A data-driven and model-agnostic approach to solving combinatorial assignment problems in searches for new physics — Anthony Badea, Javier Montejo Berlingen

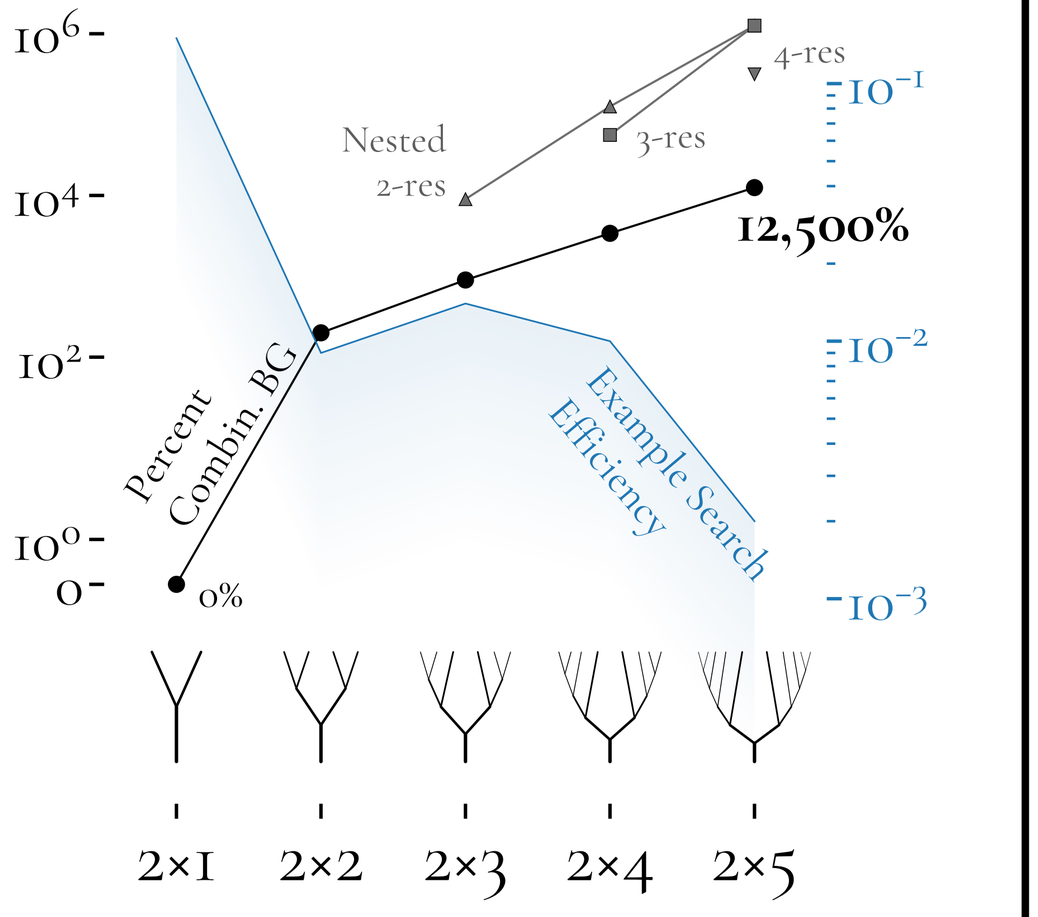
Motivation

Interesting physics processes often lead to many object final states

Full event reconstruction is a challenging combinatorial problem

ML solutions so far are trained on specific signal models, will fail if data contains a different signal

