[2307.02405]

ν^2-Flows

Fast and improved neutrino reconstruction in multi-neutrino final states

ML4Jets, 6th November 2023

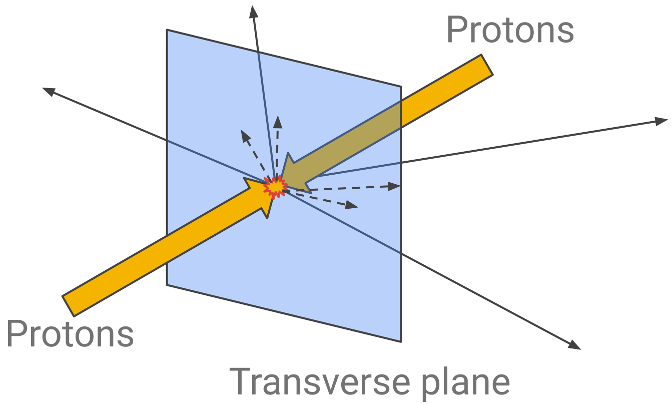
Johnny Raine, Matthew Leigh, Knut Zoch, Tobias Golling
University of Geneva

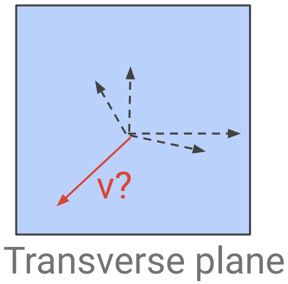
ν-Flows - Neutrino Regression

Goal: Predicting neutrino momentum

Neutrinos don't interact at all with detector

Can only infer their presence from **conservation of momentum**



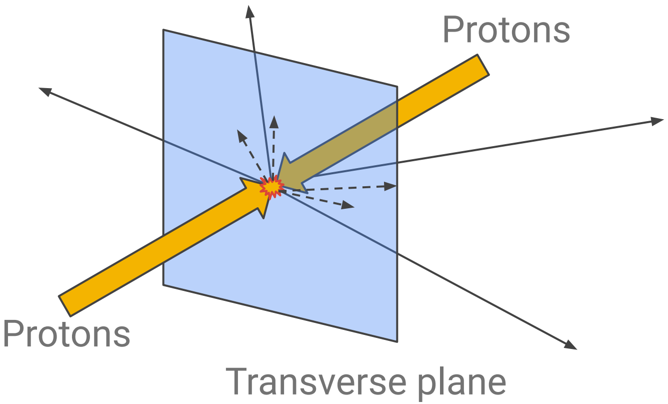


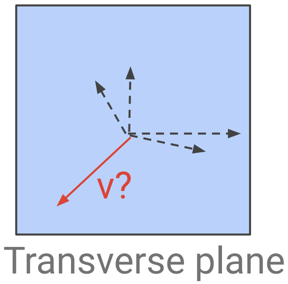
ν-Flows - Neutrino Regression

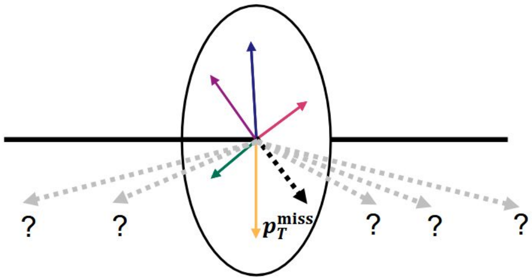
Goal: Predicting neutrino momentum

Momentum only conserved in transverse plane

Only have a **2D projection** of neutrino momentum





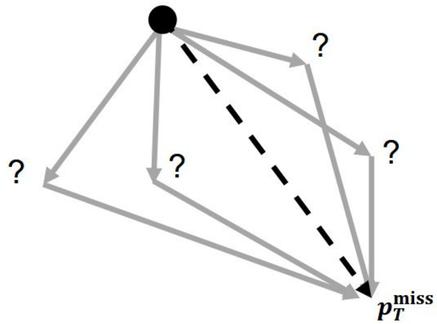


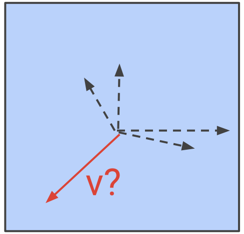
ν-Flows - Neutrino Regression

Goal: Predicting neutrino momentum

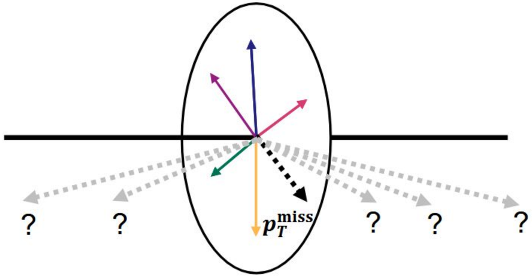
Momentum conservation from sum of vectors

Under constrained in transverse plane when multiple neutrinos

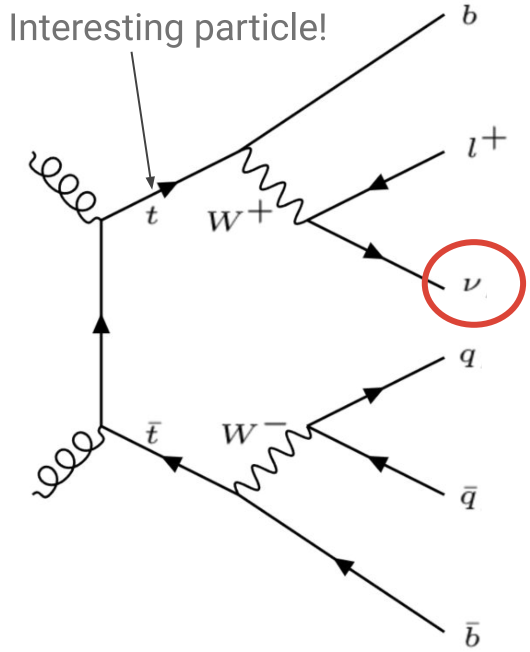




Transverse plane



ν-Flows - Neutrino Regression

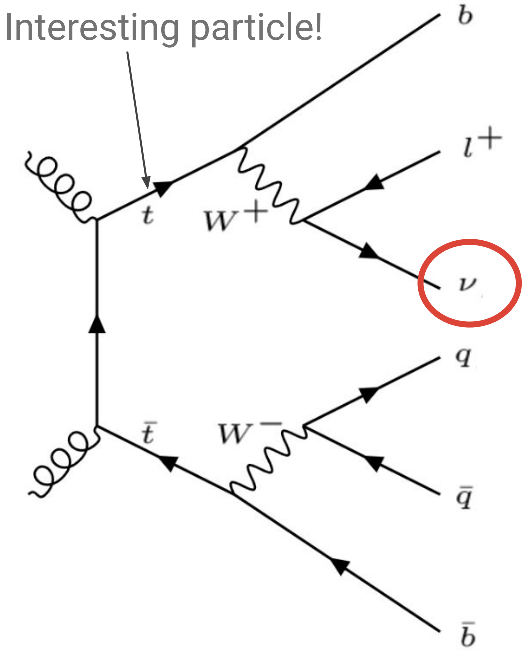


Why do we care about neutrinos?

