Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: Measurements and searches performed with the ATLAS detector at the CERN LHC often involve signatures with one or more prompt leptons. Such analyses are subject to ‘fake/non-prompt’ lepton backgrounds, where either a hadron or a lepton from a hadron decay or an electron from a photon conversion satisfies the prompt-lepton selection criteria. These backgrounds often arise within a hadronic jet because of particle decays in the showering process, particle misidentification or particle interactions with the detector material. As it is challenging to model these processes with high accuracy in simulation, their estimation typically uses data-driven methods. Three methods for carrying out this estimation are described, along with their implementation in ATLAS and their performance.

KEYWORDS: Analysis and statistical methods; Particle identification methods

ArXiv ePrint: 2211.16178
Contents

1 Introduction 1

2 Methods 1
   2.1 Matrix method 2
      2.1.1 Asymptotic matrix method 3
      2.1.2 Poisson likelihood matrix method 3
   2.2 Fake-factor method 4
   2.3 Generalisation for multi-lepton final states 5
   2.4 Use with weighted events 6
   2.5 Performance studies 7

3 The ATLAS detector 10

4 Lepton selection criteria 11
   4.1 Electron reconstruction and identification 11
   4.2 Muon reconstruction and identification 12
   4.3 Lepton isolation 12
   4.4 Removing overlaps between jets and leptons 14

5 Monte Carlo simulation samples 14

6 Sources of fake/non-prompt leptons 15

7 Measurement of real and fake/non-prompt lepton efficiencies 18
   7.1 Real-lepton efficiencies 19
   7.2 Fake/non-prompt lepton efficiencies 19

8 Systematic uncertainties 21
   8.1 Statistical uncertainties in the measured efficiencies 22
   8.2 Systematic uncertainties in the measured efficiencies 22
   8.3 Uncertainties in the modelling of real-lepton processes 23
   8.4 Uncertainties due to biases in the Poisson likelihood matrix method 23

9 Examples of application in ATLAS analyses 23
   9.1 Measurement of the $t\bar{t}Z$ cross-section in final states with three or four leptons 23
      9.1.1 Real-lepton efficiencies 24
      9.1.2 Fake/non-prompt lepton efficiencies 24
      9.1.3 Results in the fake/non-prompt lepton validation regions 27
   9.2 Model-independent search for new phenomena in multi-lepton final states 29
      9.2.1 Fake/non-prompt lepton selection 30
      9.2.2 Fake factors 31
      9.2.3 Validation regions 33

– i –
1 Introduction

Many measurements and searches for new physics performed with the ATLAS detector [1] at the Large Hadron Collider (LHC) at CERN require the presence of one or more leptons (electrons or muons) to indicate that a high-energy electroweak process occurred in the collision. Lepton candidates are reconstructed from signals in the inner tracker, calorimeters, and muon spectrometer [2, 3]. Identification criteria are then applied to suppress candidates originating from physical objects other than leptons from the hard-scattering process of the event. These background candidates fall into two categories: i) “non-prompt leptons” from the semileptonic decay of hadrons, or from photon (γ) conversions in detector material, and ii) “fake leptons” where the reconstructed object is not, in fact, due to a lepton. In contrast to the aforementioned categories, “real leptons” are defined as electrons or muons produced either directly in the hard-scattering process or directly in the decay of a short-lived non-hadronic resonance (such as a W/Z boson).

The rates at which fake or non-prompt leptons are selected are difficult to model accurately from simulation. They can depend strongly on details of the physics simulation, including in non-perturbative regions where the simulation would not be expected to be reliable. They also depend on the modelling of the material composition and response of the detector. In addition, fake leptons sometimes occur with low probability as a result of misidentifying a hadronic jet in multi-jet events. The computing resources required to simulate these processes with a sufficient sample size would be prohibitive. Therefore, “data-driven” approaches are commonly used to estimate these backgrounds.

To simplify the adoption of such methods, and to ensure that they are applied uniformly, a set of standard tools and prescriptions has been developed for use in ATLAS physics analyses that are subject to fake/non-prompt lepton backgrounds. The principles and performance of these tools are described in this paper. The motivation and mathematical basis of the methods are explained in section 2; a description of the relevant features of the ATLAS detector is given in section 3; the criteria used to select leptons are given in section 4; the simulated signal and background processes, as well as the different processes that can lead to fake/non-prompt leptons, are discussed in sections 5 and 6; the procedures used to measure the efficiencies for real and fake/non-prompt leptons are described in section 7; the systematic uncertainties associated with the methods are described in section 8; and section 9 provides examples of the application of the fake/non-prompt lepton estimation methods in two published ATLAS physics analyses, with details that are not included in the existing publications.

2 Methods

The fake/non-prompt lepton estimation methods considered in this paper depend on defining two tiers of lepton selection criteria, called the “baseline” and “tight” criteria. The tight criteria are
used to select the signal leptons in a physics analysis, while the baseline criteria accept all of the
tight lepton candidates as well as an additional set of candidates with a higher rate of fake/non-
prompt contributions. Candidates that satisfy the baseline criteria but not the tight criteria are
called “loose” leptons. If the two sets of criteria are chosen well, the fraction of real leptons in
the baseline sample that satisfy the tight criteria will be substantially higher than the corresponding
fraction for fake/non-prompt leptons. These fractions are called the “real efficiency” ($\varepsilon_r$) and “fake
efficiency” ($\varepsilon_f$), respectively.

The $\varepsilon_r$ values are generally taken from Monte Carlo (MC) simulated events that are corrected to
account for differences between data and the simulation, while the $\varepsilon_f$ values are typically measured
in a data sample consisting of different events from the ones that are used for the data analysis, as
detailed in section 7.

2.1 Matrix method

With the efficiencies known, a simple counting of the numbers of lepton candidates that satisfy
the tight and loose criteria provides an estimate for the number of fake/non-prompt leptons. In the
simplest case, where an analysis selects signal events containing exactly one tight lepton candidate
and no loose lepton candidates, the relationship between the numbers of tight and loose leptons
observed in data and the composition of the sample in terms of real and fake/non-prompt leptons is:

$$\begin{pmatrix} N^t \\ N^l \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ 1 - \varepsilon_r & 1 - \varepsilon_f \end{pmatrix} \begin{pmatrix} N^b_t \\ N^b_l \end{pmatrix},$$

(2.1)

where $N^t$ and $N^l$ are the numbers of events with tight and loose lepton candidates, and $N^b_t$ and $N^b_l$
are the unknown numbers of real and fake/non-prompt leptons in the baseline sample. In matrix
notation, the relationship is given by:

$$N_t = \mathbf{M} \varepsilon_{NL},$$

(2.2)

The fact that the unknown values ($N^b_t$ and $N^b_l$) and the observed yields ($N^t$ and $N^l$) are related
via the matrix $\mathbf{M}$ gives rise to the name of this method: the “matrix method” [4]. Inversion of the
matrix allows $N^b_t$ to be determined:

$$N^b_t = \frac{1}{\varepsilon_r - \varepsilon_f} \left[ (\varepsilon_r - 1)N^l + \varepsilon_r N^l \right].$$

(2.3)

In the typical use case, the quantity of interest is the number of events in the tight sample where
the lepton is fake/non-prompt, $N^t_f$. This is related to the number of such events in the baseline
sample, $N^b_t$, by:

$$N^t_f = \varepsilon_f N^b_t.$$

(2.4)

Similarly, the number of real leptons in the tight sample is

$$N^t_r = \varepsilon_r N^b_t,$$

(2.5)

and these can be treated as elements of a column matrix $N^t_{rf}$.

The fact that $N^t$ appears in eq. (2.1) means that information about the content of the analysis
signal region is used in the estimate, and therefore an analysis is not completely blinded when using
this approach.
Generally, the values of $\varepsilon_r$ and $\varepsilon_f$ depend on factors such as the lepton candidate’s momentum or its proximity to other objects. Details of how these variations are accounted for in the estimation are given below.

### 2.1.1 Asymptotic matrix method

In this method, events in the baseline sample are considered one at a time, and a “fake weight” $w_i$ is defined for each event, corresponding to the two terms in eq. (2.3) via eq. (2.4):

$$w_i = \begin{cases} \frac{\varepsilon_{rt}}{\varepsilon_{rt} - \varepsilon_{rt}}(\varepsilon_{rt} - 1) & \text{if lepton candidate } i \text{ is tight (so } N^t = 1 \text{ and } N^l = 0) \\ \frac{\varepsilon_{rt}}{\varepsilon_{rt} - \varepsilon_{rt}} & \text{if lepton candidate } i \text{ is loose (so } N^t = 0 \text{ and } N^l = 1) \end{cases},$$

where $\varepsilon_{rt}$ and $\varepsilon_{rt}$ are the values of $\varepsilon_r$ and $\varepsilon_f$ that are appropriate for lepton $i$. Since $\varepsilon_{rt}$ is always less than one, the weight for an event with a tight lepton candidate is negative. By extension, the total fake/non-prompt lepton background in the tight sample is then estimated by

$$N^t_f = \sum_{\text{events}} w_i.$$

This approach is convenient since the $w_i$ need only be calculated once and then can be stored with the event, allowing the distribution of the fake-lepton yield to be binned in any variable of interest, as well as a simple re-computation of $N^t_f$ if the event selection is modified. One drawback is that since the value of $w_i$ may be negative, there is no guarantee that $N^t_f$ will be positive. The value of $N^t_f$ is also sensitive to fluctuations in the input $\varepsilon_{rt}$ and $\varepsilon_{rt}$ values.

The statistical uncertainty of $N^t_f$ is given by:

$$\sigma_{N^t_f} = \sqrt{\sum_{\text{events}} w_i^2}. \quad (2.6)$$

The method that makes use of the $w_i$ is known as the “asymptotic matrix method”, since eq. (2.6) is only valid in the asymptotic limit with a large number of events.

### 2.1.2 Poisson likelihood matrix method

In this method [5] the elements of $N^t_f$ are treated as free parameters, which are varied to maximise the likelihood of the observed $N_{tl}$ values. By doing so, the Poisson-distributed nature of the $N_{tl}$ values is taken into account, so there is no need to use the asymptotic approximation. In the fit, the $N^t_{tl}$ values are converted to $N^b_{tl}$ using eqs. (2.4) and (2.5), where the entries in the matrix $M_e$ are calculated using the averages of the prompt and fake/non-prompt lepton efficiencies in the baseline sample:

$$M_e = \left( \begin{array}{c} \langle e_r \rangle \\ 1 - \langle e_r \rangle \end{array} \left( \begin{array}{c} \langle e_f \rangle \\ 1 - \langle e_f \rangle \end{array} \right) \right). \quad (2.7)$$

An earlier variant of the likelihood matrix method is described in ref. [6].
The resulting \( N_{\text{rb}} \) values are used to obtain the expectation values for \( N_t \) using eq. (2.2). These expectation values are denoted by \( N_{\text{rl}, \text{exp}} \). The \( N_{\text{rb}} \) parameters are adjusted (subject to the constraint that they must be non-negative) to maximise the joint Poisson likelihood

\[
L(N_{\text{rb}}) \equiv \prod_i P[N_{\text{rl}}|N_{\text{rl}, \text{exp}}, (N_{\text{rb}})],
\]

where the product is over the elements of \( N_{\text{rl}} \); \( N_{\text{rl}} \) and \( N_{\text{rl}, \text{exp}} \) are the \( i \)th elements of \( N_{\text{rl}} \) and \( N_{\text{rl}, \text{exp}} \), respectively; and \( P[n|\mu] \) is the Poisson probability for observing \( n \) events when \( \mu \) are expected. The output of the fit consists of an estimate of the number of fake/non-prompt leptons and the uncertainty in this quantity, which is obtained by noting the value for which \(-\ln L\) exceeds its minimum by 0.5.

The primary advantages of the Poisson likelihood approach are that the result is constrained to be non-negative, and the uncertainty is a better approximation to the range that gives 68% coverage, particularly in samples with few events. In addition, in some scenarios it can provide smaller statistical uncertainties than the asymptotic matrix method or the fake-factor method (described in section 2.2). The main drawback is that the estimated yield must be calculated for the sample as a whole rather than from a sum of individual event weights, which complicates the process of producing a distribution of the fake-lepton yield, as required for any differential measurement (to do so, the likelihood must be applied in every bin of the distribution).

### 2.2 Fake-factor method

The fact that real-lepton kinematic distributions and efficiencies are generally modelled well in simulation, and that scale factors can be applied to account for any differences observed between the values in simulation and in data control samples, leads to an alternative method that uses simulation, rather than the data, to measure the real-lepton contribution to the loose lepton sample.

The number of fake-lepton events in the loose sample is

\[
N_{l}^{f} = N_{l}^{l} - N_{l}^{r} = (1 - \varepsilon_{l})N_{l}^{b}
\]

and the number of fake-lepton events in the tight sample is

\[
N_{t}^{f} = \varepsilon_{t}N_{l}^{b} = F(N_{l}^{l} - N_{l}^{l}),
\]

where the “fake factor” \( F \) is defined as \( F \equiv \varepsilon_{l}/(1 - \varepsilon_{l}) \). Thus, in the fake-factor method, the number of tight fake/non-prompt leptons for a given analysis can be computed using the fake factor \( F \), the total number of loose lepton candidates \( N_{l}^{l} \), and the number of real leptons in the loose lepton sample \( N_{l}^{l} \), where the latter quantity can be estimated using MC simulated samples, and its contribution subsequently subtracted from the quantity \( FN_{l}^{l} \) observed in the data. In practice, the calculation is performed on an event-by-event basis to account for potential variations in \( F \) due to properties of the lepton:

\[
N_{l}^{f} = \sum_{\text{data}, i=1}^{N_{l}^{l}} F_{i} - \sum_{\text{MC}, j=1}^{N_{\text{MC}}^{l}} w_{\text{MC}, j} F_{j},
\]

where \( F_{i} \) is the fake factor appropriate for lepton \( i \), all sources of prompt leptons are considered in the sum over MC simulated events, \( N_{\text{MC}}^{l} \) is the number of MC events in the loose sample, and \( w_{\text{MC}, j} \)
is the weight assigned to simulated event $j$, based on the cross-section of the simulated process and any corrections to the selection efficiency that may be needed to reflect the performance on data events.

