Differentiable Vertex Fitting for Jet Flavour Tagging

Rachel E. C. Smith[1], Inês Ochoa[2], **Rúben Inácio**[2], Jonathan Shoemaker[1], and Michael Kagan[1]

[1]SLAC National Accelerator Laboratory [2] Laboratory of Instrumentation and Experimental Particle Physics, Lisbon

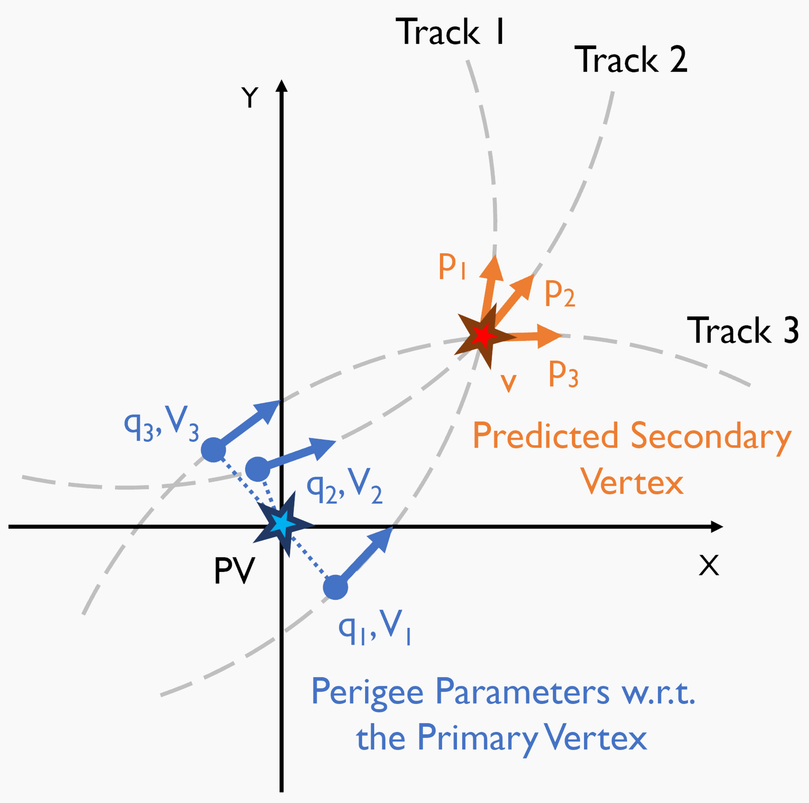
https://arxiv.org/abs/2310.12804

1. Motivation

Flavour tagging is essential for studying a wide array of physical processes at the LHC. It relies on the unique properties of heavy quark hadrons, including the presence of a secondary vertex (SV) displaced from the primary collision. The current state-of-the-art models use modern neural networks (NNs) that do not explicitly fit SVs. **Can we integrate vertex fitting into end-to-end ML trainable models?**

2. Secondary Vertexing

- **Estimate a common vertex** that originated a set of tracks.
- Formulated as an inclusive vertexing task, using the **Billoir algorithm**.
- Least square objective, denoted as $\mathcal{S}$:
  - $\mathbf{q_i}$: primary vertex
  - $\mathbf{V_i}$: covariance matrix
  - $\mathbf{v}$: vertex position
  - $\mathbf{p}_i$: track momentum at the vertex
  - $\mathbf{h}_i(\mathbf{v}, \mathbf{p}_i)$: track model
  - w_i: weight of the track between 0 and 1 that represents representing how much it contributes to the fit



$$\mathcal{S} = \chi^2 = \sum_{i=1}^{N} w_i(\mathbf{q}_i - \mathbf{h}_i(\mathbf{v}, \mathbf{p}_i))^T \mathbf{V}_i^{-1}(\mathbf{q}_i - \mathbf{h}_i(\mathbf{v}, \mathbf{p}_i))$$

