Returning CP-observables to the frames they belong

Rahool Barman
Oklahoma State University

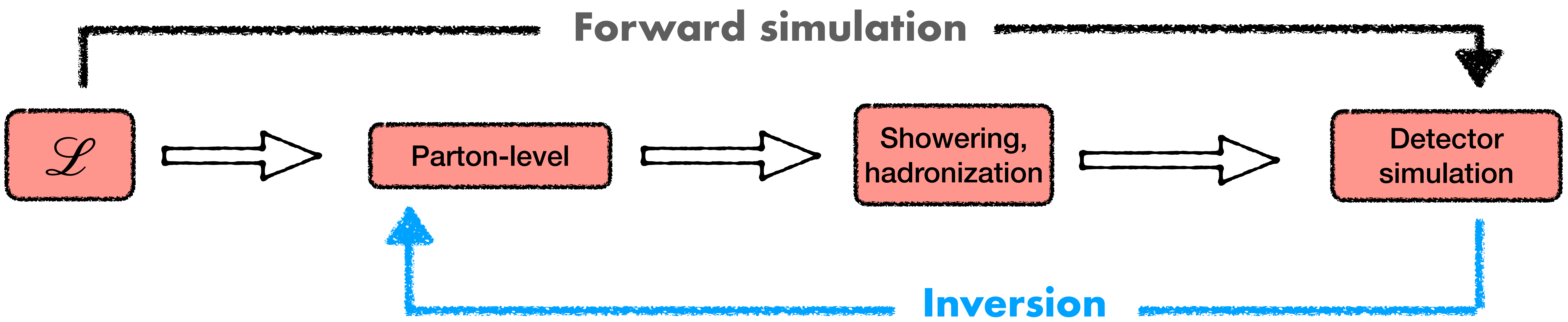
With
Jona Ackerschott, Dorival Gonçalves, Theo Heimel, Tilman Plehn

Based on
arXiv: 2308.00027

Joint meeting on CPV in extended Higgs sector
CERN

September 26, 2023

Unfolding

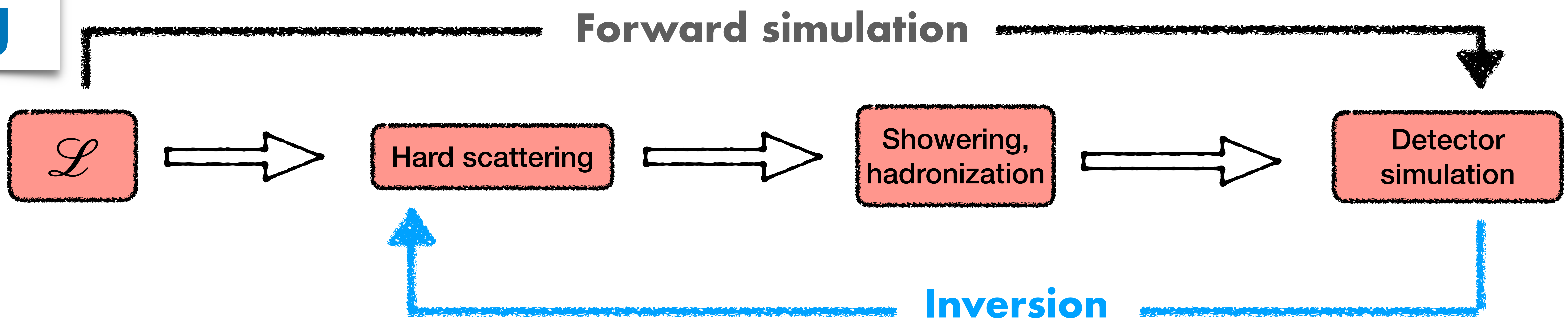


- Conventional LHC analysis involves comparing measured data with MC events simulated under NP hypothesis.
  - Reconstructed LHC events present a convoluted version of the true underlying physics.
  - Forward simulation chain can be highly resource intensive.

Invert simulation chain → apply on measured data → reconstruct parton-level

→ compare new physics hypotheses at the parton-level.

Unfolding



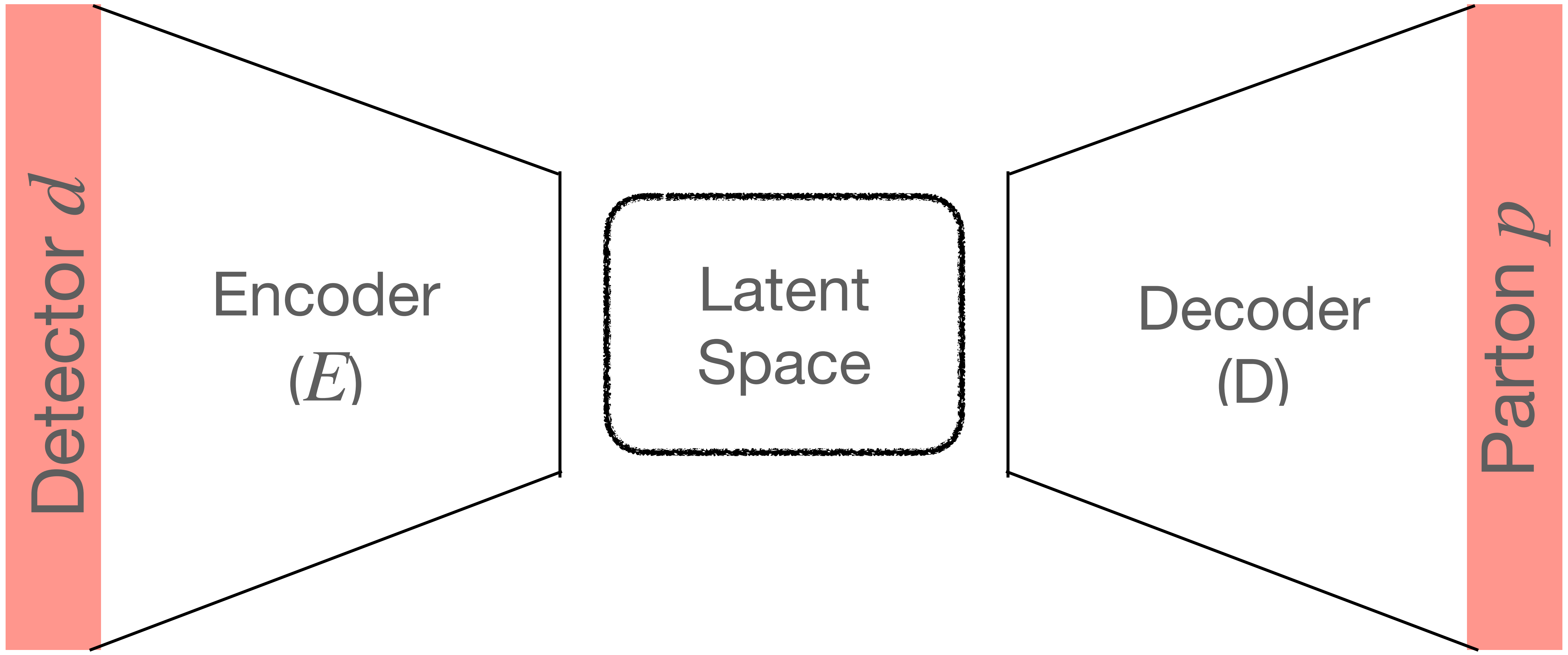
✦ Bin-independent ✦ Able to invert multi-dimensional d.o.f.

Possible with machine learning based generative models.

- Variational Auto Encoders (VAE)
- Normalizing Flows (NF)
- Generative Adversarial Networks (GAN)

[Bellagente, Butter, Kasieczka, Plehn, Winterhalder (2020)]
[Bellagente, Butter, Kasieczka, Plehn, Rousselot, Winterhalder, Ardizzone, Kothe (2020)]
[Andreassen, Komiske, Metodiev, Nachman, Thaler (2020)]
[Komiske, McCormack, Nachman (2021)]

Variational Auto Encoder (VAE)



- The Encoder maps the input detector data d to a more tractable latent space $z = E(d)$ while preserving the essential features.

