Unveiling Hidden Physics at the LHC

Oliver Fischer†,1, Bruce Mellado†,2,3, Stefan Antusch4, Emanuele Bagnaschi5, Shankha Banerjee6, Geoff Beck2, Benedetta Belfatto7,8, Matthew Bellis9, Zurab Berezhiani10,11, Monika Blanke12,13, Bernat Capdevila14,15, Kingman Cheung16, Andreas Crivellin5,6,17, Nishita Desai18, Bhupal Dev19, Rohini Godbole20, Tao Han21, Philip Harris22,23, Martin Hoferichter24, Matthew Kirk25,26, Suchita Kulkarni27, Clemens Lange28, Kati Lassila-Perini29, ZHEN Liu30, Farvah Mahmoudi5,31, Claudio Andrea Manzari5,17, David Marzocca32, Biswarup Mukhopadhyaya33, Antonio Pich34, Xifeng Ruan2, Luc Schnell35,36, Jesse Thaler22,23, and Susanne Westhoff27

1Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZL, UK
2School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa.
3iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa.
4Department of Physics, University of Basel, Klingelbergstr. 82, CH-4056 Basel, Switzerland
5Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland
6CERN Theory Division, CH-1211 Geneva 23, Switzerland
7Dipartimendto di Fisica “E. Fermi”, Università di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy
8INFN Sezione di Pisa, Largo Bruno Pontecorvo 3, I-56127 Pisa, Italy
9Siena College, 515 Loudon Road, Loudonville, NY 12211-1462, United States
10Dipartimento di Fisica e Chimica, Università di L’Aquila, 67100 Coppito, L’Aquila, Italy
11INFN, Laboratori Nazionali del Gran Sasso, 67100 Assergi, L’Aquila, Italy
12Institute for Astroparticle Physics (IAP), Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, D-76344 Eggenstein-Leopoldshafen, Germany
13Institute for Theoretical Particle Physics (TTP), Karlsruhe Institute of Technology, Engesserstrasse 7, D-76128 Karlsruhe, Germany
14Dipartimento di Fisica, Università di Torino, Via P. Giuria 1, Torino I-10125, Italy
15INFN Sezione di Torino, Via P. Giuria 1, Torino I-10125, Italy
16Department of Physics, National Tsing Hua University, Hsinchu 300, Taiwan
17Physik-Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
18Tata Institute of Fundamental Research, 1 Homi Bhabha Road, Mumbai 400005, India
19Department of Physics and McDonnell Center for the Space Sciences, Washington University, St. Louis, MO 63130, USA
20Indian Institute of Science (IISc), Bangalore 560012, Karnataka, India
21Pittsburgh Particle Physics, Astrophysics, and Cosmology Center, Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15206, USA
22Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
23The NSF AI Institute for Artificial Intelligence and Fundamental Interactions
24Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
25Dipartimento di Fisica, Università di Roma “La Sapienza”, Piazzale Aldo Moro 2, 00185 Roma,
26 INFN Sezione di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
27 Institute of Physics, NAWI Graz, University of Graz, Universitätsplatz 5, A-8010 Graz, Austria
28 Experimental Physics Department, CERN, Switzerland
29 Helsinki Institute of Physics (HIP), P.O. Box 64, 00014 University of Helsinki, Finland
30 School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
31 Univ Lyon, Univ Lyon 1, CNRS/IN2P3, Institut de Physique des 2 Infinis de Lyon, UMR 5822, F-69622, Villeurbanne, France
32 INFN, Sezione di Trieste, SISSA, Via Bonomea 265, 34136, Trieste, Italy
33 Indian Institute of Science Education and Research, Kolkata, India
34 IFIC, Universitat de València – CSIC, E-46980 Paterna, Spain
35 Departement Physik, ETH Zürich, Otto-Stern-Weg 1, CH-8093 Zürich, Switzerland
36 Département de Physique, École Polytechnique, Route de Saclay, FR-91128 Palaiseau Cedex, France
37 Institute for Theoretical Physics, Heidelberg University, 69120 Heidelberg, Germany

Editors
Abstract

The field of particle physics is at the crossroads. The discovery of a Higgs-like boson completed the Standard Model (SM), but the lacking observation of convincing resonances Beyond the SM (BSM) offers no guidance for the future of particle physics. On the other hand, the motivation for New Physics has not diminished and is, in fact, reinforced by several striking anomalous results in many experiments. Here we summarise the status of the most significant anomalies, including the most recent results for the flavour anomalies, the multi-lepton anomalies at the LHC, the Higgs-like excess at around 96 GeV, and anomalies in neutrino physics, astrophysics, cosmology, and cosmic rays.

While the LHC promises up to 4 ab$^{-1}$ of integrated luminosity and far-reaching physics programmes to unveil BSM physics, we consider the possibility that the latter could be tested with present data, but that systemic shortcomings of the experiments and their search strategies may preclude their discovery for several reasons, including: final states consisting in soft particles only, associated production processes, QCD-like final states, close-by SM resonances, and SUSY scenarios where no missing energy is produced.

New search strategies could help to unveil the hidden BSM signatures, devised by making use of the CERN open data as a new testing ground. We discuss the CERN open data with its policies, challenges, and potential usefulness for the community. We showcase the example of the CMS collaboration, which is the only collaboration regularly releasing some of its data. We find it important to stress that individuals using public data for their own research does not imply competition with experimental efforts, but rather provides unique opportunities to give guidance for further BSM searches by the collaborations.

Wide access to open data is paramount to fully exploit the LHCs potential.
# Contents

1 Introduction 6

2 Anomalies 9

2.1 Flavour Anomalies ................................................. 9
2.1.1 $b \rightarrow s \ell^+ \ell^-$ ........................................ 10
2.1.2 Tauonic $B$-meson decays .................................. 16
2.1.3 Anomalous magnetic moments of charged leptons .......... 16
2.1.4 Cabibbo Angle Anomaly ....................................... 20
2.1.5 Lepton flavour universality in the charged current ....... 23
2.1.6 Non-resonant di-electrons .................................. 24
2.1.7 Forward-backward asymmetry in semi leptonic $B$ decays .. 27
2.1.8 Combined explanations ....................................... 27
2.2 Multi-lepton anomalies at the LHC ................................ 28
2.2.1 Opposite sign di-leptons .................................... 29
2.2.2 SS and 3$\ell$ with $b$-quarks ................................. 29
2.2.3 SS and 3$\ell$ without $b$-quarks .............................. 30
2.3 Higgs-like excess at $\approx 96$ GeV ................................. 30
2.3.1 Interpretation as SM Higgs-like scalar boson ............... 31
2.3.2 Two Higgs doublet model .................................... 31
2.3.3 Supersymmetric models ..................................... 31
2.3.4 Other explanations .......................................... 32
2.4 Neutrino Anomalies .............................................. 32
2.4.1 Appearance anomalies ....................................... 32
2.4.2 Disappearance anomalies .................................... 33
2.4.3 Explanation via light neutrinos ............................... 33
2.4.4 Non oscillatory explanations of the appearance anomaly .. 33
2.5 Astrophysical Anomalies .......................................... 34
2.5.1 Positron excesses ........................................... 34
2.5.2 Anti-protons ................................................... 35
2.5.3 Galactic Centre gamma-ray excess .......................... 35
2.5.4 3.5 keV X-ray line ........................................... 35
2.5.5 511 keV gamma-ray line ..................................... 35
2.5.6 21-cm absorption line ........................................ 35
2.5.7 XENON1T electron excess .................................... 35
2.5.8 DAMA/LIBRA annual modulation ............................ 36
2.6 Cosmological Anomalies .......................................... 36
2.6.1 Dark energy .................................................. 36
2.6.2 Hubble trouble ............................................... 36
2.6.3 Clustering anomalies ......................................... 38
2.7 Anomalies in ultra-high energy cosmic rays and neutrinos .... 38
2.7.1 The muon puzzle ............................................. 38
2.7.2 Earth-emergent EeV events .................................. 38
3 Hidden physics at the LHC

3.1 Properties of hidden signatures
   3.1.1 Soft Particles in the Final State
   3.1.2 SUSY without missing transverse energy
   3.1.3 Associated production processes
   3.1.4 QCD-like final states
   3.1.5 Close-by SM resonances

3.2 Hidden SUSY scenarios
   3.2.1 Wino and Higgsino LSP
   3.2.2 Stealth SUSY
   3.2.3 $R$-parity violation
   3.2.4 Long-lived NLSP
   3.2.5 $\nu$CMSSM with a long-lived stau

3.3 Hidden portal models
   3.3.1 The Neutrino portal
   3.3.2 The Higgs portal
   3.3.3 The vector portal

3.4 Long-Lived Particles
   3.4.1 The experimental view
   3.4.2 Signatures
   3.4.3 Theoretical motivation
   3.4.4 Backward moving objects from TeV LLP

3.5 Summary

4 Open data in particle physics

4.1 Challenges
   4.1.1 The problem with triggers
   4.1.2 The role of the collaborations
   4.1.3 Technical challenges
   4.1.4 The ultimate aim

4.2 Existing framework
   4.2.1 The FAIR principles
   4.2.2 Available data from the LHC collaborations
   4.2.3 CERN open data policy
   4.2.4 CERN Open Data Portal
   4.2.5 REANA and HEPdata

4.3 CMS open data
   4.3.1 CMS data releases
   4.3.2 Data complexity
   4.3.3 Examples and documentation
   4.3.4 Usability

4.4 Recommendations from External Users
   4.4.1 Missing information
   4.4.2 Validation samples
   4.4.3 Industry standard file format

4.5 Summary
   4.5.1 Current status
   4.5.2 Future improvements
1 Introduction

Section editors: Oliver Fischer and Bruce Mellado

The discovery of a scalar resonance that resembles the Higgs boson of the Standard Model (SM) \cite{1,2,3,4} at the Large Hadron Collider (LHC) by the ATLAS \cite{5} and CMS \cite{6} collaborations has opened a new chapter in particle physics. The combined measurements show that this discovered particle has properties that are compatible with those predicted by the SM \cite{7}, which makes its discovery a great triumph for experiment and theory. Under the assumption that the discovered scalar particle is indeed the predicted Higgs boson, the SM has exhausted all its predictions pertaining to fundamental particles. While the LHC collaborations continue to measure properties of the Higgs boson and other known particles and processes, the chief focus is now on the observation of new phenomena beyond the SM.

The motivation for the existence of New Physics (NP) is no less than before the discovery of the Higgs boson. The SM itself raises the question of naturalness, i.e. why the Electroweak scale is so much smaller than the Planck scale, which is addressed by Supersymmetry (SUSY) in an elegant way. The observation of Dark Matter (DM) in the Universe, interpreted as a fundamental particle, can be addressed with minimal frameworks beyond the SM (BSM) and also with theories that often introduce an entire dark sector with new particles and forces. The observation of Neutrino Oscillations implies that the neutrinos are massive, which requires a mass-generating mechanism and therefore an extension of the SM. The above list of arguments is incomplete but suggests convincingly that the existence of NP is an established fact. No indication exists, however, in what form the NP manifests itself.

As the nature and energy scale of NP remains unknown, new phenomena could emerge at any experiment. In recent years the field of particle physics has experienced a growing litany of anomalous experimental results. Many of them are statistically significant and continue to grow, and remain unexplained by state-of-the-art calculations based on the SM. In most cases the latter have become increasingly precise and reliable, where vast data sets have been used to provide extensive testing grounds. While some anomalies might eventually find explanations within the framework of the SM, persisting ones could contain hints of NP and may serve as a guide to model building and experimental searches.

While there is overwhelming evidence that the SM is incomplete, guidance is required to resolve the issue as to how the SM will conclusively breakdown in laboratory conditions. Hundreds of BSM models have been proposed over the years, motivated by the big open questions as well as by the various combinations of experimental anomalies. At the same time, extensive exploration of LHC data with respect to inclusive and model dependent signatures performed to date indicates that no striking resonances have been observed in the accessible dynamic range. The absence of clear BSM signatures in LHC data indicates that NP is either inaccessible at the LHC, or that it is driven by more subtle topologies and therefore hidden in the backgrounds.

The question arises what the absence of BSM resonances at the LHC implies for our search strategies. Many models have been identified that are not captured by current searches, such that reinterpretation of experimental limits for different models became an important topic of discussion \cite{8,9}, and analysis strategies were developed that are less model-dependent. Since model building is the driving force behind gaining insight into new signatures, the model-centric and the model-independent approach are both necessary. The community is
elaborating on data analysis methodologies that display less model dependencies. The use of Machine Learning may play a significant role.

It is high time to scrutinise existing LHC data with respect to clues on NP in the non-strongly interacting sector in order to prepare for the high-luminosity era and beyond. When the High Luminosity LHC starts operating, the data coming out of it has to be used to glean every possible information about existing NP models. Clear guiding principles need to come from a combination of both experimental and theoretical inquiries. Given the present discussions on future colliders all over the world, this guidance is now more crucial than ever before. Guidance for the future could come from NP that is currently hidden in LHC data, but might be accessible with new search strategies.

New search strategies can only be developed from the communication between theorists and experimentalists. One remarkable example are the new means to search for hypothetical new long-lived particles, where discussions between experimentalists and theorists led to the development of new triggers, new external detectors, and influenced the planning of future experiments. To stimulate similar discussions, a testing ground is needed, where physicists can develop new strategies with a quick turnaround. Open data might constitute such a testing ground, as it provides a platform for knowledge and data exchange that has the potential to unleash the discovery potential of the LHC data, where experiments can be given pointers of what corners of the phase-space need particular scrutiny.

CERN has committed to an open data policy in support of open science, promising to “make scientific research more accessible to the community”. Open data is a relatively new experience in the field of particle physics. As open data policies are not intended to compete with experimental efforts, it is essential to establish in this context a well-defined framework, within which knowledge and data are exchanged. In this community effort the experimentalists should remain the competent authority to set down the data access guidelines, while the role of theorists would be to provide directions with new insights and ideas. The theorist’s insight thus may keep the most crucial questions in front of the eyes of everyone involved in the effort. Of course, such questions are numerous and multi-dimensional.

In the workshop “Unveiling Hidden Physics beyond the Standard Model at the LHC”¹, the existing motivation and insight into possible manifestations of NP were discussed. This included a state-of-the-art overview of significant anomalous experimental results, lessons from theories including model classes that are challenging to detect at the LHC, computing methods, and the CERN open data. The central aim of the workshop was to highlight the fact that the usage of open LHC data allows the community to test a much larger range of NP than ever before.

The present document includes a review of most significant anomalies in particle physics and a review of model classes that can be hidden in LHC data. The anomalies are grouped appropriately and potential explanations are summarised, where possible interconnections are explored. The review of hidden model classes is non-exhaustive and constitutes an example of the opportunities that model building continues to provide to the physics programme of the LHC.

Workshop discussion

Contributions: Nishita Desai, Biswarup Mukhopadhyaya

¹https://indico.tlabs.ac.za/event/100/
The discussions at the workshop were held after individual talks, in dedicated discussion sessions, and a plenary discussion. There was also some exchange on a dedicated mattermost channel.

The main points were collected and brought into this paper in many different places. Below is a list of few illustrative questions that emerged during the discussions in the opening session of the workshop, which may serve as rudders for future discussion.

- Hundreds of models have been explored in our quest for new physics, and we have little guidelines yet as to how many additional directions are worthy of exploration. Till such guidelines emerge in firmer outlines, it may be advisable to carry on prediction and analysis of new physics signatures at the LHC in model-independent ways as far as possible, so that we do not miss interesting possibilities due to any bias. At the same time, a virtue of model-based studies also looms up. The pros and cons of the theoretical characteristics of certain scenarios often become clear when their consequences are pitted against experimental data, and enrich us with wisdom that goes beyond the ambit of those specific scenarios. Thus we derive our ‘lessons from theory’ by occasionally resorting to models as well.

- The 125-GeV scalar has been observed, and whether it is ‘the Higgs’ or ‘a Higgs’ is still an open question. On the other hand, the electroweak symmetry-breaking sector has led to a good many questions about limitations of the standard model. When the High Luminosity LHC starts operating, the data coming out of it should be used to glean every speck of additional information about this scalar, and look for effects that may serve to unveil physics beyond the standard model.

- Euclidean continuation is important in understanding global Higgs behaviour. In this context, both time-like and space-like probes at high energies should be complementary. $h \rightarrow ZZ$ would be a good start in this connection, because it interferes destructively with the SM box diagram. If the intervention of new physics makes on-shell $h \rightarrow ZZ$ small, it will enhance sensitivity of off-shell Higgs signal. This is also applicable to di-Higgs production via triple Higgs coupling.

- High-$p_T$ Higgs boson physics is complementary to the off-shell Higgs boson signal. Momentum transfer to the Higgs boson production vertex in such events is space-like. So experiments should pay attention to the relatively small number of events in the high-$p_T$ range, which may accentuate the role of off-shell Higgs boson, and any trace of BSM physics contained there.

- The running effect of $m_t(\mu)$, too, depends on features of the deep Euclidean region related to the top Yukawa coupling, although it is difficult to observe the effects at the LHC, because of the uncertainty in the measurement of top quark mass. This, however, can be taken up as a challenge at the high-luminosity runs, since (a) the top Yukawa coupling is related to the issue of naturalness, and (b) the precise relationship of the top pole mass with the running mass at some energy can reveal information of BSM contribution to the relevant renormalisation group equations.

- Since theoretical scenarios may exist just a little beyond the on-shell reach of the LHC, it is important to think not only about off-shell effects but also in terms of higher-dimensional effective operators which, after all, may turn out to be our major
handles. The scale of such operators gets reflected in high-$p_T$ events which should therefore be probed with great emphasis during the high-luminosity runs. High $p_T$ events can also serve as probes of top physics in the deep Euclidean region.

- Observations on dark matter indicate potent new physics options. This includes, as major components, theoretical scenarios with symmetries such as $Z_2$, or those containing long-lived particles (either DM candidates themselves or others belonging to the dark sector). Since dark matter is a concrete reality, such scenarios should constitute high-priority search areas.

- On a more theoretical note, the issue of naturalness of the electroweak scale is yet unresolved, especially when no evidence of supersymmetry is found yet, in regions of the parameter space with ‘sensible’ values of naturalness criteria. It is high time to investigate whether the high-luminosity data contain any clue on this in the non-strongly interacting sector.

2 Anomalies

Section editors: Andreas Crivellin, Oliver Fischer and Bruce Mellado
Contributions: Emanuele Bagnaschi, Geoffrey Beck, Benedetta Belfatto, Zarab Berezhiani, Monika Blanke, Bernat Capdevila, Bhupal Dev, Oliver Fischer, Martin Hoferichter, Matthew Kirk, Farvah Mahmoudi, Claudio Andrea Manzari, David Marzocca, Bruce Mellado, Antonio Pich and Luc Schnell

This section summarises a cohort of anomalies in the data that currently do not appear to be explained by the SM. See Ref. [10] for an up-to-date summary of all existing anomalies. The section is structured as follows: Sec. 2.1 gives an overview of flavour anomalies; Sec 2.2 details the multi-lepton anomalies at the LHC; Sec. 2.3 describes the Higgs-like excess at 96 GeV; Sec. 2.4 discusses anomalies in the neutrino sector; finally, Secs. 2.5, 2.6 and 2.7 touch upon anomalies in astrophysics, cosmology and in Ultra-High energy cosmic rays, respectively.

2.1 Flavour Anomalies

Intriguing indirect hints for BSM physics have been accumulated in flavour observables within recent years: Semi-leptonic bottom quark decays ($b \rightarrow s\ell^+\ell^-$); Tauonic $B$ meson decays ($b \rightarrow c\tau\nu$); The anomalous magnetic moment of the muon ($\alpha_\mu$); The Cabibbo angle anomaly (CAA); Non-resonant di-electrons ($q\bar{q} \rightarrow e^+e^-$); The difference of the forward-backward asymmetry in $B \rightarrow D^{\ast}\mu\nu$ vs $B \rightarrow D^{\ast}\epsilon\nu$ ($\Delta A_{FB}$); Low-energy lepton flavour universality violation (LFUV) in the charged current, including leptonic tau decays ($\tau \rightarrow \mu\nu\nu$). Interestingly, all these observables admit an interpretation in terms of LFUV, i.e., NP that distinguishes between muon, electrons and tau leptons. While some of the anomalies are by construction measures of LFUV, also the other observables can be interpreted in this context (see Fig. 1). This unified view suggests a common origin of the anomalies in terms of BSM physics, which reinforces the case for LFUV with important theoretical and experimental implications. In the following, we will review these flavour anomalies and related processes.


Figure 1: Summary of the experimental hints for LFUV beyond the SM.

### 2.1.1 $b \to s \ell^+ \ell^-$

Consistent hints of New Physics have been observed in semileptonic $B$-meson decays involving $b \to s \ell^+ \ell^-$ transitions. Different experimental collaborations at the LHC, with LHCb playing the leading role, and the Belle experiment have reported deviations from SM expectations at the 2–3$\sigma$ level in several channels mediated by these transitions. The most relevant discrepancies include observables characterising the $B^0 \to K^{\ast 0} \mu^+ \mu^-$ \cite{11} and $B^+ \to K^{\ast +} \mu^+ \mu^-$ \cite{12} decay distributions, in particular the so-called $P'_5$ observable in two adjacent anomalous bins in the low-$q^2$ region,

\begin{align}
P'_5(B^0 \to K^{\ast 0} \mu^+ \mu^-)_{\text{LHCb}}^{[4.0,6.0]} & = -0.439 \pm 0.111 \pm 0.036 \quad (2.5\sigma), \quad (1) \\
P'_5(B^0 \to K^{\ast 0} \mu^+ \mu^-)_{\text{LHCb}}^{[6.0,8.0]} & = -0.583 \pm 0.090 \pm 0.030 \quad (2.9\sigma), \quad (2)
\end{align}

the $R_K$ \cite{13} and $R_{K^*}$ \cite{14} ratios, defined as \cite{15}

\begin{equation}
R_K^{\ast(*)} = \frac{\text{Br}(B^{+(0*)} \to K^{+(0*)} \mu^+ \mu^-)}{\text{Br}(B^{+(0*)} \to K^{+(0*)} e^+ e^-)},
\end{equation}

which measure LFUV in the $B \to K \ell^+ \ell^-$ and $B \to K^* \ell^+ \ell^-$ modes,

\begin{align}
R_{K_{\text{LHCb}}}^{[1.1,6]} & = 0.846^{+0.042+0.013}_{-0.039-0.012} \quad (3.1\sigma), \quad (4) \\
R_{K_{\text{LHCb}}}^{[0.045,1.1]} & = 0.66^{+0.11}_{-0.07} \pm 0.03 \quad (2.2\sigma), \quad (5) \\
R_{K_{\text{LHCb}}}^{[1.1,6]} & = 0.69^{+0.11}_{-0.07} \pm 0.05 \quad (2.6\sigma), \quad (6)
\end{align}
and the $B_s \to \phi \mu^+ \mu^-$ branching ratio \cite{16,17},
\[
\left\langle \frac{d\text{Br}(B_s \to \phi \mu^+ \mu^-)}{dq^2} \right\rangle_{\text{LHCb}}^{[1.16]} = (2.88 \pm 0.15 \pm 0.05 \pm 0.14) \times 10^{-8} \quad (3.6\sigma).
\] (7)

In addition, the branching ratio of the leptonic decay $B_s \to \mu^+ \mu^-$ shows some tension with respect to its SM prediction:
\[
\text{Br}(B_s \to \mu^+ \mu^-) = 2.85^{+0.34}_{-0.31} \times 10^{-9} \quad (2.15\sigma),
\] (8)

where the quoted value corresponds to the average of the latest LHCb measurement \cite{18,19} with the results from CMS \cite{20} and ATLAS \cite{21} (see Refs. \cite{22,25} for further details).

