UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Unfolding without Iterations, Adversaries, or Surrogates

Ayodele Ore, *EuCAIFCon 2026 Heidelberg*

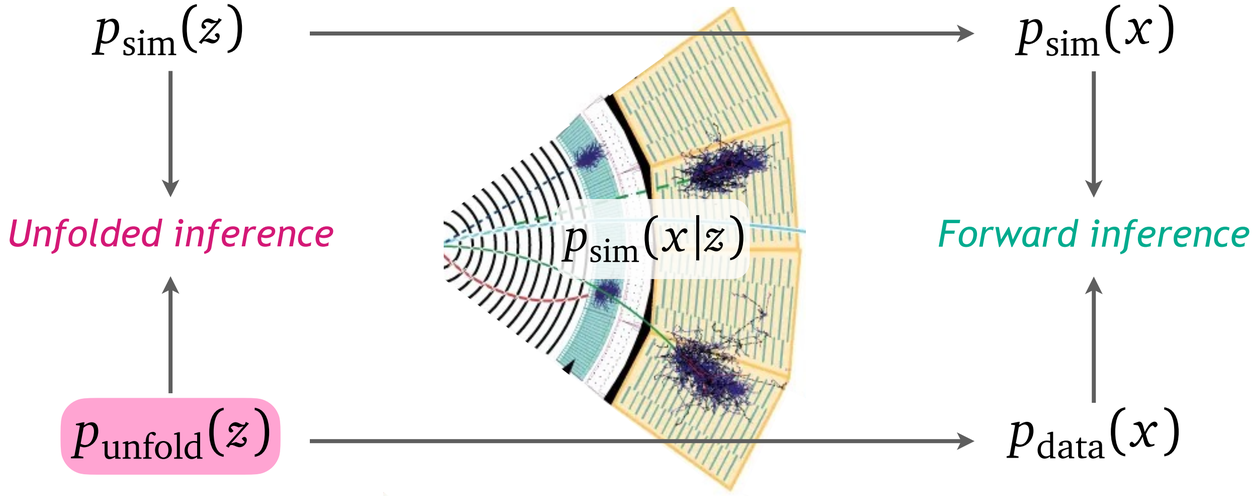
Based on arXiv:2602.24282 with Tilman Plehn

Poster #90

Unfolding

- ***Unfolding =*** *correcting data for detector effects*
- *Inverse problem: distribution-level deconvolution*
- *By unfolding:*
  - *Cost moves from repeated simulation to a once-off inference.*
  - *Theorists can test **any** model later, outside experimental collaboration*

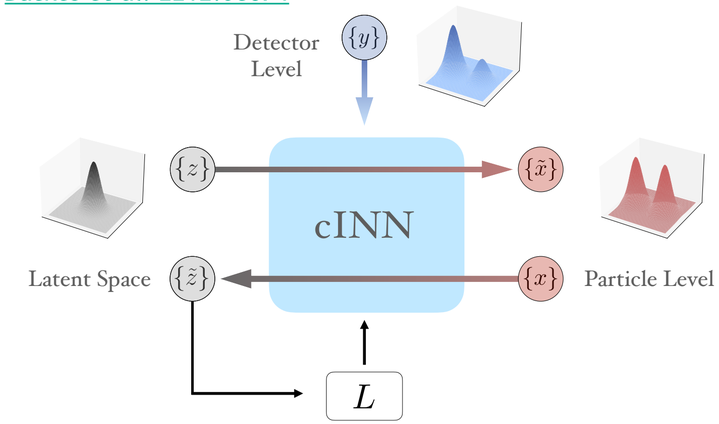
$$p_{\text{data}}(x) = \int dz\; p_{\text{sim}}(x|z)\, p_{\text{unfold}}(z)$$



