Unfolding without Iterations, Adversaries, or Surrogates

Ayodele Ore

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

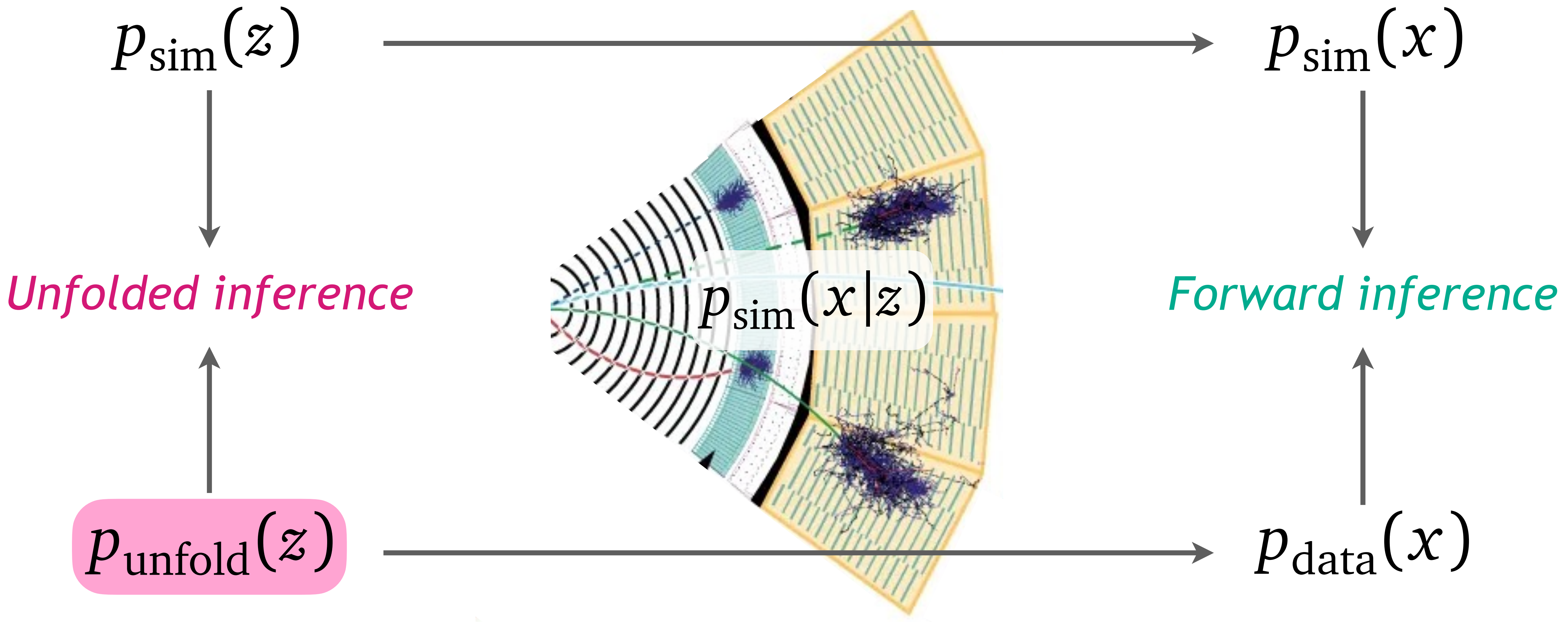
Based on arXiv:2602.24282 with Tilman Plehn