- The decoder maps z to the parton level $p' = D(z) = E(l(d))$.

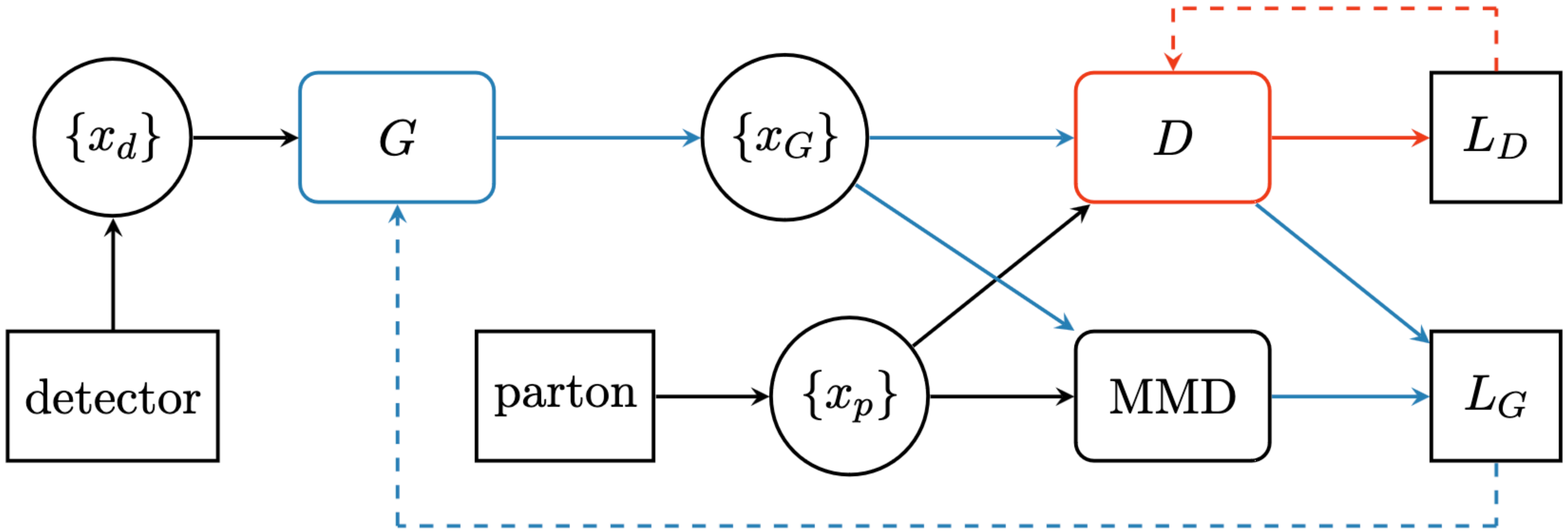
- Training goals involve minimizing the reconstruction error (how well the Decoder can generate parton data)

[Otten, Caron, Swart, Beekveld, Hendriks, Leeuwen, Podareanu, Austri, Verheyen (2019)]

Generative Adversarial Network (GAN)

[Bellagente, Butter, Kasieczka, Plehn, Winterhalder(2019)]
[Butter, Plehn, Winterhalder(2019)]

In GANs, the generator and discriminator network competes against each other.



[Image from Bellagente, Butter, Kasieczka, Plehn, Winterhalder (2019)]

- Discriminator works to distinguish generated data $\{x_G\}$ from truth data $\{x_p\}$. $[D(x_P) \to 1, D(x_G) \to 0]$

- Generator works to fool the discriminator such that $D(x_G) \to 1$.

$$L_{\text{Discriminator}} = \langle -logD(x) \rangle_{x\sim P_p} + \langle -log(1 - D(x)) \rangle_{x\sim P_G}$$

$$L_{\text{Generator}} = \langle -logD(x) \rangle_{x\sim P_G}$$

Normalizing flows

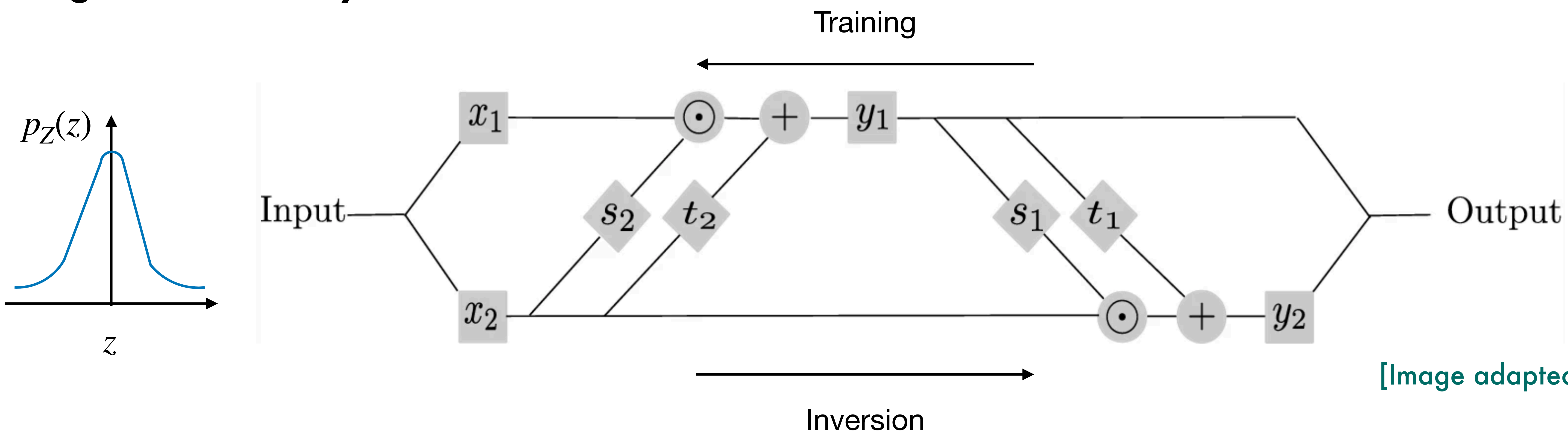
- Series of bijective layers that transform complex (Y) to simple probability distributions (Z).

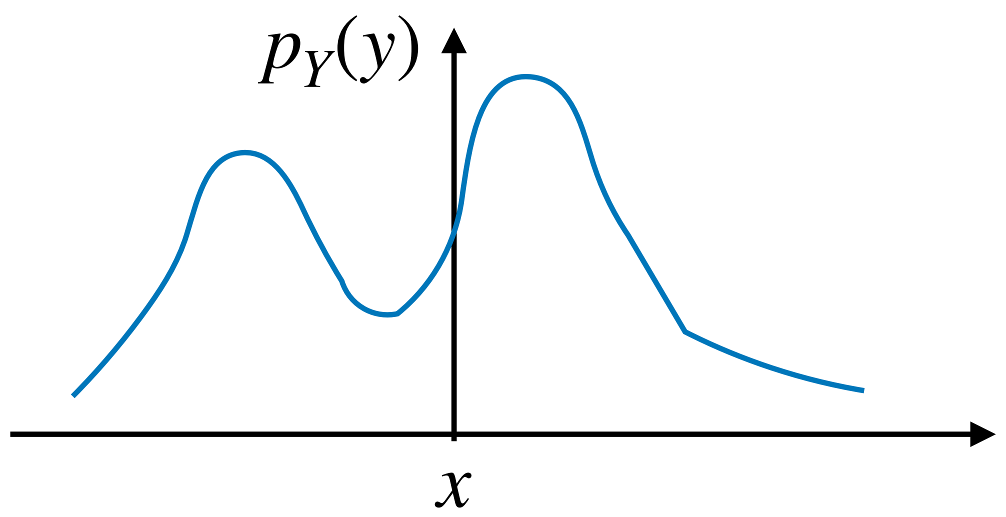
- Tractable Jacobian J : $p_Y(y) = p_Z(g(y))J$

  $Z = g(Y)$

  g: Invertible function

☑ Sampling and density estimation.





