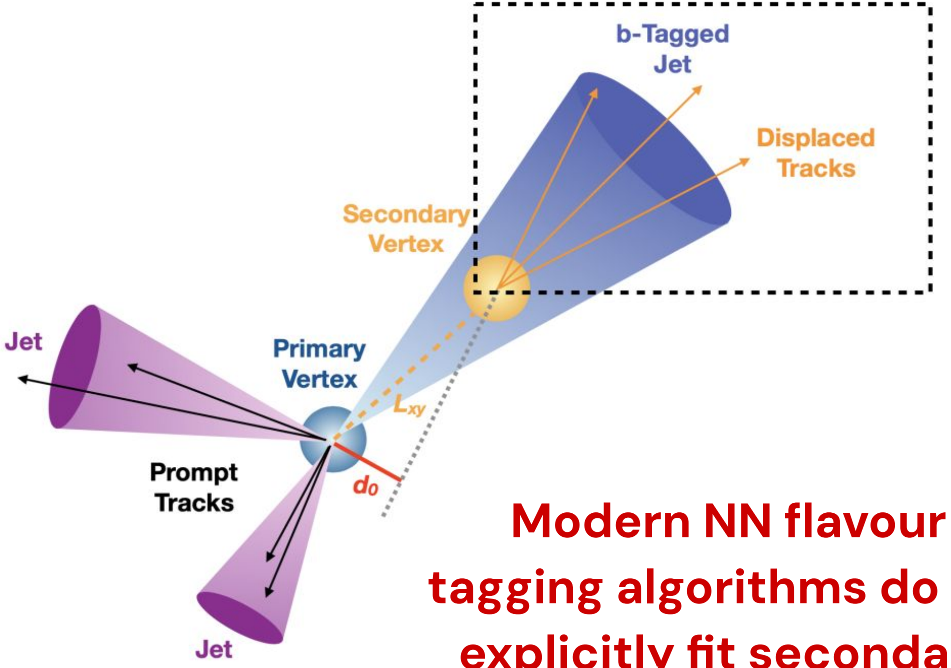
Differentiable Vertex Fitting for Jet Flavour Tagging

Rachel Smith, Inês Ochoa, Rúben Inácio,
Jonathan Shoemaker, Michael Kagan

EuCAIFCon24, Amsterdam
30 April 2024

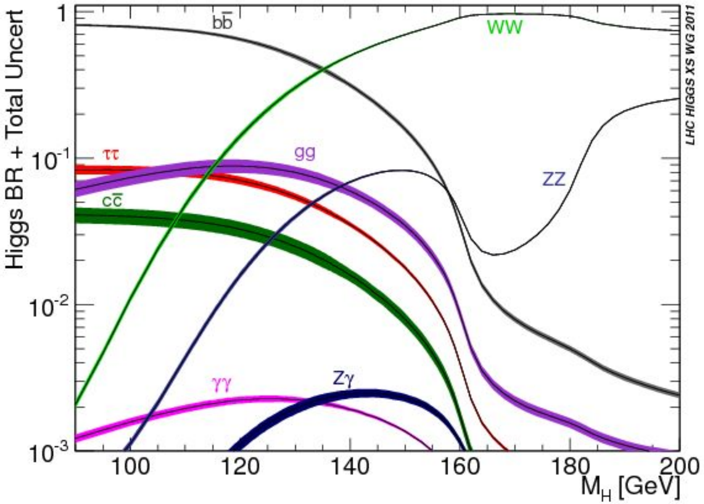
fct Fundação para a Ciência e a Tecnologia
"la Caixa"
LIP
SLAC
NSF
U.S. DEPARTMENT OF ENERGY

