Decoding the nature of Dark Matter at current and future experiments

Alexander Belyaev
School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
Particle Physics Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK
E-mail: a.belyaev@soton.ac.uk

Abstract. Decoding the nature of Dark Matter (DM) as a crucial part of Beyond-the-Standard-Model (BSM) theory is one of the most important problems of modern particle physics. DM potentially provides unique signatures at collider and non-collider experiments. These signatures are quite generic, however their details could allow us to delineate various BSM models and the properties of DM. While there are many comprehensive studies of the phenomenology of various appealing BSM models, exhibiting “top-bottom” approach, there is no clear strategy for the reverse task of identifying the underlying theory from the new signatures. To solve this problem one should consider the comprehensive set of signatures, database of models and use modern methods, including machine learning and artificial intelligence, to decode the underlying theory from potential signals of new physics we are expecting from the coming experimental data. One of the tools which could be helpful to solve the problem is High Energy Physics Model Database (HEPMDb) which was created to make a step forward towards solving this problem. It is aimed to facilitate connection between HEP theory and experiment, to store, validate and explore BSM models and to collect their signatures. DM decoding is based on the unique complementarity of Large Hadron Collider (LHC) potential as well as on the potential DM direct and indirect detection experiments to probe DM. The combination of our knowledge on this complementarity, modern analysis methods, comprehensive database of BSM models and their signatures is the key point of decoding the nature of DM and the whole underlying theory of Nature.

1. Introduction
Understanding the nature of Dark Matter (DM) is one of the greatest puzzles of modern particle physics and cosmology. Although overwhelming observational evidence from galactic to cosmological scales point to the existence of DM [1, 2, 3], after decades of experimental effort only its gravitational interaction has been experimentally confirmed. Currently, no information is available about the DM properties, such as its spin, mass, interactions other than gravitational, symmetry responsible for its stability, number of states associated to it, and possible particles that would mediate the interactions between DM and the standard model (SM) particles.

If DM is light enough and interacts with SM particles directly or via some mediators with a strength beyond the gravitational one, its elusive nature can be detected or constrained from several different observables: a) from direct production at colliders, resulting in a signature exhibiting an observed SM object, such as jet, Higgs, Z, W, photon or top-quark(s) that recoils against the missing energy from the DM pair [4, 5, 6, 7, 8]; b) via the relic density constraint
obtained through the observations of cosmic microwave background (CMB) anisotropies, such as those of WMAP and PLANCK collaborations \[9, 1, 10\]; c) from DM direct detection (DD) experiments, which are sensitive to elastic spin independent (SI) or spin dependent (SD) DM scattering off nuclei \[11, 12, 13, 14\]; d) from DM indirect detection searches, that look for SM particles produced in the decay or annihilation of DM present in the cosmos, both with high energy observables (gamma-rays, neutrinos, charge cosmic rays) produced in the local Universe \[15, 16, 17, 18, 19, 20\], and by studying the effects of energy produced by DM annihilation in the early universe on the properties of the CMB spectrum \[21, 22, 1\].

Obviously, decoding the nature of DM requires the respective signal at least in one of the search experiments. We do not have one yet. However, at the moment we can already conclude about what kind of DM models are excluded. Moreover, we have time now to create the framework which would help us to perform bottom→up decoding using the power of correlation and interplay of different signatures in the space of DM models. Such framework should combine elements of the model database, signature database and machine learning to realise identification of Dark Matter. An example of the prototype of such a framework has been created at created at Southampton University in the form of High Energy Physics Model DataBase (HEPMD) (https://hepmdb.soton.ac.uk) \[23\], potential of which is discussed at the end of this contribution.

2. Contact interactions

Let us start our discussion with the three simplest scenarios for the DM particles: complex scalars (\(\phi\)), Dirac fermions (\(\chi\)) and complex vectors (\(V_\mu\)) within the effective field theory (EFT) approach. In the EFT approach we parametrize the DM interactions with the SM quarks and gluons with the effective coupling and the scale describing operators of dimension six or five. In Table 1 we have summarised a minimal set of independent dimension-5 and dimension-6 operators for complex scalar, Dirac fermion and complex vector DM coupling to quarks and gluons, adopting the widely used notations of \[24, 25, 26\]. Since different operators have different energy behaviour and respective different invariant mass distributions: typically softer for majority operators with scalar DM, intermediate for fermion DM and the hardest for vector DM and because of relation of \(M_{\text{inv}}(DM,DM)\) and \(E_{\text{T}}^\text{miss}\) slope one can distinguish several operators and related underlying theories using the shape of the \(E_{\text{T}}^\text{miss}\) signal: C1-C2,C5-C6,D9-D10,V1-V2,V3-V4,V5-V6 and V11-12 pairs among each other \[20\].

