First steps towards the reconstruction of the squark flavour structure

Jordan Bernigaud, Björn Herrmann
Univ. Grenoble Alpes, USMB, CNRS, LAPTh, F-74000 Annecy, France
E-mail: bernigaud@lapth.cnrs.fr, herrmann@lapth.cnrs.fr

ABSTRACT: Assuming the observation of a squark at the Large Hadron Collider, we investigate methods to access its flavour content and gain information on the underlying flavour structure of the theory. Based on simple observables, we construct a likelihood inference method to determine the top-flavour content of the observed particle. In addition, we employ a multivariate analysis in order to classify different flavour hypotheses. Both methods are discussed within a simplified model and the more general Minimal Supersymmetric Standard Model. We conclude that the likelihood inference provides an estimation of the top-flavour content if additional knowledge on the gaugino sector is available, while the multivariate analysis identifies different flavour patterns and can accommodate a more minimalistic set of observables.
1 Introduction

One of the main goals of the Large Hadron Collider (LHC) is the quest for signals of physics beyond the Standard Model of particle physics. Currently in its Run-2, however, no direct evidence pointing towards the existence of new states has been found. Among the numerous extensions addressing the shortcomings of the Standard Model, Supersymmetry ranks among the most attractive solutions. However, if Nature is indeed supersymmetric, it is reasonable to assume that its exact realization is situated beyond the “vanilla” Minimal Supersymmetric Standard Model (MSSM) or its simplified realizations which are typically searched for in current experimental analyses [1–15]. To give an example, experimental studies typically are based on the Minimal Flavour Violation (MFV) paradigm, which is based on the assumption that all flavour-violating interactions stem from the Yukawa couplings alone, as it is the case in the Standard Model. However, there is no apparent reason that this paradigm is respected beyond the Standard Model. In the MSSM, additional flavour-violating terms may be present in the Lagrangian, leading to a modified phenomenology. This possibility is labelled as Non-Minimal Flavour Violation (NMFV).

The assumption of NMFV in the squark sector has received considerable attention throughout the last decade. Numerous studies have addressed flavour precision observables [16–21], dark matter aspects [22–24], and most importantly collider signatures related to squark generation mixing [25–31]. In particular, it has recently been shown that non-minimal flavour mixing between the second and third generation squarks can easily be accommodated with respect to current experimental constraints from flavour and precision data [32–34]. Even more recently, it has become apparent that such a configuration is likely to escape detection at current searches at the LHC, and that in consequence a dedicated
search for characteristical signatures of non-minimal flavour violation in the squark sector is necessary [35, 36].

Assuming the discovery of a squark-like state at the LHC, it will be crucial to understand its exact nature and in particular reveal its flavour content. This information will give important hints towards the flavour structure of the underlying theory and will hint towards possible Grand Unification frameworks [37–41]. It is the main goal of the present Paper to investigate different methods for reconstructing the flavour content of an observed squark state. To simplify this first attempt, we concentrate on squarks containing top and charm flavour. This situation is less constrained by flavour and precision data [34] as compared to mixing with first generation flavours [17]. Moreover, squarks containing top flavour are easier to access from the experimental point of view.

Our study will rely on the pair production of a flavour-mixed squark and its subsequent decays into either top or charm quarks plus missing transverse energy, or into bottom quarks and charginos. A direct reconstruction of the squark rotation matrix would basically be possible, provided that we have access to the corresponding branching ratios, potentially with the help of top-polarization measurements [42–45], plus complete information on the neutralino and chargino sector. In practice, having precise access to these information is not an option. We therefore develop methods to infer the top and charm content of the observed squark and obtain information about the flavour structure requiring as few information as possible. More precisely, we will present two methods: the first based on a likelihood inference, the second relying on multi-variate analysis techniques.

The Paper is organized as follows: in Sec. 2, we review the model of our interest, namely the MSSM with NMFV in the squark sector. In Sec. 3 we discuss observables which are measurable at LHC and which we will base our analyses on. Sec. 4 is then devoted to the first method, a likelihood inference of the top-flavour content of the squark. The second method, the multivariate analysis, is then presented in Sec. 5 for the simplified setup before discussing it for the more realistic framework of the MSSM in Sec. 6. Finally, Sec. 7 contains our conclusions.

2 Model and parameters

As discussed in the Introduction, the model of our interest is the Minimal Supersymmetric Standard Model (MSSM) with R-parity conservation and the most general squark flavour structure. In the super-CKM basis, i.e. in the basis \((\tilde{u}_L, \tilde{c}_L, \tilde{t}_L, \tilde{u}_R, \tilde{c}_R, \tilde{t}_R)\), the hermitian up-type squark mass matrix is given by

\[
M_u^2 = \frac{v_u}{\sqrt{2}} T_u - m_u \tan \beta M_{\tilde{U}}^2 + D_{\tilde{u},L} \left( V_{\text{CKM}} M_Q^2 V_{\text{CKM}}^\dagger + m_u^2 + D_{\tilde{u},L} \left( \frac{v_u}{\sqrt{2}} T_u - m_u \tan \beta \right) \right).
\]

The most important terms with respect to our study are the soft mass matrices \(M_Q^2\) and \(M_{\tilde{U}}^2\), together with the trilinear coupling matrix \(T_u\). The remaining parameters, not directly related to the squarks, are the Higgsino potential \(\mu\), the up-type quark mass matrix \(m_u\), and the ratio of the vacuum expectation values of the two Higgs doublets \(\tan \beta = v_u/v_d\). Finally,
the $D$-terms are given by $D_{\tilde{u},L} = m_Z^2 (T_3^u - e_u s_W^2) \cos 2\beta$ and $D_{\tilde{u},R} = m_Z^2 e_u s_W^2 \cos 2\beta$, with $m_Z$ being the $Z$ boson mass, $s_W$ and $c_W$ are the sine and cosine of weak mixing angle and $T_3^u$ and $e_u$ the weak isospin and electric charge of the up-type quarks.