Without neutrino information no longer able to **fully reconstruct top quark**

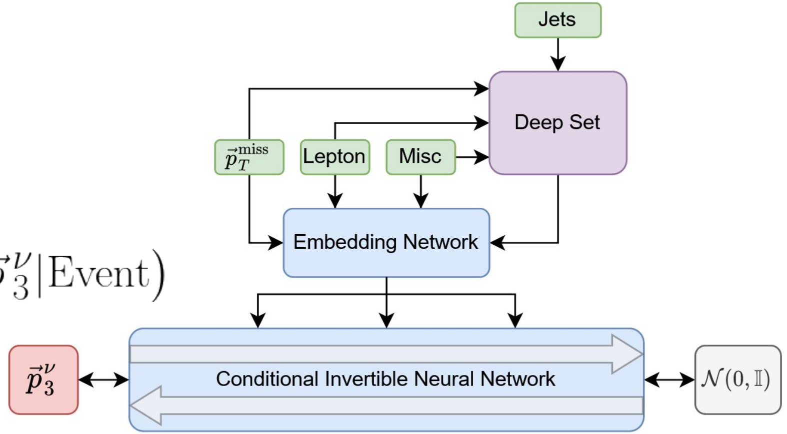
Neutrino information helps solve combinatorics: Assignment of jets to partons easier if you know the W system

ν-Flows - Neutrino Regression



Use normalizing flows to learn the conditional probability of solutions given observation from training dataset

$$p\left(\vec{p}_3^{\,\nu}|\text{Event}\right)$$



At inference obtain distribution over all possible neutrino momenta for the observed event, sample from posterior

ν-Flows - Neutrino regression

SciPost | SciPost Phys. 14, 159 (2023)

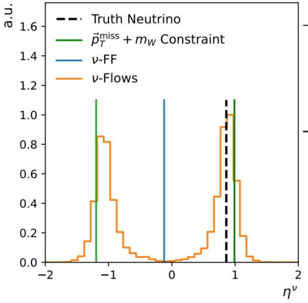
ν-flows: Conditional neutrino regression

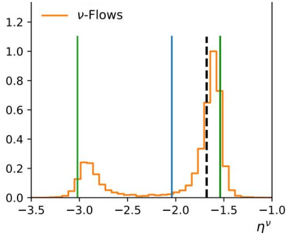
Matthew Leigh*, John Andrew Raine†, Knut Zoch and Tobias Golling

University of Geneva, Geneva, Switzerland

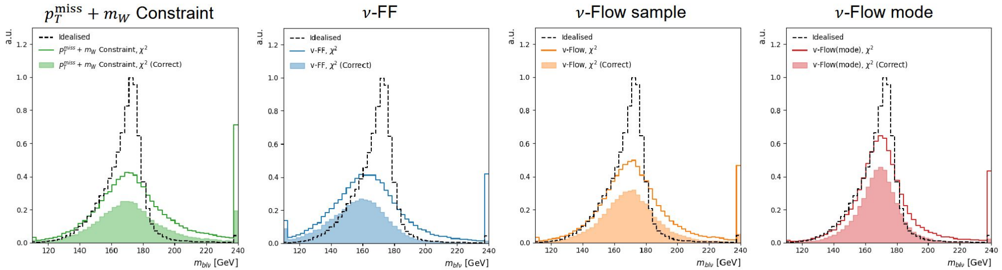
* matthew.leigh@unige.ch , † john.raine@unige.ch

Neutrino Type	Number of Jets 4	5	6	7	8	9
Truth Neutrino	0.864	0.753	0.686	0.641	0.611	0.587
$\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ and m_W Constraint	0.790	0.576	0.476	0.398	0.366	0.286
ν-FF	0.754	0.533	0.410	0.353	0.300	0.302
ν-Flows(sample)	0.803	0.624	0.515	0.457	0.391	0.357
ν-Flows(mode)	**0.813**	**0.664**	**0.575**	**0.508**	**0.481**	**0.405**





Demonstrated substantial gains for top quark pairs with one neutrino

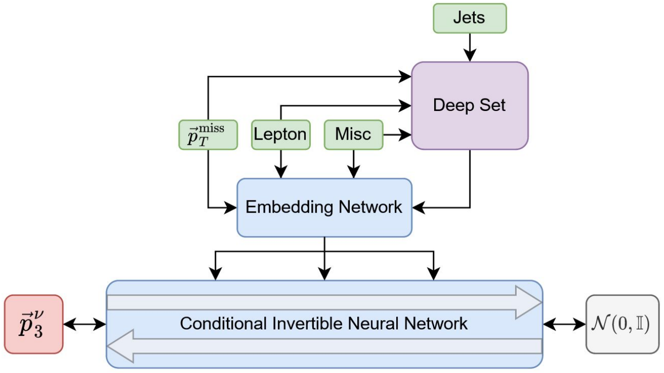


But how does it scale to multiple neutrinos?

ν^2-Flows - Moving to multiple neutrinos

As before don't want to enforce link between objects - network learns combinatorics

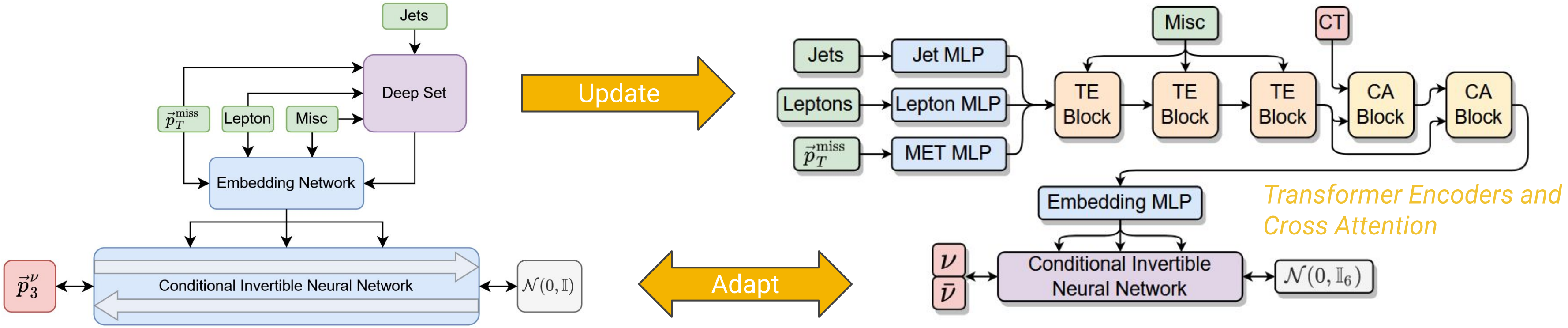
Multiplicity and permutation invariant in jets



ν^2-Flows - Moving to multiple neutrinos

As before don’t want to enforce link between objects - network learns combinatorics

Multiplicity and permutation invariant in jets and leptons
Can easily add any other objects e.g. photons and taus



Normalizing flow remains constant
Scale dimension with number of neutrinos

Moving to multiple neutrinos

For benchmarking consider top quark pair events with dilepton final state

Observe two leptons, at least two (b-)jets, missing momentum

Reconstructing top quark pair crucial for many measurements

- Top quark spin correlation - quantum entanglement sensitive at low m_{tt}
- Top quark mass (often template fit measuring b-lepton mass)
- Top quark polarisation, charge asymmetry, differential cross sections

Moving to multiple neutrinos

Compare to two standard approaches (relying on hard assumptions on mass)