Notice the presence of the coupling \(g^*\) in the definition of the effective operators, which we insert according to the Naive Dimensional Analysis \[27\]. Moreover, for the vector DM case we choose the parametrisation suggested in Ref. \[20\] that takes into account the high energy behaviour of the scattering amplitudes that are enhanced by an energy factor \((E/m_{DM})\) for every longitudinal vector DM polarisation. These operators are gauge invariant and provides the minimal and simplistic description of underlying theory of DM. The scale \(\Lambda\) is related to the mass of the mediator, while coupling \(g^*\) is related to the product of DM and SM couplings to the mediator.

These operators provide monojet-signature, the shapes of \(E_{\text{T}}^\text{miss}\) distributions for which is presented in Fig. 1 from Ref. \[20\] for DM mass of 10 GeV. One can observe a big difference in \(E_{\text{T}}^\text{miss}\) shapes of the groups of the operators, primarily split into groups of operators with scalar, fermion and vector DM. The origin of the different \(E_{\text{T}}^\text{miss}\) shapes from different operators can be related to a combination of effects. First, for a fixed Lorentz structure of the SM part of the EFT operators, the same invariant mass distribution of the DM pair, \(M_{\text{inv}}(DM,DM)\), uniquely defines the shape of the \(E_{\text{T}}^\text{miss}\) distribution. Moreover, with the increase of \(M_{\text{inv}}(DM,DM)\), the \(E_{\text{T}}^\text{miss}\) shape falls less and less steeply (again, for a given SM component of the EFT operator).

It was found in Ref. \[20\] that the reason why the bigger invariant mass of DM is correlated with flatter \(E_{\text{T}}^\text{miss}\) behaviour is related to phase space and parton density effects: when
Table 1: Minimal basis of operators of dimension six or less involving only complex scalar DM ($\phi$), Dirac fermion DM ($\chi$) or complex vector DM ($V^\mu$) interacting with SM quarks ($q$) or gluons. Here we denote the field strength tensor of the gluons as $G^{\mu\nu}$ and its dual as $\tilde{G}^{\mu\nu}$.

$M_{\text{inv}}(DM,DM)$ is small, the radiation of a high $p_T$ jet will “cost” a large relative shift in $x$, the transferred momentum of the parton, leading to a rapidly falling $E_T^{\text{miss}}$ distribution; on the contrary, when $M_{\text{inv}}(DM,DM)$ is large, the radiation of a high $p_T$ jet will “cost” a small relative shift in $x$, which will lead to a more slowly falling $E_T^{\text{miss}}$ distribution in comparison to the first case. Therefore if one theory predicts higher values of the invariant mass of DM-pair, $M(DM,DM)$, than the other theory, one expects the flatter $E_T^{\text{miss}}$ distribution for the first one. The shapes of $M(DM,DM)$ distributions for EFT operators from Table 1 indicate that the mean values of $M(DM,DM)$ distributions for vector DM operators are larger than those for most of fermion DM operators which are in their turn have higher mean value of $M(DM,DM)$ than most of scalar DM operators. So one can observe the connection of $M(DM,DM)$ distributions shape and the slope of the $E_T^{\text{miss}}$ which was presented in Fig. 1.

One should stress that non-collider DM searches play an important complementary role in probing DM parameter space. As an example in Fig. 2 (left) we present the non-collider constraints for the operators D2, which exhibit pseudo-scalar interactions of fermion Dirac DM with quarks.

One can see that even for momentum-suppressed operator D2 (because of its pseudo-scalar nature) DM DD constraints from Xenon play an important role which is comparable to collider constraints, presented in Fig. 2 (right). It is important to stress that both LHC and DM DD searches set an upper limit on value of $\Lambda$. The LHC limit is of the order of 1 TeV.
Figure 1: $E_T^{\text{miss}}$ parton level distributions for a representative subset of the EFT operators from Table I for 13 TeV LHC energy and $M_{DM} = 10$ GeV.