These tensions, together with the discrepancies observed in $b \to c\ell\nu$ modes, see Sec. \ref{sec:2.1.2} are commonly referred to in the literature as “B anomalies”. The current situation is exceptional since all deviations in $b \to s\ell^+\ell^-$ channels are consistent with a deficit in muonic modes and form coherent patterns in global fits, some of which are preferred over the SM with a very high significance. State-of-the-art global analyses of $b \to s\ell^+\ell^-$ data can be found in Refs. \cite{22,29}. These global fits differ in the treatment of theoretical uncertainties, with the most important differences being the choice of form factors \cite{30,32}, the parametrisation used to include factorisable and non-factorisable hadronic uncertainties \cite{33,36} and the approach used in the statistical analysis itself \cite{37,42}.

However, all the abovementioned global analyses share the same model-independent framework based on the effective Hamiltonian of the Weak Effective Theory (WET), in which heavy degrees of freedom with characteristic scales above the $W$ boson mass – including any potential heavy new particles – are integrated out in short-distance Wilson coefficients $C_i$,
\[
\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i \mathcal{O}_i.
\] (9)

Even though NP could generate further effective operators with structures not present in the SM, because of the processes included in the global fits, most analyses focus their attention to the electromagnetic and semileptonic operators (including their chirally-flipped counterparts) \cite{4}
\[
\mathcal{O}_7 = \frac{e^2}{16\pi^2} m_b (\bar{s} \sigma_{\mu\nu} P_L b) F^{\mu\nu}, \quad \mathcal{O}_7' = \frac{e^2}{16\pi^2} m_b (\bar{s} \sigma_{\mu\nu} P_L b) F^{\mu\nu},
\]
\[
\mathcal{O}_{9\ell} = \frac{e^2}{16\pi^2} (\bar{s} \gamma_{\mu} P_L b) (\bar{\ell} \gamma^{\mu}\ell), \quad \mathcal{O}_{9\ell}' = \frac{e^2}{16\pi^2} (\bar{s} \gamma_{\mu} P_R b) (\bar{\ell} \gamma^{\mu}\ell),
\]
\[
\mathcal{O}_{10\ell} = \frac{e^2}{16\pi^2} (\bar{s} \gamma_{\mu} P_L b)(\bar{\ell} \gamma^{\mu}\gamma_5\ell), \quad \mathcal{O}_{10\ell}' = \frac{e^2}{16\pi^2} (\bar{s} \gamma_{\mu} P_R b)(\bar{\ell} \gamma^{\mu}\gamma_5\ell),
\] (10)

where $\ell = \mu, e$, $P_{L,R} = (1 \mp \gamma_5)/2$ and $m_b = m_b(\mu_b)$ is the running $b$-quark mass in the $\overline{\text{MS}}$ scheme at the characteristic scale of the process $\mu_b \sim 4.8\text{GeV}$. The SM values of the relevant Wilson coefficients are $C_{7,9\ell,10\ell}(\mu_b) =$ \text{−0.29, 4.07, −4.31} and $C_{7,9\ell,10\ell}(\mu_b) \sim 0$, for both $\ell = \mu$ and $\ell = e$. In this language, NP effects are parametrised as shifts from their SM values $C_{it} = C_{it}^{\text{SM}} + C_{it}^{\text{NP}}$.

Since no deviations have been observed in channels with electrons in the final state, NP contributions to the electronic Wilson coefficients are assumed to be negligible. Then, the

\footnote{Global fits with a more general set of 20 operators can be found in Refs. 24,43,44.}
most updated global fits to the muonic coefficients reveal the vectorial $C_{9\mu}^{\text{NP}}$ and left-handed $C_{9\mu}^{\text{NP}} = -C_{10\mu}^{\text{NP}}$ structures as the favourite NP scenarios according to current $b \to s\ell^+\ell^-$ data \cite{22,27}. Additionally, restricted fits to LFUV observables and $B_s \to \mu^+\mu^-$ show a NP signal in $C_{10\mu}^{\text{NP}}$ with high significance \cite{23,25}. Also, scenarios including right-handed couplings (RHC) have been recently found to provide very competitive descriptions of the data \cite{22,45}. The statistical significance of these scenarios, as measured by the so-called Pull$_{\text{SM}}$, ranges from roughly to well-above 5$\sigma$ depending on the particular details of each analysis.

Notice that the NP scenarios discussed so far are all based on the underlying assumption of LFUV NP, where the NP is entirely attached to the muons. However, some analyses have also started exploring scenarios with lepton flavour universal (LFU) NP effects in addition to LFUV contributions to muons only \cite{26,45,47}. In order to account for these contributions, one possible parametrisation reads,

$$C_{i\epsilon}^{\text{NP}} = C_i^U, \quad C_{i\mu}^{\text{NP}} = C_i^U + C_i^V,$$

with $i = 9(\mu), \ 10(\mu)$. The basis redefinition in Eq. (11) provides a new description of the data with a concrete NP structure, namely, that $b \to s\ell^+\ell^-$ transitions get a common LFU NP contribution for all charged leptons (electrons, muons and tau leptons), opening new directions and extending the possible interpretations of the global fits. Interestingly, when allowing for LFU NP, the scenario ($C_{9\mu} = -C_{10\mu}, C_9^U$) with an $SU(2)_L$ LFUV structure emerges as an acceptable NP solution \cite{22,26}. Also, scenarios with $C_{10(\mu)}^U$, like ($C_{9\mu}, C_{10}^U$) and ($C_{9\mu}, C_{10}^U$), get selected with very high significance \cite{22,45}.

It is also important to discuss the implications of the global $b \to s\ell^+\ell^-$ fits on popular NP models. Now we briefly review those that are able to generate the preferred structures suggested by the global fits.

**C$^{\text{NP}}_{9\mu}$**: $Z'$ models with vectorial couplings to leptons preferably yield $C_{9\mu}^{\text{NP}}$-like solutions in order to avoid gauge anomalies. In this context, $L_{\mu} - L_{\tau}$ models \cite{48,52} are popular since they do not generate effects in electron channels. Fits including $R_{K^*}$ are also very favourable to models predicting $C_{9\mu}^{\text{NP}} = -3C_{9\mu}^{\text{NP}} \cite{53}$. Concerning leptoquarks (LQs), a $C_{9\mu}^{\text{NP}}$ solution can only be generated by adding two scalar (an $SU(2)_L$ triplet and an $SU(2)_L$ doublet with $Y = 7/6$) or two vector representations (an $SU(2)_L$ singlet with $Y = 2/3$ and an $SU(2)_L$ doublet with $Y = 5/6$).

**C$^{\text{NP}}_{9\mu}$**: This pattern can be achieved in $Z'$ models with loop-induced couplings \cite{54} or with heavy vector-like fermions \cite{55,56}. Regarding LQ models, here a single representation (the scalar $SU(2)_L$ triplet or the vector $SU(2)_L$ singlet with $Y = 2/3$) can generate a $C_{9\mu}^{\text{NP}} = -C_{10\mu}^{\text{NP}}$ solution \cite{57,63}. This pattern can also be obtained in models with loop contributions from three heavy new scalars and fermions \cite{64,68} and in composite Higgs models \cite{69}.

**RHC**: with a value of $R_K$ closer to one, scenarios with right-handed currents, namely $C_{9\mu}^{\text{NP}} = -C_{9\mu}$, ($C_{9\mu}, C_{9\mu}^U$) and ($C_{9\mu}, C_{10\mu}^U$), seem to emerge. The first two scenarios are naturally generated in $Z'$ models with certain assumptions on its couplings to right-handed and left-handed quarks, as it was shown in Ref. \cite{48} with the context of a gauged $L_{\mu} - L_{\tau}$ symmetry with vector-like quarks. One could also obtain $C_{9\mu}^{\text{NP}} = -C_{9\mu}$ by adding a third Higgs doublet to the model of Ref. \cite{51} with opposite $U(1)$ charge. On the other hand, generating the aforementioned contribution in LQ models requires one to add four scalar representations or three vector ones.

($C_{9\mu} = -C_{10\mu}^U, C_9^U$): this scenario can be realised via off-shell photon penguins in a LQ model explaining also $b \to c\tau\nu$ data \cite{70} (see Sec. 2.1.8). Remarkably, as we will discuss
below, a NP contribution with this structure allows for a model-independent combined explanation of $b \to s\ell^+\ell^-$ and $b \to c\tau\nu$ data with very high statistical significance \cite{22,45,70}.

$C_{10}^U$: NP solutions with $C_{10}^U$ (see scenarios 9–13 from Refs. \cite{22,45}) arise naturally in models with modified $Z$ couplings. In this case, $C_{9(7)}^{U}$ contributions are also generated but to a good approximation can be neglected. The $(C_{9}^{U} = -C_{10\mu}, C_{10}^{U})$ pattern also occurs in Two-Higgs-Doublet models \cite{71}. For scenarios $(C_{9\mu}, C_{10}^{U})$ and $(C_{9\mu}, C_{10}^{U})$, one can also invoke models with vector-like quarks, where modified $Z$ couplings are even induced at tree-level. The LFU effect in $C_{10}^{U}$ can be accompanied by a $C_{9}^{U}$ effect from $Z'$ exchanges \cite{72}. Vector-like quarks with the quantum numbers of right-handed down quarks (left-handed third-generation leptons explains $R_{D^{(*)}}$ expressed in terms of gauge-invariant dimension-6 operators \cite{79,80}. The operator involving third-generation leptons explains $R_{D^{(*)}}$ and the one involving the second generation gives a LFUV effect in $b \to s\ell^+\ell^-$ anomalies and the deviations in $b \to c\tau\nu$ transitions \cite{78}. This connection arises in the SMEFT scenario where $C^{(1)} = C^{(3)}$ expressed in terms of gauge-invariant dimension-6 operators \cite{79,80}. The operator involving gives a LFUV effect in $b \to s\ell^+\ell^-$ anomalies and the deviations in $b \to c\tau\nu$ transitions \cite{78}. The constraint from $b \to c\tau\nu$ and $SU(2)_L$ invariance leads to large contributions enhancing $b \to s\ell^+\ell^-$ processes \cite{80} whereas the mixing into $O_{9\mu}$ generates $C_{9}^{U}$ at $\mu = m_t$ \cite{70}. Therefore, this NP structure correlates $C_{9}^{U}$ and $R_{D^{(*)}}$ in the following way \cite{70,80}:

$$C_{9}^{U} \approx 7.5 \left(1 - \sqrt{\frac{R_{D^{(*)}}}{R_{D^{(*)}}^{SM}}} \left(1 + \frac{\log(A^2/(1\text{TeV}^2))}{10.5}\right)\right), \quad (12)$$

where $\Lambda$ is the typical scale of NP involved. In Fig. 2 we show the global fit of the pattern $(C_{9\mu}^{U} = -C_{10\mu}, C_{9}^{U})$ without and with the additional input on $R_{D^{(*)}}$ from Ref. \cite{78}, taking the scale $\Lambda = 2$ TeV. This connection between neutral and charged anomalies is remarkable as it offers a NP solution that is able to accommodate both sets of data simultaneously, and hence one finds a very high Pull$_{SM}$ of 8.1$\sigma$ for the combined fit \cite{22}.

\subsection*{2.1.2 Tauonic $B$-meson decays}

In addition to the neutral-current $b \to s\ell^+\ell^-$ transitions discussed above, also charged-current $b \to c\tau\nu$ data exhibit tensions with the SM predictions. Of particular interest are the lepton flavour universality (LFU) ratios

$$R(D^{(*)}) = \frac{BR(B \to D^{(*)}\ell\nu)}{BR(B \to D^{(*)}\ell\nu)} \quad (\ell = e, \mu), \quad (13)$$

for which measurements from BaBar \cite{81,82}, Belle \cite{83,86} and LHCb \cite{87,89} exist, see Fig. 3.

The latest HFLAV average combining these data \cite{90}:

$$R(D) = 0.340 \pm 0.027 \pm 0.013,$$

$$R(D^{*}) = 0.295 \pm 0.011 \pm 0.008, \quad (14)$$

\begin{align*}
\end{align*}
Figure 2: Preferred regions at the 1, 2 and 3 σ level (green) in the $(C_V^{9\mu} = -C_V^{10\mu}, C_U^{9})$ plane from $b \to s\ell^+\ell^-$ data. The red contour lines show the corresponding regions once $R(D^{(*)})$ is included in the fit (for $\Lambda = 2$ TeV). The horizontal blue (vertical yellow) band is consistent with $R(D^{(*)})$ ($R_K$) at the 2 σ level and the contour lines show the predicted values for these ratios.

deviates by 3.1σ from the SM prediction. Furthermore, the data for the analogous ratio $R(J/\psi)$ also seem to hint at an enhancement relative to the SM [91].

NP contributions to $b \to c\tau\nu$ transitions can be parametrised by the Wilson coefficients $C_i$ in the effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = 2\sqrt{2} G_F V_{cb} [ (1 + C_L^V) O_L^V + C_S^L O_S^L + C_S^R O_S^R + C_T O_T ] ,$$

assuming the absence of light right-handed neutrinos. Here $O_L^V$ is the left-handed current-current operator present already in the SM, $O_S^L$ and $O_S^R$ are the left- and right-handed scalar operators, and $O_T$ is the tensor operator, as defined, e.g., in Ref. [92].

Global fits to the data, including polarisation observables in $B \to D^*\tau\nu$, have been performed in Refs. [47, 93–95]. From the results of these fits, several simplified NP models can be identified as potential candidates for an explanation of the $b \to c\tau\nu$ anomalies. Due to the rather large size of the required NP contribution with respect to the SM, in all cases new particles contribute to $b \to c\tau\nu$ at the tree level, and for the sake of simplicity we restrict our attention to models with a single new state:

**Charged W' bosons:** A good fit to the available $b \to c\tau\nu$ data is obtained by a shift $C_L^V \neq 0$ of the SM $(V - A) \otimes (V - A)$ contribution, which could originate from a heavy charged $W'$ gauge boson coupling to left-handed quarks and leptons [96, 97]. This model, however, is challenged by LHC high-$p_T$ di-τ data [98] as well as by precision measurements of Z-pole observables [99].

**Charged Higgs boson $H^\pm$:** This scenario [100, 103], leading to non-zero $C_L^{L,R}$, currently provides the best fit to the low-energy $b \to c\tau\nu$ data, as – in contrast to the other simplified models – it allows one to accommodate the measured $D^*$ polarisation, $F_L(D^*)$ [104].
at the 1σ level. However, this solution is in tension with the LHC mono-τ data \cite{105}, and it induces a large branching ratio \( \text{BR}(B_c \rightarrow τν) > 50\% \). While no direct experimental bound on the latter exists, upper limits of 30% \cite{106} and even 10% \cite{107} have been estimated in the literature. On the other hand, a critical reassessment reached the conclusion that values even as large as 60% cannot be excluded at present \cite{92, 95}. A recent update on the SM prediction of the \( B_c \) lifetime supports the latter reasoning \cite{108}.

**Scalar leptoquarks:** The scalar \( SU(2)_L \)-singlet leptoquark \( S_1 \)\cite{109–111}, giving rise to the scenario \( C^L_V, C^L_S = -4C^T \neq 0 \), offers a good fit to the \( b \rightarrow cτν \) data, predicts only modest contributions to the decay \( B_c \rightarrow τν \), and passes the mono-τ test. The scalar \( SU(2)_L \)-doublet leptoquark \( S_2 \), inducing \( C^L_S = 4C^T \), on the other hand, can be brought in agreement with the \( b \rightarrow cτν \) data only in the presence of complex, i.e., CP-violating couplings \cite{112}. The latter scenario predicts a significant contribution to \( \text{BR}(B_c \rightarrow τν) \sim 20\% \), and its best-fit point is on the verge of being tested by the mono-τ searches. There are also stringent LHC constraints on these LQs from their pair-production and t-channel mediated dilepton processes \cite{113,114}.

**Vector leptoquark:** Last but not least also an \( SU(2)_L \)-singlet vector leptoquark \( U_1 \)\cite{58, 60, 61, 115–119} provides a good fit to the \( b \rightarrow cτν \) data, both with only left-handed couplings \( (C^L_V \neq 0) \) and in the presence of an additional small right-handed \( bτ \) coupling \( (C^R_V, C^R_S \neq 0) \). As in the case of the scalar \( SU(2)_L \)-singlet leptoquark, also here the contributions to \( B_c \rightarrow τν \) are small and the model evades the current LHC mono-τ searches. One of the most stringent constraints on models with an \( SU(2)_L \)-singlet vector leptoquark stems instead from LHC searches for colour-octet resonances, which are often introduced together with the leptoquark in UV-complete models \cite{120–122}.

In addition to these simplified models parametrised by the effective interactions in Eq. \cite{15}, models with light right-handed neutrinos have been examined in the literature \cite{123, 127}. While it is possible to accommodate the low-energy \( b \rightarrow cτν \) data in this case, a very large NP contribution is required due to the absence of interference with the SM contribution. Consequently the constraints from direct LHC searches, particularly mono-τ, tend to be even more severe.

To further disentangle the NP structure at work, a major role will be played by the
measurement of differential and angular observables \cite{125,133}, such as the $D^*$ and $\tau$ polarisations $F_L(D^*)$ and $P_\tau(D^{(*)})$, whose correlations turn out to discriminate well between the different scenarios. To fully exploit their model-discriminating potential, both precise measurements and a better theoretical understanding of the underlying form factors are necessary. A measurement of the baryonic LFU ratio

$$R(\Lambda_c) = \frac{BR(\Lambda_b \to \Lambda_c \tau \nu)}{BR(\Lambda_b \to \Lambda_c \ell \nu)} \quad (\ell = e, \mu),$$

will instead provide an experimental consistency check for the $R(D^{(*)})$ anomaly, thanks to a model-independent sum-rule \cite{92} relating $R(\Lambda_c)$ to $R(D)$ and $R(D^*)$, with the current prediction \cite{95}

$$R(\Lambda_c) = R_{SM}(\Lambda_c)(1.15 \pm 0.04) = 0.38 \pm 0.01 \pm 0.01. \quad (17)$$

In addition to the constraints mentioned above, further tensions may arise in concrete UV completions. For example, in certain models electroweak $SU(2)_L$ symmetry implies large contributions to the decays $B \to K^{(*)}\nu\bar{\nu}$, $B_s \to \tau^+\tau^-$ and $B \to K^{(*)}\tau^+\tau^-$ \cite{61,134}, and significant rates for $\Upsilon \to \tau^+\tau^-$ or $\psi \to \tau^+\tau^-$ are expected \cite{135}. In summary it is fair to say that stringent constraints on all NP scenarios for the $R(D^{(*)})$ anomaly exist, challenging a full resolution of the latter in the context of NP.

2.1.3 Anomalous magnetic moments of charged leptons

Ever since Schwinger’s famous prediction $a_\ell = (g - 2)_\ell/2 = \alpha/(2\pi)$ \cite{136} (and its experimental verification \cite{137}), the anomalous magnetic moments of the electron and muon have been critical precision tests of the SM. For the electron, the current best direct measurement \cite{138}

$$a_{e}^{\text{exp}} = 1159652180.73(28) \times 10^{-12} \quad (18)$$

can be contrasted with its SM prediction once independent input for the fine-structure constant $\alpha$ is specified. With the mass-independent 4-loop QED coefficient known semi-analytically \cite{139}, the dominant uncertainties now arise from the numerical evaluation of the 5-loop coefficient \cite{140} and hadronic corrections \cite{141}, both of which enter at the level of $10^{-14}$ (for the 5-loop QED coefficient there is a $4.8\sigma$ tension between Refs. \cite{140,142} regarding the contribution of diagrams without closed lepton loops). However, the current most precise measurements of $\alpha$ in atom interferometry, using Cs \cite{143} and Rb \cite{144} atoms, respectively, differ by $5.4\sigma$,

$$a_{e}^{\text{SM}}[\text{Cs}] = 1159652181.61(23) \times 10^{-12},$$

$$a_{e}^{\text{SM}}[\text{Rb}] = 1159652180.25(10) \times 10^{-12}, \quad (19)$$

resulting in a difference to Eq. \cite{18} of $-2.5\sigma$ and $+1.6\sigma$.

The world average of the muon $g - 2$ is determined by the Run 1 results from the Fermilab experiment \cite{145,148} and the Brookhaven measurement \cite{149}

$$a_{\mu}^{\text{exp}} = 116592061(41) \times 10^{-11}, \quad (20)$$

with a combined precision of 0.35 ppm. Comparison with the current SM prediction \cite{150}

$$a_{\mu}^{\text{SM}} = 116591810.43(10) \times 10^{-11} \quad (21)$$
then reveals a $4.2\sigma$ tension. Experimental efforts to corroborate or refute this tension are underway at subsequent runs at Fermilab [151] and at J-PARC [152], with a precision goal of 0.14 ppm and 0.45 ppm, respectively, and the J-PARC experiment pioneering a new experimental technique that does not rely on the magic momentum in a storage ring, see Ref. [153] for a more detailed comparison of the two methods. The SM prediction in Eq. (21), currently at 0.37 ppm, represents a coherent theory effort organised in the Muon $g - 2$ Theory Initiative [150], and is mainly based on the underlying work from Refs. [140, 141, 154–171].

The uncertainty is completely dominated by hadronic contributions, with hadronic vacuum polarisation (HVP) and hadronic light-by-light scattering (HLbL) at 0.34 ppm and 0.15 ppm, respectively. Improvements on both HVP and HLbL will continue over the next years, including new $e^+e^- \to \text{hadrons}$ data, lattice-QCD calculations at a similar level of precision, and direct input on space-like HVP from the proposed MUonE experiment [172,173]. Some recent developments include the first lattice calculation of HVP reporting subpercent precision [174], with subsequent work exploring the consequences of the emerging $2.1\sigma$ tension with the data-driven determination [175–179], new $e^+e^- \to \pi^+\pi^-$ data from SND [180], improved radiative corrections [181], a lattice-QCD calculation of HLbL at a similar level of precision as the phenomenological evaluation [182], and work aimed at refining the subleading contributions to HLbL [183,188].

New Physics explanations The absolute value of the difference between measurement and theory prediction exceeds the size of the EW contribution of the SM. Therefore, some form of enhancement mechanism is required to explain the current $4.2\sigma$ tension with BSM physics, and well-motivated scenario do exist. One possibility is that NP involves heavy particles at or above the EW scale, with an enhanced chirality flip originating from an interaction between new particles with the SM Higgs boson, with a coupling strength that is larger than the muon Yukawa. Depending on the model, this type of chiral enhancement allows for viable solutions for new particles with masses up to tens of TeV. Such an enhancement can be achieved in models with new scalars and fermions [189] with the MSSM being a specific example. Alternatively, the anomaly can be explained by new, light (or very light) weakly coupled states, such an axion-like particles (ALPs) or a dark photon $Z_d$. For a more detailed overview of various models in light of the most recent measurement, we refer the reader to Ref. [190].

The Minimal Supersymmetric Standard Model One possible theoretical framework of NP above the electroweak scale is the MSSM. Here, the chiral enhancement is provided by the factor $\tan \beta \equiv v_u/v_d$, where $v_u$ and $v_d$ are the vacuum expectation values of the two Higgs doublets of the model, $H_u$ and $H_d$ (which give mass to up-type and down-type fermions respectively). A large value of $\tan \beta \approx 50$ can be motivated by top–bottom Yukawa coupling unification [191,192], and thus this would provide a natural explanation for a large enhancement factor [193,195].

The leading MSSM contributions arise from one-loop diagrams involving a loop of either a neutralino and smuon ($a_\chi^0\tilde{\mu}$) or a chargino and a sneutrino ($a_\chi^\pm\tilde{\nu}$) (see the left plot of Fig. 4) [196,203].

