Contribution 22

Recasting Long-Lived Particles Searches

G. Cottin, N. Desai, J. Heisig, A. Lessa

Abstract

Long-lived particles (LLPs) arise in several beyond the Standard Model (BSM) theories, and they provide striking (non-standard) signatures at colliders. Several LHC searches look for LLP models in a broad range of final states, and limits have been presented for specific BSM models. However, extrapolating such limits to other scenarios often proves to be a difficult task outside the experimental collaborations. This note discusses the recasting of three types of LLP signatures: displaced vertices, displaced leptons and heavy stable charged particles. Several conclusions are obtained from these recasting attempts and recommendations to the experimental collaborations are made.

1 INTRODUCTION

The key scientific goal for the second run of the LHC is to explore physics beyond the Standard Model (BSM). Motivated by a variety of BSM theories, there has been a growing interest in non-standard signatures, such as long-lived particles (LLP). A large variety of LLP signatures have been explored by the experimental searches and the re-interpretation of these results within the context of new BSM theories is extremely relevant to exploit the full potential of the LHC experiment. However, in most cases it is not possible to reinterpret LLP searches using fast detector simulation and a cut-and-count based analysis, as it is usually done for prompt searches. In particular, efficiencies for object reconstruction (such as displaced vertices) and event selection are much more analysis dependent in LLP searches and difficult to reproduce using fast simulation. Furthermore, trigger and pile-up vetoes included in prompt searches are difficult to reproduce and can invalidate the extrapolation of prompt search results and limits to long-lived (displaced) scenarios.

In this note we discuss the difficulties of recasting LLP searches. In order to make the discussion concrete, we present results for the recasting of three distinct LLP signatures: displaced vertices, displaced leptons and heavy stable charged particles. As a way to discuss the typical issues encountered when recasting LLP searches, we try to reproduce the official exclusion curves presented in the 13 TeV ATLAS displaced vertex plus missing energy [598], the 8 TeV CMS displaced lepton [599] and the 13 TeV CMS heavy stable charged particle [120] searches. Although the issue of recasting prompt searches within the context of LLP models is a very relevant one, we do not discuss it here.

The first difficulty related to recasting LLP searches concerns the detector simulation of such signatures. While the relevant signatures for stable particles (in detector scales) are charged tracks and missing energy, for a particle decaying within the detector several signatures are possible, depending on the LLP nature and its decay. The lifetime of the particle and its boost are also essential features, since only certain parts of the detector are capable of observing specific decay products. Furthermore, the Standard Model background typically decreases as the LLP decay moves further away from the primary vertex. In general the LLP signatures can be classified as follows:

- *charged track* (stable charged particle)
- *disappearing track* (charged LLP decaying to a neutral/soft final state)
- *displaced vertex* (charged or neutral LLP decaying to charged final states)
- *track kink* (charged LLP decaying to neutral and charged final states)
- *trackless jets, displaced leptons* (neutral LLP decaying to charged final states)
- *missing energy*[1] (stable neutral particle)

Typical fast detector simulators do not yet include the information required for dealing with the above signatures. As a result, dedicated recasting tools must be developed to deal with LLP searches. In this note we mainly make use of MADGRAPH5 [95] and PYTHIA [97, 145] to simulate hadron level events and no fast detector simulation is employed. Several approaches are then discussed in order to emulate the experimental selection and reconstruction efficiencies and reproduce the official exclusion curves presented in the corresponding analyses.

The recasting of the ATLAS displaced vertex search is presented in Section 2, where two methods are employed in order to reproduce the exclusion curves for a simplified long-lived gluino scenario. While the first method makes use of the limited information provided by the corresponding conference note, the second uses the full efficiencies provided by the ATLAS auxiliary material. The goal of comparing these two approaches is to illustrate how lack of experimental information drastically decreases the recasting performance. Sections 3 and 4 discuss searches based on isolated tracks, which, in principle, are much simpler to recast. Although CMS has provided detailed efficiencies for the 8 TeV searches, we will show that these can not be easily extrapolated to the 13 TeV results. We illustrate this by recasting two CMS searches — displaced lepton search (Section 3) and the long-lived charged particle search (Section 4). Finally, in Section 5 we present the overall conclusions and recommendations drawn from the recasting of these particular searches.

2 DISPLACED VERTEX SEARCH

The ATLAS displaced vertex + missing energy analysis presented in Ref. [598] investigate an important BSM scenario: long-lived gluinos decaying to jets and missing energy. The analysis searches for displaced vertices (DV) in association with large missing transverse energy, a signature present in several BSM models (long-lived stops, hidden valley scenarios and others). Hence it is relevant to investigate how well it is possible to recast this analysis and extend its constraints to other LLP models. One important feature of the ATLAS search is that it does not rely on the gluino (or R-hadron) charge, so it can be directly applied to both charged and neutral LLPs.

[1]The missing energy signature is usually not classified as a LLP search, since it is covered by several prompt searches. Nonetheless, we include it here for completeness.

In Ref. [598] the results are interpreted in the long-lived gluino scenario with R-Parity conservation, the lightest neutralino being the LSP and the gluino the NLSP. Hence, after being pair produced, the gluinos hadronize and then decay to jets and the LSP with a 100% branching ratio:

$$pp \to \tilde{g}\tilde{g} \to (jj\tilde{\chi}_1^0) + (jj\tilde{\chi}_1^0) \tag{1}$$

Limits are presented for gluino lifetimes between 0.02 ns and 20 ns, gluino masses of 1.4 and 2 TeV and several values of $m_{\tilde{\chi}_1^0}$.

The search imposes the following criteria for selecting displaced vertices:

1. Missing energy selection in the event: $E_{\rm T}^{\rm miss} > 250$ GeV.
2. *Base vertex selection* (at least one DV in the event):
   - The DV coordinates must satisfy: $R_{\rm DV} = \sqrt{x^2 + y^2} < 300$ mm and $|z_{\rm DV}| < 300$ mm.
   - The DV must not fall into a material rich area. This criterion corresponds to discarding approximately 42% of the fiducial volume.
   - The vertex must be separated by more than 4 mm from all primary vertices.
3. Signal region selection:
   - the invariant mass of the DV must be $m_{\rm DV} > 10$ GeV, assuming all its tracks have the pion mass.
   - $n_{\rm tracks} \geq 5$, where $n_{\rm tracks}$ corresponds to the number of tracks originating from the vertex and satisfying: $p_{\rm T} > 1$ GeV and $|d_0| > 2$ mm.