Neutrino Weighting

Ellipse method

$$(\ell_{1,2} + \nu_{1,2})^2 = m_w^2 = (80.38\,\mathrm{GeV})^2,$$

$$(\ell_{1,2} + \nu_{1,2} + b_{1,2})^2 = m_t^2 = (172.5\,\mathrm{GeV})^2,$$

Scan eta values for both neutrinos
Choose solution which maximises a weight

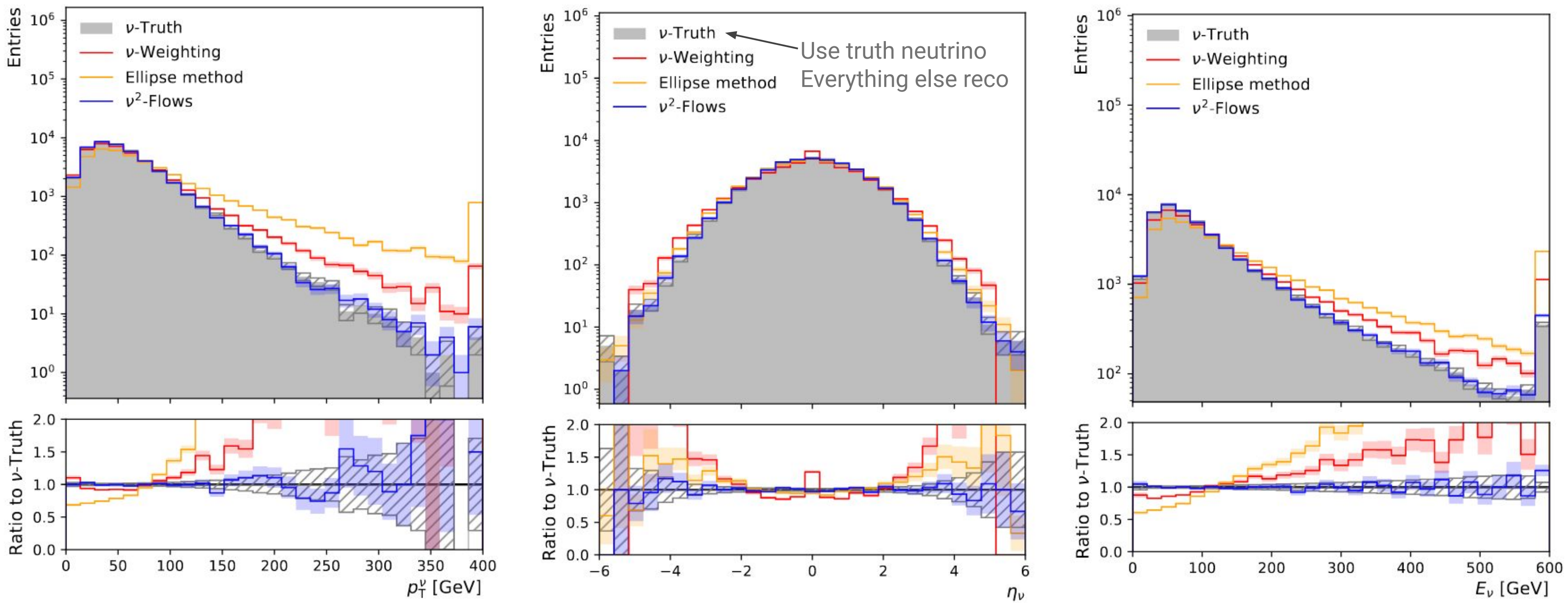
$$w = \exp\left(-\frac{||\vec{p}_{\mathrm{T}}^{\mathrm{miss}} - \vec{p}_{\mathrm{T}}^{\nu\bar{\nu}}||_2^2}{2\sigma^2}\right)$$

Additionally scan top quark mass values to improve acceptance

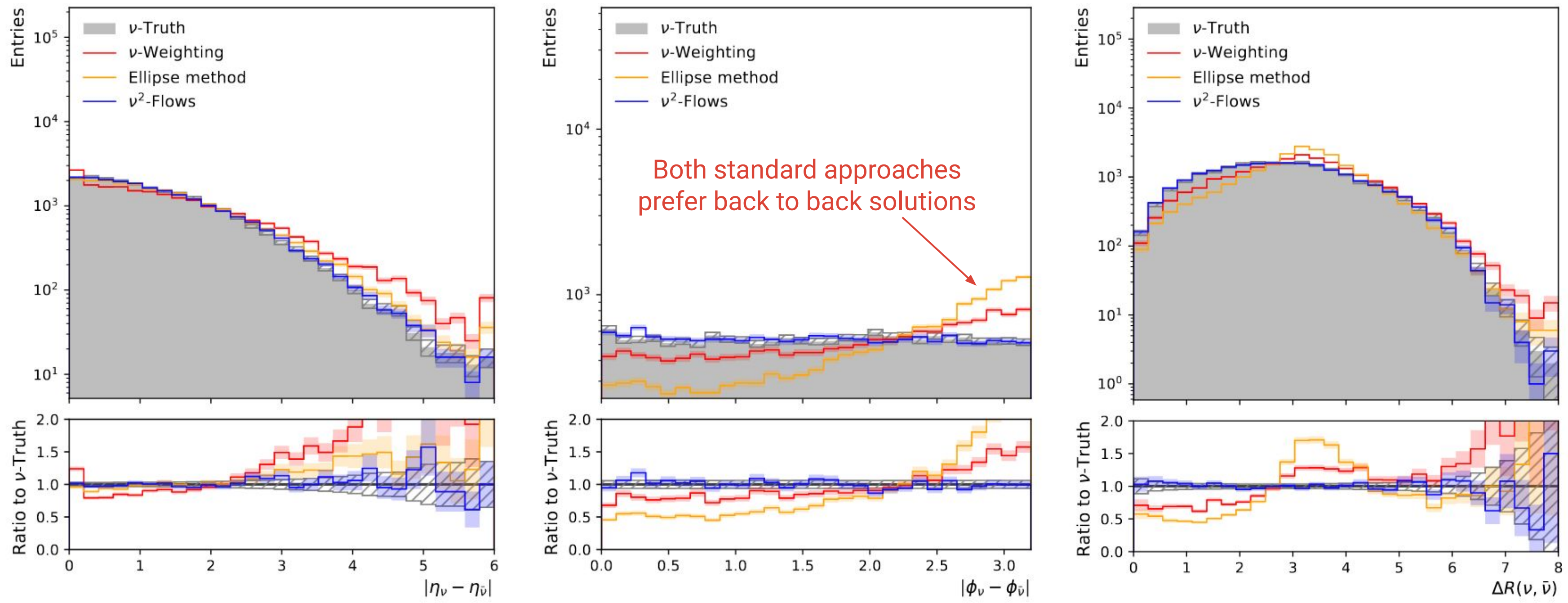
Use observed missing momentum to constrain solution further