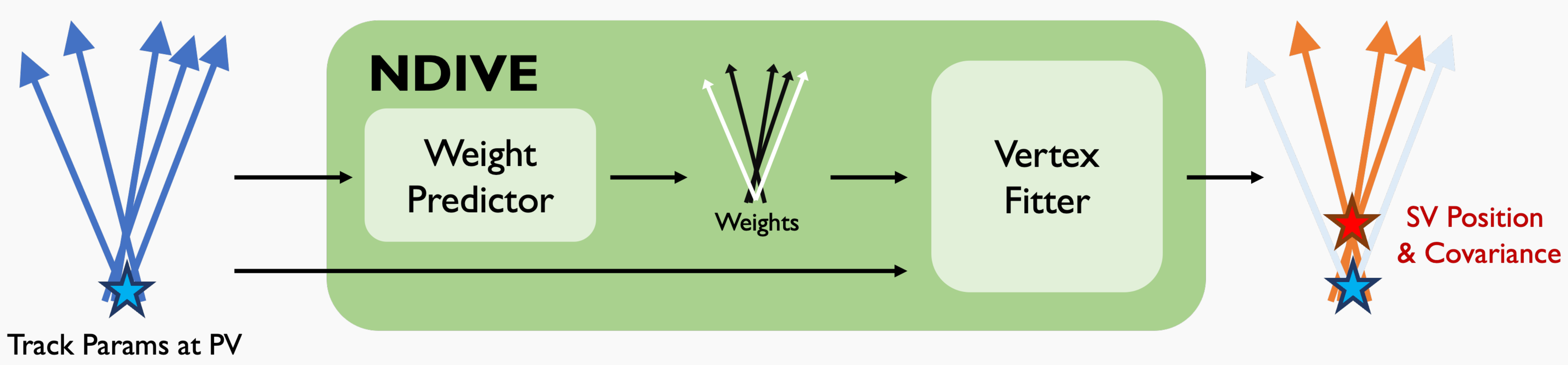
3. Differentiable optimization with custom derivatives

- We minimize an objective function $\mathcal{S}(\mathbf{x}, \alpha)$ by optimizing the value of $\mathbf{x}$ given the set of parameters α.
- Since $\mathcal{S}$ is continuously differentiable with non-singular Jacobian, we can use the **implicit function theorem**.
- Assuming $\hat{\mathcal{G}} \equiv \partial_{\mathbf{x}}\mathcal{S}(\hat{\mathbf{x}}, \alpha)$:

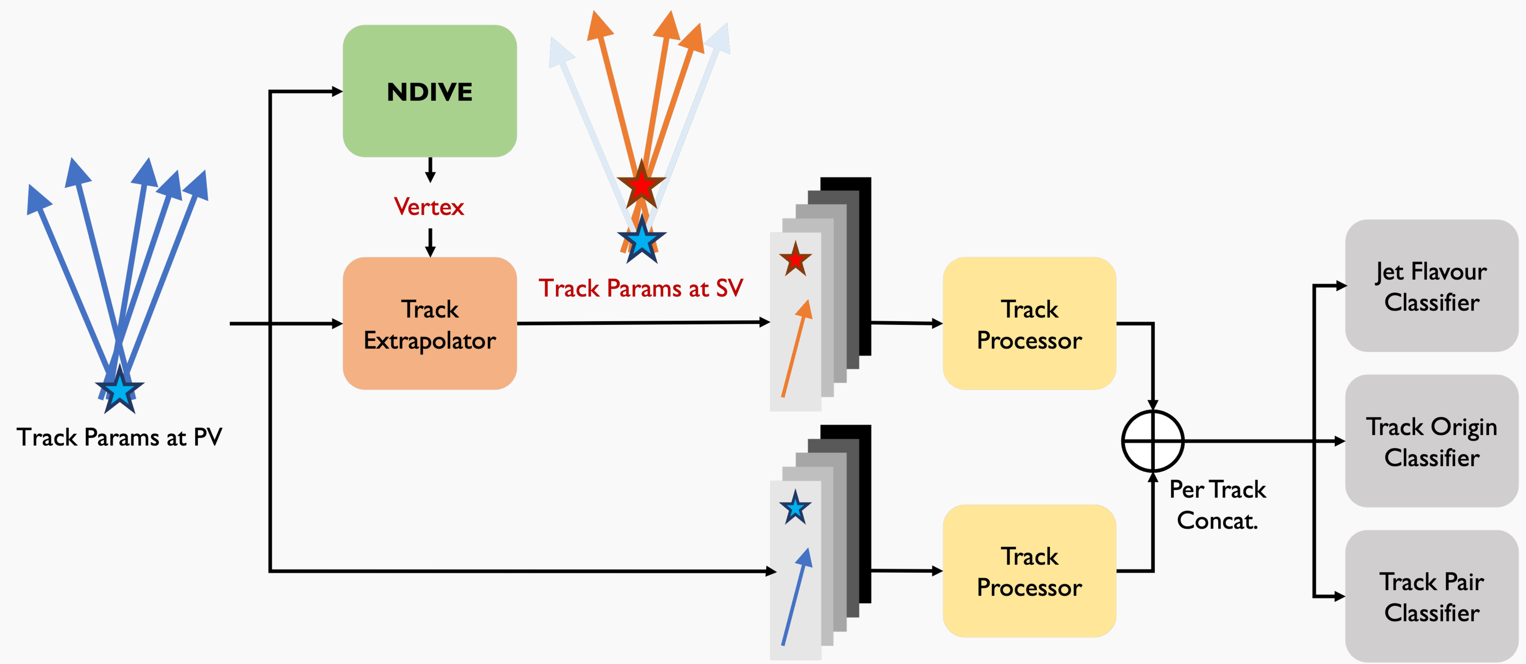
$$0 = \frac{d}{d\alpha}\hat{\mathcal{G}} = \frac{\partial\hat{\mathcal{G}}}{\partial\alpha} + \frac{\partial\hat{\mathcal{G}}}{\partial\mathbf{x}}\frac{\partial\mathbf{x}}{\partial\alpha} \Leftrightarrow \frac{\partial\mathbf{x}}{\partial\alpha} = -\left(\frac{\partial\hat{\mathcal{G}}}{\partial\mathbf{x}}\right)^{-1}\frac{\partial\hat{\mathcal{G}}}{\partial\alpha}$$