After applying the above selections, no displaced vertices were observed in 32.1 fb^{-1} of data at $\sqrt{s} = 13$ TeV. The number of expected background displaced vertices is:

$$N_{\rm DV}^{\rm BG} = 0.02 \pm 0.02 \tag{2}$$

This analysis is particularly interesting because it provides a large set of information useful for recasting.[2] Specially useful are the efficiency grids provided for DV reconstruction and event selection as a function of the relevant (truth level) variables: number of tracks, DV mass, DV position and missing energy. We point out that these detailed efficiencies are usually not provided for most LLP searches. The corresponding 8 TeV search in Ref. [147], for instance, only provided reconstruction and event-level efficiencies as a function of a single parameter. Therefore, we will discuss below how recasting performs when distinct levels of information are available. In particular, we will discuss two approaches:

- *Method 1*: recasting using correction functions for the vertex reconstruction efficiency and the track efficiency.
- *Method 2*: recasting using the efficiency grids provided by ATLAS in Ref. [598].

[2]We point out that the conference note ATLAS-CONF-2017-026, which has been superseded by Ref. [598], included only a subset of the information present in the publication.

2.1 Method 1: Recasting using correction functions

The first approach only makes use of the displaced vertex reconstruction efficiency provided by ATLAS and shown as red points in Fig.1. We point out that this data is provided only for $m_{\tilde{g}} = 1.2$ TeV and $\tau = 1$ ns. The method discussed below tries to construct functions which aim to approximate the experimental track and vertex reconstruction efficiencies. These will then be applied to the event selection and used to compute upper limits for the same gluino scenarios considered by ATLAS.

In order to recast the ATLAS displaced vertex + missing energy search, we use PYTHIA 8 [97] for the simulation of gluino production, hadronization and decays. No smearing is applied and the truth-level $E_{\mathrm{T}}^{\mathrm{miss}}$ is considered. All charged particles generated by the R-hadron decay are considered as potential charged tracks and only these are included when computing m_{DV}. Furthermore, d_0 is calculated with the assumption of a zero magnetic field. The veto of decays in material rich area is implemented simply as an overall fiducial volume cut. Since a dedicated algorithm was used by the collaboration to identify and select DV candidates, it can not be easily reproduced. However the vertex reconstruction efficiency as a function of R_{DV} has been provided in Ref. [598] for $\left(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}, \tau\right) = (1200\ \mathrm{GeV}, 100\ \mathrm{GeV}, 1\ \mathrm{ns})$ and $\left(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}, \tau\right) = (1200\ \mathrm{GeV}, 1170\ \mathrm{GeV}, 1\ \mathrm{ns})$.

As a first step in the recasting procedure, we test how well the vertex reconstruction efficiency (ϵ_{DV}) can be reproduced under the assumption that all displaced vertices satisfying the *base vertex selection* cuts are reconstructed. The efficiencies provided by ATLAS and the result obtained from recasting are shown in Fig.1. As we can see, ϵ_{DV} decreases with R_{DV}, except for the first bin (0 mm $< R_{\mathrm{DV}} <$ 5 mm). The slope seen in the recasting curve is purely due to the *base vertex selection* cuts applied. The same behavior is seen in the official data, although the efficiency falls much faster, likely due to detector effects and the reconstruction algorithm. For the low bins (except for the first one) the two curves differ by $\simeq 20\%$, while for the largest bins the difference is almost an order of magnitude. The high level of agreement in the first bin ($R_{\mathrm{DV}} <$ 5 mm) is an artificial effect due to the requirement that the displaced and primary vertices must be separated by at least 4 mm.

As shown in Fig.1, the reconstruction efficiency obtained by recasting can be overestimated by more than an order of magnitude for high values of R_{DV}. One is then tempted to directly apply the vertex reconstruction efficiencies provided by ATLAS (red points in Fig.1) to the hadron level events generated in PYTHIA. However these have been derived for the benchmark point $m_{\tilde{g}} = 1.2$ TeV, $m_{\tilde{\chi}_1^0} = 100$ GeV, $\tau = 1$ ns and are not necessarily valid for other values of the gluino and LSP masses or the lifetime. In fact, as shown in Ref. [598], the efficiencies can be affected by the mass difference $m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$. We have also computed the reconstruction efficiency using distinct values of the gluino lifetime. The results are shown in Fig.2. Despite the fluctuations at high R_{DV} (due to limited statistics), we can see that the efficiency also has a strong dependence on the gluino lifetime. Therefore we conclude that the ATLAS efficiencies shown in Fig.1 can not be directly applied to any input model. Also we can not neglect the large impact of detector response and the reconstruction algorithm, as illustrated by the difference between the blue and red points in Fig.1. In order to proceed with the recasting (using only the data from Fig.1) we will adopt the following *strategy*:

i. Apply the *base vertex selection* cuts.

ii. Rescale the DV reconstruction efficiencies obtained from the vertex selection cuts above by a factor $r\,(R_{\mathrm{DV}})$. This "correction factor" (r) is defined by the ratio of the red and blue

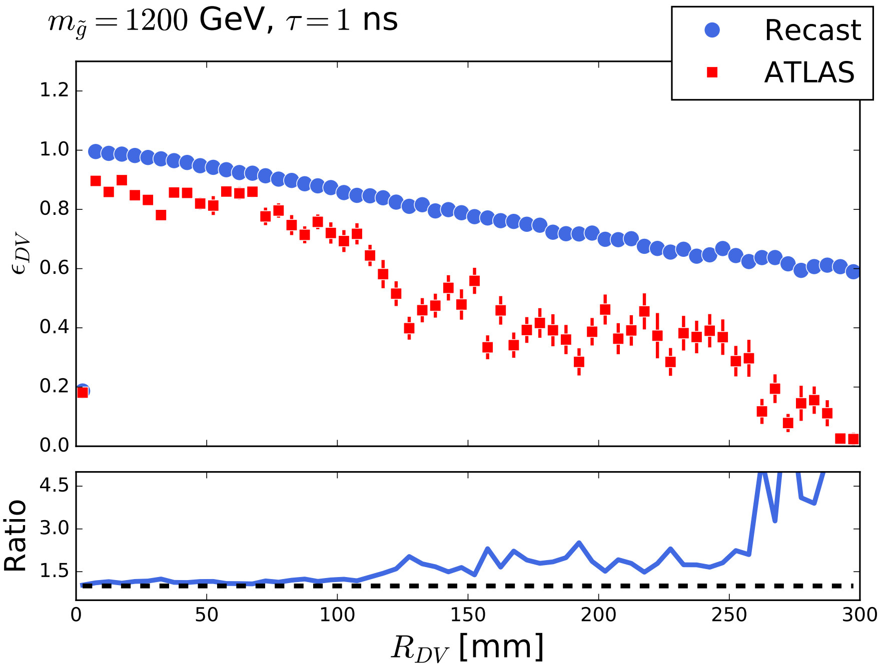

Figure 1: Vertex reconstruction efficiency as a function of the displaced vertex transverse position, $R_{\rm DV} = \sqrt{x^2 + y^2}$. The red points correspond to the results obtained by ATLAS in Ref. [598], while the blue points correspond to the results obtained through recasting after the *base vertex selection* cuts have been applied. The results refer to the benchmark point $\left(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}, \tau\right) = (1200\,\text{GeV}, 100\,\text{GeV}, 1\,\text{ns})$.

points in Fig.1. This factor aims to encapsulate the experimental features which are not captured by the base vertex selection cuts and is assumed to be model independent.

iii. Apply a constant track efficiency[3], which aims to approximate how many of the truth level charged tracks are actually reconstructed at detector level. This efficiency impacts the event selection efficiency, since events are required to have $n_{\rm tracks} \geq 5$.

iv. Apply the missing energy and signal region cuts to the surviving events.

