check the track's compatibility with the primary vertices in the event.

For Run-2 we recommend to change the default representation **to express track parameters with respect to the beamline position**. The beamline position is determined by the beamspot fit during the calibration loop and accessed via the *InDetBeamSpotService* in the ATLAS software. We encourage to adopt a similar choice at trigger level, using the available online beamspot position returned by the same ATLAS software service. The coherent usage of the beamspot service will allow also an harmonize treatment of special cases (e.g. beamspot not available). Such an approach is motivated by the transverse size of the beamspot which is expected to be below 10 μm in Run-2, well below the reconstructed hard-scattering vertex resolution for most physics processes [2]. Such a representation will allow for an unbiased

[2] The transverse vertex resolution in the best-case scenario ($t\bar{t}$ signal events) is expected to just approach 10 μm. In an average minimum bias event the vertex resolution 60-80 μm [7]

transverse impact parameter measurement without the need to update the primary vertex position, and for a longitudinal impact parameter estimation with respect to any vertex by a simple calculation, $|z_0^{\mathrm{trk}} - z_0^{\mathrm{vtx}}|$, taking into account the small beamline tilt with respect to the ATLAS detector.

2.1.3 Impact parameter usage

The default track parameter representation, as described in the previous section, will be defined with respect to the beamline and as such all impact parameters related selections should be made using this representation for Run-2.

Cuts on impact parameters are applied for three main purposes:

- To select tracks compatible with a primary vertex;
- To separate prompt and non-prompt tracks;
- To reject cosmics.

The selection of tracks compatible with a particular primary vertex will be discussed in Section 2.3. Rejecting cosmic tracks via a cut on the impact parameter alone is **not recommended**. We **recommend** that the Muon CP group provides guidelines on standard ways for analyses to evaluate the contamination from cosmic rays in analyses.

The efficient separation of non-prompt from prompt tracks is, in general, very process dependent with the optimal cut value usually being determined by the lifetime and the decay product kinematics. No explicit recommendation will be provided for this case. In the particular case of the reconstruction of photon conversion, however, a centrally-provided algorithm is run on a pre-optimized set of tracks and maintained by the tracking and e/γ CP groups.

The removal of non prompt tracks from a prompt signal is instead a very common task and was usually done in Run-1 analyses, most frequently applied to muon and electrons, with the intention of identifying tracks that originate from a particular primary vertex to reduce contamination of leptons originating from long-lived particle decays. A variety of selections was used mostly targeting very efficient reconstruction of a prompt signal. Differences in the selections were small and can easily be harmonised for most analyses.

A typical selection for muons (electrons) consisted in requiring a transverse impact parameter significance $|d_0|/\sigma(d_0) < 3(5)$ and a transverse impact parameter $|z_0 \cdot \sin(\theta)| < 1.5$ mm [5]. For both requirements the reference is usually chosen to be the primary vertex.

The usage of the d_0 significance **is recommended** over a simple d_0 selection, as it takes into account measurement uncertainty and other p_{T} and η dependent effects. The longitudinal impact parameter is less accurate than the transverse one; for this reason it offers poorer rejection for secondaries and a much looser selection is often applied. The introduction of the IBL will allow for a more accurate measurement of the longitudinal impact parameter, but it is not expected that it will play a significant role beyond associating tracks to a primary vertex (see Section 2.3).

In Run-2 we expect to **keep the recommendations** based on the same impact parameter definitions (but with respect to the beamline, adapting the longitudinal requirement to be the difference with respect to the z_0 position of the primary vertex of reference). Example selections for prompt muons and electrons from Z boson decays for a highly efficient working point (e.g. > 99% efficient) should be provided to act as a guide on how to use the variables.

2.2 Primary vertex

The primary vertex reconstruction runs on a subset of the reconstructed tracks corresponding to the "tight-primary" selection discussed in Section 2.1. This track preselection has been aimed at reducing contributions from fake vertices during Run-1, and we do expect it to also be used in Run-2 (with the corresponding redefinition of “tight-primary”). Each vertex is reconstructed with a minimum of two tracks and using the beamspot constraint. Although it would in principle be possible to require only a single track together with the beamspot constraint, this has not been used so far due to an increased fake and split vertex rate. After reconstructing the primary vertices, the hard-scattering vertex is identified and flagged. The algorithm used in Run-1 for this purpose has been the sum of squared transverse momenta of tracks associated to the vertex. Such a simple selection has provided satisfying performance, and the lack of further studies does not allow us to foresee any alternative to it for Run-2.

Various quality selections have been applied during Run-1 to the hard-scattering primary vertex. Most of them require a minimum number of associated tracks ranging from 2 (equivalent to no additional quality requirement except that of the vertex’ existence) to 5 [5]. Such a requirement is useful in a low pileup environment with the need to ensure that a hard proton–proton interaction has occurred in a given bunch crossing (such as to suppress beam backgrounds). The efficiency of such a requirement for signal events is usually larger than 99%, but does depend on the physics process of interest. For the 2012 data and with the pileup conditions expected for Run-2, however, the probability of no proton–proton interaction during a bunch crossing is very small, rendering such a selection unnecessary; still it is good practice to ensure a vertex has been reconstructed in order to have well defined (longitudinal) impact parameters with respect to it. We therefore **recommend** to require that at least one vertex is reconstructed, without further quality requirements, in Run-2 analyses.

Reconstructed pileup vertices can be used as an event-by-event parametrisation of the number of in-time pileup interactions. For the best correlation of such quantities, we **recommend** to count pileup primary vertices without further quality selections since their reconstruction is already optimised for a low fake and split vertex rate.

A target benchmark performance for the vertex reconstruction will be to ensure that the split and fake vertex rate at mu=40 to be below 1% and vertex reconstruction efficiency to be above 99% for Z→ $\mu\mu$ events where the muons are within the ID acceptance. These requirements should be met while while trying to minimise contamination from other vertices [8].

2.3 Track-vertex association

A growing number of applications has recently focused on selecting tracks compatible with a given reconstructed interaction vertex, and removing tracks that are not. Such applications range from jet-vertex-association to track-based missing transverse momentum reconstruction to particle-flow, etc. The selections vary considerably and often without a clear benefit over a more standard choice.

A more uniform approach is thus desirable for Run-2, also implemented in the form of a dual-use tool [9], to be fully part of the reconstruction and analysis work flow. Two main use cases may be considered:

- “**Loose-matching**”, in the form of a very efficient algorithm to ask if a given track is compatible with a given interaction;
- “**Tight-matching**”, when pileup rejection becomes a limiting factor, a more aggressive approach could be desirable.

Building on Run-1 experience, the loose-matching recommendation will be based on a simple impact parameter selection, similar to that described for leptons in Section 2.1.2, as for example a $|z_0 - z_0^{\mathrm{vtx}}| <$ 1.5 mm selection.

For the tight-matching, the recommendation will employ association based on the compatibility of the tracks to vertices, stored as “weight” of the track with respect to the vertex the track was used for in the primary vertex reconstruction. The primary vertex reconstruction always ensures one-to-one association. A dedicated requirement for non-associated tracks (e.g. from heavy-flavour decays) is also expected, similarly to the recent studies performed for jet vertex association in Ref. [10].

Until it has been established that the detector performance is well described by simulation, caution should be taken when applying the tight-matching. Once that condition is met, the tight-matching requirement is to be preferred.

2.4 Track collection disk usage

The track collection is one of the largest contributors to the size of the stored xAOD reconstruction output. At high pileup the size of the track collection grows to an unacceptable level. Although some increase is inevitable as the collection grows approximately linearly with pileup, we outline in the following possible ways to reduce the size of the output track collection.

2.4.1 TrackParticle xAOD content

For each track particle, the following substantial information is stored in the xAOD:

- Track parameters and covariance matrix at the beamline (30 floats);
- Position at which the track parameters are expressed (3 floats);
- Track parameters and covariance matrix at the first and last measurement (62 floats + 2 8-bit ints);
- Track fit quality (2 floats);
- Reconstruction information (3 8-bit ints and 1 long int);
- Detailed information about the number and type of hits in the three sub-detectors (32 8-bit ints);
- Particle identification information (4 floats);
- ElementLinks to associated vertex and Trk::Track (4 ints);
- ElementLinks to truth associated particle (2 ints);
- Truth matching information from MCTruthClassifier (2 ints).

The two largest items on the list are the track parameters and their covariance matrices followed by the detailed hit level information. At present three sets of track parameters (including full covariance matrices) are stored for every track. The first is the most widely used, the track parameters at the beamline, and we would be reluctant to remove this information for any track. The second set is located at first measurement. These track parameters are predominately used for secondary vertex reconstruction where the knowledge of the location of the first measurement helps to reduce fake vertices. The final set of track parameters is located at the last measurement. Although this information is required at reconstruction level, it is not overly used during final analysis, expect for electron tracks where it is used to determine the energy loss along the track and thus could be removed for all non-electron tracks with little to no impact on physics analyses.

The detailed hit information is already quite compressed but could be removed if the results of predefined track selections (for example those defined in Section 2.1) were calculated at the time of writing.

Our **recommendation** is to remove the last measurement track parameters from all tracks expect those associated to electrons. And if necessary, for non-vital tracks, remove the first measurement and condense the hit level information by storing the results of the track passing a predefined cuts.

2.4.2 Track selection for slimmed tracks

It would also be possible to reduce the disk size of the track collection by only writing the most used information for the least used tracks in the event (previous referred to as non-vital tracks).

A combination of the following selections could be used to select tracks for which detailed information is written while for the remainder less information is saved:

- Tracks originating from/consistent with the first "n" primary vertices. This will reduce the pileup dependence of the track collection size and is relatively simple to implement. However in the case of an event topology where the Σp_{T}^2 vertex is often not the signal interaction, it may lead to significant performance degradation.

- Tracks associated to leptons. Tracks associated leptons are often the most scrutinised and it is therefore useful to have as much information as possible about them, while reducing the information for the others.

- High p_{T} tracks. As the p_{T} of tracks is steeply falling, significant gain can be made by increasing the p_{T} threshold for the tracks that are written to disk. For example in a $\mu = 41$ $t\bar{t}$ sample, 20% of the silicon track collection have $400 < p_{\mathrm{T}} < 500$ MeV, 50% have $400 < p_{\mathrm{T}} < 600$ MeV, and 70% have $400 < p_{\mathrm{T}} < 800$ MeV. The impact of such loss of tracks has not being studied yet.

Both options, track selection and information pruning, need to be further discussed but it should be assured that the information that is removed from the tracks does not limit the analysers ability to recalculate other variables at the xAOD level. As part of this discussion a full list of the variables and quantities that use tracks and can be recalculated at xAOD level needs to be compiled, and the performance benefit of their recalculation needs to be evaluated.

2.4.3 Number of tracks in the collection

One of the most obvious ways to reduce the track-related disk usage is to reduce the number of tracks in the collection. However, it is difficult to do so without impinging physics performance, in particular: flavour-tagging, electron and photon reconstruction, tau reconstruction, muon reconstruction. It goes without saying that detailed studies will be needed if one wants to understand the performance implications of any significant reduction of the track collection.

It should also be noted that to ensure a fully consistent reconstructed event, it is not desirable to remove tracks after the reconstruction stage. We thus **recommend** that any reduction of the track collection is applied at the reconstruction level.

The following approaches could be considered:

- Tightening the quality criteria of silicon seeded tracks. The silicon seeded track reconstruction is currently optimised for efficiency with relatively loose basic track quality requirements (7 Silicon Hits and a $\chi^2/\mathrm{nDoF} < 25$) and considerable fake contamination at high pileup (roughly 25% at $\mu = 41$)[4]. Tightening these requirements could have an impact on efficiency-sensitive object reconstruction (e.g. leptons) that needs to be studied.

- Improve TRT-only tracking. TRT-only tracks make up one third of the track collection at $\mu = 41$ [11]. The vast majority of these tracks consists of combinatorial fakes and without of the use of additional information (e.g. TRT high-threshold information) they cannot be used for physics analysis with any confidence. Currently these tracks are only used in the reconstruction of photon conversion vertices. An effort is underway to reduce the fake contamination of TRT-only tracks. An electromagnetic cluster seeded version of the algorithm has already been implemented.

- Narrowing the parameter space search window. Currently the track reconstruction will search for tracks within a relatively large window of phase space ($p_T > 400$ GeV, $d_0 < 10$ mm, $z_0 < 250$ mm). For example, reducing the transverse impact parameter down to 3 mm would reduce the collection size of silicon tracks by one third at $\mu = 41$. Tracks with high transverse impact parameter normally originate from hadronic interactions, photon conversions or long-lived particle decay, or are combinatorial fake tracks. Therefore reducing the IP cut could have an impact on the performance of photon conversion reconstruction, flavour-tagging and tau reconstruction.

Improving the track reconstruction setup (as described in the first two item) is an ongoing effort not only for disk space issues, but for CPU performance and to have better physics performance overall. Reducing the parameter space search window is clearly going to have a negative impact on the physics performance and is something we would only consider as a last resort solution.

3 Harmonisation Study Group 2: Particle Identification and Isolation

3.1 Information collection

A questionnaire related to PID and isolation usage for electrons, muons, taus, and photons was sent to the Group 2 analysis contacts on Jan 24/14, to forward to the physics groups. Some feedback (found at this link and this link) was received related to Run-1 experience as well as desired features for Run-2. Based on this feedback, five meetings were held to discuss possible paths for PID and isolation harmonisation for Run-2:

- Fri Feb 7/14: isolation with cone variables (→ meeting link)
- Wed Feb 19/14: PID methodology and working points for Run-2 (→ meeting)
- Fri Mar 7/14: isolation developments for Run-2 (→ meeting)
- Tues Mar 18/14: trigger considerations for harmonisation (→ meeting)
- Thurs April 10/14: CP impact of harmonisation (→ meeting)

This section summarises some recommendations on the harmonisation of PID and isolation based on these discussions. These recommendations do not extend to e.g. exotic long-lived particles, which have their own particular PID and isolation issues.

One of the goals of the harmonisation exercise is to reduce the usage of disk space and limit CPU consumption. The reduction of the numbers of supported PID and isolation flavours and working points will naturally do this, but is unlikely to have a huge impact on e.g. the size of the future xAODs. The real benefit of the recommendations of this harmonisation will be the creation of centrally supported variables along with associated data-to-MC scale factors and uncertainties which will enable consistency across analyses as well as simplify the peer-reviewing of papers and CONF notes. The proposals made here are likely to significantly impact the amount of work required by the CP groups. However, if viewed from the total amount of effort that was required in Run-1 by the individual analyses to provide their own recommendations, centralisation should result in a net reduction of overall ATLAS work load. The physics groups should be encouraged to provide the labour required by the CP groups to make this harmonisation successful. Without this extra input of human resources, the proposals made in this section will not be possible.

3.2 Harmonisation of PID

The harmonisation of PID for Run-2 in the Egamma, Muon, and Tau CP communities is already well in hand. There is a solid plan by the CP communities to limit the number of PID methodologies to a single type per particle and limit the proliferation of working points that resulted from Run-1 to approximately three per particle type. As well, work is ongoing to keep the online and offline PID as similar as possible, though some inevitable differences are foreseen, as discussed in Section 3.4. As we continue to take data in 2015 and improve offline PID, it will be important to keep the offline and online PID selections as close as possible to avoid the proliferation of working points such as PID to be used for physics analyses vs. PID to be used for the trigger.