ML4Jets 2026, Vienna

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*

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



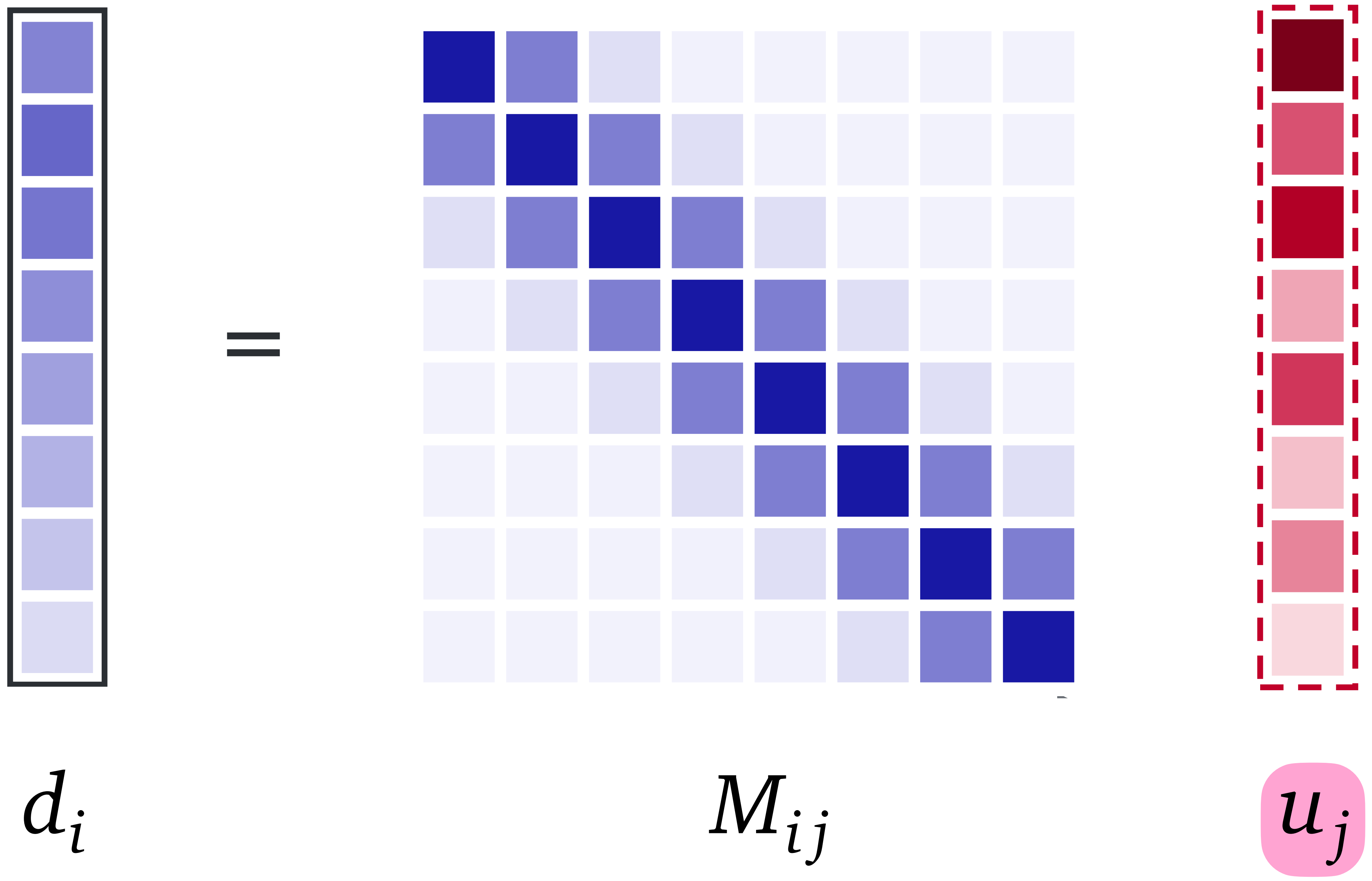
Classical approach: Matrix unfolding