With the above procedure we aim to capture the impact of detector effects and the reconstruction algorithm in a model independent way. The validity of this approach clearly relies on strong assumptions about the detector performance and the relation between truth level and detector level observables. However, without further information we believe it is not possible to significantly improve the recasting. Once we apply all the cuts and the correction factor to the events, it is possible to compute signal efficiencies for any input model. These efficiencies can then be used to extract 95% CL upper limits on the total visible cross section using the number of observed events ($N_{\rm obs} = 0$) and the expected background ($N_{\rm DV}^{\rm BG} = 0.02 \pm 0.02$). For reference, a 100% efficiency corresponds to the upper limit $\sigma_{\tilde{g}\tilde{g}} < 0.091$ fb.

[3]The track efficiency clearly depends on the track $p_{\rm T}$ and production position, hence a constant efficiency is an oversimplification. We have tried distinct functional forms for $\epsilon_{\rm track}$ following a procedure similar to the one described in Refs. [600, 601]. None of these, however, perform much better (for all values of τ) than the results presented here assuming a constant efficiency.

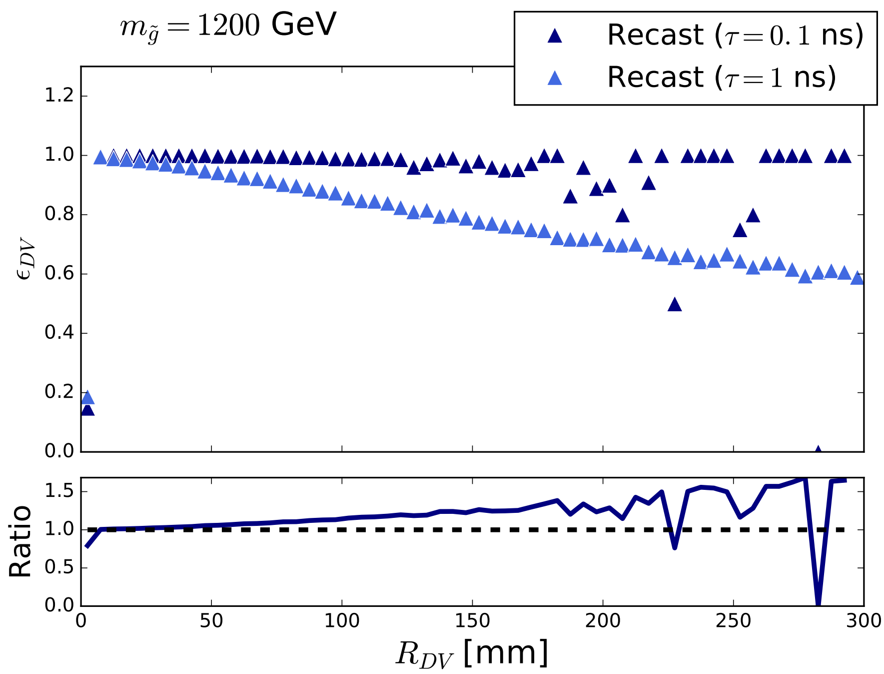

Figure 2: Vertex reconstruction efficiency as a function of the displaced vertex transverse position ($R_{\rm DV}$). The results were obtained through recasting after the *base vertex selection* cuts have been applied. The plot compares the recasting results for a fixed gluino and LSP mass ($m_{\tilde{g}} = 1200$ GeV, $m_{\tilde{\chi}_1^0} = 100$ GeV) and two lifetime values, $\tau = 1$ ns and $\tau = 0.1$ ns.

The results obtained by the above approach for the benchmark points $m_{\tilde{g}} = 1.4$ TeV and $m_{\tilde{g}} = 2$ TeV are shown in Fig. 3. We present recasting curves for three distinct values of the (constant) track efficiency: $\epsilon_{\rm track} = 15\%, 25\%, 100\%$. As we can see, a 25% efficiency can reproduce the official exclusion curve (solid black line) within $\sim 50\%$ for most of the lifetime values. However, in the regions where the efficiency drops significantly ($\tau < 10^{-2}$ ns), the recasting curve is wrong by more than an order of magnitude. Therefore, we conclude that this procedure is not satisfactory for a general purpose recasting of the search presented in Ref. [598]. In the next section we discuss how the recasting improves once we include the detailed efficiencies provided by ATLAS.

2.2 Method 2: Recasting using ATLAS efficiency grids

This approach makes use of the full information provided in the auxiliary material[4] of Ref. [598]: efficiency grids for the event selection (as a function of $R_{\rm DV}$ and $E_{\rm T}^{\rm miss}$) and for the vertex reconstruction efficiency (as a function of $R_{\rm DV}$, $m_{\rm DV}$ and $n_{\rm track}$). These parametrized efficiencies are also given for different regions in the detector, encapsulating the effect of the material veto cut.

According to the note provided by ATLAS, these efficiencies can be applied at truth level once some fiducial cuts have been applied. We use truth level missing energy and identify the truth $R-$hadron decay position and decay products.