In what follows, a *working point* is considered to be a set of selection criteria, optimised by a CP group, to identify a lepton or a photon. Examples are *tight* electrons or *medium* taus. A *supported working point* is one where, in addition, the CP groups provide the appropriate data-to-MC scale factors and associated uncertainties. In what follows, not all working points are proposed to be supported working points. Ideally, there would be approximately two supported working points per particle type (the exact number must be driven by the physics needs of analyses, but an example could be a tighter PID for analyses requiring a single good quality lepton/photon and a looser PID for analyses with multiple leptons/photons) and an additional much looser working point for background evaluations (this latter PID would not necessarily be a supported working point as is already the case for example for *loose* photons.).

3.2.1 Muon PID

The Run-1 muon PID recommendations consisted of STACO and MUID, with the later development of the so-call muon third chain MUONS which was designed to provide the best of the two chains. The third chain was not used extensively in Run-1 analyses and so the experience with it is still limited. There are a few Run-1 analyses planning to use the third chain for their final 2012 publications. *It will be important to gather their experience with MUONS since the third chain is the only planned muon identification method for Run-2.* At the moment, there are no plans to move away from a cut-based PID towards MVA-type selections.

Recommended path for Run-2: Only the third chain will be supported. The plan is to have three supported working points (equivalent to *medium* and *tight* and an additional working point at very high p_{T} for analyses like those used in exotics which require good-resolution muons). *A few representative Run-1 analyses using STACO/MUID should also perform their analyses with MUONS and benchmark the performance, to ensure that the third chain can be safely used for Run-2.*

3.2.2 Tau PID

Both likelihood (LLH) and boosted decision tree (BDT) and well as cut-based PID, each with several working points, were developed and used extensively for Run-1. However, the final recommendations consist of only BDT for jet and electron rejection and a cut-based PID for the muon veto.

The recent offline developments are focused on using substructure techniques to better classify tau decays with the goal to improve energy resolution. Real improvements seem possible; however, some limitations may have been reached in terms of the improvement of fake rates. The many developments ongoing for Run-2 are not really subject to harmonisation as these will be accomplished by the Tau CP group and should not lead to an increase in diversification. The baseline technique for tau identification in Run-2 will likely be very similar to the final Run-1 version. In addition to the traditional fixed working points, early Run-2 might see in parallel a possibility for a continuous PID option which could well become the default some time into Run-2.

Recommended path for Run-2: as for the final Run-1 recommendations, entirely BDT for electron and jet rejection with a cut-based muon veto with the three traditional supported working points (*loose*, *medium*, *tight*). In practice, a looser working point might also be needed in order to develop the vetos. The tau trigger will also undergo significant changes for Run-2, bringing improved performance and better harmonisation with the offline (see Section 3.4.2).

3.2.3 Electron PID

Run-1 saw two electron PID methodologies, cut-based and likelihood-based, for central electrons and a cut-based, superceded in 2012 by an MVA-type selection based on Fisher discriminants, in the forward region. There also was an almost unmanageable proliferation of working points as shown in Ref. [12] (central: *relaxedLLH*, *verylooseLLH*, *loose*, *looseLLH*, *multilepton*, *medium*, *mediumLLH*, *tight*, *verytightLLH*, most supported with official recommendations for scale factors, and forward: *looseFWD*, *mediumFWD*, *tightFWD*).

The Egamma CP community has a clear path towards Run-2, which is to make the LLH-based PID the only supported methodology since, for a given efficiency working point, the LLH PID has a factor of two better background rejection than the cut-based PID. The Run-1 experience with LLH electrons, as reported in the questionnaires, was mostly limited to Higgs analyses. *More experience is needed, in particular with precision analyses and analyses at high electron E_T, to make a definitive statement on the appropriateness of LLH electrons for PID. It is therefore important that some of the cut-based Run-1 analyses systematically benchmark the LLH alternative.* In addition, the 2012 data-to-MC scale factors for both cut-based and LLH have not been established to the same precision as the cut-based ones with the 2011 data. *It is important that the final 2012 data-to-MC PID scale factors soon be made available for both cut-based and LLH PIDs, to demonstrate that these two can be determined at the same precision, and then be subsequently used in final Run-1 analyses requiring the extra precision.*

A detailed plan for the transition in 2015 from cut-based to LLH electron PID will be worked out at the Egamma CP workshop in the fall of 2014. *At this time, it should also be discussed how in later years LLH electrons can be commissioned at the start of new LHC running-periods without the benefit of developing an optimised cut-based menu targeting the specific running conditions for that year (e.g. increased pileup or higher beam energy).* The ability to make such cut-based menus will still be available as the isEM bits needed to generate them will continue to exist. However, the Egamma CP group does not have the resources to build two commissioning menus every year as is being done for 2015.

The decision in Run-1 to develop and implement LLH electrons was to a large extent based on the interests of a small group of people to develop this particular methodology. It is possible that other MVA-type selections could be developed in Run-2. However, until such new PIDs exist, are benchmarked, and demonstrate an improvement over LLH, they are not part of the harmonisation discussion.

With the move of electron PID to LLH in Run-2, the idea of "continuous PID" (as is being discussed in the Tau CP community) was raised. It is a possibility for early Run-2, but has not yet been actively investigated if it would be more useful for the physics communities to have such a PID and whether or not the CP community could support these continuous working points with associated scale factors and uncertainties. A possible use-case for continuous PID could be to enable analyses to get, for untriggered objects, arbitrarily loose electron samples for background studies. In any case, even if developed, it is not proposed at this time that such a continuous PID be supported with data-to-MC scale factors and uncertainties.

Recommended path for Run-2: have 3 supported working points (*loose*, *medium*, *tight*) plus perhaps one *veryloose* (which could be useful for background studies but may not need to be supported with scale factors) for both cut-based and LLH at the start of Run-2, then the cut-based option is dropped, once LLH is well understood and commissioned with early data. The trigger would follow the same path, in order to keep the definition of electrons at trigger and offline as close as possible. The important differences between online and offline electrons are addressed in Section 3.4.3. The LLH electrons should be reoptimised for Run-2 so that the three working points (*loose*, *medium*, *tight*) are subsets of each other, which was not strictly the case for Run-1. The LLH variables that are responsible for efficiency loss at high E_T in Run-1 could be loosened or dropped so as to provide a flatter efficiency and smaller systematic error at high E_T, particularly useful for exotics analyses. *A few representative Run-1 analyses using cut-based PID should also perform their analyses with LLH and benchmark the performance, to ensure that LLH provides comparable performance.*

3.2.4 Photon PID

Run-1 analyses use almost exclusively the cut-based photon PID, for which *loose* and *tight* working points exist, though only *tight* is currently fully supported with scale factors and uncertainties. A neural-net-based (NN) selection also exists and was used with the 2011 data set, but not in 2012 since no gain beyond the cut-based photon PID was observed. It was found that larger uncertainties were associated to the measurement of efficiencies with NN. *This issue would need to be further investigated, if NN or other MVA-type PID were planned to be used in Run-2.* In the questionnaire, the question was asked if the two current working points are sufficient for analyses. Arguments were made that a *medium* working point might be useful since this is used at the trigger level. As well, several analyses (missing transverse momentum γ+jet JES, SUSY di-photon, and SM Z+γ) indicated that such a working point would be helpful. At least two groups (SUSY single-photon and SM Z+γ) also stated a wish for a *verytight* working point. Even if new working points were added of the photon PID menu, not all would necessarily need to be supported with scale factors and uncertainties, as is currently the case. The NN (or perhaps some other MVA-type PID) will continue to be developed for Run-2 but it is not foreseen to be the default method in time for, nor even at the beginning of, Run-2.

Recommended path for Run-2: Three cut-based working points could exist (*loose*, *medium*, *tight*, perhaps reoptimised to provide the equivalent of a *verytight* working point). Only *medium* and *tight* need be supported working points. The photon trigger will remain cut-based as well, following a similar path to the offline photons. The differences between online and offline photons are addressed in Section 3.4.3.

3.2.5 Other points

Some analyses apply impact-parameter requirements on tracks (d_0 or z_0) on top of the normal PID recommendations. There are two primary reasons for this: 1) cosmics veto and 2) reducing background e.g. from heavy flavour. The question was raised as to whether there could there be harmonised recommendations for electrons and muons in both cases. This aspect is discussed in Section 2.1.

The question was raised as to how to consistently recalculate the $E_{\mathrm{T}}^{\mathrm{miss}}$ quantity when changing the definition of an electron (from e.g. cut-based to LLH). A tool for this is currently being developed by the jet-$E_{\mathrm{T}}^{\mathrm{miss}}$ group as discussed in Section 4.5.

3.3 Harmonisation of isolation

No group within ATLAS has truly claimed ownership and support for isolation variables. For Run-1, analysers are expected to decide on their own how to use and optimise isolation for their specific analyses, including the derivation of data-to-MC correction scale factors and associated uncertainties. This has lead to a variety of isolation variables, cone sizes, and selection requirements in analyses, which has made it difficult to review these within ATLAS and apply some uniformity in publications across similar analyses. It is a **recommendation** that the CP groups take ownership of isolation and provide recommendations for variables, isolation working points, and support a limited set of centralised scale factors and uncertainties, as is currently done for PID. Clearly, all of these possibilities presented below need significant *manuspower* from all physics and CP communities to be implemented. However, a centralised isolation would likely require much less effort to create and maintain than what was used by the many analyses in Run-1 to provide their own recommendations.

As described in more detail below, **the harmonisation proposal comes at several levels:**

1. The development of a harmonised tool/interface for accessing all track and calorimeter isolation variables, regardless of particle type. This technical proposal would reduce the mistakes that users can make when selecting the isolation variables and applying the various isolation corrections for e.g. leakage and pileup.

2. This tool would only support the isolation variables recommended by the CP groups, in consultation with the physics groups. As mentioned below, this implies one variable type per track/calorimeter isolation and would ideally limit the number of cone sizes per variable type to two (one for analysis and an alternate for systematics). This would limit the proliferation of additional isolation

variables for use in analyses. However, the common interface would easily enable new variables to supplant old ones, should continued R&D in these areas come up with better solutions than the ones proposed below.

3. This tool would make use of centrally-derived "lookup tables" from specific physics benchmarks (such as $Z \rightarrow ee, \mu\mu$ tag and probe). These tables would be used to determine the cut value for a given isolation efficiency. Such centralised tables imply that a common definition of efficiency can be used across all physics groups using similar benchmark physics process.

4. Though this tool could trivially return the cut values needed for an extensive range of isolation efficiencies e.g. as a function of p_{T} (see below), the CP groups would only provide the data-to-MC scale factors and associated uncertainies for very few isolation working points, ideally two such as a *loose* and a *tight* isolation. The full support for such a limited set of working points would encourage most users to use just these working points in their analyses. However, the flexibility of the tool to provide a continuous range of unsupported ones would still enable analyses in particular topologies to use the tool (but then they have to derive their own scale factors).

Much information on the current usage of track and calorimeter isolation was gathered through the questionnaire. Some additional notes on isolation usage in ATLAS are appended to the agenda of the Feb 7/14 meeting [13], since most of the current documentation on isolation in CDS is outdated. In general, analyses isolate muons with a $p_{\mathrm{T}}^{\mathrm{cone}}$ variable and apply a fixed percent cut on the requirement. Some analyses use $p_{\mathrm{T}}^{\mathrm{cone}}$ as part of an electron isolation strategy but this is often accompanied by a calorimeter isolation. The use of $E_{\mathrm{T}}^{\mathrm{cone}}$-type variables is usually associated with electron/photons but sometimes also used with muons. The usage of $E_{\mathrm{T}}^{\mathrm{cone}}$-type variables in analyses is less uniform in terms of requirements (fixed GeV cut, fixed percent cut, and mixture of both are used, the latter especially in exotics analyses due to efficiency loss and background issues at high and loss E_{T}, respectively). As well, a wide variety of cone sizes was reported for both track and calorimeter isolation. Some of this large selection of choices seemed to be largely historical in nature and a thorough comparison within specific analyses would have to be made to be able decide if fewer cone sizes (e.g. two for $p_{\mathrm{T}}^{\mathrm{cone}}$ and two for $E_{\mathrm{T}}^{\mathrm{cone}}$-type) would be sufficient. *Some representative Run-1 analyses that use isolation should also do a systematic evaluation of the impact of using alternate isolation definitions/cone sizes on their results, including systematics, in view of reducing the supported number of cone sizes to no more than two for each of calorimeter and track-based isolation (one as a primary variable and an alternate for systematic studies).* This does not need to be done for all Run-1 analyses, but rather a few representative analyses with specific kinematic features (e.g. high and low E_{T}) and environments (e.g. busy vs. quiet).

The following subsections describe the more commonly used isolation types in ATLAS and provide some harmonisation recommendations.

3.3.1 Track isolation

The default track isolation variable for Run-2 is likely to remain $p_{\mathrm{T}}^{\mathrm{cone}}$. This variable will likely not be impacted much by potential developments for Run-2 related to track and quality definition and track-vertex association, as discussed in Section 2. There will be minimal track quality selection changes due to the insertion of the IBL. The minimum p_{T} threshold for $p_{\mathrm{T}}^{\mathrm{cone}}$ is expected to be 400 MeV, unless hard-pressed by CPU constraint. Some ideas are being developed to try to harmonise the way tracks are associated to vertices and there are some investigations related to online and offline with respect to how the track is extrapolated back to the choice of primary vertex since it might not find the same primary vertex. All of these are likely to impact $p_{\mathrm{T}}^{\mathrm{cone}}$, but not significantly. The calculation of $p_{\mathrm{T}}^{\mathrm{cone}}$ could still be refined. For example, there may be some residual issues related to a mismatch in how the p_{T} of electrons (other than the primary electron candidate) within the isolation cone is calculated (using the pion hypothesis or using the GSF). Where ever possible, a common definition for track, quality definition, and track-vertex association should be used for both online and offline.

There will be some standard default $p_{\mathrm{T}}^{\mathrm{cone}}$ (with a yet-to-be-determined number of cone sizes) in the xAODs and there will be standard software available if an analyser wishes to rebuild a $p_{\mathrm{T}}^{\mathrm{cone}}$ variable for specific applications in analyses. The information required to do so should be in the xAODs, unless these become just too large. It is important that this possibility to rebuild $p_{\mathrm{T}}^{\mathrm{cone}}$ not be limited to the default hard-interaction primary vertex.

3.3.2 Calorimeter isolation

Two types of calorimeter isolation variables were extensively used in Run-1: $E_{\mathrm{T}}^{\mathrm{cone}}$ corrected for leakage and pileup (PU) (the latter using the number of primary vertices (NPV) for the correction at 7 and 8 TeV and sometimes ambient energy density corrections at 8 TeV) and $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ corrected for leakage and PU (the latter using ambient energy density corrections). The $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ variable is not yet supported for muon isolation, but work is ongoing to find the best way to subtract the muon from the $E_{\mathrm{T}}^{\mathrm{cone,topo}}$, to apply leakage corrections, and to use the same ambient energy correction as is used for electrons.

It was shown [13] that $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ with the ambient energy correction is in many ways better behaved than $E_{\mathrm{T}}^{\mathrm{cone}}$ with the NPV dependent corrections: narrower distributions, better data/MC agreement, and less dependence on PU. The tentative **recommendation** is to pursue the possibility to make $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ with ambient energy the new default for Run-2. Significant improvements are planned for Run-2, as were outlined at the Mar 21/14 Egamma meeting [14]: *i*) Improvement of the subtraction of the candidate electron/photon in $E_{\mathrm{T}}^{\mathrm{cone,topo}}$. Currently, a 5×7 sliding-window approach is used. This is being replaced with the removal of the full topo cluster associated to the electron, which should significantly remove the bias observed with the sliding-window approach. *ii*) Improve the computation of the area of the cone used in the ambient energy correction. Currently, this is a fixed cone size. The possibility of kT-jet building on top of the electron candidate and then using this jet area as part of the ambient energy correction is being investigated. It is to be noted that it is unlikely that calorimeter isolation variables can be recomputed at the xAOD level given the amout of cell-level information required.