[Image adapted from Nguyen, Ardizzone, Kothe (2019) and talk by A. Butter at Pheno-2022]

- Bijective map between parton-level and detector-level phase space

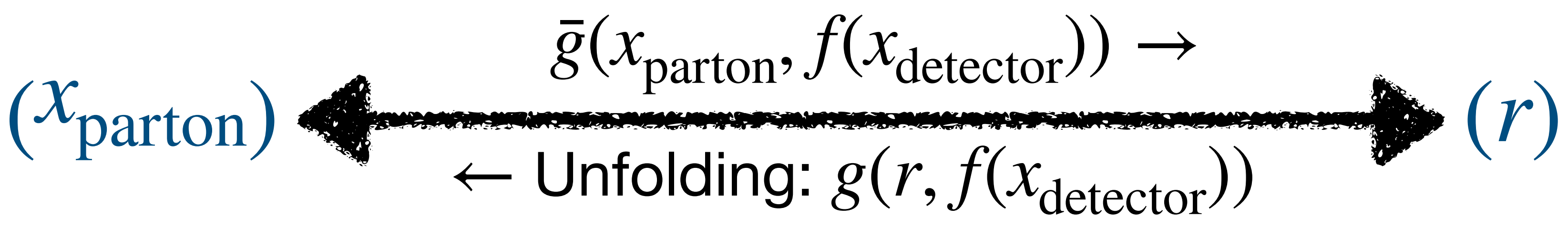
$$\begin{pmatrix} x_{\text{parton}} \\ r_{\text{parton}} \end{pmatrix} \longleftrightarrow \begin{pmatrix} x_{\text{detector}} \\ r_{\text{detector}} \end{pmatrix}$$

Forward simulation: $\bar{g}$ →

← Unfolding: g

[Bellagente, Butter, Kasieczka, Plehn, Rousselot, Winterhalder, Ardizzone, Kothe (2020)]

Conditional INN

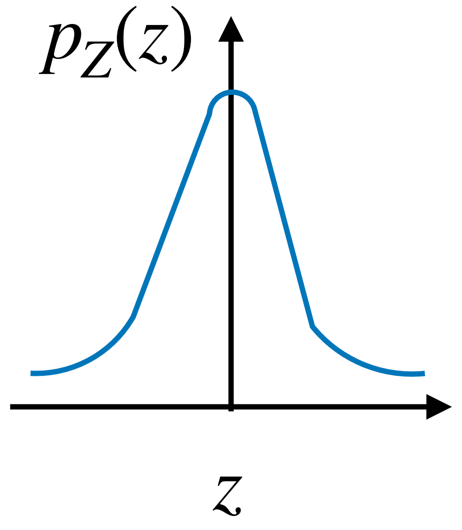


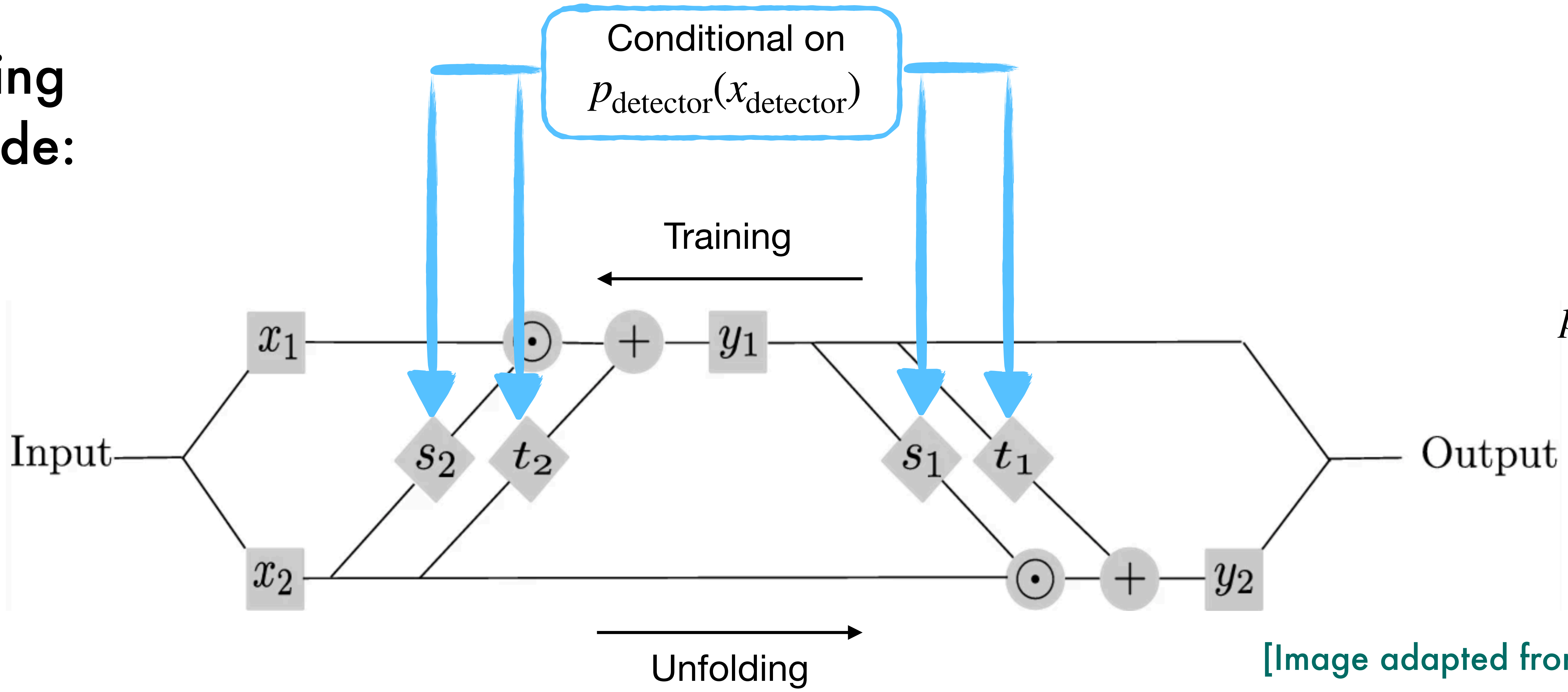
- Generate probability distributions at the parton-level, given detector-level events x_{detector}
  - $\longrightarrow$ Conditional on x_{detector}
  - $\longrightarrow$ x_{parton} mapped with r

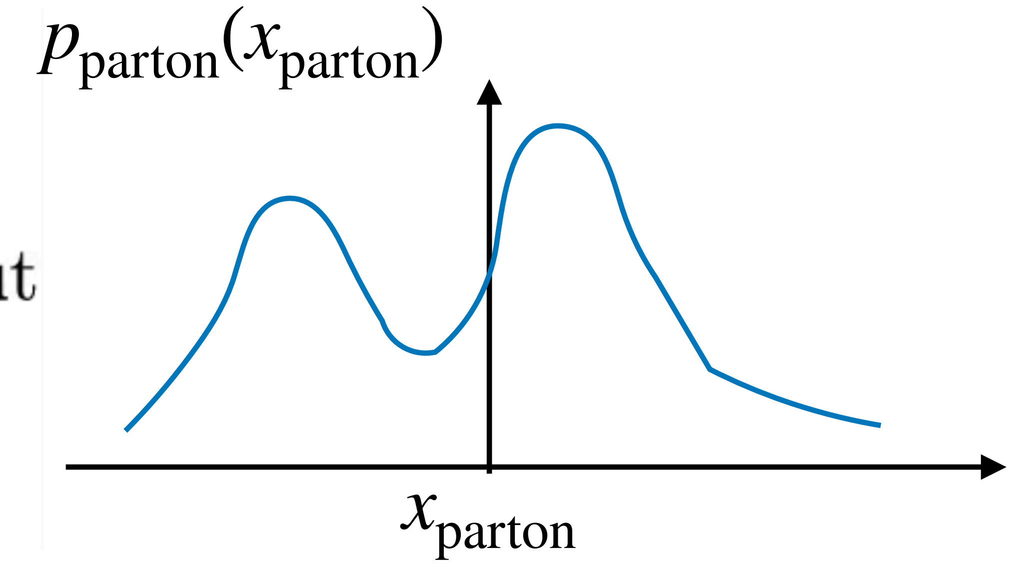
[Bellagente, Butter, Kasieczka, Plehn, Rousselot, Winterhalder, Ardizzone, Kothe (2020)]