[4]This material can be directly access from:
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/SUSY-2016-08/hepdata_info.pdf

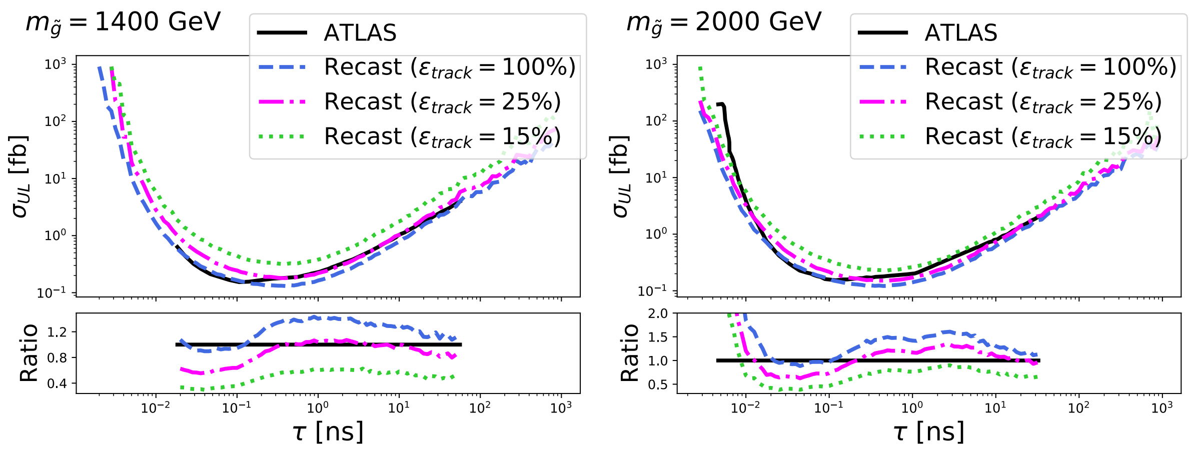

Figure 3: Comparison between the official exclusion curve and the one obtained by recasting the analyses with Method 1 for two values of gluino masses.

The selection of events requires:

- truth level missing energy $E_T^{\text{miss}} > 200$ GeV.
- one trackless jet with $p_T > 70$ GeV, or two trackless jets with $p_T > 25$ GeV. A trackless jet is defined as a jet for which the scalar sum of the p_T of all charged particles inside the jet does not exceed 5 GeV. These jet requirements are applied to 75% of the data. The remaining 25% do not need to satisfy any jet cuts.

In addition, events must have at least one displaced vertex with:

- distance between the interaction point and the decay position > 4 mm.
- the decay position must lie in the fiducial region $R_{\text{DV}} < 300$ mm and $|z_{\text{DV}}| < 300$ mm.
- the number of selected decay products must be at least 5, where selected decay products are charged and stable, with $p_T > 1$ GeV and $|d_0| > 2$ mm.
- the invariant mass of the truth vertex must be larger than 10 GeV, and is constructed assuming all decay products have the mass of the pion.

After imposing the above fiducial cuts, the vertex reconstruction and event selection efficiencies provided by ATLAS can then be applied to compute the final signal efficiencies. The results for the two benchmark points are shown in Fig. 4. Finally, using these efficiencies, we can extract 95% CL upper limits on the total visible cross section using the same procedure described in the previous Section. The results for the exclusion curves are shown in Fig. 5, where we can see that the recasting reproduces the official exclusion curves fairly well for most of the lifetime values. The largest discrepancies are within $\sim 40\%$, which corresponds to a major improvement with respect to the results obtained using only the limited information provided by the ATLAS conference note, as discussed in Section 2.1.

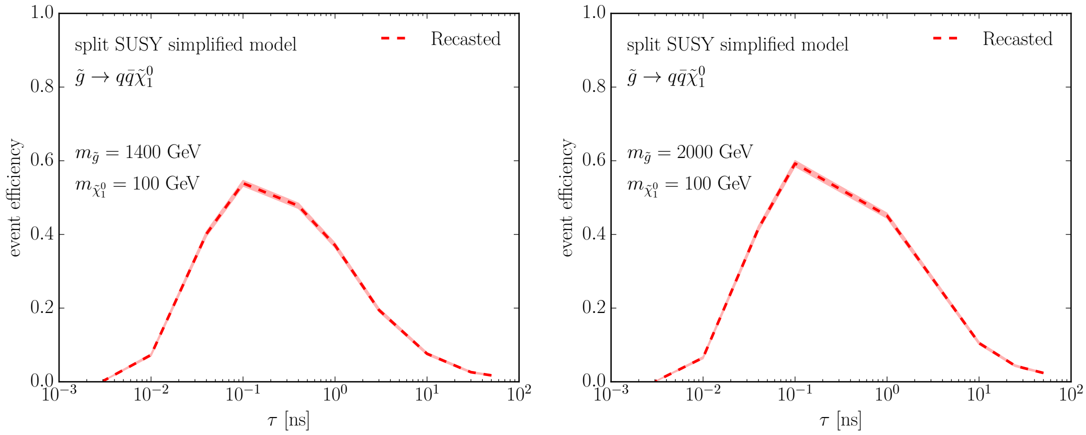

Figure 4: Recasted event-level efficiencies against gluino proper decay lifetime for two values of gluino masses.

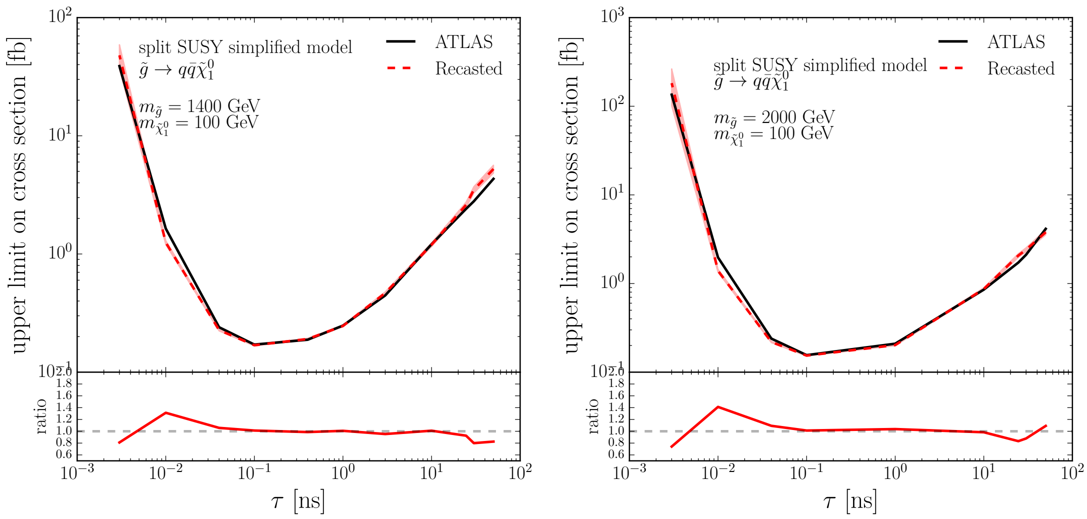

Figure 5: Comparison between the official exclusion curve (solid black line) and the one obtained by recasting the analyses with Method 2 (dotted red line) for two values of gluino masses.