- *Traditional unfolding first* ***discretises into histograms***

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

$$\downarrow$$

$$d_i = \sum_j M_{ij} u_j$$

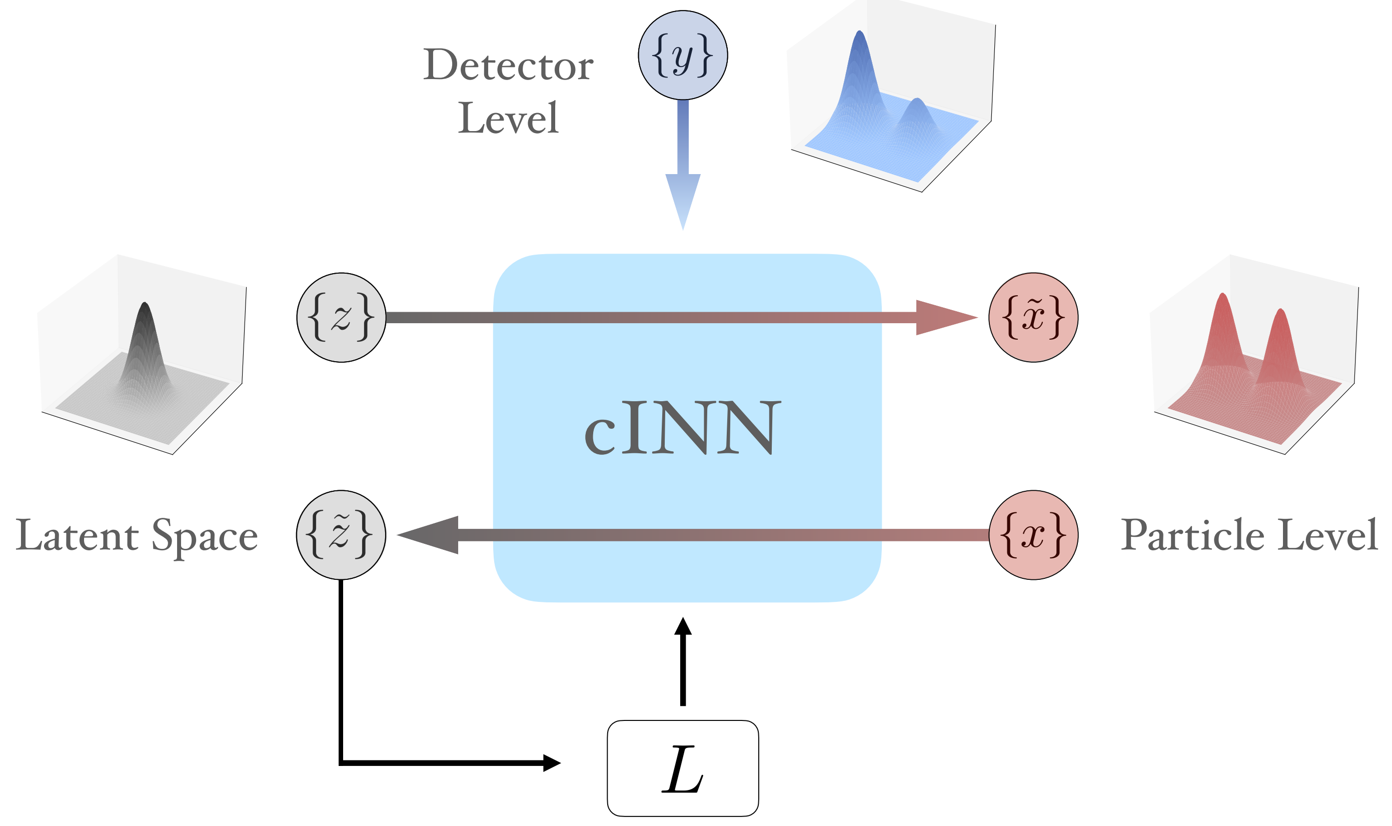