---

3The two-loop contributions have been calculated in Refs. [204,212] and the resummation of $\tan \beta$ has been achieved [209,213,214]. These results have been implemented in codes such as SuperIso [215,217], micrOMEGAs [215], FeynHiggs [206,207], (SARAH-)Spheno [219] and FlexibleSUSY [220] to the dedicated software such as SUSYFLAVOR [221] and GM2Calc [222].
Figure 4: Left: one-loop contributions to \((g - 2)_\mu\) from MSSM particles; the external photon line can be attached to any charged particle. Right: example of a two loop Barr-Zee contribution involving a charged sfermion.

The relevant phenomenological question is how to account for relatively light SUSY particles (and explain \((g - 2)_\mu\)) while meeting constraints from the LHC searches, DM phenomenology, and other observables. Incorporating LHC Run 2 limits [190, 223–257], with a varying degree of sophistication and diverse focuses (model building, collider, DM etc.). Furthermore, several global studies, trying to incorporate and correlate LHC Run 2 bounds with limits coming from different sectors, have been performed as well, both in the context of scenarios with universal and minimal SUSY breaking mechanisms [258–261], and for more phenomenological oriented models [262, 263]. In the former case, LHC constraints push the mass of the SUSY states to the TeV scale, such that the \(\tan\beta\) enhancement is insufficient to provide a sufficiently large contribution to \(a_\mu\). For the latter scenarios, if sufficiently freedom is allowed, SUSY contributions can still account for the observed discrepancy [263].

The possibility of explaining the anomaly in non-minimal supersymmetric extensions of the SM, in light of recent LHC constraints, has also been extensively studied in the literature, cf., e.g., Refs. [265–283].

**Leptoquarks** Another possible explanation, which also provides a viable solution to the hints for lepton flavour universality violation in semi-leptonic \(B\) decays, is given by leptoquarks. Indeed, two scalar LQ representations can provide a chiral enhancement factor of \(m_t/m_\mu \approx 1600\) [284–289]. This allows for a TeV scale explanation with perturbative couplings that is not in conflict with direct LHC searches. It is furthermore very predictive as it involves, besides the LQ mass only two couplings, whose product is fixed by requiring that \(g - 2\) be explained. Therefore, correlated effects in \(h \to \mu^+\mu^-\) [289], and to a lesser extent in \(Z \to \mu^+\mu^-\) [288, 290], arise which can be used to (indirectly) distinguish the two LQ representations at future colliders. In fact, correlations with \(h \to \mu^+\mu^-\) and \(Z \to \mu^+\mu^-\) become of interest for a wide range of chirally enhanced scenarios, see Ref. [291].

**Interplay with the lepton EDMs** A consequence of explanations via chiral enhancement concerns the phase of the Wilson coefficient of the dipole operator, which emerges as a free parameter. In particular, such scenarios in general violate the scaling expected from MFV [189, 292], which may result in a large muon EDM well above MFV projections derived
Figure 5: Left: plot from Ref. [249] showing the neutralino-smuon mass range still allowed after considering LHC constraints (shaded blue areas) for a MSSM $(g-2)_\mu$ solution at 1σ ($2\sigma$) in orange (yellow), with a bino-like LSP, degenerate smuons and $\tan\beta = 30$. Right: plot from Ref. [264] showing in green (light blue) the preferred region by the current $(g-2)_\mu$ $(\pm (g-2)_\mu)$ measurement in the plane of the kinetic mixing parameter $\epsilon$ and of the mass the dark $Z_d$ mediator; the dotted and dashed lines corresponds to the limits from the $Q_{\text{WEAK}}$ and APV experiments respectively, while the solid lines represent their combinations; the $\delta$ parameter is related to the $Z-Z_d$ mass mixing.

from the limit on the electron EDM [293]. A large part of the parameter space in which the phase is $O(1)$, as well possible from an EFT perspective [189, 294, 295], could be covered by a proposed dedicated muon EDM experiment at PSI [296]. In fact, the corresponding non-MFV flavour structure is not at odds with naturalness arguments, since, in the limit of vanishing neutrino masses, lepton flavour is conserved, and thus it is possible to completely disentangle the muon from the electron EDM via a symmetry, meaning that no fine tuning is necessary. This could for example be achieved via an $L_{\mu} - L_{\tau}$ symmetry [297, 299], which can naturally give rise to the observed neutrino mixing matrix [300, 302], and, even after its breaking, protects the electron EDM and $g-2$ from BSM contributions [50].

The other solutions via heavy NP In addition, there exists a plethora of alternative BSM explanations of the muon $g-2$, including composite or extra-dimensional models [303, 305] or models with vector-like leptons [50, 68, 189, 291, 292, 306, 309], including in addition a second Higgs doublet [310, 312].

Also a pure 2HDM can provide a solution. This is either possible via Barr-Zee diagrams in the 2HDM-X [103, 190, 310, 313, 323], where the external photon couples to an internal charged fermion loop, which then couples to the muon line via one of the new Higgs bosons and a photon (as in Fig. 4 right, but with the sfermion replaced by a fermion) or including vector-like leptons. Alternatively, a lepton flavour violating $\tau\mu$ couplings can provide a $m_\tau/m_\mu$ enhancement [71, 324, 326], which is however strongly constrained from $h \to \tau\mu$ searches.
Weakly coupled models with new light states Another possibility to explain the anomaly is to have weakly coupled new states (sometimes called generally Feebly Interacting Particles – FIPs) that however can provide a significant contribution to their small mass, a case for which a rich literature is available. Below we mention only a small selection of these studies. In the case of a spin-1 explanation, an example is given by dark Z models. In Fig. 5 we show a plot taken from Ref. 264 where a dark Zd mediator model was studied. In the plot, the interplay with other current and future low-energy experiments is also shown. Concerning the possibility of an axion-like explanation, a recent paper 332 points out that this prospect poses some problems due to the fact that it seems to require an axion decay constant of $\mathcal{O}(10)$ GeV, which in turn implies the existence of new states at low scales, creating phenomenological issues which are not easily addressable. Another interesting possibility is given by ALP-portal explanations, where the ALP assumes the role of mediator with a dark sector 331. The possibility of explaining simultaneously $(g - 2)_\mu$ and the flavour anomalies using FIPs has been recently presented in Ref. 333.

2.1.4 Cabibbo Angle Anomaly

One of the fundamental predictions of the SM is the unitarity of the CKM matrix. In particular, for the first row of CKM elements it implies the condition

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1,$$

which in practice reduces to the Cabibbo universality ($|V_{ud}| \approx \cos \theta_{12}$, $|V_{us}| \approx \sin \theta_{12}$), since the last entry is negligibly small: $|V_{ub}|^2 < 2 \times 10^{-5}$ 334. At present, with the improved control of theoretical uncertainties in the determinations of $|V_{us}|$ and $|V_{ub}|$, anomalies are emerging that could be a signal of NP at the TeV scale 335–340. The present situation is shown in Fig. 6 and can be summarised as

$$A : |V_{us}| = 0.22326(55), \quad B : |V_{us}/V_{ud}| = 0.23130(49), \quad C : |V_{ud}| = 0.97355(27).$$

The first two results A and B are extracted from data on kaon semileptonic $K_{3\ell}$ and leptonic $K_{\mu\ell}$ decays 334, respectively, using the most accurate lattice QCD calculations for the vector form factor $f_+(0)$ and for the decay constants ratio $f_K/f_\pi$ 341. The precision of the third result C crucially depends on the knowledge of radiative corrections to be applied in $\beta$ decays 342–348. Using the value of the Fermi constant $G_F = 1.1663787(6) \times 10^{-5}$ GeV$^{-2}$ from muon decay 349, the value of $|V_{ud}|$ is then obtained from the latest update of $\mathcal{F}t$ values in superallowed $0^+\to 0^+$ nuclear transitions, $\mathcal{F}t = 3072.24(1.85)$ s 350, but affected by additional nuclear corrections 351. The extracted value of $|V_{ud}|$ is consistent with a determination via neutron decay (included in C above), based on the average of the neutron lifetimes measured by the eight latest experiments using the neutron trap method, $\tau_n = 879.4(6)$ s, and employing the latest experimental average $g_A = 1.27625(50)$ for the axial coupling 351. Even when using the currently most precise measurements, $\tau_n = 877.75(0.28)^{+0.22}_{-0.16}\times 357$ and $g_A = 1.27641(45)(33)$ 358, the determination from neutron decay is not yet competitive with superallowed $\beta$ decays, but will provide a powerful independent determination in the future. In addition, note that there is also a deficit in the first-column CKM unitarity relation

$$|V_{ud}|^2 + |V_{cd}|^2 + |V_{td}|^2 = 0.9970(18),$$

Footnote:

4Let us note, however, that the neutron lifetime measured in beam experiments, $\tau_n = 888.0(2.0)$ s, is 4$\sigma$ away from the trap value, and it is also incompatible with superallowed $0^+\to 0^+$ transitions 352. This discrepancy, barring the possibility of some unfixed systematics, could be another indication for NP 353–356.
Figure 6: Updated plot of Ref. [335] for the data in Eq. (23) in $V_{us}$–$V_{ud}$ plane. The 1σ, 2σ and 3σ contours (green circles) of the fit are in tension with CKM unitarity (black solid curve). The projections on $|V_{us}|$ axis show the values $|V_{us}|_B$ and $|V_{us}|_C$ obtained from the unitarity condition.

less significant than the tension in the first row, but strengthening the possibility of NP related to the determination of $V_{ud}$.

A fit of the data in Eq. (23) shows the deviation from unitarity at the 3σ level (see Fig. 6). Alternatively, by employing unitarity, these data can be translated into three different results for the Cabibbo angle: $|V_{us}|_A = 0.22326(55)$, $|V_{us}|_B = 0.22535(45)$ and $|V_{us}|_C = 0.2284(11)$, which are in obvious tension with each other. It has also been shown that the discrepancy between the $K_{e3}$ and $K_{μ2}$ results $|V_{us}|_A$ and $|V_{us}|_B$ is unlikely to be due to radiative corrections [359, 360].

In fact, the three determinations of the Cabibbo angle do not necessarily correspond to quite the same values in the presence of NP. The amplitudes of the $K_{e3}$ and $K_{μ2}$ decays are proportional to vector $πγμs$ and axial $πγμγ5s$ currents, respectively. On the other hand, superallowed nuclear transitions are only sensitive to vector current $πγμd$, and the Fermi constant, which is fixed by the muon decay width, could also be affected by NP [361]. Therefore, the Cabibbo angle anomaly (CAA) can be a signal of BSM physics, which can also have other phenomenological implications and be related to other existing anomalies. The different possible explanations can be broadly grouped into three categories: modifications of the $W$ quark vertex, modifications to the $W$ lepton vertex, or effects in four-fermion contact interaction operators.

**Modifying the $W$ quark vertex** The $W$ couplings to quarks are modified after EW symmetry breaking by the two operators: $Q^{(3)ij}_{φq}$ and $Q^{ij}_{φud}$ (see Ref. [79] for the definitions of the operators). The latter generates right-handed $W$-quarks couplings and it has been showed that the interplay between $C^{11}_{φud}$ and $C^{12}_{φud}$ can solve the tension in the CAA [339].

[^The different tree-level extension of the SM, which give rise to these effects have been categorised in Ref. [362].]
and bring the determinations of $|V_{us}|$ from $K_{l3}$ and $K_{l2}$ into agreement. $Q_{\phi q}^{(3)ij}$ generates left-handed $W$-quark couplings and the CAA requires $C_{\phi q}^{(3)11} \approx -(9 \text{ TeV})^{-2}$. These operators can be induced via the mixing of SM quarks with vector-like quarks \[335,340,363,365\]. Note however, that because of $SU(2)$ invariance this operator generates also effects in $\Delta F = 2$ processes which would rule out this solution, unless these effects are suppressed by assuming that $Q_{\phi q}^{(3)ij}$ respect a global $U(2)^2$ flavour symmetry.

**Modifying the $W$ lepton vertex** The SMEFT coefficient $C_{\phi f}^{(3)}$ corresponds to modifications of the $Wl\nu$ and $Zl\ell$ leptonic currents after EW symmetry breaking. This is also interesting for another reason – namely that since this coefficient carries flavour indices such that NP could be related to LFUV \[336,337\], which ties in to many of the other anomalies discussed in this white paper. In order to explain the CAA, this coefficient must be approximately $C_{\phi f}^{(3)22} \approx (9 \text{ TeV})^{-2}$ \[366,367\].

There are four different vector-like leptons (VLLs), as well as an $SU(2)_L$ triplet vector boson, that can be responsible for generating this operator, and hence explain the CAA. The phenomenology of VLLs for the CAA has been studied in detail \[366,369\] and the vector triplet idea has been examined in Refs. \[366,370\].

For the VLLs, extra phenomenological consequences arise since the $SU(2)$ singlet operator $Q_{Hl}^{(1)}$ is generated. This operator alters leptonic $Z$ decays, which are strongly constrained by EW precision observables from LEP and the LHC. Such effects allow us to distinguish between the different models. E.g., it has been shown that extending the SM with a single VLL \[366,371\] leads to tensions between the region of parameter space favoured by the CAA and EW observables. Adding multiple new representations, each coupling to a single different SM lepton can provide a better fit to data \[367\]. For the vector triplet boson, a minimal model leads to tensions with EW precision observables \[366\], which can be eased in a less minimal setup \[370\].

**Four-fermion operators** There are several four-fermion operators in the SMEFT that can affect the determination of the Fermi constant or directly alter semi-leptonic decays \[361\]. Starting with four-lepton operators, the severe constraints from the Michel parameter, muonium–anti-muonium oscillations and the upper bounds on LFV processes lead to the conclusion that the only viable solution to the CAA proceeds via a modification of the SM operator $Q_{\ell\ell}^{(2)12}$ with a Wilson coefficient $C_{\ell\ell}^{(2)12} \approx -(8 \text{ TeV})^{-2}$. Simple models generating this contribution via a singly charged scalar have been recently proposed in Refs. \[372,374\]. This option was also proposed in Ref. \[335\] via a generic flavour-changing boson (see also Ref. \[375\]), which can be induced by gauge bosons of chiral inter-family symmetry \[376,377\]. All these possibilities lead to constructive interfere with the SM in muon decay such that the Fermi constant of the Lagrangian $G_F$ is smaller than the one measured from the muon lifetime. We note that while this type of solution resolves the tension between A/B determinations with C, it can only slightly alleviate the tension between A and B themselves and leads to additional tensions in the EW fit \[361\].

Concerning 2-quark-2-lepton operators, only $Q_{\ell\ell}^{(3)1111}$ is able to give a sizable BSM effect in $\beta$ decays via interference with the SM and the CAA requires $C_{\ell\ell}^{(3)1111} \approx (11 \text{ TeV})^{-2}$. Possible extensions of the SM that induce this operator are LQs \[365,378\] or a colour-neutral vector triplet \[370\]. It is also worth noting that the size required to explain the CAA is compatible with the one preferred by CMS searches for $pp \to e^+e^-$, see Sec. 2.1.6. Finally,
Table 1: Experimental determinations of the ratios $g_\ell/g_{\ell'}$ [334,383,387,388].

| $|g_\mu/g_e|$ | $\Gamma_{\tau\rightarrow\mu}/\Gamma_{\tau\rightarrow e}$ | $\Gamma_{\pi\rightarrow\mu}/\Gamma_{\pi\rightarrow e}$ | $\Gamma_{K\rightarrow\mu}/\Gamma_{K\rightarrow e}$ | $\Gamma_{K\rightarrow\pi\mu}/\Gamma_{K\rightarrow\pi e}$ | $\Gamma_{W\rightarrow\mu}/\Gamma_{W\rightarrow e}$ |
|----------|------------------|------------------|------------------|------------------|------------------|
| 1.0017 (16) | 1.0010 (9) | 0.9978 (18) | 1.0010 (25) | 0.998 (4) |

| $|g_\tau/g_e|$ | $\Gamma_{\tau\rightarrow e}/\Gamma_{\mu\rightarrow e}$ | $\Gamma_{\tau\rightarrow\pi}/\Gamma_{\pi\rightarrow e}$ | $\Gamma_{\tau\rightarrow K}/\Gamma_{K\rightarrow e}$ | $\Gamma_{W\rightarrow\tau}/\Gamma_{W\rightarrow e}$ |
|----------|------------------|------------------|------------------|------------------|
| 1.0011 (14) | 0.9646 (38) | 0.9857 (78) | 1.004 (16) |

| $|g_\tau/g_e|$ | $\Gamma_{\tau\rightarrow\mu}/\Gamma_{\mu\rightarrow e}$ | $\Gamma_{W\rightarrow\tau}/\Gamma_{W\rightarrow e}$ |
|----------|------------------|------------------|
| 1.0028 (15) | 1.021 (12) |

Scalar interactions are typically negligible for first-generation fermions, while most severely constrained from processes that display chiral enhancement [379,382].

### 2.1.5 Lepton flavour universality in the charged current

The recently observed anomalies in $b \rightarrow c\tau\nu$ and $b \rightarrow s\mu^+\mu^-$ transitions suggest a possible violation of lepton universality in other processes where strong constraints on the universality of the leptonic $W^{\pm}$ couplings $g_\ell$ ($\ell = e, \mu, \tau$), emerging from the measured weak decays of the $\mu, \tau, \pi$ and $K$, exist. The most accurate phenomenological tests of the universality of the leptonic charged-current couplings are summarised in Table 1, which updates Ref. [383].

The leptonic decays $\ell \rightarrow \ell'\nu\ell'\nu\ell'$ provide very clean measurements of the $W^{\pm}$ couplings. The $\tau \rightarrow \mu/\tau \rightarrow e$ ratio directly constrains $|g_\mu/g_e|$, while the comparison of $\tau \rightarrow e, \mu$ with $\mu \rightarrow e$ provides information on $|g_\tau/g_\mu|$ and $|g_\tau/g_e|$. Taking into account the different lepton masses involved and the small higher-order electroweak corrections, the current data confirm the universality of the leptonic $W^{\pm}$ couplings with a 0.15% precision.

A slightly better sensitivity on $|g_\mu/g_e|$ has been obtained from the precisely measured ratio of the $\pi^- \rightarrow e^-\bar{\nu}_e$ and $\pi^- \rightarrow \mu^-\bar{\nu}_\mu$ decay widths [334]. At this level of precision, a good control of radiative QED corrections is compulsory [384,385]. Comparable accuracies have been also reached from the corresponding $e/\mu$ ratios in $K_{\ell 2}$ and $K_{\ell 3}$ decays [386].

The comparison of the $\tau^- \rightarrow P^-\nu_\tau$ and $P^- \rightarrow \mu^-\bar{\nu}_\mu$ ($P = \pi, K$) decay widths allows for an independent determination of $|g_\tau/g_\mu|$. The radiative corrections to these ratios involve low-energy hadronic effects that have been recently re-evaluated [387], using Chiral Perturbation Theory techniques and the large-$N_C$ expansion. While this updated calculation agrees with previous evaluations, the estimated hadronic uncertainties are found to be slightly larger. Nevertheless, one obtains a quite accurate test of universality at the 0.4% (0.8%) level from the $\tau (K)$ ratios.

The decays $W \rightarrow \ell\bar{\nu}_\ell$ provide a more direct access to the leptonic $W$ couplings. However, with the limited statistics collected at LEP it was only possible to reach precisions of $O(1\%)$ [389]. The LEP data exhibited a slight excess of $W \rightarrow \tau\nu_\tau$ events, implying 2.7% and 2.4% deviations from lepton universality in $|g_\tau/g_\mu|$ and $|g_\tau/g_e|$, respectively. This was very difficult to reconcile with the much more precise indirect constraints from $\tau, \mu, \pi$ and $K$ decays [390].

The large amount of data provided by the LHC has made it possible to perform more precise tests of the leptonic $W$ decays. The recent ATLAS determination of $\Gamma_{W\rightarrow\tau}/\Gamma_{W\rightarrow\mu}$ [388] agrees well with the SM expectation. The ATLAS measurement alone would imply $|g_\tau/g_\mu| = \ldots$
Table 2: Experimental determinations of the ratios $g_\ell/g_\ell'$ from $W \rightarrow \ell\bar{\nu}_\ell$ decays.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>g_\mu/g_\ell</td>
<td>/g_\ell'$</td>
<td>0.996 (10)</td>
<td>0.999 (9)</td>
<td>1.002 (5)</td>
</tr>
<tr>
<td>$</td>
<td>g_\tau/g_\mu</td>
<td>/g_\ell'$</td>
<td>1.034 (13)</td>
<td>0.996 (7)</td>
<td>0.992 (10)</td>
</tr>
<tr>
<td>$</td>
<td>g_\tau/g_e</td>
<td>/g_\ell'$</td>
<td>1.031 (13)</td>
<td></td>
<td>0.997 (11)</td>
</tr>
</tbody>
</table>

0.996 ± 0.007. The larger error quoted in Table 1 reflects the sizable discrepancy with the old LEP value. A preliminary CMS measurement of the $W$ leptonic branching fractions [391], not yet included in Table 1, fully confirms the ATLAS result, eliminating the long-standing $W \rightarrow \tau$ anomaly. The separate results from LEP and the LHC experiments are collected in Table 2, which also displays the preliminary world averages including the Tevatron data.

Following the PDG prescription, the errors of the $|g_\tau/g_\mu|/g_\ell'$ and $|g_\tau/g_e|/g_\ell'$ averages have been increased to account for the discrepancy with the LEP values.

Clearly, the current data verify the universality of the leptonic $W$ couplings to the 0.15% level. In Table 1 one can only identify two small deviations that do not reach the 2σ level: there is a slight (1.9σ) excess of $\tau \rightarrow \mu$ versus $\mu \rightarrow e$ events, and a small deficit (1.8σ) of $\tau \rightarrow K$ versus $K \rightarrow \mu$ transitions. The relatively large hadronic uncertainty involved could easily explain the second deviation, although a systematic deficit of kaon final states seems to be present in $\tau$ decays [383], leading to a determination of $|V_{us}|$ slightly lower than the one obtained from kaon decays [392,394]. The slight excess of $\tau \rightarrow \mu$ events could be correlated with a possible explanation of the Cabibbo anomaly through a slight violation of lepton universality [337,372,373]. In any case, more precise experimental studies are needed.

The different universality tests provide complementary information, since they are sensitive to different types of NP contributions. While the decays of the $W$ boson probe directly its leptonic couplings, the indirect constraints from low-energy leptonic and semileptonic decays test the potential presence of additional intermediate particles, which could modify each analysed process in a different way. From this point of view, it is worth to mention the universality test extracted from the ratio of $B \rightarrow D(\ast)\mu\nu$ and $B \rightarrow D(\ast)e\nu$ transitions: $|g_\mu/g_e|/g_\ell' = 0.989 (12)$ [395]. Although much less precise that the other indirect determinations in Table 1 it severely restricts the type of possible explanations to the $b \rightarrow c\tau\nu$ anomaly.

### 2.1.6 Non-resonant di-electrons

A search for LFUV in the non-resonant production of di-leptons was recently performed by CMS, observing a ≈ 4σ excess in electron pairs with an invariant mass greater than 1.8 TeV in the $pp \rightarrow e^+e^-$ channel. As the muon channel agrees with the SM expectation, this measurement points towards LFUV. Also ATLAS [398] and HERA [399] found more di-electrons than expected in their studies of quark-lepton contact interactions.

In addition to the total cross-section, the CMS collaboration provided the differential cross-section ratio

$$R_{\mu\mu/ee} = \frac{d\sigma(pp \rightarrow \mu^+\mu^-)/dm_{\mu\mu}}{d\sigma(pp \rightarrow e^+e^-)/dm_{ee}},$$

for different $m_{\ell\ell}$ ($\ell = e, \mu$) bins. For each bin, they quoted two values, distinguishing the cases where zero (at least one) of the di-leptons were detected in the endcaps, corresponding
Figure 7: The double ratio values $R_{\mu\mu/ee,ij}/R_{\mu\mu/ee}^{\text{SM}}$ measured by the CMS collaboration in nine $m_{\ell\ell}$ bins between 200 and 3000 GeV are shown (black circles and squares) together with the best fit curves for different 2-quark–2-lepton operator scenarios (coloured lines).

