Unsupervised tagging of semivisible jets with normalized autoencoders in CMS

Florian Eble, Annapaola de Cosa, Roberto Seidita

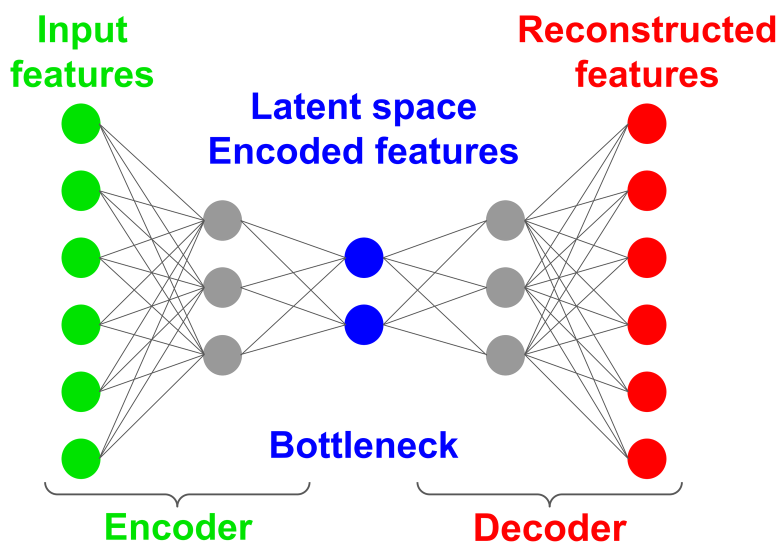
Semivisible jets

Semivisible jets [1] (SVJ) are a new physics signature arising in Hidden Valley theories where the dark sector is made of dark quarks interacting via a confining SU(N) force (dark QCD). In this family of models, dark quarks are expected to hadronize in the dark sector, forming dark bound states. A fraction of them is unstable and promptly decays back to Standard Model (SM) quarks, which then hadronize in the SM sector. The resulting SM hadrons form a jet interspersed with invisible dark hadrons, dark matter candidates. The different jet substructure of SVJs, due to the double hadronization step and invisible dark hadrons, can be exploited to classify them versus SM jets. Because the true nature of dark QCD (if it exists) is unknown, unsupervised machine learning algorithms are well suited tools to perform a model-independent search for SVJs.

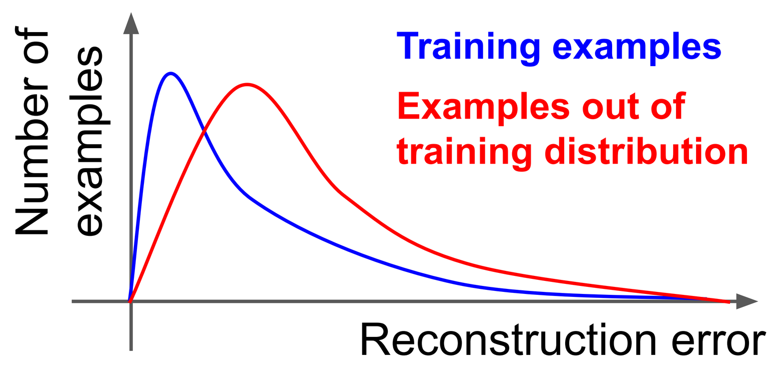
Autoencoders

Autoencoders (AE) are neural networks composed of two parts:

- an encoder, which maps the input features space to a lower dimensional latent space,
- a decoder which maps the latent space to an output space with same dimension as the input space.



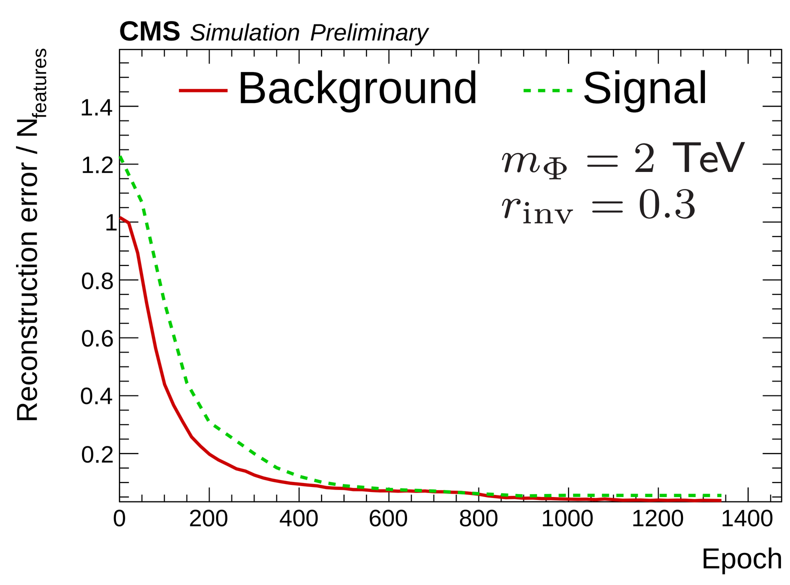
AEs are trained to minimize the reconstruction error between input and output, such that examples out of the training distribution have a higher loss. Trained on SM data, AEs can thus perform signal-agnostic searches for new physics [2, 3]. In the case of SVJs, AEs are trained on SM jets.