The underlying flavour structure enters the mass matrix through the soft mass and trilinear matrices. Under the assumption of Minimal Flavour Violation (MFV), these matrices are diagonal, such that the CKM-matrix remains the only source of quark flavour violation. Relaxing this assumption, i.e. considering the more general framework of Non-Minimal Flavour Violation (NMFV), allows for off-diagonal entries within these three matrices. Let us note that the same arguments and definitions hold in the sector of down-type squarks and sleptons, which are, however, beyond the scope of the present work.

From the up-type squark mass matrix in Eq. (2.1), the rotation to the basis of physical mass eigenstates $(\tilde{u}_1, \ldots, \tilde{u}_6)$ is done through

$$\text{diag}(m_{\tilde{u}_1}^2, \ldots, m_{\tilde{u}_6}^2) = R_{\tilde{u}} M_{\tilde{u}} R_{\tilde{u}}^\dagger.$$

(2.2)

By convention, the states $\tilde{u}_i$ are labelled to be crescent in mass. All information about the flavour structure of the up-type squarks is contained in the rotation matrix $R_{\tilde{u}}$, and couplings involving up-type squarks in the physical basis relate to the entries of this matrix.

As discussed in the Introduction, a direct reconstruction of the complete squark rotation matrix cannot be aimed at in a near or even mid-term future. We therefore introduce a somewhat less precise but still very meaningful quantity, which is the stop flavour content of the lightest up-type squark $\tilde{u}_1$. This quantity, defined as

$$x_{\tilde{t}} \equiv (R_{\tilde{u}})_{13}^2 + (R_{\tilde{u}})_{16}^2,$$

(2.3)

will be at the centre of the present study. In order to sample the parameter space, we will in addition make use of the quantities $\theta_{\tilde{t}}$ and $\theta_{\tilde{c}}$, which correspond to the top and charm helicity mixing within the lightest state $\tilde{u}_1$,

$$(R_{\tilde{u}})_{12} = (1 - x_{\tilde{t}})^{1/2} \cos \theta_{\tilde{c}}, \quad (R_{\tilde{u}})_{13} = x_{\tilde{t}}^{1/2} \cos \theta_{\tilde{t}},$$

$$ (R_{\tilde{u}})_{15} = (1 - x_{\tilde{t}})^{1/2} \sin \theta_{\tilde{c}}, \quad (R_{\tilde{u}})_{16} = x_{\tilde{t}}^{1/2} \sin \theta_{\tilde{t}}.$$  

(2.4)

The cases $x_{\tilde{t}} = 0$ and $x_{\tilde{t}} = 1$ correspond to MFV with respectively $\tilde{u}_1$ being a pure charm-flavoured or top-flavoured state. Moreover, $\cos \theta_{\tilde{t},\tilde{c}} = 0$ corresponds to a “right-handed” squark, while $\cos \theta_{\tilde{t},\tilde{c}} = 1$ corresponds to a “left-handed” squark.

3 Observables related to flavour violation at LHC

If a squark should be observed at the Large Hadron Collider or any future hadron collider, it will most likely be produced from (flavour-conserving) gluon-initiated processes and manifest through its decay into quarks and gauginos. In our setup, this corresponds to the following decay modes

$$\tilde{u}_1 \to t\tilde{\chi}_1^0, \quad \tilde{u}_1 \to c\tilde{\chi}_1^0, \quad \tilde{u}_1 \to b\tilde{\chi}_1^+,$$

(3.1)
which are simultaneously open if the squark is a mixture of the two flavours, i.e. if \( 0 < x_i < 1 \). Here, the neutralinos manifest as missing transverse energy, while the charginos will decay further into \( W^- \)-bosons and neutralinos.

Our study is based on the assumption that these decays are observed, and that we have access to the observables

\[
m_{\tilde{u}}_1, \ m_{\tilde{\chi}_1^0}, \ m_{\tilde{\chi}_1^+}, \ R_{c/t} = \frac{\text{BR}(\tilde{u}_1 \rightarrow c\chi_1^0)}{\text{BR}(\tilde{u}_1 \rightarrow t\chi_1^0)}, \ R_{b/t} = \frac{\text{BR}(\tilde{u}_1 \rightarrow b\chi_1^+)}{\text{BR}(\tilde{u}_1 \rightarrow t\chi_1^0)}, \quad (3.2)
\]

Note that the production cross-section of the squarks, as well as their branching ratios alone, are difficult to access. We therefore choose to work with the ratios defined above rather than with the pure associated event rates. Analytical expressions for the relevant decay rates in the NMFPV framework can be found in Ref. [18]. Note that in the definition of the ratios \( R_{c/t} \) and \( R_{b/t} \), we assume without loss of generality that the decay into top quarks is always open.