In this double ratio, many of the experimental and theoretical uncertainties cancel [400]. It was further normalised to one in the bin from 200 to 400 GeV to correct for the relative sensitivity to electrons and muons. The measurements are indicated by the black squares (circles) for the barrel only (endcap) measurements in Fig. 7. A trend towards values smaller than 1 is visible for large $m_{\ell\ell}$.

ATLAS followed a different strategy, measuring the differential cross-section
\[ \frac{d\sigma(pp \to \ell^+\ell^-)/dm_{\ell\ell}}{d\sigma} \quad \text{for} \quad \ell = e, \mu, \]
integrated over the signal regions [2.2, 6.0] TeV and [2.7, 6.0] TeV for searches of NP interfering constructively and destructively with the SM, respectively [398]. Comparing these results to the SM predictions, they derived exclusion limits for NP coupling equally to up and down quarks of both chiralities. Interestingly, also ATLAS measured more di-electrons than expected in the constructive channel, but their measurements are still consistent with the SM prediction.

Potential explanations to the di-electron excess in the CMS measurements, including the constraints from the ATLAS analysis, have recently been discussed in Refs. [365, 401, 402]. The presence of NP was studied within an EFT approach, considering only operators coupling to first-generation quarks and leptons. The Wilson coefficients $C_i$ of the relevant operators are reported in Table 3. For each of them, the ratio $R_{\mu\mu/ee,ij}^{\text{SM+NP}}(C_i)/R_{\mu\mu/ee,i}^{\text{SM}}$ was
Table 3: Wilson coefficient values \((\times 10^{-3} \text{ TeV}^{-2})\) that are preferred (excluded) by the CMS (ATLAS) \(pp \rightarrow e^+e^-\) measurements. For CMS we state the 1\(\sigma\) ranges as well as the pull to the SM. The asterisk labels scenarios that do not provide an improved fit compared to the SM point. For ATLAS we show the 95\% C.L. exclusion limits.

\[
\chi^2(C_i) = \sum_{i=1,\ldots,9}^{j=e,b} \left( \frac{R_{\text{Data}}^{\mu\mu/ee,ij}}{R_{\text{SM}}^{\mu\mu/ee,ij}} - \frac{R_{\text{SM}+NP}^{\mu\mu/ee,ij}(C_i)}{R_{\text{SM}}^{\mu\mu/ee,ij}} \right)^2 \sigma_{ij}^2, \tag{28}
\]

where \(\sigma_{ij}\) are the corresponding uncertainties reported in Ref. \([403]\). Additionally, the ATLAS exclusion limits were recast for the cases where the NP couples differently to up and down quarks and to the different chiralities. The preferred (excluded) values for the Wilson coefficient, given by the CMS (ATLAS) analysis are shown in Table 3 and the best fits to CMS data are displayed in Fig. \[\text{a}\]. NP interfering constructively with the SM contribution is preferred and the operators considered can improve the fit with a pull of up to \(3.3\sigma\). While the corresponding Wilson coefficient values are not excluded by the ATLAS analysis, constraints coming from kaon decays, \(K^0-\bar{K}^0\) mixing or \(D^0-\bar{D}^0\) mixing are more stringent and need to be considered when addressing the CMS excess. Finally, it is important to note that the operator \([Q_{\ell q}^{(3)}]_{1111}\) can also provide an explanation to the Cabibbo Angle Anomaly, discussed in Sec. 2.1.4. In Ref. \([402]\) it was shown that \([Q_{\ell q}^{(3)}]_{1111}\) can provide a simultaneous explanation for these two anomalies and a best-fit value of \([C_{\ell q}^{(3)}]_{1111} = 1.1/(10 \text{ TeV})^2\) was extracted.

### 2.1.7 Forward-backward asymmetry in semi leptonic B decays

The observable \(\Delta A_{FB}\) encodes the difference of the forward-backward asymmetry in \(B \rightarrow D^*\mu\nu\) vs \(B \rightarrow D^*e\nu\). Like for \(R(K^{(*)})\) the muon and electron mass can both be neglected such that the form-factor dependence cancels and the SM prediction is, to the currently relevant precision, zero. Even though the corresponding measurements of the total branching ratios are consistent with the SM expectations \([404,405]\), recently Ref. \([406]\) unveiled a \(\approx 4\sigma\) tension in \(\Delta A_{FB}\), extracted from \(B \rightarrow D^*\ell\bar{\nu}\) data of BELLE \([407]\).

A good fit to data requires a non-zero Wilson coefficient of the tensor operators. Importantly, among the set of renormalisable models, only two scalar LQ can generate this operator at tree level and only the \(SU(2)_L\) singlet gives a good fit to data \([408]\). However,
even in this case, due to the constraints from other asymmetries, $\Delta A_{FB}$ cannot be fully explained but the global fit to $b \to c\mu\nu$ and $b \to cee$ data can be improved by more than $3\sigma$ [408].

2.1.8 Combined explanations

In the presence of the deviations from the SM discussed above, it is reasonable to explore NP models that are able to address, in a coherent manner, more than one anomaly. While it is often possible to simply combine independent solutions for each anomaly, we consider here combined explanations in the sense that the mediators are either directly connected from a UV perspective or their joint contribution to some anomalous observable is crucial for a successful explanation.

Neutral and Charged-current $B$-anomalies When combining $b \to s\ell^+\ell^-$ and $R(D^{(*)})$ anomalies, the viable scenarios are dictated by those that offer a good explanation of $R(D^{(*)})$, since this observable is the one that requires the lowest NP scale. As discussed in Sec. 2.1.2, the possible solutions necessarily involve LQs, specifically the vector $U_1 \sim (3, 1, 2/3)$ or the scalars singlet $S_1 \sim (3, 1, 1/3)$ or doublet $R_2 \sim (3, 2, 7/6)$. Of these, only the vector LQ $U_1$ also generates a viable contribution to $b \to s\ell\ell$, and is thus the only single-mediator scenario for combined explanations of $B$-anomalies [20, 62, 70, 113-115, 118, 120, 409-413]. Alternatively, two scalar leptoquarks can also provide a good explanation, by adding the triplet $S_3 \sim (3, 3, 1/3)$, that mediates successfully $b \to s\ell\ell$, to $S_1$ or $R_2$.

The $S_1 + S_3$ scenario [120, 134, 414, 421] has two viable possibilities: if the $S_1$ couplings to right-handed fermions vanish, then both $S_1$ and $S_3$ contribute to $R(D^{(*)})$ via $C_V^L$; if instead also right-handed couplings are involved, then the largest contribution to $R(D^{(*)})$ arises from $S_1$ via $C_S^L \approx -4C_T$ coefficients. From a UV perspective, these two LQ could arise as pseudo-Nambu-Goldstone bosons from a strongly-coupled UV sector, together with the SM Higgs [414, 420].

In the $R_2 + S_3$ scenario the two leptoquarks contribute mostly independently to $R(D^{(*)})$ and $b \to s\ell\ell$, respectively [114, 422, 423]. However, these two scalars have been proposed as arising from the same GUT scenario [112].

$B$-anomalies and anomalous muon magnetic moment This scenario can be seen as a further phenomenological requirement on the models discussed in the previous paragraphs. Of the three setups presented as combined explanations of $B$-anomalies, only those involving the scalar leptoquarks $S_1$ and $R_2$ can also provide a satisfactorily contribution to $a_\mu$, thanks to the $m_t/m_\mu$ enhancement. In this case, however, given the presence of couplings to both left and right-handed muons and tau, a too-large contribution to $\tau \to \mu\gamma$ is induced. This can be cancelled by other contributions from the same mediator, with a tuning at the level of one part in three at the amplitude level [416, 417, 419].

$B$-anomalies, anomalous muon magnetic moment, and CAA Possible combined explanations for all these anomalies has been recently proposed involving the scalar LQ $S_1$ and a scalar charged singlet $\phi^+$ [374]. In this scenario $S_1$ contributes to $R(D^{(*)})$ and $a_\mu$ in the same way as described above, while $\phi^+$ generates a tree-level contribution to the muon decay $\mu^- \to e^-\nu_\mu\bar{\nu}_e$ that, shifting the Fermi constant, improves the fit of the Cabibbo angle $\theta_{13}$ [361, 372, 373, 424]. The contribution to $b \to s\mu\mu$ is instead induced via a one-loop box diagram involving both $S_1$ and $\phi$. While this model is able to pass all present
Final state & Characteristics & SM backgrounds & Significance \\
\( \ell^+ \ell^- + b\text{-jets} \) & \( m_{\ell\ell} < 100 \text{ GeV}, \text{ low } b\text{-jet mult.} \) & \( t\bar{t}, Wt \) & > 5\( \sigma \) \\
\( \ell^+ \ell^- + \text{jet veto} \) & \( m_{\ell\ell} < 100 \text{ GeV} \) & \( W^+ W^- \) & \( \approx 3\sigma \) \\
\( \ell^\pm \ell^\mp, 3\ell + b\text{-jets} \) & Moderate \( H_T \) & \( tW^\pm, t\bar{t}t \) & > 3\( \sigma \) \\
\( \ell^\pm \ell^\mp, 3\ell, n_b = 0 \) & In association with \( h \) & \( W^\pm h, WW \) & \( \approx 4.5\sigma \) \\
\( Z(\to \ell\ell) \ell, n_b = 0 \) & \( p_{\ell\ell} < 100 \text{ GeV} \) & \( ZW^\pm \) & > 3\( \sigma \) \\

Table 4: Summary of the status of the multi-lepton anomalies at the LHC, where \( \ell = e, \mu \).

bounds, the masses for these scalars are required to be at about 5 TeV, with some couplings reaching somewhat large values of \( \approx 3 \), which are at the threshold of limits from perturbative unitarity.

An alternative BSM scenario that can account for both neutral and charged-current B-anomalies, as well as the muon \( g - 2 \) is based on a minimal \( R \)-parity violating (RPV) SUSY framework with relatively light third-generation sfermions. Although the \( LQD \)-type RPV couplings of squarks somewhat resemble the scalar \( LQ \) couplings, there are some key differences in case of RPV, such as the possibility of explaining muon \( g - 2 \) purely via \( LLE \)-type couplings, same chiral gauge structure as in the SM, natural flavour violation, as well as many other attractive inbuilt features of SUSY, such as radiative stability of the Higgs boson, radiative neutrino masses, radiative electroweak symmetry breaking, stability of the electroweak vacuum, gauge coupling unification, (gravitino) dark matter and electroweak baryogenesis.

### 2.2 Multi-lepton anomalies at the LHC

One of the implications of a two-Higgs doublet model with an additional singlet scalar \( S \) (2HDM+\( S \)), is the production of multiple-leptons through the decay chain \( H \to Sh, SS \), where \( H \) is the heavy CP-even scalar and \( h \) is considered as the SM Higgs boson with mass \( m_h = 125 \text{ GeV} \). Excesses in multi-lepton final states were reported in Ref.\(^{430}\). In order to further explore results with more data and new final states, while avoiding biases and look-else-where effects, the parameters of the model were fixed in 2017 according to Refs.\(^{429,430}\). This includes setting the scalar masses to \( m_H = 270 \text{ GeV} \), \( m_S = 150 \text{ GeV} \), treating \( S \) as a SM Higgs-like scalar and assuming the dominance of the decays \( H \to Sh, SS \). Statistically compelling excesses in opposite sign di-leptons, same-sign di-leptons, and three leptons, with and without the presence of \( b \)-tagged hadronic jets were reported in Refs.\(^{432,434}\). The possible connection with the anomalous magnetic moment of the muon \( g - 2 \) was reported in Ref.\(^{435}\). Interestingly, the model can explain anomalies in astro-physics (the positron excess of AMS-02\(^{436}\) and the excess in gamma-ray fluxes from the galactic center measured by Fermi-LAT\(^{437}\) (see Sec. 2.5) if it is supplemented by a Dark Matter candidate\(^{438}\).

We give a succinct description of the different final states and corners of the phase-space that are affected by the anomalies. The anomalies are reasonably well captured by a 2HDM+\( S \) model. Here, \( H \) is predominantly produced through gluon-gluon fusion and decays mostly into \( H \to SS, Sh \) with a total cross-section in the rage 10-25 pb\(^{432}\). Due to the relative large Yukawa coupling to top quarks needed to achieve the above mentioned direct

\textsuperscript{6}A possible candidate of the lighter scalar has been reported in Ref.\(^{431}\).
production cross-section, the production of $H$ in association with a single top-quark\textsuperscript{7}. These production mechanisms together with the dominance of $H \rightarrow SS, Sh$ over other decays, where $S$ behaves like a SM Higgs-like boson, lead to a number of final states that can be classified into several groups of final states. There are three groups of final states where the excesses are statistically compelling: opposite sign (OS) leptons ($\ell = e, \mu$); same sign (SS) and three leptons ($3\ell$) in association with $b$-quarks; SS and $3\ell$ without $b$-quarks. In the sections below a brief description of the final states is given with emphasis on the emergence of new excesses in addition to those reported in Refs. \cite{430, 432, 434}, when appropriate. It is important to reiterate that the new excesses reported here are not the result of scanning the phase-space, but the result of looking at pre-defined final states and corners of the phase-space, as predicted by the model described above.

2.2.1 Opposite sign di-leptons

The production chain $pp \rightarrow H \rightarrow SS, Sh \rightarrow \ell^+\ell^- + X$, is the most copious multi-lepton final state. Using the benchmark parameter space in Ref. \cite{439}, the dominant of the singlet are $S \rightarrow W^+W^-, b\bar{b}$. This will lead to OS leptons with without $b$-quarks. The most salient characteristics of the final states are such that the di-lepton invariant mass $m_{\ell\ell} < 100$ GeV where the bulk of the signal is produced with low $b$-jet multiplicity, $n_b < 2$ \cite{432}. The dominant SM background in events with $b$-jets is $t\bar{t}W^\pm$. It is important to note that the $b$-jet and light-quark of the signal is significantly different from that of top-quark related production mechanisms. As a matter of fact, excesses are seen when applying a full jet veto, top-quark backgrounds become suppressed and where the dominant backgrounds is non-resonant $W^+W^-$ production \cite{430, 440, 441}. A review of the NLO and EW corrections to the relevant processes can be found in Refs. \cite{432, 441}, where to date the $m_{\ell\ell}$ spectra at low masses remains unexplained by MC tools. A measurement of the differential distributions in OS events with $b$-jets with Run 2 data further corroborates the inability of current MC tools to describe the $m_{\ell\ell}$ distribution \cite{443}. A summary of deviations for this class of excesses is given in Table 4.

2.2.2 SS and 3$\ell$ with $b$-quarks

The associated production of $H$ with top quarks lead to the anomalous production of SS and $3\ell$ in association with $b$-quarks with moderate scalar sum of leptons and jets, $H_T$. The elevated $t\bar{t}W^\pm$ cross-section measured by the ATLAS and CMS experiments can be accommodated by the above mentioned model \cite{432, 433}. Based on a number of excesses involving $Z$ bosons, in Ref. \cite{439} it was suggested that the $CP$-odd scalar of the 2HDM+$S$ model could be as heavy as $m_A \approx 500$ GeV, where the two leading decays would be $A \rightarrow t\bar{t}, ZH$. The cross-section for the associated production $pp \rightarrow t\bar{t}A$ with $A \rightarrow t\bar{t}$ would correspond to $\approx 10$ fb. This is consistent with the elevated $t\bar{t}t\bar{t}$ cross-section reported by ATLAS and CMS \cite{444, 446}. The combined significance of the excesses related to the cross-section measurements of $t\bar{t}W^\pm$ and $t\bar{t}t\bar{t}$ surpass 3$\sigma$, as detailed in Table 4. It is important

\textsuperscript{7}As the $HWW$ coupling is suppressed with respect to the top Yukawa coupling, the associated production of $H$ with a single top not suppressed.

\textsuperscript{8}While the ATLAS experiment provided unfolded cross-sections, the CMS experiment presented a comparison of the yields in the data to a MC \cite{442}. The MC does not describe simultaneously the di-lepton invariant mass and transverse momentum. As such, for the purposes of this study this data set is inconclusive, where we encourage the CMS experiment to provide differential measurements similar to those performed by ATLAS in Ref. \cite{440}.

29
to note that the ATLAS collaboration has reported a small excess in the production of four leptons with a same flavour OS pair consistent with a $Z$ boson, where the four-lepton invariant mass, $m_4 < 400$ GeV [447]. This excess can also be accommodated by the direct production of $A \to ZH$.

### 2.2.3 SS and 3$\ell$ without $b$-quarks

The production chain $pp \to H \to SS,Sh$ can give rise to SS and 3$\ell$ events, where $b$-jet activity would be depleted compared to production mechanism considered in Sec. 2.2.2. The potential impact on the measurement of the production of the SM Higgs boson in association with a $W$ boson and other measurements in the context discussed here was reported in Ref. [448]. A survey of available measurements of the signal yield of the $Wh$ production was performed in Ref. [434]. A deviation of $3.8\sigma$ with respect to the $Wh$ yield in the SM in corners of the phase-space predicted by the simplified model. The CMS experiment has recently reported the signal strength of the $Vh, V = Z, W^\pm$ production with the $h \to W^+W^-$ decay for low and high $V$ transverse momentum [449]. The signal strength for $Vh$ with the $V$ transverse momentum, $p_{TV} < 150$ GeV, where the BSM signal is concentrated, is $2.65^{+0.69}_{-0.64}$. This deviates from the SM value by an additional $2.6\sigma$. It is worth noting that in order to reconcile observed excesses in Secs. 2.2.1 and 2.2.2 with the ones described here, it is necessary to assume the dominance of the $H \to SS$ decay over $H \to Sh$ [434]. Another important prediction of the simplified model is the elevated $WWW$ cross-section. The ATLAS experiment reports a signal strength of $1.66 \pm 0.28$ [450]. The latter includes the $Wh \to WWW^*$ production, hence it is not added to the combination due to partial double counting. Lastly, another final state of interest is the production of $ZW^\pm$ events where $Z$ transverse momentum, $p_{TZ} < 100$ GeV with depleted $b$-jet activity. Excesses were reported in Ref. [432]. The CMS experiment has recently reported an important excess in events with 3$\ell$ in association with one and two jets used for the measurement of $Zh, h \to W^+W^-$ production, where $ZW^\pm$ is the dominant background [449]. As the analysis of the excess in the context of the simplified model described here is in progress, the significance of this excess is not added to the combination reported in Table 4.

### 2.3 Higgs-like excess at $\approx 96$ GeV

The LEP experiments reported a mild excess with a local significance of $2.3\sigma$ in the search for a SM Higgs boson [452] using the process $e^+e^- \to Zh(\to b\bar{b})$. The largest excess was observed for the $b\bar{b}$ invariant mass of 98 GeV. Renewed interest in this excess emerged with the CMS experiment reporting similar excesses with Run 1 and $35.9$ fb$^{-1}$ of Run 2 data [453], with a local significance of $2.8\sigma$ at 95.3 GeV. The ATLAS experiment has reported the results of a search with over $80$ fb$^{-1}$ of integrated luminosity [454]. No excess was found there, but the measured limit does not exclude the results from the CMS experiment. As such, sufficient data is available in the complete Run 2 data set to understand if the above mentioned excesses are due to statistical fluctuations. While the excess described above is not yet statistically compelling, it has been studied by a number of authors, as we will review below.

---

9CMS [451] pursues a different approach compared to the more inclusive selection performed by ATLAS. For instance, the requirement that the azimuthal separation between the vector of the three leptons and the missing transverse energy be greater than $2.5$ rad and other requirements suppress the contribution from the BSM signal considered here.
2.3.1 Interpretation as SM Higgs-like scalar boson

The LEP and CMS excesses can be interpreted as signal from a SM Higgs-like boson with a mass around 96 GeV, as was done for instance in Ref. [455]. In this interpretation the signal strength in terms of a would-be SM Higgs boson with this mass is [456,457]:

\[
\mu_{\text{LEP}} = \frac{\sigma(e^+e^- \rightarrow ZS \rightarrow Zb\bar{b})}{\sigma(e^+e^- \rightarrow Zh_{98} \rightarrow Zb\bar{b})} = 0.117 \pm 0.057
\] (29)

and

\[
\mu_{\text{CMS}} = \frac{\sigma(pp \rightarrow S \rightarrow \gamma\gamma)}{\sigma(pp \rightarrow h_{96} \rightarrow \gamma\gamma)} = 0.6 \pm 0.2
\] (30)

where \(h_X\) stands for a SM Higgs-like boson with \(m_h = X\) GeV. A few comments are in order: First, the measurement in Eq. (29) is mostly sensitive to the coupling \(h_{98}ZZ\), the size of which is compatible with the coupling measurements of the SM Higgs boson at the LHC [458–461]. Second, the masses are compatible when one takes the coarse resolution of the LEP measurement into account [462]. Third, the branching ratio of a SM Higgs-like scalar, expected to be \(\text{Br}(h_{96} \rightarrow ZZ^*) < 1 \cdot 10^{-3}\), makes the observation of an excess with 4\ell events at the LHC very difficult with the available data set [463,464].

It is important to note, that a SM Higgs-like boson with such a mass decays predominantly into \(b\bar{b}\), where the branching ratio \(\text{Br}(h_{96} \rightarrow b\bar{b}) > 0.8\), and that \(\mu_{\text{CMS}} > \mu_{\text{LEP}}\). Therefore the two signal strengths reported by CMS and LEP suggest an underlying model that is more complicated and has more than just one additional degree of freedom, which has been addressed in Ref. [455] by including additional light fields, and considering a more extended scalar sector.

2.3.2 Two Higgs doublet model

A two Higgs doublet model was used in Refs. [455,465] to explain the excess, the key being a moderately-to-strongly fermiophobic CP-even scalar boson with a mass of 95 GeV. The scalar’s fermiophobic nature suppresses the top contribution and thus enhances its di-photon rate, such that the signal strength of around 0.1 pb (cf. Eq. (30)) can be explained by the combination of VBF, \(W h\) and \(Zh\) production alone. The authors show that sizable scalar production rates can also arise from either top-quark pair and single-top production followed by decay chains involving (comparatively light) charged scalars, or associated production from the decays of a heavier CP-odd scalar bosons.

An extension of the two Higgs doublet model with an additional Higgs singlet field, the so-called “next to minimal two Higgs doublet model” can fit the CMS excess as demonstrated in Ref. [457]. In this model the additional (potentially light) CP-even scalar resonance is largely unconstrained by other phenomenological observations and therefore a more general fit is possible. The N2HDM of type II and type IV (flipped) can fit the two excesses, while type I and type III (lepton specific) cannot. The fit prefers the lowest possible masses for the charged Higgs bosons around 650 GeV and \(\tan \beta\) just above 1.

2.3.3 Supersymmetric models

It is not possible to explain the 96 GeV scalar in the MSSM [466]. A well-motivated extension of the minimal Supersymmetric Standard Model is given by the Next-to-MSSM (NMSSM), where a mostly singlet-like Higgs that mixes with a doublet component of about 10\% (mixing
squared) is a natural candidate to explain the excess \cite{462,467}. In this kind of models, the
decays of such a light Higgs boson depend on the mixing between the different Higgs states.
In particular the couplings to down- or up-type quarks can be suppressed to meet the LEP
measurement of $b\bar{b}$ in Eq. \cite{29}, while the di-photon rate can be moderately enhanced to meet
the CMS measurement in Eq. \cite{30}.