2.3 Comments on the recasting procedure

The results presented in Fig.3 show that it is not possible to recast the ATLAS displaced vertex + missing energy search making use only of the simple vertex reconstruction efficiencies provided in the conference note ATLAS-CONF-2017-026. Although different correction functions can be applied at the vertex and track level to try to capture the impact of detector effects and the vertex reconstruction algorithm, we find that the recasting of the exclusion curves is still off by an order of magnitude at small lifetimes. On the other hand, using the ATLAS efficiency grids highly improves the level of agreement, limiting the discrepancies to be under $\sim 40\%$, as shown by Fig.5. We therefore find that the parametrization of the efficiency grids in terms of truth level variables is extremely useful for recasting, being also straightforward to implement.

Finally, we note that, even though the parametrized selection efficiencies can be in principle used for any model (and are said to be model independent according to the ATLAS note), we encourage the experimental collaborations to present exclusion curves (or signal efficiencies) for a second model with a distinct event topology. This information is essential to validate the recasting procedure and accurately assess the level of model independence provided by the parametrized efficiencies.

3 DISPLACED LEPTON SEARCHES

One of the cleanest search strategies for long-lived particles decaying into leptons is to look for leptons with a non-zero impact parameter with the primary vertex. Such tracks are normally vetoed to remove contamination from underlying event, cosmics etc., and therefore may not be attributed to the right collision event in standard prompt searches. The CMS displaced lepton search ($\sqrt{s} = 8$ TeV, $\mathcal{L} = 19.7 \pm 0.5\text{fb}^{-1}$) [599] follows a simple strategy of requiring two isolated, oppositely charged, lepton tracks that have a significant impact parameter with respect to the primary vertex. The accompanying material includes efficiency for identification of electrons and muons based on the transverse impact parameter (d_0) and the transverse momentum (p_T) of the lepton. The benchmark scenario for the analysis is stop-pair production $pp \to \tilde{t}\tilde{t}^*$ ($m_{\tilde{t}} = 500$ GeV) followed by an R-parity violating decay via λ'-type coupling $\tilde{t} \to b\ell$, with equal probabilities for $\ell = e, \mu$ and τ.

The recommended recasting procedure is to apply the cuts below on generator-level leptons and reweight the event with the four identification efficiencies and an overall trigger efficiency of 0.95.

1. Select events with one e and one μ, oppositely charged, both coming from a stop
2. Require both leptons to have $|\eta| < 2.5$
3. Require decay vertex of stop to have transverse position $v_0 < 40$ mm, and z-position $v_z < 300$ mm.
4. Require $p_T^{(e,\mu)} > 25$ GeV and $\Delta R_{e\mu} > 0.5$
5. Jet isolation: for each jet (anti-kt, $R = 0.5$, $p_{\mathrm{T}}^{\mathrm{min}} = 10$ GeV), require $\Delta R_{\ell j} > 0.5$
6. Transverse impact parameter 0.1 mm $< d_0 < 20$ mm.
7. Signal regions are further defined as:

   SR3: Both leptons satisfy 1.0 mm $< d_0 < 20$ mm.

   SR2: One or both leptons fail SR3 but satisfy $d_0 > 0.5$ mm

   SR1: One or both leptons fail SR2 but satisfy $d_0 > 0.2$ mm

The updated analysis at 13 TeV ($\sqrt{s} = 13$ TeV, $\mathcal{L} = 2.6\text{fb}^{-1}$, conference note CMS-PAS-EXO-16-022) [120], applies stronger p_{T} cuts by requiring $p_T^{e(\mu)} > 42(45)$ GeV. It also improves isolation cuts on the leptons by requiring that the sum of p_{T} (of all particles) in a cone of 0.3 should be less than 3.5% (6.5%) in the barrel (encdap) for electrons. For muons, cone size is taken to be 0.4 and the sum of p_T is required to be less than 15%. Separation between the two leptons is $\Delta R_{e\mu} > 0.5$, same as before. The mass of the stop in the benchmark point is increased to 700 GeV, however, the decay branching fractions remain unchanged.

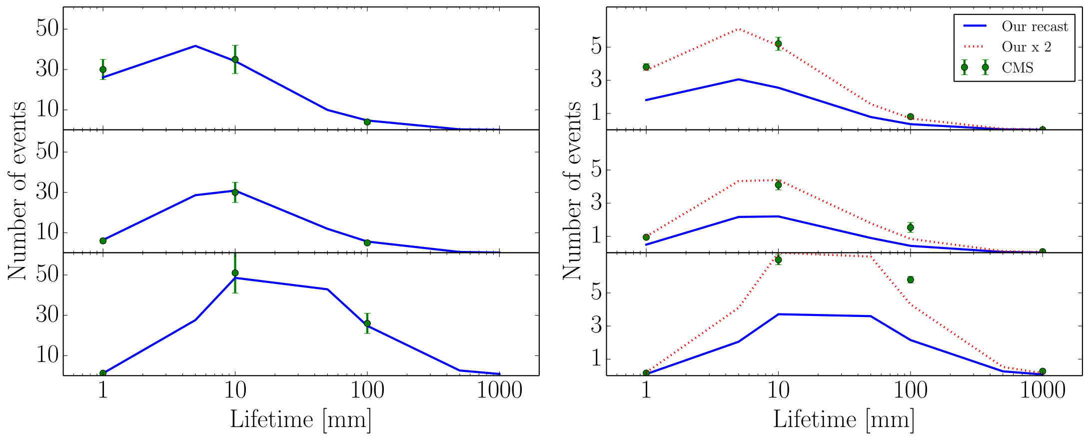

Figure 6: Result from recasting the 8 TeV (left) and 13 TeV (right) displaced lepton analysis comparing number of expected signal events from our simulation (blue) with the published CMS values (green dots). The panels refer to SR1 (top), SR2 (middle) and SR3 (bottom).

A comparison of our validation of the analysis can be seen in Fig. 6. At 8 TeV, the agreement is well within the quoted errors. Using the same efficiency parametrization as 8 TeV for the 13 TeV analysis, we are unable to reproduce the benchmark efficiencies and we find that in general, we have an overall mismatch of a factor of about 2 with respect to the published expected number of signal events. Moreover, we find that the mismatch is higher for longer lifetimes, prompting the inference that the d_0 dependence of the efficiency may have changed between the two runs.

4 CHARGED HEAVY PARTICLE SEARCH

Another important example of LLPs which have been searched for at the LHC are heavy stable charged particles (HSCPs). Here we will consider the searches performed by CMS using 8 TeV [121] and 13 TeV [120] data. These searches are targeted at detector-stable HSCPs and are based on the signature of highly ionizing tracks and anomalous time-of-flight. Below we will discuss the recasting of both of these searches. We point out, however, that while the 8 TeV analysis provided detailed efficiencies for the HSCP reconstruction, the 13 TeV results do not include this information. Therefore, for the 13 TeV search the recasting will make use of an extrapolation of the 8 TeV efficiencies.