Target phase space for unfolding can be chosen flexibly to include:

- ☑ jet radiation
- ☑ Particle decays







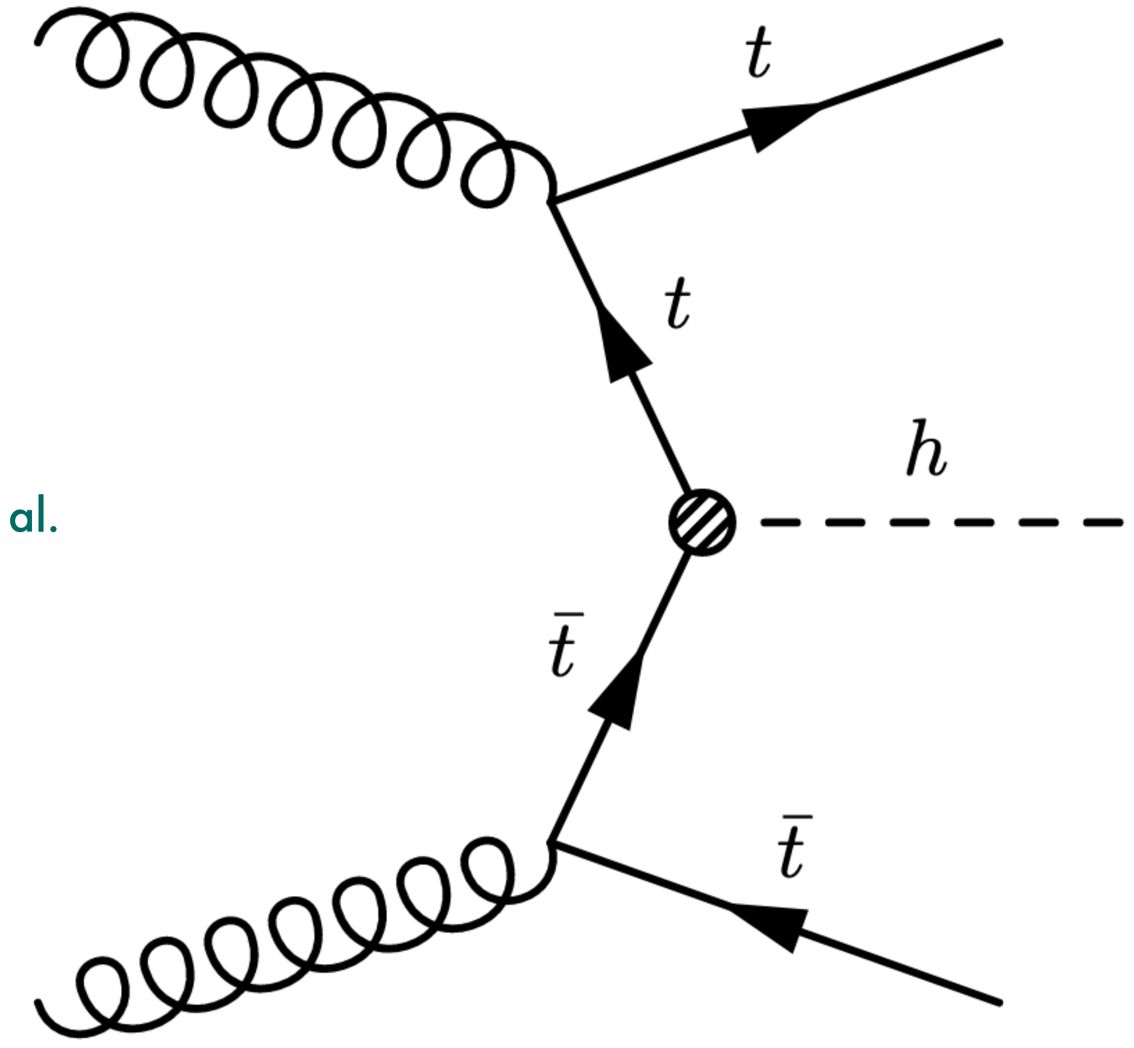
[Image adapted from Nguyen, Ardizzone, Kothe (2019) and talk by A. Butter at Pheno-2022]

- We use the Bayesian version of cINN
  - Stable network predictions
  - Allows the estimation of training-related uncertainties.

[Butter, Heimel, Hummerich, Krebs, Plehn, Rousselot, Vent (2021)]

CP measurement in Higgs-top interactions

- New sources of CPV interactions can explain the matter-antimatter asymmetry in the universe.



- CPV in hVV interactions is extensively tested at the LHC.

[See for instance: G. Aad et al. (1506.05669), G. Aad et al. (1602.04516), A. M. Sirunyan et al. (1707.00541), A. M. Sirunyan et al. (1903.06973), A. M. Sirunyan et al.(1901.00174), G. Aad et al. (2002.05315), Bernreuther, Gonzalez, Wiebusch (2010), Englert, Goncalves, Mawatari, Plehn (2012), Djouadi, Godbole, Mellado, Mohan (2013), Anderson, Bolognesi, Caola, Gao et al. (2013)]

- CPV in $hf\bar{f}$ couplings manifest at tree-level:

  $\rightarrow$ desirable choice: $ht\bar{t}$

$$\mathscr{L} = -\frac{m_t}{v}\,\kappa_t\, h\, \bar{t}\,(\cos\alpha \;+\; i\,\gamma_5\;\sin\alpha)\;t$$

- $pp \rightarrow h$ (+ jets): indirect constraints.

[Duca, Kilgore, Oleari, Schmidt, Zeppenfeld (2001), Klamke, Zeppenfeld (2007), Dolan, Harris, Jankowiak, Spannowsky (2014)]