The main advantage of the fake-factor method is that this result does not depend on $N^t$, i.e. the yield in the analysis signal region. This means that unlike the matrix method, the fake-factor method can be applied while remaining “blind” to the contents of the signal region. However, the method does have some of the same drawbacks as the asymptotic matrix method, namely the possibility of $N^t_f$ being negative, and sensitivity to fluctuations in the $\epsilon_f$ values.

### 2.3 Generalisation for multi-lepton final states

The above methods can be generalised to cases where multiple baseline lepton candidates are considered in each event. For the matrix methods, this is done by increasing the dimensionality of $M_\ell$, $N_{t\ell}$, and $N_{t\ell}^{t(b)}$ to $2^{n_b}$, where $n_b$ is the number of baseline lepton candidates in each event. The estimated fake/non-prompt yield depends on the requirements of a particular analysis in three ways: first, from the requirement placed on the desired number $n_t$ of tight lepton candidates per event; second, whether or not events with additional loose lepton candidates are vetoed; and third, on the minimum number of fake/non-prompt leptons $m_f$ defining the background to be evaluated with one of the data-driven methods.\(^2\) The consideration of $m_f$ is reflected in the transition from the number of fake/non-prompt lepton events in the baseline sample to the number in the tight sample, in a generalisation of eq. (2.4):

$$N^t_f = \sum_i g_i (\epsilon_r, \epsilon_f) N_{t\ell}^{b_i}. \quad (2.9)$$

Here, the sum is over all combinations of real and fake/non-prompt leptons that include at least $m_f$ fake/non-prompt leptons, $g_i$ is a function of the real and fake/non-prompt efficiencies that will result in the required number of tight lepton candidates for a given set of real and fake/non-prompt leptons, and $N_{t\ell}^{b_i}$ is the $i$th element of $N_{t\ell}^b$.

To address the requirements on $n_t$ and the possible presence of additional loose lepton candidates, the analysis must consider all events in the baseline sample with lepton multiplicities up to the sum of the allowed numbers of tight and loose lepton candidates in the signal region.

As an example, consider the case where $n_b = 2$. If an analysis selects signal events containing exactly two tight lepton candidates and no loose lepton candidates, and the background with $m_f \geq 1$ is being evaluated, then

$$N^t_f = \epsilon_{t1} \epsilon_{t2} N_{t1}^{b_1} b_2 + \epsilon_{t1} \epsilon_{t2} b_1 N_{t2}^{b_2} + \epsilon_{t1} \epsilon_{t2} b_1 b_2 N_{t1}^{b_2},$$

where $N_{t1}^{b_1}$ is the number of events in the baseline sample where the first lepton candidate is real and the second is fake/non-prompt, and $N_{t2}^{b_2}$ and $N_{t1}^{b_2}$ are defined correspondingly. The ordering of the lepton candidates is typically according to $p_T$, but the method does not depend on the ordering used.

\(^2\)An example of a case where $m_f$ is greater than one would be a dilepton analysis where there are backgrounds from both $W(\rightarrow \ell \nu) + j$ events where the jet forms a fake/non-prompt lepton, and dijet events where both jets form fake/non-prompt leptons. The analysers may choose to estimate the first contribution from MC simulation, and use data-driven methods for the second contribution, and therefore setting $m_f = 2$ for the data-driven approach is required to avoid double-counting.
Another example may be an analysis accepting dilepton events \((n_b = 2)\) with one tight lepton and another baseline (not necessarily tight) lepton. If, again, the background being evaluated consists of events with at least one fake lepton \((m_f \geq 1)\), the coefficients \(g_i\) in eq. (2.9) change such that the expression becomes:

\[
N^t_i = (\epsilon_{s1} + \epsilon_{s2} - \epsilon_{s1} \epsilon_{s2}) N^b_{t1t2} + (\epsilon_{s1} + \epsilon_{s2} - \epsilon_{s1} \epsilon_{s2}) N^b_{t1t2},
\]

where the additional terms are needed to account for the additional ways that fake/non-prompt leptons might satisfy the signal selection (and the terms involving products of the efficiencies are subtracted to avoid double-counting). For an analysis that imposes the same lepton candidate requirements but only the background with \(m_f = 2\) is being evaluated, the expression is:

\[
N^t_i = (\epsilon_{s1} + \epsilon_{s2} - \epsilon_{s1} \epsilon_{s2}) N^b_{t1t2},
\]

The methods can also be extended to cases where there are more than two levels of lepton selection criteria, or distinct categories of fake/non-prompt leptons, as in ref. [7].

As with the matrix method, the fake-factor method can be generalised to higher lepton candidate multiplicities. In the dilepton final state, the number of events with two tight lepton candidates, of which at least one is fake/non-prompt, is:

\[
N^t_i = N^{t1t2} - N^{t1t2} + F_1 (N^{t1t2} - N^{t1t2}) + F_2 (N^{t1t2} - N^{t1t2}) - F_1 F_2 (N^{t1t2} - N^{t1t2}),
\]

where \(N^{t1t2}\) is the number of events where both lepton candidates are real leptons, and \(F_1\) and \(F_2\) correspond to the fake factors associated with the first and second lepton candidate, respectively.

However, the algebraic simplification that leads to eq. (2.10), where the result depends simply on products of the fake factors and the observed tight and loose lepton candidate yields, with a correction term that depends only on events where all the leptons are real, does not hold for all possible event selections nor all values of \(m_f\); such a simplification is restricted to cases where the baseline and tight candidate lepton multiplicities are the same in all events, and where \(m_f = 1\).

### 2.4 Use with weighted events

In some cases, such as for self-consistency tests using simulated events, it may be advantageous to weight the events that are input to the fake/non-prompt background estimate. This is straightforward for the fake-factor and asymptotic matrix methods, since the weight returned by the method for each event can be multiplied by the event weight \(w_{\text{evt}}\). For the likelihood matrix method, this is handled by using the scaled Poisson distribution [8], in which the values of \(N^i_{\text{obs}}\) and \(N^i_{\text{evt}}\) from eq. (2.8) are scaled according to the event weights \(w_{\text{evt}}\):

\[
s_i \equiv \frac{\sum_{i=1}^{N_{\text{obs}}} w_{\text{evt}}}{\sum_{i=1}^{N_{\text{obs}}} w_{\text{evt}}^2},
\]

so that the likelihood becomes

\[
L (N^i_{\text{obs}}) = \prod_i P \left[ N^i_{\text{obs}} / s_i | N^i_{\text{evt}} \right] \left( N^i_{\text{obs}} / s_i \right)
\]
2.5 Performance studies

MC simulations of experiments are utilised to assess the statistical performance of the methods. These simulations consist of pseudoexperiments that mimic scenarios that might occur in an actual analysis. Pseudoexperiments with sample sizes of 10 or 1000 dilepton events in the baseline sample are considered. The fraction of fake/non-prompt leptons in the baseline lepton sample is varied for each pseudoexperiment, with a uniform distribution between 0 and 100%. The values of $\varepsilon_t$ and $\varepsilon_f$ for each lepton are drawn from Gaussian distributions with specified means and widths (limits are imposed such that the values are always between zero and one, and $\varepsilon_f$ is always at least 10% less than $\varepsilon_t$). Each lepton is randomly assigned as fake/non-prompt or real, in accord with the fraction of fake/non-prompt leptons assumed for the pseudoexperiment. Then each lepton is judged to either meet or fail to meet the tight selection criteria, based on whether or not it is a real lepton and the values of $\varepsilon_t$ and $\varepsilon_f$ assigned to it. The set of simulated leptons is input to the data-driven algorithms, and the estimated fake yield and its statistical uncertainty are determined for each pseudoexperiment. These values are then compared with the expectation value for the number of fake/non-prompt lepton events in each pseudoexperiment, which is determined by the numbers of real and fake/non-prompt baseline leptons, the values of $\varepsilon_t$ and $\varepsilon_f$ for each lepton, and the value of $m_f$ for the simulated analysis. For example, in a dilepton sample with $m_f \geq 1$, the expectation value is:

$$\langle N^t_f \rangle = \sum_{rf \text{ events}} \varepsilon_{t,i} \varepsilon_{t,f,i} + \sum_{fr \text{ events}} \varepsilon_{r,i} \varepsilon_{t,f,i} + \sum_{ff \text{ events}} \varepsilon_{f,i} \varepsilon_{t,f,i}.$$ 

For the Poisson likelihood matrix method, the likelihood maximisation is implemented using the MINUIT function minimisation package [9] to minimise the negative log likelihood. The interface to MINUIT is provided by the TMinuit class in ROOT [10].

The fake-factor method requires a two-step process, where the contribution from real-lepton events is subtracted in the second step. In an actual physics analysis, this subtraction is done using MC simulation of the prompt lepton contribution. For the simple MC simulation used here, the second step is modelled by running each pseudoexperiment a second time with the same parameters but a statistically independent sample of events, and running the fake-factor method only on the events that do not have fake/non-prompt leptons. The second sample has ten times more events than the primary sample, to be consistent with the usual case where the MC simulation sample has a multiple of the number of events in the data sample. The result of this second run is then scaled down by a factor of ten and subtracted from the result when using the initial set of simulated events.

As an initial example, one can investigate the performance for dilepton events under conditions that are favourable for estimating the fake/non-prompt lepton background. This means that the samples are large (1000 events per pseudoexperiment), and the values of $\varepsilon_t$ and $\varepsilon_f$ are on average well-separated (here, $\langle \varepsilon_t \rangle = 0.90$ and $\langle \varepsilon_f \rangle = 0.10$). The ratios of the estimated to true fake yields are shown versus $\langle N^t_f \rangle$ in figure 1(a). The average statistical uncertainties in the estimates for each method are shown in figure 1(b), and the fraction of pseudoexperiments in which the true fake yield lies within the uncertainty reported for each method is shown in figure 1(c). The performance of the methods for dilepton analyses with low statistical precision (10 events per pseudoexperiment) are shown in figure 2. Finally, to represent a more challenging situation, the case where there is less separation between the values of $\varepsilon_t$ and $\varepsilon_f$ (due, for example, to the application of stricter online
Figure 1. Performance of the three methods in pseudoexperiments with dilepton events where the leptons had an average $\epsilon_r = 0.9$ and $\epsilon_f = 0.1$, both values were varied according to a Gaussian distribution of width 0.1 when simulating each lepton, and there were 1000 events per pseudoexperiment. The quantity $\langle N_t \rangle$ is the expectation value for the number of fake/non-prompt lepton events in each pseudoexperiment. Plot (a) shows the ratio of the estimated to expected fake-lepton yields, (b) shows the absolute uncertainty estimate for each method (for the Poisson likelihood method the average of the upward and downward uncertainties is taken), and (c) shows the fraction of pseudoexperiments where the true fake yield lies within the reported one standard deviation ($1\sigma$) range.

Figure 2. Performance of the three methods in pseudoexperiments with dilepton events where the leptons had an average $\epsilon_r = 0.9$ and $\epsilon_f = 0.1$, both values were varied according to a Gaussian distribution of width 0.1 when simulating each lepton, and there were 10 events per pseudoexperiment. The quantity $\langle N_t \rangle$ is the expectation value for the number of fake/non-prompt lepton events in each pseudoexperiment. Plot (a) shows the ratio of the estimated to expected fake-lepton yields, (b) shows the absolute uncertainty estimate for each method (for the Poisson likelihood method the average of the upward and downward uncertainties is taken), and (c) shows the fraction of pseudoexperiments where the true fake yield lies within the reported one standard deviation ($1\sigma$) range.

lepton selection criteria that might be required when the LHC runs at higher luminosities) is also explored. Figure 3 shows the results when $\langle \epsilon_r \rangle = 0.70$ and $\langle \epsilon_f \rangle = 0.30$, and there are 10 events per pseudoexperiment.

These studies show that all three methods give accurate estimates, with nearly equivalent performance, in high-statistics samples with a large separation between $\epsilon_r$ and $\epsilon_f$ (as shown in
Figure 3. Performance of the three methods in pseudoexperiments with dilepton events where the leptons had an average $\varepsilon_t = 0.7$ and $\varepsilon_f = 0.3$, both values were varied according to a Gaussian distribution of width 0.1 when simulating each lepton, and there were 10 events per pseudoexperiment. The quantity $\langle N_{tf} \rangle$ is the expectation value for the number of fake/non-prompt lepton events in each pseudoexperiment. Plot (a) shows the ratio of the estimated to expected fake-lepton yields, (b) shows the absolute uncertainty estimate for each method (for the Poisson likelihood method the average of the upward and downward uncertainties is taken), and (c) shows the fraction of pseudoexperiments where the true fake yield lies within the reported one standard deviation ($1\sigma$) range.

One notable feature is the dip in the uncertainty values for all three methods near $\langle N_{tf} \rangle = 10$ in figure 1(b). This occurs because there are two ways for the model to produce an expectation around 10: either a very low fake fraction in the baseline to start with, or a very large fake fraction so that there are few real leptons and most of the background is from events with two fake/non-prompt leptons, which gives a minimum value of $\varepsilon_f^2 \cdot 1000 = 10$ when $\varepsilon_f = 0.10$. These two processes will have very different uncertainties. There is also a bias in the Poisson likelihood matrix method toward low values when the true number of fakes is large. This bias arises due to the averaging of efficiencies over the entire baseline sample (see eq. (2.7)) when in fact the efficiencies are on average different for real and fake leptons. Such differences in the averages occur randomly in the ‘toy’ MC tests, but can occur systematically in a physics analysis if, for example, the real and fake efficiencies have different kinematic distributions. The biases can be mitigated by binning the baseline sample in the variables for which the real-lepton and fake-lepton distributions may differ, and performing the likelihood fit separately in each bin. As an example of the effect of such a binning, the pseudoexperiments can be run with the results binned according to the values of $\varepsilon_f$.

The effect of using two such bins in the value of $\varepsilon_f$ for each lepton is shown in figure 4.