4.1 Recasting the 8 TeV search

As mentioned above, the 8 TeV CMS analysis [121] provides efficiency grids for HSCP reconstruction. In particular, probabilities for events to pass the on- and off-line selection criteria are given as a function of the truth-level HSCP kinematics. Due to the inclusive nature of the search, this provides a powerful way to reinterpret the search for arbitrary models containing detector-stable [110, 126, 602–604] or metastable [605] LLPs.

The recasting follows the procedure described in Ref. [121]. We simulate events for HSCP production using MadGraph5 [95] for the parton level process and Pythia 6 [145] for showering and hadronization. For each event we first identify isolated HSCP candidates imposing the following isolation criteria:

$$\left(\sum_{i}^{\substack{\text{charged}\\ \Delta R<0.3}} p_{\mathrm{T}}^{i} \right) < 50\,\mathrm{GeV}\ , \qquad \left(\sum_{i}^{\substack{\text{visible}\\ \Delta R<0.3}} \frac{E^{i}}{|\vec{p}|} \right) < 0.3\,, \tag{3}$$

on the truth level events. The sums include all charged and visible particles, respectively, in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.3$ around the direction of the HSCP candidate, p_T^i denotes their transverse momenta and E^i their energy. Muons are not considered as visible particles and the HSCP candidate itself is not included in either sum. Once the HSCP candidates are identified, the final signal efficiency (ϵ) is then given by:

$$\epsilon = \frac{1}{N} \sum_{i}^{N} P_{\text{on}}(\vec{k}) \times P_{\text{off}}(\vec{k})\,, \tag{4}$$

where the sum runs over all generated events and P_{on} (P_{off}) is the on-line (off-line) selection efficiency provided by the 8 TeV CMS analysis as a function of the HSCP truth level kinematics ($\vec{k}$). For events containing two HSCP candidates, the above probabilities must be replaced by [121]

$$P^{(2)}_{\text{on/off}}(\vec{k}^1, \vec{k}^2) = P_{\text{on/off}}(\vec{k}^1) + P_{\text{on/off}}(\vec{k}^2) - P_{\text{on/off}}(\vec{k}^1) P_{\text{on/off}}(\vec{k}^2)\,, \tag{5}$$

where $\vec{k}^{1,2}$ are the kinematical vectors of the HSCPs.

In order to validate the above procedure we compute exclusion curves for the same benchmark models considered by the 8 TeV CMS search [126]. The results are shown in Figure 7 for the the gauge-mediated supersymmetry breaking (GMSB) model containing long-lived staus. As we can see both the signal efficiencies (left panel) and the 95% CL upper limits on the inclusive production cross section (right panel) agree with the official CMS results within 5% or less, thus providing an excellent approximation.

4.2 Recasting the 13 TeV search

Given the excellent performance of the recasting method discussed above for the 8 TeV LHC run it would be appealing to be able to use a similar method for the 13 TeV analysis [120]. However, the provision of object efficiencies for the 13 TeV search is so far not pursued. In this section we discuss an attempt to recast the 13 TeV search based on an extrapolation of the object efficiencies from the 8 TeV to the 13 TeV run.

Our aim is to use the 8 TeV efficiencies ($P^{8\,\text{TeV}}_{\text{on/off}}$) for the 13 TeV recasting through the introduction of a correction function that accounts for the differences between both runs. Since the selection criteria for the 8 TeV and 13 TeV searches are very similar, we require the same isolation cuts listed in Eq. 3. We then assume the following ansatz for the 13 TeV efficiencies ($P^{13\,\text{TeV}}_{\text{on/off}}$):

$$P^{13\,\text{TeV}}_{\text{on}}(\vec{k}) = P^{8\,\text{TeV}}_{\text{on}}(\vec{k})\ , \ P^{13\,\text{TeV}}_{\text{off}}(\vec{k}) = F(\beta) \times P^{8\,\text{TeV}}_{\text{off}}(\vec{k})\,, \tag{6}$$

where F is the correction function. The above relations assume that the on-line probability does not change drastically and the main differences between the runs happen in the off-line (trigger) selection. This is a reasonable assumption, since the on-line cuts of both analyses are similar. Furthermore, we expect our treatment to leave enough freedom to account for small corrections in the on-line probability as well. At least for events with only one HSCP candidate there is no

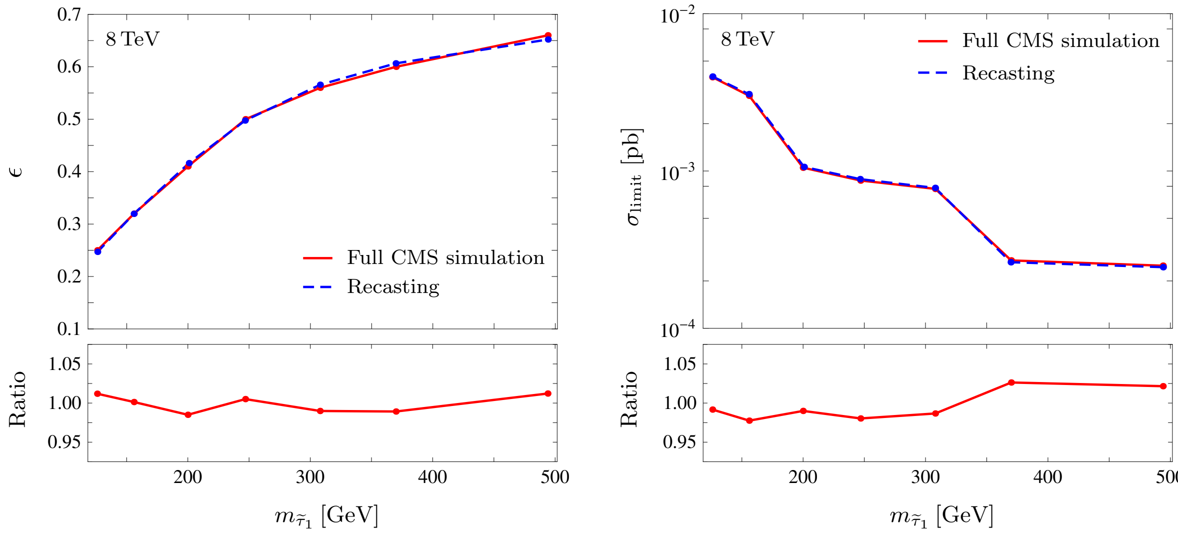

Figure 7: Signal efficiency ϵ (left panel) and 95% CL cross section upper limit (right panel) for the GMSB model as the function of the stau mass. We compare the CMS analysis [121] from the full detector simulation (red solid lines) with the recasting from [121] (blue dashed lines). In the lower frames we show the respective ratios $\epsilon^{\text{Full}}/\epsilon^{\text{Recast}}$, $\sigma_{\text{limit}}^{\text{Full}}/\sigma_{\text{limit}}^{\text{Recast}}$. Taken from [126].

distinction, as the two probabilities are just multiplied. Finally, we assume that F only depends on the HSCP velocity (β) and we parametrize this correction function by eight parameters (C_{β_n}), which corresponds to the value of the function for $\beta = 0, 0.47, 0.6, 0.7, 0.77, 0.83, 0.89, 1.0$. For other values of β we interpolate linearly.