Less flexible to resolution effects but computationally more efficient

Comparing kinematics

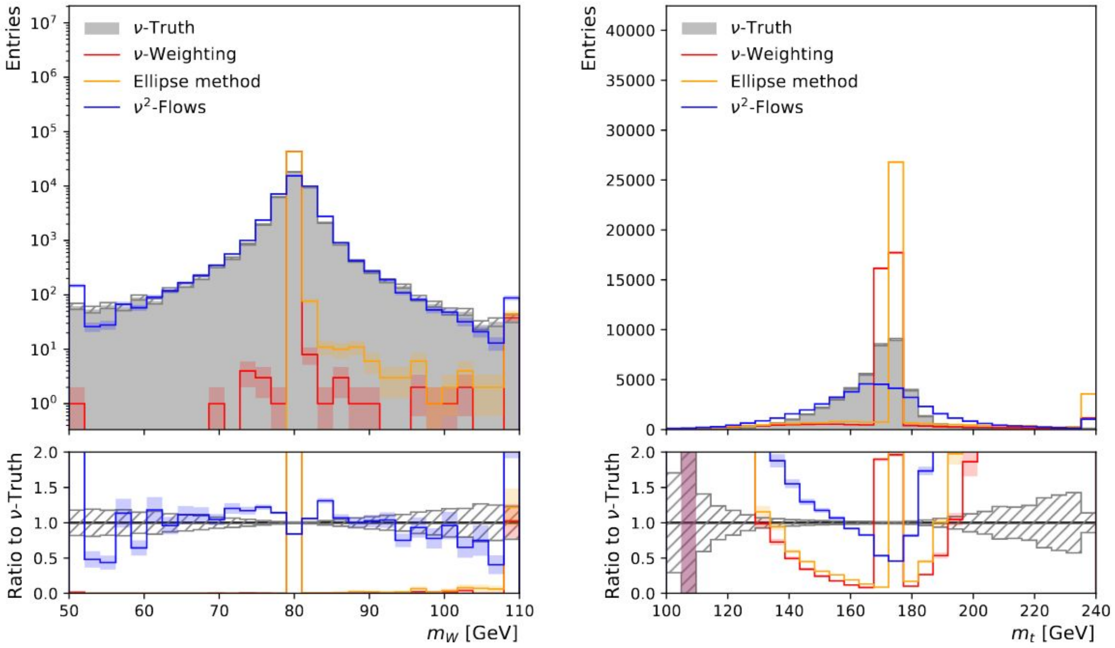


Comparing correlations



Compare impact of mass constraints

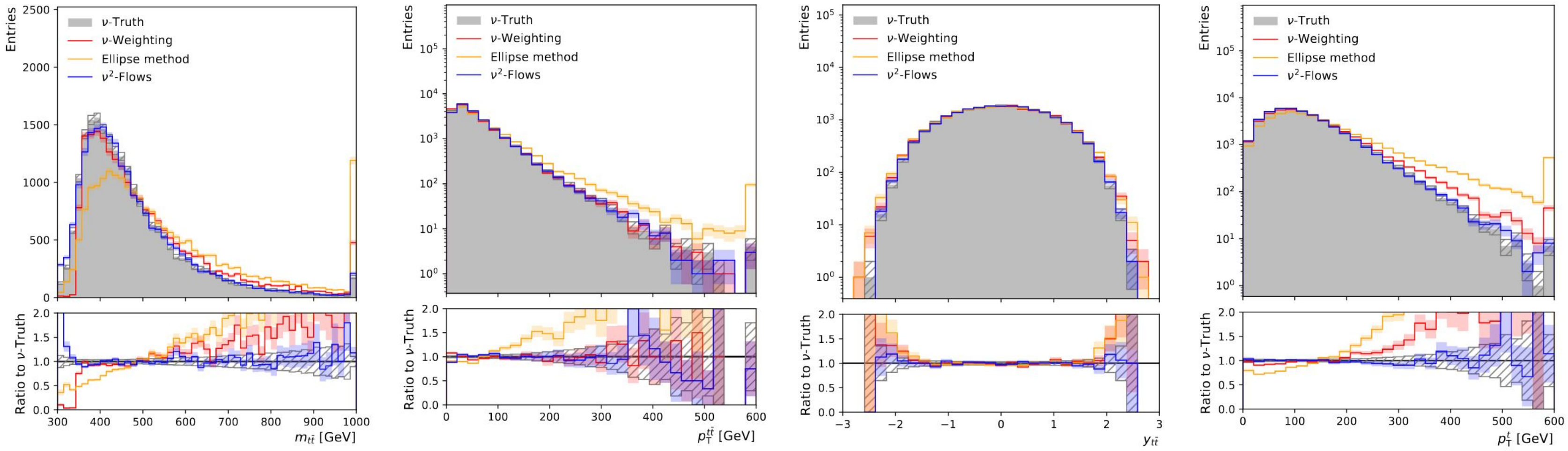
Don't see a perfect peak due to wrong top association for standard methods



Comparing top quark kinematics

Capturing complex distributions much more closely with ν^2-Flows

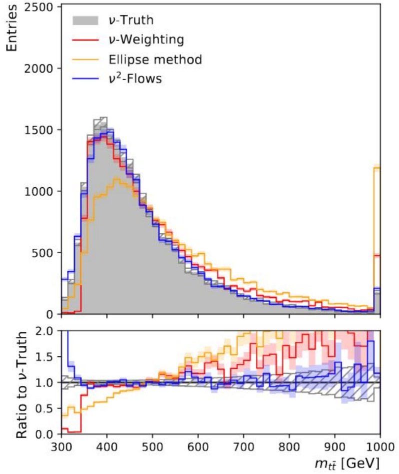
In particular m_{tt} at threshold and high values is much closer to optimum