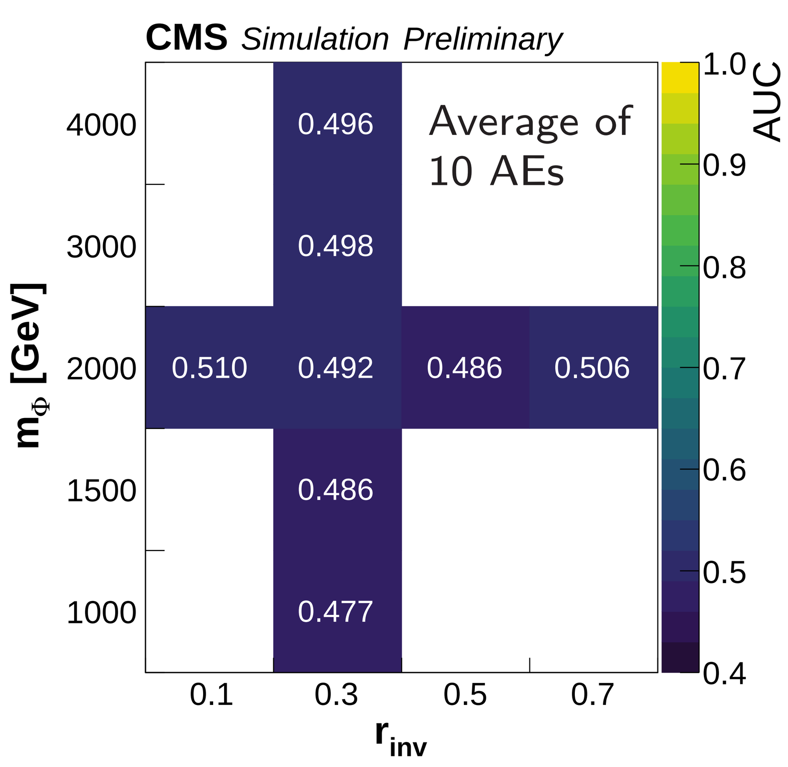
The problem of out-of-distribution reconstruction