3.3.3 Other isolation variables

Mini-isolation provides shrinking cone sizes as a function of, e.g. the lepton candidate p_{T}. It is currently mainly used for muons, but could also be applied to electrons. This type of isolation provides enhanced performance for specific topologies but is not meant for across-the-board usage as a replacement of the existing cone variables.

There also exists a track isolation variable based on the number of tracks in a cone, rather than the sum p_{T} of the tracks. Not much usage of this variable was reported in the questionnaires. This specialised variable is not considered as a primary track isolation method in the harmonisation exercise.

The implementation of other isolations variables is also under discussion. However, until they exist and have been fully validated, they are not part of the harmonisation exercise.

3.3.4 Harmonisation of the isolation tools and centrally-derived isolation efficiencies

Harmonisation might be achieved by providing the user with a common tool or interface that returns cut values for a fixed isolation efficiency. For some analyses, this efficiency will be lower, in order to control significant background-related issues, while for others it can be higher as long as the background is not a problem. Users would specify e.g. track or calorimeter isolation, cone size, particle type, desired efficiency and the tool would return the corresponding cut value. These numbers would have to be based on a fixed benchmark process like isolated leptons from W or Z. It is acknowledged that such a tool would not work for every analysis environment, but providing some standardised benchmark values would already be a significant improvement for many users.

The EisoTool in egammaAnalysisUtils already provides requirements to perform electron isolation for a fixed isolation efficiency between 70% and 99%, binned in electron η and E_{T}. It was established with a tag and probe analysis using the 2012 $Z \rightarrow ee$ data. Currently, the tool supports $E_{\mathrm{T}}^{\mathrm{cone}}$ (20,30) (with leakage and NPV corrections), $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ (20,30) (with leakage and ambient energy corrections), $p_{\mathrm{T}}^{\mathrm{cone}}$ (20,30), and mini-isolation. It was demonstrated that the functionality of this tool could easily be extended to provide cut values for any parameterisation of the desired efficiency as a function of p_{T} or E_{T}. This would enable a user to get e.g. lower efficiency at lower p_{T} where background might be an issue and higher efficiency at higher p_{T}.

The **recommendation** is to extend this tool (or an equivalent common interface) to also support muon and photon isolation. The efficiency "lookup tables" needed to populate the tool with values for isolation efficiencies should be centrally derived by the CP groups, in collaboration with the physics analysis groups. These tables, which already exist for 2012 electrons, would also need to be built from data for muons and photons. Though $Z \rightarrow \mu\mu$ (and maybe $J/\psi \rightarrow \mu\mu$) could be used for muons, it is not clear that the current benchmarks used to calculate photon ID efficiency could be used for getting as well the isolation efficiency (radiative Z's could be used, but their p_{T} range is limited). *So, further studies are required to determine the appropriate benchmark processes for photon isolation efficiency.*

There are no real show-stoppers to implement this technically, though it was reiterated that any efficiency values provided are for specific benchmarks e.g. $Z \rightarrow ee, \mu\mu$ tag and probe data. So, a user would have to be aware of the circumstances under which this tool could be used and potentially consider adding systematics if the isolation environment of an analysis is different than the benchmark used in the tool. The current tool provides a cut value for a fixed isolation efficiency binned in electron η and E_{T}. The point was raised that perhaps it will also require ϕ for muons and so perhaps the entire lepton/photon 4-vector might need to be passed to the tool.

The issue was raised of what the tool should provide for p_{T} ranges beyond e.g. the Z tag and probe data currently available (i.e. TeV-scale objects). This is not so much an issue related to the tool itself but rather how to populate the cut vs. efficiency "lookup table" (either with MC information or somehow extrapolation of data information to high p_{T}).

Recommended path for Run-2: The EisoTool (or at least a common interface that can be separately managed by the CP groups) should be expanded to provide "continuous" calorimeter and track isolation working points for electrons, photons, and muons. This would provide a harmonised interface with a common look and feel and enough flexibility that analysers could choose either their own working point, for analyses that are not too sensitive to isolation, or more complicated working points for analyses like in exotics that need to span a wide kinematic range of isolation. The CP groups would provide centrally-derived efficiency lookup tables for very specific benchmarks. Though this would not meet the needs of all analyses, it would provide valuable isolation information for many as well as a launching point for the determination of isolation cuts for more complicated topologies. Having a common interface would also enable to somewhat trivially (technically speaking) later extend the offered isolation working points to additional environments, such isolation of particles in busy events, that might be derived by analyses desiring new benchmarks. This recommendation does not encompass the data-to-MC scale factors to correct the isolation efficiencies. These are addressed in the next section.

3.3.5 CP support for isolation: isolation efficiency lookup tables and data-to-MC scale factors

It is already recommended in the previous section that the CP groups provide the efficiency lookup tables for the common isolation tools. In addition, the proposed **recommendation** is that the user could get a very few specific benchmark isolation scale factors and associated systematic uncertainties from the same tool that provides the equivalent PID scale factors from each CP community. Mention was made that this could also be unified across CP groups behind a common wrapper tool so that the users have one-stop shopping for all scale factors and associated systematic uncertainties for all CP communities.

The existing Egamma and Muon CP groups should expand the functionality of their existing PID scale factor tools to encompass isolation. For e.g., TElectronEfficiencyCorrectionTool is already instantiated several times by a user to get trigger, reconstruction, and PID scale factors. It is a **recommendation** that the CP groups provide a limited number of isolation scale factor benchmarks, for each supported PID working point (e.g. two: a looser one and a tighter one), if this is required, or preferably independent of PID working point, if this is possible. If too unwieldy, a tool or receipe to calculate these scale factors could be provided. These few reference benchmarks would be sufficient for many analyses. Analyses requiring other working points would have to provide their own scale factors. However, already having a few reference points would allow them to more easily extrapolate to their preferred working points and potentially merely add to an existing systematic error for the use of this alternate working point.

It is to be noted that background rejection using isolation is more analysis dependent. Such values would have to be derived directly by the analyses and not by the CP groups. However, it was raised that the CP communities might provide plots of key kinematic variables (e.g. number of jets, etc.) for the benchmarks that they used to determine the isolation working points and scale factors so that a user could carefully consider whether or not the offered isolation efficiencies obtained by CP are applicable to a specific analysis. There is the worry that with a harmonised tool comes the danger of analyses using it blindly when it may not be applicable to a different topology.

Recommended path for Run-2: The CP groups, in association with physics groups, provide the required lookup tables to provide the "continuous" calorimeter and track isolation working points for electrons, photons, and muons for very specfic benchmarks (such as $Z \to ee, \mu\mu$ for leptons). Theses continuous working points would not all be supported by scale factors. Instead, a limited set (perhaps two, related to *loose* and *tight* isolation working points) of well-specified isolation data-to-MC scale factor benchmarks, and associated systematic uncertainties should be made available. The CP groups would also provide guidelines or even tools to calculate isolation efficiencies such that some of the scale factors derived by analyses, and later endorsed by the CP groups, could then be made available to other analyses. *Studies are still required with the 2015 trigger and PID operating points as to whether or not these have to be provided for each PID/trigger configuration or whether isolation can be factorised from PID/trigger.*

3.4 Trigger considerations for PID and isolation harmonisation

It is important that the online and offline development/usage of PID and isolation to be harmonised as much as possible to avoid mismatches that can cause losses in efficiency. However, it is also important to state that this should only be a harmonisation where ever it is possible and makes sense. For e.g., limiting our trigger rates to be within their allocated budgets is more important than harmonisation at all costs. Here, quite some studies are still needed in the context of the DC14 data challenge at 13 TeV in order to develop the 2015 starting menus and have a final proposal ready for the start of Run-2. The documentation below should be updated by the end of 2014 to reflect the final recommendations. A few trigger considerations are listed below.

3.4.1 Muon triggers

For Run-2, offline tools will be used at HLT and the third chain will be supported. Some biases are expected due to L1 trigger detector coverage/performance and the use of dedicated fast algorithms and offline algorithms at HLT. It was shown that Run-1 isolation was already quite good, with the trigger rate being dominated by real prompt muons. Ongoing investigations showed that to further reduce trigger rates tighter isolation will be required. Several benchmarks are investigated for Run-2. However, no significant trigger issues related to the harmonisation of online and offline muon reconstruction and identification were identified.

3.4.2 Tau triggers

The feedback from the questionnaires was that the tau PID users were generally happy with Run-1 PID working points, but more harmonisation between trigger and offline would have been better. The 2012 trigger strategy resulted in degraded L1/L2 energy resolution with respect to the offline and online/offline differences in the BDT training that caused trigger efficiency loss. A concrete list of items which were considered short-comings in 2012 was identified. The 2015 trigger design will try to address:

- energy calibration constants not the same between trigger and offline,
- no pileup energy correction used online,
- different input variable set and training for the BDT,
- no equivalent of the L1 isolation used at offline, no equivalent to the L2 cut-based approach,
- tracking requirements at trigger level are different.

Significant work needs to be invested in preparation for Run-2 to mitigate the current differences between online and offline. There is a tentative plan to train online PIDs as similarly to offline as possible so that the PIDs will be compatible (no major loss requiring both) at the start of 2015. But after data-taking begins, the offline PID will be subject to studies and improvement which could reduce the overlap with online PID, as in 2012. In this case, the improved offline PID will be developed in parallel for unbiased taus while maintaining some version of the offline PID which is best for triggered taus. But much work still has to be done now and in the context of the DC14 challenge which makes it impossible to fully discuss the final harmonisation solution in this document.

3.4.3 Electron and photon triggers

Some important differences between online and offline for electrons and photons will remain in Run-2. The 2012 calibration involved a scheme of data-driven precorrections followed by a MVA-based calibration. The former will be difficult to keep in synch with the online and the latter is being developed to enable online MVA calibration. Thus, residual mismatches between online and offline will remain because of e.g. precorrections and the lack of bremsstrahlung recovery with GSF at trigger level. Though GSF was implemented at the trigger EF for electrons, it was deemed to be too CPU intensive to be used online. Hence, GSF is only used offline.

For the start of 2015, both offline and online will have cut-based photon PID and both cut-based and likelihood-based electron PID. For the Run-2 trigger, electrons will not be a subset of photons since MVA calibration is optimised separately for electrons, converted and unconverted photons and soon after the start of Run-2, electrons will be LH while photons will be cut-based. Online LH triggers currently implemented are without pileup-dependent cut values. Offline PU corrections are based on the number primary of vertices, which is not available online. Using μ, rather than the number of primary vertices is being investigated. Also, since GSF is not used for the trigger, online LH will not have access to the GSF variables used for the offline LH PID.

In Run-1, the HLT $p_{\mathrm{T}}^{\mathrm{cone}}$ was opened with respect to the primary vertex (or 0,0,0 if no vertex was found in the ROI). For Run-2, the cone will be opened with respect to the track z_0 if no vertex is found or if it far from the track. This change will put isolation more in line with what is done for muons.

The differences between offline and online are inevitable, but several ongoing studies are attempting to evaluate and to mitigate e.g. the difference in electron LH for online and offline.

3.5 Charge misidentification

The correct identification of the charge of an electron is important in many analyses, e.g. when exploiting charge correlations of the final-state particles. During Run-1, analysers would calculate this charge-flip effect for their analyses with e.g. a $Z \to ee$ tag and probe sample considering as probes the ensemble of di-electron pairs without any requirement on the reconstructed sign of the track. Analysers would then present their findings to the Egamma CP group, asking that the results be endorsed for use in publications.

Recommended path for Run-2: Calculating the charge misidentification rate using e.g. the same samples already used by the Egamma CP group to calculate PID efficiency and scale factors seems to be a straight forward extension of the CP work. The Egamma CP group should take ownership for determining this quantity for a well-specified benchmark. In fact, in the recent Egamma publication of 2011 performance [15], this measurement is included.

3.6 Determination of fakes

A number of methods is used to determine the level of fakes in an event sample. Two commonly used methodologies are as follows. The fake-determination methodology adopted by many physics groups is the "two-dimensional side-band (so-called ABCD) method", where three mutually exclusive regions near the signal region are used to determine e.g. the multi-jet background in the signal region. This method assumes that the two variables used to determine the four regions are uncorrelated. The "matrix method" is popular in e.g. top analyses [16] to evaluate the multi-jet background using background leptons. The method relies on defining e.g. "loose" and "tight" lepton samples and measuring the "tight" selection efficiencies for real leptons in data control samples dominated by real leptons and background efficiencies for "loose" leptons in control regions where the contribution of background leptons is dominant.

The optimal way to evaluate fakes is quite likely process dependent and so no one universal method is likely recommendable. In analyses where multiple fake-evaluation methods were used on a common sample, it was demonstrated in a few cases that the systematic uncertainty on a specific method was sometimes not enough to cover the different results observed by two methods. In other cases, the alternate method resulted in a much less precise determination than the primary method, resulting in the alternate method only being used as a cross check, rather as part of a systematic uncertainty. This advocates the need for additional studies to ultimately provide recommendations on the methodology to use in specific environments, including the extraction of the associated systematics. It is a **recommendation** that the concerned CP and physics groups organise a series of meetings to review the methodologies used in Run-1, and encourage Run-1 analysers to do comparisons right down to the systematic evaluation level of different methods on a common sample. Doing this now with Run-1 data would directly help the preparations for Run-2 since the early running environment will likely be not so different from that of Run-1.

3.7 Studies still required with Run-1 data

Although much information was gathered through the questionnaires, it was also evident that some comparative information that can lead to analysis harmonisation decisions was missing. What is ideally needed is for the same people to apply different methods to a same analysis and seeing the impact on the results, right to the level of evaluating the systematics. Some analyses were requested to perform specific checks to help the harmonisation effort, but were unable to do so because of their current full engagement in finalising the Run-1 papers. It is recommended that the following studies be performed with Run-1 data by the fall of 2014.

3.7.1 PID

It is urgent that the CP groups finalise their "best" performance recommendations for the 2012 dataset. It is particularly important for the Egamma and Muon CP groups to do so for the LLH electrons and third chain muons, respectively, as these methodologies are planned to be the default for Run-2. The conclusion of such a study is that the performance parameters for LLH electrons can be established to the same performance and precision as those for the cut-based electrons.

The questionnaires related limited experience with LLH electrons and third-chain muons and so the following types of analyses are recommended to be done with Run-1 data. This does not need to be done for all Run-1 analyses, but rather a few representative analyses with specific kinematic features (e.g. high E_{T}), environments (e.g. busy vs. quiet), and level of precision not yet demonstrated in previous Run-1 analyses using LLH. e.g., in order of priority:

1. High lepton E_{T} measurements such as performed in exotics analyses of Z'
2. Precision SM measurements such as cross sections of W, Z and W, Z masses
3. Precision and high E_{T} measurements that traditionally use bit-flipping of cut-based electrons to build background templates (e.g. $Z \to ee$) to see if background templates can still be built with the proposed LLH working points or if cut-based bit flipping would still be required. The possibility of building specialised LLH working points for particular types of backgrounds is being investigated, but the proliferation of supported working points should be avoided if possible.