When the situation becomes more challenging, such as in figures 2 and 3, the characteristics of each method become more distinct. For low-statistics samples, the Poisson likelihood approach tends to exhibit a bias toward high values when the true number of fakes is small, a natural consequence of the fact that it cannot return negative values. The coverage of the true value by the estimated uncertainty is, however, still reasonable. A clear distinction between the precision of the methods also appears in figures 2 and 3, where the Poisson likelihood approach has the smallest statistical uncertainty, followed by the asymptotic matrix method and then the fake-factor method. This is because the Poisson likelihood approach considers lepton efficiencies averaged
Figure 4. Ratio of the estimated to expected fake-lepton yields for the Poisson likelihood matrix method in pseudoexperiments with dilepton events where the leptons had an average $\varepsilon_r = 0.9$ and $\varepsilon_f = 0.1$, both values were varied according to a Gaussian distribution of width 0.1 when simulating each lepton, and there were 1000 events per pseudoexperiment. The triangles show the result with a single bin, while the circles show the result where the input events are binned according to the values of $\varepsilon_f$ for each lepton. Two bins are used for each of the two leptons, and the results from all four bins are summed to give the estimated $N_{\text{fake}}$ value.

over the entire sample and is therefore less susceptible to event-by-event fluctuations in the values of $\varepsilon_r$ and $\varepsilon_f$.

Despite the differences between them, all the approaches are valid, as shown by the statistical coverage plots: except in extreme cases, the confidence intervals built from the estimates and their statistical uncertainty do contain the true fake yield in at least 68% of pseudoexperiments, and the visible overcoverage is due mostly to the fact that the uncertainties are computed under the assumption that all of the $N_{\text{tl}}$ values are independent, while the pseudoexperiments were generated with a fixed number of baseline events for each, which means there was some anticorrelation among the $N_{\text{tl}}$.

When selecting which method to use, the analyser needs to consider the relative benefits and complexities of implementing the methods, along with the size of the uncertainty in the fake/non-prompt lepton background yield relative to other uncertainties in the analysis.

3 The ATLAS detector

While the above description of the matrix and fake-factor methods is general, the remainder of this paper discusses the application of these methods to ATLAS physics analyses, and therefore a brief description of the experimental apparatus follows.

The ATLAS detector [1] at the LHC covers nearly the entire solid angle around the collision point.\(^3\) It consists of an inner tracking detector surrounded by a thin superconducting solenoid,
electromagnetic (EM) and hadronic calorimeters, and a muon spectrometer (MS) incorporating three large superconducting toroidal magnets. The inner-detector system is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$.

The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer (IBL) installed before Run 2 [11, 12]. It is followed by the silicon microstrip tracker, which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, EM calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for EM and hadronic measurements respectively.

The MS comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions. Interesting events are selected to be recorded by the first-level trigger implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [13]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger reduces in order to record events to disk at about 1 kHz.

An extensive software suite [14] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

4 Lepton selection criteria

Full descriptions of the electron and muon reconstruction algorithms and available selection criteria used in ATLAS are provided in refs. [2] and [3], respectively. Here the features most relevant to the fake/non-prompt lepton background estimation are summarised briefly.

4.1 Electron reconstruction and identification

Electron candidates are reconstructed within $|\eta| < 2.47$ as tracks in the inner detector matched to energy clusters in the EM calorimeter.

points upwards. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---

2023 JINST 18 T11004
In order to separate true electrons from other unwanted reconstructed candidates, electron identification (ID) algorithms are used. These rely upon a set of variables that quantify the distribution of energy in the calorimeter, the quality of the spatial match between the calorimeter deposit and the associated track, and the transition radiation signal in the TRT (see table 1 of ref. [2] for a complete list). Rather than place individual requirements on these variables, they are combined into a likelihood discriminant based upon the probability density functions of the variables measured for prompt electrons in $Z \rightarrow e^+e^-$ events and for background candidates reconstructed in inclusive collision events.

Since different analyses have different requirements for electron selection efficiency and background rejection, several “working points” (WPs) are defined by different values of the likelihood. The likelihood threshold values are varied according to the $p_T$ and $|\eta|$ of the electron candidate, so that the selection efficiency varies smoothly with the electron $p_T$. The most commonly used ID WPs (and their average efficiencies measured for typical electroweak processes) are “Loose” (93%), “Medium” (88%) and “Tight” (80%). In addition to the listed WPs, there is another (“LooseAndBLayer”) WP that imposes the same requirement on the likelihood as the “Loose” WP, but also requires that the electron track have a hit in the IBL to suppress candidates originating from photon conversions.

Often, additional requirements are imposed on the impact parameter of the electron’s track: the impact parameter in the transverse plane, $d_0$, with respect to the centre of the beamspot must satisfy $|d_0| < 5\sigma(d_0)$, where $\sigma(d_0)$ is its estimated uncertainty, while the longitudinal separation $z_0$ between the point where $d_0$ is measured and the chosen primary vertex of the event, multiplied by a moderating factor $\sin(\theta)$ which accounts for reduced $z_0$ accuracy for more forward tracks, cannot exceed 0.5 mm in absolute value.

### 4.2 Muon reconstruction and identification

Muon candidates are reconstructed in the region $|\eta| < 2.5$ by combining MS tracks with matching inner-detector tracks. The muon reconstruction efficiency is approximately 98% per muon in simulated $Z \rightarrow \mu^+\mu^-$ events. After reconstruction, high-quality muon candidates used for physics analyses are selected by a set of requirements on the number of hits in the different inner subdetectors and different MS stations, on the track fit properties, and on variables that test the compatibility of the individual measurements in the two detector systems, as detailed in ref. [3]. These criteria reduce the background from in-flight decays of light-flavour hadrons, which often result in kinked tracks. The most commonly used muon ID WPs (and their efficiencies measured in $t\bar{t}$ MC events) are “Medium” (98%) and “HighPt” (80%), the latter optimised to offer the best momentum resolution for $p_T > 100$ GeV. The same impact parameter requirements as defined for electrons are also often applied to muon candidates, with a tighter condition in the transverse plane: $|d_0| < 3\sigma(d_0)$.

### 4.3 Lepton isolation

In addition to the ID criteria mentioned above, most analyses place requirements on the isolation of the lepton from other detector activity. This is especially helpful in reducing the contribution from...

---

*Those efficiencies, as well as those quoted for muons in the next section, do not include requirements for the leptons to be successfully identified in the context of the trigger decision, which may include small additional inefficiencies [15, 16].*
leptons produced in heavy-flavour decays, or muons from $\pi^\pm$ or $K^\pm$ decays within jets\textsuperscript{5} as there are often other components of the jet that are near the lepton in these cases. In all of the methods described in section 2, the baseline lepton selection usually does not require isolation, while the tight lepton selection usually does. However, many of the single-lepton triggers\textsuperscript{[15, 16]} used in ATLAS require isolation, so analyses that depend on such triggers cannot avoid applying isolation requirements for baseline leptons.

The calorimeter isolation\textsuperscript{[2, 3]} is calculated from the sum of transverse energies of calorimeter energy clusters within $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = XX/100$ of the lepton candidate, not including the contribution expected from the candidate itself. Typical values of XX are 20, 30 or 40. Expected residual contributions to the isolation from the lepton candidate, as well as expected contributions from particles produced by additional proton–proton (‘$pp$’) interactions, are subtracted\textsuperscript{[20]}, resulting in the variable $E_{\text{coneXX}}$.

The track-based isolation\textsuperscript{[2, 3]}, denoted by $p_{\text{T coneXX}}$, is based on tracks near the lepton candidate with either $p_T > 0.5$ or 1 GeV that satisfy basic track-quality requirements and are spatially consistent with the primary vertex of the event. The scalar sum of the transverse momenta of such tracks, excluding tracks associated with the lepton candidate, is compared with the $p_T$ of the candidate to assess the isolation. For muon candidates, only the single track associated with the candidate is excluded; for electron candidates, additional tracks consistent with pair-production from a bremsstrahlung photon are also excluded. The track isolation can also be defined with a variable cone size ($p_{\text{T varconeXX}}$). In this case, the size of the cone around the lepton candidate within which tracks are considered is varied as a function of the $p_T$ of the candidate:

$$\Delta R = \min \left( \frac{10 \text{ GeV}}{p_T \text{ [GeV]}}, R_{\text{max}} \right),$$

where $R_{\text{max}}$ is the maximum cone size (typically 0.2 or 0.3).

Combining selections on track-based and calorimeter-based isolation provides even better fake/non-prompt lepton background rejection, as the two isolation variables use complementary information. Track-based isolation was found to be less sensitive to detector noise and pile-up effects than calorimeter-based isolation, and the inner detector provides a better $p_T$ measurement than the calorimeters for individual soft hadrons. On the other hand, calorimeter-based isolation includes neutral particles as well as some particles below the inner detector’s track-$p_T$ threshold, which are ignored when computing track isolation. However, track and calorimeter isolation variables measure hadronic activity in a redundant manner, since charged particles are measured by both the calorimeters and the inner detector, and simple selection cuts applied independently to those two variables may not achieve optimal rejection power. To avoid this, an analysis can use a “particle-flow” algorithm, which allows removal of overlapping contributions from the track-based and calorimeter-based isolation, decreasing the correlation between the two variables. For the time being, particle-flow-based isolation variables are defined only for muons, and discussed in ref.\textsuperscript{[3]}.

For analyses where the fake/non-prompt lepton background may be dominated by non-prompt electrons and muons from the decays of $b$- and $c$-hadrons\textsuperscript{[21, 22]}, isolation WPs using a boosted jets are reconstructed from clusters of topologically connected calorimeter cells (topo-clusters), as described in ref.\textsuperscript{[17]}. The anti-$k_T$ algorithm\textsuperscript{[18, 19]} is used to form jets from the topo-clusters, with the radius parameter $R$ usually set to 0.4; when reconstructing jets formed from the merged decay products of boosted resonances, a larger $R$ value, typically 1.0, is used. Typically, jets with $p_T > 20 \text{ GeV}$ and $|\eta| < 4.5$ are considered in physics analyses.
decision tree (BDT) discriminant based on isolation and secondary vertex information, referred to as the non-prompt lepton BDT, are also used.

Several isolation WPs based on tracking, a combination of calorimetry and tracking, particle-flow, or a non-prompt lepton BDT are defined to allow for consistency across analyses that require different levels of lepton isolation. They are described in refs. [2] and [3].

4.4 Removing overlaps between jets and leptons

In some cases the same object can result in multiple signatures in the detector. For example, an electron will deposit energy in the calorimeter, and that energy will generally also be clustered into a jet. In addition, sometimes objects will be spatially correlated due to the underlying physics, as when a muon is produced by heavy-flavour decay within a jet. To avoid double-counting, and to select only the isolated leptons that are of interest in physics analyses, an overlap removal procedure is applied to resolve these ambiguities. The procedure used for most physics analyses is as follows, where the lepton candidates are those that satisfy the baseline ID criteria for the analysis:

- All electron candidates that are within $\Delta R = 0.01$ of a muon candidate (or share a track with a muon candidate) are removed.
- All jets that are within $\Delta R = 0.2$ of any remaining electron candidates are removed.
- All electron candidates that are within $\Delta R = 0.4$ of any remaining jet are removed.
- Cases where a remaining jet is within $\Delta R = 0.4$ of a muon candidate are examined to determine the number of tracks associated with the jet. If there are more than two such tracks the muon candidate is removed; otherwise the jet is removed.

Variations of this procedure are also supported, primarily for analyses that focus on heavy-flavour jets or that select boosted massive particles using large-$R$ jets.

5 Monte Carlo simulation samples

The studies of fake/non-prompt leptons that are presented in sections 6 and 7 make use of large MC samples of simulated events. The production of $t\bar{t}$ events at next-to-leading order (NLO) in the quantum chromodynamics (QCD) coupling constant $\alpha_s$ is described in ref. [23] and relies on the Powheg Box v2 event generator [24] interfaced with Pythia 8.230 [25] for parton showering and subsequent steps, with the A14 set of tuned parameters [26]. The parton distribution functions used for matrix element calculation and parton showering are NNPDF3.0nlo [27] and NNPDF2.3lo [28] respectively. Generated events were filtered such that at least one of the top quarks decays semileptonically. The EvtGen 1.2.0 program [29] was used to model heavy-flavour hadron decays.

The production of Drell–Yan $Z/\gamma^* \to \ell^+\ell^-$ events ($\ell = e, \mu, \tau$) at NLO in $\alpha_s$ is described in ref. [30] and relies on the Sherpa 2.2 event generator [31] with the dedicated set of tuned parameters and the NNPDF3.0nnlo [27] parton distribution function. Events were generated according to a partition of the phase space described in ref. [30], resulting in a set of orthogonal samples which
were combined with weights corresponding to the NLO cross-section calculated by the generator. The $\ell^+\ell^-$ invariant mass was required to be at least 40 GeV.

A full ATLAS detector simulation [32] based on Geant4 [33] was then used to faithfully reproduce particle interactions with the detector and its response. Additional $p p$ interactions in the same or neighbouring bunch crossings were also simulated in order to reproduce the conditions of data-taking during the LHC Run 2 operation. Simulated events are weighted to reproduce the measured performances of various selection criteria, such as electron identification [2]. In addition, reconstructed lepton and jet momenta are transformed [2, 34, 35] to match the real detector resolution.

The source of a lepton candidate in simulation, discussed in section 6, is identified by using preserved generator-level information to match the track of each reconstructed lepton candidate to a suitable stable particle produced by the generator or the Geant4 simulation, among those preserved in the event record (sections 3.6 and 5.3 in ref. [32]). The matching is performed by evaluating the contribution of each stable particle to the simulated ID and (for muons) MS hits used to reconstruct the track, and requiring that the matched particle accounts for at least half of those hits (section 6.3 in ref. [36]). While the latter is a general procedure used for all tracks, for leptons specifically the result is discarded if the reconstructed track has less than three combined hits in the pixel and SCT detectors, or if the track and the stable particle do not satisfy $\Delta R < 0.2$. If the stable particle is a charged lepton (or, in the case of electron candidates, may also be a photon), its origin is checked so that the different sources of non-prompt leptons can be distinguished.

In section 9, other simulated processes are used. The corresponding MC samples are obtained with a similar workflow. Complete information about the MC generators and their configurations, as well as the cross-section calculations used to normalise the simulated samples, is available in the references provided in that section.