In order to compute the correction function F introduced above we will make use of the official signal efficiencies reported in 13 TeV LHC analysis [120] for specific benchmark points. We determine the parameters C_{β_n} of the correction function in a global fit to these efficiencies using the χ^2 defined as:

$$\chi^2(C_{\beta_n}) = \sum_m \frac{(\epsilon_m(C_{\beta_n}) - \epsilon_m^{\text{CMS}})^2}{\sigma_\epsilon^2}\,, \tag{7}$$

where $\epsilon_m(C_{\beta_n})$ is the efficiency for the benchmark point m using the correction function parameters C_{β_n}, ϵ_m^{CMS} is the respective efficiency reported in [120] and σ_ϵ is the characteristic size of the uncertainty which we (arbitrarily) set to 0.02. We include in the fit the 6 direct stau benchmark points used in the CMS analysis and minimize the χ^2 using MULTINEST [606, 607] for an efficient exploration of the parameter space. The best-fit correction function, $F_{\text{best-fit}}$, and its 1σ uncertainty[5] are shown in figure 8. The deviation of F from 1 implies a decrease or increase of the respective detector and signal efficiencies between the 8 and 13 TeV analysis. We find a slight decrease of efficiencies for large velocities $\beta \gtrsim 0.85$, which is, however, not significant. More surprisingly, the efficiencies at low velocities ($\beta \simeq 0.5$) contain a large and significant increase. From these results it appears that the CMS detector in Run 2 performs significantly better at low velocities. We also point out that the uncertainties for low values of β are quite large, which illustrate the fact that the signal efficiencies provided by the 13 TeV

[5] As stated above the 1σ uncertainty corresponds $\sigma_\epsilon = 0.02$ which is roughly the level of accuracy we aim at in the fit. Note, however, that intrinsic systematic uncertainties in the determination of the efficiencies might be larger, up to around 10%.

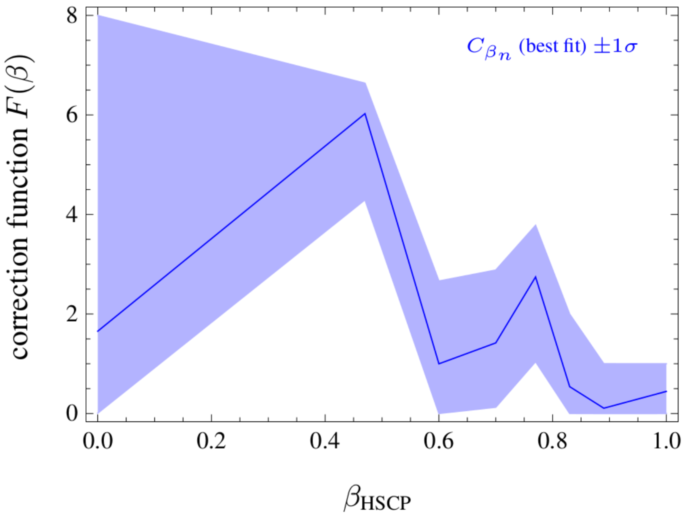

Figure 8: Best-fit correction function $F(\beta)$ and its $\pm 1\sigma$ band.

CMS results are not sufficient to fully constraint the correction function F. However, since no additional information is provided by the 13 TeV CMS analysis, we will use the best fit for F in the following.

Using the off- and on-line probabilities defined in Eq. 6 and the best fit for the correction function ($F_{\text{best-fit}}$) shown in Fig. 8 we can then compute the 13 TeV signal efficiencies for any input model. We first compute efficiencies for the same 6 benchmark points used to obtain F. These benchmarks correspond to direct production of long-lived status with distinct values of the stau mass. The results for the final signal efficiencies for these six benchmark points are shown in Fig. 9 as a function of the HSCP (stau) mass. The solid black line shows the efficiencies reported by CMS in Ref. [120], while the dashed blue curve shows the efficiencies obtained by recasting using the extrapolation of the 8 TeV probabilities and the best fit for F. We also show the recasting efficiencies obtained without the inclusion of the correction function (dashed magenta curve). As expected, $F_{\text{best-fit}}$ reproduces the efficiencies for the six mass points well within the expected fit uncertainties (2%). In particular, it significantly improves the agreement with respect to the naive extrapolation without a correction, *i.e.* for $F = 1$. Formally, in our fit this is reflected in a decrease in the χ^2 from around 130 for $F = 1$ to 0.1 for $F_{\text{best-fit}}$.

The good agreement shown in Fig. 9 between the official CMS results and $F_{\text{best-fit}}$ is expected, since the same benchmark points were used to fit the correction function F. Therefore, a crucial test of the validity of the recasting is its application to an independent set of models. Fortunately, in Ref. [120] CMS has also reported the signal efficiencies for six GMSB points with long-lived staus. These points include both direct production of staus and production through cascade decays of heavier sparticles, resulting in a broader spectrum of event topologies. Using again the extrapolation of the 8 TeV efficiencies and the best fit for the correction function from Fig.8, we compute the final signal efficiencies for these six GMSB points. The results are shown in Fig. 10. As we can see, the efficiencies using $F_{\text{best-fit}}$ now deviate by up to 20% for large HSCP masses, where our recasting undershoots the efficiencies reported by CMS. Although the overall agreement is improved by the correction function, $\chi^2(F_{\text{best-fit}}) \simeq 88$ versus $\chi^2(F = 1) \simeq 240$, the result is much worse than the ones obtained at 8 TeV, where the uncertainties were below