At the Egamma workshop in the fall of 2014, a concrete plan for the switch over from cut-based to LLH electrons (for both online and offline) should be established. This workshop should include discussions as to whether commissioning LLH electrons with the 50 ns data early in 2015 is sufficient to make the decision to switch or whether 25 ns data will also be needed. As well, it should be addressed how to commission later years of LHC running without optimised cut-based menus (i.e. as we move to higher luminosity and different centre-of-mass energy).

3.7.2 Isolation

Some specific questions were added to the questionnaire to understand the usage of isolation in analyses and to find out for whom isolation was a dominant systematic. Investigations were also made into how the isolation working point/thresholds were determined in analyses. Annecdotal evidence revealed that some choices were historical in nature (i.e. the decision was made to automatically adopt what was already being used in the same or other working groups) or studies were made to find the best working point, but often not right down to the systematics evaluation level (i.e. deciding on the working point with the best performance without evaluating if alternate working points would also be sufficiently adequate). For the analyses most sensitive to isolation in their systematics (e.g. exotics high-mass di-photon and SM low-mass Drell Yan were two analyses mentioned in the questionnaire) and for those with middle sensitivity, the following studies, ordered by priority, are recommended:

1. Benchmark the impact of replacing their usual track and/or calorimeter isolation cone size to an alternate value (e.g. switching a default $R = 0.3$ to a 0.2 or 0.4). What is the impact on the final systematic uncertainty? Can the analysis make do with a different cone size? The hope is to limit the number of supported cone sizes to no more than two for each of track and calorimeter isolation.

2. What range of isolation threshold values would produce an acceptable systematic to their analyses (in order to provide feedback as to what would be appropriate values for a "loose" and a "tight" fully-supported isolation working points).

The questionnaires also reported rather differing treatment of electron and muon isolation within a same analysis. Studies should be made as to whether a different treatment of electrons and muons is really justified, once the candidate lepton has been properly removed from the cone. This should be studied in a few analyses with the following characteristics:

- Measurements with quieter and busier environments (e.g. in SM measurements)
- Analyses at high E_{T} since issues were reported with track isolation for high E_{T} electrons (e.g. SM and exotics)

Such studies could be made on a more common footing once $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ for muons has been developed and the studies for the improved $E_{\mathrm{T}}^{\mathrm{cone,topo}}$ for electrons have been completed. These two projects should be completed by the Muon and Egamma CP groups as soon as feasible.

3.7.3 Fakes determination

There is a lot of experience in ATLAS in fakes determination. In a first instance, the experience of Run-1 should be shared across analyses in different working groups. The concerned CP and physics groups should organise a series of meetings to review the methodologies used in Run-1, and encourage Run-1 analysers to do comparisons right down to the systematic evaluation level of using different methods on a common sample. This should be performed on a wide spectrum of analyses since the optimal way to evaluate fakes is likely process dependent. Example analyses that could contribute to these studies are:

- Di-lepton top (i.e. reasonably busy environment enriched by fakes from b decays)
- High and low mass Drell Yan (i.e. spans the p_{T} spectrum, with low mass being a difficult environment)
- $H \to WW$ (which already has documented significant experience) and the equivalent SM WW. Inclusive W (which is a quiet environment where missing transverse momentum is often used to build templates)
- SM di-photon (which in the past has documented significant discrepancies when applying two methodologies to a same sample)

The idea would be to come up with a set of recommendations on the types of methodologies to use for particular environments and as well as prescriptions on how to evaluate the systematics associated with each method. The final outcome should be a centralised tool to calculate fake-rates, based on the most commonly used/needed methodology.

4 Harmonisation Study Group 3: Jets and Missing Transverse Momentum

This section summarises the conclusions reached on the harmonisation of jet and missing transverse momentum algorithmic definitions, calibrations, selections and cleaning. The study group recognises certain aspects for which harmonisation is possible while retaining the flexibility required to provide the best available performance at all times and to all analyses. Given the constraints imposed by these requirements, preliminary guidelines on resource allocation in the new analysis model [17, 9, 18, 19] and how they impact the flexibility of Jet/Etmiss analyses and the harmonisation effort are also discussed. On aspects where harmonisation is not possible at this time, a path towards harmonisation at the beginning of Run-2 is also outlined.

The trigger is not discussed specifically, but harmonisation decisions on reconstruction and calibration should be propagated to the trigger whenever possible to ensure optimal turn on curves. Since most harmonisation studies will be quite advanced or concluded during DC14, trigger studies that address the impact on the turn-on curves of the different harmonisation recommendations should be performed in the early stages of DC14. Such studies would guarantee that harmonisation decisions (such as the choice of jet calibration) do not impact the trigger turn on curves in a significant way. Feedback from these studies would help in the menu definition, which will be finalised before the end of the year. Ideally, these decisions would be made using the $\sqrt{s}$ = 13 TeV DC14 simulations, since the conditions and geometry will be those to be used during Run 2. However, some of these trigger studies can be performed using the Run 1 part of DC14, and these could be used to make certain decisions if the Run 2 part of DC14 is delayed.

4.1 Inputs to jet and missing E_{T} reconstruction and studies

Jets and missing E_{T} in Run-1 have been built and studied using uncalibrated (EM) clusters, calibrated (LCW) clusters, tracks and p-flow objects. During Run-1, LCW weights have shown a non-negligible sensitivity to pileup which is still under study [20]. Particle flow objects are currently the subject of extensive R&D. Since the outcome of these studies remains to be known, all these inputs may remain useful in Run-2.

The selection used for tracks and p-flow objects has been different, depending on the specific purpose for which it was used. An effort is ongoing to standardise the track selection used by all jets and missing E_{T} clients. A result of that effort is the common selection used by the purely track-based missing E_{T} reconstruction and the new `METRefFinal` with track-based soft components for the missing E_{T}. This effort will have to continue through DC14 to guarantee that the impact of the new geometry, with the IBL, is understood and considered.

Low-level objects, such as tracks or truth particles, will be associated to jets following a ghost-matching association procedure [21]. This procedure (as opposed to the ΔR association still in use for some studies during Run-1) follows the geometry of the jet more accurately, and has been shown to provide critical performance in certain studies within the Jet/Etmiss group [22]. In the spirit of harmonisation, it is expected that the overlap removal and flavor tagging study groups will understand the potential for using ghost matching too.

Recommended path for Run-2: Track selections will be studied in detail in the Run-2 part of DC14. One set of baseline selections for jet and missing E_{T} studies will be then proposed based on the current experience, the tracking CP recommendations (see Section 2) and dedicated DC14 studies. Variations on these selections may be used upon justification and given dedicated studies. One set of p-flow objects will be used and maintained for both jet and missing E_{T} studies. Tracks and truth particles will be associated to jets using ghost-matching and ΔR matching will no longer be supported by the Jet/Etmiss group.

4.2 Jet selection and cleaning

The p_T thresholds, pileup rejection and jet cleaning cuts have been studied. The use of different p_T thresholds by different analyses is driven by signal and background selections and cannot be subject to a hard harmonisation recommendation. Therefore, the Jet/Etmiss group will continue to provide recommendations down to jets of $p_T = 20$ GeV or any p_T cut that is determined to be of high efficiency and for which uncertainties can be estimated. In the context of the missing E_T, the p_T thresholds can play an important role, since they determine which energy depositions contribute to the soft term and which do not, and the definition and systematic uncertainties on the soft term need to be then carefully evaluated. This will be discussed in detail in Section 4.5.

During Run-1, the Jet/Etmiss group has provided several cuts on the jet vertex fraction (JVF) [23] to allow analyses to obtain different levels of pileup jet rejection. Three cuts were proposed for 2012 data while only two cuts were part of the 2011 data recommendation. Most analyses have used the tightest cut. Certain groups, such as the SUSY and Higgs groups, have, however, performed detailed optimizations. From the SUSY group optimization studies using 2012 data all loose ($|JVF| > 0$), medium ($|JVF| > 0.25$) and tight ($|JVF| > 0.5$) operating points have been used. Those studies have found the loose cut to remove the jet multiplicity dependence on N_{PV} for analyses with requirements on hard objects (jets or missing E_T), while the tight cut would introduce a negative trend in the average jet multiplicity as a function of N_{PV}. The tight cut was, however, necessary to achieve a flat average jet multiplicity as a function of N_{PV} in final states with less hard activity [24, 25]. This topology dependence is not fully understood at this time.

An interplay between efficiency for signal selection and rejection of pileup jets is expected, in such a way that qualitatively one can expect some level of topological dependence. This is why the Jet/Etmiss group has provided several operating points. However, the study group should provide recommendations on harmonisation based on quantitative estimates of the importance of having several operating points. Such quantitative estimates require studies from each physics working group, but also state-of-the-art pileup rejection tools. The Jet/Etmiss combined performance group has recently commissioned a new jet vertex tagging algorithm (JVT) [26] that improves significantly the rejection power for pileup jets and eliminates the N_{PV} dependence of the signal efficiency, which could be behind some of the observations made by the SUSY group. Since the JVT variable is not yet available for analyses to use, it is not possible to determine at this time whether one operating point will be sufficient to cover most analysis needs. Such a study will be pursued as part of the DC14 exercise. Even if it is determined that one operating point is sufficient for all analyses participating in the DC14 exercise, the Jet/Etmiss combined performance group will still support other operating points in the beginning of Run-2 for analyses that were not represented in DC14. However, a generic recommendation could be made available for most analyses based on the studies based on DC14.

The Jet/Etmiss group has also provided four levels of jet cleaning cuts during Run-1. The cuts defining the looser, loose, medium and tight recommendations were first defined in 2010 [27] and updated in 2011 [28] for the new detector conditions and updated luminosity. Incremental updates happened in 2012 [29]. In 2012, analyses had to redefine some of the recommendations that had not been updated by the Jet/Etmiss group, in particular in the context of exotic searches. Such redefinitions varied from analysis to analysis, even though they were trying to address the same effects. The loss in coherence across analyses came from a lack of updates to the medium and tight recommendations and from those studies being performed outside of the Jet/Etmiss group. This should be easily avoidable in Run-2 by emphasising the mandate of the Jet/Etmiss group to provide cleaning recommendations for all analyses.

The four recommendations were found to be useful for different analyses: most analyses using the looser recommendation, with the highest efficiency, while exotic and Standard Model analyses requiring very clean samples at high p_T and willing to take the efficiency loss at low p_T used the medium and tight recommendations. A generic loose recommendation would satisfy most analyses and could be the result of the harmonisation effort, but other recommendations will be maintained by the Jet/Etmiss group. The looser and loose recommendation, on one side, and the medium and tight recommendation could be merged into two recommendations, but such decisions will have to await the beginning of Run-2, once the size of the issues addressed by the Run-1 recommendations and some potentially new effects caused by the smaller bunch spacings are understood.

Recommended path for Run-2: Several operating points were necessary in Run-1 to handle the topology dependence of pileup suppression techniques. The development of more sophisticated taggers may allow for a more generic recommendation. This will be studied in DC14. The Jet/Etmiss group will support (if not recommend) several operating points for pileup suppression cuts at the beginning of the run. JVT or other newer variables will be recommended for pile-up suppression and JVF will no longer be maintained. The Jet/Etmiss group will provide several jet cleaning recommendations as in Run-1. The looser recommendation was used by most analyses, and can be expected to again be the mainstream recommendation. The medium and tight recommendations could be merged into one to simplify the selection recommendations, but any such decisions will need to be postponed to the beginning of Run-2.

4.3 Jet reconstruction and calibration

During Run-1 the Jet/Etmiss group has used many jet calibrations and jet reconstruction algorithms for a variety of studies. The reconstruction algorithms and calibrations in use by physics groups for the final Run-1 publications are summarized in Table 1. The algorithms and calibrations used by the Jet/Etmiss

	EM+JES	EM+JES+GS	LCW+JES
anti-k_t $R = 0.4$	X	X	X
anti-k_t $R = 0.6$	X		X
anti-k_t $R = 0.2,0.3,0.5,0.7,0.8$			X
trimmed ($R_{\mathrm{subjet}} = 0.3$, $f = 0.05$) anti-k_t $R = 1.0$			X
filtered ($\mu = 0.67$, $y_{\mathrm{filt}} = 0.09$) Cambridge/Aachen $R = 1.2$			X

Table 1: Jet algorithms and calibrations in use by physics groups at the end of Run-1.

group in R&D studies exceeds those listed in the table. The proliferation of algorithms and calibrations has physics and historical reasons. The EM+JES calibration was originally designed as a commissioning calibration and was supposed to be phased out during 2011–2012, either by the EM+JES+GS calibration or the LCW+JES calibration. The worse resolution found at low p_{T} for the LCW+JES calibration in 2012 [30], and the lengthened process of validation that the GS calibration has undergone has made the EM+JES calibration survive in certain analyses sensitive to low-p_{T} resolution until the end of Run-1[25, 31]. The limitations of the LCW+JES calibration at low p_{T} have also driven the need to commission the EM+JES+GS calibration, currently the calibration with the best resolution across all jet p_{T}s. The GS calibration, which improves both EM- and LCW-based JES, will be a default part of the JES for the final 2012 calibration. This is expected to be true in Run-2 as well.

The use of $R = 0.4$ and $R = 0.6$ anti-k_t jets has proven an interesting exercise in the inclusive jet measurement, revealing differences with respect to next-to-leading-order calculations[32, 33]. These differences highlight that jets cannot be considered simple 4-vectors for either experimental measurements or theoretical calculations. In order to shed more light into these discrepancies an inclusive jet analysis with anti-k_t $R = 0.2$–0.8 jets is currently in preparation[34]. The use of different jet radii cannot be excluded in Run-2 as a source of additional event structure information in other analyses too.

The use of large-R jet collections has proven key in extracting features of boosted topologies, important for specific final states in the SM, Exotics and Top physics working groups. Large-R jet collections underwent in 2011 a large harmonisation exercise. The result of that was the choice of two algorithms and one grooming configuration per algorithm. Further reduction is not possible at this time, since the usage of certain substructure variables for top and boson tagging is still under investigation and the calculation of some substructure variables requires using specific jet clustering algorithms.

Despite all considerations above, the vast majority of analyses used exclusively anti-k_t $R = 0.4$ jets. The kinematic regime available during Run-2 will increase the use for algorithms focusing on boosted topologies, and their usage is expected to increase. However, this does not preclude concluding at this time that anti-k_t $R = 0.4$ jets will be the recommended algorithm for most analyses.

Given the advances made in understanding the jet calibration during Run-1 it is also foreseeable that the Jet/Etmiss group will be able to recommend one calibration only for anti-k_t $R = 0.4$ jets which will be optimal at both low and high p_{T}s. From the experience in Run-1, we know that the calibration may need to change during the course of the run. However, and as long as the corresponding collection is stored in the xAOD, this should be readily applicable in the derivation framework [17, 19] within a short timescale, which would allow all analyses to use the recommended calibration uniformly at all times. As a consequence, the xAOD will be required to store inputs to all calibrations: EM topological clusters, LCW topological clusters and particle-flow objects, and the corresponding anti-k_t $R = 0.4$ jets.

The Jet/Etmiss group needs to continue supporting jets reconstructed with other radius parameters, as required by the physics output of current SM analyses and potentially also searches in Run-2. Given a reconstruction time of $\approx$ 100 ms per jet collection, the reconstruction of 5–10 additional jet collections in the derivation framework should be possible, as long as only a fraction of the data is considered. One of the main bottlenecks for the usage of any new collection is the *in situ* jet energy scale determination analyses [35], which include dijet and Z/γ+jet final states. The trigger selection used in these analyses in Run-1 selected about 10% of the data, which validates the possibility of doing an R-scan in the derivation framework for the commissioning of these jets for general usage in ATLAS. This does not guarantee that all physics analyses will be able to use the R-scan jet collections, but the most foreseeable usage of such collections should fit well within the CPU budget allocated for the derivation framework. Saving all these collections in the xAOD is not considered to be possible at this time due to disk constrains. However, the size of the derivations for the analyses interested in the R-scan should be small enough that adding a few collections should have a negligible impact on the total disk usage. Only one calibration (EM, LCW or p-flow) will be recommended for such analyses.