For the further study, it is interesting to examine those expressions in order to find the \( x_i \)-dependence of the observables in certain limits concerning the nature of the involved neutralinos and charginos. For example, assuming a pure higgsino-like neutralino and neglecting the neutralino mass with respect to the squark mass, we obtain

\[
R_{c/t} \bigg|_{\chi_1^0 = B^0, \ m_{\tilde{u}}_1 > m_{\tilde{\chi}_1^0}} = \frac{m_\tau^2}{m_t^2} \frac{1 - x_i}{x_i}, \quad (3.3)
\]

As a second example, we assume a pure bino-like neutralino and obtain

\[
R_{c/t} \bigg|_{\chi_1^0 = B^0, \ m_{\tilde{u}}_1 > m_{\tilde{\chi}_1^0}} = \frac{1 - x_i + \kappa_c (R^a)^2_{15}}{x_i + \kappa_t (R^a)^2_{16}} \rightarrow \frac{1 - x_i}{x_i}, \quad (3.4)
\]

where \( \kappa_q = c_q^2 / (e_q - T_q^2)^2 - 1 = 15 \) for \( q = c, t \), and the last expression holds for a pure “left-handed” or a pure “right-handed” squark. Finally, for a pure wino-like neutralino, the ratio becomes

\[
R_{c/t} \bigg|_{\chi_1^0 = W^0} = \frac{B_c \lambda_c^{1/2} (R^a)^2_{12}}{B_t \lambda_t^{1/2} (R^a)^2_{13}} \rightarrow \frac{B_c \lambda_c^{1/2}}{B_t \lambda_t^{1/2}} \frac{1 - x_i}{x_i}, \quad (3.5)
\]

where \( \lambda_q = m_{u_1}^4 + m_{\tilde{\chi}^0_1}^4 + m_q^4 - 2(m_{u_1}^2 m_{\tilde{\chi}^0_1}^2 + m_{u_1} m_q^2 m_{\tilde{\chi}^0_1} m_q) \) denotes the usual Källén function associated to the squark decay and \( B_q = m_{u_1}^2 - m_{\tilde{\chi}^0_1}^2 - m_q^2 \) for \( q = c, t \). Here, the last expression holds for a pure “left-handed” squark.

In order to gain a better understanding of these ratios, we start by randomly scanning over the parameters governing the lightest squark, neutralino, and chargino. More precisely, we vary the physical squark mass \( m_{\tilde{u}_1} \), and the parameters \( x_i, \theta \), and \( \theta_L \) defining its flavour decomposition. In the gaugino sector, we vary the bino, wino, and Higgsino mass parameters \( M_1, M_2, \) and \( \mu \). The physical gaugino masses are obtained by diagonalizing the mass matrices at the tree-level.

As the expressions in Eqs. (3.3) – (3.5) do not exhibit a dependence on \( \tan \beta \), we conclude that this parameter only has a mild impact on the observables of our interest. We
therefore fix $\tan \beta = 10$ throughout the presented analyses. All parameters are scanned over in a uniform manner according to the ranges given in Table 1. The corresponding parameter distributions are illustrated in Fig. 1 for the relevant physical masses and Fig. 2 for the corresponding mixing parameters, respectively. The shape of the mass distributions are explained by the fact that we require the decay modes mentioned above to be kinematically allowed, which favours larger squark and smaller gaugino masses. Since we impose a flat distribution of the stop content $x_{\tilde{t}}$, the elements $(R_{\tilde{u}}^2)_{1i}$ ($i = 2, 3, 5, 6$) of the up-squark rotation matrix follow a parabolic distribution. As the distributions of $(R_{\tilde{u}}^3)_{1i}$ for $i = 3, 5, 6$ are similar to the one of $(R_{\tilde{u}}^3)_{12}$, they are not shown separately in Fig. 2.

For each parameter point, the gaugino masses and the ratios $R_{c/t}$ and $R_{b/t}$ of our interest are computed using the full analytical expressions of Ref. [18]. The results are depicted in Fig. 3, where we indicate as colour code the dominant component of the involved neutralino as well as the nature of the decaying squark. As expected from Eqs. (3.3) – (3.5), distinct regions are observed in the distributions of $R_{c/t}$. The same kind of feature appears for the ratio $R_{b/t}$. More precisely, the two ratios depend strongly on the neutralino decomposition and the “chirality” (expressed in terms of $\theta_{\tilde{t}}$ and $\theta_{\tilde{c}}$ defined in Eqs. (2.4)) of the decaying squark. The width of each band is due to the fact that the majority

| Variable | Range  | Variable | Range  |
|----------|--------|----------|--------|
| $m_{\tilde{t}_1}$ | [700, 2000] | $M_1$ | [600, 2000] |
| $x_{\tilde{t}}$ | [0, 1] | $M_2$ | [600, 2000] |
| $\cos \theta_{\tilde{t}}$ | [0, 1] | $\mu$ | [600, 2000] |

Table 1: Scanned ranges of the parameters associated to the squark (left) and gaugino sector (right). All masses are given in GeV.