6 Sources of fake/non-prompt leptons

The relative contributions of fake/non-prompt leptons from different sources to the sample of selected lepton candidates depend on the energy and spatial location of the candidate in the detector, the identification, impact parameter and isolation criteria applied, the overlap removal procedure, and the nature of the selected final state (e.g. the presence or absence of heavy-flavour hadrons).

Figures 5 and 6 present for illustration the relative contributions of the different sources of fake/non-prompt muons and electrons respectively, as a function of the transverse momentum of the candidate, as measured in MC simulated events. They are shown for two different processes: $t\bar{t}$ production, leading to final states rich in heavy-flavour hadrons, and Drell–Yan production of $e^+e^-$ or $\mu^+\mu^-$ pairs; for $t\bar{t}$ events at least one of the two $W$ bosons produced in the $t\bar{t}$ decay is required to decay leptonically ($W \rightarrow \ell\nu$). For these particular figures, the reconstruction/ID of electron and muons and the general event selection follow those described in ref. [37] for baseline lepton candidates, which correspond to rather loose criteria (in particular, no isolation nor transverse impact parameter requirements are applied, and no overlap removal is done). For both processes, events are considered only if they contain a pair of reconstructed baseline leptons with identical charges, a signature for which fake/non-prompt leptons usually represent a non-negligible source of background.
Non-prompt leptons are those arising from electroweak decays of hadrons. Heavy-flavour $b$- and $c$-hadrons decay close to the interaction point and the resulting leptons are distinguishable from real leptons mostly by isolation and impact parameter.

The “prompt $\gamma$-conversion” category only includes electron candidates where the photon is separated by $\Delta R > 0.1$ from any generator-level high-$p_T$ electron from the hard-scatter interaction (photons emitted at smaller separation are generally reconstructed as part of the electron candidates).

Non-prompt leptons are those arising from electroweak decays of hadrons. Heavy-flavour $b$- and $c$-hadrons decay close to the interaction point and the resulting leptons are distinguishable from real leptons mostly by isolation and impact parameter.

Figure 5. Relative contributions of fake or non-prompt muons from different sources as a function of $p_T$ in simulated (a) $t\bar{t}$ and (b) Drell–Yan processes. Events are required to contain at least two leptons with identical charges. Punch-through hadrons are charged hadrons reaching the MS and leading to the reconstruction of a fake muon candidate. Some other minor sources are not displayed; the sum of their contributions is less than 2% in every bin of (a), and less than 4% in bins of (b) up to $p_T = 100$ GeV. The muons are required to satisfy minimal track-quality criteria specified in the text. Error bars represent statistical uncertainties of the simulation.

Figure 6. Relative contributions of fake or non-prompt electrons from different sources as a function of $p_T$ in simulated (a) $t\bar{t}$ and (b) Drell–Yan processes. Events are required to contain at least two leptons with identical charges. Some other minor sources are not displayed; the sum of their contributions is less than 0.5% in every bin. These do not include electron candidates corresponding to high-energy muon ionisation signal in the calorimeters, or from conversions of final-state radiation (FSR) photons with small angular separation from a genuine electron or muon. The electrons are required to satisfy loose ID criteria specified in the text and ref. [37]. Error bars represent statistical uncertainties of the simulation.
Light-flavour hadrons can also be a major source of fake leptons via decay-in-flight in the tracker volume. This happens mainly in final states for which QCD multi-jet production is a significant contributor. A charged hadron stopping early in the calorimeter, and generating a narrower-than-average shower, can mimic the experimental signature of an electron. Electron ID criteria are particularly powerful in rejecting these candidates rather than those from other sources, based notably on three-dimensional profiles of the shower, but since many orders of magnitude more hadrons than leptons are produced in collisions at the LHC, a substantial number of such fake electrons may be selected in a physics analysis. With regard to muons, the depth of the ATLAS calorimeter is sufficient to stop most pions and kaons before they can reach the MS. Muon candidates arising from kaons decaying semileptonically in flight before reaching the MS are a more important contribution to the fake/non-prompt lepton background.

Among hadrons faking electrons, one significant class is neutral pion decays into photons ($\pi^0 \rightarrow \gamma \gamma$); the collimated photons create a single energy deposit in the EM calorimeter, while the associated track might be provided by the conversion of one of the photons in the upstream detector material. Due to the importance of this phenomenon, the electron ID criteria specifically discriminate against these by attempting to identify a two-peak structure in the distribution of the cell energies matched to the electron cluster [2], which is not present for real electrons. Other contributions such as Dalitz decay of pions [38] also create comparable experimental signatures.

The conversion of photons into electron–positron pairs represents the last important class of non-prompt electrons. These $\gamma$-conversions must typically occur early on (e.g. in the beam pipe) and be largely asymmetric in the splitting of the momentum between the two electrons; otherwise, the conversion vertex can be reconstructed, or the candidate electron’s track lacks hits in the first layers of the inner detector, both leading to the proper classification of the reconstructed object as a photon instead of an electron [2]. The origin of the photon itself influences the characterisation of the candidate as non-prompt or real: photons emitted close to a real electron, either due to bremsstrahlung or as higher-order quantum electrodynamic (QED) corrections to the production process, are typically considered part of the electron candidate (the calorimeter energy deposits tend to overlap to the extent that a single cluster is reconstructed); furthermore, from the perspective of quantum field theory, well-defined electrons must include extra radiation (“dressed leptons” [39]). The electron reconstruction procedure [2] accounts for bremsstrahlung, in particular by allowing kinks in the track consistent with bremsstrahlung emission in dense material regions. In contrast, photons from other origins, such as initial-state radiation, QED processes not involving leptons or where photons are sufficiently separated from leptons, or hadronic jet fragmentation, may be considered as sources of fake electrons.

It can be seen in figure 5 that non-prompt muons constitute the only substantial contribution to the fake/non-prompt muon background, while for electrons figure 6 shows more variety: in general, non-prompt electrons are particularly represented in the lower $p_T$ range, especially in processes involving the production of heavy-flavour hadrons, while hadron fakes and converted photons populate the higher $p_T$ range.\footnote{However, for the range $p_T < 10\text{ GeV}$, not shown in the figure, hadron fakes are also the dominant contribution.}

The different sources of fake/non-prompt leptons have distinct probabilities to satisfy the tight lepton selection criteria described in section 4. Figure 7 illustrates those differences for the particular...
example of signal leptons definitions used in ref. [37] (including lepton–jet overlap removal) and simulated $t\bar{t}$ events containing two leptons with identical charges. Such variability between sources is unwelcome, as the fake efficiencies required for the application of the methods described in section 2 then depend upon the relative contributions of each source to the regions of interest, which may not be easy to assess. It is therefore desirable to measure the efficiencies in regions similar in composition to the regions where the background estimate is needed; otherwise large extrapolation uncertainties may apply.

Precise simulation of these various sources of background, including their relative contributions, is indeed very challenging, as it relies heavily on the modelling of the soft-QCD regime by event generators, including modelling of fragmentation and hadronisation processes, hadron decay modelling, and soft emissions and detector modelling. Another issue is that only a small fraction of fake or non-prompt lepton candidates survive the ID and isolation requirements, so the simulation of a very large number of events is needed to obtain a statistically accurate prediction. For inclusive processes with large cross-sections (e.g. multi-jet production), this is often impractical.

For these reasons, many of the fake/non-prompt lepton background predictions used in ATLAS publications are based on methods using the data, such as the ones described in this paper. They rely on common properties shared to some extent by the different sources of fake/non-prompt leptons that differentiate them from real leptons, such as a high likelihood to not meet the combination of ID and isolation criteria.

7 Measurement of real and fake/non-prompt lepton efficiencies

The methods described in section 2 both rely on knowledge of the efficiency for leptons that pass the baseline selection to also pass the tight selection. For the fake-factor method, only the efficiency for fake/non-prompt leptons is used explicitly in the calculation, while for the other methods the efficiency for real leptons must also be measured. In many cases these efficiencies depend on the
properties of the lepton (such as its $p_T$ or angular distance from a jet) or on the event in which it is found (such as the overall activity of the event, as measured for example by the number of reconstructed primary vertices).

7.1 Real-lepton efficiencies

The efficiencies of specific working points of lepton selections are calibrated precisely in ATLAS for general purposes, primarily with “tag-and-probe methods” based e.g. on $Z \rightarrow \ell\ell$ events. By performing these measurements on both data and MC-simulated events, “scale factors” (SFs) that account for differences between the efficiencies observed in data and simulation are derived. These SFs can then be applied to simulated events using the selection criteria that are relevant to a given analysis to determine the appropriate real-lepton efficiencies. This MC-based approach is valid as long as both the baseline and tight lepton selection criteria are taken from the set for which SFs have been measured; only in extraordinary cases would an analysis utilise different selection criteria. Since the efficiencies depend more strongly on the environment than the SFs do, the main advantage of this approach over purely data-based measurements is to allow efficiencies to be obtained directly in the desired environment (i.e. the region in which the fake/non-prompt background estimate is needed), rather than being extrapolated from a more distant region which would be needed for reliable measurements in data.

Details of the real-electron and real-muon efficiency measurements can be found in refs. [2] and [3], respectively. The real efficiencies are often parameterised with respect to the $p_T$ and $|\eta|$ of the leptons, and measured separately for electrons and muons.

7.2 Fake/non-prompt lepton efficiencies

The fake/non-prompt lepton efficiencies are specific to each analysis, primarily since there are several sources for such leptons (see section 6), which will contribute with different weights depending on the chosen selection criteria. In general, though, the first step in the efficiency measurement is to identify a region that has a large contribution from fake leptons. Two approaches are commonly used. In the first, events with a pair of leptons with the same electric charge are selected. Since such lepton pairs are only rarely produced at the LHC (via processes such as $WZ + \text{jets}$, $ZZ + \text{jets}$, and $t\bar{t} + X (X = W/Z)$ production) it is likely that one of the two leptons is fake/non-prompt. By placing stringent quality criteria on one lepton in these events, the probability that the remaining lepton is fake/non-prompt is enhanced. The second approach is to use single-lepton events, where criteria are imposed to suppress the contribution from real leptons. Examples of such criteria are requiring the missing transverse momentum\(^7 E_{T}^{\text{miss}}\), or transverse mass\(^8 m_T\), to be below specific thresholds, thereby reducing the contribution from $W+\text{jets}$ or $t\bar{t}$ events, or requiring via the track impact parameters that the lepton originate from a position inconsistent with the primary event vertex, thereby enhancing the contribution of leptons from heavy-flavour decay.

---

\(^7E_{T}^{\text{miss}}\) is defined as the negative vector sum of the $p_T$ of the reconstructed and calibrated objects in the event, with a correction applied for inner detector tracks that originate from the primary collision vertex and are not associated with any other objects, and $E_{T}^{\text{miss}}$ is defined as the magnitude of $E_{T}^{\text{miss}}$ [40].

\(^8\)Here $m_T \equiv \sqrt{2 p_T^\ell p_T^{\text{miss}} (1 - \cos \Delta \phi_{\ell,E_{T}^{\text{miss}}})}$, where $\Delta \phi_{\ell,E_{T}^{\text{miss}}}$ is the azimuthal angle between the lepton and $E_{T}^{\text{miss}}$ directions.
Figure 8. Probability for a muon produced in the decay of a $b$-hadron to satisfy a track-based isolation requirement, in simulated $t\bar{t}$ events, shown as a function of (a) the muon $p_T$; and (b) the sum $p_T + E_{\text{cone}}^{\text{40}}$ of the muon $p_T$ and nearby calorimeter energy deposits. In each case, the universality of these probabilities is gauged by changing the kinematic properties of the parent hadron (markers of various colours), achieved indirectly by selecting events in different ranges of global transverse energy $H_T$. Precise definitions of the isolation criteria, the $p_T + E_{\text{cone}}^{\text{40}}$ and $H_T$ variables, as well as descriptions of the muon and jets selections, are given in section 7.2.

In either approach, there will be a residual contribution from events with only real leptons in the selected sample. This contribution is typically estimated using MC simulation and subtracted separately from both the tight and baseline samples before the ratio of these samples is taken to measure the efficiency.

As with the real-lepton efficiencies, the fake/non-prompt lepton efficiencies depend on properties of the lepton candidates or of the event in which they are found. Therefore, it is generally helpful to bin the efficiencies in the lepton $p_T$ and $|\eta|$, and possibly in terms of other quantities as well. The optimal binning to be used is chosen in the context of each physics analysis, considering the given numbers of events and potential changes in the relative contributions from the different sources of fake/non-prompt leptons; illustrative examples are provided in section 9.

Adopting a parameterisation of the efficiency with respect to variables other than $p_T$ and $\eta$ can sometimes be beneficial. For example, the probability that a fake or non-prompt lepton satisfies the isolation criteria is correlated with the fraction of the jet’s visible momentum carried by the lepton. Therefore parametrising the fake efficiency as a function of the lepton $p_T$ amounts to assuming that, for a particular lepton $p_T$, the distribution of the parent jet momentum is similar in the regions where the efficiencies are measured and the regions where the background estimates are needed. If this assumption does not hold, it can be useful to adopt instead a parameterisation as a function of the parent jet momentum. Since this quantity is not easily accessed experimentally (unlike the lepton $p_T$), proxy observables are used in practice. An example of successful application is the analysis in ref. [41], which employed the sum of the lepton’s $p_T$ and the transverse energy in a cone around the lepton as a proxy.

The preceding discussion is also illustrated in figure 8 for non-prompt muons produced in the decay of $b$-hadrons in simulated $t\bar{t}$ events. The probabilities for such muons to satisfy a track-based isolation requirement, as defined in section 4, are shown for two alternative parameterisations: one
based on the muon $p_T$, and the other on the scalar sum of the muon $p_T$ and the transverse energy deposited in calorimeter-cell clusters within a cone of size $\Delta R = 0.4$ around the muon (referred to as $E_T^{\text{cone40}}$ in section 4). This scalar sum serves as a proxy for the parent jet’s transverse momentum.