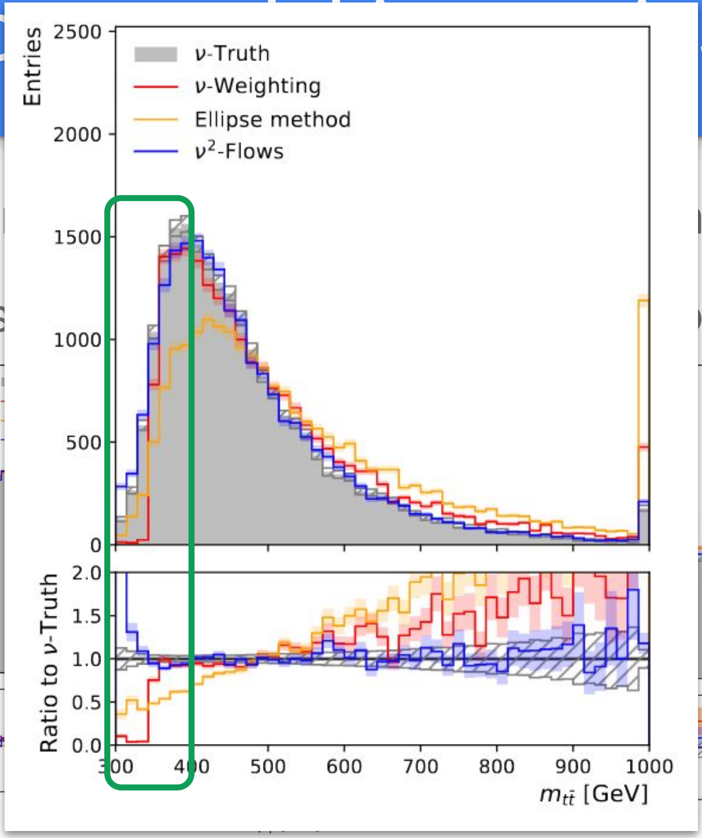


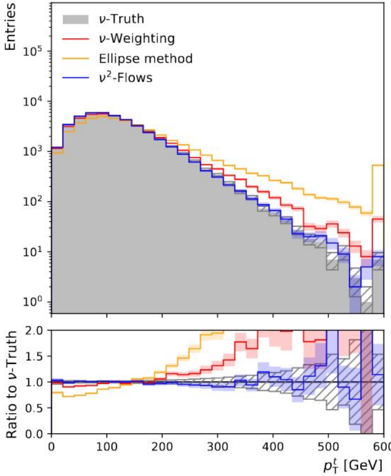
Comparing to ... s

Capturing complex dist... ν^2-Flows

In particular m_{tt} at thres... oser to optimum



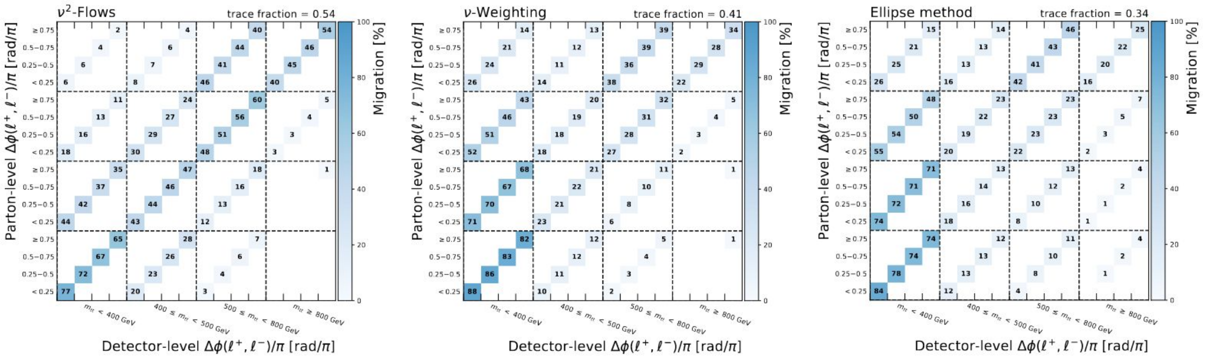




Quantifying improvement

Substitute in each method into a double-differential measurement

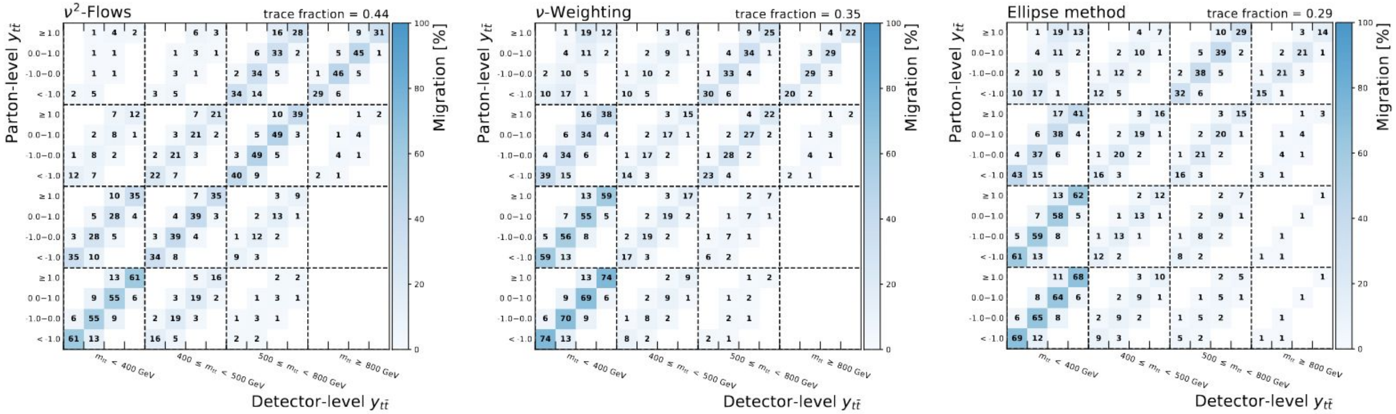
- Look at response matrix - is it diagonal?
- Look at statistical precision in each bin after unfolding
- Look at efficiency of method in each truth level bin



Observable	Bin edges
$m_{t\bar{t}}$	[0, 400, 500, 800, inf] GeV
$\Delta\phi(\ell^+\ell^-)$	[0.0, 0.25, 0.5, 0.75, 1.0] rad/π
p_T^t	[0, 75, 125, 175, inf] GeV
$p_T^{t\bar{t}}$	[0, 70, 140, 200, inf] GeV
$y_{t\bar{t}}$	[-inf, -1.0, 0.0, 1.0, inf]

ν^2-Flows - Extending to multiple neutrinos

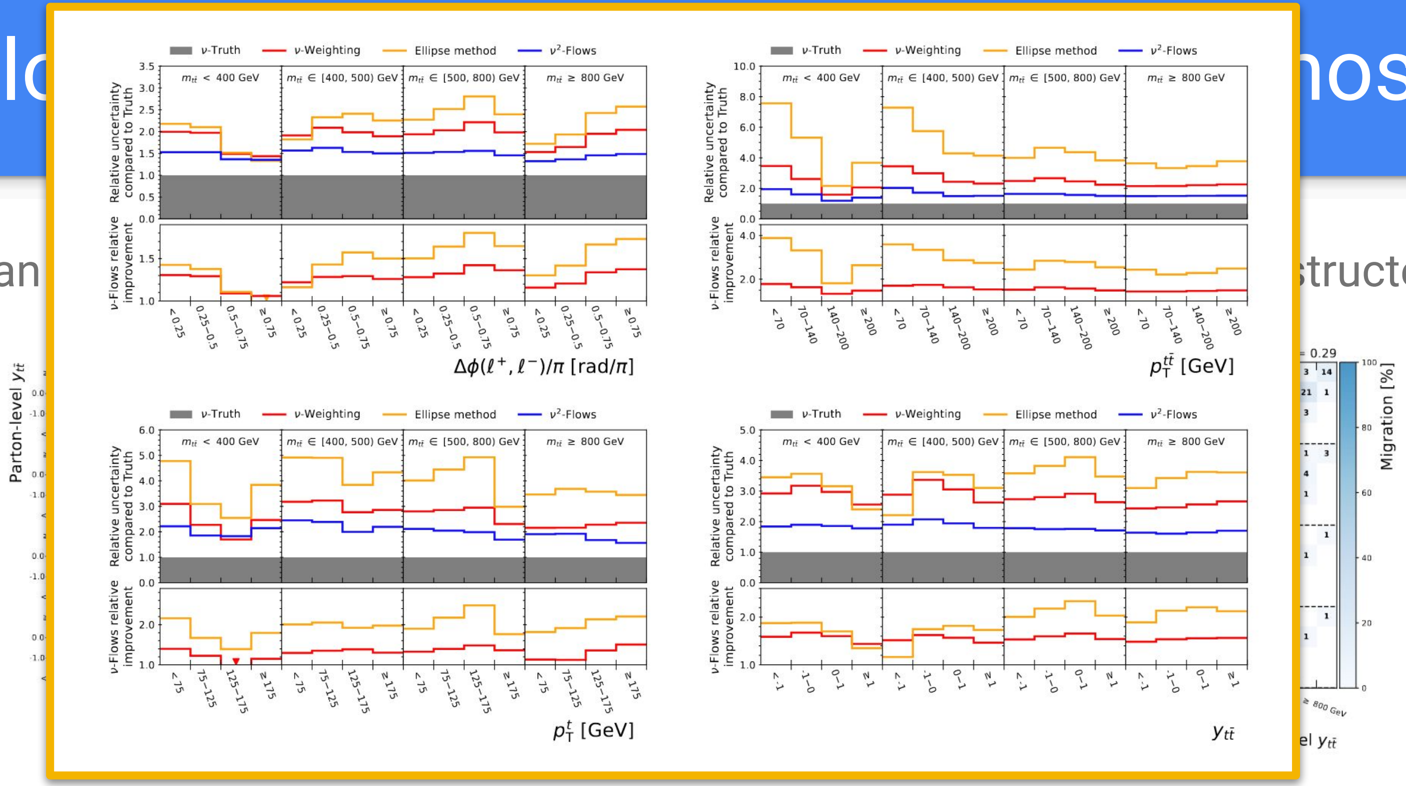
Use in an unfolding measurement - compare migration of truth to reconstructed