The $\mu\nu$SSM is a supersymmetric extension of the Standard Model that incorporates
neutrino physics. In this model the additional light scalar can be interpreted as the CP-even
right-handed sneutrino and sizable mixing of the sneutrino and the doublet Higgses
are possible. In general, it is easier to reproduce the LEP observed signal strength in
supersymmetric models than the CMS one \cite{468}. This model shares many features with the
NMSSM, however, the link to the neutrino sector breaks the $R$-parity and thus avoids Dark
Matter bounds, freeing up the parameter space to accommodate the excess.

### 2.3.4 Other explanations

The multiphoton signals have been explained in the context of other models too. For instance
the lightest new particle predicted within the Randall Sundrum (RS) model, the radion, was
discussed as possible explanation of the CMS excess in Ref. \cite{469}. A model with an extra
$U(1)$ and heavy vector-like quarks that also addresses the flavour anomalies was discussed in
Ref. \cite{470}. The anomaly-free $U(1)_Y'$ extension of the SM with two complex scalar singlets
were discussed in Ref. \cite{471}.

### 2.4 Neutrino Anomalies

In the SM active neutrinos carry isospin charge $\pm \frac{1}{2}$ under the weak interaction, where
they have left-handed helicity. However, due to the establishment of neutrino oscillations
neutrinos are known to have mass, even if small. The observation of solar and very-long-
baseline reactor neutrino oscillations and the observation of atmospheric and long-baseline
accelerator neutrino oscillations give vigorous evidence of three-neutrino mixing \cite{472,473}.
This poses a problem for the SM. Currently, the origin of mass of neutrinos remains an
open question in particle physics. Hence, the neutrino sector is of invaluable significance for
particle physics.

A number of anomalies beyond the three-neutrino mixing framework have emerged in
short-baseline neutrino experiments with $L/E_\nu \approx 1$ m/MeV, where $L$ is the distance that the
neutrino traverses and $E_\nu$ is the neutrino energy. These can be classified into two groups:
neutrino appearance and disappearance. This section covers anomalies in measurements
at fixed-target experiments that may point to physics beyond the standard three-flavour
framework.

#### 2.4.1 Appearance anomalies

The appearance anomalies pertain to excesses in the production of $\nu_e$ and $\bar{\nu}_e$ via charge-
current quasi-elastic production reported by the LSND \cite{474} and MiniBooNE \cite{475} experiments. The MiniBooNE experiment observed a total of $460.5 \pm 99 \nu_e$ and $\bar{\nu}_e$ excess events
in the energy range $200 < E_\nu < 1250$ GeV. This excess corresponds to a significance of $4.7\sigma$
and a recent update confirms previous findings \cite{476}. The two-neutrino oscillation interpretation
of the MiniBooNE data is consistent with the corresponding allowed region obtained
with data from the LSND experiment. The combined significance of the MiniBooNE and
LSND excesses reaches 6σ. The MiniBooNE excess has been studied in terms of the two-neutrino mixing, where \((\Delta m^2, \sin^2 2\theta) = (0.041 \text{ eV}^2, 0.96)\) \([475]\). The MiniBooNE excess will be further studied by the Fermilab short-baseline neutrino programme \([477]\).

### 2.4.2 Disappearance anomalies

Anomalies pertaining to the disappearance of \(\nu_e\) and \(\bar{\nu}_e\) are observed in reactor neutrino experiments \([478-480]\) and Gallium radioactive source experiments anomalies \([481, 482]\). The latter emerge from the GALLEX \([483-485]\) and the SAGE \([486-489]\) experiments. Gallium radioactive source experiments are solar neutrino experiments with intense artificial \(^{51}\text{Cr}\) and \(^{37}\text{Ar}\) radioactive sources inside the detectors, where the inverse beta reaction \(\nu_e + ^{71}\text{Ga} \rightarrow ^{71}\text{Ge} + e^-\) is used. These experiments indicate that the ratio of measured to predicted \(^{71}\text{Ge}\) occurrences is less than unity. The significance of the deficit is about 3σ, and it is referred to as the Gallium anomaly.

This anomaly has been studied within the framework of two-neutrino mixing, where \(\Delta m^2 > 0.1 \text{ eV}^2\), also orders of magnitude greater than those of the solar and atmospheric oscillations \([481, 482]\). The compatibility of the Gallium and the reactor anomalies has also been studied within the context the two-neutrino mixing, where reasonable agreement between the two disappearance anomalies is found \([481]\).

### 2.4.3 Explanation via light neutrinos

Both types of anomalies indicate that the mass splitting, \(\Delta m^2\), has to be much larger than the three-neutrino mixing scheme inferred from the observation of neutrino oscillations in solar, very-long-baseline reactor, atmospheric and long-baseline accelerator experiments.

Light sterile neutrinos with masses around a few eV have been considered as a possibility to explain the above mentioned anomalies. That said, when interpreting all the anomalies simultaneously in terms of sterile neutrinos through global fits, it is not possible to provide a consistent interpretation \([490, 493]\). In this light, a comprehensive global programme is under way to study the dependence of the anomalies on distance and energy in order to conclusively ascertain the origin of the anomalies \([494]\).

### 2.4.4 Non oscillatory explanations of the appearance anomaly

Possible non-oscillatory explanations of the MiniBooNE anomaly were discussed in a model-independent way in Ref. \([495]\). Most of the existing explanations make use of shortcomings in particle identification of the MiniBooNE detector. Not all explanations address the LSND anomaly, and few produce a good fit to the observed excess spectra.

The largest class of models explaining the MiniBooNE anomaly introduces heavy neutrinos with masses \(\mathcal{O}(100) \text{ MeV}\) that are produced in the detector from upscattering of the muon neutrinos, and which subsequently decay into various final states that give rise to electromagnetic radiation \([496, 504]\).

Another class of models introduces heavy neutrinos with masses around \(\mathcal{O}(1) \text{ MeV}\) that are produced at the target and in the decay pipe, which rapidly decay into a final state including an electron neutrino and thus enhance the electron neutrino flux at the detector \([505, 508]\).

A variation of the above explanations, where heavy neutrinos are produced at the target and travel into the detector to decay into electromagnetic radiation was discussed in Ref. \([509]\). A different explanation was given in Ref. \([510]\) where short baseline oscillations
were explained via a strong medium potential, including a new resonance scattering of neutrinos on the local overly dense relic neutrino background.

2.5 Astrophysical Anomalies

The field of astroparticle physics has been replete with observational anomalies over the past decades.

2.5.1 Positron excesses

The first case we will discuss will be the positron excesses observed by multiple experiments and at multiple energies. Observations of consistent excess positrons above the expected backgrounds, at both GeV and TeV energy ranges, by the PAMELA and ATIC instruments respectively [511–513] as well as the HESS telescope [514] (at TeV scales) generated initial speculation that the consequences of BSM physics (dark matter in particular) had been observed [515–517]. These excesses were strengthened by further data collected by Alpha Magnetic Spectrometer (AMS and AMS-02) collaboration [436] as well as the Fermi and Dark Matter Particle Explorer (DAMPE) satellites [518, 519]. This has resulted in strong evidence of unexplained positrons at both the GeV and TeV energy scales. The AMS-02 and DAMPE data have been thoroughly examined for consistency with various dark matter models. In particular the DAMPE excess is considered to favour a leptophilic scenario where dark matter couples to electrons and/or muons via heavy mediator [520–522]. In the case of PAMELA, Fermi-LAT and AMS-02 the treatment has often been “model independent” (where annihilation/decay channels are considered independently) such as in Refs. [523, 524]. In particular, AMS-02 offers one of the most powerful indirect dark matter probes. Authors that study AMS-02 as an observed dark matter signature again find the data favours leptophilic dark matter [525–527], coupling either directly to electrons and muons or via scalar/gauge mediators. The amplitude of the AMS-02 and DAMPE fluxes typically requires some form of boosting in annihilating cases, such as Sommerfeld enhancement [528], or a local over-density [520–522]. The continued existence of these leptonic excesses in astroparticle physics is clearly fertile hunting ground for new phenomena. For most of the studied models, collider searches remain difficult as the mediator mass is necessarily of the order of few TeV.

2.5.2 Anti-protons

Anti-proton spectra also display excesses as observed by PAMELA and then AMS-02 [529, 530]. The impact of dark matter on the anti-proton content of cosmic-rays had been studied prior to these experimental results in [531]. The significance of the excess and it favouring DM explanations has been extensively argued [532–539]. However, it has recently been shown that the excess significance is almost entirely erased when correlated errors are taken into account [540]. In addition, simple DM models face difficulties maintaining consistency across different targets and frequency bands [541, 542]. A common difficulty faced in this regard is that annihilation/decay channels that produce copious anti-protons also result in large gamma-ray fluxes that can be difficult to reconcile with observational data.

2.5.3 Galactic Centre gamma-ray excess

An anomalous excess of gamma-rays in the Milky-Way galactic centre, observed with the Fermi-LAT instrument, has long been studied as a potential signature of BSM physics [437]
This particular excess has been controversial, with competing “mundane” explanations, as well as concerns over modelling and the effect of unresolved sources. Despite these concerns there are still indications that the excess is significant and favours DM explanations, provided the Milky-Way halo is a contracted Navarro-Frenk-White profile. Similar to the anti-proton case, simple DM models face difficulties in that their indirect emissions in other sources and in the radio frequency range often exceed existing observational data.

### 2.5.4 3.5 keV X-ray line

An unidentified X-ray emission line near 3.5 keV has been detected by a number of groups using XMM-Newton, Chandra and Suzaku datasets. The most popular explanation so far is in terms of a decaying keV-scale sterile neutrino dark matter. The dark matter interpretation was ruled out with blank-sky observations, but these results have been disputed in Refs. To add to this confusion, there exist various other astrophysical as well as non-astrophysical explanations for the line, and in particular, there are many atomic lines in this energy range, which could mimic the signal. Hopefully, the situation will be clarified with more precise data from future X-ray missions, such as Athena.

### 2.5.5 511 keV gamma-ray line

The INTEGRAL satellite has detected an excess of 511 keV photon emission from electron-positron pair annihilation in the central region of Milky Way relative to astrophysical expectations. This was recently confirmed by COSI. The origin of the positrons in the Galactic Centre is still actively debated and it is plausible that it could originate from sources associated with BSM physics, such as pair-annihilation of MeV-scale dark matter particles. Next-generation MeV gamma-ray telescopes such as eASTROGAM with better point source detection sensitivity would provide a clearer picture.

### 2.5.6 21-cm absorption line

The EDGES experiment has reported an anomalous absorption profile centred at 78 MHz (corresponding to a redshift $z = 17$) in the sky-averaged 21-cm spectrum. This indicates an unexpected cooling of the hydrogen gas during or prior to the so called Cosmic Dawn era, as compared to the ΛCDM model expectations. The cosmological interpretation of the EDGES anomaly is subject to significant instrumental and systematic uncertainties. Nevertheless, milli-charged sub-GeV dark matter has emerged as a popular BSM candidate to explain this $3.8\sigma$ anomaly. Some issues regarding this interpretation were raised in Refs., which however can be overcome. The new HERA upper limits on the power spectrum of 21-cm fluctuations in the early Universe neither support nor disfavour the EDGES detection, but future measurements from HERA at lower frequencies will be important for further testing this anomaly.

### 2.5.7 XENON1T electron excess

The XENON1T collaboration has reported an excess of electron recoil events over the known backgrounds in the 2-3 keV recoil energy range. There have been several BSM interpretations of this anomaly, such as solar axions, boosted dark matter.
and neutrino magnetic moment \[595,596\]. However, the possibility of a tritium beta-decay background in the detector can neither be confirmed nor excluded so far as the cause of this anomaly \[587\].

### 2.5.8 DAMA/LIBRA annual modulation

The DAMA/LIBRA collaboration has reported an annual modulation signal in their NaI detector with a high statistical significance of \(12.9\sigma\) \[597\]. Although the signal is compatible with general expectations from halo dark matter in the Milky Way \[598\], the DAMA/LIBRA-preferred DM parameter space is now strongly disfavoured by many other direct-detection experiments \[599\]. There exist a handful of ideas \[600–602\] to reconcile the DAMA result with other null results; however, it is still a contentious issue whether the DAMA/LIBRA signal is indeed due to dark matter in the first place; see, e.g., Refs. \[603, 604\]. This needs to be independently verified using the same target material, which is the main goal of two currently running experiments, COSINE \[605\] and ANAIS \[606\].

### 2.6 Cosmological Anomalies

There are several cosmological anomalies that have slowly become more robust as available data has increased.

#### 2.6.1 Dark energy

The need for dark energy \[607,608\] may be considered as one of the major anomalies in modern cosmology, as there is no clear candidate to explain this phenomenon. There are many suggested solutions to this problem, from scalar fields, chaplygin gas, holographic dark energy, and modified gravity, cf. \[609\].

#### 2.6.2 Hubble trouble

A recently emerging anomaly has been the Hubble tension, where late-Universe and early-Universe measurements of the Hubble constant disagree by as much as \(3.46\sigma\) \[610\]. This tension emerges in the comparison of late-Universe measurements such as cosmic shear \[613,614\], weak lensing \[615,616\], galaxy clustering \[617\], and cepheid distance calibrations \[612\]. A whisker plot displaying a large selection of both early and late-Universe data from Ref. \[610\] is displayed in Fig. 8. There are many conjectured solutions to this problem, including models with an early dark energy component \[618,621\] with alternate models argued to be more natural, achieved via coupling to neutrinos \[622,626\] or via a \(k\)-essence scenario \[627\]. Additionally, a time-varying dark energy equation of state in the late-Universe has been argued for in Refs. \[628,629\]. A slew of additional models have been suggested in the literature that function by increasing the number of relativistic degrees of freedom at recombination, from sterile neutrinos \[630,632\], to thermal axions \[633\], to decaying dark matter \[634,636\] as well as several others \[610\]. Additionally, some authors have considered models where dark matter interacts with photons \[637,638\] or baryons \[639\]. Otherwise the effect of decaying ultra-light scalars has been considered in Ref. \[640\] and primordial black holes in Ref. \[641\]. A caveat on this topic is the sensitivity to systematics, as illustrated in Ref. \[642\], where the authors use a more robust distance calibration for nearby galaxies to demonstrate that late-Universe Hubble measurements, in this case via extinction in red giant branch stars.
Figure 8: Hubble tension whisker plot with references from Ref. [610]. The pink band represents Planck results [611] and the cyan band is for cepheids from Ref. [612].
within the Large Magellanic Cloud (LMC), can be largely reconciled with early-Universe measurements.

2.6.3 Clustering anomalies

A second emerging issue is the disagreement in the clustering properties of dark matter between early-Universe measurements from Planck [611] and late-Universe measurements via cosmic shear [643], as well as direct clustering measurements [644]. These newer measurements confirm earlier results on the same topic [645, 646]. It is as yet unclear whether this anomaly can be explained via systematics or requires additional physics. Some suggestions are that such data can probe dark energy models [647], or that the discrepancy can be ameliorated by dark matter decaying into dark radiation [648], or dark matter self interactions [649]. Finally, this anomaly has been used as a test of modified gravity scenarios [650], which indicated that general relativity was preferred over the studied forms of modified gravity.

2.7 Anomalies in ultra-high energy cosmic rays and neutrinos

The flux of cosmic ray particles impinging on the terrestrial atmosphere provides hadron-nucleus collisions at energies covering many orders of magnitude, including a regime well beyond the LHC’s energy. The interpretation of these observations relies on accurate models of air shower physics, which is a challenge and an opportunity to test QCD under extreme conditions.

2.7.1 The muon puzzle

High-energy cosmic rays can be observed via the extensive air showers in Earth’s atmosphere, which are hadronic cascades that eventually decay into muons. The muons reaching the detectors at ground-level are the key observable to infer the mass composition of cosmic rays. The muon puzzle constitutes an excess of muons at the ground level, measured in extensive air showers stemming from primaries with energies above 10 PeV. This excess is seen compared to state-of-the-art simulations, and has a significance of $8\sigma$ [651]. This muon deficit becomes apparent in cosmic ray interactions with centre-of-mass energies around the TeV scale, suggesting that the origin of this excess could be observable at the LHC.

2.7.2 Earth-emergent EeV events

At laboratory energies beyond PeV the mean free path of neutrinos becomes smaller than the Earth radius, which implies that beyond this energy, no particle should be able to penetrate the Earth. Upgoing showers are expected as a result of astrophysical tau neutrinos that convert to tau leptons while passing through the Earth, but the observed exit angles are restricted to be small for ultra-high energy neutrinos within the SM. Therefore the observations of a couple of Earth-emergent EeV-scale upgoing shower events by the ANITA experiment [652, 653] and the Extremely High Energy Northern Track neutrinos by IceCube [654] cannot be explained with astrophysical sources [655] nor within the SM [656]. These anomalies were addressed qualitatively in models involving long-lived particles, leptoquarks and heavy dark matter [119, 657, 659].
3 Hidden physics at the LHC

Section editors: Kingman Cheung, Rohini Godbole, Zhen Liu and Tao Han
Contributions: Shankha Banerjee, Kingman Cheung, Oliver Fischer and Zhen Liu

After the Run 2 at the LHC, ATLAS, CMS, and LHCb experiments have placed constraints on many models Beyond the Standard Model, either by direct searches for new particles or interactions, or through precision measurements. For example, searches for supersymmetric partners have pushed weak-scale SUSY to some uncomfortable corners in the parameter space [660–662]; searches for $W'$, $Z'$ bosons, and leptoquarks have placed limits of order $3 - 4$ TeV on their masses; searches for extra scalar bosons have restricted their masses to be heavier than $600 - 700$ GeV.

No convincing sign of NP has been detected in form of a resonance, however. This could mean that hypothetical new particles have masses above the LHC energy threshold or tiny production rates, and are therefore inaccessible by the experimental analyses. Here we consider the exciting possibility that NP is accessible at the LHC in principle, but that the corresponding signatures are hidden from experimental detection.

In this section we discuss the properties that the NP has to have in order to be hidden from current analyses. Last, we provide specific scenarios and models that are considered hidden in the data due to some of these properties.

3.1 Properties of hidden signatures

New particles with masses below the LHC energy threshold and non-negligible production rates may not have been covered by experimental searches, despite being testable in principle. Such new particles may not have been detected or studied for any combination of the following possible reasons:

- The triggers employed in Run 1 and 2 don’t respond to the final state;
- The new particle only decays hadronically and is buried in QCD backgrounds;
- The new particle is only produced in association with other particles;
- The mass of the new particle is very close to an SM particle.

Many models exist that are hidden from current searches for any of the above reasons, and it is not a problem to identify signatures and tailor dedicated searches for each of them individually. In the following we detail and discuss certain properties that render hypothetical NP signatures ‘hidden’ at the LHC.

3.1.1 Soft Particles in the Final State

Particles with transverse momentum below the trigger thresholds are called soft particles. An event that includes only soft particles will typically not be recorded at all. Therefore, NP with signatures that include only soft particles in the final state constitutes an important class of hidden NP. Studying this kind of event is extremely challenging as they are drowned in QCD backgrounds, which are the very reason for the triggers and their thresholds.

A generic scenario that can give rise to such signatures includes two or more particles that have almost degenerate masses, i.e. the mass spectrum is compressed. The production of the heavier new particle states, followed by decays to the lightest new particle state,
leaves little phase space for the SM particles that are radiated off in this process, such that their transverse momenta are suppressed. If the lightest new particle state is neutral, it will escape the detector, leaving a signature with missing energy and soft SM particles. Typical examples are SUSY scenarios with wino- or Higgsino-LSP, where the next-to-lightest SUSY particle is a chargino, which are discussed below.

A practical approach to studying such signatures is to include initial state radiation, e.g. a photon, jet, or other SM particles, see, e.g. Refs. [663, 664]. In the case of the lightest new particle being neutral, this leaves a mono-X signature, where X is the radiated-off SM particle. The disadvantages of this approach are the reduced signal cross-section, and the limited amount of information that can be used to identify the underlying model.

In the case where the next-to-lightest new particle is charged and has a macroscopic lifetime, its track could be detected. In this case, even if the decay leaves a soft pion or lepton plus a neutral particle that escapes detection, this decay could be detected via the disappearing track signature, cf. Ref. [665], as described in more detail below.

Neutral particles with macroscopic lifetimes may lead to signatures with soft displaced leptons or jets, which can be detected with dedicated triggers [666].

In some scenarios a spectrum may exist that allows for cascade decay of new particles into one another. An example is given by SUSY, where such cascade decays of gluinos or squarks can give rise to multi-jet and/or multi-lepton events. For instance, in the well-motivated scenarios of stop-bino coannihilation [667–672], the stop decays into soft jets plus LSP bino, one can rely on a hard ISR jet to trigger and select signal events [673].

3.1.2 SUSY without missing transverse energy

Because of $R$-parity conservation, SUSY particles are usually considered to be produced in pairs, and to promptly decay via chains of varying lengths into the LSP, radiating off SM particles. Therefore the traditional SUSY analyses search for the missing transverse energy in an event containing a number of SM particles.

Various SUSY scenarios exist, however, where this signature does not apply. One possible reason is a compressed mass spectrum, which may preclude the event to be registered at all, as described in the previous subsection. Another plausible reason exists for scenarios where the LSP is not stable, e.g. in SUSY scenarios where $R$-parity conservation is violated. In this case, provided the LSP decays inside the detector, no missing transverse energy is generated from the escaping LSP, and the relevant signature is instead given by the LSP decay products.

3.1.3 Associated production processes

New particles that are singly or doubly produced usually give rise to a well understood signature in the detector and can be discovered via a resonance in a given final state. On the other hand, processes including new particles that are produced in association with additional SM particles could skip the selection filters in experimental analyses. In such a scenario, a statistically significant resonance may be hidden in the data but invisible to the analyses, e.g. because its associated production leads to particles in the final state against which analyses veto.

This situation is theoretically well motivated, for instance in models where the new particles couple to the SM only after mixing of gauge bosons. This case, where BSM particles are produced in association with $\gamma, W, Z$ gauge bosons has been studied for quite a number of exotic particles beyond the SM, e.g. for heavy scalar bosons or axions.
Associated production with a Higgs boson as a unique window to BSM physics is well motivated, for instance via the Higgs-portal models. Especially the Higgs boson coupling to a dark sector, including DM candidates, can give rise to the mono-Higgs signature. Furthermore, the heavy Higgs boson(s) in many BSM scenarios could provide new production avenues to discover new particles. For instance, the charged Higgs boson produced in association with top quark and bottom quark can decay into tau slepton and sneutrino, making a discovery possible [674].

Associated production with a single top quark, a top quark pair, or the Higgs boson pose further challenges to experimental searches, as the decay products from the heavy SM particles significantly modify the final state, thus obviating present searches for this type of NP. The heavy SM particles may therefore be a relevant avenue to BSM searches, despite the associated NP production rates often being suppressed.

One interesting example is given by models where the dark matter couples only to the top quark [675, 676] such that it can escape all the existing direct and indirect searches of dark matter. The signature at the LHC would be a top-quark pair plus large missing energies.

### 3.1.4 QCD-like final states

Scenarios exist where new particles decay exclusively into QCD final states, i.e. quarks and/or gluons. Such scenarios are motivated, for instance, with $R$-parity violating (RPV) SUSY models [677] or baryonic $Z'$ [678]. Hadronisation of such a final state degrades much of its kinematic structure. Even resonances in invariant mass spectra can be entirely washed out. Combined with the huge QCD backgrounds and pile up this type of signal is notoriously difficult to study at the LHC.

Searching for large missing energy in this kind of signatures is not very useful, and the possibility to reconstruct a resonance depends crucially on the jet resolution and reconstruction, and is sensitive only to large signal rates. As traditional analyses are not very effective, other means of detection are necessary. For instance recent developments in jet substructure can improve identification of fast moving heavy resonances [679]. In general, machine learning is a very promising avenue to studying this kind of NP signature.