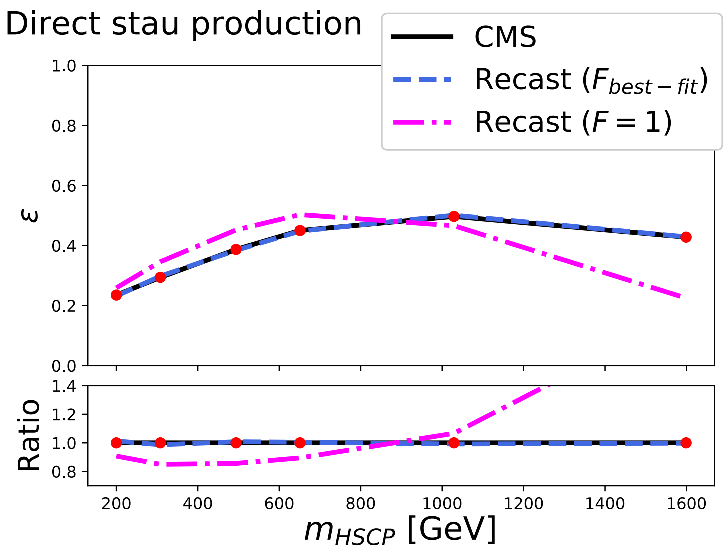

Figure 9: Efficiencies for the six benchmark masses in the direct stau production scenario. The solid black curve shows the efficiencies from [120], while the dashed curves show the efficiencies obtained through recasting. The blue curve corresponds to the best-fit for the correction function, while the magenta shows the efficiencies for $F = 1$.

5%.

The differences between the recasting efficiencies and the ones reported by CMS for the GMSB model might arise from several shortcomings in our recasting procedure. First, we assume F to only dependent on β whereas the full probability maps are parametrized as a function of three kinematic variables: β, p_{T} and $|\eta|$. A parametrization in terms of all three variables, however, is clearly not feasible given the limited amount of information provided in Ref. [120]. As illustrated by Fig. 8, even the parametrization in terms of a single variable can not be fully constrained using only the signal efficiencies reported by CMS. Second, assigning the correction function to the off-line probabilities only might be an over-simplification. For events containing two HSCP candidates, a single correction function F may not be sufficient to parametrize the differences between the two runs. Again, only a better understanding of the underlying changes (*e.g* in the trigger settings) can resolve these ambiguities.

4.3 Comments on the recasting procedure

As illustrated by the results in Section 4.1, the uncertainties in the recasting of HSCP searches can be reduced to the few percent level if efficiencies for the reconstruction and selection of HSCP tracks are provided. This is the case for the 8 TeV CMS analysis, where we were able to successfully recast the HSCP results with a high degree of accuracy [126]. The situation is drastically distinct if these officiencies are not provided by the experimental collaboration, as illustrated by the 13 TeV results in Section 4.2. Even though the CMS analyses for both runs are very similar, the 8 TeV object efficiencies cannot be easily extrapolated to the run 2. Therefore, in order to establish robust reinterpretations of HSCP searches, further information on detector efficiencies are required.

We also point out that the recasting uncertainties for the 13 TeV search could only be properly

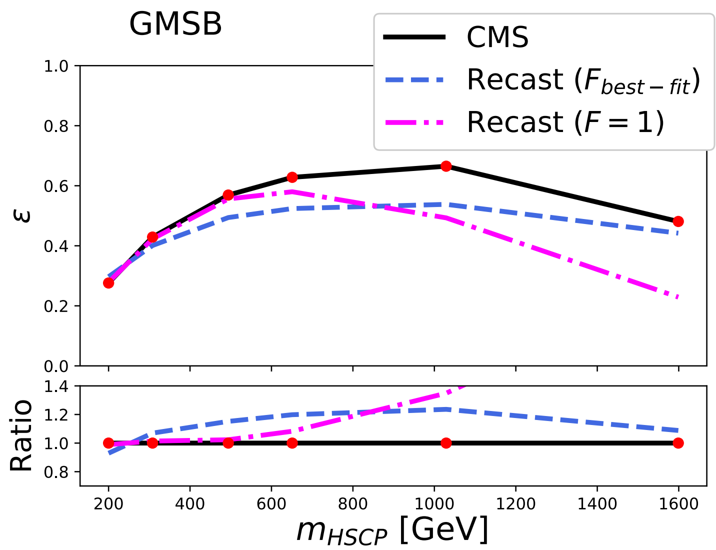

Figure 10: Efficiencies for the six benchmark masses in the GMSB scenario. The solid black curve shows the efficiencies from [120], while the dashed curves show the efficiencies obtained through recasting. The blue curve corresponds to the best-fit for the correction function, while the magenta shows the efficiencies for $F = 1$.

assessed using the official CMS signal efficiencies for two distinct scenarios (the direct stau and GMSB models). This is an important point, since several LLP searches present results for a single model or type of signal topology (see Section 2, for instance). The second model is essential for estimating the validity of the recasting procedure and if it can indeed be extrapolated to other models containing distinct event topologies. As shown by the results in Figs. 9 and 10, the recasting uncertainties would have been highly underestimated if the CMS values for the GMSB scenario were not available.

5 CONCLUSIONS

In this note we have investigate the feasibility of recasting LLP searches within the context of three distinct signatures: displaced vertices, displaced leptons and charged tracks. Although each signature presents its own challenges for recasting, we showed that without detailed object reconstruction and selection efficiencies a satisfactory recasting can not be performed outside the experimental collaborations. The final signal efficiencies or limits in this case can be inaccurate by almost an order of magnitude. On the other hand, for the cases where the object efficiencies were available, such as the ATLAS displaced vertex and the 8 TeV CMS displaced lepton and HSCP searches, we were able to reproduce the official results within 5% to 40%.

In summary, our overall conclusions regarding the relevant information required for a proper recasting of LLP searches are as follows:

1. *Object efficiencies* for reconstruction of the LLP signature must be provided in terms of the relevant truth level observables. As discussed above, without these efficiencies the recasting uncertainties can be over 100%.

2. *Cut-flow tables for expected signal events* would greatly improve the ability to recast these

searches. In the event of a mismatch in the calculation of the final expected signal yield, it is at the moment impossible to pinpoint the source of the problem.

3. *Limits for at least two models (or signal topologies)* should be published so a sanity check can be made before using object efficiencies on a different model. Current standard practice has been to provide object efficiencies with respect to Monte-Carlo truth-level objects for the benchmark model. It is therefore still difficult to understand how well these can be trusted when applied to models or topologies other than the benchmark (e.g. see Figs. 9 and 10).

ACKNOWLEDGEMENTS

This research was supported in part by the São Paulo Research Foundation (FAPESP), projects 2015/20570-1 and 2016/50338-6. GC acknowledges support by the Ministry of Science and Technology of Taiwan under grant No. MOST-106-2811-M-002-035. JH acknowledges support by the German Research Foundation DFG through the research unit "New physics at the LHC". ND acknowledges the support of the OCEVU Labex (ANR-11-LABX-0060) and the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French government program managed by the ANR.