Our work removes the signal models and only assume symmetric pair production



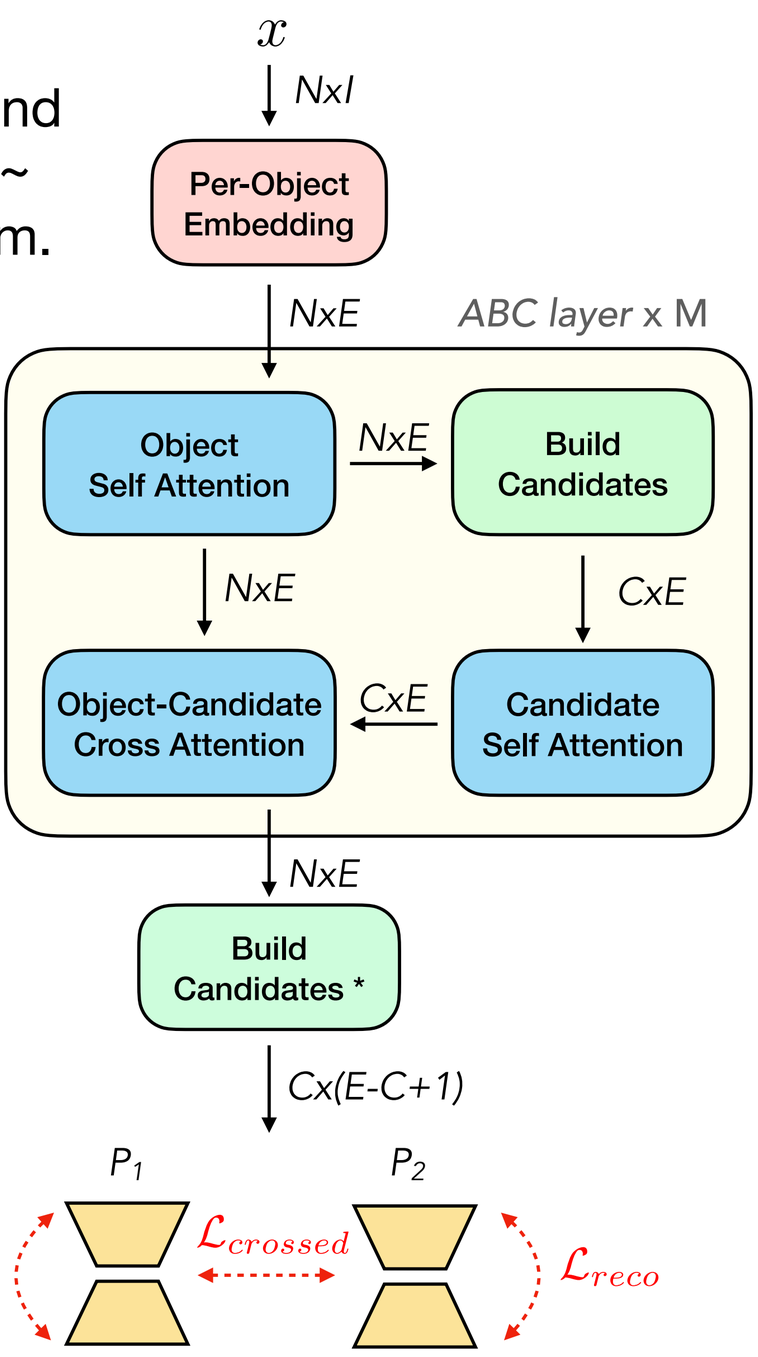
Phys. Rev. D 106, L016001

Setup

Input: background enriched data ~ QCD MC 4-mom.