Current limit (ATLAS: 2004.04545):
$|\alpha| < 43^0$ at 95 % CL

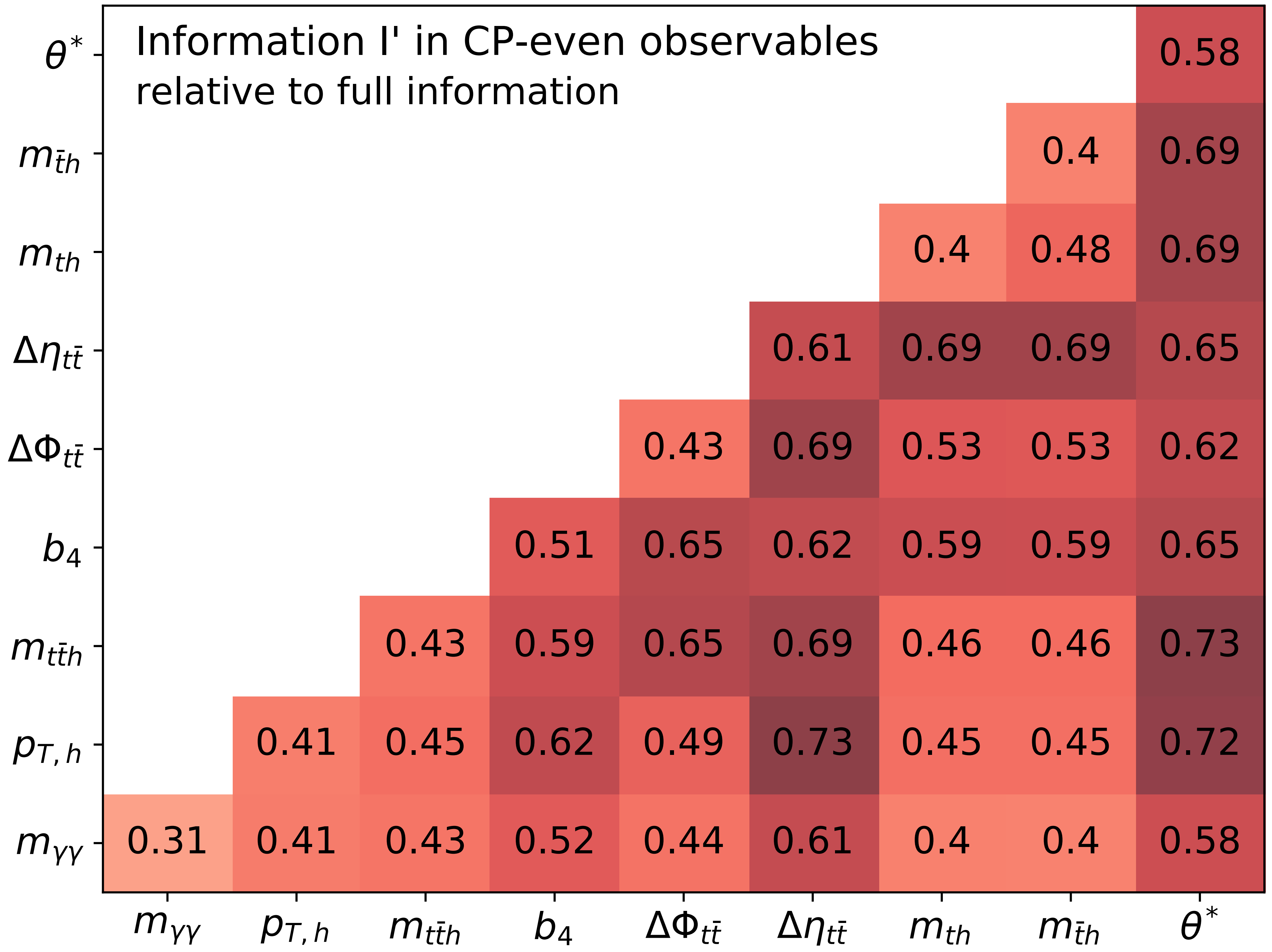
- $pp \rightarrow t\bar{t}h$: opportunity to directly probe α and κ_t

[Buckley, Goncalves (2016), Azevedo, Onofre, Filthaut, Goncalo (2017)]

Improved statistics @ HL-LHC paves the pathway for precision studies.

$t\bar{t}(h \to \gamma\gamma)$ @ HL-LHC

Importance matrix at the **non-linear level**



[RKB, Goncalves, Kling (2021)]

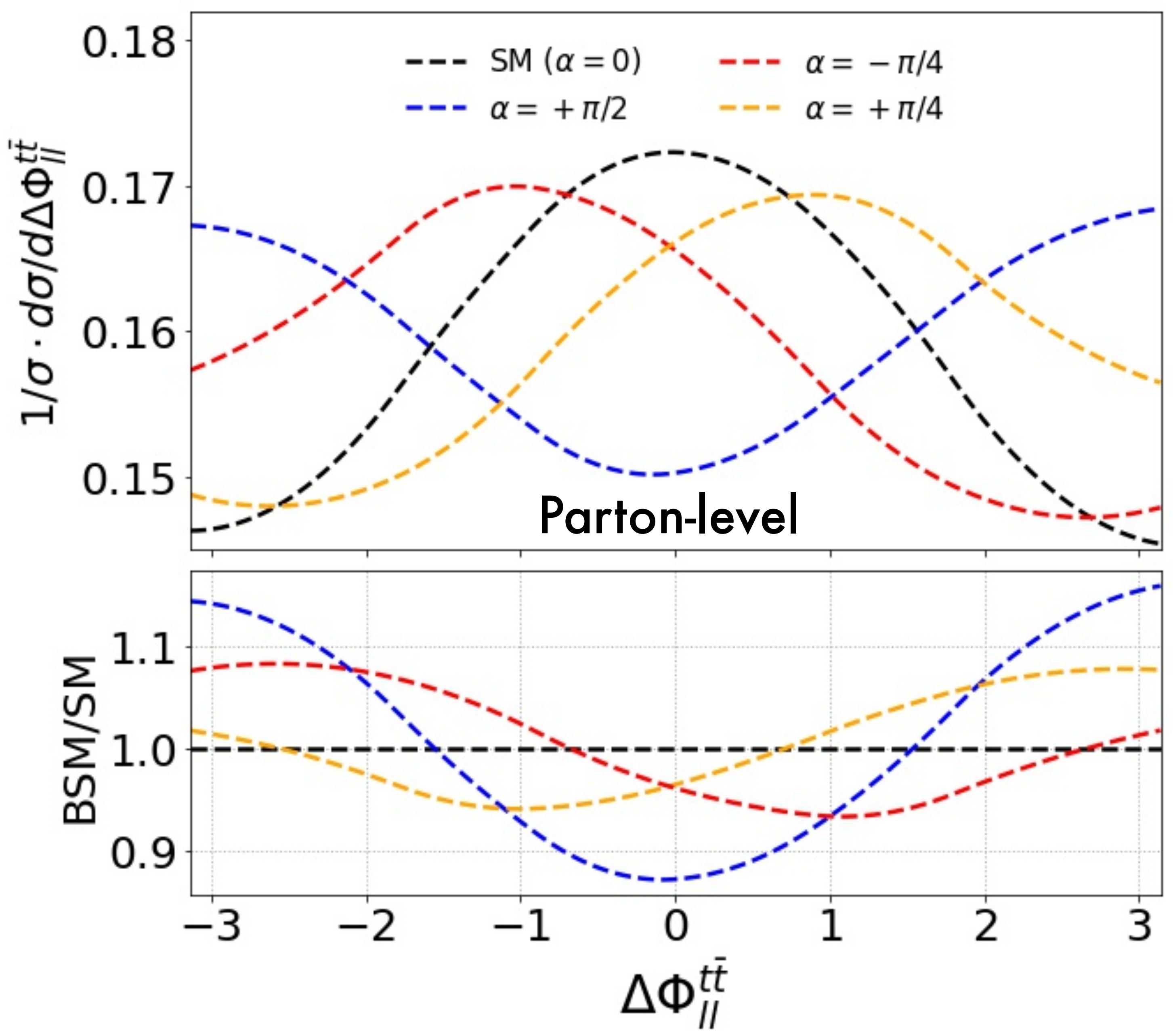
Sensitive only to non-linear new physics effects.