Much improved diagonal response! Relationship between truth and reco is well captured

ν^2-Flo ... nos

Use in an ... structed



Improvements in all double-differential distributions with ν^2-Flows -> direct statistical benefit in analysis!

Using SVD binned unfolding

ν^2-Flows always 100%

Observables		Method	Efficiency at finding a solution (per bin) [%]															
$m_{t\bar{t}}$	$y_{t\bar{t}}$	ν-Weighting	97	98	97	97	96	96	97	96	93	94	94	94	89	90	89	89
		Ellipse	91	92	92	91	84	82	84	82	67	68	68	70	58	54	51	50
	$p_T^{t\bar{t}}$	ν-Weighting	99	95	91	89	98	95	90	84	97	92	89	82	94	88	81	81
		Ellipse	96	86	74	59	88	79	66	50	75	66	52	38	60	52	34	37
	p_T^t	ν-Weighting	98	96	92	86	97	97	96	91	95	95	95	93	96	89	90	89
		Ellipse	95	88	72	51	89	86	80	63	77	74	72	63	61	56	55	51
	$\Delta\phi(\ell^+\ell^-)$	ν-Weighting	96	97	98	99	96	96	97	97	93	94	94	94	90	90	91	88
		Ellipse	90	92	92	94	84	84	83	82	76	74	68	64	57	57	57	48

In many bins even more **statistical gains** due to ν^2-Flows always having a solution!

And what about the top quark mass?

Can use neutrino solutions in top quark mass reconstruction

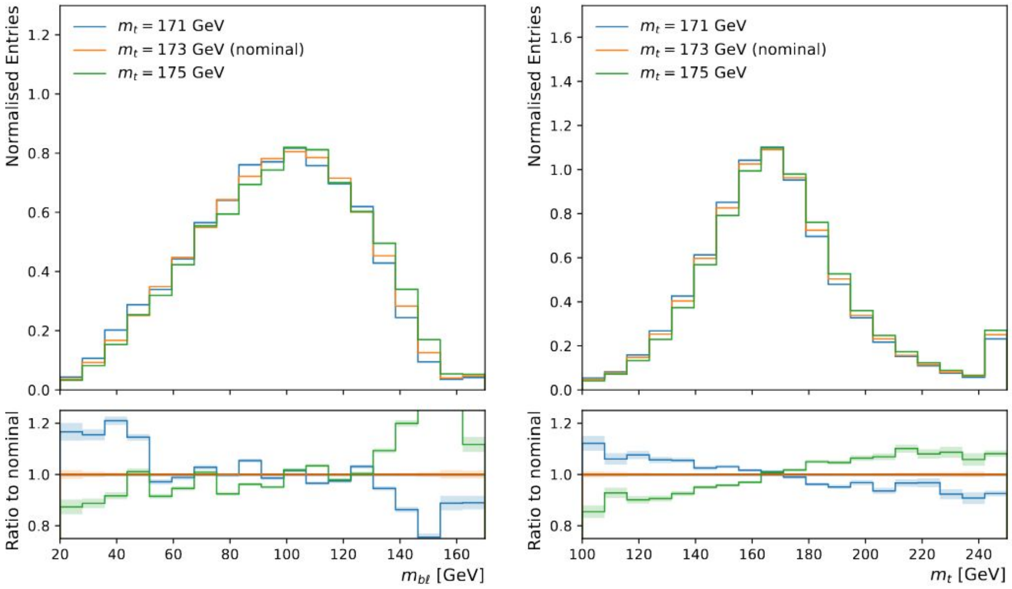
When using ν^2-Flows retain sensitivity to top mass!

Train on one value of mass

Evaluate on samples with different truth mass values

No neutrino

With ν-Flows



Conclusions

ν^2-Flows…

outperforms standard approaches still used in current ttbar dilepton analyses

can accommodate any neutrino and event object multiplicities

- Takes advantage of modern transformer architectures with conditional normalizing flows
- Shown on ttbar but framework designed to be adaptable to any final state

has fast inference and is competitive with fast analytical methods

can even be used in a mass analysis

Backup

Links

ν-Flows paper: https://arxiv.org/abs/2207.00664

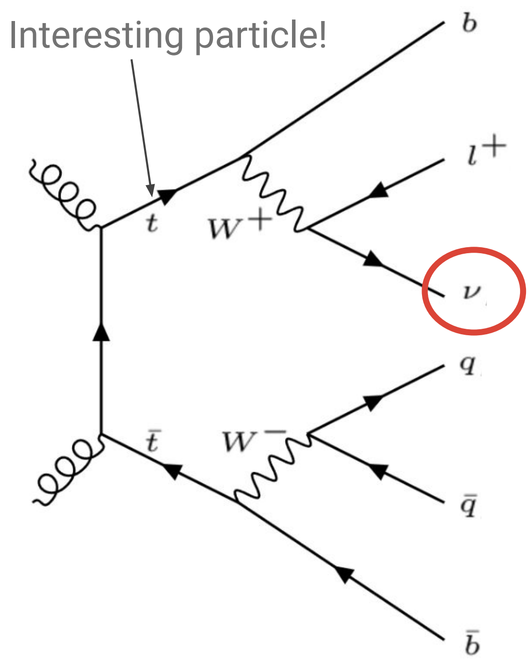
ν^2-Flows implementation: https://github.com/rodem-hep/nu2flows

Dataset: https://zenodo.org/record/8113516

Neutrino Weighting implementation from: https://arxiv.org/abs/1903.07570

Ellipse implementation from: https://github.com/betchart/analytic-nu

ν-Flows - Neutrino Regression



Standard approach in HEP for 1ℓ

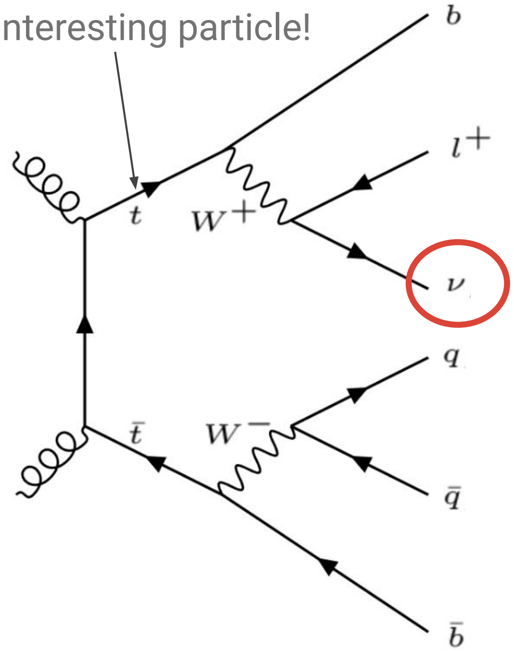
Assume mass of W exactly 80.46 GeV

Given $p_4^W = p_4^\nu + p_4^\ell$ $\quad$ $p_4 = (p_x, p_y, p_z, m)$

- Use missing transverse momentum for p_T^ν
- Solve quadratic equation for p_z^ν

ν-Flows - Neutrino Regression

Interesting particle!