### 3.1.5 Close-by SM resonances

To avoid contamination from SM resonances, experimental analyses often exclude signal regions that are close to the masses of these known particles. Therefore, new particles with masses that are very close to known SM particles could have escaped detection up to now, provided that their decay channels are overlapping, and their signal rates and widths would not leave a visible imprint.

So far, we have not encountered such scenarios in collider searches. Nevertheless, we cannot easily exclude the possibility that new particles are hidden around the $Z$ boson peak [680] or Higgs boson peak [681]. While all decay modes of the $Z$ boson are tested thoroughly with high precision, those of the Higgs boson still leave plenty of room for the existence of a new particle with a very similar mass and with one or more decay channels the same as the Higgs boson. One then has to carefully scrutinise the shape and the height of the Higgs boson resonance and separate it from the interference effects [682].

High-dimensional models can give rise to a non-trivial spectral density of a new particle, which in turn affects its signature in the detector. The so-called HEIDI models [683] allow single or double peaks in the invariant mass spectrum, or a continuum of states, which require very high precision measurements in order to be studied.

41
3.2 Hidden SUSY scenarios

The general minimal supersymmetric standard model (MSSM) has more than 100 soft SUSY-breaking parameters. Some GUT-motivated scenarios have greatly reduced the number of parameters, e.g., mSUGRA, and thus a lot of experimental searches were based on such simplified scenarios. Conventional searches for SUSY rely on the final states with multi-leptons and/or multi-jets plus large missing energies. The current searches have pushed these scenarios restrictively to a small part of the parameter space. Nevertheless, there are some scenarios or SUSY breaking models that are still hidden from detection. We are going to describe a few of these scenarios.

3.2.1 Wino and Higgsino LSP

In wino LSP and Higgsino LSP scenarios [664,666,684], the lightest chargino and neutralino are degenerate or very close in mass, such that the decay of the chargino into the lightest neutralino plus the charged lepton or pion would be very soft. Nevertheless, some higher order corrections may be able to lift the degeneracy to some extent [685, 686]. Also in the Compressed SUSY scenario [687], all SUSY partners are very close in mass. In these examples, the decay products are usually too soft to pass detection thresholds, which renders this subclass of SUSY models hidden from detection.

3.2.2 Stealth SUSY

In this type of scenarios, the SUSY particles have weak-scale masses that feel SUSY breaking only through couplings to the MSSM, which theoretically motivates the small mass splitting between fermion/boson pairs.

The resulting mass spectrum is compressed, which leaves very little phase space for the missing transverse energy after the decay of the SUSY partner into its SM counterpart [688], and the conventional strategy of searching for large missing transverse energies is not effective. Experimental collaborations instead search for multi-jets, multi-leptons, and/or multi-photons in the final state, which are expected to be soft [689].

3.2.3 R-parity violation

In so-called R-parity violating (RPV) SUSY [677] the R-parity is not conserved, which leads to the LSP not being stable. Depending on the magnitude of the RPV couplings, the LSP decays can be prompt or long-lived. In most cases, there is no missing energy in the final state and thus the conventional searches for SUSY fail. In addition, if only $U^c D^c D^c$ RPV couplings exist, the decays are totally hadronic and in addition the signature is buried in QCD background.

3.2.4 Long-lived NLSP

In some SUSY scenarios the next-to-lightest SUSY particle (the NLSP) can have a suppressed decay rate in to the LSP plus some SM particles, which renders its lifetime macroscopic. Examples are gauge-mediated SUSY breaking (GMSB) models [690,692], where the LSP is given by the gravitino. The decay of the NLSP, either a neutralino or a stau, can take a long time. The current triggers for SUSY may entirely miss such a scenario. In GMSB models,
the gravitino mass is given by

\[ m_G = \frac{F}{\sqrt{3}M_p} \simeq 2.5 \left( \frac{F}{(100\text{TeV})^2} \right) \text{eV}, \]  

(31)

where \( F \) is the SUSY breaking scale and \( M_p \) is the reduced Planck mass. If \( F \) is of order \( 10 - 1000 \text{ TeV} \), the gravitino is the LSP. Also, the interaction of the gravitino with the NLSP is suppressed by the scale \( F \) such that the NLSP has a long decay into the LSP.

In the case that the lightest neutralino is the NLSP, its decay width into \( \gamma G \) is

\[ \Gamma(\tilde{\chi}_1^0 \rightarrow \gamma G) = |\cos \theta_W N_{11} + \sin \theta_W N_{12}|^2 \frac{m_{\tilde{\chi}_1^0}}{16\pi F^2}. \]  

(32)

On the other hand, if the stau is the NLSP, its decay width is given by

\[ \Gamma(\tilde{\tau}_1 \rightarrow \tau G) = \frac{m_{\tilde{\tau}_1}}{16\pi F^2}. \]  

(33)

In the former case, if the decay length of the lightest neutralino is too long, the signature would be the same as the conventional missing energy search. Otherwise, one can search for non-pointing photons in the final state. In the slepton-NLSP case, the stau carries charge and leaves visible tracks in the inner tracker, which is much easier for detection.

In the scenario of gluino LSP [694,695], the gluinos will be copiously produced by QCD interactions. Subsequently, they will hadronise into \( R \)-hadrons, either electrically neutral or charged, or changing charges through nuclear interaction with detector material. When charged, the \( R \)-hadrons could be detected as stable charged particles, but when neutral, it would be difficult to detect it, because the energy loss in collisions with detector would be small.

### 3.2.5 \( \nu \)CMSSM with a long-lived stau

In Ref. [696] the constrained minimal SUSY (CMSSM) was extended with right-handed superpartners, a scenario that we call the \( \nu \)CMSSM. As we are well aware, the CMSSM is strongly affected by the LHC direct searches, the Higgs boson mass constraint as well as from other dark matter experiments. The evidence of neutrino masses ensuing from neutrino oscillations requires the extension of the SM with at least right-handed neutrinos with a Dirac mass term. Thus, in case of the CMSSM, we extend the theory with right-handed sneutrinos. Here, we consider that our next-to-lightest SUSY particle (NLSP) is the \( \tilde{\tau} \). Even with this minimal extension of the CMSSM, we get striking signatures of heavy charged metastable particles. We consider several bounds. The most important bounds come from the neutrino mass and from the big bang nucleosynthesis (BBN).

To be more specific, we extend the MSSM potential by just a single term, for each family

\[ W^R_\nu = y_\nu \hat{H}_u \hat{\nu}_R, \]  

(34)

where \( y_\nu \) is the neutrino Yukawa coupling, the left-handed lepton superfield, \( \hat{L} = (\hat{\nu}_L, \hat{\ell}_L) \), and \( \hat{H}_u = (\hat{H}_u^+, \hat{H}_u^0) \) is the Higgs superfield. This gives rise to the masses of the \( T_3 = +1/2 \) fermions. Finally, the superfield for the right handed neutrinos is \( \hat{\nu}_R \). From the global fits of the neutrino oscillation parameters from solar, atmospheric, reactor and accelerator neutrino
data and from the combination of the Planck temperature and polarisation data, we obtain
the following bound
\[ y_H^u \sin \beta \subset [2.8, 4.4] \times 10^{-13}, \quad (35) \]
where \( \tan \beta = \langle H_u^0 \rangle / \langle H_d^0 \rangle \). Furthermore, if we neglect any inter-family mixing, then the additional mass term for the sneutrinos can be written as
\[ -L_{\text{soft}} \supset M^2_{\tilde{\nu}} |\tilde{\nu}_R|^2 + (y_\nu A_\nu H_u \tilde{L} \tilde{\nu}_R^c + \text{h.c.}), \quad (36) \]
where \( A_\nu \) is responsible for the left-right mixing in the scalar mass matrix. The left-right mixing in the sneutrino sector can be written as
\[ \tan 2\tilde{\Theta} = \frac{y_\nu v \sin \beta |\cot \beta - A_\nu|}{m^2_{\tilde{\nu}_L} - m^2_{\tilde{\nu}_R}}, \quad (37) \]
The mass eigenstates are
\[ m^2_{\tilde{\nu}_L} = M^2_{\tilde{\nu}} + \frac{1}{2} m^2_Z \cos 2\beta \quad \text{and} \quad m^2_{\tilde{\nu}_R} = M^2_{\tilde{\nu}} , \quad (38) \]
where \( M_{\tilde{\nu}} \) is the soft scalar mass for the left-handed (right-handed) sleptons (neutrinos).

The \( \tilde{\tau}_1 \) finally decays into the right-handed sneutrinos via \( \tilde{\tau} \to W^{(*)} \tilde{\nu}_R \) and the two body decay width (assuming \( m_{\tilde{\tau}} > m_{\tilde{\nu}_R} + m_W \)) can be written as
\[ \Gamma_{\tilde{\tau}} \simeq \Gamma_{\tilde{\tau} \to \tilde{\nu}_R W} = \frac{g^2 \tilde{\Theta}^2}{32\pi} |U^{(\tilde{\tau})}_{L1}|^2 \frac{m^2_{\tilde{\tau}}}{m^2_W} \left[ 1 - \frac{2(m^2_{\tilde{\nu}_R} + m^2_W)}{m^2_{\tilde{\tau}}} + \frac{(m^2_{\tilde{\nu}_R} - m^2_W)^2}{m^4_{\tilde{\tau}}} \right]^{3/2} , \quad (39) \]
where \( g \) is the \( SU(2)_L \) coupling, \( m_W \) is \( W \)-boson mass and \( U^{(\tilde{\tau})} \) is the mixing matrix of the staus \( (m_{\tilde{\tau}} \leq m_{\tilde{\nu}_L}) \), which relate the mass and the gauge eigenstates as
\[ \begin{pmatrix} \tilde{\tau}_L \\ \tilde{\tau}_R \end{pmatrix} = U^{(\tilde{\tau})} \begin{pmatrix} \tilde{\tau}_1 \\ \tilde{\tau}_2 \end{pmatrix} , \quad (40) \]

\( L1 \) indicates the \((1,1)^{th}\) element of this matrix. When the two body decays are forbidden, the dominant three body decays are \( \tilde{\tau} \to \tilde{\nu}_R \bar{\nu}_R \tilde{\nu}_R q^q \). The stau lifetime depends on the decay modes and on the mixing in the stau and the sneutrino sectors. Typical lifetimes vary between a few second to up to \( 10^{11} \) seconds.

The NLSP’s lifetime is not long enough to ensure that its decay occurs well after its freezeout. However, \( \tilde{\nu}_R \) contains all good properties of cold dark matter. It is stable due to \( R \)-parity conservation and because it evades direct detection constraints owing to suppressed interactions due to tiny Yukawa coupling. The density parameter of \( \tilde{\nu}_R \) can be written as
\[ \Omega_{\tilde{\nu}_R} = \frac{m_{\tilde{\nu}_R}}{m_\chi} \Omega_\chi , \quad (41) \]
where \( \Omega_\chi \) is the present density parameter of the \( \tilde{\tau}_1 \) NLSP, assuming it to be stable.

Below, in Fig. 9 we show the allowed parameter space for two different Yukawa values assuming at least 10% relic contribution. We also show the allowed parameter region abiding all other collider and cosmological constraints in Fig. 10. We explicitly show the constraints ensuing from the BBN in Fig. 11 where the visible energy is \( E_{\text{vis}} = \frac{m^2_{\tilde{\tau}} + m^2_W - m^2_{\tilde{\nu}_R}}{2m_{\tilde{\tau}}} \).
$B_{\text{had}} = 2/3$ which corresponds to the hadronic branching ratio of the $\tilde{\tau}_1$ for two body decays. Lastly, $Y_{\text{NLSP}}$ is the ratio of the number density to the entropy density at the $\tilde{\tau}_1$ freeze-out.

Before concluding this section, we wish to mention the LHC prospects of this model. We studied the potential of the following channels:

- $2 \tilde{\tau} + N$ hard jets ($N \geq 2$),
- $2 \tilde{\tau}$ (two stable charged tracks),
- passive detection of highly-ionising (slow) particles.

For this, we consider BPs following the trend

$$m_{\tilde{\nu}_R} < m_{\tilde{\tau}} < m_{\tilde{\chi}^0_1} < m_{\tilde{\epsilon}_1, \tilde{\mu}_1} < \ldots < m_{\tilde{g}}$$

The "stable" $\tilde{\tau}_1$ will behave like a slow muon, with velocity $\beta = p/E$ much less than 1 (as can be seen in Fig. 12).

The $2 \tilde{\tau} + N$ hard jets ($N \geq 2$), comes with $E_T$ and is of no concern in this section. The category with 2 stable charged tracks is something that is worth saying a few words in.
Figure 10: Allowed parameter region in the $m_0 - m_{1/2}$ plane satisfying existing collider, low energy, relic and BBN constraints for the ‘hierarchical’ (green) and degenerate (red) neutrino masses. Here, $m_{0,1/2} < 2500 \text{ GeV}$, $|A_0| < 3000 \text{ GeV}$, $5 < \tan \beta < 40$, $0 < m_{\tilde{\nu}_R} < m_0$ and $\text{sign}(\mu) > 0$.

this section. The most dominant final state is with two muons. For the four BPs listed in Ref. [696] we list the results in Table 5.

3.3 Hidden portal models

Portal models consider NP that interacts with the SM via one specific (class of) SM particles, namely the neutrinos, the Higgs boson, and the vector bosons. Usually, also pseudoscalar particles are considered, and often new fermions are introduced. This approach allows for an effective categorisation of observed signatures and for some degree of model independence. For a recent discussion, see for instance Ref. [697]. Below we highlight the portals’ properties that are relevant here.

3.3.1 The Neutrino portal

The observation of neutrino oscillation has firmly established the non-vanishing masses of the active neutrinos. New physics is required to invoke certain mechanisms for neutrino mass generation. One of the most celebrated theoretical means is the so-called type I seesaw
Figure 11: Allowed parameter region in the lifetime-injected hadronic energy plane which satisfies every existing constraint for the 'hierarchical' (green) and degenerate (red) neutrino masses. The two curves are for the constraint from $^4$He (magenta dashed) and $^2$H/H (cyan solid) abundance. The dotted (blue) curve denotes the impact of assuming a tightened $^2$H/H determination.

The mechanism, which generates small neutrino mass by introducing right-handed neutrinos, which couple to the active neutrino and the Higgs fields through a Dirac mass term, plus a Majorana mass term, and can be described in a simple way as

$$\mathcal{L}_Y = -Y_D \bar{L} H N - M_N \bar{N} N + \text{h.c.},$$

(42)

where $H$ is the SM Higgs field, $N$ are the right-handed neutrinos with Majorana mass $M_N$ and Yukawa coupling $Y_D$. In the type I seesaw approximation the small neutrino mass is given by $M_D^2/M_N$, where $M_D = Y_D \langle H \rangle$ is the Dirac mass term.

The right-handed and active neutrinos mix, thus creating a number of light and heavy mass eigenstates, with dominant active and sterile neutrino components, respectively. The smallness of the light neutrinos' masses can be achieved either by a very large Majorana mass of order $10^{11-13}$ GeV, by a very small Yukawa coupling $Y < 10^{-5}$, or by invoking additional symmetries [698].

These kinds of models are often called 'heavy neutral leptons', 'sterile neutrinos', or 'neutrino-portal models', which have been discussed often and in many different settings, cf. [47].
Figure 12: Velocity distribution of the $\tilde{\tau}$-NLSP for BP3.

<table>
<thead>
<tr>
<th>Cut set</th>
<th>Benchmark point</th>
<th>$N_S$</th>
<th>$N_B$</th>
<th>$N_S/N_B$</th>
<th>$\mathcal{S}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>BP1</td>
<td>526</td>
<td>5684</td>
<td>0.09</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>BP2</td>
<td>358</td>
<td>12772</td>
<td>0.06</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>BP3</td>
<td>258</td>
<td>12772</td>
<td>0.05</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>BP4</td>
<td>47</td>
<td>12772</td>
<td>0.01</td>
<td>0.6</td>
</tr>
<tr>
<td>B</td>
<td>BP1</td>
<td>1337</td>
<td>12772</td>
<td>0.10</td>
<td>11.3</td>
</tr>
<tr>
<td></td>
<td>BP2</td>
<td>1069</td>
<td>12772</td>
<td>0.08</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>BP3</td>
<td>826</td>
<td>12772</td>
<td>0.06</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>BP4</td>
<td>232</td>
<td>12772</td>
<td>0.02</td>
<td>2.0</td>
</tr>
<tr>
<td>C</td>
<td>BP1</td>
<td>1543</td>
<td>3481</td>
<td>0.44</td>
<td>21.8</td>
</tr>
<tr>
<td></td>
<td>BP2</td>
<td>1014</td>
<td>3481</td>
<td>0.29</td>
<td>15.1</td>
</tr>
<tr>
<td></td>
<td>BP3</td>
<td>715</td>
<td>3481</td>
<td>0.21</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>BP4</td>
<td>211</td>
<td>3481</td>
<td>0.06</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Table 5: Table shows the number of signal and background events after the selection cuts (Table 6), the ratio $N_S/N_B$ and the statistical significance. $\mathcal{L} = 3000 \text{ fb}^{-1}$.

e.g. Refs. [699–704].

At the LHC, heavy neutrinos are produced predominantly in Drell Yan processes, and they have been searched for traditionally via lepton number violating signatures, which are striking because of absent SM backgrounds [705]. Symmetry protected models allow for in principle large production cross-sections, but even their most promising signatures are not very effective due to the signal rates being suppressed by other constraints, and the ubiquitous backgrounds [706]. Heavy neutrinos that are lighter than the $W$ boson have long lifetimes. They may give rise to displaced vertex signatures in any of the detector
Table 6: Three sets of selection cuts applied in the \( \tilde{\tau} \) pair analysis.

<table>
<thead>
<tr>
<th>Cut on</th>
<th>Cut set A</th>
<th>Cut set B</th>
<th>Cut set C</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>&gt; 0.85</td>
<td>–</td>
<td>&lt; 0.95</td>
</tr>
<tr>
<td>( p_T^{\mu_1} )</td>
<td>&gt; 200 GeV</td>
<td>&gt; 200 GeV</td>
<td>&gt; 70 GeV</td>
</tr>
<tr>
<td>( \sum</td>
<td>p_T^{\text{vis.}}</td>
<td>)</td>
<td>&gt; 700 GeV</td>
</tr>
<tr>
<td>(</td>
<td>y(\mu_{1,2})</td>
<td>)</td>
<td>&lt; 2.4</td>
</tr>
<tr>
<td>( M_{\mu_1,\mu_2} )</td>
<td>&gt; 1200 GeV</td>
<td>&gt; 1000 GeV</td>
<td>–</td>
</tr>
<tr>
<td>( \Delta R(\mu_1,\mu_2) )</td>
<td>&gt; 0.2</td>
<td>&gt; 0.2</td>
<td>–</td>
</tr>
<tr>
<td>( \Delta R(\mu,j) )</td>
<td>&gt; 0.4</td>
<td>&gt; 0.4</td>
<td>–</td>
</tr>
<tr>
<td>( \Delta R(j,j) )</td>
<td>&gt; 0.4</td>
<td>&gt; 0.4</td>
<td>–</td>
</tr>
</tbody>
</table>

components thus be hidden from the current searches, which will be discussed below.

### 3.3.2 The Higgs portal

Here we describe a simple Higgs-portal model with an additional real SM-singlet scalar field \( X \) that mixes with the SM Higgs doublet field \( \Phi \) in the presence of a new \( Z_2 \) symmetry. The new scalar field \( X \) is odd under the \( Z_2 \) such that no \( X \) or \( X^3 \) terms appear, while all the SM fields are even. The Lagrangian is given by

\[
L = \frac{1}{2} \partial_\mu X \partial^\mu X + \frac{1}{2} \mu_X^2 X^2 - \frac{1}{4} \lambda_X X^4 - \frac{1}{2} \lambda_{\Phi X} (\Phi^\dagger \Phi) X^2 + L_{\text{SM}}.
\]

After electroweak symmetry breaking both the SM Higgs doublet field \( \Phi \) and the new scalar singlet field \( X \) are expanded around their vacuum-expectation values \( \langle \phi \rangle \approx 246 \text{ GeV} \) and \( \langle \chi \rangle \):

\[
\Phi(x) = \frac{1}{\sqrt{2}} \left( \begin{array}{c} 0 \\ \langle \phi \rangle + \phi(x) \end{array} \right),
\]

\[
X(x) = \langle \chi \rangle + \chi(x).
\]

The mass matrix of the two scalar fields is

\[
L_m = -\frac{1}{2} (\phi \chi) \left( \begin{array}{cc} 2\lambda \langle \phi \rangle^2 & \lambda_{\Phi X} \langle \phi \rangle \langle \chi \rangle \\ \lambda_{\Phi X} \langle \phi \rangle \langle \chi \rangle & 2\lambda_X \langle \chi \rangle^2 \end{array} \right) \left( \begin{array}{c} \phi \\ \chi \end{array} \right).
\]

It is possible to rotate \( (\phi \chi)^T \) to \( (h_h)^T \) through an angle \( \theta \)

\[
\left( \begin{array}{c} h \\ h_s \end{array} \right) = \left( \begin{array}{cc} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{array} \right) \left( \begin{array}{c} \phi \\ \chi \end{array} \right),
\]

where \( h \) is the scalar Higgs boson observed at 125 GeV while \( h_s \) is the new scalar boson of the model. The mixing angle \( \theta \) is constrained to be very small due to various experimental datasets including the Higgs boson strength data \[707\]. The new scalar boson \( h_s \) originating from a hidden sector can decay back into SM particles via the mixing angle, which is suppressed by \( \sin^2 \theta \). For example, the partial width into a pair of SM fermions is given by

\[
\Gamma(h_s \rightarrow f \bar{f}) = N_f \sin^2 \theta \frac{m_f^2 m_{h_s}}{8\pi \langle \phi \rangle^2} \left( 1 - \frac{4m_f^2}{m_{h_s}^2} \right)^{3/2},
\]

49
where $N_f$ is the colour of the fermion. The total decay width of $h_s$ can be obtained by summing over all kinematically allowed fermion pairs. For small enough $\theta$ and light $h_s$ the lifetime of $h_s$ can travel a macroscopic distance before decay, which may complicate its detection.

To detect a new physical scalar through mixing with the Higgs boson at the LHC is more complicated than one would think. Depending on the mass of the new particle, most of its decays leave hadronic final states and have to contend with towering backgrounds involving top quarks and multiple vector boson production, cf. e.g. Ref. [708].

The four-lepton final state is often referred to as the ‘golden channel’ due to small and controllable SM backgrounds, and used by the ATLAS [709] and CMS [710] collaborations to search for heavy scalars. However, the tiny total cross-section of this process is suppressed further by the necessarily small scalar mixing and by additional decay channels, such as di-top, or Higgs $Z$, and possible decays involving (invisible) new particles.

### 3.3.3 The vector portal

Generic BSM theories often contain additional gauge symmetries, especially $U(1)'$ symmetries. Additional spin-1 or vector particles commonly arise from the breakdown of a larger gauge symmetry factor, or when the vector is a composite state. Such vector particles may have interactions with SM particles, and also with possible new particles. In addition, these vectors may mix kinetically with the $U(1)$ factor in the SM gauge group. For example, after the electroweak symmetry breaking and diagonalisation of the gauge kinetic terms, a dark photon ($A'$) theory may have the following Lagrangian:

$$L_{A'} \supset -\frac{1}{4} F'_{\mu\nu} F'^{\mu\nu} + \frac{1}{2} m^2_{A'} A'_\mu A'^\mu + \epsilon \epsilon A'_\nu J^\nu_{\mathrm{EM}},$$

(49)

where $m_{A'}$ is the dark photon mass, $\epsilon$ the small kinetic mixing parameter, $J^\mu_{\mathrm{EM}}$ the standard model electromagnetic current, and $F'_{\mu\nu}$ the standard field strength operator for $A'$. While $A'$ could have more interactions with the SM particles, the Lagrangian above can be considered the minimal dark photon scenario for phenomenological purposes.