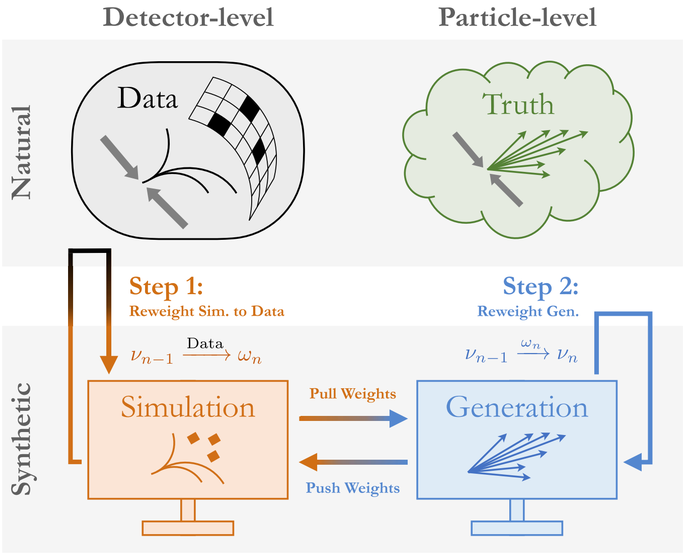
ML methods: Iterations, Adversaries, Surrogates

Iterative

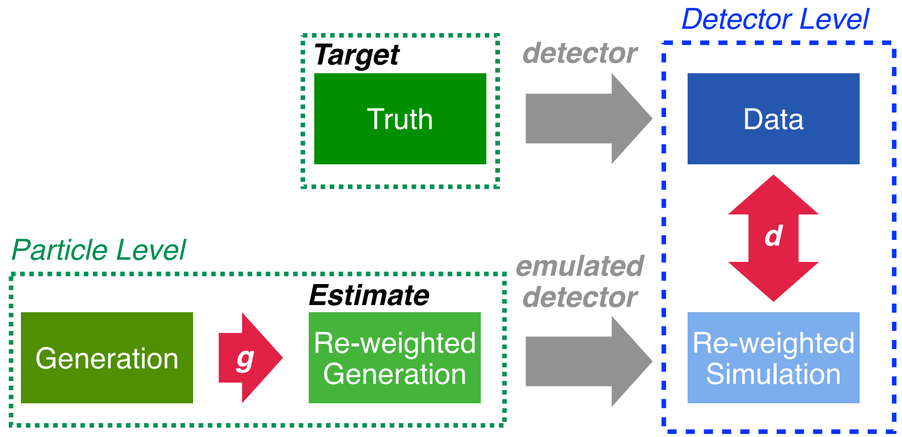
Backes et al. 2212.08674



Andreassen at al. 1911.09107

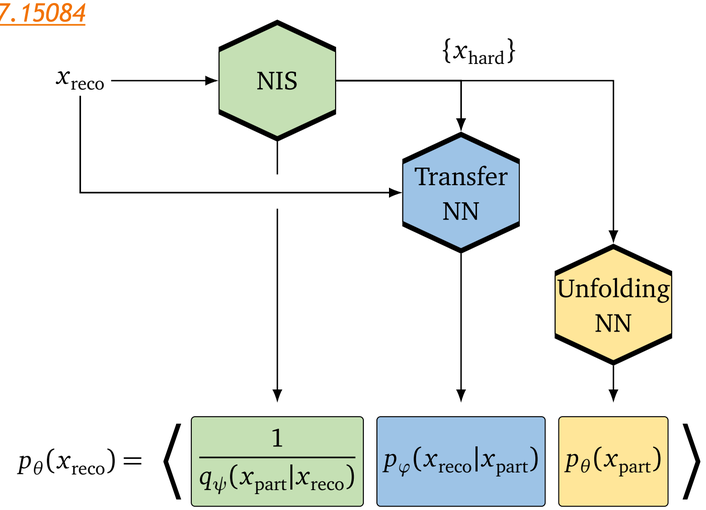


Desai et al. 2407.11284



Adversarial

Butter et al. 2507.15084



Surrogate-based

AUSSIE

- *We introduce:*

  ***A**dversary-free **U**nfolding **S**an**S** **I**teration or **E**mulation*

  (AUSSIE)