In the second parameterisation, the fake efficiency is the fraction of jets with a non-prompt muon and visible momentum $p_T + E_T^{\text{cone40}}$ in which the muon is mostly isolated, i.e. $p_T \gg E_T^{\text{cone40}}$. While one might consider jets with arbitrarily soft muons in the denominator of this fraction, for practical reasons figure 8(b) only includes events where muons satisfy $p_T > 10$ GeV. To study the dependence of these efficiencies on the momentum distribution of the underlying jet, different regions of jet momentum are emphasised by imposing different requirements on the global transverse energy $H_T$ in the event, defined for this purpose as the scalar $p_T$ sum of all jets with $p_T > 25$ GeV and $|\eta| < 2.8$ that are a distance $\Delta R > 0.6$ from the muon. This quantity is indeed partially correlated with the kinematics of the muon’s parent jet, via the momentum of the top quarks producing all these jets. The reconstruction, calibration and selection of jets and muons for this figure are otherwise those detailed in ref. [37].

It can be observed that for the case of a $p_T$-dependent parameterisation, the efficiencies vary strongly with $H_T$, although large differences occur mostly for $p_T > 40$ GeV. Since most non-prompt muons are produced at low $p_T$, the overall impact of this non-captured dependency might be small, unless regions of interest in the analysis specifically select high-$p_T$ leptons. In contrast, the parameterisation as a function of $p_T + E_T^{\text{cone40}}$ is much less influenced by $H_T$, making the measured efficiencies less dependent on the event topology. In practice, a compromise has to be found between this observation and other elements evoked above justifying a $p_T$-dependent parameterisation, especially for electrons because ID criteria are usually employed in addition to isolation. The direct dependency of efficiencies on other variables that are also correlated with the event topology (e.g. $E_T^{\text{miss}}$) may also be reduced by such a parameterisation.

As detailed in the next section, suitable uncertainties must be assigned to the use of the fake efficiencies in different regions than those in which they are measured, in particular to account for potential differences in relative contributions of the different sources of fake/non-prompt leptons. To minimize these uncertainties, it has sometimes been found beneficial to use an approach closer to that of section 7.1, in which efficiencies are evaluated in the simulation and supplemented by data-driven correction factors that depend on the source of the fake/non-prompt lepton. These correction factors are derived using dedicated control regions that are enriched in a particular source of fake/non-prompt leptons. The main assumptions are then the universality of the correction factors across different processes, and the ability of the simulation to adequately predict the relative contributions of each source in the regions of interest. Such an approach has for example been used in ref. [42].

8 Systematic uncertainties

Systematic uncertainties in the fake/non-prompt lepton background estimates from the matrix method and the fake-factor method arise from uncertainties in the values of $\epsilon_i$ and $\epsilon_f$. These uncertainties can be traced to statistical uncertainties from the samples used to measure the efficiencies, to potential biases that may cause the efficiencies in the signal region for a particular analysis to differ from the values obtained from control samples (such as differences in the origin of fake/non-prompt leptons between these regions), and to uncertainties in the modelling of contamination from real-
lepton processes in the samples used to measure the fake efficiency. Details of these uncertainties and their estimation are provided below. The overall impact of variations in \( \varepsilon_r \) and \( \varepsilon_f \) depends on the characteristics of the analysis. For example, in the matrix method Equations 2.3–2.5 imply that for analyses with

\[
N_t^r \times \frac{\Delta \varepsilon_t}{\varepsilon_t^2} > N_t^f \times \frac{\Delta \varepsilon_t}{\varepsilon_t^2}
\]

the uncertainty in \( \varepsilon_t \) will dominate, and vice versa.

8.1 Statistical uncertainties in the measured efficiencies

Statistical uncertainties in the real-lepton and fake-lepton efficiencies can be accounted for either by analytically propagating the uncertainties through to the estimated event yields or by varying the efficiencies input to the nominal yield calculation by their statistical uncertainties and using the resulting difference in the estimated fake/non-prompt background yield as the resulting uncertainty. The real- and fake-efficiency uncertainties are generally uncorrelated since the efficiencies are measured in statistically independent samples. In the usual case where the efficiencies are measured in bins of one or more quantities, the bins are uncorrelated, so the variations in each bin are applied separately. The total statistical uncertainty in the fake/non-prompt lepton background estimate is given by the sum in quadrature of the uncertainties from all variations. This source of systematic uncertainty is usually not dominant.

8.2 Systematic uncertainties in the measured efficiencies

Systematic uncertainties in \( \varepsilon_f \) are generally larger and more challenging to assess than for \( \varepsilon_r \):  

1. Real-lepton efficiencies have only slight variations due to event environment effects since real leptons have small contributions to their measurements from underlying event and jet activity. This also means there are a wide variety of samples with which they can be calibrated in great detail.

2. There are several sources of fake/non-prompt leptons, and the efficiencies may differ between these sources. Therefore, any differences in the fake/non-prompt lepton composition between the sample used to measure the efficiencies and the signal region for an analysis may lead to a bias in the efficiencies.

Several methods are used to estimate systematic uncertainties in the fake-lepton efficiencies. One is simply to vary the selection criteria for events in the control region used to measure the fake-lepton efficiencies, since the composition of fake/non-prompt leptons in the standard and alternative control regions may differ. A more sophisticated approach is to use MC simulation to estimate the fake/non-prompt lepton compositions in both the control and analysis regions. This information, combined with the MC-estimated selection efficiencies for each source of fake/non-prompt leptons, can be used to provide an estimate of the uncertainty.
8.3 Uncertainties in the modelling of real-lepton processes

When measuring the fake-lepton efficiencies, a correction must be applied to account for contamination from processes with real leptons in the control sample used for the measurement:

$$\varepsilon_{\ell} = \frac{N_{t} - N_{t}^{l}}{(N_{t} + N_{l}) - (N_{t}^{l} + N_{l}^{l})}$$

where $N_{t}$ and $N_{t}^{l}$ are the numbers of real leptons in the selected tight and loose samples, respectively. MC simulation of real-lepton processes is generally used to estimate $N_{t}$ and $N_{t}^{l}$, with corrections applied to account for known differences in object selection efficiencies between simulation and data. Nonetheless, several sources of systematic uncertainty in the real-lepton contamination remain: uncertainties in the cross-sections of the real-lepton processes, uncertainties in the correction factors, and uncertainties in the parameters, e.g. parton distribution functions (PDFs) and factorisation/renormalisation scales ($\mu_{f}, \mu_{r}$), used in the simulation.

8.4 Uncertainties due to biases in the Poisson likelihood matrix method

The biases that occur in some situations for the Poisson likelihood matrix method (see section 2.5) may also be considered as a source of systematic uncertainty, especially when the signal region contains many bins with few events in each. The magnitude of the bias can be estimated either by repeating the analysis with coarser bins, or by constraining the total fake/non-prompt lepton yield estimate to be the value returned by applying the Poisson likelihood matrix method to the entire unbinned sample. Any resulting differences in the binned estimates can be taken as a systematic uncertainty.

9 Examples of application in ATLAS analyses

In this section the application of the fake/non-prompt lepton background estimation methods is described using two example ATLAS analyses. The first is a measurement of the $t\bar{t}Z$ differential cross-section using events that contain three or four lepton candidates, and the second is a model-independent search for “beyond the Standard Model” (BSM) phenomena in events with three or more lepton candidates. In both cases, the high lepton multiplicity suppresses the Standard Model (SM) backgrounds, which makes the relative contribution of the fake/non-prompt lepton background larger.

Data from $\sqrt{s} = 13$ TeV $pp$ collisions recorded by the ATLAS detector between 2015 and 2018 are used to perform these analyses. In this period, the LHC delivered colliding beams with a peak instantaneous luminosity of $L = 2.1 \times 10^{34}$ cm$^{-2}$s$^{-1}$, achieved in 2018, and an average number of $pp$ interactions per bunch crossing of 33.7. After applying beam, detector and data-quality criteria, the total integrated luminosity of the dataset is 139 fb$^{-1}$ [43]. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [44], obtained using the LUCID-2 detector [45] for the primary luminosity measurements.

9.1 Measurement of the $t\bar{t}Z$ cross-section in final states with three or four leptons

This analysis measured the differential production cross-section of the $t\bar{t}Z$ process in final states with a total of three or four electron and muon candidates. The tools described above are used for
the lepton efficiency measurements and the application of the Poisson likelihood matrix method (described in section 2.1.2). In addition to the fake/non-prompt lepton estimation in the signal regions of the analysis, the results are checked in several validation regions enriched in fake/non-prompt leptons. The Poisson likelihood matrix method was chosen for this analysis since the number of fake/non-prompt leptons in the signal regions with three or four leptons is expected to be very low, and the Poisson likelihood matrix method provides more stable results for binned estimations, which are necessary for differential background predictions in these low-statistics signal regions. The fraction of events with more than one fake/non-prompt lepton in the signal or validation regions has been checked and found to be negligible. More details about the analysis and the fake/non-prompt lepton background estimation can be found in ref. [46].

9.1.1 Real-lepton efficiencies

The first step in the application of the matrix method is to measure the efficiencies for real and fake/non-prompt leptons that satisfy the baseline criteria to also satisfy the tight criteria. In the $t\bar{t}Z$ cross-section measurement, baseline electrons are required to satisfy the “LooseAndBLayer” ID WP, whereas tight electrons are required to satisfy the stricter “Medium” ID criteria and to be isolated from nearby tracks and calorimeter energy deposits. Both the baseline and tight muons are required to satisfy the “Medium” WP, and tight muons are in addition required to be isolated from nearby tracks.

As discussed in section 7.1, the real-lepton efficiencies are obtained using MC simulation, corrected to match the performance seen in data control samples. Those efficiencies are shown in figure 9, binned in lepton $p_T$ and $|\eta|$. To check for potential dependencies on the number of additional jets in the events used for the measurements, the efficiencies are derived for different jet multiplicities. No significant differences between the real-lepton efficiencies are observed.

9.1.2 Fake/non-prompt lepton efficiencies

The fake/non-prompt lepton efficiencies are measured with same-charge electron–muon ($e\mu$) or muon–muon ($\mu\mu$) data events using a tag-and-probe method, where one “tag lepton” with very stringent requirements on momentum and isolation ($p_T > 40\text{ GeV}$, $\max(\rho^{\text{cone}20}_{p_T}, E^{\text{cone}20}_{p_T})/p_T < 0.01$) is selected and the remaining “probe lepton” is used for the efficiency measurement. Events with more than two leptons are not considered. Only the aforementioned baseline electron or muon requirements are used to select the probe lepton. The samples are dominated by events containing at least one fake/non-prompt lepton, and are orthogonal to the signal regions, which require a minimum of three tight lepton candidates. In addition, $t\bar{t}Z$ production is negligible in these samples.

The definition of the regions used for the fake-efficiency measurements is summarised in table 1. For the electron fake efficiencies an $e\mu$ signature is used, with the muon being used as the tag lepton. For the muon fake efficiencies, the $\mu\mu$ region is used, since the $e\mu$ region also contains unwanted events where an electron with an incorrectly measured charge is selected as the tag lepton and paired with a real muon.

Real-lepton background processes (primarily diboson and $t\bar{t}W$ production) leading to the same-charge dilepton signature are estimated using MC-simulated events and are subtracted from data.

---

9In $\mu\mu$ events with both muons satisfying the tag-selection, the one with the higher $p_T$ is chosen as the tag lepton.
Figure 9. The two-dimensional real-lepton efficiencies obtained for (a) electrons and (b) muons, in bins of \( p_T \) and |\( \eta \)| of the leptons. The last \( p_T \) bin is inclusive. The real-lepton efficiencies are obtained using MC simulation, but corrected to match the performance seen in data control samples.
Figure 10. The two-dimensional fake/non-prompt lepton efficiencies measured for (a) electrons and (b) muons, in bins of $p_T$ and $|\eta|$ of the leptons. The last $p_T$ bin is inclusive. The indicated uncertainties show only the statistical errors in the given bins. Compared to the real-lepton efficiencies shown in figure 9, the fake/non-prompt lepton efficiencies depend much more on the specifications of the analysis.
Table 1. Definition of the fake-lepton control regions used for the electron ($e$-fakes-CR) and muon ($\mu$-fakes-CR) fake-efficiency measurements. $N_\ell$ is the number of leptons ($\ell$, while $N_{\text{jets}}$ ($N_{b\text{-jets}}$) is the number of jets ($b$-tagged jets, see text), respectively. For additional details of the definitions of physics objects, see ref. [46].

| Region       | $N_\ell$ ($\ell = e, \mu$) | $p_T (\ell)$ | $p_T$ (jets) | $N_{\text{jets}}$ | $N_{b\text{-jets}}$ |
|--------------|-----------------------------|--------------|--------------|-------------------|-------------------|
| $e$-fakes-CR | $= 2 \ (e^+\mu^-)$        | $> 7$ GeV    | $> 25$ GeV   | $\geq 1$          | $\geq 1$          |
| $\mu$-fakes-CR | $= 2 \ (\mu^+\mu^-)$     | $> 7$ GeV    | $> 25$ GeV   | $\geq 1$          | $\geq 1$          |

to obtain an unbiased efficiency measurement. The contribution from electrons with misassigned charge in the same-charge $e\mu$ region is also subtracted using estimates from MC-simulated events.

After this subtraction, the dominant source of fake/non-prompt leptons is found to be heavy-flavour hadron decays. The fake efficiencies, binned in lepton $p_T$ and $|\eta|$, are shown in figure 10.

It is assumed that since the loose and tight lepton selection criteria depend on quantities related to the lepton itself or to its immediate surroundings, the chosen parameterisation captures the main variations in the fake efficiencies, and residual dependencies on the event environment can be covered by systematic uncertainties (see section 8). Indeed, in the simulation, the dependence of the fake efficiencies on the number of light-flavour jets or $b$-tagged jets is mild. The uncertainties are evaluated by comparing, in the simulation, the $\varepsilon_f$ values in the control regions described in table 1 to the values in the signal regions. These differences, evaluated as a function of $p_T$ and $|\eta|$, are applied as a systematic uncertainty of the fake-efficiency measurement (as discussed in section 8), and are of the order of 10%–20%, except for muons with $p_T > 50$ GeV, for which they reach 40%. Furthermore, normalisation uncertainties are considered for the real-lepton background processes, which are subtracted in the fake-efficiency measurement. They are evaluated by scaling the real-lepton background processes upwards and downwards within their cross-section uncertainties before the subtraction and using the differences between the modified and nominal efficiencies as uncertainties, which are added in quadrature to the aforementioned uncertainties.