Figure 2: Left: Non-collider constraints on D2 operator with fermion DM: (i) SI DM DD searches (shaded blue region below the lowest blue contour), (ii) constraints from relic density (above the yellow dashed line), (iii) constraints from the CMB (shaded green area) and (iv) constraints from the validity of the EFT ($\Lambda > 2m_{DM}$). Right: LHC monojet constraints on D2 EFT operator. The area inside the red, orange and blue solid curves is excluded by current LHC data at 95% CL for $\mu = 1, 6$ and $4\pi$, respectively. The projected LHC limits for 300 fb$^{-1}$ are indicated by dashed thin lines. The combined exclusion regions from CMB and DM DD searches for $\mu = 1$ are given by the light-purple area. See details and complete set of plots in Ref [28].
for present LHC data while DM DD searches the limit strongly depend on the operator. For example for non-suppressed operators conserving parity the limit on $\Lambda$ is about 3 orders of magnitude above the LHC one. On the other hand LHC limit is beyond DM DD searches for operators with suppressed elastic scattering cross sections on the nuclei (C2,C4,C6,D2,D3,D4,D6-10,V2,V4-V10). Moreover for operators with pseudo-vector currents which have suppressed DM DD rates, one should take into account the effect of their running from TeV energy scale at the LHC down to low energy scale at DM DD experiments, due to which an operator acquires non-negligible vector component [30, 31, 32].

3. Beyond EFT

The analysis of $E_T^{miss}$ shape presented here can be applied to different scenarios, beyond the EFT approach in general, where the DM mediator is not produced on-the-mass-shell, such as the case of t-channel mediator or mediators with mass below $2M_{DM}$, where the $M_{inv}(DM,DM)$ is not fixed. This case covers a wide range of theories.

![Figure 3: Left: Normalized signal (dotted blue and dashed red) and $Zj$ background (solid black) parton-level $p_T^j$ distributions for the 13 TeV LHC for the NSUSY scenario: normalised signal and $Zj$ background distributions. See details in Ref [33]. Right: Normalized $E_T^{miss}$ from $h_1 h_2$ i2HDM signal vs background for $\sqrt{s} = 13$ TeV. See details in [34].](image)

As an example in Fig. 3(left) the normalised shape for $E_T^{miss}$ distribution from $pp \to \chi_1^+ \chi_1^- / \chi_1^0 / \chi_1^0 \to \chi_1^0 \chi_1^0 +$ soft leptons/jets Minimal Supersymmetric Model(MSSM) signal and its dominant irreducible background $Z + jet \to \nu \bar{\nu} + jet$ ($Zj$) is presented for LHC@13TeV [33]. The model parameter space in the compressed chargino-neutralino scenario driven by small $\mu$ parameter is essentially characterised only by DM ($\chi_1^0$) and chargino ($\chi_1^\pm$) masses and mildly depends on the value of $\tan \beta$.

In the Fig. 3(right) we present $E_T^{miss}$ from $h_1 h_2 j$ inert two Higgs doublet model (i2HDM) signal alongside the estimated (by CMS) experimental background for $\sqrt{s} = 13$ TeV. The parameter space of the model, details of which can be found for example in [34], is characterised by DM mass ($h_1$), the mass of the second neutral scalar ($h_2$), the mass of the charged scalar ($h^+$) and the DM-Higgs boson coupling. However the $h_1 h_2 j$ production cross section depends only on two parameters – $h_1$ and $h_2$ masses in analogy to the above SUSY scenario. An important feature of the signal versus background shapes in these completely different theory cases is that
the background falls more rapidly with $E_T^{miss}$ and the difference in the slope with respect to the signal is bigger for higher DM mass. This behaviour has the same explanation as for EFT study case above – it is related to the bigger invariant mass of the invisible system for the signal – $M(DM, DM)$ than for the background – $M_Z$. This feature is very instrumental to increase signal-to-background ratio (S/B) (which is typically below 1% for low $E_T^{miss}$ cuts) by increasing the value of $E_T^{miss}$ or by performing the signal-background shape analysis \[34\].

The complementary role of non-collider DM searches is also crucial in case of these two complete and consistent models. In case of MSSM LHC will be able to cover the region inaccessible by Xenon1T in small $\Delta M$ region, while Xenon1T is able to cover $m_{DM}$ well beyond the LHC reach for $\Delta M > 3 - 5$ GeV \[33\]. In case of i2HDM model collider sensitivity with mono-jet signature is even more limited because of the lower production rates of the scalar DM, $h_1$, or its inert partners ($h_2$ and $h^+$) and expected LHC reach is below 100 GeV for $M_{h_1}$.