- *Based on **density ratios**, similar to OmniFold*

Step 1: *Estimate ratio $R(x)$ on observables using a classifier*

Step 2: *Learn $\overline{R}(z)$ directly using a **new loss** targeting the true relationship with $R(x)$*

$$p_{\text{data}}(x) = \int dz\; p_{\text{sim}}(x|z)\; p_{\text{unfold}}(z)$$

$$\frac{p_{\text{data}}(x)}{p_{\text{sim}}(x)} = \int dz\; p_{\text{sim}}(z|x)\; \frac{p_{\text{unfold}}(z)}{p_{\text{sim}}(z)}$$

$R(x)$ $\qquad\qquad$ $\overline{R}(z)$

AUSSIE Step 2

- *How can we train* $\overline{R}_\varphi(z) \approx \overline{R}(z)$ *?*
- ***OmniFold**'s second step is essentially:* $\left(R(x) - \overline{R}_\varphi(z)\right)^2$
  - *Mismatched function spaces!* ⇒ *Iteration needed*

Target

$$R(x) = \int dz\, p_{\text{sim}}(z|x)\, \overline{R}_\varphi(z)$$

AUSSIE Step 2

- *How can we train* $\overline{R}_\varphi(z) \approx \overline{R}(z)$ *?*
- ***OmniFold**'s second step is essentially:* $\left(R(x) - \overline{R}_\varphi(z)\right)^2$
  - *Mismatched function spaces! ⇒ Iteration needed*
- ***AUSSIE** targets the integral relation directly*
  - *Plain MSE between LHS and RHS not tractable*
  - *Instead use **kernel norm** of the residual:*

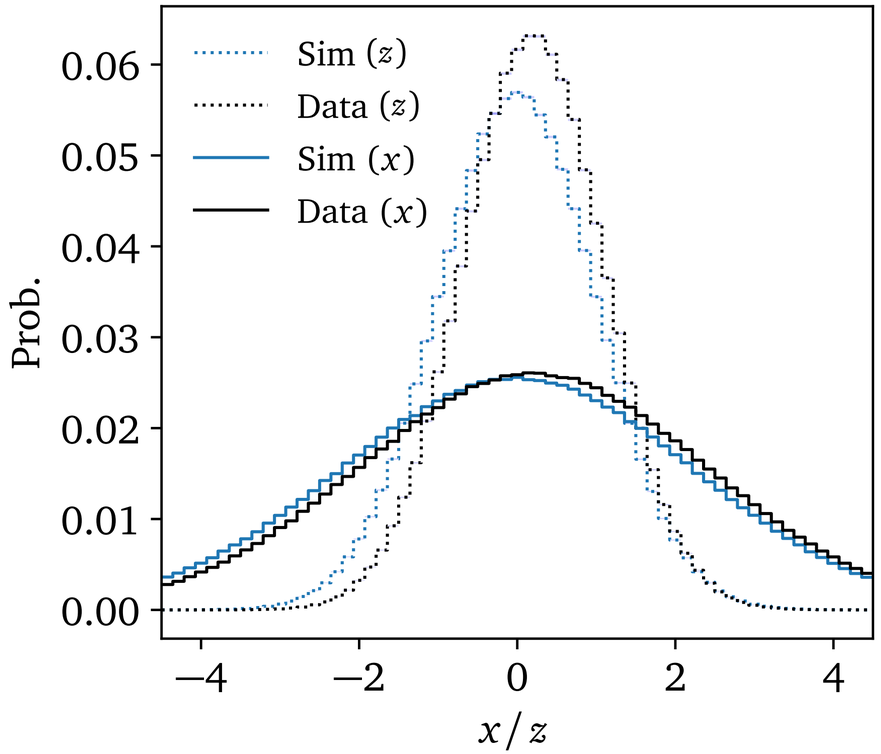
Target

$$R(x) = \int dz\, p_{\text{sim}}(z|x)\, \overline{R}_\varphi(z)$$

$$\mathcal{L}_{\text{AUSSIE}}[\overline{R}_\varphi] = \left\langle \left(R(x) - \overline{R}_\varphi(z)\right) K(x,x') \left(R(x') - \overline{R}_\varphi(z')\right) \right\rangle_{p_{\text{sim}}(x,z)p_{\text{sim}}(x',z')}$$