9.1.3 Results in the fake/non-prompt lepton validation regions

To validate the performance of the method, predictions are obtained and compared with data in two dedicated validation regions called “VR-3$\ell$-1$b$3$j$” and “VR-3$\ell$-1$b$3$j$-noZ”, which have a larger proportion of fake/non-prompt leptons than is expected in the signal regions. The definitions of these two validation regions are summarised in table 2. No charge requirements are placed on the reconstructed lepton candidates in these regions.

The variable $m_{\ell\ell}^{\text{SF}}$ refers to the invariant mass of the same-flavour opposite-charge (SFOC) lepton pair with the invariant mass closest to the $Z$ boson mass. VR-3$\ell$-1$b$3$j$ is a region similar to the actual signal regions defined in ref. [46], but without a requirement on the $Z$ mass for the (SFOC) lepton candidate pair. Therefore, it contains a higher fraction of fake/non-prompt leptons after the selection. To further enhance the fake/non-prompt lepton contribution, the third-highest-$p_T$ jets containing $b$-hadrons are identified (tagged) by the MV2c10 $b$-tagging algorithm [47]. The algorithm uses a multivariate discriminant with quantities such as the impact parameters of associated tracks, and well-reconstructed secondary vertices.
Table 2. Definition of the fake/non-prompt lepton enriched validation regions, VR-3ℓ-1b3j and VR-3ℓ-1b3j-noZ. \( N_\ell \) stands for the number of leptons (\( \ell \)), while \( N_{\text{jets}} \) (\( N_{b\text{-jets}} \)) is the number of jets (\( b\)-tagged jets), respectively. The leptons are ordered by decreasing \( p_T \) (\( \ell_1, \ell_2, \ell_3 \)). The mass (\( m \)) variables are discussed in the text. For additional details of the object definitions, see ref. [46].

| Region                  | \( N_\ell \) (\( \ell = e, \mu \)) | \( p_T \) (\( \ell_1, \ell_2, \ell_3 \)) | \( p_T \) (jets) | \( |m_{\ell\ell}^{\text{SF}} - m_Z| \) | \( N_{\text{jets}} \) | \( N_{b\text{-jets}} \) |
|-------------------------|------------------------------------|------------------------------------------|------------------|--------------------------------------|----------------|----------------|
| VR-3ℓ-1b3j             | 3                                  | > 27, 20, 20 GeV                        | > 25 GeV        | -                                    | = 3           | = 1           |
|                         | (2 tight, 1 loose)                 |                                          |                  |                                      |                |                |
| VR-3ℓ-1b3j-noZ         | 3                                  | > 27, 20, 20 GeV                        | > 25 GeV        | > 10 GeV                             | = 3           | = 1           |
|                         | (all tight)                        |                                          |                  |                                      |                |                |

A lepton that satisfies the baseline selection criteria must not satisfy the tight criteria. An additional validation region, VR-3ℓ-1b3j-noZ, is defined by requiring all three leptons to satisfy the tight selection criteria, but placing a veto on SFOC lepton pairs that have an invariant mass consistent with the \( Z \) boson, thereby enhancing the fake/non-prompt lepton fraction in this region. Both regions are orthogonal to the analysis signal regions and intended to validate the predictions of the matrix method for different levels of fake/non-prompt lepton contamination.

Some example distributions are shown in figure 11 for VR-3ℓ-1b3j and figure 12 for VR-3ℓ-1b3j-noZ. The processes with three real leptons (modelled with MC simulations) plus the prediction from the matrix method can be compared with data in these regions.

The hatched bands in figures 11 and 12 show only the statistical uncertainties of the MC prediction and the uncertainties associated with the fake/non-prompt lepton estimates (i.e. no theoretical or detector-related systematic uncertainties are included). The total uncertainty associated with the fake/non-prompt lepton estimate itself contains a systematic component, which is evaluated from variations of the input fake/real efficiencies (described in the previous section), and the statistical uncertainty of the data sample to which the Poisson likelihood matrix method is applied.\(^{11}\)

There is generally good agreement between the data and the total background estimate, except that the background is overestimated at low \( \Delta R(\ell_1, \ell_2) \). One contribution to that difference is that the two leading (two highest-\( p_T \)) lepton candidates are likely to be real, yet when they are near each other they have a lower efficiency for satisfying the isolation criteria, and thus are misinterpreted as fake/non-prompt leptons by the matrix method. Analyses that are sensitive to such issues may benefit from imposing a minimum \( \Delta R \) requirement between leptons. A discrepancy is also observed in the higher \( p_T \) bins of figure 11(d). Events in these bins have three leptons with \( p_T \) above 50 GeV. As shown in figures 9 and 10, only a single \( p_T \) bin above 50 GeV is available for measuring \( \varepsilon_t \) and \( \varepsilon_\ell \), due to the limited number of events in the control regions, so variations above 50 GeV may be missed. This point was not investigated thoroughly since only a small fraction of events in the analysis were impacted.

\(^{11}\)The total uncertainties of the fake/non-prompt lepton estimates may be different from the uncertainties reported in ref. [46], as the number of loose leptons in these validation regions is larger than in the signal region of the \( t\bar{t}Z \) analysis and, therefore, the statistical uncertainties are smaller.
9.2 Model-independent search for new phenomena in multi-lepton final states

Many interesting new models for BSM physics predict final states with three or more leptons. The general multi-lepton search for new phenomena [48] agnostically considers such final states. Its aim is to be sensitive to BSM phenomena in often-overlooked corners of phase space. The background estimation for the multi-lepton search uses MC predictions to account for events that contain only real leptons, and the fake-factor method for events containing at least one fake/non-prompt lepton. The dominant sources of fake/non-prompt leptons are semileptonic heavy-flavour decays (primarily of $b$-hadrons), light-hadron decays, and misidentification of light hadrons as leptons. These mainly arise in $Z$ + jets and $t\bar{t}$ events.
Figure 12. Comparisons of the predicted and observed yields in VR-3ℓ-1b3j-noZ, with respect to the $p_T$ of the (a) leading and (b) subleading lepton candidates, (c) the scalar sum of the lepton and jet transverse momenta, $H_T$, and (d) the missing transverse momentum in the event, $E_T^{\text{miss}}$. The processes with three real leptons are modelled with MC simulation, while the contribution from fake/non-prompt leptons (dark red) comes from the Poisson likelihood matrix method as described above. The hatched band shows the uncertainty from MC statistics and the fake/non-prompt background estimate. The rightmost (leftmost) bins are inclusive and contain all events above (below) the $x$-axis ranges. The lower panel shows the ratio of data to the total SM prediction (sum of the real-lepton background contributions estimated with MC samples and the fake/non-prompt lepton contribution estimated with the matrix method). Further details are available in ref. [46].

9.2.1 Fake/non-prompt lepton selection

The fake factors are measured using events with a single lepton candidate, where the selection targets QCD dijet events. In order to prevent a bias in the fake factor due to trigger selection criteria [15, 16], the selected events are required to have fired a loose single-lepton trigger where isolation requirements are not imposed. However, due to the high rate of events that pass such triggers, a prescale factor is applied, which reduces the number of events available for measuring the fake factors (the luminosity of the prescaled samples is $\sim 0.5$ fb$^{-1}$ for electrons and $\sim 3.7$ fb$^{-1}$ for muons). Baseline lepton candidates must pass a common object selection, as detailed in ref. [48]. Electron candidates are required to pass either the “Loose” ID WP with calorimeter-
track-based isolation requirements, or the “Tight” ID WP with no isolation requirement. Baseline muon candidates are required to pass the “Medium” ID WP (“HighPt” for $p_T > 300$ GeV). Only leptons with $p_T > 25$ GeV and satisfying the longitudinal and transverse track impact parameter requirements (see sections 4.1 and 4.2) are considered.

More stringent selection criteria are imposed for tight lepton candidates. For muons, where fake/non-prompt candidates are mainly muons from semileptonic heavy-flavour decays, only the isolation criteria is modified: the tight selection requires that muons satisfy track-based isolation criteria. For electrons, where both light- and heavy-flavour hadrons are a non-negligible source of fake/non-prompt candidates, both the ID and isolation criteria are modified. Tight electrons must satisfy both the “Tight” ID WP and calorimeter- and track-based isolation requirements.

Additional selection requirements are imposed on the single-lepton sample to ensure a high purity of fake/non-prompt leptons, which reduces the statistical uncertainty of the computed fake factor: the $E_T^{\text{miss}}$ is required to be $< 25$ GeV ($< 40$ GeV) for events with electron (muon) candidates, and the number of jets in the event is required to be $\geq 1$ ($2$) for events with electron (muon) candidates. For muon-candidate events, there must also be at least one jet with $p_T > 35$ GeV (the “tag jet”), and the azimuthal angle between the muon candidate and this jet is required to be $> 2.7$ radians.

The fake factor is binned in both the $p_T$ and $|\eta|$ of the lepton candidate. The bins are defined to have tolerable statistical uncertainties while preventing sizeable differences between the fake-factor values in adjacent bins. The highest muon-$p_T$ bin includes all muon candidates with $p_T > 80$ GeV, as there are very few fake/non-prompt muons at higher transverse momenta.

### 9.2.2 Fake factors

The fake factors calculated for the multi-lepton analysis are shown in figures 13 and 14 for electrons and muons. The uncertainties in the fake factors are categorised and evaluated as described below.

There are statistical uncertainties due to the data and MC sample sizes in each fake-factor bin, which due to their small size are summed in quadrature into a single uncertainty.

In the regions enriched in fake/non-prompt leptons, MC predictions are used to subtract the real-lepton contribution from the data. The uncertainties in the MC contributions are propagated to the fake factors. The main contributions from real leptons in the single-lepton regions are from $Z + \text{jets}$, $W + \text{jets}$ and $t\bar{t}$ events. For the $Z/W + \text{jets}$ processes, an uncertainty of 5% is applied to the cross-section. For the $t\bar{t}$ process, uncertainties on the cross section due to renormalization and factorization scales ($\sim 3\%$) and PDF ($\sim 4\%$) are applied.

Extrapolating the fake factors from the single-lepton sample to the multi-lepton samples used in the analysis introduces an uncertainty because these samples differ in the kinematic distributions of fake/non-prompt leptons, and possibly also in the fake/non-prompt lepton composition. Two uncertainties are included to address the bias caused by imposing a $E_T^{\text{miss}}$ upper bound in the fake/non-prompt lepton estimation sample, and by imposing a $p_T$ requirement on the tag jet in the fake/non-prompt muon estimation sample. These uncertainties are estimated by varying the requirements on these variables upwards and downwards by 10 GeV.\footnote{Only the $E_T^{\text{miss}}$ variation is displayed separately in the plots, as the variation of the jet $p_T$ results in an uncertainty that is too small to be visible.} Plots showing the impact of these systematic effects on the fake factors are given in figures 13 and 14.
The composition uncertainty in the total event yields is then derived from the difference between estimates calculated with either the nominal fake factors oblivious to the presence of $b$-tagged jets in the events, or the alternative set that requires such a jet. This uncertainty is found to have a negligible impact on the analysis, since most of the jets in the signal region and in the sample used to measure the fake factors are not $b$-tagged.
9.2.3 Validation regions

Validation regions are defined using appropriate sub-selections of $e_{\text{ee}}\mu$ and $e\mu_{\mu}$ events. These are used to check that the computed fake factors extrapolate correctly from the regions where they are calculated to the regions in which they are applied. The “on-$Z$” validation region requires an SFOC lepton pair with a dilepton mass within 10 GeV of the $Z$ boson mass. The “off-$Z$” validation region also requires a SFOC pair of leptons, but requires the dilepton mass to fall outside of the $Z$-mass window. Only mixed-flavour final states are selected for these validation regions so that which lepton to choose as the third one, assumed to be the fake/non-prompt lepton, is unambiguous. The sources of fake/non-prompt leptons contribute in different ratios to the on-$Z$ and off-$Z$ validation regions: the on-$Z$ validation region is more sensitive to $Z + \text{jets}$ events than the off-$Z$ region, while the inverse is true for $t\bar{t}$ events, although in absolute terms, $Z + \text{jets}$ events are more numerous than $t\bar{t}$ events in both cases. Both validation regions target, through a $m_T$ requirement of $m_T(\ell, E_T^{\text{miss}}) < 40$ GeV, a third lepton that is likely to be fake/non-prompt. The union of the on-$Z$ and off-$Z$ validation regions is called the “fakes validation region”.

The variables of primary importance for this analysis are the invariant mass of all lepton candidates in the event ($m_{\text{inv}}$) and the $E_T^{\text{miss}}$. The signal regions are separated according to the
values of these quantities, as discussed in ref. [48]. The $m_{\text{inv}}$ distributions in the two validation regions are shown in figure 16, while the $E_T^{\text{miss}}$ distributions are shown in figure 17. Lastly, the electron and muon candidate $p_T$ distributions are shown in the fakes validation region in figure 18. The comparisons in figure 16 were also presented in ref. [48] (albeit in logarithmic scale) and thus include fitted normalisation factors for the $WZ + \text{jets}$ and $ZZ + \text{jets}$ backgrounds from their respective control regions (“post-fit”). On the other hand, the complementary distributions shown in figures 17 and 18 were obtained independently of this statistical analysis, and thus employ the unconstrained SM background normalisations and uncertainties (“pre-fit” distributions). The background estimate is consistent with the data within the statistical and systematic uncertainties.

Figure 15. Fake factors measured in bins of the lepton candidate’s $p_T$ and $|\eta|$ for events with (a,c) electron candidates and (b,d) muon candidates in events (a,b) without and (c,d) with $b$-tagged jets.

---

2023 JINST 18 T11004
**Figure 16.** Comparison between data and prediction for the $m_{\text{inv}}$ distribution in the (a) on-Z and (b) off-Z validation regions, after fitting the normalisation factors for the $WZ + \text{jets}$ and $ZZ + \text{jets}$ backgrounds and systematic uncertainties [48]. All uncertainties, systematic and statistical, are included. The leftmost (rightmost) bin is inclusive and contains all events with $m_{\text{inv}} < 200\,\text{GeV}$ ($> 500\,\text{GeV}$). The hatched grey area in these figures shows the total uncertainty. Further details are available in ref. [48].