4. Beyond mono-$X$ signature
While mono-$X$ (with $X$ being jet, $\gamma$, $Z$, $H$, $t$ etc.) DM signatures at colliders are the most general ones, their rates are typically very low (usually at the percent level or even lower). Besides several other interesting but model-specific DM signature studies, one should stress one signature which can be also considered as quite generic one. In case when DM, $D^0$, is embedded into electroweak multiplet and its mass split from the charged odd particle(s), $D^+$, is generated only radiatively (preserving gauge invariance), the one can find that the value of this mass split is of the order of 0.2 GeV. In this case $D^+$ has a very small width and respectively large life-time. $D^+$ being long lived particle (LLP) dominantly decays into DM and very soft pion: $D^+ \rightarrow D^0 \pi^0$. Production of $D^+$ in pairs or in association with DM leads then to the typical signature from charged LLP: disappearing charged track (DCT) as soon as the track from LLP is long enough (from few cm to a meter). In case of such signature the S/B ratio is much higher than in case of mono-jet signal and therefore, substantially bigger DM masses can be probed with charged LLPs from DM sector \[33\] \[36\] \[37\]. As an example, we would like to present here results for the minimal vector triplet DM ($V^0$) model \[37\] which predicts the right amount of DM for $M_{DM}$ in the 3-4 TeV range depending on DM coupling to the Higgs boson.

In this model the SM is supplemented by a new massive vector boson $V_\mu$ in the adjoint representation of $SU(2)_L$, e.g. by two new massive vector particles: $V^0$ and $V^\pm$. If $V_\mu$ transforms homogeneously (i.e. $V_\mu \rightarrow g_{M} V_\mu g_L$ where $g_L \in SU(2)_L$) and $Z_2$ symmetry is imposed (which links the quartic $V$ coupling to the gauge coupling constant and makes theory unitary before EW symmetry breaking and in the absence of the Higgs boson as found in Ref. \[38\]) then Lagrangian can be written as:

$$\mathcal{L} = \mathcal{L}_{SM} - Tr \{D_\mu V_\nu D^\mu V^\nu\} + Tr \{D_\mu V_\nu D^\mu V^\nu\} - \frac{g^2}{2} Tr \{[V_\mu, V_\nu] [V^\mu, V^\nu]\} - igTr \{W_\mu [V^\mu, V^\nu]\} + \tilde{M}^2 Tr \{V_\mu V^\nu\} + a (\Phi^+ \Phi) Tr \{V_\mu V^\nu\}$$

where $D_\mu = \partial_\mu - i g [W_\mu,]$ is the usual $SU(2)_L$ covariant derivative in the adjoint representation and $\mathcal{L}_{SM}$ represents the SM Lagrangian. The main difference with respect to the model in Ref. \[38\] is that the $SU(2)_L$ symmetry is broken by the Higgs mechanism and the associated gauge bosons have mass. We thus allow for a coupling of $V$ to the Higgs scalar field $\Phi$.

Due to the $Z_2$ symmetry the neutral new vector boson, $V^0$ is stable and therefore is the perfect DM candidate. The mass split, $\Delta M$, between $V^0$ and $V^\pm$ is generated radiatively and its value is just above the pion mass which makes $V^\pm$ long lived \[37\].

In Fig. 4 we present results for spin-independent cross-section for $V^0$-nucleon elastic scattering as a function of $M_V$ and for representative values of $a$. It is very important to note that Xenon1T experiment combined with DM relic density constraints excludes DM mass above 4 TeV.
Figure 4: Spin-independent cross-section for $V^0$-nucleon elastic scattering as a function of $M_{V}$ and for representative values of $a$. The continuous black curve represents the elastic cross-section computed with the values of $M_{V}$ and $a$ that saturate the measured DM relic density. The grey dashing highlights the parameter space where perturbative unitarity loss occurs at too low scale. See details in [37].

At the same time the production rate $\sigma_{eff}$ for $\sigma(pp \to V^\pm V^0) + 2\sigma(pp \to V^+ V^-)$ process which leads to the disappearing charge track signatures (that is why $pp \to V^+ V^-$ process comes with coefficient 2 which is equal to the number of disappearing charged tracks it provides) are high enough to probe this process at colliders. In particular, the current LHC@13TeV limit on $M_{V}$ is about 1.4 TeV [37]. Moreover, it was found [37] that 100 TeV collider will be able to exclude DM mass below 4 TeV, thus allowing to probe the entire parameter space of the model. One should also note that in case of i2HDM, DCT signature also allows to substantially enhance LHC potential and probe DM mass upto about 500 GeV [35] which is much higher than 100 GeV – the maximum DM mass which can be probed via mono-jet signature.