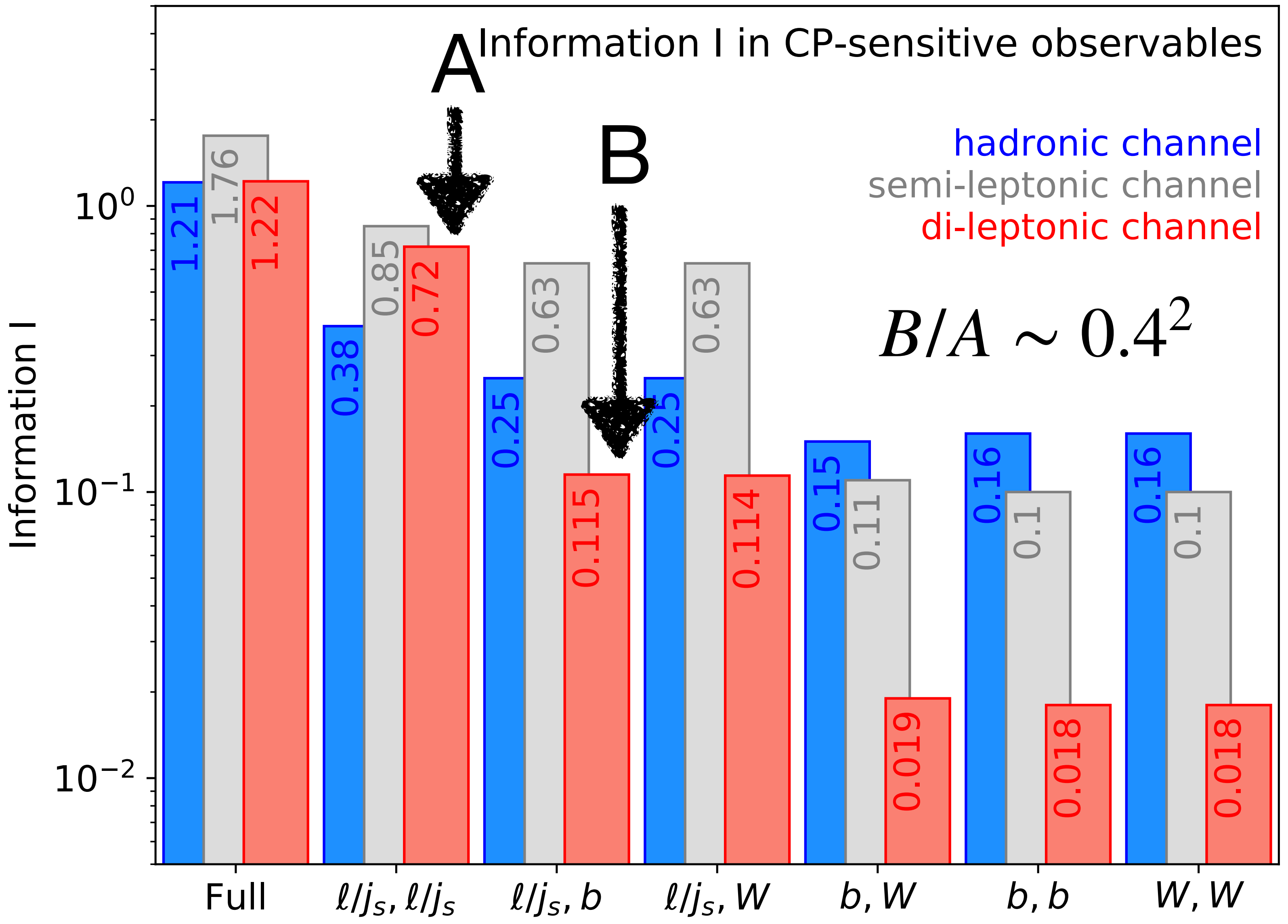
CP-odd observables

- Short lifetime for t $(10^{-25}\ s)$ → Spin correlations can be traced back from their decay products.

- CP-odd observables constructed from antisymmetric tensor products $\epsilon(p_t, p_{\bar{t}}, p_i, p_j) \sim \epsilon_{\mu\nu\rho\sigma} p_t^\mu p_{\bar{t}}^\nu p_i^\rho p_j^\sigma$:

$$\Delta\phi_{ij}^{t\bar{t}} = \textbf{sgn}\left[\vec{p}_t \cdot (\vec{p}_i \times \vec{p}_j)\right] \arccos\left[\frac{\vec{p}_t \times \vec{p}_i}{|\vec{p}_t \times \vec{p}_i|} \cdot \frac{\vec{p}_t \times \vec{p}_j}{|\vec{p}_t \times \vec{p}_j|}\right]$$





[RKB, Goncalves, Kling (2021)]

← Spin correlations scale with the spin analysing power β_i.

[Mileo, Kiers, Szynkman, Crane, Gegner (2016); Goncalves, Kong, Kim (2018)]; RKB, Goncalves, Kling (2021)]

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\xi_i} = \frac{1}{2}\left(1 + \beta_i P_t \cos\xi_i\right)$$

$$\text{Fisher Info} = \mathbb{E}\left[\frac{\partial \log p(x\,|\,\kappa_t, \alpha)}{d\alpha}\frac{\partial \log p(x\,|\,\kappa_t, \alpha)}{d\alpha}\right]$$

- Kinematic reconstruction efficiency is limited at the detector level

Use Machine learning techniques to maximize the extraction of NP information from CP observables.

Unfolding semileptonic $t\bar{t}h$ events

$$pp \to t\bar{t}h \to (t \to \ell\nu b)(\bar{t} \to jj\bar{b})(h \to \gamma\gamma)$$

➡ Parton-level:

$$1\ell + 2b + 2\gamma + \nu + 2j$$

➡ Detector-level:

$$1\ell + 2b + 2\gamma + MET + \leq 6 \text{ jets inclusive}$$

Acceptance cuts

$$|\eta_b| < 4\,, \quad |\eta_j| < 5\,, \quad |\eta_\ell| < 4, \quad |\eta_\gamma| < 4$$

$$p_{T,b} > 25 \text{ GeV}\,, \quad p_{T,j} > 25 \text{ GeV}\,, \quad p_{T,\ell} > 15 \text{ GeV}\,, \quad p_{T,\gamma} > 15 \text{ GeV}$$

Questions:

★ Can the unfolding model correctly reconstruct the two hard jets at the parton level from a variable number of jets at the detector level?

★ How well can the dedicated observables be reconstructed?

★ How model-dependent is the training?

Event parametrization

- Event information at the parton level can be parametrised through the 4-momentum of the final state particles → may include redundant d.o.f.

- Reconstruction of sharp kinematic features like mass peaks can be challenging:
  - ✓ Can be improved by adding targeted maximum mean discrepancy loss:
    - ☑ Affects only the target distributions
    - ☑ Avoids large model dependence
    - ☒ Complications in training and performance limitations.

[Butter, Plehn, Winterhalder (2019)]
[Bellagente, Butter, Kasieczka, Plehn, Rousselot, Winterhalder, Ardizzone, Kothe (2020)]

<u>Alternative approach:</u>

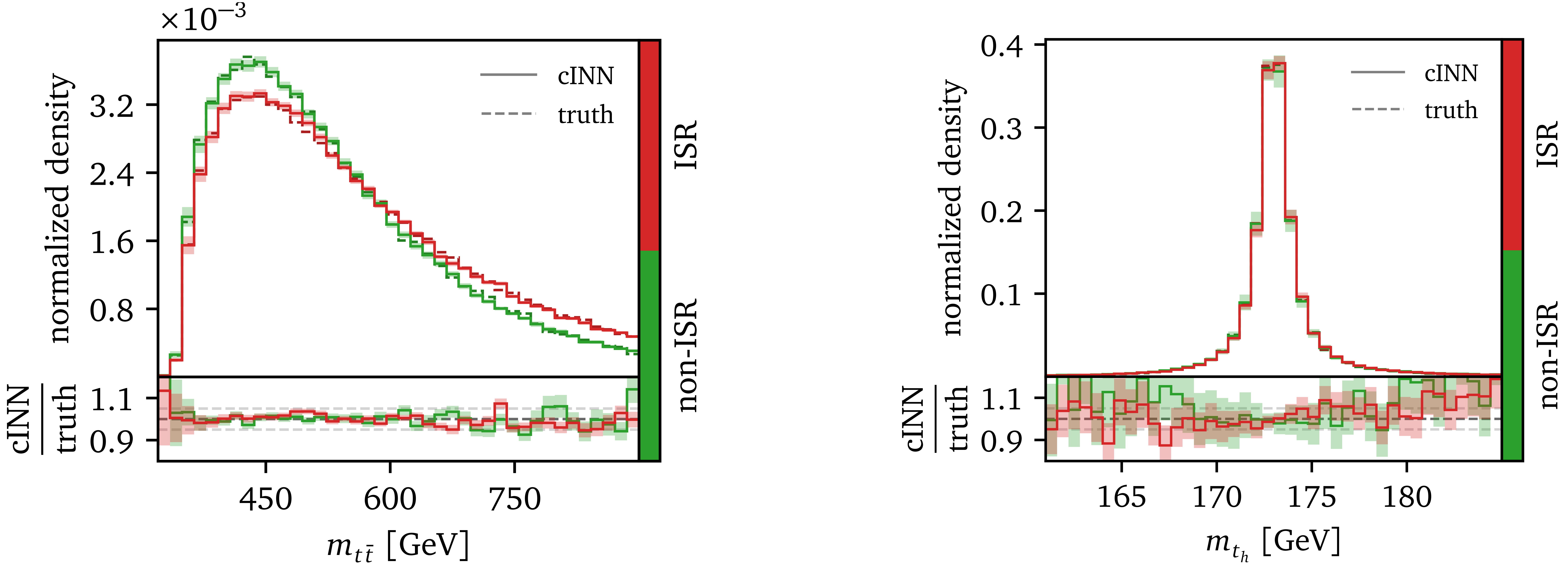
→ directly learn invariant mass features and important observable with appropriate phase-space parametrization.