Figure 1: Distributions of the squark (left) and gaugino (right) masses obtained from the scan summarized in Table 1. The masses are given in GeV. The distributions are shown on a linear scale.
Figure 2: Distributions of the squark (upper row) and neutralino (lower row) mixing parameters associated to the masses shown in Fig. 1. The distributions are shown on a linear scale.

of the parameter points feature mixed gauginos and squarks rather than corresponding to the limit cases discussed above. Nevertheless, the presence of the observed rather distinct regions is an important feature which will turn out to be crucial in the identification of the squark flavour decomposition from the observables given in Eq. (3.2).
Figure 3: Distributions of the ratios $R_{c/t}$ (left) and $R_{b/t}$ (right) of the decay modes defined in Eq. (3.2) in dependence of the stop composition $x_{\tilde{t}}$ of the decaying squark. The colour code refers to different combinations of neutralino compositions and squark “chiralities”.

4 Likelihood inference in a simplified model

In order to infer the stop component $x_{\tilde{t}}$ of the observed squark, we start by constructing a maximum likelihood estimator. For a given set of data

$$D = \{m_{\tilde{u}_1}, m_{\tilde{\chi}^0_1}, m_{\tilde{\chi}^\pm_1}, R_{c/t}, R_{b/t}\}$$

supposed to be obtained at the Large Hadron Collider, we associate a likelihood value to each point of an ensemble of random parameter points. Assuming uncorrelated parameters and thus a Gaussian distribution, this likelihood takes the form

$$\ln L(\theta) = -\frac{1}{2} \sum_i \left( \frac{\theta_i - D_i}{\sigma_i} \right)^2$$

with $\theta$ being the set of parameters associated to the parameter point under consideration and $\sigma_i$ being the error associated to the observable $D_i$.

We now divide the interval $x_{\tilde{t}} \in [0; 1]$ into $N$ bins of equal size. For each bin $j = 1, \ldots, N$, we then compute the average likelihood $\hat{L}_j(x_{\tilde{t}})$ of all random parameter points having their value of $x_{\tilde{t}}$ inside the given bin. From the obtained values of $\hat{L}_j(x_{\tilde{t}})$ over the interval $x_{\tilde{t}} \in [0; 1]$, we can fit a Gaussian distribution in order to find the maximum of likelihood corresponding to the inferred value of the stop component $x_{\tilde{t}}$. The associated uncertainty $\sigma(x_{\tilde{t}})$ is then based on the standard deviation value of the fit.

Several subtleties appear and have to be addressed when performing this analysis. The first one is of technical nature: in order to fill the full $x_{\tilde{t}}$ interval in an appropriate way, we need a rather large number of parameter points. Note that this number depends on the chosen bin size. The second, and more tricky part, lies in the uncertainties $\sigma_i$ entering the likelihood in Eq. (4.2). Usually this standard deviation correspond to the uncertainty associated with the measurement of the observables $D_i$. However, in the present case,
Figure 4: Likelihood fit four test data sets featuring a fixed gauginos composition as in Eq. (4.3). The resulting inferred values of the stop component are listed in Table 2. The distributions are shown on a linear scale.

since different parameter points \( \theta_i \) can lead to the same observation, there is an additional feature, which we shall discuss later on.

For the sake of simplicity, and in order to illustrate the proposed inference method, we fix in a first step the parameters associated to the neutralino and chargino decomposition as

\[
N_{1l} = 0.5, \quad U_{11} = V_{12} = 1, \quad U_{12} = V_{11} = 0, \quad (4.3)
\]

i.e. we consider a maximally mixed neutralino. For the present example, we have performed a random scan over the five parameters of Eq. (4.1) leading to an ensemble of \( 5 \cdot 10^8 \) parameter points. Moreover, we assign a common value of \( \sigma_i = 0.25 D_i \) to the uncertainties entering the likelihood calculation.

Assuming four different test parameter points \( P_i \) (\( i = 1, \ldots, 4 \)), we perform the analysis described above and infer the stop component \( x_i \) using a Gaussian likelihood fit. The results are illustrated in Fig. 4 and summarized in the upper part of Table 2. More precisely, for each test parameter point, we show the average likelihood \( \hat{L}_j(x_i) \) obtained for each bin and the Gaussian fit. As can be seen, our method manages to recover the actual stop component within the resulting uncertainty from the Gaussian fit.
If almost all of the points in a given bin lead to a rather low value of \( L(\theta) \) and if the values of \( \sigma_i \) are rather low, one of the points may have a great importance. This in turn can lead to a bad estimation of the stop component \( x_\tilde{t} \). The conclusion of this is that the uncertainties \( \sigma_i \) in Eq. (4.2) have in fact two components: the uncertainty of the observable itself and in addition an uncertainty coming from the potential degeneracy of the observables with respect to the stop component \( x_\tilde{t} \). Such a degeneracy may lead to a variation of the density of points between different regions of parameter space.

For completeness, we illustrate the impact of the uncertainty \( \sigma_i \) associated to the observables \( D_i \). The four graphs shown in Fig. 5 present the likelihood fit of a single parameter point for four different values of \( \sigma_i/D_i \) as indicated in Table 2. For \( \sigma_i/D_i \) being set to 0.35, the fit converges well and the resulting value for \( x_\tilde{t} \) is close to the assumed one. However, the gaussian width is rather large leading to a sizeable uncertainty \( \sigma(x_\tilde{t}) \) on the inferred value. Going to smaller values for the uncertainties \( \sigma_i \) leads to a less accurate fit and drives the inferred value away from the assumed one. The associated gaussian width is smaller, and thus the inferred uncertainty is less important. The explanation of this feature is that the actual uncertainty on the inferred stop composition has in reality two components, the first coming from the width of the fitted Gaussian function, the second coming from the overall quality of the fit. Finally, pushing \( \sigma_i/D_i \) to even lower values makes the fit break down, as can be seen in the last panel of Fig. 5. This is the actual tricky point of the present analysis. The choice of \( \sigma_i/D_i \) must therefore be made with precaution because it has more impact than the Gaussian width and it also depends on the density of points in a certain area of the scan. Using this method to infer the actual stop composition will need specific care and deeper understanding of the uncertainties, which is beyond the first step presented in this work.