Jet flavour tagging in high energy physics

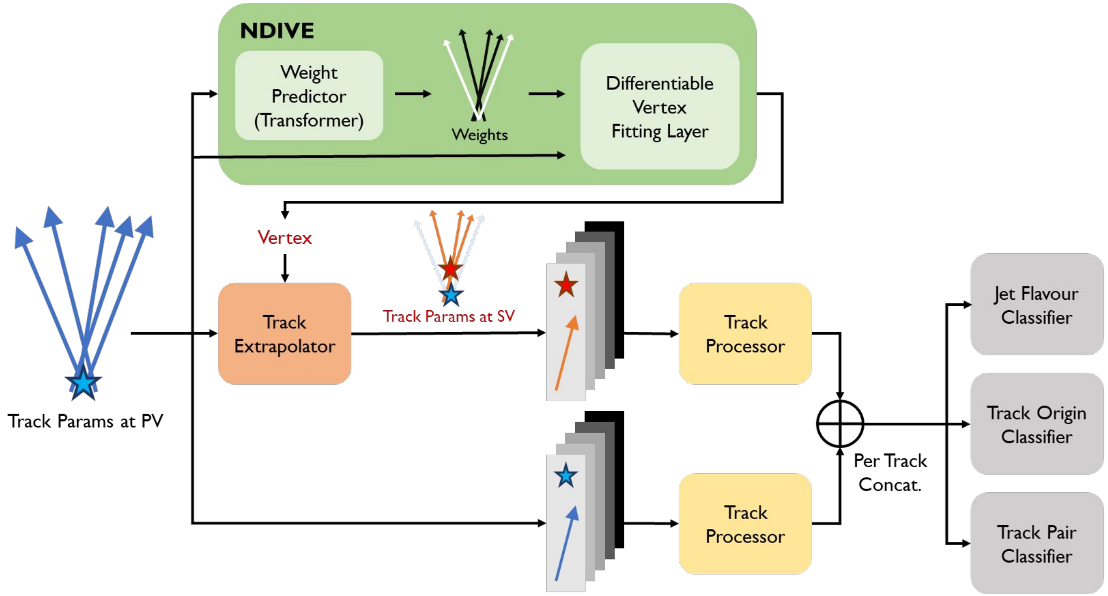


Modern NN flavour tagging algorithms do not explicitly fit secondary vertices...

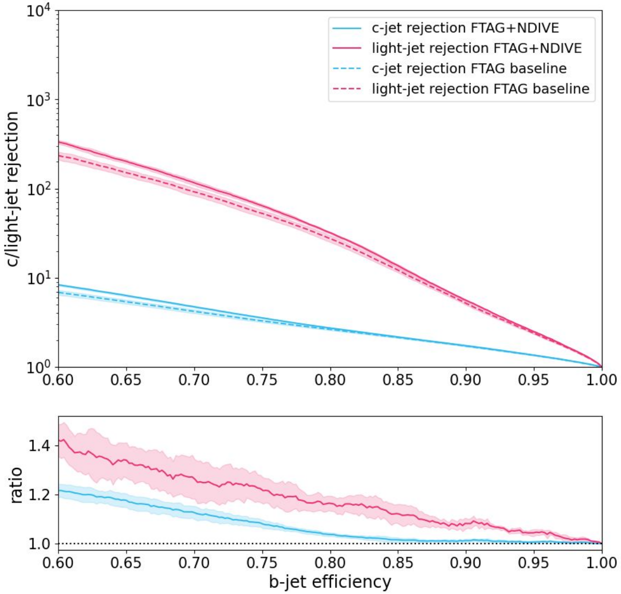
Standard Model Higgs boson decays preferentially to a pair of b-quarks



NDIVE (Neural DIfferentiable VErtexer)



NDIVE integration into NN flavour tagging model improves performance:

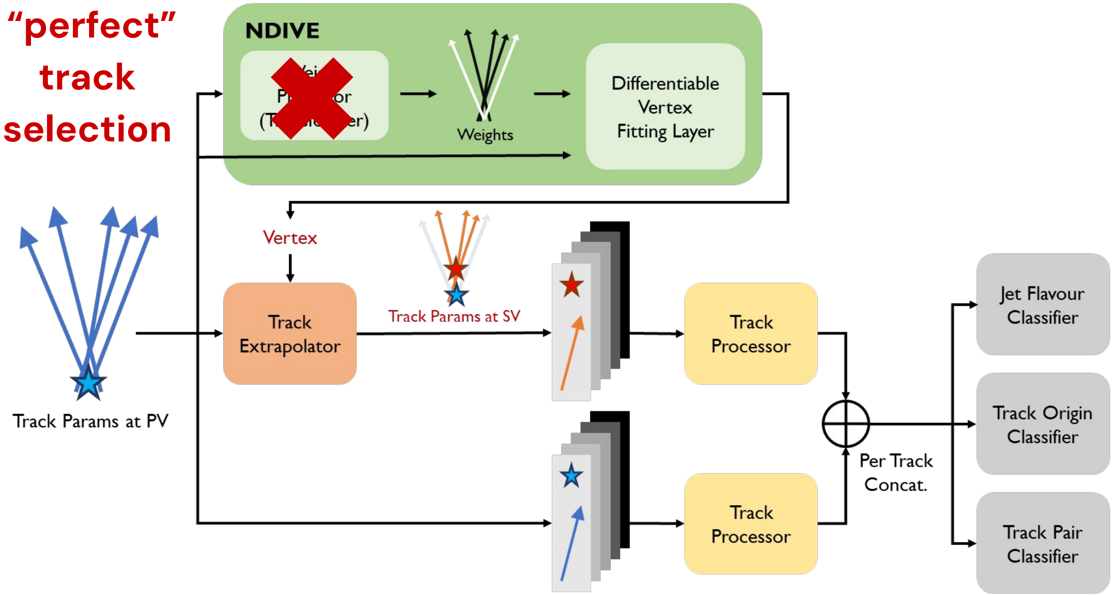


fully integrated and jointly optimizable;

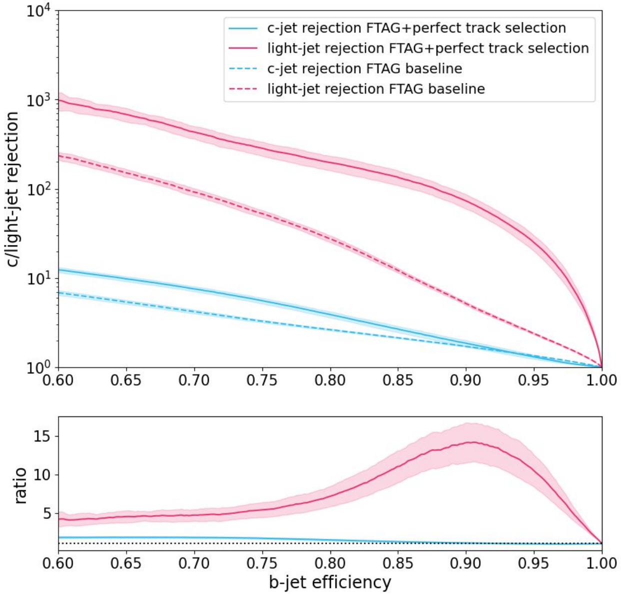
explicitly introduce physics knowledge into NNs!

NDIVE (Neural DIfferentiable VErtexer)

“perfect” track selection



NDIVE integration into NN flavour tagging model improves performance:



Significant improvements possible with better track selection!

Introduction

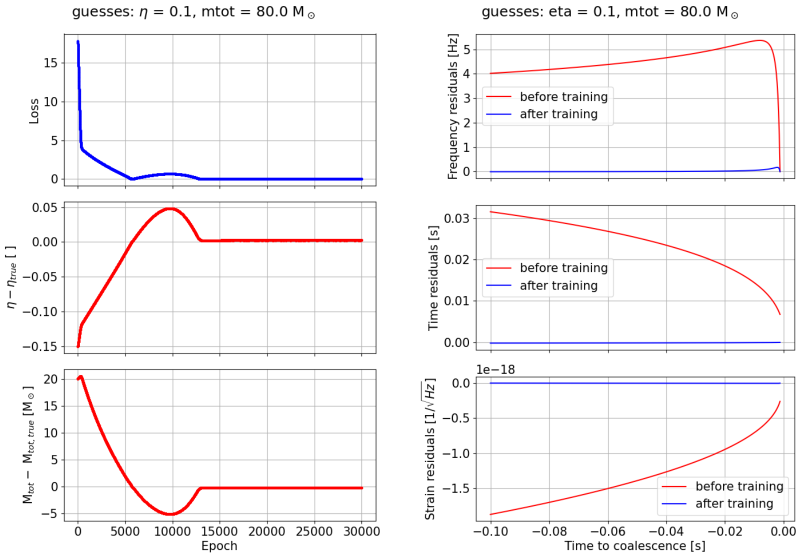
- **PINNGraPE** is a PyTorch algorithm which **does PE for a Gravitational-Wave (GW) signal's source thanks to a Physics-Informed Neural Network (PINN) [?].**