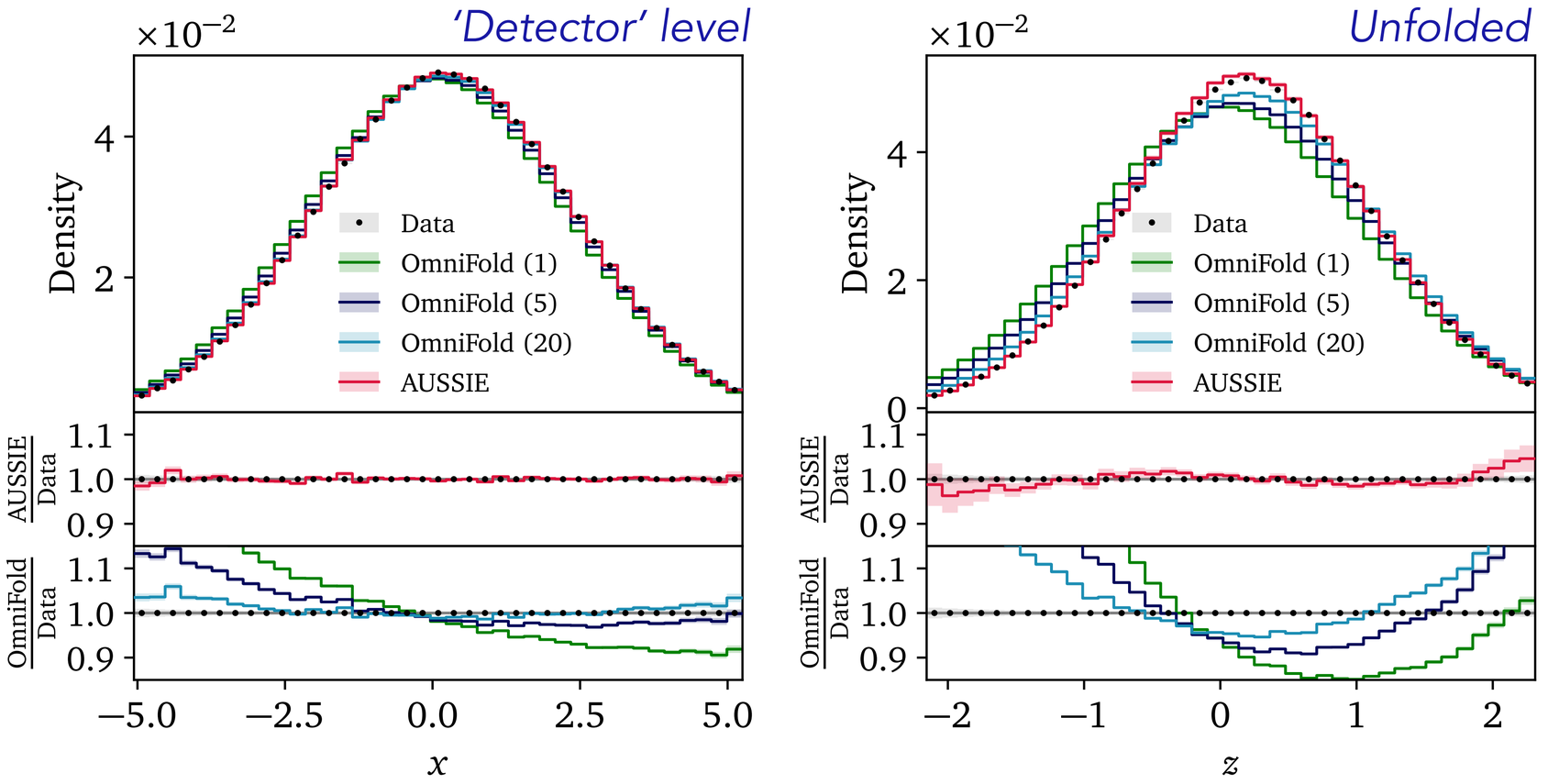
Gaussian example

- *Toy model with large smearing exemplifies advantage of AUSSIE*

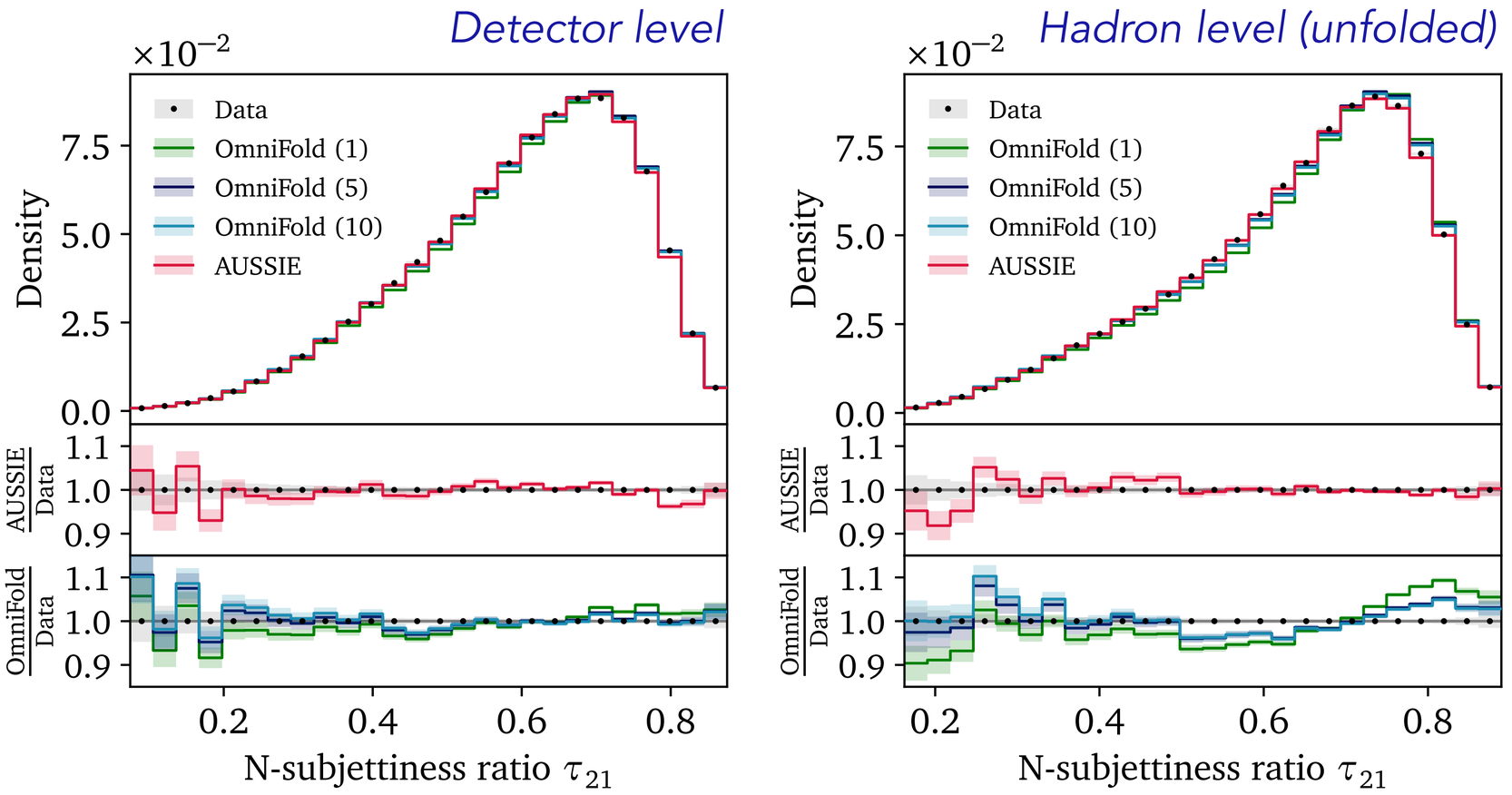


Gaussian example

- *Gaussian kernel works excellently in low dim*
- *AUSSIE matches while OmniFold struggles*



Jet substructure (constituent-level)