In this case, the partial width of $A'$ to SM fermions follows:

$$\Gamma_{A' \to f\bar{f}} = \frac{\epsilon^2 \alpha_{\mathrm{EM}}^2}{3} m_{A'} \left( 1 + \frac{2m_f^2}{m^2_{A'}} \right) \sqrt{1 - \frac{4m_f^2}{m^2_{A'}}}.$$  

(50)

Here $\kappa \equiv 3Q^2$ for SM quarks with charge $Q$ and $\kappa \equiv 1$ for SM charged leptons. The kinetic mixing parameter $\epsilon$ is expected and constrained to be tiny.

Dark photons can be produced at the LHC via their mixing with the SM photon. The strong experimental limits on the mixing parameter $\epsilon$ render the production rates tiny, and thus difficult to test. The limits on the mixing also necessitate a small decay rate, and hence enforce a long lifetime for the dark photon, further complicating the discovery of its signatures.

### 3.4 Long-Lived Particles

The negative results at the LHC have raised the question of whether there is a systemic shortcoming in detecting new physics. Indeed, one of the possibilities is that new physics may manifest itself in the form of long-lived particles (LLP), which might have escaped from
detection in the current design of the experimental triggers or due to the size of detectors or negligible interactions with the detectors.

This particular class of NP models raised strong interest in the HEP community. Experimentalists and theorists are working hard to overcome the challenge of detecting LLP signatures, which include disappearing tracks, emerging jets and leptons, kinks in tracks and which depend strongly on the explicit considered model and particle content.

3.4.1 The experimental view

In recent years, there are rising interests in searches for LLPs in both theoretical and experimental communities to search for LLP signatures at the LHC which is, however, a very challenging task. Current hardware and software triggers and analyses have been focused on promptly decaying new particles, such as squarks or gluinos in supersymmetry, top partners in composite models, and leptoquarks in GUT models.

Efforts by the collaborations in recent years led to tremendous progress to cover a large variety of signatures from long lived particles, see e.g. Ref. [711]. The LHC collaborations have developed a broad programme of LLP searches [712], which grew from discussions with theorists, cf. e.g. Refs. [712–714].

It is worthwhile to comment on the scope of the programme for LLP searches, and its future perspective. Improved triggers and new components in the detector have been suggested, discussed, and implemented at the ATLAS and CMS experiments to accommodate the searches for LLPs [715–719]. New search strategies have been suggested and discussed, and will be implemented already in the next run [712, 714, 716, 718, 720]. Specific experiments are planned, under construction, or being commissioned, to add complementary search capacity to the big LHC collaborations. Examples are MilliQan [721], FASER [722], MoeDAL [723], and MATHUSLA [720].

3.4.2 Signatures

Particles with long lifetimes give rise to distinct signatures in the LHC detectors, which crucially depend on their electric charge, lifetime, velocity, and interactions with the detector. The search for the LLPs at the LHC can make use of the tracker detector, calorimeters, muon spectrometer, and/or the new timing detector, depending on the decay length and decay products of the LLP.

Charged LLPs (cLLP) interact with the detector components, and in particular they leave a track in the tracker detector, possibly with a very characteristic ionisation signature. A cLLP’s signature can be very different, depending on its lifetime: In case of the decay taking place outside the detector, one has a muon-like ionised track throughout all the detector components; When a cLLP decays into a number of soft final states it can give rise to a disappearing track, provided the charged daughter particle is missed; A cLLP that decays into one charged daughter and one (or more) neutral particles can give rise to a charged track with a kink.

Neutral LLP can be observed when they decay, possibly via a chain, into charged particles inside the detector. If the decay takes place in the tracker, i.e. if the tracks of the charged daughter particles can be reconstructed, it is in principle possible to also reconstruct the point of decay, the so-called displaced vertex.

In general, displaced decays make misinterpretation of the daughter particles possible, in particular if the LLP decay takes place outside the tracker, and only partial detector information on the daughter particles (in particular their charges) is available. Then the
reconstruction of the displaced vertex is not always possible but, depending on the LLP mass, the decay products could be delayed, which would make them observable via the time delay feature, as proposed in Ref. [716].

3.4.3 Theoretical motivation

Lifetimes are inversely proportional to the total decay rate of the decaying particle, and can therefore become larger when the coupling constant(s) are tiny and/or when the phase space is suppressed. In SUSY models with a very small $R$-parity-violating (RPV) coupling, the lightest supersymmetric particle (LSP) can travel a macroscopic distance before decay, thus giving rise to LLP signatures. One can identify displaced vertices [713, 724, 725] in the tracker detector or make use of the timing detectors to detect less relativistic and heavy LLPs [716, 726]. In gauge-mediated SUSY breaking models, the photino-NLSP case is the more difficult scenario. It decays into a non-pointing photon and missing energy. One has to rely on the EM calorimeter to determine the displaced vertex. If the decay falls outside the EM calorimeter, the event would be easily lost.

In the Higgs-portal models hadronic LLP signatures can generically be produced [716, 718, 727, 729]. If $m_{h_s} \sim O(1)$ GeV the $h_s$ can decay into a pair of muons or pions with a displaced vertex [707]. Such muon pairs appear as energetic collimated muons (so-called “muon-jets”), and collimated pion pairs appear as so-called “fat-jets”. One can also make use of muons and jets with a displaced vertex detected in the inner tracker or the muon spectrometer [730, 731].

Other light portals have similar LLP signatures but with different kinematic distributions. For instance, searches for Axion-Like-Particles (ALPs) have posed a great challenge at the LHC due to its low production rate and large hadronic background. It is possible to cover new ground for ALPs through a displaced vertex search [732], or through new production and decay modes that are beyond the minimal model [733]. Note that LHCb has certain advantage in searching for low mass prompt and displaced dilepton resonances [732, 735], allowing for complementary coverage in the low mass regime with respect to the ATLAS and CMS experiments. In models of sterile neutrinos, once they are produced, they can travel a macroscopic distance before they decay into leptons and/or hadrons [703, 736], thus giving rise to displaced vertex or emergent leptons or jets. The signals can be detected at the tracker and/or calorimeters.

In addition, the hidden strong dynamics [737–742] have gained increasing attention due to the distinct feature of dark showers. Such dark shower signatures require close examination of the experimental capabilities and they are under active development.

Complementing the searches with the existing detectors and upgrades, new peripheral experiments and low energy particle physics experiments are being evaluated and under construction to look for these exotic long-lived signatures [743].

3.4.4 Backward moving objects from TeV LLP

Of particular interest is a less-studied signature of backward moving objects (BMOs) [744]. While we are more used to studying fast moving particles whose decay products are collimated along the direction of the parent, for slow moving particles, such decay products have a wider distribution. With a few examples, we show that for heavy LLP (around TeV mass scale) searches at the LHC, the particles ensuing from the secondary vertex can be at large angular separations with respect to the direction of motion of the LLP. A fraction of such particles can even go in the backward direction, giving rise to striking signatures as these
particles traverse the various layers of the detector outside-in, towards the direction of the beam pipe. This can be translated to the energy deposited in the tracker. The particles can come from as far as the hadron calorimeter (HCAL) or even ones that can come from outside the detector and into the muon chamber. We see that the most prominent effect comes from LLPs which come to rest inside the detector, where the example being studied is the $R$-hadrons. We also see similar results when the LLPs are lighter than the TeV scale or when some of the available energy is carried away by a massive invisible daughter. The four benchmarks studied are the following:

- $X \rightarrow qg$, where $X$ is the LLP and $q$ is a massless quark. Such decays can be seen in $R$-parity violating (RPV) supersymmetric models. We can have processes like $\tilde{q} \rightarrow qg$ or $\tilde{l} \rightarrow qg$. We classify this channel as $2BM0$.

- $X \rightarrow qqq$: This is another example of a RPV process where an example process is $\tilde{\chi}^0_1 \rightarrow qqq$. We call this category as $3BM0$.

- $X \rightarrow qDM$, where $DM$ is a heavy invisible daughter. Such processes are possible through $R$-parity conserving scenarios with channels such as $\tilde{q} \rightarrow q\chi^0_1$ or $\tilde{g} \rightarrow g\chi^0_1$, where $\chi^0_1$ is the lightest neutralino. We categorise this channel as $2BM$.

- $X \rightarrow qqDM$: This can also come about in an $R$-parity conserving scenario and the following three body decay encapsulates such a process, $\tilde{g} \rightarrow q\bar{q}\chi^0_1$. We call this the $3BM$ category.

In Fig. 13, we show the angle $\theta$ that 1 or DM makes with the direction of the mother LLP, for several benchmark points (BPs). We choose various values for the LLP, $M_X$ and the mass of the invisible particle, $M_{DM}$. We can clearly see that for many of these BPs, the fraction of (light) particles going in the backward direction, is significant.

We now consider two scenarios to show the energy fraction that traverses back. In the first scenario we consider particles that can decay within the HCAL and traverse towards the tracker. We perform our analysis by assuming a very simple geometry with $L_{\text{tracker}} = 600$ cm along the $z$-axis, $R_{\text{tracker}} = 100$ cm and the transverse distance of the last layer of the HCAL to be 300 cm from the $z$-axis. We show the energy fraction $E_{\text{in}}/E_{\text{LLP}}$ provided the LLP decays somewhere between 100 cm and 300 cm in the transverse direction, somewhere between the tracker and the HCAL. Figure 14 shows this energy fraction for the 4 categories. For the $2BM0$ and $3BM0$ scenarios, the observations are striking. For the $2BM0$ case, the fraction of energy coming back inside the tracker for the massless two-body decay scenario are 25.9% for the stopped $R$-hadrons and 12.2% for the moving LLP. For the $3BM0$ scenario, these numbers respectively become 34.2% and 14.2%.

In a similar vein we consider the scenario when a particle decays just outside the muon chamber and comes back inside. For this, we again follow a simplified geometry with $R_{\text{muon-chamber}} = 750$ cm and $L_{\text{muon-chamber}} = 1300$ cm along the $z$-direction. The energy fractions, at least for the massless cases, are similar to the one that we discussed for the tracker scenario.

Before concluding this part, we will briefly touch upon the experimental considerations as what we have discussed until now is mostly at the parton-level analysis. We have not discussed the backgrounds but the dominant backgrounds would ensue from cosmic ray events. One way of suppressing such backgrounds is to tag the backward moving objects only in the lower half of the detector. Moreover, there can be backgrounds ensuing from beam-induced noise, overlapping events as well as instrumental noise. Furthermore, shower shapes for the
Figure 13: Angle $\theta$ between the direction of $X$ and one of the quarks, $q$ or the massive daughter (DM).
Figure 14: Normalised distribution of $E_{\text{in}}/E_{\text{LLP}}$: $M_X = 2$ TeV, $M_{DM} = 1.5$ TeV. For the definition of the 2BM/3BM decays, see the text. For the first bin ($E_{\text{in}}/E_{\text{LLP}} < 0.1$) $E_{\text{in}} = 0$.

Figure 15: As in Fig. 14 but for the case of the muon chamber.

ECAL would be a good identifier as it is expected that the shapes would be different for the inside-out conventional jets versus the outside-in jets. Using ECAL timing information can help in drastically reducing such backgrounds. When the LLP decays in one of the outer layers of the HCAL, we will be more interested in the HCAL shower shape variables. Provided the upgrade of the HCAL has depth segmentation, the energy, $E(D_i)$ deposited in the $i$th layer of the HCAL, can be used as inputs in a BDT to discriminate between backward moving signal jets and forward moving background jets. The proposed high-granularity calorimeter within CMS might help us address several of these issues. Furthermore, gaining timing information from the muon chamber as well as the upgrades in the tracker (with the inclusion of additional timing layers) will be important for understanding such signatures, better. $BMOs$ which are heavily displaced with respect to the primary vertex can have large impact parameters and will mostly be not recognised by the present jet algorithms. The jet algorithms can be tuned to catch such displaced jets with large impact parameters but this can be extremely resource intensive. Ideas like data scouting and parking can be used to improve this situation. Moreover, reconstructing the $BMOs$ in the tracker can be very challenging and one has to make modifications in the track reconstruction algorithms by relaxing the requirements on the impact parameter.

Finally, before concluding this section, we want to refer to a study [745] that shows the potential of constraining the proper lifetimes of LLPs provided they are discovered. This study considers the prospects of the high-luminosity runs of the LHC. High pile-up is considered with the various upgrades that are proposed. Model-dependent and model-
Figure 16: Model-dependent $\chi^2$ as a function of the reconstructed decay length $c\tau$.

Independent methods are utilised and machine learning algorithms employed to reconstruct the proper lifetimes of neutral LLPs decaying into leptons (may be also accompanied with missing energy). The proper lifetimes of charged LLPs decaying into leptons and missing energy is also considered. Neutral LLPs decaying into displaced jets is discussed along with the challenges faced in high PU environments. As an example, we show the lifetime estimates for the model-dependent displaced lepton category in Table 7 and Fig. 16.

3.5 Summary

For many known and unknown reasons, BSM physics may be hidden under the current searches at the LHC. There are many notoriously difficult scenarios in some well-known frameworks, such as supersymmetry and hidden-sector models. Known reasons include non-prompt decays of the new particles, very soft particles in the final states, overlapping with existing particles, and/or very small production rates.

For these difficult scenarios with known reasons, theorists and experimenters have been working together to formulate useful strategies to cover them. New triggers have been put in the trigger system to identify non-prompt or long-lived particles, covering various signatures for the LLPs in various parts of the detector. New techniques including machine-learning or initial-state radiation are employed to improve the possibility of detecting such soft particles in the final state and for those signals with small production rates. For those scenarios that are hidden due to unknown reasons or unexplored theories, unconventional ideas such as anomaly detection with machine learning would help.
<table>
<thead>
<tr>
<th>$M_X$ (GeV)</th>
<th>DL (cm)</th>
<th>cross-section (fb)</th>
<th>Rec. DL (cm)</th>
<th>$1\sigma$ LL (cm)</th>
<th>$1\sigma$ UL (cm)</th>
<th>$2\sigma$ LL (cm)</th>
<th>$2\sigma$ UL (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>8</td>
<td>12</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>11</td>
<td>4</td>
<td>23</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>7</td>
<td>3</td>
<td>30</td>
<td>2</td>
<td>35</td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>1</td>
<td>39</td>
<td>23</td>
<td>83</td>
<td>21</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>-</td>
<td>9</td>
<td>&gt;150</td>
<td>8</td>
<td>&gt;150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>-</td>
<td>5</td>
<td>&gt;150</td>
<td>4</td>
<td>&gt;150</td>
</tr>
<tr>
<td>1000</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>10</td>
<td>7</td>
<td>16</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>11</td>
<td>6</td>
<td>22</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>1</td>
<td>42</td>
<td>34</td>
<td>69</td>
<td>33</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>-</td>
<td>19</td>
<td>&gt;150</td>
<td>17</td>
<td>&gt;150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>-</td>
<td>10</td>
<td>&gt;150</td>
<td>9</td>
<td>&gt;150</td>
</tr>
</tbody>
</table>

Table 7: Lifetime estimates by model-dependent $\chi^2$ fitting of the $d\tau$ distribution for the displaced leptons signature.

4 Open data in particle physics

Section editors: Nishita Desai and Suchita Kulkarni
Contributions: Matthew Bellis, Philip Harris, Clemens Lange, Kati Lassila-Perini and Jesse Thaler

The goal of discovering new physics at the LHC rests on our ability to collect and analyse all interesting collision data. The detectors at the LHC are intricate machines that require extraordinary levels of technical sophistication. Although the abstract concept of a “collision event” is clear enough to theorists and experimentalists alike, the actual data readout and its translation into a “reconstructed” event from which physics may be extracted is extremely complicated and the algorithms employed to do this are perpetually under improvement. The question of data preservation therefore is also extremely complicated and fraught. Indeed for nearly all particle physics experiments before now, no one outside of the respective experimental collaborations has had either access or ability to make any use of data collected at the said experiment. The reliance on computer software to store and analyse the data further means that it is usable only as long as the software is runnable (which unsurprisingly is no longer viable after a couple decades in most cases).

The High Energy Physics community is historically strongly in support of initiatives to keep all scientific research available to the public at no cost (i.e. the cost of publication and the infrastructure to maintain access is borne not by the individuals who perform or access research but by participating educational institutes, governmental or non-governmental organisations). The SCOAP3 initiative, for example, is a partnership of over three thousand libraries, major funding agencies and research centres in 44 countries and 3 intergovernmental organisations to ensure open access to research published as journal articles. Nearly all articles published in research journals are also voluntarily and independently uploaded by authors on the arXiv preprint server. The LHC experimental collaborations have also increasingly made the data from published papers — histograms, figures and tables — available in human readable and machine readable formats via the HepData portal. In this context the LHC reinterpretation forum and the Les Houches recommendations have been important. All major software used in scientific computation is made available under a
public licence (most commonly the GNU Public licence, but occasionally Creative Commons or similar licences may be used.) Zenodo [740] is another sister repository, also maintained by CERN with the aim to provide citable DOIs to preserve research software and data products and in this sense is complementary to HEPData. The idea that data from publicly funded experiments might still remain behind a curtain is therefore highly unpalatable to most high-energy physicists. However, the question of what to release as data and how to ensure integrity in its future use remain questions that are under active and vigorous discussion.

As a first step towards open data, CERN launched the Open Data portal in November 2014 [747]. At the moment, this portal contains rich content involving collision and simulated datasets for research, derived datasets for education, configuration files and documentation, virtual machines and container images and software tools and analysis examples. As of March 2021, this data contains over 7600 bibliographic records and over 900k files or with 2.4 petabytes of data.

The availability of Open Data has also enabled novel theoretical research. There are currently over thirty papers citing the open data framework and multiple new studies in the realm of new searches, QCD jet studies, and Machine Learning [748–750] have been performed. At least one such work has then prompted further work by the publishing experiment [751], completing the experiment to theory and back to experiment cycle and providing evidence to the claim that Open Data would result in genuine scientific advance.

4.1 Challenges

Open data efforts at the LHC need to overcome several technical and philosophical challenges. Some of these are related to the fact that only a subset data is stored and hence the lost data can never be recovered. Data collected at the LHC is also released with some delay allowing collaborations to exploit first. Challenges relating to preserving technology used to process this data by collaborations, data documentation and necessity for data validation mechanisms also need to be considered.

4.1.1 The problem with triggers

What shall be deemed an “interesting” collision event at the LHC is determined based on our current understanding of the Standard Model. The unprecedented number of proton collisions per second — or luminosity — means that not every “event” can be stored to the disk. The experiments therefore choose what to store using certain criteria called triggers. These triggers — combining hardware level and software level decision making — aim to select events that either have large enough energy signatures (e.g. by requiring high-momentum particles) or unusual combination of detectors firing at the same time, with the assumption that anything that “happens” would be caught by one of these. The first of these strategies has been honed over time. Over the last few years, considerable work has also been done to improve triggers to include unusual events with long-lived particles or radiation from hidden sectors. This second kind of triggering, however, necessarily needs physicists to know beforehand what kind of non-standard signature is expected. Therefore, it is obvious that only theories that the field already knows are visible in standard triggers, or are popular enough to have a trigger designed for that specific signature.
4.1.2 The role of the collaborations

Experimental physics collaborations currently have not only the privilege of data access but also the responsibility of ensuring the accuracy of their interpretation. Every result announced by an experimental collaboration is painstakingly cross checked by several independent internal groups who each use their own algorithms and strategies to ensure that the final calibration and statistical analysis is accurate.

Once data is publicly released, this quality control is out of the hands of the experimental collaborations and therefore it would be impossible for the experiments to ensure the quality of results claimed based on their data. This has for a long time been the primary reason cited to avoid releasing data publicly. Furthermore, the resulting publicity from a spectacular claim often proceeds under its own steam and fraudulent or over-enthusiastic claims that are then proven to be false would then result in erosion of public trust in the scientific process and further endanger funding of future fundamental physics experiments which clearly cannot be afforded by any single university, or even country. It is important to point out that this concern is largely acknowledged by both experimental and theory community. The theory community is hence especially careful in using open data.

4.1.3 Technical challenges

First, the experiments need to decide whether they intend to release raw or processed data, both of these choices would come with their own issues. In general, if we wish to have a reliable and usable open data framework, much thought and work is needed in at least three directions:

- Reliable storage and access technology. Preservation of software used for data analysis and hardware capable of running said software.

- Detailed meta-data and documentation:
  - In case of raw data, all relevant calibration information, preservation of associated algorithms etc. to obtain the reconstructed events
  - In case of releasing reconstructed events only, extensive documentation and internal information characterising the reconstruction and explaining limitations in its usability

- Mechanisms for validation of results. How to ensure sanity of results derived using this data? Educating the public on what is “real” and what fake.

4.1.4 The ultimate aim

When the signatures of new physics have been captured and saved, long-term access to data by the scientific community could use it to test all possible theories. The case for an Open Data platform is therefore motivated not just by arguments of democratic access to data obtained by a publicly funded experiment, but makes solid scientific sense.

Such open data policy has long existed in the field of astrophysics, where e.g. data from publicly funded telescopes is released regularly. The main challenge in doing this for particle physics experiments comes from the complexity in triggering and calibrating the data.
4.2 Existing framework

Experimental data being openly available is a well established philosophy in HEP. However, making large amounts of data available is not the same as it being useful. This scientific data management and stewardship is taken care of by the FAIR principles, which are also embraced by CERN open data. Moreover, keeping in mind the needs of different users, data are made available in different levels via different platforms as well. We will briefly review implementation of FAIR principles in CERN open data portal and associated efforts to make data more user-friendly.

4.2.1 The FAIR principles

This independent effort is the articulation of principles needed to publish and maintain the quality of data released. This is summarised by the FAIR \cite{752} guiding principles for scientific data management, which emphasises that all data published should be Findable, Accessible, Interoperable, and Reusable. The CERN Open Data effort aligns with the FAIR principle by ensuring:

- **Findability:** Assigning a “record ID” and optionally a DOI to every data product. Rich context description and associated documentation is provided. All histograms and numerical data are required to be machine readable. A search interface is provided for searching and identifying correct datasets.

- **Accessibility:** A graphical user interface for manually downloading the data and an automated cern opendata client is available. This command line client supports downloading the dataset and metadata via HTTP and XRootD access protocols.

- **Interoperability:** CERN open data portal offers several data formats and vocabularies as a community standard. In addition, data is also offered under common classification rather than formal vocabularies. The latter assists in physics interpretation, while the former is applicable in designing appropriate data processing chains. In addition, the data variables are equipped with detailed semantics description. This helps identifying the variables which are of utmost importance for analysis design. For some examples of analysis development based on the principle of interoperability see Refs. \cite{753,754}.

- **Reusability:** A detailed record of data provenance i.e., how the data was generated is available in JSON format. Instructions for processing of both RAW and reconstructed datasets (e.g., AOD format used by CMS) are provided and lastly, computing environment is preserved in the form of docker and singularity images.

4.2.2 Available data from the LHC collaborations

The CERN Open data portal currentlycatalogues data from six experiments — ATLAS, ALICE, CMS, LHCb, OPERA and PHENIX. Focusing on the LHC experiments, of the four, CMS has the largest repository of research-level data available (currently at over 1.1k datasets). ATLAS currently only has outreach-level data. ALICE has released 15 datasets and documentation from 2.76 TeV and 7 TeV runs. Whereas LHCb has 4 datasets of very limited events to be used with some published analysis software. Several data releases for research-level data are available. The latest release in this series was in December 2020 and contains data from the CMS experiment including first 2010–11 heavy-ion data samples and
reference proton-proton datasets (214 TB). We will discuss the CMS Open data format in detail in the coming sections, however, we outline general principles here.

4.2.3 CERN open data policy

The CERN open data policy recommends data releases at four different levels:

- **Level 1** consists of results which are released in the form of publications. Access to these publications is enabled under the SCOAP3 agreement for high energy physics.