As a final step, we relax the assumption on the gaugino decompositions given in Eq. (4.3), and include the gaugino mixing parameters in the random scan. Again, we generate
Figure 5: Same as Fig. 4 for a single test parameter point but varying uncertainty \( \sigma_i/D_i \).

Figure 6: Same as Fig. 4 for two test parameter points obtained by scanning in addition over the parameters related to the gaugino sector.

an ensemble of \( 5 \cdot 10^8 \) parameter points with \( \sigma_i = 0.35D_i \), and apply our reconstruction method to two data sets \( P_5 \) and \( P_6 \). The results are shown in Fig. 6 and summarized in the lower part of Table 2. Even if the true stop components lie within the inferred intervals, the uncertainties are much larger in this case, such that the results become rather meaningless.
in certain cases. In addition, from Fig. 6 we can see that the likelihood is no longer Gaussian. This is due to the fact that here different regions of the parameters present a concentration of points able to explain the data.

5 Multivariate analysis in a simplified model

In order to go beyond the likelihood inference presented in the previous Section, especially in a more realistic setup as discussed, e.g., in Ref. [34], we now employ a multivariate analysis (MVA) classifier. In the present Section, we will start by presenting results obtained from a multi-layer perceptron (MLP) provided by ROOT through the TMVA package [46] for the simplified setup already used in Sec. 3. The discussion of the complete MSSM with squark generation mixing of Ref. [34] will follow in Sec. 6.

In this context, the goal of our analysis is slightly different with respect to the previous Section. While the likelihood inference aims at estimating the actual stop component of the observed squark, a multivariate analysis is designed to efficiently classify different configurations. In order to provide a simple illustration, we define two categories based on the stop composition $x_t$, which remains the key quantity of our interest. We will divide the parameter space into “top-flavoured” squarks and “charm-flavoured” squarks according to

\[
\begin{align*}
  x_t < 0.5 & \iff \text{“charm – flavoured”}, \\
  x_t > 0.5 & \iff \text{“top – flavoured”}.
\end{align*}
\]

Let us note that these categories are for the moment rather arbitrary and aim at the illustration of the method rather than representing specific physical regions. In particular, additional categories can be defined in order to refine the analysis. Based on the two categories, the MLP can be trained on the parameter points obtained from the random scan, and subsequently tested on a subset of points, the test sample, in order to compute the efficiency and the misidentification rate of the classifier. The analysis presented here is based on a training sample of $10^6$ points, which have been obtained by uniformly scanning as indicated in Table 1.

The classifier basically combines the set of observables given in Eq. (3.2), i.e. $R_{c/t}$, $R_{b/t}$, $m_{\tilde{u}_1}$, $m_{\tilde{t}_1^0}$ and $m_{\tilde{t}_1^+}$, into one single variable called the MLP response. The algorithm will associate an MLP value to each parameter point of the scan, depending on the set of observables that maximize the separation between the two categories. Several points need to be kept in mind when performing this type of analysis. A key point is the so-called “overtraining”, meaning that training the algorithm on a too small dataset may enforce the identification of unphysical regions, i.e. statistical fluctuations, as physical ones. We have performed an overtraining check by comparing the classification performance on the training sample and on the test sample. The behavior of the algorithm being the same on the two samples, we conclude that there are no statistical fluctuations having an impact on the classification.

The rather simple situation of having only two categories will also serve to study the influence of the underlying prior distribution, in particular of the stop component $x_t$. 

\[\text{– 11 –}\]
We start from the same setup as in Sec. 3, where the random parameter scan has been performed such that the stop component $x_{i\tilde{t}}$ exhibits a flat distribution. For this case, we show the obtained MLP response for the two categories in Fig. 7, together with the prior distribution of the stop component (see also Fig. 2).

As can be seen, the classifier manages to separate the two categories with a rather good efficiency. For a given misidentification rate, the associated efficiency, i.e. the number of points of a chosen class surviving the misidentification cut, of the classifier can be computed. In the present case, for a misidentification rate of 10%, we obtain an efficiency of 54% for the “top-like” squark region and of 64% for the “charm-like” case. In other words, we can tag respectively approximately 54% and 64% of the points at 90% confidence level.

As a second example, we employ the classifier to the case of a non-uniform prior distribution of the stop-content $x_{i\tilde{t}}$. Inspired by the results of Ref. [34], we choose a prior distribution peaking at its “MFV-like” extremities $x_{i\tilde{t}} \approx 0$ and $x_{i\tilde{t}} \approx 1$. Apart from the prior distribution (and thus the squark rotation matrix elements), the sample has the same characteristics as the previous one. The prior distribution and the resulting MLP response are shown in Fig. 8. While it is approximately symmetric in the case of a flat prior, the MLP response associated to the two categories is clearly non-symmetric in the present case. This can be traced to the fact that the observables used to classify are non-symmetric with respect to “top-flavoured” and “charm-flavoured” squarks.