  - *Poor scaling to large N_{bins}*
  - *Undercuts the "unfold once" opportunity*

ML methods: Iterations, Adversaries, Surrogates

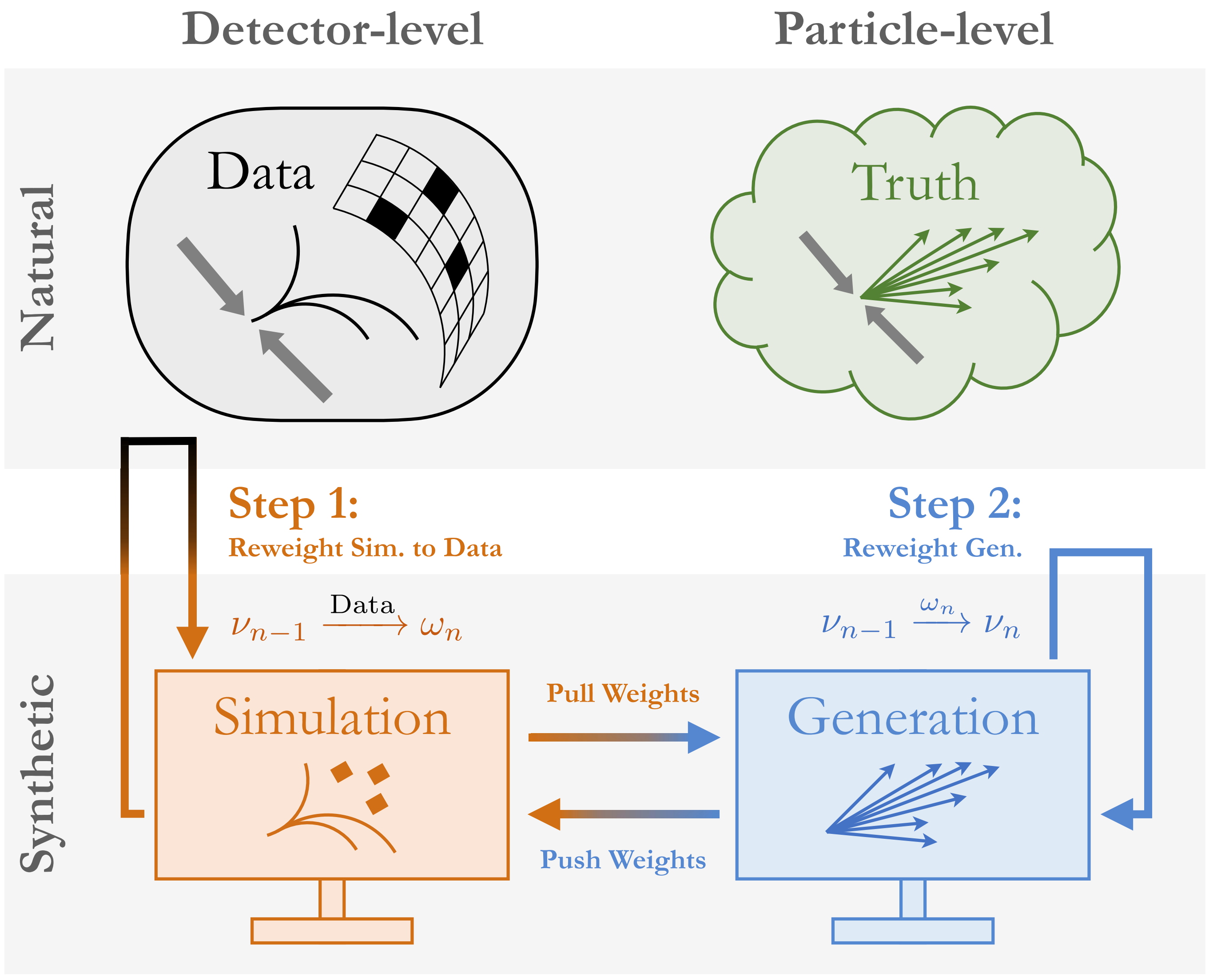
ML methods: Iterations, Adversaries, Surrogates