- This result defines a custom derivative function of the NN optimization in the **backward pass**.

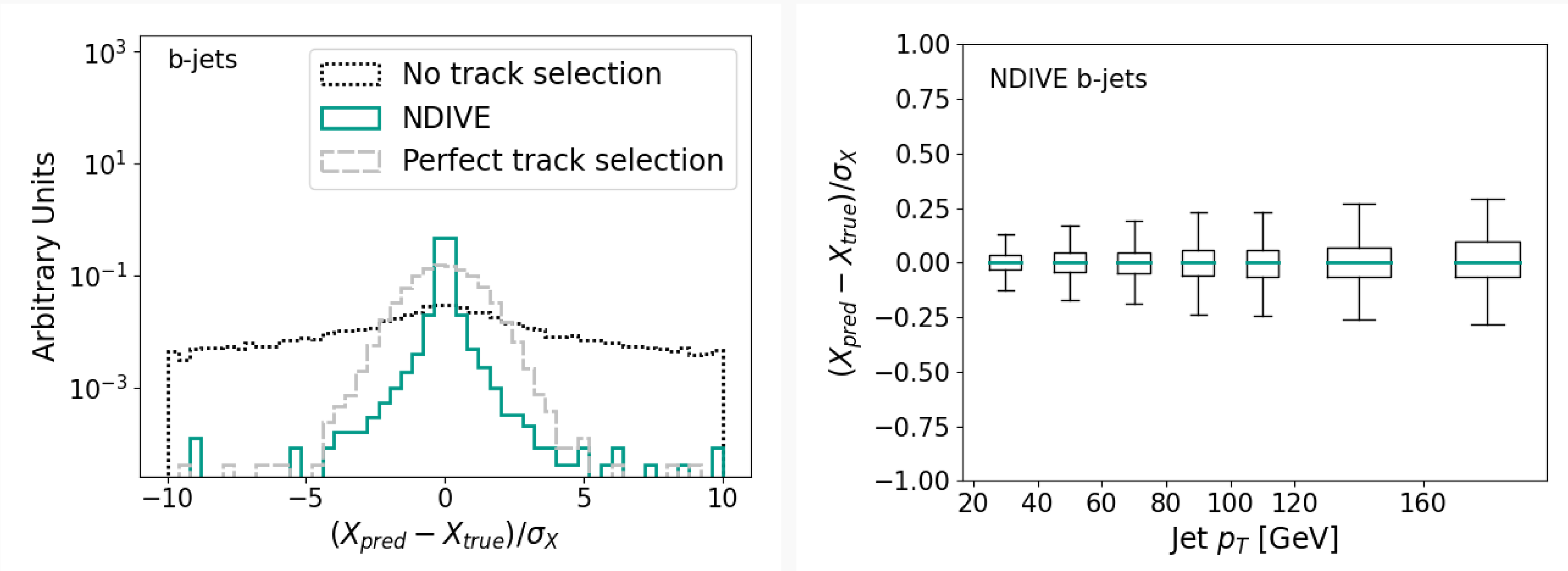
4. NDIVE



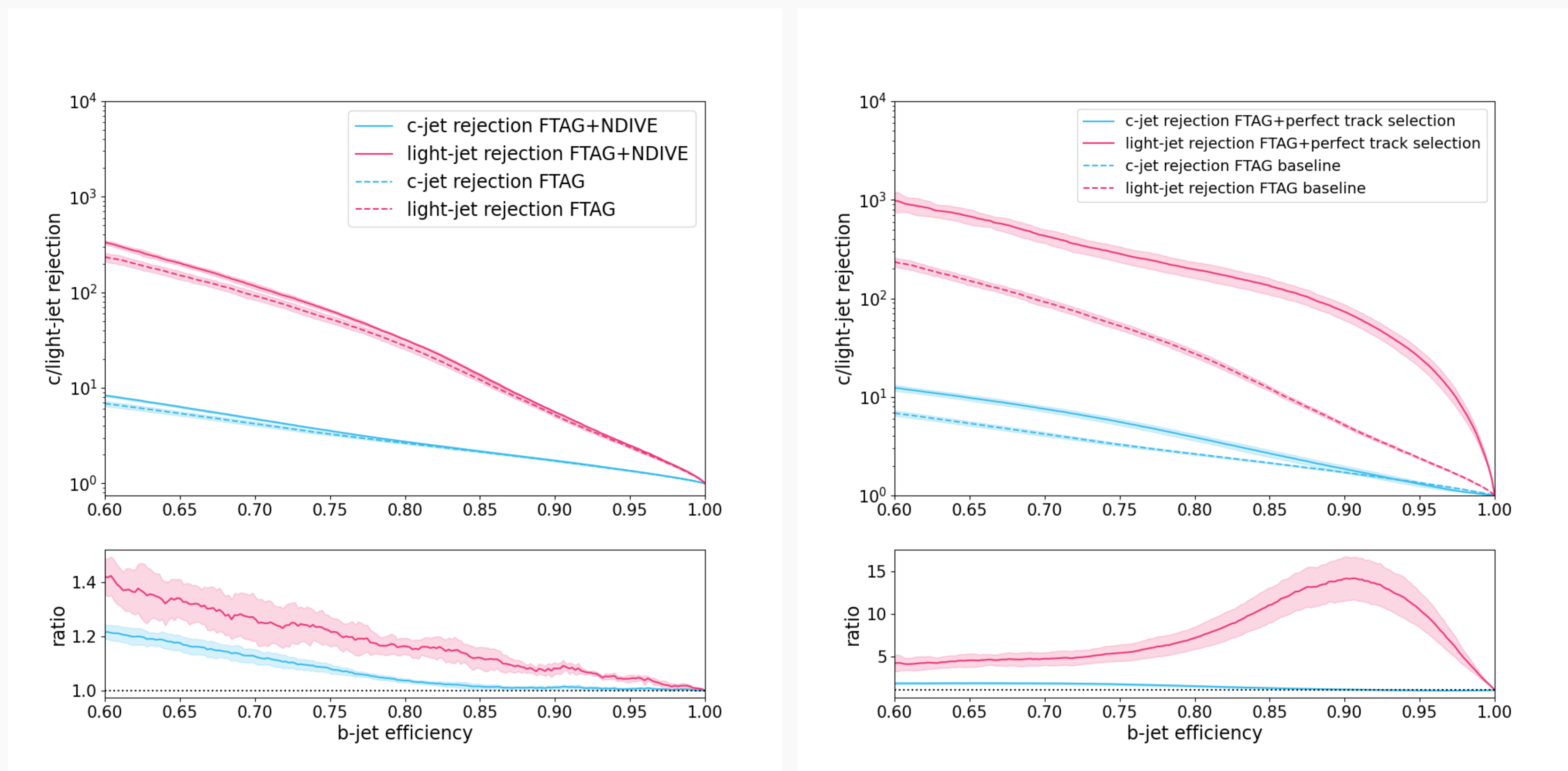
5. FTAG+NDIVE



6. Performance: vertex fitting and flavour tagging



- NDIVE is able to accurately estimate the SVs by providing unbiased predictions.



- NDIVE integration into FTAG provides improvements in the rejection of both c- and light-jets.
- Further room for improvement with better weight prediction.

7. Conclusion

- We **introduce the differentiable vertex fitting algorithm NDIVE** that can readily be integrated and jointly optimized in a larger flavour tagging NN model.
- **These methodological developments are generic**, applicable to other vertex fitting algorithms and other schemes for integrating vertex information into NNs.

Fundação para a Ciência e a Tecnologia

Fundação "la Caixa"

SLAC NATIONAL ACCELERATOR LABORATORY