Form intermediate candidates

Output: object pairings

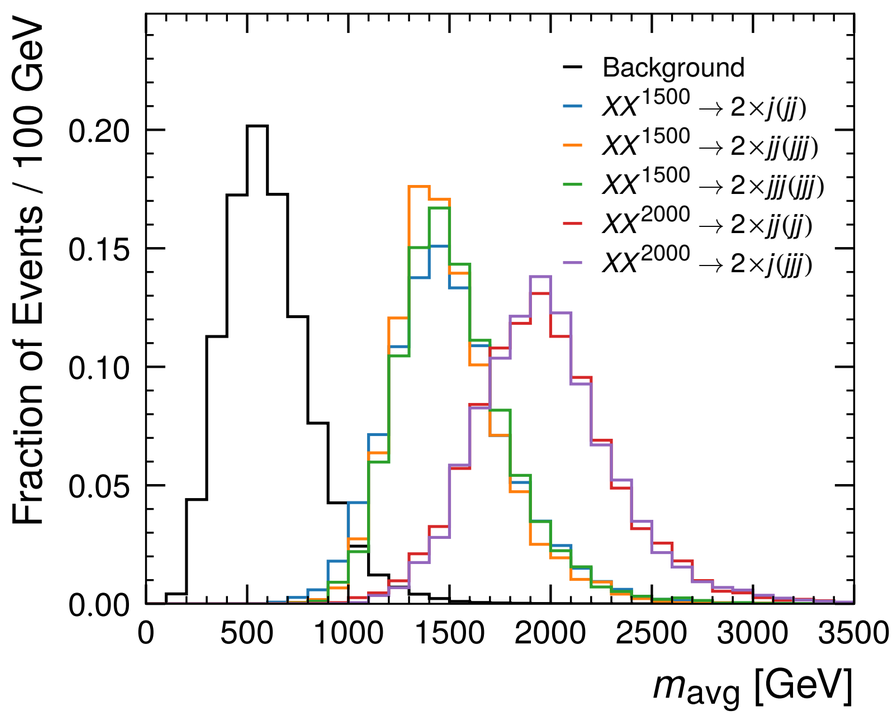


Loss

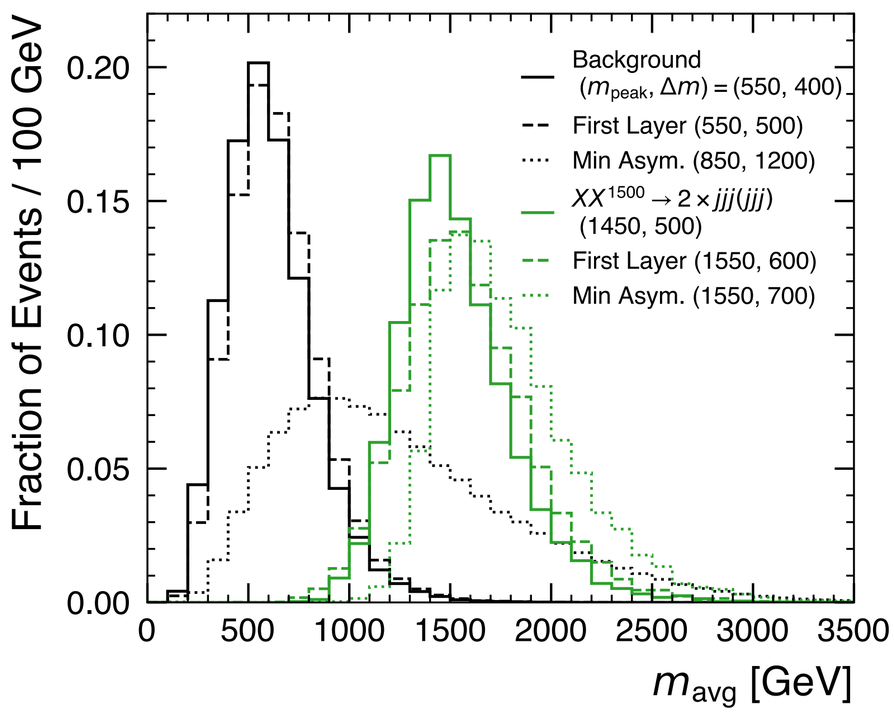
$$\mathrm{L} = \mathrm{L}_{reco} + \mathrm{L}_{crossed} + \mathrm{L}_{random}\ (+\mathrm{L}_{ISR})$$
$$\mathrm{L}_{reco} = \lambda_{reco} \cdot \left(\|p_1 - \hat{p}_1\| + \|p_2 - \hat{p}_2\| \right)$$
$$\mathrm{L}_{crossed} = \lambda_{crossed} \cdot \|z_1 - z_2\|$$
$$\mathrm{L}_{rand} = \lambda_{rand} \cdot \max\left(0, 1 + \|z_1 - z_2\| - \|z_1^{rand} - z_2^{rand}\|\right)$$
$$\mathrm{L}_{ISR} = \lambda_{ISR} \cdot E_T(P_{ISR})/\mathrm{GeV}$$

Results

Clear mass peaks for different multiplicity final states, despite model being trained purely on background



Strong performance in comparison to classical methods



Phys. Rev. D 109, L011702

AB (University of Chicago, Schmidt Futures)
JMB (Instituto de Física de Altas Energías)