Backes et al. 2212.08674



Iterative

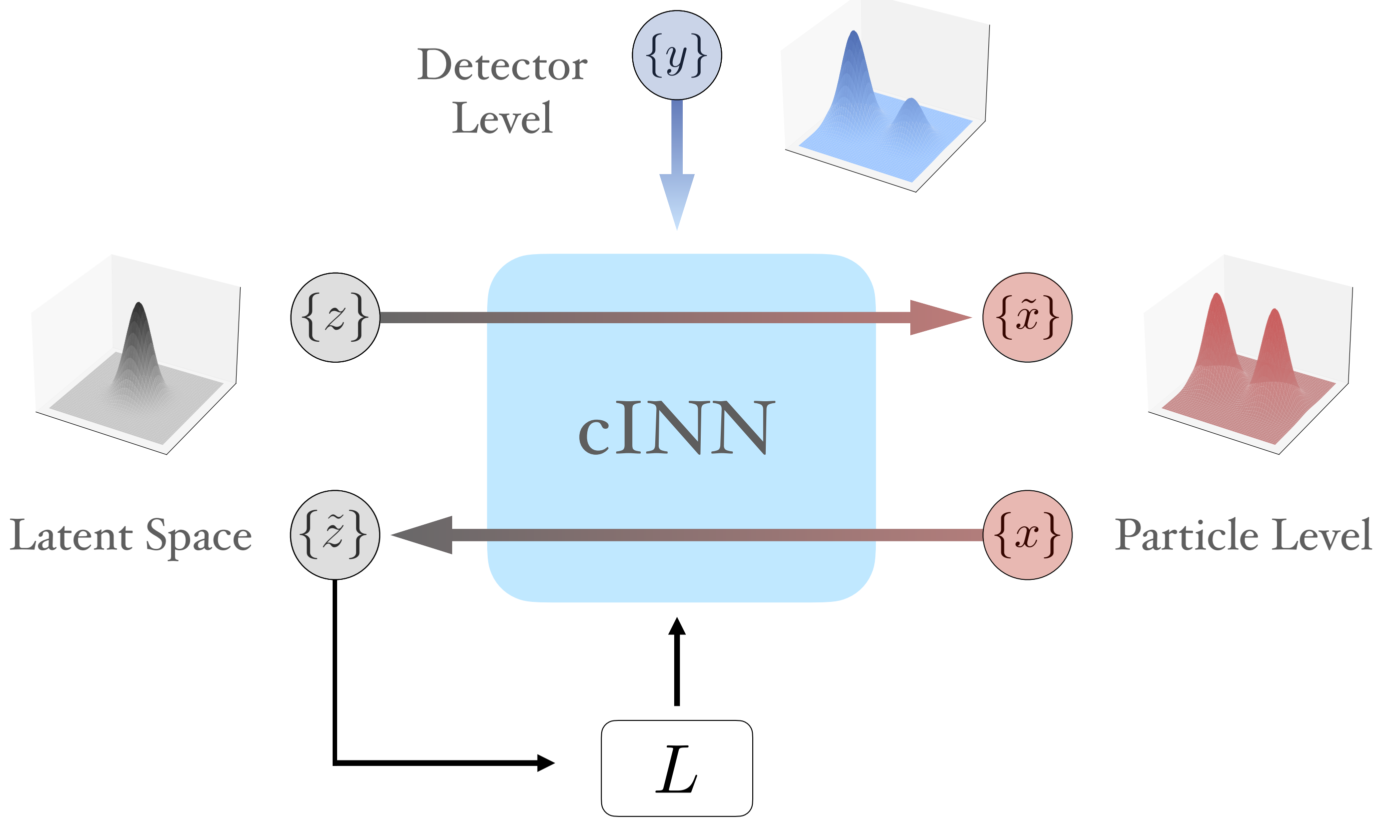
Andreassen at al. 1911.09107