The Jet/Etmiss group also needs to be able to support an always-evolving ever-richer jet substructure programme. Such a programme involves the capability to quickly study new techniques as well as the issuance of generic recommendations to allow a wide variety of physics analyses to use coherently a well-validated set of jet collections and substructure techniques. While this impacts only a fraction of ATLAS analyses ($\approx$ 10%), their use across different physics groups makes the discussion of large-R jets and substructure variables relevant for this document.

The goal of harmonisation across physics groups was achieved through a grooming harmonisation effort in 2011 [36]. That effort will be repeated in 2014 using the first $\sqrt{s} = 13$ TeV MC, available as part of the Run-2 DC14 exercise. The outcome of that effort should be, as in 2011, one or two collections and grooming parameters. To allow for straightforward use across analyses, those collections will be written out to the xAOD. This will also simplify monitoring tasks for triggers using large-R jets. Given the little CPU time involved in the grooming step, the stored collections will not be groomed, allowing for additional flexibility for certain taggers that require the links to all the original constituents. Preferred grooming recommendations will be issued, as in Run-1. R&D efforts for substructure objects can involve a very large number of jet collections (20–30), built from a smaller set of jet collections with different grooming parameters. Most generally, however, these R&D efforts can be performed in a small subset of the available MC and data samples. The latter can, in fact, be heavily skimmed to focus on boosted topologies. It is, thus, foreseen that such R&D efforts can be performed using the tools made available in the derivation framework.

Recommended path for Run-2: Based on the Run-1 experience, anti-k_t $R = 0.4$ jets with one calibration at any given point in time will be recommended for use by most analyses. The derivation framework should allow for updates in the calibration recommendation, as long as anti-k_t $R = 0.4$ jets built with EM, LCW and particle flow constituents are stored in the xAOD. The R-scan using anti-k_t jets can be commissioned using the derivation framework. Its usage for all foreseen (not necessarily foreseeable) analyses is also possible using the derivation framework. A harmonisation of the large-R jet collections used in boosted topologies will be performed for the beginning of Run-2 during the DC14 exercise, as was done in 2011 for the Run-1 analyses. Two or three ungroomed jet collections will be saved in tier 0 for monitoring purposes and to ease the use of the recommended collections across analyses. Recommended grooming will be performed in the derivation framework. R&D carried out within the Jet/Etmiss group using many jet collections and grooming settings will be done using the derivation framework.

4.4 Jet energy scale systematic uncertainties

Several recommendations existed at the end of Run-1 for the application of jet energy scale systematic uncertainties. These different recommendations differed in the number of nuissance parameters representing the uncertainties arising from *in situ* tests of the jet energy scale [35, 37]. In particular, a recommendation with the full set of nuissance parameters (50) and one with a reduced set of parameters comprising the most important nuissance parameters (9) obtained through the singular value decomposition of the correlation matrix, were proposed. In addition, a recommendation which kept nuissance parameters of different origin separate in the reduction step, to be able to treat correlations with CMS and across years, was also provided (and this resulted in 15 nuissance parameters from the *in situ* analyses). In addition, alternative sets of uncertainties with stronger correlations were provided for systematic studies of p_{T} and η correlations which are not fully understood for certain terms. For large-R jets, the systematics were mostly obtained with different methodologies, and no recommendation on the correlation between large-R jet and small-R jet systematic uncertainties was provided.

All these uncertainties have been used by different analyses in Run-1. In addition, certain legacy analyses have used a single nuissance parameter, despite the fact that such use has been deprecated since the beginning of 2011. In order to make it possible to combine across years, analyses and CMS with a reasonable treatment of correlations, the Jet/Etmiss group will generically recommend only the equivalent of the recommendation with 14 nuissance parameters to analyses in Run-2. The other recommendations will remain available, but only to analyses that show the need to use these recommendations. These recommendations will extend to the large-R collections, which are currently in the process of being integrated into the main calibration framework. This will allow for a trivial treatment of correlations between small-R and large-R jet systematic uncertainties, except for specific topological systematic uncertainties, that will likely be more complex for certain uses of large-R jets. The use of a single nuissance parameter will not be recommended and its deprecation should be enforced.

The study group recognizes that part of the reason to resort to the use of a single nuissance parameter was due to issues in the convergence of fits that include small nuissance parameters that may be affected most strongly by statistical uncertainties in the MC simulation or the data. No centralised solution to this problem was provided by the Jet/Etmiss group, which resulted in analysis groups with little manpower using the simplest solution: a single nuissance parameter. Analysis groups with more manpower studied this in detail and typically used a methodology that has come to be known as "pruning" [38, 31], or dropping nuissance parameters that do not seem to change the relevant distributions in a statistically significant way. For Run-2, the Jet/Etmiss group will provide, as part of the recommendation, one or several methodologies to assess the uncertainty due to small nuissance parameters, to replace "pruning" and harmonise the treatment of such nuissance parameters. Two such methodologies have already been proposed in the study group, but are not described here because they may evolve quickly. They should be exercised by the DC14 analyses to test their adequacy for generic application.

Recommended path for Run-2: The Jet/Etmiss group will recommend the minimal set of nuisance parameters that allows to correlate across years and experiments for all analyses. Additional sets of nuissance parameters will be available but only for use when it is demonstrated that the generic recommendation is not usable for a specific analysis. The generic recommendation will also provide one or several methodologies to handle small nuissance parameters, if their size is comparable to that of statistical uncertainties in analyses. These methodologies will be tested in the DC14 analyses to obtain the appropriate feedback from physics groups.

4.5 Missing transverse momentum reconstruction and cleaning

It has been recognized by the study group that during Run-1 there was a proliferation of missing E_{T} versions. Some of that proliferation was related to dedicated object selections used by different groups in their analyses, while additional complications were added by the existence of several variants for the reconstruction of the soft term of the missing E_{T}. Several analyses also made used of a purely track-based missing E_{T} for background rejection, and this needs to remain a supported variant during Run-2, even if it will not be the default variant of the missing E_{T}.

The use of dedicated object selections in the missing E_{T} is justified by the sample dependence of fake rates and efficiencies for the different reconstructed physics objects used in different analyses. It is expected that in the new analysis model the selection used in the missing E_{T} reconstruction will be performed in the derivation framework, and for each derivation one fully-reconstructed missing E_{T} (and potentially a fully track-based missing E_{T}) will be saved. Physics groups will be responsible for understanding the appropriateness of the missing E_{T} versions saved in their derivations, and this should lead to a small number of missing E_{T} versions to be commissioned within each physics group. The definitions of each of these versions will be formally described in meetings of the Etmiss subgroup of the Jet/Etmiss group. A standard fully-reconstructed missing E_{T} can be created in the Tier-0 for monitoring purposes, but it is expected to be rebuilt in the derivation framework to take into account the correct calibration constants.

The fully track-based missing E_{T} can also be saved at Tier-0, since its usage is widespread and it does not require updates for calibration constants. The selection involved in the track-based missing E_{T} can be expected to evolve during Run-2, but that evolution is likely to be slow enough to fit well within one of the xAOD-to-xAOD reprocessing campaigns.

Redoing the fully-reconstructed missing E_{T} for $\mathcal{O}(50)$ derivations could be one of the CPU bottlenecks of the framework. In order to overcome this, work is ongoing in parallel to improve the CPU performance of the missing E_{T} reconstruction in release 19 and build an extended composition map for release 20 that would allow rebuilding the missing E_{T} without needing to access the cluster container. This new composition map could help improve the CPU performance in the derivation framework and would allow a totally consistent treatment of the missing E_{T} objects in the estimation of systematic uncertainties, something not possible up to now.

The selection of jets within the missing E_{T} deserves a special treatment since, unlike for other combined performance objects, fake rates and efficiencies are not strongly sample-dependent, and the calibration is of major concern. In particular, the jets that are not considered as hard objects in the missing E_{T} are considered as soft objects, and thus not calibrated. It makes then little sense to leave the jet selection for the missing E_{T} fully to analyses. The optimization of the jet p_{T} cut for jets that are included in the missing E_{T} has been studied in some detail in certain physics working groups [39]. These studies suggest that its optimization might be somewhat topology dependent. Detailed studies are currently ongoing in the Jet/Etmiss group to understand whether a generic recommendation is possible, or whether more than one recommendation will be necessary to suit all analysis needs. Similarly, the use of pileup suppression or cleaning cuts have shown negligible effects in some studies of the missing E_{T}, but more detailed studies are necessary. The jet selection impacts the definition of the soft terms. Therefore, if more than one set of jet p_{T} cuts are considered, that will lead to several soft term definitions available to physics groups. If that were the case, the Jet/Etmiss group would provide general guidelines on how to choose between the different sets.

The expertise gained during Run-1 should allow, as for jets, to provide one baseline methodology for the definition of the soft term. This will simplify the selection of the soft term definition by analyses. This definition currently rests on the use of only tracks and dropping all calorimeter information not associated to hard objects. This definition could, however, evolve during Run-2 based on R&D studies. However, given the rate of evolution observed in Run-1, the relevant modifications could be accommodated within an xAOD-to-xAOD reprocessing campaign, or even the derivation framework. These conclusions refer to the generic recommendations provided by the Jet/Etmiss group. Simultaneously, the `LocHadTopo` variant of the soft terms will be built in Tier-0 to help the monitoring of the trigger, and R&D studies that may use other versions of the soft term will be built in dedicated Jet/Etmiss derivations.

The cleaning of fake missing E_{T} during Run-1 was developed primarily within different analyses or physics groups. The goal of such cleaning cuts was primarily to remove events with missing E_{T} caused by large energy depositions in dead calorimeter modules [24]. Several cuts were defined, according to the analysis that used them, even though the goal of each set of cuts was not very different. In Run-2, the work of analysers will remain key in spotting problems in the missing E_{T} reconstruction and coming up with potential solutions. However, the Jet/Etmiss group and the Etmiss subgroup within it should become the common place for the discussions of those problems/solutions. This will allow for generic recommendations on missing E_{T} cleaning to be produced and harmonisation of how such problems are handled across analyses in ATLAS.

Recommended path for Run-2: The selection of hard objects going into the missing E_{T} will reflect the analysis selection, except for jets, and it is to be performed in the derivation framework. The selections used in the different missing E_{T} definitions will be the responsibility of physics groups, which should consult the Etmiss subgroup of the Jet/Etmiss group about these definitions and any problems encountered in their study. Work is ongoing in release 19 and towards release 20 to make this fit within the CPU resources available for the derivation framework. The selection of jets going into the hard terms of the missing E_{T} will not be left to analyses. The optimal choice of jet cuts is under study, but if the sample dependence is too great, it may lead to more than one recommendation. The Jet/Etmiss group will provide one recommendation only on the methodology to be used for the reconstruction of the soft terms. The Jet/Etmiss group will support and validate the recommended definition(s) of the soft terms, as done during Run-1. R&D studies will require the study of other versions of the soft terms and will be performed in dedicated Etmiss derivations. A few collections, including `LocHadTopo` and a fully track-based version of the missing E_{T} will be saved in the xAOD for monitoring purposes and/or widespread use across analyses. Missing E_{T} cleaning recommendations will be centrally provided by the Jet/Etmiss group with the help and feedback from analyses across the collaboration.

5 Harmonisation Study Group 4: Flavour Tagging

This section summarises the conclusions reached on the harmonisation of flavour tagging truth definitions, algorithms, calibrations, systematic uncertainties and nuisance parameters. The study group recognises certain aspects for which harmonisation is possible while retaining the flexibility required to provide the best available performance at all times and to all analyses. On aspects where harmonisation is not possible at this time, a path towards harmonisation at the beginning of Run-2 is also outlined.

In all casees two aspects are considered for the harmonisation. The baseline recommendations which relate to the usage carried out in $\sim$ 80% of analyses and more specialist usage which will relate to a smaller subset of analyses.

5.1 Heavy Flavour Modelling and Labelling

In Run 2 five different parton showering generators were used in ATLAS MC production: Pythia6, Pythia8, fHerwig, Herwig++ and Sherpa. Each of these generators was found to have significantly different heavy flavour modelling from truth level comparison [?], which consequentially lead to significantly different behaviour in terms of the b-tagging performance. EvtGen re-decays the heavy flavour hadrons, replacing the original decays, and thereby ensures that consistent lifetimes and decay tables are used in the generation of MC events. Without using EvtGen, differences in b-tagging efficiency of up to 10% are seen between the different parton shower generators. After using EvtGen, all generators agree at the $1-2\%$ level. Furthermore, in all cases the agreement with the efficiencies measured in data is significantly improved, such that the residual MC to data scale factor are centered around one. While this constitutes a dramatic improvement in the modeling of heavy flavor hadrons, unfortunately due to the different fragmentation models implemented in the various MC generators a residual difference still exists after using EvtGen: these differences are related mainly to the initial production fraction of heavy flavor hadrons, to the number of tracks not from the HF-hadron decay and to the fragmentation function. Longer term these generators can be tuned to data to reduce these dependences, but whilst they exist generator dependent data/MC SFs will still be necessary.

The recommendation for Run 2 is to use EvtGen for all MC production as default.

For Run 2 the flavour labelling will be changed from the parton ΔR cone-based labelling used in Run 1 to the ghost matching of weakly decaying hadrons as outlined in the truth task force report [?]. In the parton ΔR cone-based labelling, the labelling of the flavour of a jet in simulation is done by spatially matching the jet with generator level partons after final state radiation (FSR): if a b quark is found within $\Delta R < 0.3$ of the jet direction, the jet is labelled as a b jet. If no b quark is found the procedure is repeated for c quarks and τ leptons. A jet for which no such association can be made is labelled as a light-flavour jet. The new procedure changes both the objects used for the matching, and the criteria according to which they are matched to a jet. In terms of objects, weakly decaying heavy flavor hadrons replace the use of final state partons: hadrons are well defined physics objects, while partons are ill-defined objects whose definition depends on the MC generator being considered, and the new definition follows what was already done for several unfolded measurements of the cross section of heavy flavor jet production. In terms of matching criteria, instead of a simple ΔR matching, the use of the ghost association procedure is suggested. Ghosts are additional particles with nearly vanishing momentum which are added during jet clustering, and are therefore clustered together with the remaining jets: as a result of the procedure, each ghost is unambigously associated to a jet. From a practical standpoint, the main differences in the new procedure with respect to spatial ΔR matching are that:

- The effective association radius varies with the jet size, in particular for AntiKt jets with $R = 0.4$, the association radius will be $\Delta R \approx 0.4$.
- For jets with irregular shapes, the matching will follow the shape of the jet.

- While the previous matching scheme allowed two different close-by jets to be matched to the same heavy flavor parton, with the new scheme the assignemnt of a heavy flavor hadron to a jet is unambigous.

The change in the object definition from partons to weakly decaying hadrons will harmonise the flavour labelling between what was used in many analyses to flavour label events and that used to derive and parameterise the b-tagging performance and data/MC scale factors. Additionally the flavour labelling will be extended to include pile-up jets.

The recommendation for Run 2 is to use ghost-matching of hadrons as default.