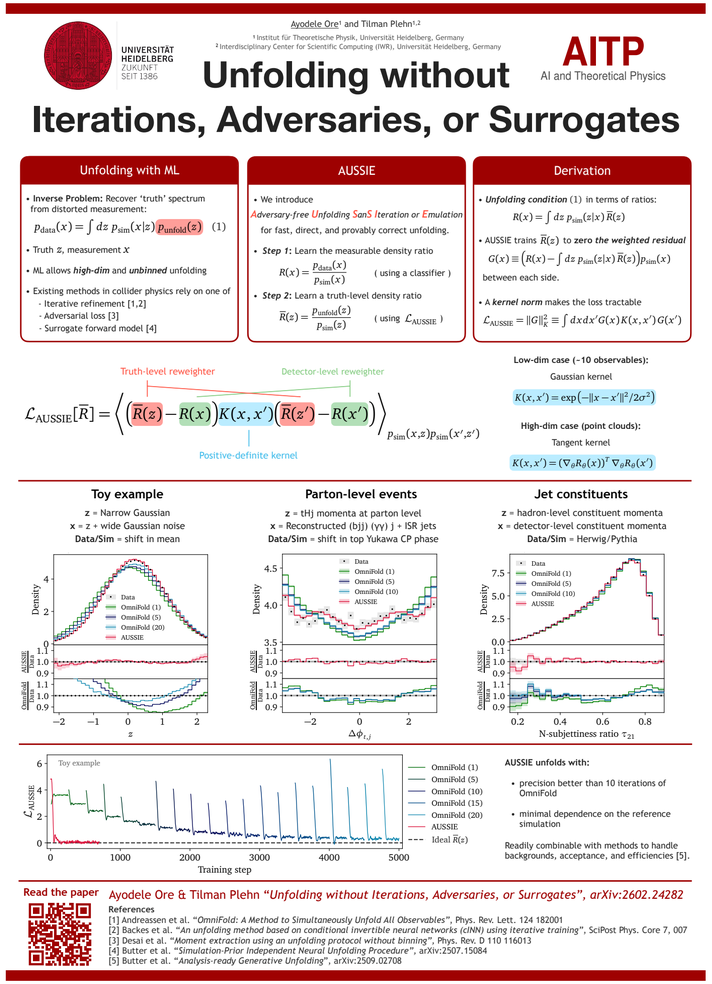


- ***Full phase space unfolding:*** *particle cloud up to N=150*
- *Adaptive kernel from* ***Neural Tangent Kernel*** *of step-1 classifier:*

$$K(x, x') = \left(\nabla_\theta R_\theta(x) \right)^T \nabla_\theta R_\theta(x')$$

- *AUSSIE mitigates residual bias toward simulated dataset*

Summary

- ***AUSSIE*** *is a new reweighting-based unfolding method, requiring*
  - ✓ *No iterative refinement*
  - ✓ *No adversarial training*
  - ✓ *No surrogate model*
- *Excellent performance across a range of dimensions:*
  - ✓ *Minimal dependence on reference sim*
  - ✓ *Readily extendible to account for efficiency / acceptance effects*
- *Links:*

Poster #90