→ may provide direct access to the most important CP-even and CP-odd observables.

$$\vec{p}_{t\bar{t}}, m_{t_\ell}, |\vec{p}_{t_\ell}^{\,\text{CS}}|, \theta_{t_\ell}^{\text{CS}}, \phi_{t_\ell}^{\text{CS}}, m_{t_h},$$
$$\text{sign}(\Delta\phi_{\ell\nu}^{t\bar{t}})\, m_{W_\ell}\, |\vec{p}_\ell^{\,t\bar{t}}|, \theta_\ell^{t\bar{t}}, \phi_\ell^{t\bar{t}}, |\vec{p}_\nu^{\,t\bar{t}}|$$
$$\text{sign}(\Delta\phi_{du}^{t\bar{t}})\, m_{W_h}, |\vec{p}_d^{\,t\bar{t}}|, \theta_d^{t\bar{t}}, \Delta\phi_{\ell d}^{t\bar{t}}, |\vec{p}_u^{\,t\bar{t}}|$$

Jet combinatorics

Parton level truth and unfolded top invariant masses m_{t_ℓ} and m_{t_h}

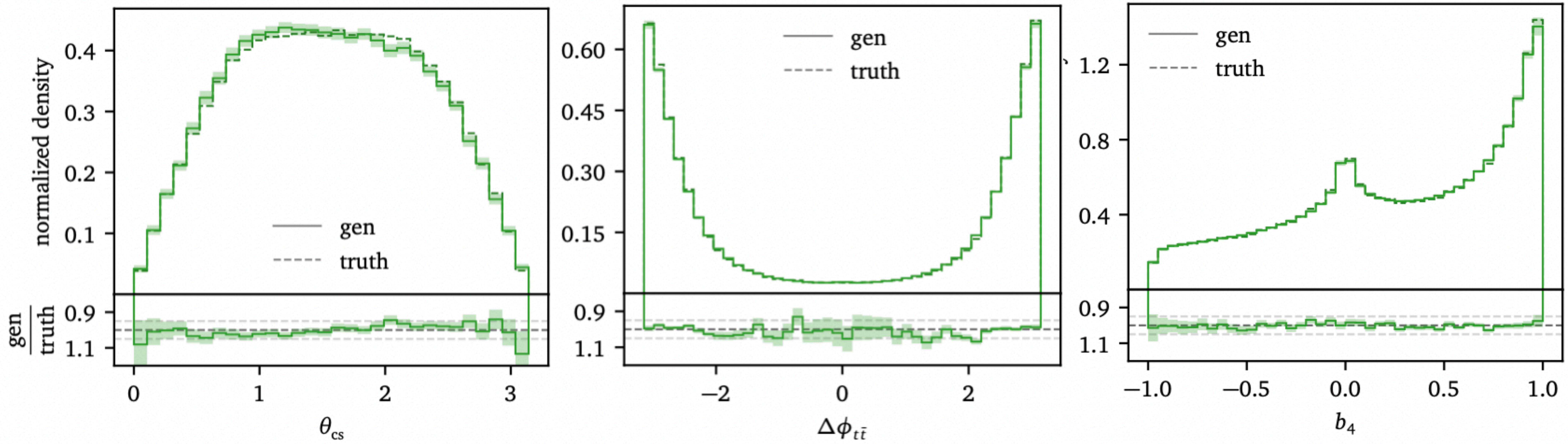


★ Unfolded distributions in good agreement with parton level truth despite added combinatorial ambiguity at the detector level.

★ Truth distributions within 1σ error bands.

Reconstruction of dedicated observables

Parton level truth and unfolded SM for θ_{CS}, $\Delta\phi_{t_\ell t_h}$ and b_4.



★ Unfolded distributions in close agreement with truth:

✓ Close agreement even for observables not included in event parametrization.

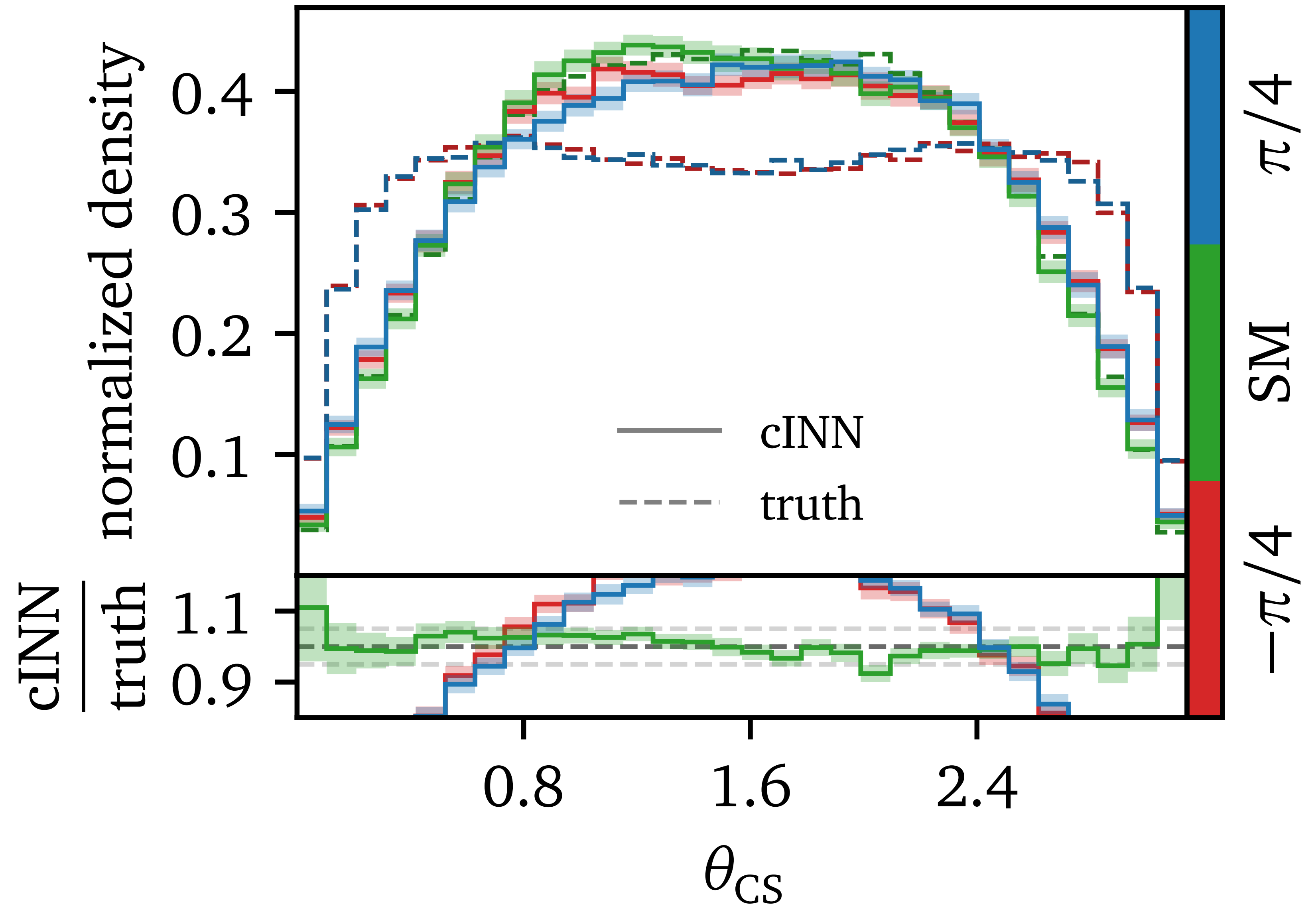
✓ Full phase space reconstruction.