Standard approach in HEP for 1ℓ

Ignore width of resonance
Ignores detector resolution

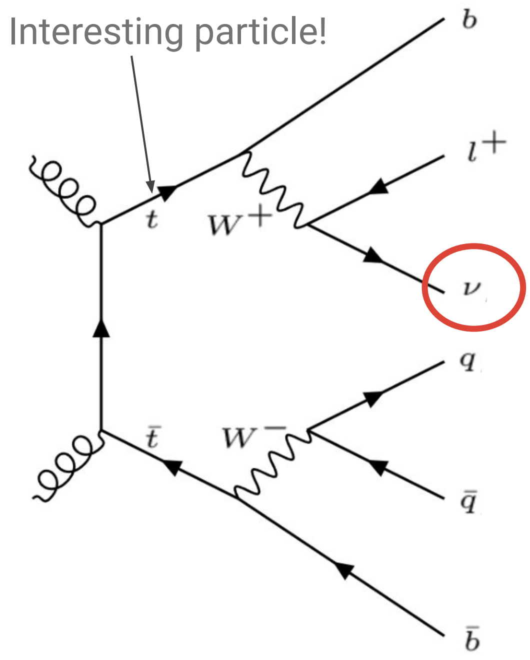
Assume mass of W exactly 80.46 GeV

Given $p_4^W = p_4^\nu + p_4^\ell$ $\quad p_4 = (p_x, p_y, p_z, m)$

- Use missing transverse momentum for p_T^ν
- Solve quadratic equation for p_z^ν

Can give two solutions with no preference
Can yield no real solutions

ν-Flows - Neutrino Regression



Standard approach in HEP for 1ℓ

Assume mass of W exactly 80.46 GeV

Given $p_4^W = p_4^\nu + p_4^\ell$ $\quad$ $p_4 = (p_x, p_y, p_z, m)$

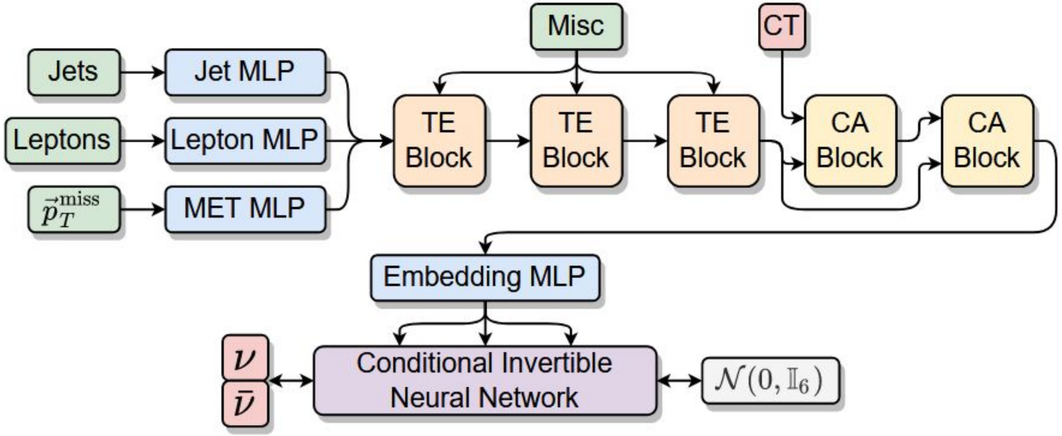
- Use missing transverse momentum for p_T^ν
- Solve quadratic equation for p_z^ν

For 2ℓ also assume top quark mass

- Improve efficiency by smearing m_{top} and scanning η^ν
- Use $p_T^{\nu\nu}$ / p_T^{miss} to select best solution

ν^2-Flows - Extending to multiple neutrinos

Move to 2 neutrinos - permutation invariant architecture for all objects



Transformer encoder

Truth level information, can enforce an order as long as consistent in training

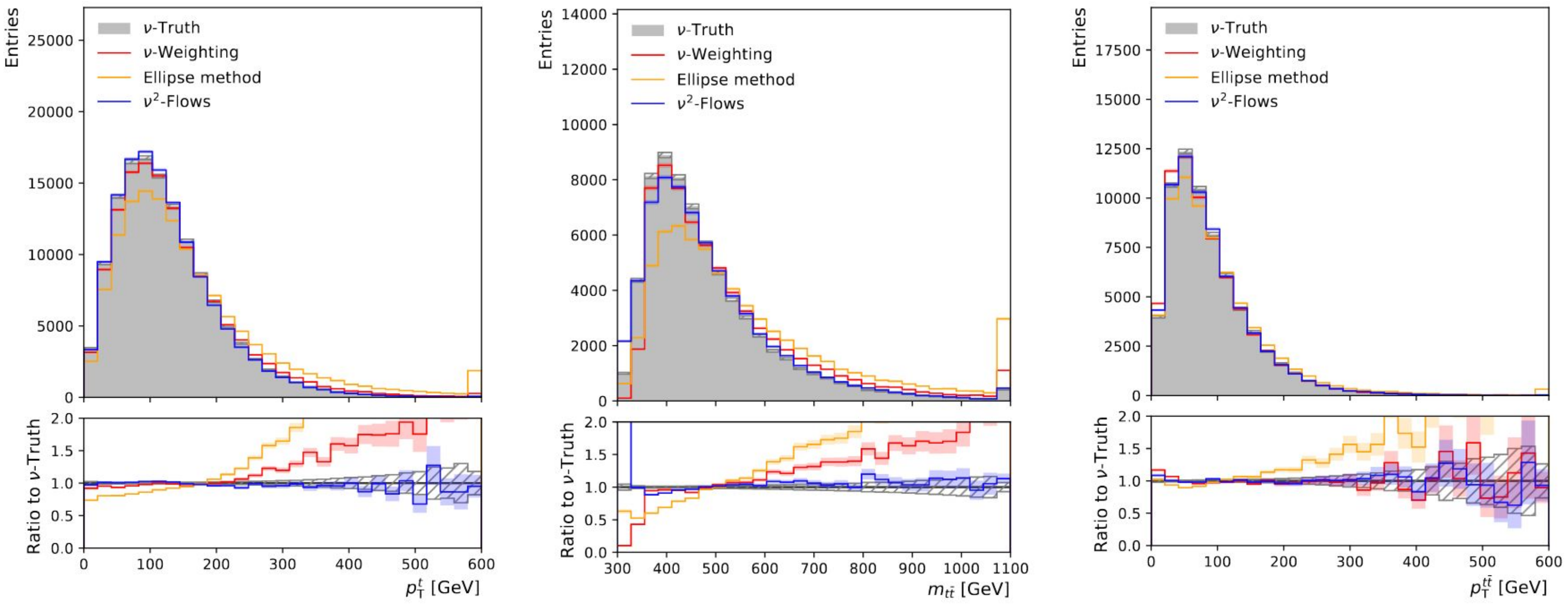
Inputs

Category	Variables	Description
$\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$	p_x^{miss}, p_y^{miss}	Missing transverse momentum 2-vector
Leptons	p_x^ℓ, p_y^ℓ, p_z^ℓ, $\log E^\ell$	Lepton momentum 4-vector
	q^ℓ	Lepton charge
	ℓ^{flav}	Whether lepton is an electron or muon
Jets	p_x^j, p_y^j, p_z^j,$\log E^j$	Jet momentum 4-vector
	isB	Whether jet passes b-tagging criteria
Misc	N_{jets}, N_{bjets}	Jet and b-jet multiplicities in the event