- **We solve (1) thanks to a Recurrent Neural Network** (RNN) with a Runge-Kutta integrator at 4^{th} order implemented inside.

$$\frac{df}{dt} = \mathcal{F}[f, \eta, M_{tot}] \tag{1}$$

$$\begin{aligned} \mathcal{L} =& \frac{\beta_f}{N}\sum_{k=1}^{N}|f_k - f(t_k)| + \\ &+ \frac{\beta_t}{N}\sum_{k=1}^{N}|t_k - t(f(t_k))| + \\ &+ \frac{\beta_h}{N}\sum_{k=1}^{N}|h_k - h(f(t_k))| \end{aligned} \tag{2}$$

Results

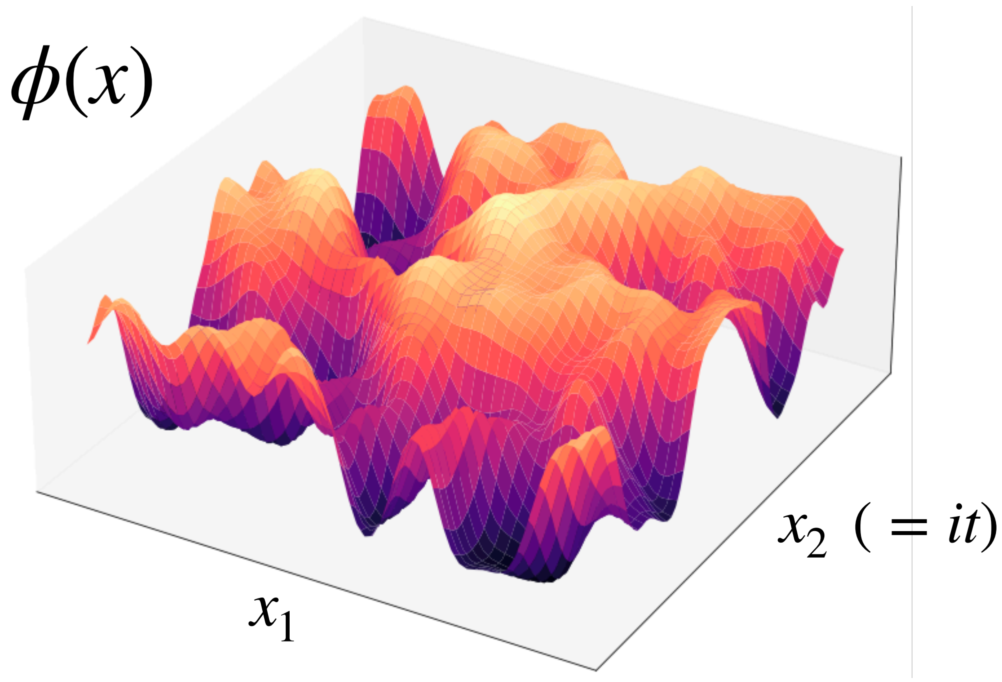

guesses: η = 0.1, mtot = 80.0 M⊙
guesses: eta = 0.1, mtot = 80.0 M⊙
Loss
η − η_true [ ]
M_tot − M_tot, true [M⊙]
Epoch
Frequency residuals [Hz]
Time residuals [s]
Strain residuals [1/√Hz]
1e−18
Time to coalescence [s]
before training
after training

Conclusions

- **PINNGraPE is able to infer η and M_{tot} values with 10^{-2} relative error from frequency and strain data, implementing 1.5PN formalism.**
- **Near future steps**:
  - to build a real dataset spanning a physical parameter space;
  - to test robustness against noise and glitches;
  - to extend the number of parameters to infer.
- (Not so) **remote future step**:
  - use of cWB real outputs,
  - apply PINNs approach to TOV equations, in order to constrain NS's equation of state.