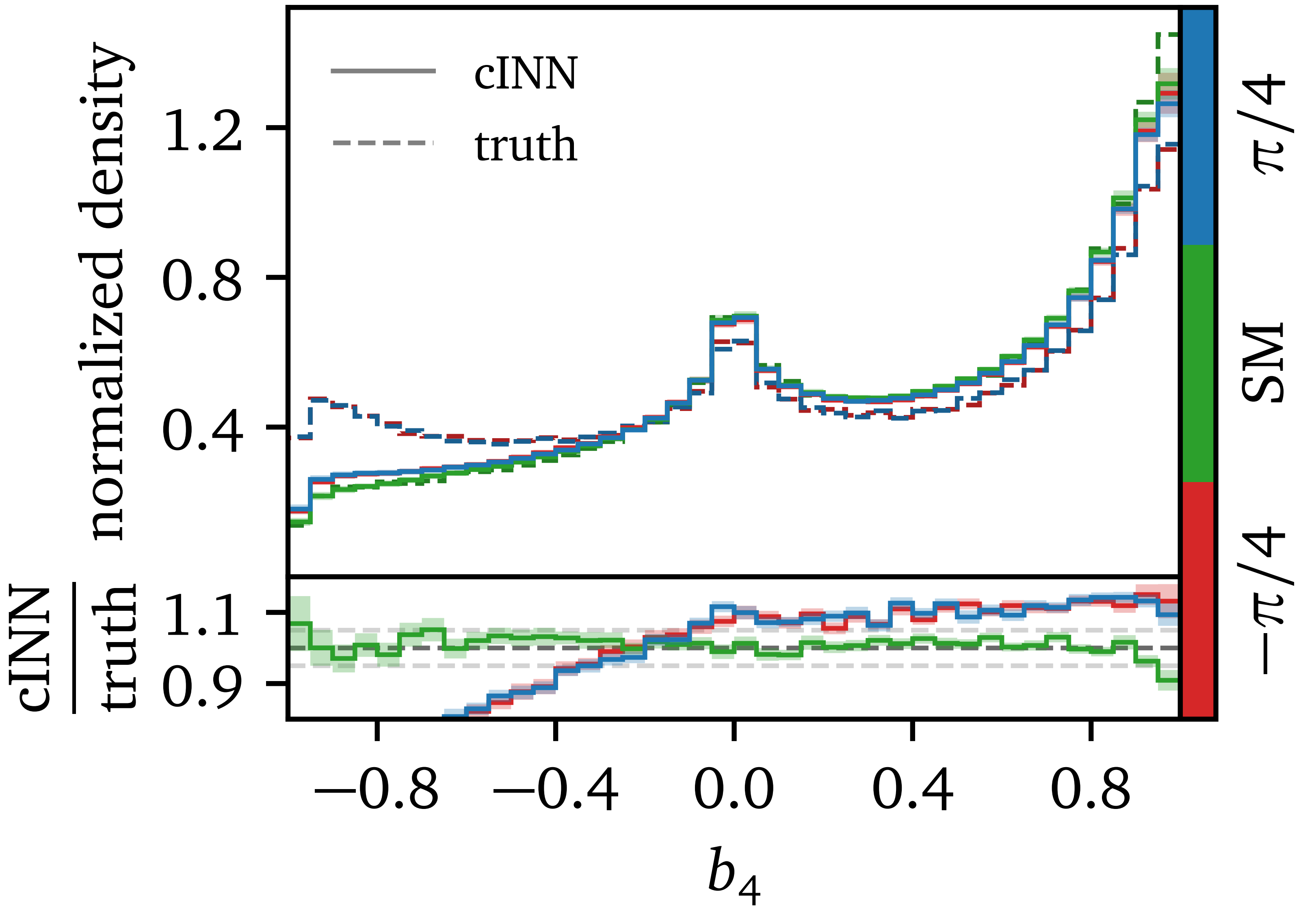
★ Potential differences from the truth are covered by the uncertainty estimates of the Bayesian network.

Model dependence

Unfolding SM events using networks trained on events with different amounts of CP-violation.

We train 3 networks on $\alpha = +\pi/4, -\pi/4$ and SM, respectively
Unfold SM dataset





★ Networks trained on $\alpha = \pi/4$ and $-\pi/4$ show only a slight bias towards broader θ_{CS} and flatter b_4 distributions.

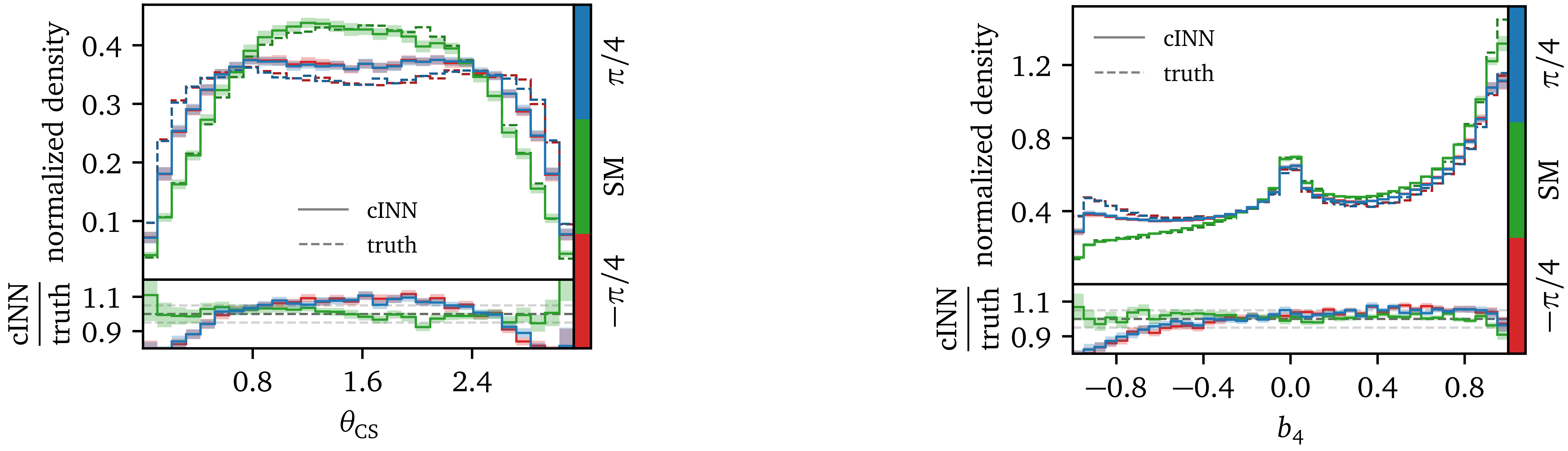
★ $\sim 10 - 20\,\%$ bias → much smaller than the changes at parton truth from varying α.

Model dependence

Unfolding events with CP-violation using a network trained on SM events.

Train network on SM dataset

Unfold $\alpha = +\pi/4, -\pi/4$ and SM dataset



★Again, the effect of bias is much smaller than the effect of α on the data.

Outlook

- Generative unfolding makes it possible to invert high-dimensional distributions and full phase-space reconstruction.

- The trained cINN behaves as an efficient kinematic reconstruction algorithm capable of tackling complex reconstruction challenges.

- The trained unfolding network was able to
  - extract various CP observables at the parton level with appropriate phase space parametrization.
  - resolve jet combinatorial ambiguity.
  - absolve any large model-dependence.

- Promising outlook for an experimental study, with a proper treatment of statistical limitations, continuum backgrounds, calibration, and iterative improvements of the unfolding network.

Thank you