In this example, for the misidentification rate of 10%, we obtain an efficiency of 64% (60%) for the “top-flavoured” (“charm-flavoured”) category. It appears that the efficiency depends on the prior distribution. More precisely, the classifier becomes more efficient in identifying the “top-flavoured” category, but slightly less performant concerning the “charm-flavoured” category.

The increasing classification power coming from the prior distribution can intuitively be understood as the two categories are now more different. The border between the two...
cases, i.e. $x_\tilde{t} \sim 0.5$, where it is phenomenologically difficult to assign a given point to a single category, are less populated in the second case with non-uniform prior. It is therefore easier to maximize the separation. After this first analysis within a simplified setup, we now aim at applying the MLP method to a more complete model.

6 Application to the MSSM with mixed top-charm squarks

As announced in the previous Section, we finally apply the multivariate analysis (MVA) classifier to the Minimal Supersymmetric Standard Model (MSSM) with non-minimal flavour mixing between charm- and top-flavoured squarks. In order to work with a rather “realistic” setup, we choose to use the parameter points obtained in Ref. [34] as basis of our study. These parameter points defined at the TeV scale have been shown to fulfill all relevant constraints coming from flavour and precision measurements, in particular the Higgs-boson mass, the decays $B \rightarrow X_s \gamma$ and $B \rightarrow X_s \mu \mu$, and the meson oscillation parameter $\Delta M_{B_s}$, to name the most important ones. For all details on the applied constraints and the related Markov Chain Monte Carlo study of the MSSM with non-minimal flavour violation in the squark sector, the reader is referred to Ref. [34].

Following the preliminary study of the simplified setup in Sec. 5, it is interesting to examine the prior distribution of the quantity that we want to address, i.e. the stop component $x_i$ of the lightest up-type squark. As can be seen from its representation in Fig. 9, the distribution strongly peaks at the “MFV-like” ends. Moreover, flavour and precision data tend to favour a high charm content with respect to top content in the lightest squark. Note that this situation is similar to the non-uniform prior tested in Sec. 5, which turned out to yield a higher efficiency than the simpler uniform prior. We now perform the same MLP classification using a training sample containing about $6 \cdot 10^5$ points obtained from the MCMC analysis of Ref. [34].

\footnote{For the present study, we have extended the sample resulting from the analysis presented in Ref. [34] using exactly the same computational setup.}
Figure 9: Prior distribution (linear scale) of the stop composition $x_{\tilde{t}}$ from the MCMC analysis of Ref. [34].

Figure 10: MLP response on the NMFV-MSSM of Ref. [34] for the separation the “charm MFV” (upper left), “charm NMFV” (upper right), “top NMFV” (lower left), and “top MFV” (lower right) categories (red) from the remaining parameter points (blue).
Starting from the prior distribution shown in Fig. 9, we divide the ensemble of points into four categories defined as follows:

\[
\begin{align*}
0.00 \leq x_\tilde{t} &< 0.05 \iff \text{"charm MFV"} \\
0.05 < x_\tilde{t} &< 0.50 \iff \text{"charm NMFV"} \\
0.50 < x_\tilde{t} &> 0.95 \iff \text{"top NMFV"} \\
0.95 < x_\tilde{t} &\leq 1.00 \iff \text{"top MFV"}
\end{align*}
\]

Note that, although the given definition of the above categories is again somewhat arbitrary, the exact value of the cuts between MFV and NMFV does not have a major impact on the methods presented in the following. It might, however, affect the efficiency of the proposed analysis, and the exact definition of the categories may in practice depend on the problem under consideration.

The MVA classifier is used to separate each of the four categories from its complement, i.e. the ensemble comprising the three other classes. In Fig. 10, we show the MLP responses obtained for the four cases. As expected from the overpopulated prior region, the “charm MFV” category is rather well identified. However, the identification is less efficient for the two NMFV categories, which are underpopulated in the prior distribution. For the sake of a numerical comparison between the categories, and also to the cases presented in Sec. 5, we summarize the obtained efficiencies of the classifier in Table 3.

Overall, the performance of the classifier is better than for the simplified situations presented in Sec. 5. This can be traced to the underlying prior distribution of the stop content \( x_\tilde{t} \) (see Fig. 9). The categories which are most difficult to identify, i.e. the two NMFV categories, are less populated in this particular model. The algorithm is therefore less performant in distinguishing these categories. The small bump observed around MLP \( \sim 0.7 \ldots 0.8 \) in both NMFV categories is an artefact of the employed multi-class MLP due to the presence of phenomenologically different regions.

Let us finally mention that we have also tested the likelihood inference method discussed in Sec. 4 on the present case of the NMFV-MSSM of Ref. [34]. However, for this method it turns out that inferring in a region of rather low density is quite difficult (contrary to the case of a uniform prior applied in Sec. 4). In addition, the strongly peaked prior distribution of the stop component \( x_\tilde{t} \) leads to a certain bias, such that the obtained results are not reliable any more. We therefore do not discuss this method further for the given model.

| Categories         | Efficiency |
|--------------------|------------|
| “charm” MFV        | 95%        |
| “charm” NMFV       | 51%        |
| “top” NMFV         | 41%        |
| “top” MFV          | 69%        |
7 Conclusion

We address the question to which extend the flavour decomposition of a squark-like state produced at the Large Hadron Collider can be reconstructed. As a starting point, we have considered a rather simple but typical set of collider observables related to inter-generational mixing between top- and charm-flavoured squarks. The quantity of our interest is the top-flavour content of the observed squark state, since it may give valuable information on the flavour structure of the theory.