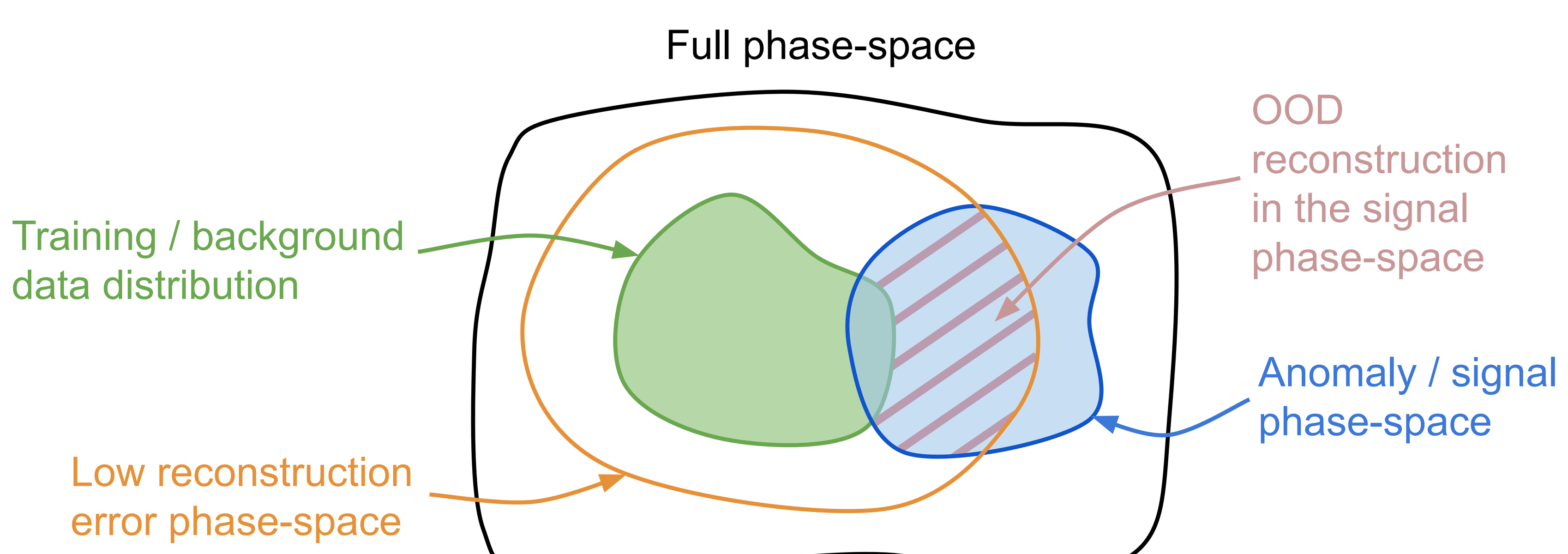
AEs were proven to well perform anomalous detection of SVJs versus QCD [4] but achieve poor classification of SVJs versus top-quark jets.

10 AEs were trained on a dataset of top-quark jets until minimal validation loss. They take 8 jet substructure input features, mapped to a normal distribution. The architecture is a fully connected network with 10, 10, 6, 10, 10 neurons.[1]



The AEs generalize (reconstruct with low error) out of the training phase-space (out-of-distribution, OOD), in particular in regions where SVJs are present: the average reconstruction error for background (SM) and signal (SVJ) jets is the same. This results in low anomaly detection performance.





[1] Input features and technical details available in this note:

Normalized autoencoder

Normalized autoencoders [5] (NAE) suppress OOD reconstruction by learning the training data probability distribution $p_{\rm data}$. The NAE model probability p_θ is defined to assign high probability to low reconstruction error (E_θ) examples:

$$p_\theta(x) = \frac{1}{\Omega_\theta}\exp\left(-E_\theta(x)\right)$$

Examples following p_θ are obtained by sampling via a Langevin Markov Chain Monte Carlo (MCMC) ("negative examples"). The loss function is the difference between the reconstruction error of the training ("positive") examples and of the negative examples:

