Loss functions

Stage 1 and 2

- **Pre-Train, Encoder**   *(Based on cosine Similarity)*

  - 1. Local Cohesion | Ltrk | "Connect the dots" of a single particle track. | Prototype Contrastive |

  - 2. Global Grouping |Lpri| Group all fragments of a single shower by origin. | Prototype Contrastive |

  - 3. Semantic Clustering| Lpid| Identify the physical type of hit. | Prototype Contrastive |

- **Pre-Train, Decoder**

  - 4. Occupancy | Locc| Predict if a masked voxel was empty or contained a hit. | Focal Cross-Entropy |

  - 5. Regression | Lreg | Predict the energy of a masked hit. | Huber Loss (Charge-Weighted)|

- **The Fine-tuning**

  - **Classification:** Flavour - Cross-Entropy, Charm - BCEWithLogits

  - **Regression**: Smooth L1 (Huber)

  The total loss is a weighted sum with learnable uncertainty parameter, σ_i

$$\mathcal{L}_{\text{total}} = \sum_{i=1}^{5} \left(\frac{1}{2\sigma_i^2} \mathcal{L}_i + \log \sigma_i \right)$$

Neutrino Production

From ATLAS IP to FASER

- At $s = 13$ TeV, the inelastic cross section is approximately, σ_inel 13 TeV ~ 75 mb

- For LHC Run 3 (250 fb-1), this corresponds to ~10^16 light particles produced

  - Angular spread of meson decays $\theta \sim m\pi/E(TeV)$~ mrad

Particle	Decay mode	Branching fraction (approx.)
π^+	$\pi^+ \to \mu^+ \nu_\mu$	99.9877%
	$\pi^+ \to e^+ \nu_e$	1.23×10^{-4} (helicity suppressed)
K^+	$K^+ \to \mu^+ \nu_\mu$	63.56%
	$K^+ \to \pi^0 e^+ \nu_e$ (Ke3)	5.07%
D^0	Inclusive semileptonic	$\mathcal{B}(D^0 \to X e^+ \nu_e) \approx 6.46\%$
D^+	Inclusive semileptonic	$\mathcal{B}(D^+ \to X e^+ \nu_e) \approx 16.13\%$
D_s^+	$D_s^+ \to \tau^+ \nu_\tau$	5.36%