**Figure 17.** Comparison between data and prediction for the $E_T^{\text{miss}}$ distribution in the (a) on-Z and (b) off-Z validation regions. All uncertainties, systematic and statistical, are included. The rightmost bin is inclusive and contains all events with $E_T^{\text{miss}} > 475\,\text{GeV}$. The hatched grey area in these figures shows the total uncertainty. Further details are available in ref. [48].
10 Conclusions

For physics analyses exploring signatures with one or more prompt leptons, background contributions due to fake/non-prompt leptons are often difficult to estimate in simulations. Therefore, data-driven methods are commonly used. Three related methods have been adopted by the ATLAS Collaboration as recommended tools: the asymptotic matrix method, the Poisson likelihood matrix method, and the fake-factor method. All three approaches depend on defining two categories of leptons, one of which (“tight”) is subject to the same identification and selection criteria as are used in the analysis. The other category (“loose”) adds additional lepton candidates with less stringent selection requirements. The union of the two sets is called the “baseline” sample. The criteria are typically defined such that the probability for a real baseline lepton to satisfy the tight criteria is substantially higher than the corresponding probability for a fake/non-prompt lepton. Then, the relative numbers of loose and tight leptons in the analysis sample can be used to estimate the contribution of fake/non-prompt leptons, either inclusively or differentially in any variables of interest.

Despite their similarities, the methods each have their own strengths and drawbacks. The asymptotic matrix method and fake-factor method provide a fake/non-prompt lepton weight for each event, which is convenient for analyses. However, these methods are subject to large uncertainties if the efficiency for baseline fake/non-prompt leptons to satisfy the tight criteria is large in parts of the analysis phase space. The Poisson likelihood matrix method returns a smaller uncertainty in such cases, and avoids any possibility of producing a negative estimate for the event yield, but does not provide a per-event weight, introducing difficulties e.g. for differential estimations. The fake-factor method uses simulation rather than data to incorporate the contribution from events
where all leptons are tight. This method can therefore be employed while the signal region for an analysis is fully blinded, although it may induce additional simulation-related uncertainties in the background estimate.

The systematic uncertainties for all three methods arise from similar sources, with the largest contributions related to the extrapolation of the efficiencies measured in the control samples to events in the analysis sample. Differences in the fake/non-prompt lepton composition in the samples must be accounted for and appropriate uncertainties must be assigned to this extrapolation.

The performance of the Poisson likelihood matrix method and the fake-factor method has been demonstrated in a differential $t\bar{t}Z$ cross section measurement and in a model-independent search for BSM phenomena in multi-lepton final states, respectively. In both cases, the chosen method was shown to provide a reliable estimate of the fake/non-prompt lepton background, as measured in validation regions.

**Acknowledgments**

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNR and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MNE/IFA, Romania; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DS/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TENMAK, Türkiye; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; PRIMUS 21/SCI/017 and UNCE SCI/013, Czech Republic; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristea programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and MINERVA, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafsson’s Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [49].
References

[1] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, 2008 JINST 3 S08003 [nSPIRE].

[2] ATLAS collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, 2019 JINST 14 P12006 [arXiv:1908.00005] [nSPIRE].

[3] ATLAS collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C 81 (2021) 578 [arXiv:2012.00578] [nSPIRE].

[4] D0 collaboration, *Extraction of the width of the $W$ boson from measurements of $\sigma(p\bar{p} \to W + X) \times B(W \to e\nu)$ and $\sigma(p\bar{p} \to Z + X) \times B(Z \to ee)$ and their ratio*, Phys. Rev. D 61 (2000) 072001 [hep-ex/9906025] [nSPIRE].

[5] J. Erdmann et al., *Reformulation of a likelihood approach to fake-lepton estimation in the framework of Bayesian inference*, Nucl. Instrum. Meth. A 1021 (2022) 165939 [arXiv:2106.13628] [nSPIRE].

[6] T.P.S. Gillam and C.G. Lester, *Improving estimates of the number of ‘fake’ leptons and other mis-reconstructed objects in hadron collider events: BoB’s your UNCLE*, JHEP 11 (2014) 031 [arXiv:1407.5624] [nSPIRE].

[7] ATLAS collaboration, *Observation of inclusive electrons in the ATLAS experiment at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-073 (2010).

[8] G. Bohm and G. Zech, *Statistics of weighted Poisson events and its applications*, Nucl. Instrum. Meth. A 748 (2014) 1 [arXiv:1309.1287] [nSPIRE].

[9] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. 10 (1975) 343 [nSPIRE].

[10] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A 389 (1997) 81 [nSPIRE].

[11] ATLAS collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-2010-19 (2010).

[12] ATLAS IBL collaboration, *Production and integration of the ATLAS Insertable B-Layer*, 2018 JINST 13 T05008 [arXiv:1803.00844] [nSPIRE].

[13] ATLAS collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C 77 (2017) 317 [arXiv:1611.09661] [nSPIRE].

[14] ATLAS collaboration, *The ATLAS collaboration software and firmware*, ATL-SOFT-PUB-2021-001 (2021).

[15] ATLAS collaboration, *Performance of the ATLAS muon triggers in Run 2*, 2020 JINST 15 P09015 [arXiv:2004.13447] [nSPIRE].

[16] ATLAS collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C 80 (2020) 47 [arXiv:1909.00761] [nSPIRE].

[17] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C 74 (2014) 3071 [arXiv:1407.5063] [nSPIRE].

[18] M. Cacciari, G.P. Salam and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, JHEP 04 (2008) 063 [arXiv:0802.1189] [nSPIRE].
[19] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C 72 (2012) 1896 [arXiv:1111.6097] [SPIRE].
[20] M. Cacciari and G.P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B 659 (2008) 119 [arXiv:0707.1378] [SPIRE].
[21] ATLAS collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D 97 (2018) 072003 [arXiv:1712.08891] [SPIRE].
[22] ATLAS collaboration, *Search for chargino–neutralino pair production in final states with three leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C 81 (2021) 1118 [arXiv:2106.01676] [SPIRE].
[23] ATLAS collaboration, *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*, ATL-PHYS-PUB-2018-009 (2018).
[24] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP 06 (2010) 043 [arXiv:1002.2581] [SPIRE].
[25] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. 191 (2015) 159 [arXiv:1410.3012] [SPIRE].
[26] ATLAS collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021 (2014).
[27] NNPDF collaboration, *Parton distributions for the LHC Run II*, JHEP 04 (2015) 040 [arXiv:1410.8849] [SPIRE].
[28] R.D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B 867 (2013) 244 [arXiv:1207.1303] [SPIRE].
[29] D.J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A 462 (2001) 152 [SPIRE].
[30] ATLAS collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006 (2017).
[31] SHERPA collaboration, *Event Generation with Sherpa 2.2*, SciPost Phys. 7 (2019) 034 [arXiv:1905.09127] [SPIRE].
[32] ATLAS collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C 70 (2010) 823 [arXiv:1005.4568] [SPIRE].
[33] GEANT4 collaboration, *GEANT4 — a simulation toolkit*, Nucl. Instrum. Meth. A 506 (2003) 250 [SPIRE].
[34] ATLAS collaboration, *Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C 83 (2023) 686 [arXiv:2212.07338] [SPIRE].
[35] ATLAS collaboration, *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C 81 (2021) 689 [arXiv:2007.02645] [SPIRE].
[36] ATLAS collaboration, *Charged-particle distributions in $\sqrt{s} = 13$ TeV pp interactions measured with the ATLAS detector at the LHC*, Phys. Lett. B 758 (2016) 67 [arXiv:1602.01633] [SPIRE].
[37] ATLAS collaboration, *Search for squarks and gluinos in final states with same-sign leptons and jets using 139 fb$^{-1}$ of data collected with the ATLAS detector*, JHEP 06 (2020) 046 [arXiv:1909.08457] [SPIRE].
[38] R.H. Dalitz, *On an alternative decay process for the neutral pi-meson*, Letters to the Editor, *Proc. Phys. Soc. A* 64 (1951) 667 [SPIRE].

[39] ATLAS collaboration, *Proposal for particle-level object and observable definitions for use in physics measurements at the LHC*, ATL-PHYS-PUB-2015-013 (2015).

[40] ATLAS collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at \( \sqrt{s} = 13 \) TeV*, *Eur. Phys. J. C* 78 (2018) 903 [arXiv:1802.08168] [SPIRE].

[41] ATLAS collaboration, *Measurement of \( W^+W^- \) vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector*, *Phys. Rev. D* 96 (2017) 012007 [arXiv:1611.02428] [SPIRE].

[42] ATLAS collaboration, *Search for supersymmetry in events with four or more charged leptons in 139 fb\(^{-1}\) of \( \sqrt{s} = 13 \) TeV pp collisions with the ATLAS detector*, *JHEP* 07 (2021) 167 [arXiv:2103.11684] [SPIRE].

[43] ATLAS collaboration, *ATLAS data quality operations and performance for 2015-2018 data-taking*, 2020 *JINST* 15 P04003 [arXiv:1911.04632] [SPIRE].

[44] ATLAS collaboration, *Luminosity determination in pp collisions at \( \sqrt{s} = 13 \) TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021 (2019).

[45] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, 2018 *JINST* 13 P07017 [SPIRE].

[46] ATLAS collaboration, *Measurements of the inclusive and differential production cross sections of a top-quark-antiquark pair in association with a Z boson at \( \sqrt{s} = 13 \) TeV with the ATLAS detector*, *Eur. Phys. J. C* 81 (2021) 737 [arXiv:2103.12603] [SPIRE].

[47] ATLAS collaboration, *ATLAS b-jet identification performance and efficiency measurement with t\bar{t} events in pp collisions at \( \sqrt{s} = 13 \) TeV*, *Eur. Phys. J. C* 79 (2019) 970 [arXiv:1907.05120] [SPIRE].

[48] ATLAS collaboration, *Search for new phenomena in three- or four-lepton events in pp collisions at \( \sqrt{s} = 13 \) TeV with the ATLAS detector*, *Phys. Lett. B* 824 (2022) 136832 [arXiv:2107.00404] [SPIRE].

[49] ATLAS collaboration, *ATLAS computing acknowledgements*, ATL-SOFT-PUB-2021-003, (2021).
The ATLAS collaboration