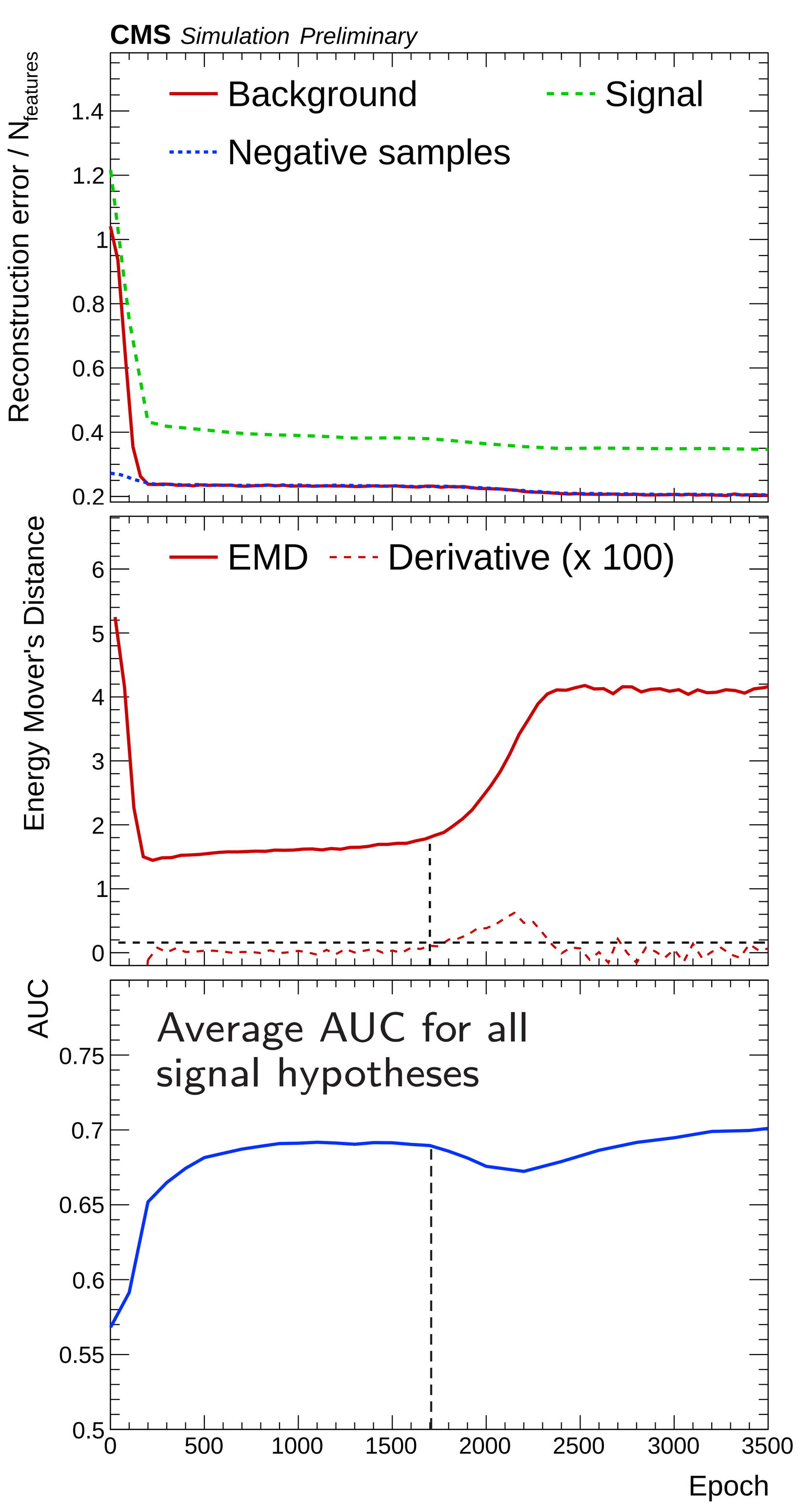
$$\mathbb{E}_{x\sim p_{\rm data}}\left[L_\theta(x)\right] = \underbrace{\mathbb{E}_{x\sim p_{\rm data}}\left[E_\theta(x)\right]}_{\text{positive energy } E_+} - \underbrace{\mathbb{E}_{x'\sim p_\theta}\left[E_\theta(x')\right]}_{\text{negative energy } E_-}$$