We first have employed a likelihood inference method, which basically allows to infer the top-flavour content of the observed squark. With the help of a simplified model incorporating non-minimal flavour violation between the top and charm squarks, we have obtained viable information on the squark flavour structure assuming that additional information, in particular concerning the gaugino sector, is provided. In absence of information on the neutralino and chargino nature, the likelihood inference is less viable. However, the more additional information is available, e.g. on the gaugino sector (even if not fully determined), the more efficient this method will be. We also tried to use the likelihood inference method to the more general situation of the Minimal Supersymmetric Standard Model (MSSM) with additional top-charm mixing in the squark sector. However, it turns out to be inapplicable due to the somewhat extreme prior distribution of the top-flavour content and the available number of parameter points in the test sample.

The second method consists of a multivariate analysis classifier, which can efficiently separate two categories among a sample making use of a given set of observables. Performing this analysis on both the simplified setup and on the more general MSSM framework has led to promising results concerning the separation between the Minimal and Non-Minimal Flavour Violation hypotheses. It turns out that this method can better deal with the strongly peaked prior distributions as it is the case in the MSSM with top-charm flavour mixing.

We want to emphasize the fact that the two methods are complementary. While the multi-variate analysis does not return an actual value for the top-flavour content of the squark, the likelihood inference can provide a reasonable estimation. However, the likelihood inference needs additional information, especially on the gaugino sector, and cannot handle very extreme prior distributions. These inconveniences can be avoided by the use of the multivariate analysis, which already allows to gain valuable information on the flavour structure. As this is a first attempt of the reconstruction of the squark flavour structure, the presented analysis relies on rather simple observables. Designing improved analyses inspired from this work should lead to a considerable improvement of the performances. As an example, one might consider additional observables related to the same parameters, such as, e.g., such as the top polarization from the squark decay or event rates stemming from gluino production and decay. From the machine-learning point of view, many algorithms exist for parameter-fitting problems and with a specific analysis it may be possible to access the actual value of the top-flavour content in a generic gaugino sector. Furthermore, considering new types of algorithms and additional observables may give access to the actual entries of the squark rotation matrix.
Acknowledgments

The authors would like to thank M. Reboud for useful discussions. The work of J.B. is supported by a Ph.D. grant of the French Ministry for Education and Research. This work is supported by Investissements d’avenir, Labex ENIGMASS, contrat ANR-11-LABX-0012. The figures presented in this Paper have been generated using MatPlotLib. [47].

References

[1] ATLAS collaboration, G. Aad et al., Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s}=8$ TeV pp collisions with the ATLAS detector, JHEP 11 (2014) 118, [arXiv:1407.0583].

[2] ATLAS collaboration, G. Aad et al., Search for pair-produced third-generation squarks decaying via charm quarks or in compressed supersymmetric scenarios in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector, Phys. Rev. D90 (2014) 052008, [arXiv:1407.0608].

[3] ATLAS collaboration, G. Aad et al., Search for Scalar Charm Quark Pair Production in pp Collisions at $\sqrt{s}=8$ TeV with the ATLAS Detector, Phys. Rev. Lett. 114 (2015) 161801, [arXiv:1501.01325].

[4] CMS collaboration, A. M. Sirunyan et al., Search for new phenomena with the $M_{T2}$ variable in the all-hadronic final state produced in proton-proton collisions at $\sqrt{s}=13$ TeV, Eur. Phys. J. C77 (2017) 710, [arXiv:1705.04650].

[5] ATLAS collaboration, M. Aaboud et al., Search for supersymmetry in final states with two same-sign or three leptons and jets using 36 fb$^{-1}$ of $\sqrt{s}=13$ TeV pp collision data with the ATLAS detector, JHEP 09 (2017) 084, [arXiv:1706.03731].

[6] CMS collaboration, A. M. Sirunyan et al., Search for top squark pair production in pp collisions at $\sqrt{s}=13$ TeV using single lepton events, JHEP 10 (2017) 019, [arXiv:1706.04402].

[7] CMS collaboration, A. M. Sirunyan et al., Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s}=13$ TeV, JHEP 10 (2017) 005, [arXiv:1707.03316].

[8] CMS collaboration, A. M. Sirunyan et al., Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in protonproton collisions at $\sqrt{s}=13$ TeV, Phys. Lett. B778 (2018) 263–291, [arXiv:1707.07274].

[9] ATLAS collaboration, M. Aaboud et al., Search for direct top squark pair production in final states with two leptons in $\sqrt{s}=13$ TeV pp collisions with the ATLAS detector, Eur. Phys. J. C77 (2017) 898, [arXiv:1708.03247].

[10] ATLAS collaboration, M. Aaboud et al., Search for supersymmetry in events with b-tagged jets and missing transverse momentum in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector, JHEP 11 (2017) 195, [arXiv:1708.09266].

[11] ATLAS collaboration, M. Aaboud et al., Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}=13$ TeV with the ATLAS detector, JHEP 12 (2017) 085, [arXiv:1709.04183].
[12] CMS collaboration, A. M. Sirunyan et al., Search for supersymmetry in proton-proton collisions at 13 TeV using identified top quarks, Phys. Rev. D97 (2018) 012007, [arXiv:1710.11188].