G. Aad102, B. Abbott120, D.C. Abbott103, K. Abeling135, S.H. Abidi29, A. Abouhorma35e, H. Abramowicz151, H. Abreu150, Y. Abulaiti117, A.C. Abusleme Hoffman137a, B.S. Acharya49a, 49b, 49c, A. Adam Bourdarios1, L. Adamczyk85a, L. Adamek135, S.V. Addepalli26, J. Adelman115, A. Adiguzel21c, S. Adorni56, T. Adye134, A.A. Affolder136, Y. Afik36, M.N. Agarwal13, J. Agarwala73a, 73b, A. Aggarwal100, C. Agheorghiesei27c, J.A. Aguilar-Saavedra130f, A. Ahmad36, F. Ahmadov38, 38b, W.S. Ahmed100, S. Ahuja95, X. Ai48, G. Aielli76a, 76b, M. Ai Tamliah135e, B. Aitbenkhish35a, I. Aizenberg169, M. Akhiyik100, T.P.A. Åkesson98, A.V. Akimov37, K. Al Khoury41, G.L. Alberghi23b, J. Albert165, P. Albicocco53, S. Alderweireldt52, M. Alekseev36, I.N. Aleksandrov38, C. Alexa27b, T. Alexopoulos10, A. Alfonsi114, F. Alfonso23b, M. Albroob120, B. Ali132, S. Ali148, M. Arie37, G. Alimonti71a, W. Alkakhi55, C. Allaire66, B.M.M. Allbrooke146, C.A. Allendes Flores137f, P.P. Allport20, A. Aloisio72a, 72b, F. Alonso90, C. Alpigiani138, M. Alvarez Estevez99, M.G. Alviggi72a, 72b, M. Aly101, Y. Amaral Coutinho82a, A. Ambl器104, C. Amelung36, M. Ameri01, C.G. Ames109, D. Amidei106, S.P. Amor Dos Santos130a, K.R. Amos163, V. Ananiev125, C. Anastopoulos139, T. Andeen11, J.K. Anders36, S.Y. Andreae47a, 47b, A. Andreazza71a, 71b, S. Angelidakis67, A. Angerami41, 41e, A.V. Anisenkov37, A. Anov84, C. Antel56, M.T. Anthony139, E. Antipov121, M. Antonelli53, D.J.A. Antirini17a, F. Anulli75a, M. Aoki83, T. Aoki153, J.A. Aparisi Pozo163, M.A. Aparo146, L. Aperio Bella48, C. Appelt18, N. Aranzabal36, V. Araujo Ferraz82a, C. Arcangeletti53, A.T.H. Arce51, E. Arena92, J-F. Arguin108, S. Argyropoulos54, J.-H. Arling48, A.J. Armbruster36, O. Arnaez155, H. Arnold114, Z.P. Arrubarrena Tame109, G. Artoni75a, 75b, H. Asada111, K. Asai118, S. Asai153, N.A. Asbab61, J. Assalhah135d, K. Assamagan29, R. Astalos28a, R.J. Atkin31a, M. Atkinson162, N.B. Atlay18, H. Atmani62, P.A. Atmasiddha106, K. Augsten132, S. Auricchio72a, 72b, A.D. Auriol20, V.A. Austrup171, G. Avner150, G. Avolio36, K. Axioti36, M.K. Ayoub14c, G. Azuelos108, 108a, 108j, D. Babal28a, H. Bachacou135, K. Bachas152a, A. Bachiu34, F. Backman47a, 47b, A. Badea61, P. Bagnaia75a, 75b, M. Bahmann18, A.J. Bailey163, V.R. Bailey162, J.T. Baines134, C. Bakalis10, O.K. Baker172, P.J. Bakker114, E. Bakos15, D. Bakshi Gupta8, S. Balaji147, R. Balasubramanian114, E.M. Baldwin37, P. Balek135, E. Labarbari71a, 71b, F. Balli135, L.M. Baltes63a, W.K. Balunas32, J. Balz100, E. Banas66, M. Bandieramonte129, A. Bandoypadhayay24, S. Bansal24, L. Barak151, E.L. Barberio105, D. Barberis57, 57a, M. Barbero102, G. Barbour96, K.N. Barends33a, T. Barillari110, M-S. Barisits36, T. Barklow143, R.M. Barnett17a, P. Baron122, D.A. Baron Moreno101, A. Baronecchi62, G. Barone29, A.J. Barr126, L. Barranco Navarro47a, 47b, F. Barreiro99, J. Barreiro Guimarães da Costa14a, U. Barron151, M.G. Barros Teixeira130a, S. Barsov37, F. Bartels63a, R. Bartoldus143, A.E. Barton91, P. Bartos28a, A. Basa48, A. Basan100, M. Baselga49, I. Bashta77a, 77b, A. Bassalat66, M.J. Basso155, C.R. Basson101, R.L. Bates59, S. Batlamous35e, J.R. Batley32, B. Batool141, M. Battaglia136, D. Battiluga18, M. Baur75a, 75b, P. Bauer24, A. Bayirli21a, J.B. Beacham51, T. Beau127, P.H. Beaucourn158, F. Becherer54, P. Bechtel24, H.P. Beck19, K. Becker167, A.J. Beddall21d, V.A. Bednyakov38, C.P. Bee145, L.J. Beemster15, T.A. Beermann16, M. Begalli82d, M. Begel29, A. Behera145, J.K. Behr48, C. Beirao Da Cruz E Silva36,
H. Torres, E. Torró Pastor, M. Toscani, C. Tocci, M. Tost, D.R. Tovey, A. Traeff, I.S. Trandafir, T. Trefzger, A. Tricoli, I.M. Trigger, S. Trincav-Duvoit, D.A. Trischuk, B. Trocmé, A. Trofymov, C. Troncon, L. Truong, M. Trzebinski, A. Trzupek, F. Tsai, M. Tsai, A. Tsiamis, P.V. Tsiareishkha, S. Tsigaridas, A. Tsirigotis, V. Tsiskaridze, E.G. Tskhadadze, M. Tsupolou, Y. Tsigiakawa, I.I. Tsukerman, V. Tsvaia, S. Tsuno, O. Tsur, D. Tsuybychev, Y. Tu, A. Tudorache, V. Tudorache, A.N. Tuna, S. Turchikhin, I. Turk Cakir, R. Turra, T. Turtuvsbin, P.M. Tuts, S. Tzamarias, P. Tzanis, E. Tzovara, K. Uchida, F. Ukegawa, P.A. Ulloa Poblete, E.N. Umaka, G. Unal, M. Unal, A. Undrus, G. Unel, J. Urban, P. Urquijo, G. Usai, R. Ushioda, M. Usman, Z. Uysal, L. Vacavant, V. Vacek, B. Vachon, K.O.H. Vadia, T. Vafeiadis, A. Vaitkus, C. Valderanis, S. Valentinetti, A. Valero, A. Vallier, J.A. Valls Ferrer, D.R. Van Arnum, T.R. Van Daalen, P. Van Gemmeren, M. Van Rijnbach, S. Van Stroud, I. Van Vulpen, M. Vanadria, W. Vandelli, M. Vandenbroucke, E.R. Vandewall, D. Vannicola, L. Vannoli, R. Vartiainen, E.W. Varnes, C. Varzi, T. Varol, D. Varouchas, L. Varriale, K.E. Varvell, M.E. Vasile, F. Vazeille, T. Vazquez Schroeder, J. Veatch, V. Vecchio, M.J. Veen, I. Velisek, L.M. Veloce, F. Veloso, S. Veneziano, A. Ventura, A. Verbytskyi, M. Verducci, C. Vergis, M. Verissimo De Araujo, W. Verkerke, J.C. Vermeulen, C. Vernieri, P.J. Verschuuren, M. Vessella, M.C. Vetterli, A. Vgenopoulos, N. Viaux Maira, T. Vici, O.E. Vickey Boeriu, G.H.A. Viehbauer, L. Viganò, M. Villa, M. Villaplana Perez, E.M. Villhauer, E. Viluchti, M.G. Vincer, G.S. Virdee, A. Vishwakarma, C. Vittori, I. Vivarelli, V. Vladimirov, E. Voevodina, F. Vogel, P. Vokac, J. Von Ahnen, E. Von Toerne, B. Vormwald, K. Vorobev, M. Vox, K. Voss, J.H. Vossebold, M. Vozak, L. Vozdecky, N. Vranjes, M. Vranjes Milosavljevic, M. Vreeswijk, R. Vuillermet, M. Vranjes Milosavljevic, R. Vuckovic, L. Vujkovic, S. Vada, C. Wagner, W. Wagner, S. Wahdan, H. Wahlberg, R. Wakasa, M. Wakida, V.M. Walbrecht, J. Walder, R. Walker, W. Walkowiak, A.M. Wang, A.Z. Wang, C. Wang, H. Wang, J. Wang, R. Wang, R. Wang, S.M. Wang, S. Wang, T. Wang, W. Wang, X. Wang, X. Wang, Y. Wang, Y. Wang, Z. Wang, A. Warburton, R.J. Ward, N. Warrack, A.T. Watson, H. Watson, M.F. Watson, G. Watts, B.M. Waugh, A.F. Webb, C. Weber, H.A. Weber, M.S. Weber, S.M. Weber, C. Wei, Y. Wei, A.R. Weidberg, J. Weingarten, M. Weirich, C. Weiser, C.J. Wells, T. Wenaus, B. Wendland, T. Wengler, N.S. Wenke, N. Wermes, M. Wessels, K. Whalen, A.M. Wharton, A.S. White, J.M. White, D. Whiteson, L. Wickremasinghe, W. Wiedenmann, C. Wiel, M. Wielers, C. Wiglesworth, L.A.M. Wiik-Fuchs, D.J. Wilbern, H.G. Wilkens, D.M. Williams, H.H. Williams, S. Williams, S. Willcocks, P.J. Windischhofer, F. Winklmeier, B.T. Winter, J.K. Winter, M. Wittgen, M. Wobisch, R. Wölker, J. Wollrath, M.W. Wolter, H. Wolters.
53 INFN e Laboratori Nazionali di Frascati, Frascati; Italy
54 Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany
55 II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen; Germany
56 Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
57 (a) Dipartimento di Fisica, Università di Genova, Genova; (b) INFN Sezione di Genova; Italy
58 II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen; Germany
59 LPSC - School of Physics and Astronomy, University of Glasgow, Glasgow; United Kingdom
60 INFN e Laboratori Nazionali di Frascati, Frascati; Italy
61 Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA; United States of America
62 (a) Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei; (b) Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao; (c) School of Physics and Astronomy, Shanghai Jiao Tong University, Key Laboratory for Particle Astrophysics and Cosmology (MOE), SKLPPC, Shanghai; (d) Tsung-Dao Lee Institute, Shanghai; China
63 (a) Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg; (b) Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg; Germany
64 (a) Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong; (b) Department of Physics, University of Hong Kong, Hong Kong; (c) Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; China
65 Department of Physics, National Tsing Hua University, Hsinchu; Taiwan
66 JCU, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay; France
67 Centro Nacional de Microelecetrónica (IMB-CNM-CSIC), Barcelona; Spain
68 Department of Physics, Indiana University, Bloomington IN; United States of America
69 (a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine; (b) ICTP, Trieste; (c) Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine; Italy
70 (a) INFN Sezione di Lecce; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce; Italy
71 (a) INFN Sezione di Milano; (b) Dipartimento di Fisica, Università di Milano, Milano; Italy
72 (a) INFN Sezione di Napoli; (b) Dipartimento di Fisica, Università di Napoli, Napoli; Italy
73 (a) INFN Sezione di Pavia; (b) Dipartimento di Fisica, Università di Pavia, Pavia; Italy
74 (a) INFN Sezione di Pisa; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa; Italy
75 (a) INFN Sezione di Roma; (b) Dipartimento di Fisica, Sapienza Università di Roma, Roma; Italy
76 (a) INFN Sezione di Roma Tor Vergata; (b) Dipartimento di Fisica, Università di Roma Tor Vergata, Roma; Italy
77 (a) INFN Sezione di Roma Tre; (b) Dipartimento di Matematica e Fisica, Università Roma Tre, Roma; Italy
78 (a) INFN-INFN-TIFPA; (b) Università degli Studi di Trento, Trento; Italy
79 Università Innsbruck, Department of Astro and Particle Physics, Innsbruck; Austria
80 University of Iowa, Iowa City IA; United States of America
81 Department of Physics and Astronomy, Iowa State University, Ames IA; United States of America
82 (a) Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora; (b) Universidade Federal do Rio de Janeiro COPPE/EE/IF, Rio de Janeiro; (c) Instituto de Física, Universidade de São Paulo, São Paulo; (d) Rio de Janeiro State University, Rio de Janeiro; Brazil
83 KEK, High Energy Accelerator Research Organization, Tsukuba; Japan
84 Graduate School of Science, Kobe University, Kobe; Japan
85 (a) AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow; Poland
86 Institute of Nuclear Physics Polish Academy of Sciences, Krakow; Poland
87 Faculty of Science, Kyoto University, Kyoto; Japan
88 Kyoto University of Education, Kyoto; Japan
89 Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka; Japan
90 Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata; Argentina
91 Physics Department, Lancaster University, Lancaster; United Kingdom
92 Oliver Lodge Laboratory, University of Liverpool, Liverpool; United Kingdom
93 Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana; Slovenia
Department of Physics, Santiago; Instituto de Alta Investigación, Universidad de Tarapacá, Arica; Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso; Chile

Department of Physics, University of Washington, Seattle WA; United States of America

Department of Physics and Astronomy, University of Sheffield, Sheffield; United Kingdom

Department of Physics, Shinshu University, Nagano; Japan

Department Physik, Universität Siegen, Siegen; Germany

Department of Physics, Simon Fraser University, Burnaby BC; Canada

SLAC National Accelerator Laboratory, Stanford CA; United States of America

Department of Physics, Royal Institute of Technology, Stockholm; Sweden

Departments of Physics and Astronomy, Stony Brook University, Stony Brook NY; United States of America

Department of Physics and Astronomy, University of Sussex, Brighton; United Kingdom

School of Physics, University of Sydney, Sydney; Australia

Institute of Physics, Academia Sinica, Taipei; Taiwan

also affiliated with an institute covered by a cooperation agreement with CERN

E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi; High Energy Physics Institute, Tbilisi State University, Tbilisi; University of Georgia, Tbilisi; Georgia

Also at An-Najah National University, Nablus; Palestine

Also at Borough of Manhattan Community College, City University of New York, New York NY; United States of America

Also at Bruno Kessler Foundation, Trento; Italy

Also at Center for High Energy Physics, Peking University; China

Also at Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Thessaloniki; Greece

Also at Centro Studi e Ricerche Enrico Fermi; Italy

Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Department de Fisica de la Universitat Autonoma de Barcelona, Barcelona; Spain

Also at Department de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departement de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departamento de Física de la Universitat Autonoma de Barcelona, Barcelona; Spain

Also at TRIUMF, Vancouver BC; Department of Physics and Astronomy, York University, Toronto ON; Canada

Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba; Japan

Department of Physics, Tokyo Institute of Technology, Tokyo; Japan

Department of Physics, University of Toronto, Toronto ON; Canada

Department of Physics, Tsukuba University, Tsukuba; Japan

United States of America

United Arab Emirates

United Kingdom

United States of America

United States of America

Germany

University of Victoria, Victoria BC; Canada

Waseda University, Tokyo; Japan

Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot; Israel

Department of Physics, University of Wisconsin, Madison WI; United States of America

Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal; Germany

Department of Physics, Yale University, New Haven CT; United States of America

Department of Physics, University of British Columbia, Vancouver BC; Canada

Department of Physics and Astronomy, University of Illinois, Urbana IL; United States of America

Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia - CSIC, Valencia; Spain

Department of Physics and Astronomy, Tufts University, Medford MA; United States of America

United Arab Emirates University, Al Ain; United Arab Emirates

Department of Physics and Astronomy, University of California Irvine, Irvine CA; United States of America

Department of Physics and Astronomy, University of Uppsala, Uppsala; Sweden

Department of Physics, University of Illinois, Urbana IL; United States of America

Department of Physics, Technion, Israel Institute of Technology, Haifa; Israel

Department of Physics, Aristotle University of Thessaloniki, Thessaloniki; Greece

Department of Physics, Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv; Israel

International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo; Japan

Department of Physics, University of Washington, Seattle WA; United States of America

Department of Physics, Royal Institute of Technology, Stockholm; Sweden

Department of Physics, Simon Fraser University, Burnaby BC; Canada

Department of Physics and Astronomy, University of Sheffield, Sheffield; United Kingdom

Department of Physics, Shinshu University, Nagano; Japan

Department of Physics, Simon Fraser University, Burnaby BC; Canada

Department of Physics, University of Sheffield, Sheffield; United Kingdom

Department of Physics, Tokyo Institute of Technology, Tokyo; Japan

Department of Physics, University of Sussex, Brighton; United Kingdom

Department of Physics, Technion, Israel Institute of Technology, Haifa; Israel

Also at An-Najah National University, Nablus; Palestine

Also at Borough of Manhattan Community College, City University of New York, New York NY; United States of America

Also at Bruno Kessler Foundation, Trento; Italy

Also at Center for High Energy Physics, Peking University; China

Also at Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Thessaloniki; Greece

Also at Centro Studi e Ricerche Enrico Fermi; Italy

Also at CERN, Geneva; Switzerland

Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departament de Física de la Universitat Autonoma de Barcelona, Barcelona; Spain

Also at Department de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departement de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departamento de Física de la Universitat Autonoma de Barcelona, Barcelona; Spain

Also at Department de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departamento de Física de la Universitat Autonoma de Barcelona, Barcelona; Spain

Also at Department de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland

Also at Departamento de Física de la Universitat Autonoma de Barcelona, Barcelona; Spain