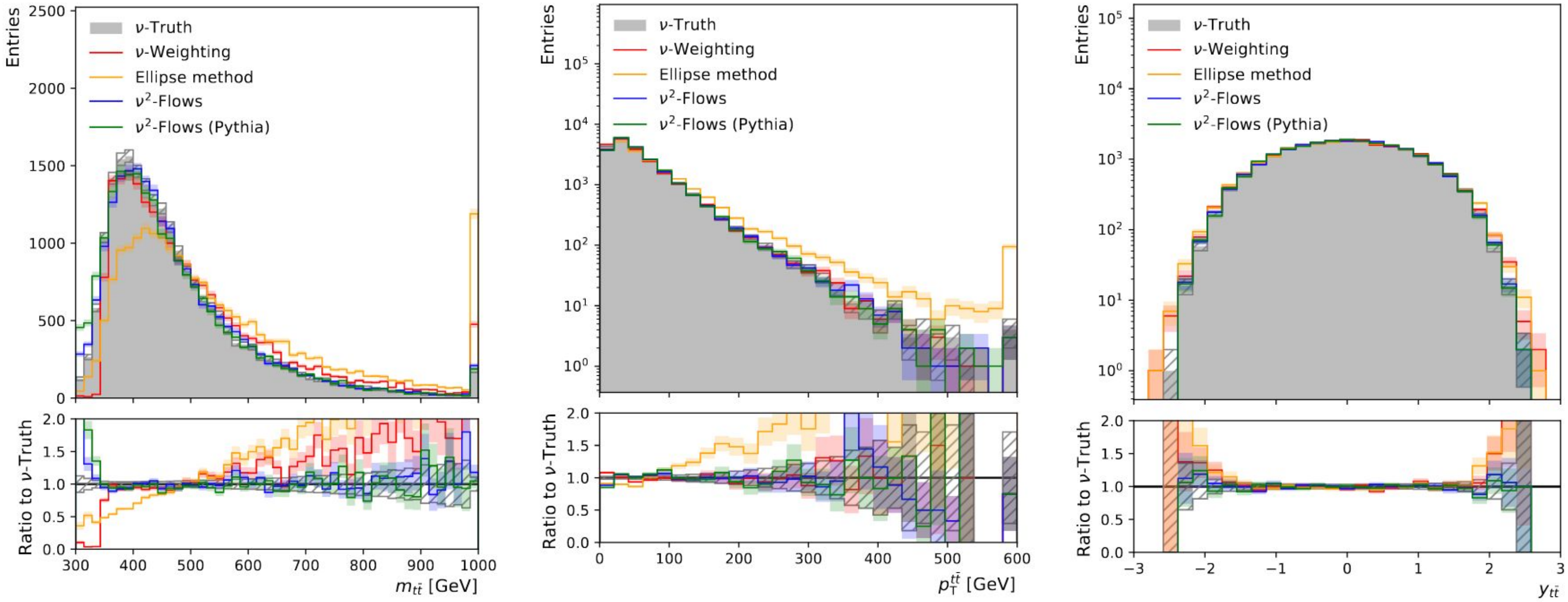
Observables

Observables	
$m_{t\bar{t}}$	Invariant mass of $t\bar{t}$ system
$\Delta\phi(\ell^+\ell^-)$	Separation in ϕ between the two leptons
p_T^t	Transverse momentum of the top quark
$p_T^{t\bar{t}}$	Transverse momentum of the $t\bar{t}$ system
$y_{t\bar{t}}$	Rapidity of the $t\bar{t}$ system

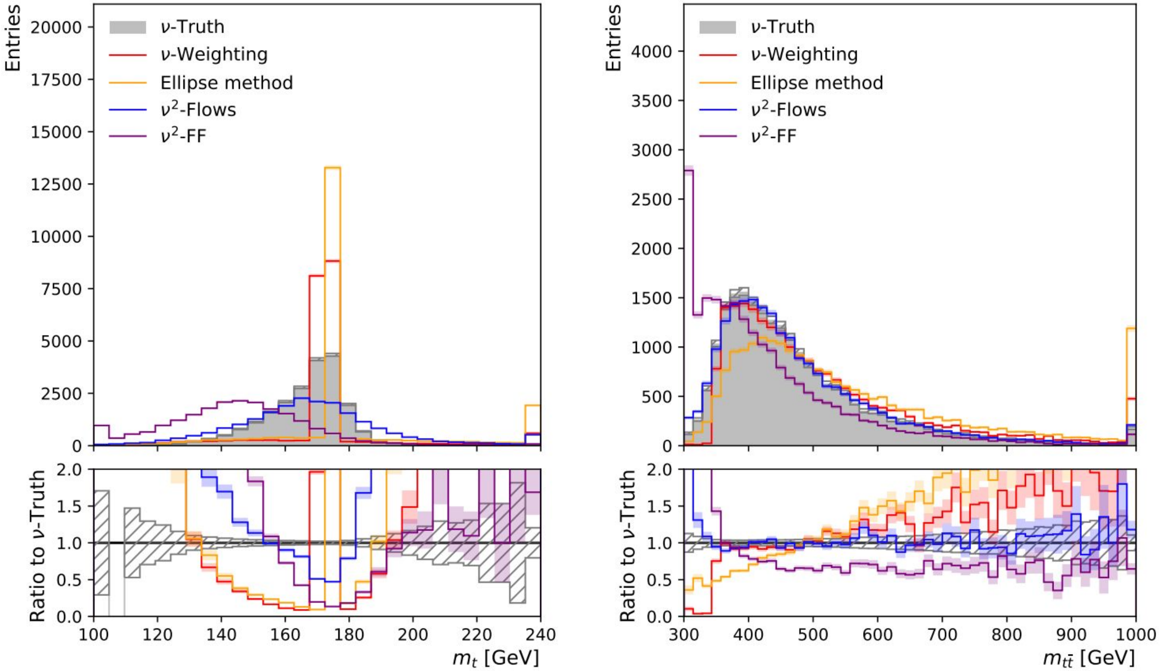
Evaluating on extra ISR/FSR sample



Training on Pythia LO instead of MG5



Comparison to a feed-forward network



Origin of improvement in distributions

Better distributions everywhere

Not just improving due to low scoring NW events

Where NW has large w weight

Where NW has small w weight

Performance seems independent of NW weight