5.2 Flavour Tagging Algorithms

5.2.1 b-Tagging for general use

In Run I the vast majority of analyses used the 70% operating point of the MV1 b-tagging tool. MV1 is a general purpose tool, which takes as inputs the output of the likelihood tools IP3D and SV1 along with JetFitterCOMBNN, which combines the output of the JetFitter and IP3D b-tagging tools. It was trained on $t\bar{t}$ jets and so has optimal performance in the range 20-300 GeV. Feedback from Run I analyses also noted that they were limited by the c-jet rejection and requested that a tool with better c-jet rejection would be desirable in Run II.n

The main recommended tagger for Run-II analyses is **MV2**. This tagger is based on inputs from three kind of more basic taggers, impact parameter based (IP2D, IP3D), inclusive SV finding (SV1) and b+c decay chain fitting with a flight axis constraint (JetFitter). The input variables can be found on the following twiki: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTaggingMV2`. This tagging tool has two main advantages over previous . Firstly it signficantly simplifies the b-tagging workflow as it takes inputs directly from the more basic tools (SV and JetFitter) which removes the need for the intermediate tagging tools as JetFitterCOMBNN and SVx to be both tuned and ran. Secondly, it has been shown to have significantly better performance than MV1, with 30% better light-jet rejection for the 70% b-jet efficiency operating point.

MV2 also provides three variations trained with a significant fraction of c-jets among the background of non b-jets (10 or 20%). We expect to calibrate the same working points as for $MV2$. The tentative recommendation is for MV2c20 to be used as the default b-tagging tool, which provides the same light-rejection as MV1 but rejects 25% more c-jets. **T**he optimal light vs c-jet rejection still has to be determined and to determine this we need further feedback from the physic gropus.

Four main working points will be calibrated, corresponding to b-jet efficiencies of 50, 60, 70 and 85%. This is enough to cover both the integrated (single working points) and continuous cases (which will correspond to a 5-bin MV2 distribution). The exact values of the working points have not been established yet, we plan to define them based on MC samples produced with Athena release 20.1. **T**he physics groups are asked to cross-check the proposed working points in their analyses and give feedback as soon as possible. *Readiness: MV2 is ready and available in release 20.*

While we expect this algorithm and these working points to be used by $>$ 80% of physics analysis needs, other algorithms might be used in specific domains. Not all of these might be ready at the beginning of Run-II data taking.

5.2.2 Dense jet environemnt

The MVb tagger improves the b-tagging performance, especially in environments with overlapping jets. The main feature is that it reduces the dependence of the b-jet efficiency on how displaced the b-hadron is with respect to the clustered jet. We will support this tagger for such special applications. The number of working points will depend on the analysis needs. *Readiness: the code is available and running in*

Release 20, but a new training is required: this is work in progress. Involvement of interested analysis groups would be very welcome.

5.2.3 *c*-jet tagging algorithm

We plan to have a dedicated *c*-jet tagging algorithm, based on the Run-I experience, and most likely two working points to calibrate. In Run-I we first developed the JetFitterCharm algorithm, which was calibrated and used in two SUSY analyses, and later GAIA, based on replacing a conventional Neural Network algorithm with a deep neural network, with partially unsupervised training. The novelty was mainly in a dedicated version of JetFitter and in an improved Neural Network which has as ouput the posterior probability of a jet to be a *b*-, *c*- or light-jet flavored jet. *Readiness: GAIA has only partially made it into release 20 and the dedicated version of JetFitter for c-tagging has not been ported ro release 20, so at this point we foresee this to become available only in release 21. If there is sufficient interest, we plan temporarily to train a special version of MV2 with c-jets as a signal. No change to the code is required, but an additional dedicated training and changes to the job options. The help of interested physics groups would be very important.*

5.3 Calibrated Objects

5.3.1 Jets

Two jet collections were calibrated in Run 1, LC and EM jets, with the vast majority of analyses using LC jets for *b*-tagging. The requirement to calibrate two jet collections was a significant additional burden on the flavour tagging group. For Run 2 we propose to only calibrate the *b*-tagging for the AntiKt4 jet collection recommended by the JetEtMiss group. The recommendation for Run 2 will be to calibrate the AntiKt4 jet collection recommended as default by the JetEtMiss group.

5.3.2 Boosted Jets

For boosted environments the recommendation will be to use AntiKt3 ghost matched track-jets to the recommended large R-jet.

5.4 JVT

During Run 1 two JVF operating points were calibrated, corresponding to no JVF cut or JVF ¿ 0.5 (or 0.75 for 7 TeV). The vast majority of analyses used the JVF ¿ 0.5 operating point, whereas most SUSY analyses used no JVF cut on the jets. For Run 2 we propose to only calibrate one JVT operating point, which will be one of the operating points provided by the JetEtMiss group.

Jets fail the JVx cut for two main reasons, they have no tracks associated to them or there is a signficiant probability they are not from the primary interation. In both cases there is little reason to try *b*-tagging them. Studies are ongoing to ascertain the efficiency of a *b*-tagged jet to also pass the JVF/JVT operating point. Studies show that for a 70% MV1 *b*-tag 97% of the jets also pass the JVF requirement. First *b*-tagging studies for JVT have just become available and it turns out that the equivalent efficiency for JVT is $\approx$ 99%. As such if a jet is going to be *b*-tagged then applying a JVT cut should both not harm the efficiency and should also provide a more robust selection of jets to consider for *b*-tagging. Further investigations are ongoing to understand if there is anything left which can be recovered in the residual $\approx$ 1 %. In the event that an analysis demonstrates that a different JVT cut is necessary, an additional uncertainty can be derived.

The recommendation for Run 2 is to use one JVT cut, to be recommended jointly by the JetEtMiss and *b*-tagging groups, on any jets which are to be considered for *b*-tagging.

5.5 Calibrations

The main calibrations will be based on ttbar or di-jet events (b-jets), multi-jet events with fully reconstructed D^* mesons (c-jets) and with the negative tag method in di-jet events (light-jets). For b-jets calibrations in the $p_T = [20 - 300]$ GeV kinematic region will rely on ttbar events (as established in Run-I), while above that a high p_T calibration using template fits and based on di-jet events will be used (this is being presently finalized on Run-I data). The ttbar calibration will be mainly based on di-lepton events, but later on also the 1-lepton ttbar calibration should become important, in particular for p_T in the range of [200 − 400] GeV (as demonstrated in Run-I). *Readiness: for Run-I all these calibrations have been finalized, except for the high p_T calibration, which is going to be finalized on Run-I data in time for the Run-II pre-recommendations.*

5.5.1 Application of scale factors

The physics analysis user will apply the calibration in terms of MC-to-data scale factors for the jets depending on whether they pass or not a certain tagging requirement (integrated calibration) or whether they land in a certain bin of the b-tagging discriminant distribution (continuous calibration). At the beginning of Run-II the data-based calibration will only be provided in bins of p_T, for all three flavors (b,c and light-jet). This is a change w.r.t. Run-I, where the light-jets were also calibrated as a function of η. However the dependence on η will be checked and the strategy changed if necessary. No calibration will be provided for τ jets, although the calibration for c-jets will be used with an inflated uncertainty: this procedure is -to a good extent- arbitrary, so any analysis which is crucially sensitive to the b-tagging efficiency of τ jets will require a more dedicated procedure.

Despite the fact that the dependence of the b-tagging efficiency in simulation on the specific MC generator used will be reduced by the use of EvtGen for decaying b and c-hadrons for almost all generators in MC15, a residual % level dependence is expected and therefore, when applying the data calibration, additional MC-to-MC scale factors will be applied, to take such dependence into account. These factors will be provided by the b-tagging group for all available generators. An uncertainty will be provided as well.

The possible topological dependence of the MC efficiency will be dealt with differently in the integrated and continuous cases. In the integrated case, the usual *inefficiency* scale factors will be used. These absorbe the non-normalization in the overall b-tagging distribution, after the scale factors are applied, by modifying the weight of the jets which have not been tagged. In the continuous case, no correction will be applied as a default, but a systematic uncertainty will be provided which is expected to cover for the analogous effect in the continuous case. These corrections/systematics require knowledge of the differences between the b-tagging efficiencies of the reference sample (for which the calibration has been derived) and for the sample on which b-tagging is applied. While this is analysis specific, the b-tagging group will provide maps for the main different samples (e.g. W+jets, Z+jets, ttbar, dijets).

We plan to smooth the calibration results and systematic uncertainties as a function of p_T. Later on, we plan to do this also for the continuous calibration. Results using a Gaussian kernel smoothing look promising and a first test calibration file is being produced (see Fig. 1. We foresee to have this in place by the time of the pre-recommendations.

Both the integrated and the continuous calibrations will be supported, at least for the default tagger and operating points.

Readiness: Most of the combination code to prepare the analysis calibration file is ready, in the configuration used in Run-I. Presently integrated and continuous calibration make use of different framework. A combination of the two is planned, but there is no manpower available to do this before the start of Run-II data taking; so this is planned for release 21. The work on smoothing is ongoing, first results are ready for the integrated calibrations and the plan is to use the smoothing for the integrated

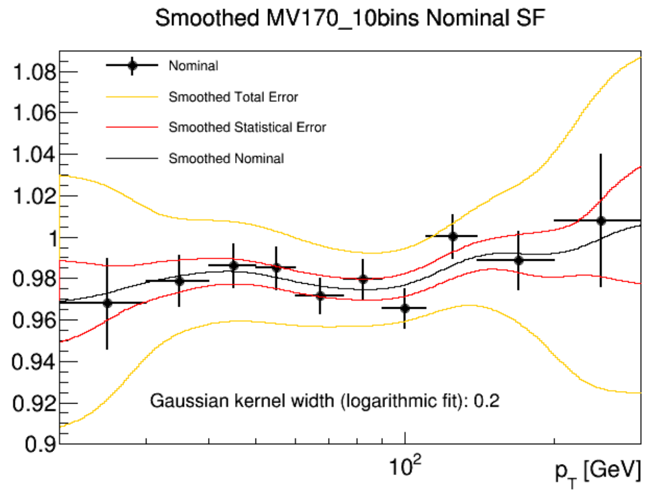

Figure 1: The result of the Gaussian kernel smoothing procedure is shown here in terms of the smoothed calibrated MC to data curve (black curve), the statistical uncertainty band (red) and the total uncertainty band (yellow). This is compared to the original binned data to MC scale factors (data points), which are shown with their statistical only uncertainties. The plot refers to the standard 70% b-jet efficiency working point for the default MV1 algorithm, for 2012 data and for EM jets.

calibrations from the beginning of Run-II.

5.6 Uncertainties and Nuisance Parameters

The scale factors applied to MC to calibrate the b-tagging to the data based measurement are varied according to a set of uncertainties as provided by the calibration analyses.

For Run-II the main recommendations is a two-step approach:

1. For all systematics which are considered in a physics analysis already and which also affect b-tagging (typically JES, JER, top theory uncertainties if considered), correlate the variation in the b-tagging SF with the variation already considered in the analysis.

2. After removing such uncertainties, the eigenvector method is applied to the residual b-tagging systematic uncertainties.

This approach requires a harmonization of the set of NPs to be considered in the domains of JetEtMiss (one default set) and theory uncertainties (PMG, top related uncertainties). In the b-tagging calibration one only set of these will be considered, the default one. This approach has been tested succesfully in the context of the 2011 top mass measurement, with the 2D and 3D methods. It is simple enough to apply that it will be recommended for all analyses: once a certain common analysis systematics is considered (e.g. a certain JES variation), the use of the nominal b-tagging scale factor is replaced by the use of the scale factor corresponding to that specific systematic variation. All residual systematic uncertainties are considered within the eigenvector approach: for a physics analysis this means considering an additional set of b-tagging specific systematics corresponding to the set of eigenvector variations, as described in the next subsection.

For the integrated calibrations, for each flavor (b,c,light and τ-jets) it is possible to consider an envelope of uncertainties, corresponding to the sum in quadrature of all uncertainties. This does correspond to 4 nuisance parameters. However, unless b-tagging only plays a minor role in a certain analysis, this procedure is not recommended, and at least a check using the more complete recommended procedure should always be performed. In this method, infact, the right degree of correlation across the p_T spectrum of the calibrated scale factors are ignored (the assumption is that the correlation is always 100%). An exception to this recommendation can be made for light-jets. For the continuous calibration it is not possible to get an envelope of uncertainties.

Readiness: the CDI tool, which is meant to aid the user in applying the data based calibration scale factors, is fully ready for the procedure recommended for Run-II. On the calibration side, need to agree with other CP groups and PMG about the set of uncertainties to be considered for correlation and about their naming.

5.6.1 Eigenvector decomposition

Even after correlating common systematics within the physics analysis, a sizable number of uncertainties still needs to be considered. Since these are typically larger than the number of p_T (and $\eta|$) bins of the calibration, instead of considering the original uncertainties one-by-one, the full covariance matrix of the scale factors is built and then an eigenvector decomposition is run, resulting in a number of systematic uncertainties equal to the number of bins of the calibration. The uncertainties in common with the physics analysis are removed before building the covariance matrix. The analysers is left with a set of b-tagging specific variations to the scale factors to apply, which are ordered according to importance (eigenvalue) and where the number depends on the flavor and the calibration considered (since these have different binnings). For the integrated calibrations the number of needed NPs will be 7 (6 p_T bins + 1 extrapolation) for b-jets, 5 for c-jets (4+1) and 12 for light-jets. c-jet and τ-jets are considered as fully correlated. Given the high number of NPs for light-jets and the high degree of correlation among bins, for analyses which are not particularly sensitive to the light-jets p_T distribution, it is proposed to use one single NP, corresponding to the envelope. For the integrated calibration, this corresponds therefore to 14 nuisance parameters. *Readiness: the full procedure is ready. The extrapolation procedure is going to be refined for the Run-II pre-recommendations. The exact number of NPs parameter might change slightly as a result of introducing the smoothing procedure: this will also happen in time for the Run-II pre-recommendations.*

5.6.2 Continuous calibration

In the continuous case, everything is analogous to the integrated case, although the binning is extended also to the tag weight distribution, where the independent number of bins is the total number of tag weight bins minus one. The eigenvector decomposition results therefore in a much higher number of NPs. For the continuous calibration, in the default proposed calibration (4 working points), the user will therefore have 24 NPs for b-jets, 16 for c-jets and 48 for light-jets. According to previous studies, not all of these however do give rise to sizable systematic variations, and some can be removed without any visible effect on the total systematic uncertainty (pruned). The NPs are all oredered by their size according to the eigenvalue. In measurements dominated by statistical uncertainties NPs can be pruned less (e.g. c-jets), while much more pruning is possible in the case of light-jets. However, how many of the NPs need to be considered remains analysis specific and an evaluation within the analysis is needed. For the $VH \to Vbb$ analysis it turned out that only $\approx$ 10 NPs were needed for the b- and light-jets, and $\approx$ 15 NPs for c-jets. An additional pruning, in particular for the light-jets, should however be possible without reducing the accuracy in the assesment of systematics. Merging the residual uncertainties, despite desirable, is not conceptually easy to do, because there is not a well defined way to add the remaining set of eigenvector variations in quadrature (i.e. simple addition in quadrature of all uncertainties affecting the tag weight distribution would result in a systematic uncertainty which doesn't preserve the overall normalization). *Readiness: the full procedure is ready, but no changes are planned during the initial phase of Run-II. No* *extrapolation procedure to higher p_T is presently available. The extension of the smoothing procedure to the continuous calibration and a more proper extrapolation procedure are only foreseen on the timescale of release 21.*

6 Harmonisation Study Group 5: Overlap Removal (OR)

6.1 Information collection

Overlap removal addresses both the topic of *duplication*, i.e. the reconstruction of one physical object as two different objects, and the topic of *isolation*, i.e. the treatment of two seperate but close-by objects. The two cases can overlap if reconstruction combines two different objects into a single object. There is also a connection with the object ID, in particular if the (isolating) overlap requirement is used to reject fake objects.