- **Level 2** is aimed at outreach and education and access is enabled via CERN open data portal.

- **Level 3** consists of reconstructed data which will be useful for reproduction of physics analysis as well as for new physics analyses to be designed later if needed.

- **Level 4** contains raw collision data which may not be suitable for external consumption as experiment-specific knowledge of calibration or resolution may be needed to correctly interpret this data.

In addition, all data published by CERN experiments will try to abide by FAIR requirements. However, aside from publicly available data, some amount of internal or “restricted” analysis knowledge needs to be preserved. This is also done in accordance of the FAIR principles, on the CERN Analysis Preservation portal [755]. It should be remembered that FAIR principles do not necessarily require open access.

4.2.4 CERN Open Data Portal

The CERN Open Data Portal [10] is the gateway to accessing the CMS open data. For CMS open data users, it provides access to the data themselves, the software needed access these data, and a number of examples and mini-tutorials on how to use these tools both for analysis and education. It also has an incredibly detailed amount of information about the provenance of these datasets, both collision data and Monte Carlo. In a sense, the “documentation” provided is complete in the same way a dictionary is complete. But while a dictionary will help you define any word you run across, it is perhaps not the best resource if you want to compose a sonnet in the style of Shakespeare. For a more concrete example, the CMS Open Data Guide will show a user how to access the data and explore the different types of “physics objects”, but it does not give much explanation on these objects or why there might be multiple definitions of a ”Muon”.

As described in Sec. 4.4 in a 2017 paper from Tripathee et al. [756], the non-CMS-member authors used CMS open data to explore jet substructure. That paper contains an appendix, “Advice to the community”, in which the authors detailed their experience with the open data, both good and bad, and provided advice for improving access for other users. While the authors were able to produce new scientific measurements with these data, it was not without challenges. The success of their group motivated the CMS open data team to provide a better experience for open data users by significantly improving the documentation.

It should be pointed out that this did not mean getting rid of the information on the Portal. That information is complete and necessary for users who want the details of what triggers were used for collision datasets or the details of generator parameters for the Monte Carlo simulations.

---

[10]: https://opendata.cern.ch/
Carlo. Instead, the decision was made to develop a guide for new users and to organise a workshop to provide a hands-on experience for these users.

4.2.5 REANA and HEPdata

In the context of open access to analysis information, an excellent resource has also been developed in the form of the REANA framework [757]. This is a platform for reproducible analysis and aims at providing an integrated access to the data, computing environment, and recipes. The framework can be deployed by using containerised workflows on Kubernetes, HTCondor, Slurm back-ends. It is possible to process both CMS and ATLAS data via the REANA framework.

Another very notable resource is the HEPData repository. HEPData [758] hosts publication level data. It also provides an interactive interface to explore and download publication-level data behind plots and tables. So far HEPData hosted only results in the form of plots, however now it has started to also host likelihood information, which is a very welcome step towards creating a central place for repository of a publication level data. Along with HEPData, it also offers a bridge to GitHub and hence preserves older software releases. Zenodo is widely used by the machine learning community among other users at the LHC. Finally, some amount of internal or “restricted” analysis knowledge still needs to be preserved. This is also done in accordance of the FAIR principles, on this portal [755]. It should be remembered that FAIR principles do not incorporate open access.

4.3 CMS open data

The CMS experiment established a data preservation, re-use and open access policy in 2012 [759] and started regular releases of research-quality data in 2014. All proton-proton data collected during 2010 – 2011 and a half of those from 2012 are now in the public domain, and the latest release in December 2020 includes heavy-ion data from 2010–2011. This is a substantial amount of data, resulting, together with associated artifacts, in a total volume of 2.3 PB.

4.3.1 CMS data releases

The CMS data releases take place regularly, after an embargo period of six years following data taking. The releases include 50% of the collision data and the corresponding simulated datasets. The collision data release is completed within 10 years. However, the amount of open data will be limited to 20% of data with the similar centre-of-mass energy and collision type while such data are still planned to be taken.

CMS releases a full reprocessing of data at Level 3 (reconstructed data good for physics analysis) from each data-taking period and the data are released in the same format and with the same data quality requirements from which analyses of the CMS collaboration start. For the Run 1 data (data taking 2010 –2012) the format is the Analysis Object Data (AOD) format, based on the ROOT framework [760] and processed through the CMS software CMSSW [761]. This format contains reconstructed “physics objects” such as electrons, muons, jets, and their properties, and keeps the most relevant lower-level information such as hits in the tracking system and calorimeter clusters corresponding to the physics objects.

The CMS open data follow FAIR principles and are provided with rich associated metadata. Due to the complexity of experimental particle physics data, the FAIR principles alone do not guarantee the re-usability of these data, and additional effort is needed to pass the
knowledge needed to interpret them. The interplay between the CMS experiment open data team and the CMS open data users is of utmost importance in this context.

Figure 17 provides a simplified flowchart of an analysis, as well as what hardware, software, and documentation resources are needed/provided for the CMS open data. Grey boxes indicate the resources that are provided by CERN open data portal and CMS open data team. Light green boxes indicate steps where the procedures, software, or hardware and storage is left to the CMS open data user.

The recommendations from users of CMS open data who are external to the CMS collaboration (see Section 4.4) are valuable to the CMS open data team, and acts a guideline for the future directions, in the limits of possible. The challenges and measures taken to address them by the CMS open data team are discussed in Secs. 4.3.2–4.3.4.

4.3.2 Data complexity

The CMS data are released in the format that allows their widest possible use, without special preparation for their use in the public domain. This is also the format from which the collaboration members start the analysis. However, as underlined in the preceding section, the complex data format is a challenge for open data users.

The complexity originates from several factors. First, there is no single definition of a physics object, there can be multiple instances of an object, such as jets defined with different algorithms, and the object type is chosen to match the requirements of the analysis. Furthermore, a signal can be interpreted as several different object candidates, for example, a signal interpreted as a muon can also appear in the list of electron candidates. Therefore, physics object identification criteria are always applied in the data analysis, and the criteria are adapted balancing between the efficiency of selecting an object and the purity of the
selected object, and depend on the analysis being carried out.

Final corrections and fine-tuning to the objects are often applied in the analysis phase. This happens because corrections and algorithms are developed at the same time as the first analysis of the data, and do not necessarily make their way to the final reprocessed data. Some of these corrections are available from the condition database, and some others as separate “recipes” to be applied to the objects.

Another challenge are the triggers and procedures related to data taking. A single dataset consists of events passing one or more of hundreds of different trigger paths. In some cases triggers are prescaled so that only a predefined fraction of events passing the trigger selection is recorded, to enable collecting data for processes that occur often and otherwise would fill the entire data-taking bandwidth. One event may also end up in two different datasets, and this eventual overlap needs to be accounted for in the analysis. All this needs to be handled in the data analysis to properly scale the number of selected events with the cross-sections being measured.

The use of CMS data also requires some knowledge of the CMS experiment software CMSSW, built on top of the HEP-specific ROOT data structure. The software is openly available, and provided as software containers to open data users. The efforts for facilitating its use are discussed in Sec. 4.3.3.

Part of these difficulties will be overcome when the slimmer miniAOD and nanoAOD formats will be made available. These formats are both based on the ROOT data structure, but the nanoAOD format does not require using CMSSW software, and will therefore be of particular interest to CMS open data users. These formats are routinely produced for Run 2 data, the nanoAOD format starting from the data taking of 2016. However, this slimming has a price and it is estimated that roughly half of the analyses done in CMS can be done using nanoAOD. Therefore, miniAOD format will also be made available.

4.3.3 Examples and documentation

At the time of publication of this document, users of the CMS Open Data have had three options for learning about how to access and analyse these data: the information provided on the CERN Open Data Portal, CMS Open Data Guide, and the (to-date) two workshops run by the CMS open data team. The Open Data Portal is the original source of documentation, but suffers from a lack of any sort of “roadmap” for new users, cf. Sec. 4.2.4. The CMS Open Data Guide is very much a work-in-progress at the time of this document (Summer 2021). The workshops have been very successful and have provided person-power and feedback to improve the documentation, in addition to being a pinned source of information.

**CMS Open Data Guide:** Within CMS, one of the most widely used onboarding tools is the [CMS Offline Workbook](https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook) a series of wiki pages that walk users through all the software requirements of becoming a CMS member. From getting your computing accounts to how to start data analysis to generating your own Monte Carlo data. Much of it is public, but it also links to internal pages that require CMS membership to access. It also tends to be focused on the latest data releases, and not older datasets like are released on the open data portal, though the documentation is still there.

The Workbook is laid out like a roadmap, walking the user from the first steps of finding and accessing data to more complicated procedures like jet energy corrections. It is a great resource for CMS users and the CMS open data team is currently working on an analogous
site for the open data, to be referred to as the CMS Open Data Guide. It is very much a work-in-progress but the overall structure is there, and it mirrors the Workbook in that it tries to provide a starting point for a brand new user, eager to work with the open data.

It would be a waste of person-power to try and completely rewrite the documentation that exists on the Open Data Portal and it would be near impossible to keep them in sync, should things change. Instead, the goal is to have the Open Data Guide link to documentation on the Portal that already exists, while providing contextualizing information and guidance on how that documentation can be used. What additional information is necessary is informed by the paper from Thaler et al. [756] and feedback from participants in the 2020 and 2021 workshops.

CMS Open Data Workshops: The CMS open data group has conducted two workshops in 2020 and 2021: “CMS Open Data Workshop for Theorists at the LPC”\textsuperscript{12} and “CMS Open Data Workshop”\textsuperscript{13} which ran for 3 and 4 days, respectively. There were about 25 engaged participants in the first and about 50 in the second. Because of the pandemic, both were held virtually and attracted scientists from all over the world. These workshops built on the model of a hands-on workshop where participants were required to actually do the exercises and run the code themselves. The group opted to build the workshop lessons on a framework developed by the Software Carpentry\textsuperscript{14} organisation, a framework also in use by other CERN workshops.

The workshops followed a similar format on a coarse level. Participants were shown how to find the data of interest on the Portal. Lectures and examples were given about how to locate and apply different triggers and how to access different physics objects, as well as what those objects meant. Examples were also given of how to apply different energy corrections (e.g. jets) and how to use those in systematic uncertainty calculations. Both workshops ended with a hands-on example of how users could leverage Google Cloud computing to process the open data to scale and store the skimmed data on Google’s Cloud infrastructure.

Time was also allotted to gather feedback from the participants on their experience with the workshop and their interest in using the open data, and a survey was sent out after the workshop to supplement this real-time feedback. This feedback has been and will be used to improve the next versions of these workshops as well as to inform the development of the Open Data Guide. For example, some pages of the Guide might point to the lessons from the Workshops, all of which are accessible on the web still along with recordings of the lectures.

4.3.4 Usability

There are several challenges when it comes to the usability of open data. The complexity of the data on its own as discussed in Sec. 4.3.2 often makes it difficult to get started in the first place. In addition, the original data sets are huge and require substantial computing power for processing. Derived and therefore simplified data sets make it easier to obtain a first meaningful result. The examples, documentation, and workshops discussed in Secs. 4.2.4 and 4.3.3 lower the barrier of entry. However, the examples provided are far from a realistic and complete physics analysis.

\textsuperscript{12} https://indico.cern.ch/event/882586/
\textsuperscript{13} https://indico.cern.ch/event/1031398/
\textsuperscript{14} https://software-carpentry.org/
In order to e.g. generate new Monte Carlo simulation samples to test a new physics model, change the underlying reconstruction to evaluate new reconstruction methods, or assess systematic uncertainties one needs to use the experimental software, i.e. CMSSW in the case of CMS, and for the Run 1 data the AOD format. For a few events, this can be done on a local computer, but the data set required for an analysis including simulation samples typically consist of millions to billions of events. In order to be able to process these, one needs to have access to high-throughput computing resources. In addition, these need to provide the possibility to execute CMSSW. The use of software containers make this significantly easier as long as the underlying processor architecture is the same. Thus, CMS open data releases are accompanied by software container images that contain the respective CMSSW release needed for processing the data.

Besides using these software images for Monte Carlo event generation, simulation and reconstruction, they can be used to extract desired event information and convert it into CMSSW-independent data formats as discussed in Sec. 4.4. Furthermore, by extending the existing software, new algorithms can be developed and integrated. The updated software can be added to the container image and shared via container registries such as Docker Hub [763].

There are currently no computing resources provided by the CMS Collaboration for the use of CMS open data by non-CMS members. These therefore need to be provided by the open data users themselves, e.g. at their institution or by using commercial public cloud computing services. To help open data users make use of public cloud compute resources, tutorials for “CMS open data in the cloud” have been conducted at the CMS open data workshops, cf. Sec. 4.3.3.

Since individual data and simulation events are independent of each other, processing of the data sets can be parallelised. A typical analysis workflow therefore splits or scatters the files belonging to a data set into smaller chunks, which are then merged or gathered after the processing step. The input data can hereby be streamed directly from the open data servers or first copied to local disks. Since public cloud computing services often provide the possibility to quickly scale up computing resources and then scale them down immediately after processing, this scatter-gather step can be sped up significantly. These steps can be implemented as a simple workflow, for example using Argo Workflows [764], for which examples are provided at the CMS open data workshops. After this first processing step, the results can usually be analysed further on smaller computing clusters or even local computers.

The use of software containers has also been instrumental for the development and validation of CMS open data workflows in general since they provide a possibility to use CMSSW independently of the standard high-energy physics computing environments as well as in continuous integration systems. The latter are particularly important to ensure that the open data remain usable. However, while in particular Docker has made it significantly easier to use software containers, there are still some technical hurdles that need to be overcome, which often depend on the operating system used. Feedback from the CMS open data workshops was very useful to further improve documentation and usability.

4.4 Recommendations from External Users

In the first application of CMS Open Data, based on the release of 2010 data, the authors of Ref. [756] highlighted a number of challenges and recommendations for the community. Some of these issues have been resolved by subsequent CMS Open Data releases, such as
the inclusion of simulated Monte Carlo data sets. Other issues are a challenge not only for external users, but also for internal CMS application, such as the lack of centralised documentation.

4.4.1 Missing information

Because the CMS Open Data is stored in the AOD format, it is in principle possible to reproduce any CMS analysis that does not rely on lower-level information. In practice, though, there are numerous technical challenges to outsiders using the AOD format as well as important knowledge about the CMS experiment that is not fully archived. Despite these challenges, there have been a number of successful uses of the CMS Open Data in the literature, cf. Refs. [748–750,756,765–777].

4.4.2 Validation samples

The first recommendation is to provide reference validation examples that includes all steps of published CMS analyses. For example, the authors of Ref. [748] attempted to repeat the measurement of the Z boson cross-section to validate their treatment of the di-muon final state; a reference Z boson measurement from CMS would have been helpful in this context. Progress towards establishing some benchmarks were made in Ref. [770], though the code for that study is not yet public.

4.4.3 Industry standard file format

The second recommendation is to release data in an industry standard file format. The current pipeline involves running the CMS software framework on a virtual machine, which can become unwieldy for analyses that require a large number of events or need to run in a cluster/cloud environment. The authors of Ref. [771] translated a subset of the information from select AOD files into the standard HDF5 format [778] and posted them on Zenodo [779–787] along with example code [788–790]. This format allows the use of external analysis software, and it provides a benchmark data sample for the jet physics community. This issue may eventually be resolved when CMS releases data in the nanoAOD format.

4.5 Summary

4.5.1 Current status

The particle physics community has spent much time and effort to articulate and design principles and practices to enable publication of data at various levels of openness. The SCOAP3 agreement provides platforms in terms of journals, books and a document repository that ensures all material published under this agreement will be available as open access, without cost to individuals who would wish to access it. Open access to publication-level material however does not automatically mean open access to data and this has to be negotiated separately. Although releasing data publicly has long been the norm in fields like astrophysics, it is not common in particle physics due to the extremely complex nature of data and the extraordinary amount of processing needed to bring it to a level from which physics analyses can be done. Two main efforts in this regard are the publication of FAIR principles for scientific data and the CERN Open Data project. All CERN experiments also encourage that published analyses are accompanied by digitised tables and plots uploaded to
the HepData repository. The implementation of this last requirement is still patchy, however compliance is improving.

The premise and usage of open data is multi-fold, therefore data releases correspond to four different levels ranging from publication-level results up to raw collision data. The two intermediate levels, one that allows for education and outreach, and one that enables real data preservation and re-analysis, are the levels that require the most thought and sophistication in implementation.

It is not enough just to have open data available, it is important that it is findable, accessible, interoperable and reusable. This is ensured by means of the FAIR principles on basis of which open data portals are designed. Currently, we find only the CMS experiment has research-level data available on the CERN OpenData portal.

4.5.2 Future improvements

Despite advances of open data in particle physics, its usage remains challenging for any person who does not have training in the vocabulary used by the experimental collaboration. The data format, the methods for selecting and downloading datasets and the preprocessing needed are quite opaque and require extensive documentation to be made usable. Publication of step-by-step validation of analyses and using industry-standard file formats (as opposed to home-grown ones) are two suggestions we find from current non-CMS users of CMS Open Data. The availability of simulated datasets (not just collision data) would also help in disambiguation of complicated processes. It is also worth noting that apart from CMS, data from no other experimental collaboration is currently at a level that can be used for physics studies. It will be much easier to have a fruitful discussion about improvements once a few more implementations of the open data principles become available that can provide illustrations of best practices (or lack thereof).

Another important factor that hampers usability, at least for theorists, is the non-availability of computing resources. Given the size of the data-files and associated frameworks, it is nearly impossible to run it on a laptop or a personal computing machine. Access to high-performance computing clusters necessary to process data is expensive and currently limits accessibility to individuals who belong to universities or institutes that already have some provision for computing. Simplifying access to powerful computing machinery will improve open data usage beyond experiments and a handful of theory groups. Therefore, while open data signals the beginning of a very important journey towards open science, there is still some progress to be made before it is widely used and exploited.

5 Summary

The field of particle physics is at the crossroads. The discovery of a Higgs-like boson, a major accomplishment for the field almost fifty years in the making, completes the predictions for fundamental particles and interactions in the SM. At present no clear guidance is available as to how it will break down. On the other hand, the motivation for the existence of New Physics, given by, for example, the established existence of Dark Matter and Neutrino Oscillations, has not diminished since the discovery of the Higgs-like boson.

The immense and far-reaching scope of the LHC physics programme promises to ultimately unveil the New Physics. The exploration of the phase-space by the LHC experiments is far from exhausted even with the available data sets, where Run 3 and the HL-LHC promise to deliver up to $4\text{ab}^{-1}$ of integrated luminosity. Unfortunately, the exploration of
the LHC data by way of inclusive or model dependent searches performed to date indicates that no striking resonances have been observed in the accessible dynamic range. In this paper, we posit that New Physics is accessible at the LHC in principle, but that it is inaccessible by current analysis strategies for a variety of reasons, and that its signatures are hidden in the data.

We summarise the status of the most significant anomalous experimental results in particle physics, including the most recent results for the flavour anomalies, the multi-lepton anomalies at the LHC, the Higgs-like excess at around 96 GeV, and anomalies in neutrino physics, astrophysics, cosmology, and ultra-high energy cosmic rays. The anomalies corroborate the need for extensions of the SM, and we provide overviews over possible BSM models. The fact that many anomalies can be explained within the same theoretical framework is pivotal, as it stimulates model building, which already gave rise to a plethora of BSM models and classes of models.

The known systemic shortcomings of the LHC and its search strategies allowed us to identify some of these reasons, including: final states consisting in soft particles only, associated production processes, QCD-like final states, close-by SM resonances, and SUSY scenarios where no missing energy is produced. We find that new strategies are necessary to unveil the hidden NP signatures, which have to strike a careful balance between the model-centric and less model-dependent approaches. It is generally understood that Machine Learning can play a significant role here, with unsupervised and semi-supervised learning and a wide range of algorithms becoming invaluable assets. Another very promising avenue is presented by CERN’s open data policy, which provides a testing ground for new search strategies with a quick turnaround.

We discussed the challenges for open data in particle physics, including preservation and access of data and software, documentation, and validation mechanisms. The CERN open data policy and the independent FAIR principles are meant to ascertain that open data will be useful to the community. A specific example of open data is given by the CMS collaboration, which releases reconstructed data after six years to be used widely for analyses and is so far the only collaboration sharing its data in this way. CMS open data can be accessed via the CERN Open Data Portal, the CMS Open Data Guide, and two workshops run by the CMS open data team. We summarised CMS users’ and external users’ recommendations on data complexity and gave an overview over the challenges and measures taken by the CMS open data team to address them.

We conclude that wide access to open data by individuals is necessary to fully exploit the potential of the LHC and that, despite the advances in CERN open data, its public usage remains challenging for individuals. Improvements of data formats, the documentation, and availability of computing resources, are required to enable the community. We find that individuals using public data for their own research does not imply competition with experimental efforts, but rather provides unique opportunities to give guidance for further NP searches by the collaborations. The communication between theorists and experimentalists is paramount, possibly now more than ever.

Acknowledgements

We thank S. Kraml for useful comments. SK is supported by the Austrian Science Fund Elise-Richter grant project number V592-N27. ND acknowledges the support of Department of Science and Technology of the Government of India via the Ramanujan Fellowship SB/S2/RJN-070/2018. BB is supported by the ERC research grant NEO-NAT no. 669668.
ZB is supported in part by the MIUR grant PRIN 2017X7X85K and in part by the SRNSF grant DI-18-335. TH is supported in part by the U.S. Department of Energy under grant No. DE-FG02-95ER40896. KC is supported in part by Taiwan Ministry of Sciences and Technology with grant number MoST-110-2112-M-007-017-MY3. JT is supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, [http://iaiff.org/]), and by the U.S. DOE Office of High Energy Physics under grant number DE-SC0012567. A.C. and C.A.M. acknowledge financial support by the Swiss National Science Foundation, Project No. PP00P2_176884. M.H. is supported by the Swiss National Science Foundation, Project No. PCEFP2_181117. MB is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant 396021762 – TRR 257. B.C. is supported by the Italian Ministry of Research (MIUR) under the Grant No. PRIN 20172LNEEZ. A.P. is supported by the Spanish Government and ERDF funds from the EU Commission [grant FPA2017-84445-P] and by the Generalitat Valenciana [grant Prometeo/2017/053]. BM and XR are grateful for support from the South African Department of Science and Innovation through the SA-CERN programme and the National Research Foundation for various forms of support. MK was supported by MIUR (Italy) under a contract PRIN 2015P5SBHT and by INFN Sezione di Roma La Sapienza and partially supported by the ERC-2010 DaMESyFla Grant Agreement Number: 267985. Contribution by MB is based upon work supported by the National Science Foundation under Grant No. PHY-1913923. DM acknowledges support by MIUR grant PRIN 2017L5W2PT and the INFN grant SESAMO. The work of BD is supported in part by the U.S. Department of Energy under Grant No. DE-SC0017987.

References

[10] [https://github.com/hepcomm/hepmist]


[334] P. A. Zyla et al. [Particle Data Group], PTEP 2020, 083C01 (2020).


[391] [CMS], 2021. http://cds.cern.ch/record/2758905


89


[665] [ATLAS], 2021. http://cds.cern.ch/record/2759676


[710] [CMS], 2018. http://cds.cern.ch/record/2621419


[717] [CMS], 2018. http://cds.cern.ch/record/2647987


[723] [MoEDAL], J. Pinfold et al., 2009. http://cds.cern.ch/record/1181486


[735] P. Ilten, Y. Soreq, M. Williams, and W. Xue, JHEP 06, 004 (2018),
    arXiv:1801.04847 [hep-ph].
[740] Z. Chacko, P. J. Fox, R. Harnik, and Z. Liu, JHEP 03, 112 (2021),