Lattice Quantum Field Theory

as a sampling problem



(Euclidean) action

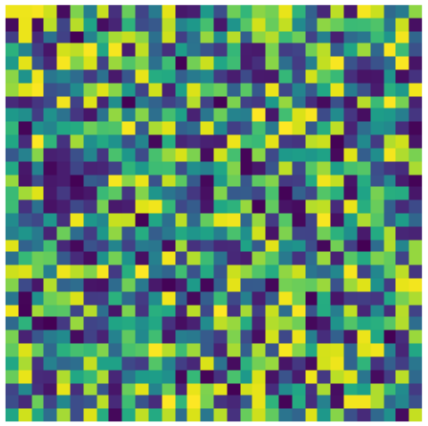
$$S[\phi] = \int d^D x \, \frac{1}{2} \left((\partial_\mu \phi)^2 + m^2 \phi^2 \right) + \lambda \phi^4$$

$$p(\phi) \propto e^{-S[\phi]}$$

Samples: $\phi \sim e^{-S[\phi]}$ —**estimate observables**→ $\langle \mathcal{O}[\phi] \rangle$

Generative models

trivial theory



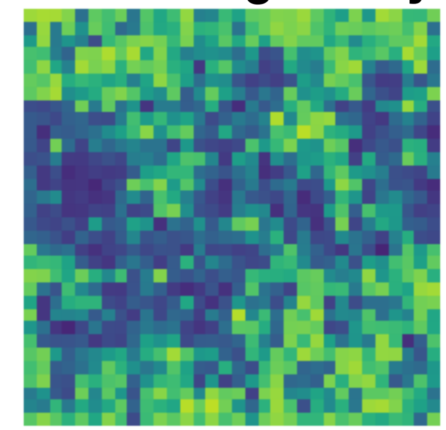
$\mathcal{N}$

bijection f

"Normalizing flow"

Computable likelihood

interacting theory



$e^{-S[\phi]}$

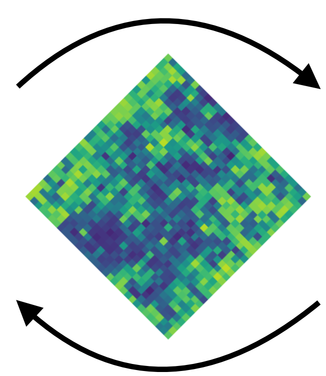
---

continuous flow

$$\dot{f} = g_\theta[\phi]$$

$$\phi_0 \longrightarrow \phi_1$$

built-in symmetry



transfer learning

$\lambda = 0.7$ ⬇ $\lambda = 0.8$

L=6 ⬇ L=12

Importance nested sampling with normalizing flows

(for gravitational-wave inference)

Michael J. Williams, John Veitch, Chris Messenger

arXiv:2302.08526

UNIVERSITY OF PORTSMOUTH
University of Glasgow

Can we accelerate nested sampling with machine learning?

nessai

Nested sampling + normalizing flows

➕ Improved sampling efficiency
➖ Limited by nested sampling design

Poster 38

Importance nested sampling with normalizing flows

(for gravitational-wave inference)

Michael J. Williams, John Veitch, Chris Messenger

arXiv:2302.08526

UNIVERSITY OF PORTSMOUTH

University of Glasgow

What if we design a nested sampling algorithm around normalizing flows?

i-nessai

Importance nested sampling + normalizing flows

➕ Addresses the main bottlenecks

➕ Further improvements to sampling efficiency

Poster 38