In the week of Jan 14th to Jan 17th, the OR Harmonization group contacts gathered information from their respective analysis groups about current OR practise and about studies performed during Run-1 dedicated to the optimization of the corresponding requirements. Fortnightly meetings were then scheduled, dedicated to specific topics within the OR area.

This document summarises the topical discussions and resulting recommendations on the OR harmonisation.

Disclaimer: The performance of an OR requirement depends on the definition of the objects involved. This suggests a close connection to the recommendations in the previous sections. If not explicitely stated otherwise, the statements below are valid for central isolated electrons and muons connected with a medium-level ID requirement (see Section 3). Jets are AntiKt4Topo jets with the usual pileup and bad-jet rejection applied with a well defined minimum transverse momentum, typically in the range between 20 GeV and 30 GeV (see Section 4). Tau leptons are reconstructed from AntiKt4Topo jets, with additional ID requirements based on tracks and calorimeter shower shape (see Section 3).

We **recommend** as a baseline to apply overlap procedures between the final objects defined in the analysis for selection or veto purpose.

In the OR we need a figure of merit to quantify the compatibility of two objects. The classical approach uses the geometrical proximity. A suitable variable to quantify the proximity of objects is the distance in azimuthal angle ϕ and in rapidity y, both invariant under a Lorentz boost in z direction. In Run-1, the pseudorapidity η has been used instead of the rapidity, which for jets is in general not a good approximation of the rapidity. It is hence **recommended** to use $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2}$ as a measure of the distance of two objects. This change would, however, require a repetition of some of the lepton-jet OR studies to check that the optimized thresholds from Run 1 are still applicable. The same is valid for the improvements in ΔR resolution expected with the correction of the jet 4-vector to the primary vertex.

We encourage studies of alternative measures of compatibility, for example multi-dimensional requirements or the ghost association of electrons and muons to jets.

Overlap removals are also applied in the reconstruction of the missing E_{T} (see Section 4). Since the primary goals of the overlap removals are different in both use cases, we do not aim to harmonize the approaches used in the two sectors. Nevertheless, new techniques used in the missing E_{T} sector should be checked for potential application in the analysis OR.

6.2 Electron-jet overlap removal

The primary goal of the electron-jets OR is to remove reconstructed jets that are identical with reconstructed prompt electron candidates while preserving heavy-flavour (HF) jets with semileptonic decays as well as light-flavour (LF) jets that can fake a loose electron. A secondary goal is to remove events (or objects) where a close-by prompt electron and a hadronic jet might bias each other's position or energy reconstruction. Depending on the details of the requirement the latter is conceptually rather an isolation requirement, i.e. a good-event or good-object selection than a strict OR but it will also be covered in this section since the boundaries between the two regimes are not sharp.

6.2.1 Discussion

An electron is usually clustered as well as a jet. In addition, the presence of the electron disrupts the clustering of hadronic energy in its vicinity.

A typical distribution of electron-jet pairs before an electron-jet OR would show two disjoined populations in the $\Delta R(e, j)$ vs. $E(j)/E(e)$ space. The first population consists of electrons that are reconstructed as jets. The electron-seeded jets can incorporate varying amounts of hadronic energy, leading to a small tail of increased $\Delta R(e, j)$ correlated with an increased $E(j)/E(e)$. The second population consists of jets which are clustered separate from the electron-seeded jet. The two populations are separated by a strongly depleted region around $\Delta R(e, j) = 0.35$. Usually these distributions, together with residual biases on key observables depending on the overlap criterion are used to optimize the OR configuration.

From these results follows that by removing all jets closer to a well identified electron than $\Delta R(e, j) < 0.2$, the plain duplication of electrons as jets can be efficiently removed. The distribution of the population indicates that it contains already a tiny fraction of real hadron jets which have included a close-by electron. Conclusive studies of a possible reduction of the overlap regime to address these issues are missing.

The ΔR regime outside the duplication region (typically $0.2 < \Delta R(e, j) < 0.6$) is characterized by events with real hadronic jets in the proximity of an electron, with two populations: (a) A small fraction of hadron jets which have included the electron clusters, resulting in E_{jet}/E_e significantly above the average and a jet axis moved towards the electron; (b) a population of hadron jets which have lost a fraction of their energy to the jet seeded by the electron, which for a certain fraction of these jets resulted in dropping below the lower p_{T} threshold [40, 41, 42, 43]. For $0.2 < \Delta R(e, j) < 0.35$, also the electron reconstruction is biased [44]. The potential mismodelling of this delicate region involves additional systematic uncertainties on electron and jet reconstruction [45].

The two most common approaches used in ATLAS for this second regime are the electron isolation and the jet isolation, respectively.

In the first approach, the electron is classified as non-isolated and removed, which usually leads to the removal or reclassification of the event. This isolation cannot replace the genuine electron isolation criterion (see Section 3), since its performance depends on the lower jet p_{T} threshold and the choice of the size of the inner duplication cone. This procedure contributes to a more reliable reconstruction of both jets and leptons. In addition, it avoids the reclassification of the hadronic part of the event topology by removing real hadronic jets from the event record.

The second approach removes jets close to the electron in $\Delta R(e, j)$. This strategy is tailored to ensure a precise measurement of cross sections for high jet multiplicities, where lepton identification systematics is only a minor nuisance. Typical OR criteria are $\Delta R(\ell, j) > 0.4(0.5)$. A removal of one jet would usually not discard the event but simply reclassify it as lower multiplicity. In these measurements OR requirements are as well part of the theory reference jet phase space with respect to which cross sections are quoted.

The definition of the affected region in terms of the outer $\Delta R(e, j)$ threshold has been subject of several studies [42, 43, 40, 44, 45]. $\Delta R(e, j) \approx 0.35$ marks the upper boundary for jets composed of the electron and real hadronic activity in the vicinity and corresponds to the region maximally depleted of original hadronic jets. In the regime of $0.35 < \Delta R(e, j) < 0.6$ the jet reconstruction efficiency is reduced by energy loss to the second jet seeded by the electron. Since it seems we can trust the simulation to reproduce the physics in this region to a large extent and since the fraction of these events in an inclusive analysis is usually small, an additional removal of jets or electrons with $0.2 < \Delta R(e, j) < 0.4$ allows to reduce the potential simulation bias on the cross section to a permille level per jet.

The OR will always induce a small inefficiency in cases where a prompt electron arbitrarily overlaps with a real hadronic jet. In selections that enrich soft electrons without isolation cut together with HF jets, an additional inefficiency is introduced from the cases where an electron from a semileptonic decay within the jet is mistaken as a prompt electron and induces the removal of the HF jet [46]. This can be avoided by selecting isolated electrons and by removing only jets that overlap with an isolated electron. In this approach, a data-driven electron-fake estimate has to be parametrized properly in order to avoid shape biases due to the different OR in the fake estimate [47].

The increasing significance of boosted topologies with higher center-of-mass energies leads to an increased fraction of events with prompt electrons in the vicinity of real hadron jets, which are removed by the classic OR approaches [48]. In specific cases, a modification of the OR decision based on the 4-vector of an auxiliary jet resulting from subtraction of the 4-vector of the close-by electron provides a temporary means to avoid the efficiency loss [48].

New techniques that subtract the detector signal of a prompt electron (tracks in track jets or topo clusters in topo jets) before the jet clustering show potential to significantly increase the selection efficiency for these topologies [49, 40].

6.2.2 Recommended path for Run-2

As a default approach in Run-2, we recommend a 2-step procedure with the removal of the jet that duplicates the electron in an inner cone in $\Delta R(e, j) < 0.2$ and the removal of the electron or the entire event for events with $0.2 < \Delta R(e, j) < 0.4$. In the selection of events with prompt electrons and HF jets, we recommend to use an isolation requirement in the electron selection and to remove only jets that overlap with isolated electrons.

Removal of jets instead of events is still an option in explicit measurements of the jet activity provided that the simulation of the electron reconstruction and identification efficiency is validated as a function of the jet kinematics and that the reduced jet phase space is properly mirrored in the MC truth reference.

There are cases in which a removal of the event is preferable to the removal of the electron that fails the jet-isolation. This is in particular the case where additional background is introduced by a miss-classification of events due to the removal of prompt electrons close to real hadron jets.

Recommended further studies

The ΔR requirement for the electron-jet duplication removal should be revisited with the updated jet calibration and the change from an η based to a rapidity-based ΔR definition. The reoptimization of the ΔR regime should also aim to minimize the fraction of events with hadron jets above the analysis p_{T} threshold which are merged with the electron cluster.

The ghost association of electrons to jets might provide a more efficient tool to avoid duplication than a mere geometrical ΔR requirement and should be studied in detail in preparation for Run-2.

In view of the increasing importance of boosted topologies in Run-2, we recommend to further study techniques that subtract the detector signal of prompt electrons before the jet clustering. These techniques should be favored over the classic overlap removal and over techniques which subtract the 4-vector of the reconstructed electron from the calibrated jet momentum.

We recommend to ensure that the information needed for these studies is stored in the new Run 2 data format and propagated into the respective derivations.

6.3 Muon-jet overlap removal

The primary goal of a muon-jet OR is to (a) separate prompt muons from muons originating from the decay of hadrons within a jet and (b) to remove jets originating from FSR or brems photons. The OR is usually based on already cone-isolated muons with a low impact parameter significance (see Section 3), such that a jet-based muon isolation (a) is only an auxiliary criterion. In contrast, jets originating fom photon radiation in a muon final state (b) constitute a similar problem as the double-counting of electrons as jets (see Section 6.2) but much less frequent. It is hence desirable to efficiently separate an FSR/brems jet from a real hadron jet.

6.3.1 Discussion

The following constellations can produce a muon overlapping with a jet in the reconstructed data [50].

- pileup: pileup can accidentally produce a jet and a muon in the same area of the detector either from different bunch crossings or another pp interaction in the same bunch crossing. This combinatorial background is largely reduced by requiring that the tracks originate from the primary vertex.
- Light meson decays: muons are produced in decays of $\pi^{\pm}$ and $K^{\pm}$ which due to the relatively long lifetime can occur anywhere in the detector. The other particles produced together with the meson often form a jet in the same region. This can be reduced by requiring large p_{T}^{μ}, muon isolation and a vertex requirement. Most analyses prefer to keep the jet and reject these muons.
- HF decays: charm and beauty quarks often decays to muons. The same techniques as for light mesons can be applied here, but typically the muon momentum is larger, the secondary vertex is closer to the primary vertex and the $\Delta R(\mu, j)$ is typically smaller than for light mesons. Some analyses treat such muons as part of their signal, so whether they should be removed or not is analysis dependent.
- Final state radiation: collinear final state radiation (FSR) produces a photon very close to the muon track. The ID track is then combined with the photon energy deposit in the EM calorimeter, which is reconstructed both as an electron and as a jet. Detailed discussion of this signature follows later in this section. Most analyses prefer to keep these muons and reject the jets.
- Bremsstrahlung: Muons at very high energy will loose energy due to bremsstrahlung as they pass through matter. Since most of the matter is located in the calorimeters these muons are sometimes reconstructed as jet. The energy deposits are typically located within a cone of $\Delta R < 0.1$ around the muon track, with little to no energy spilling into the surrounding hollow cone, allowing the muons to pass calorimeter isolation criteria (see Section 3). Most analysis prefer to keep these muons and reject the jets.

The FSR/brems jets can be identified by several characteristics.

The most striking feature is that the jet consists of one or a very small amount of ID tracks. Thus, requiring that jets overlapping with isolated muons should have a minimum number of tracks (usually three) provides discriminating power between muons from LF or HF jets versus muons with FSR or bremsstrahlung. Since a small fraction of HF jets contains only one or two tracks, the HF rejection is slightly reduced. [51, 52].

Another method for identifying FSR is the EM fraction of the jet. Since the photon deposits energy in the EM calorimeter, the muon+photon appears electron-like. This method is not applicable in general for very hard muons, which can produce bremsstrahlung in the hadron calorimeter, resulting in a reduced EM fraction. The performance of this criterion is hence analysis dependent [50, 53].

The relative transverse momentum of the jet and the muon can also be used. A muon produced inside a jet will typically have lower transverse momentum than the jet itself, while the opposite is true for a jet reconstructed from a radiating muon. The performance of this criterion is again dependent on the typical muon and jet momentum distributions in the analysis [50, 52].

Bremsstrahlung from high-energetic muons could be identified using the energy deposition in the TileCal D-layer, which has been calibrated using cosmic muons. This criterion is however not applicable to FSR photons. In addition, it would imply increasing the size of the xAOD and the reconstruction time [52].

Typically a muon produced from a HF decay is found at larger distance from the jet compared to a jet reconstructed from an FSR/brems photon, where the angular distance is small. Because the opening angle depends on the boost of the mother particle, analysis dependent optimisation is required [50, 52, 53].

Once established as not being initiated by an FSR/brems photon, the jet close to the muon can be used for an additional muon isolation requirement. Studies in inclusive final states with isolated prompt muons, b-jets and light jets show that the significance of a signal selection is flat wrt to a modification of the $\Delta R(\mu, j)$ in a range $0.25 < \Delta R(\mu, j) < 0.4$. Using a cone of $\Delta R(\mu, j) < 0.4$ for the muon isolation provides a symmetry to the electron final states (see Section 6.3).

In boosted final states, with a large fraction of prompt muons close to real hadron jets, a large jet isolation cone induces substancial inefficiencies on the selection [54]. In these cases, a reduction of the jet isolation cone has been demonstrated to increase the signal significance. Both a fixed reduced OR regime of $\Delta R(\mu, j) < 0.1$ and a dynamic regime $\Delta R(\mu, j) < 0.04 + 10\ \mathrm{GeV}/\mathrm{p_T}(\mu)$ have been studied. The dynamic OR has the advantage of providing a flat efficiency over $p_T(\mu)$.

6.3.2 Recommended path for Run-2

As a baseline for inclusive analyses, we recommend to apply a muon-jet OR in a region of $\Delta R < 0.4$. The decision to keep the muon (prompt muon emitting a photon) or the jet (muon from hadron decay) should be temporarily based on the number of tracks N_{tracks} associated to the jet: The jet should be kept and the muon removed for $N_{\mathrm{tracks}} > 2$.

Depending on the analysis, we recommend to study the tightening of the selection of muon-FSR jets by including requirements on the EM-fraction (EM-fraction $>$ 0.8), the distance to the muon ($\Delta R(\mu, j) < 0.1$) or the $\sum p_T$ of the tracks associated to the jet in addition to the N_{tracks} requirement.

For boosted analyses, we recommend to reduce the OR cone to $\Delta R \lesssim 0.1$ either fixed or dynamically, provided the mis-identification of bremsstrahlung as hadron jets outside the reduced cone, as well as the mis-identification of muons from hadron decays as prompt muons are properly studied.

Recommended further studies

Considering the increased significance of boosted final states in Run-2, the various approaches of reducing the overlap removal cone (fixed or dynamic) should be studied in more detail and for a wider range of final states, in particular in view of the increased HF background and of jets created by muon bremsstrahlung. The studies should result in a baseline recommendation for boosted final states.