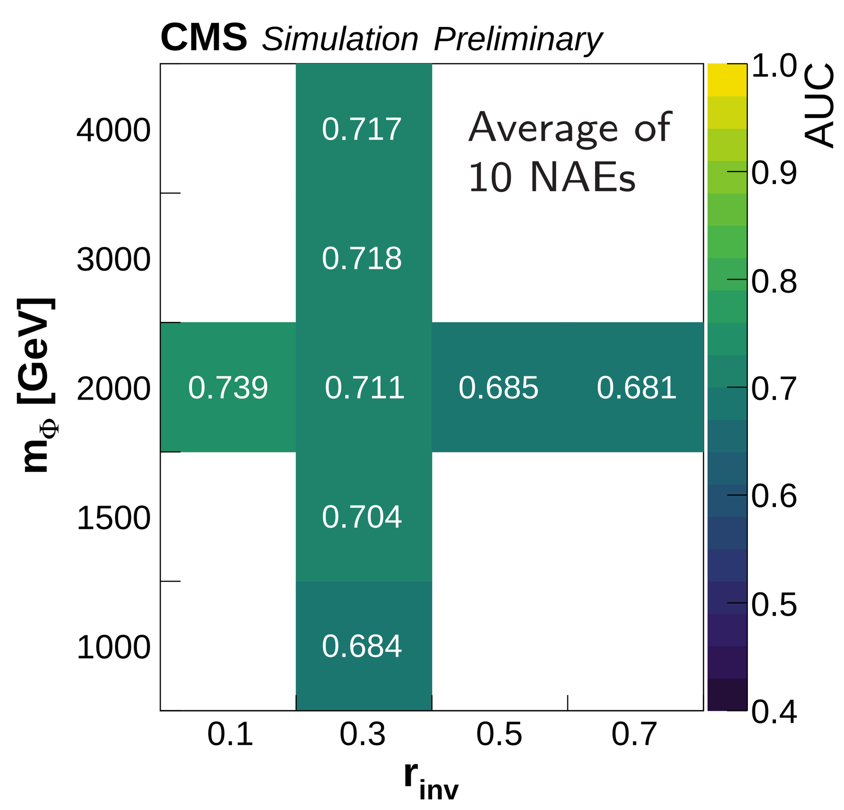
Unsupervised SVJ tagging versus top jet

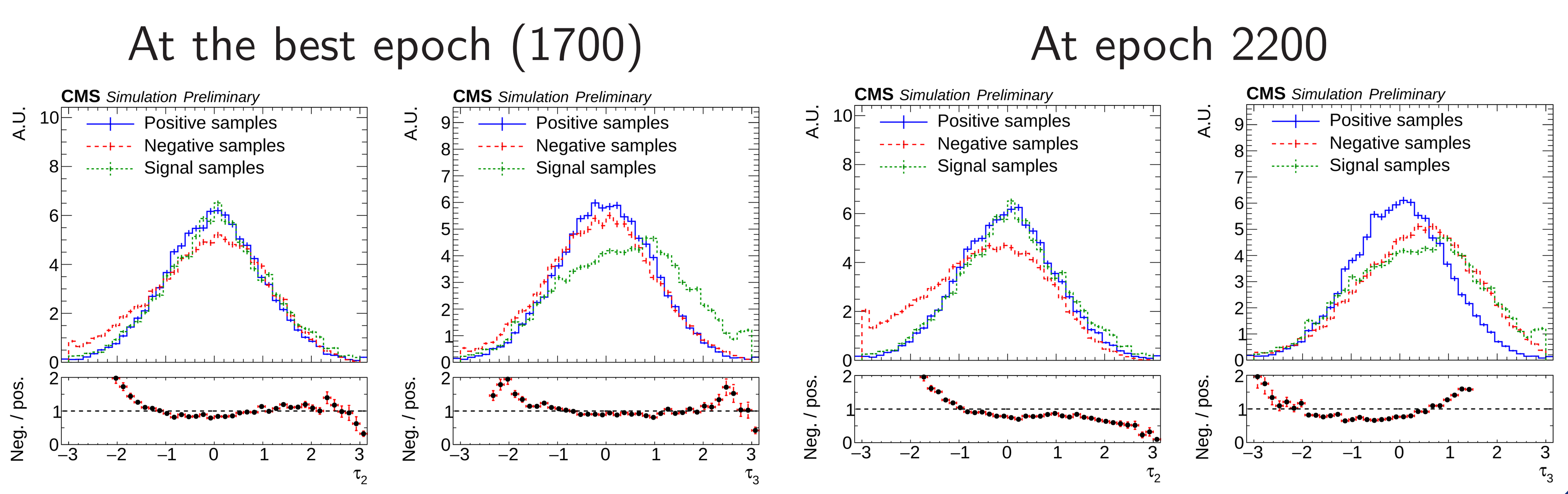
The loss function was modified to prevent the divergence of negative energy and minimize positive energy while the energy difference is close to 0:

$$L = \log\left(\cosh\left(E_+ - E_-\right)\right) + \alpha E_+$$

The Energy Mover's Distance (EMD) is used to quantify the distance between the training and the negative samples in the input feature space. As the positive energy is minimized beyond a certain value, the EMD increases: the network cannot better reconstruct training examples and suppress OOD reconstruction at the same time. The best epoch is just before the EMD increase: minimal OOD reconstruction and maximal training examples reconstruction. **This is a fully signal-agnostic procedure to train a NAE, not using signal SVJs simulation.**







References

[1] T. Cohen, M. Lisanti, H. K. Lou, and S. Mishra-Sharma. LHC Searches for Dark Sector Showers. *JHEP*, 11:196, 2017.

[2] T. Heimel, G. Kasieczka, T. Plehn, and J. M. Thompson. QCD or What? *SciPost Phys.*, 6(3):030, 2019.

[3] M. Farina, Y. Nakai, and D. Shih. Searching for New Physics with Deep Autoencoders. *Phys. Rev. D*, 101(7):075021, 2020.

[4] F. Canelli, A. de Cosa, L. L. Pottier, J. Niedziela, K. Pedro, and M. Pierini. Autoencoders for semivisible jet detection. *JHEP*, 02:074, 2022.

[5] S. Yoon, Y.-K. Noh, and F. Park. Autoencoding under normalization constraints. In *ICML*, pages 12087–12097. PMLR, 2021.