5. Towards decoding framework

While there are many comprehensive studies of the phenomenology of various appealing BSM models, exhibiting “top-bottom” approach, there is no framework at the moment which can solve the reverse engineering task – the task of decoding the nature of DM. It is not surprising why – we are all eagerly looking for the signal first of all and busy with the interpreting and exploring our own models. A huge amount of work has been done on the model building, phenomenology and experimental searches as well as on building different tools, examples of which has been given above. And there is really huge potential of combining different methods and signatures to probe different models. What is missing is the framework which joins all these pieces in one tool which would help us to decode underlying theory, in particular its part related to DM. The task of decoding of the whole underlying theory sounds probably too ambitious to the author, while decoding of its DM part sound more realistic since it contain specific and possibly much smaller piece of the theory.

This framework requires the database of models, database of various signatures and set of tools which will be able to effectively explore not only parameter space of each particular model, but also the model parameter space and compare predicted signatures with the observed ones. Such a framework would allow objectively judge about preferred model or set of models which would fit signal best of all. An example of the prototype of such a framework actually already exists in the form of High Energy Physics Model DataBase (HEPMDB).
HEPMDB is created as a web-server accessible to everybody and is able to:

(i) collect HEP models for all multipurpose Matrix Element (ME) generators in the form of Feynman rules and parameters written in the format specific for a given package;
(ii) collect models’ sources which can be used to generate HEP models for various ME generators using FeynRules [39] or LanHEP [40];
(iii) allow users to perform simulations for their own models or models available at HEPMDB using the full power of the High Performance Computing (HPC) IRIDIS cluster standing behind the HEPMDB itself. Connection to HPC cluster is one of the key features of the HEPMDB: it provides a web interface to ME generators such as CalcHEP[41], Madgraph [42] and Whizard [43] which can be easily run on the HPC cluster. This allows users to avoid problems related to installing the actual software, which can sometimes be quite cumbersome;
(iv) collect simulated events and plot distributions using web interface.

Though the signature database at HEPMDB is at the development stage, users can indicate some essential features of the signatures which model can provide, such as presence of resonance, \(E_T^{\text{miss}}\) etc. Probably the most important feature of HEPMDB is that it can be efficiently developed by the whole HEP community. There are more than 300 registered users at HEPMDB and more than 50 models. Any registered user can add his/her own model and signature, use HEPMDB for his/her event simulation. The next step of HEPMDB development is the addition of various packages for event and DM analysis(e.g. microMEGAs[44, 45]), comprehensive database of signatures. This will make another step on development of HEPMDB towards the framework for identification of underlying theory when the experimental signal will be observed.

6. Conclusions
In the absence of a DM signal we can still do a lot – we can prepare ourselves for its discovery and identification. \(E_T^{\text{miss}}\) shape is quite instrumental in understanding the underlying theory at colliders, while direct and indirect DM searches are very powerful in complementing collider searches especially in the parameter space with large DM mass. We also advocate the usage of new DM signatures such as disappearing charge tracks which allows to substantially extend collider exploration of large DM mass. We show that collider and non-collider DM searches have a unique power to decode the nature of Dark Matter on the examples of several appealing DM theories. Such complementarity and usage of different signatures would allow us to decode the nature of DM, signals from which we are expecting in the near future. However, for this decoding we need an efficient framework which would combine the experience of HEP community and would allow to identify effectively the underlying theory of DM from the experimental signal which we will hopefully observe in the near future. We present one of the tools which could be part of this framework – High Energy Physics Model Database (HEPMDB) which was created to make a step forward towards solving this problem.

Acknowledgments
Author would like to thank the organisers of ACAT2019 workshop for the invitation and warm hospitality and acknowledges partial support from the STFC grant ST/L000296/1 and NExT Institute.

References
[1] Ade P A R et al. (Planck) 2016 Astron. Astrophys. 594 A13 (Preprint [1502.01589])
[2] Blumenthal G R, Faber S M, Primack J R and Rees M J 1984 Nature 311 517–525 [96(1984)]