The discrimination of hadron jets from FSR/bremsstrahlung jets should be improved in order to maximize both the efficiency of prompt muons and the rejection of HF background. Several variables should be studied with the aim of constructing a discriminat that is indpendent of the particular kinematics of a certain final state and that is applicable to jets due to both FSR and bremsstrahlung. We recommend to study the following variables in combination with the N_{tracks} requirement:

- jet EM fraction (as a function of $p_T(\mu)$),
- $\Delta R(\mu, j)$ (as a function of $p_T(\mu)$),

- muon energy loss: $(p_{\mathrm{T}}(\mathrm{reco}) - p_{\mathrm{T}}(\mathrm{MS}))/p_{\mathrm{T}}(\mathrm{reco})$,
- $\sum p_{\mathrm{T}}(\mathrm{tracks})$ of the jet in relation to other momenta, e.g. $p_{\mathrm{T}}(\mathrm{jet})$.

In the long run, it would be desirable, instead of discarding the FSR/brems photon, to use it in order to correct the muon 4-momentum. Studies in that direction should be encouraged. As a first step, the existing software developed for this purpose in Run-1 (residing in `egammaAnalysisUtils`) should be enabled to work with the Run-2 data format.

6.4 Lepton-lepton overlap removal

6.4.1 Discussion

Duplicate muons are removed during reconstruction based on the `TrackSelectorProcessorTool`. The inefficiency for real very close-by muons has yet to be checked, in particular in view of very boosted final states.

Electron duplication is not supposed to be present with the standard commissioned electron authors (see Section 3). When combining standard electrons with soft electrons, duplication can be addressed by removing a second (softer or equally energetic) lepton that is closer than 0.05 in ΔR to the leading lepton of the same flavour or that shares a track with the leading lepton [55].

Some groups perform an electron-electron isolation in a larger cone (e.g. $\Delta R < 0.2$ or larger), followed by object or event removal, thus avoiding a potential bias in the simulation of the reconstruction efficiency for two real, close-by same-flavour leptons [56]. When using isolation or standard electron identification criteria, the latter phase space regime is only scarsely populated [56].

A duplication of a muon as an electron is possible when the muon radiates a hard photon (FSR, bremsstrahlung). In these cases, the two objects are closer than 0.01 in ΔR or share the same ID track. The criterion of a shared ID track provides a better discrimination power against pairs of real close-by muons and electrons [55, 57]. The usual approach in this case is to remove the electron. Some groups remove the event instead of the electron to protect against a bias in the reconstruction of the muon momentum in presence of hard photon radiation. This measure should not be necessary with improved FSR recovery in the muon sector (see Section 3).

The muon-electron duplication is usually accompagnied by a muon-jet duplication. An unsophsticated muon-jet isolation (see Sections 6.3) before the muon-electron removal would lead to a misclassification of the object as electron [57].

6.4.2 Recommended path for Run-2

We recommend to remove an electron if it shares a track with an identified muon.

The current muon and electron reconstruction software eliminates duplication for the comissioned object types. In case of changes in the reconstruction software we expect the CP groups to verify duplication.

Because of the lack of studies regarding the simulation of reconstruction and calibration of close-by leptons, it could be considered to remove events with close-by leptons (e.g. $\Delta R < 0.2$) depending on the aspired statistical and systematic precision and the fraction of these events in the selected phase space regime.

Some analyses want to discard muons with hard FSR/bremsstrahlung due to the possible impact on the reconstructed muon momentum. We recommend not to use the duplication of muons to electrons as a means to identify these events but to design a more reliable criterion. Ideally the FSR recovery based on the existing software should be considered.

Recommended further studies

With higher center-of-mass energies, boosted topologies with two close-by light leptons will become more prominent. In this regard we recommend to perform dedicated studies on the simulation of the reconstruction efficiency and calibration of close-by electrons and muons, as soon as the event statistics allows for a sufficiently precise assessment of the performance.

6.5 Tau-X overlap removal

6.5.1 Discussion

The OR of taus with light leptons and jets is strongly connected to the tau ID.

The overlap between taus and AntiKt4Topo jets in Run 1 can be considered an identification step rather than OR. The jet connected to a selected tau can be identified by matching in $\Delta R < 0.2$. In analyses with tau final states, jets matching to well-identified selected taus are usually removed from the collection of selected hadronic jets. There is no known use case where final states with taus and jets of different sizes are selected.

Muons can fake 1-prong taus in case of anomalous energy loss in the calorimeter, which usually also leads to a worse match between MS and ID tracks. Two approaches are employed to remove these fakes: (a) A muon veto based on cuts on tau quantities: $p_{\mathrm{T}}/E_{\mathrm{T}}$ and f_{EM}. (b) An OR within $\Delta R < 0.2$ with an identified muon (combined, $p_{\mathrm{T}} > 4$ GeV, requirements on ID track). A more aggressive OR using very loose muons (standalone or combined, $p_{\mathrm{T}} > 2$ GeV, no requirement on the ID track) can be considered if muon fakes (e.g. from $Z\rightarrow \mu\mu$) are a serious issue. The agressive OR shows inefficiencies for very hard taus due to punch-through pions that give rise to standalone muon candidates. In these cases, the default OR can recover the efficiency loss.

Electrons can fake 1-prong and sometimes 3-prong taus. A multivariate electron veto: EleBDT-Loose/Medium/Tight, based on tau properties has been designed to remove these fakes. Usually, the EleBDTMedium working point is used.

The electron veto is combined with an OR in $\Delta R < 0.2$ to a loosely identified electron. The electron ID used for the selection of these electrons has progressed from a cut-based `LoosePP` to a likelihood-based `VeryLooseLH` with better performance.

6.5.2 Recommended path for Run-2

In the following, we specify base line recommendations which have been optimized based on the Run 1 tau reconstruction. With the change in Run 2 to a tau-substructure based method [58], we recommend to revisit these recommendations. We expect though no major surprises, since the current OR regimes are already quite conservative in design.

In analyses with tau final states, we recommend to remove the jet that overlaps in $\Delta R < 0.2$ with an identified tau.

We recommend to remove a tau if it overlaps in $\Delta R < 0.2$ with an identified muon or if it fails the muon veto.

We recommend to remove a tau candidate if it is flagged by the EleBDTMedium or if it overlaps in $\Delta R < 0.2$ with an electron that passes the likelihood-based `VeryLooseLH` ID criterion.

With these tau-ID related procedures in place, a classic overlap removal with tigher muons or electrons used in the analysis selection is obsolete.

6.5.3 Recommendation for specific cases

For selections of soft/medium tau leptons ($p_{\mathrm{T}} < 100 - 150$GeV) and a large expected background from processes with muon final states, we recommend to consider dropping the muon-veto and applying a more aggressive OR in $\Delta R < 0.2$ with a combined or standalone muon, $p_{\mathrm{T}} > 2$ GeV, no ID track requirement.

6.6 Photon-X overlap removal

6.6.1 Discussion

Most analyses are using a photon isolation cone (see Section 3) of $\Delta R < 0.4$. Consequently they perform an OR within that cone to any other object (including other photons) according to the hierarchy: e/μ $\rightarrow$ photon $\rightarrow$ jet.

6.6.2 Recommended path for Run2

As a base line, we recommend to perform an OR within a cone of $\Delta R < 0.4$ to any other object and to other photons, with the following hierarchy:
e/μ $\rightarrow$ photon $\rightarrow$ jet.

We recommend to optimize these approaches based on more detailed studies in various final states with photons. In particular, we recommend to pursue the harmonization of the photon-jet OR and the electron-jet OR.

6.7 Overlap removal sequence

The OR should be applied between objects selected in the analysis. The recommended OR sequence depends on these objects. We recommend the following hierarchical sequence:

1.a) Tau ID and OR with specific loose electrons and muons

1.b) ID and isolation of electrons and muons

2.) OR muon $\longleftrightarrow$ electron

3.) OR lepton $\longleftrightarrow$ photon

4.) OR lepton, photon $\longleftrightarrow$ jet

References

[1] https://indico.cern.ch/event/271433/.

[2] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AnalysisHarmonisation.

[3] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Geneva, Sep, 2010.

[4] ATLAS Collaboration Collaboration, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012.

[5] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AnalysisHarmonisationGroup1.

[6] ATLAS Collaboration Collaboration, ATLAS Collaboration, New J. Phys. **13** no. arXiv:1012.5104. CERN-PH-EP-2010-079, (2010) 053033. 70 p. Comments: 57 pages plus author list (82 pages total), 19 figures (85 independent .eps files), 10 tables.

[7] S. Pagan Griso, K. Prokofiev, A. Andreazza, K. Grimm, E. Guido, F. Meloni, M. Rudolph, A. Salzburger, and A. Wildauer, Tech. Rep. ATL-COM-PHYS-2012-561, CERN, Geneva, May, 2012.

[8] S. Pagan Griso, K. Prokofiev, A. Andreazza, K. Grimm, E. Guido, F. Meloni, M. Rudolph, A. Salzburger, and A. Wildauer, Tech. Rep. ATL-COM-PHYS-2012-474, CERN, Geneva, Apr, 2012.

[9] D. Adams, P. Delsart, M. Elsing, K. Koeneke, E. Lancon, W. Lavrijsen, S. Strandberg, W. Verkerke, I. Vivarelli, and M. Woudstra, Tech. Rep. ATL-COM-SOFT-2013-125, CERN, Geneva, Dec, 2014.

[10] P. Nef and A. Schwartzman, Tech. Rep. ATLAS-COM-CONF-2014-025, Apr, 2014.

[11] A. Morley,. https://indico.cern.ch/event/315625/contribution/3/material/slides/0.pdf.

[12] egamma CP group, Tech. Rep. ATL-COM-PHYS-2013-1287, Sept, 2013.

[13] https://indico.cern.ch/event/298645.

[14] https://indico.cern.ch/event/163461.

[15] ATLAS Collaboration, submitted to Eur. Phys. J. (2014), `arXiv:1404.2240 [hep-ex]`.

[16] ATLAS Collaboration, JHEP **02** (2014), `arXiv:1311.6724 [hep-ex]`.

[17] J. Catmore, M. Elsing, E. Lipeles, D. Rousseau, and I. Vivarelli, Tech. Rep. ATL-COM-SOFT-2013-005, CERN, Geneva, Apr, 2013.

[18] W. Verkerke, I. Vivarelli, and N. Krumnack, Tech. Rep. ATL-COM-SOFT-2014-005, CERN, Geneva, Mar, 2014.

[19] P. J. Laycock and A. Krasznahorkay, Tech. Rep. ATL-COM-SOFT-2014-022, CERN, Geneva, Apr, 2014.

[20] C. Young,. https://indico.cern.ch/event/281979/contribution/4/material/slides/0.pdf.

[21] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804** (2008) 005, arXiv:0802.1188 [hep-ph].

[22] H. Reisin Carretero, D. Miller, M. Swiatlowski, G. Otero y Garzon, R. Piegaia, A. Schwartzman, and T. Carli, Tech. Rep. ATL-COM-GEN-2012-013, CERN, Geneva, Jul, 2012.

[23] ATLAS Collaboration Collaboration, G. Aad et al., Eur.Phys.J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[24] https://indico.cern.ch/event/300511/.

[25] H. HSG3, Tech. Rep. ATL-COM-PHYS-2013-1504, CERN, Geneva, Nov, 2013.

[26] P. Nef and A. Schwartzman, Tech. Rep. ATLAS-COM-CONF-2014-025, CERN, Geneva, Apr, 2014.

[27] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets.

[28] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011.

[29] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2012.

[30] S. Batista, R. Camacho, D. DeMarco, S. Gupta, and J. Taenzer,. https://indico.cern.ch/event/293406/contribution/2/material/slides/0.pdf.

[31] G. Aad, B. Allbrooke, D. Buescher, A. Buckley, D. Cinca, Y. Coadou, P. Conde Muo, I. Connelly, B. Cooper, A. Davison, C. Debenedetti, Y. Enari, G. Facini, S. Fracchia, P. Francavilla, G. Gaycken, V. Giangobbe, R. Goncalo, G. Gonzalez, H. Gray, J. Grivaz, C. Gwilliam, S. Hagebck, G. Halladjian, M. Jackson, D. Jamin, K. Kiuchi, V. Kostyukhin, J. Lee, W. Lockman, K. Lohwasser, D. Lopez Mateos, L. Ma, A. Maio, J. Maneira, M. Martinez, U. Mallik, A. Mehta, K. Mercurio, K. Mochizuki, N. Morange, Y. Ming, Y. Nagai, J. Nielsen, I. Ochoa, H. Otono, G. Piacquadio, E. Pinto, M. Proissl, M. Sanders, T. Scanlon, B. Smart, P. Sommer, V. Sorin, J. Therhaag, J. Thomas-Wilsker, P. Thompson, L. Vacavant, J. Wang, C. Wang, S. Wang, C. Weiser, R. Zaidan, and L. Zhang, Tech. Rep. ATL-COM-PHYS-2013-465, CERN, Geneva, Apr, 2013.

[32] ATLAS Collaboration Collaboration, G. Aad et al., Phys.Rev. **D86** (2012) 014022, arXiv:1112.6297 [hep-ex].

[33] ATLAS Collaboration Collaboration, G. Aad et al., arXiv:1312.3524 [hep-ex].

[34] https://indico.cern.ch/event/276171/contribution/2/material/slides/0.pdf.

[35] P. Loch, T. Carli, A. Schwartzman, and M. Begel, Tech. Rep. ATL-COM-PHYS-2013-161, CERN, Geneva, Feb, 2013. On behalf of the jet/etmiss group.

[36] ATLAS Collaboration Collaboration, G. Aad et al., JHEP **1309** (2013) 076, arXiv:1306.4945 [hep-ex].

[37] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties.

[38] Tech. Rep. ATLAS-CONF-2013-108, CERN, Geneva, Nov, 2013.

[39] https://indico.cern.ch/event/300484/.

[40] S. Grinstein, M. Martinez, E. Meoni, C. Ochando, and E. Perez, Tech. Rep. ATL-COM-PHYS-2010-021, Dec, 2009.

[41] https://indico.cern.ch/event/221981/material/slides/0?contribId=4.

[42] https://indico.cern.ch/event/77358/session/0/contribution/1/material/slides/0.pdf.

[43] https://indico.cern.ch/event/77358/session/0/contribution/3/material/3/0.pdf.

[44] https://indico.cern.ch/event/276702/contribution/1/material/slides/0.pdf.

[45] https://indico.cern.ch/event/276706/material/1/0?contribId=1.

[46] https://indico.cern.ch/event/301464/contribution/3/material/slides/0.pdf.

[47] K. Becker, T. Cornelissen, F. Derue, A. Henrichs, D. Hirschbuehl, X. Lei, O. Nackenhorst, F. O'Grady, D. Pelikan, M. Pinamonti, S. Pires, J. Sjoelin, and P. Tepel, Tech. Rep. ATL-COM-PHYS-2013-1100, Aug, 2013.

[48] https://indico.cern.ch/event/310262/contribution/3/material/slides/0.pdf.

[49] https://indico.cern.ch/event/277776/material/slides/0?contribId=3.

[50] https://indico.cern.ch/event/303658/contribution/3/material/slides/0.pdf.

[51] https://indico.cern.ch/event/226208/material/slides/0?contribId=5.

[52] https://indico.cern.ch/event/137299/contribution/0/material/slides/0.pdf.

[53] https://indico.cern.ch/event/298322/contribution/3/material/slides/0.pdf.

[54] https://indico.cern.ch/event/267390/session/0/contribution/1/material/slides/0.pdf.

[55] https://indico.cern.ch/event/310262/contribution/7/material/slides/1.pdf.

[56] https://indico.cern.ch/event/311383/contribution/9/material/slides/0.pdf.

[57] https://indico.cern.ch/event/102781/material/slides/0?contribId=2.

[58] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TauSubstructure.