[13] CMS collaboration, A. M. Sirunyan et al., Search for top squarks and dark matter particles in opposite-charge dilepton final states at \( \sqrt{s} = 13 \) TeV, Phys. Rev. D97 (2018) 032009, [arXiv:1711.00752].

[14] ATLAS collaboration, M. Aaboud et al., Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector, JHEP 01 (2018) 126, [arXiv:1711.03301].

[15] ATLAS collaboration, M. Aaboud et al., Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb\(^{-1}\) of \( \sqrt{s} = 13 \) TeV pp collision data with the ATLAS detector, JHEP 06 (2018) 108, [arXiv:1711.11520].

[16] S. Heinemeyer, W. Hollik, F. Merz and S. Penaranda, Electroweak precision observables in the MSSM with nonminimal flavor violation, Eur. Phys. J. C37 (2004) 481–493, [hep-ph/0403228].

[17] M. Ciuchini, A. Masiero, P. Paradisi, L. Silvestrini, S. K. Vempati and O. Vives, Soft SUSY breaking grand unification: Leptons versus quarks on the flavor playground, Nucl. Phys. B783 (2007) 112–142, [hep-ph/0702144].

[18] G. Bozzi, B. Fuks, B. Herrmann and M. Klasen, Squark and gaugino hadroproduction and decays in non-minimal flavour violating supersymmetry, Nucl. Phys. B787 (2007) 1–54, [arXiv:0704.1826].

[19] F. del Aguila, J. Aguilar-Saavedra, B. Allanach, J. Alwall, Y. Andreev et al., Collider aspects of flavour physics at high Q, Eur. Phys. J. C57 (2008) 183–308, [arXiv:0801.1800].

[20] B. Fuks, B. Herrmann and M. Klasen, Flavour Violation in Gauge-Mediated Supersymmetry Breaking Models: Experimental Constraints and Phenomenology at the LHC, Nucl.Phys. B810 (2009) 266–299, [arXiv:0808.1104].

[21] B. Fuks, B. Herrmann and M. Klasen, Phenomenology of anomaly-mediated supersymmetry breaking scenarios with non-minimal flavour violation, Phys.Rev. D86 (2012) 015002, [arXiv:1112.4838].

[22] B. Herrmann, M. Klasen and Q. Le Boulc’h, Impact of squark flavour violation on neutralino dark matter, Phys.Rev. D84 (2011) 095007, [arXiv:1106.6229].

[23] M. Blanke and S. Kast, Top-Flavoured Dark Matter in Dark Minimal Flavour Violation, JHEP 05 (2017) 162, [arXiv:1702.08457].

[24] M. Blanke, S. Das and S. Kast, Flavoured Dark Matter Moving Left, JHEP 02 (2018) 105, [arXiv:1711.10493].

[25] T. Hurth and W. Porod, Flavour violating squark and gluino decays, JHEP 08 (2009) 087, [arXiv:0904.4574].

[26] A. Bartl, K. Hidaka, K. Hohenwarter-Sodek, T. Kernreiter, W. Majerotto and W. Porod, Impact of squark generation mixing on the search for gluinos at LHC, Phys. Lett. B679 (2009) 260–266, [arXiv:0905.0132].

[27] M. Bruhnke, B. Herrmann and W. Porod, Signatures of bosonic squark decays in non-minimally flavour-violating supersymmetry, JHEP 1009 (2010) 006, [arXiv:1007.2100].
[28] A. Bartl, H. Eberl, B. Herrmann, K. Hidaka, W. Majerotto et al., Impact of squark generation mixing on the search for squarks decaying into fermions at LHC, *Phys. Lett.* **B698** (2011) 380–388, [arXiv:1007.5483].

[29] M. Blanke, G. F. Giudice, P. Paradisi, G. Perez and J. Zupan, Flavoured Naturalness, *JHEP* **1306** (2013) 022, [arXiv:1302.7232].

[30] A. Bartl, H. Eberl, E. Ginina, K. Hidaka and W. Majerotto, $h^0 \rightarrow \bar{c}c$ as a test case for quark flavor violation in the MSSM, *Phys. Rev.* **D91** (2015) 015007, [arXiv:1411.2840].

[31] M. Arana-Catania, S. Heinemeyer and M. J. Herrero, Updated Constraints on General Squark Flavor Mixing, *Phys. Rev.* **D90** (2014) 075003, [arXiv:1405.6960].

[32] M. Backovi, A. Mariotti and M. Spannowsky, Signs of Tops from Highly Mixed Stops, *JHEP* **06** (2015) 122, [arXiv:1504.00927].

[33] M. Dimou, S. F. King and C. Luhn, Phenomenological implications of an SU(5) $\times S_4 \times U(1)$ SUSY GUT of flavour, *Phys. Rev.* **D93** (2016) 075026, [arXiv:1512.09063].

[34] A. Prasath V, R. M. Godbole and S. D. Rindani, Longitudinal top polarisation measurement and anomalous $Wtb$ coupling, *Eur. Phys. J.* **C75** (2015) 402, [arXiv:1405.1264].

[35] “Toolkit for multivariate data analysis with root (tmva).” [https://root.cern/tmva](https://root.cern/tmva).
[47] J. D. Hunter, Matplotlib: A 2D graphics environment, *Computing In Science & Engineering* 9 (2007) 90–95.