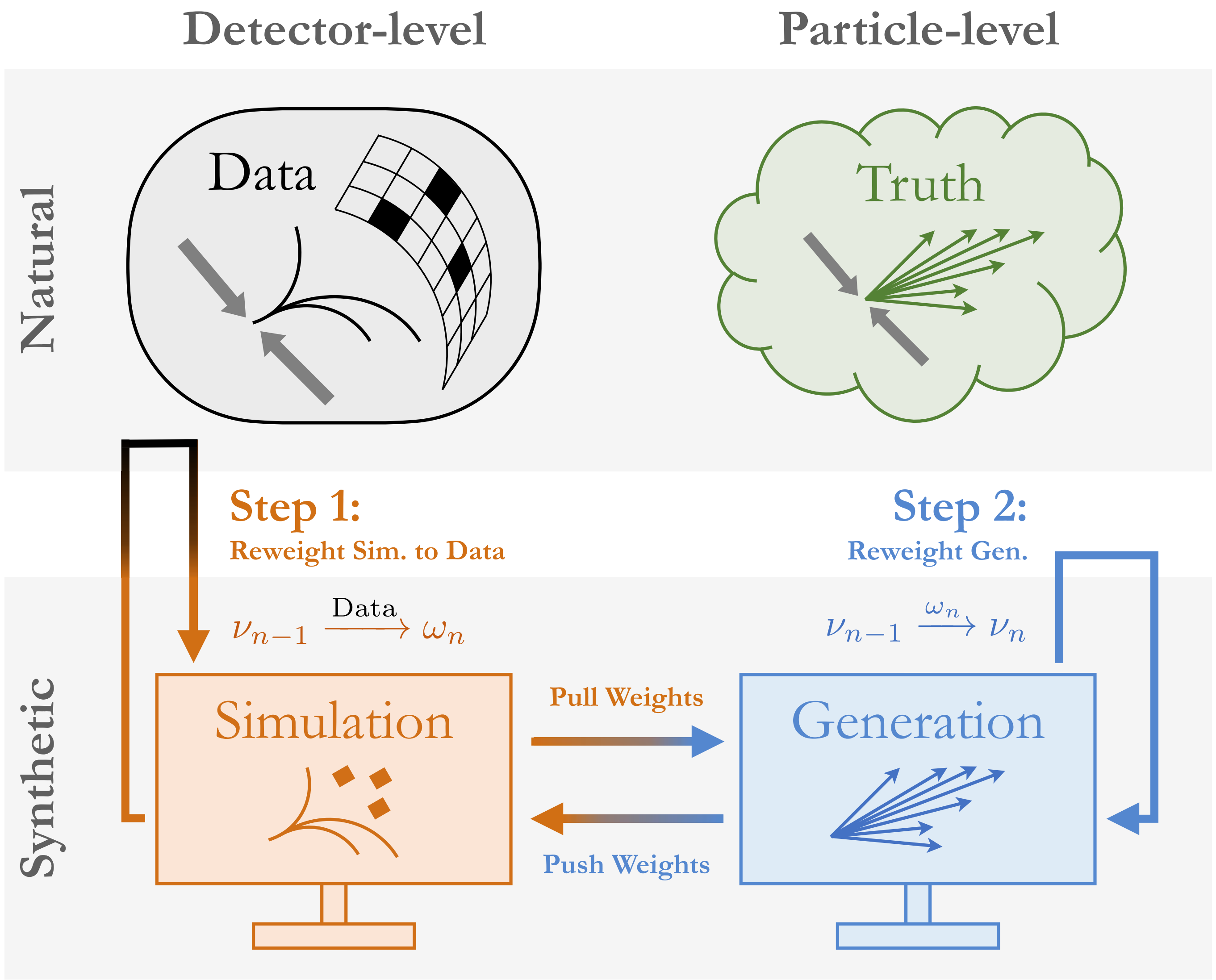
ML methods: Iterations, Adversaries, Surrogates

Iterative

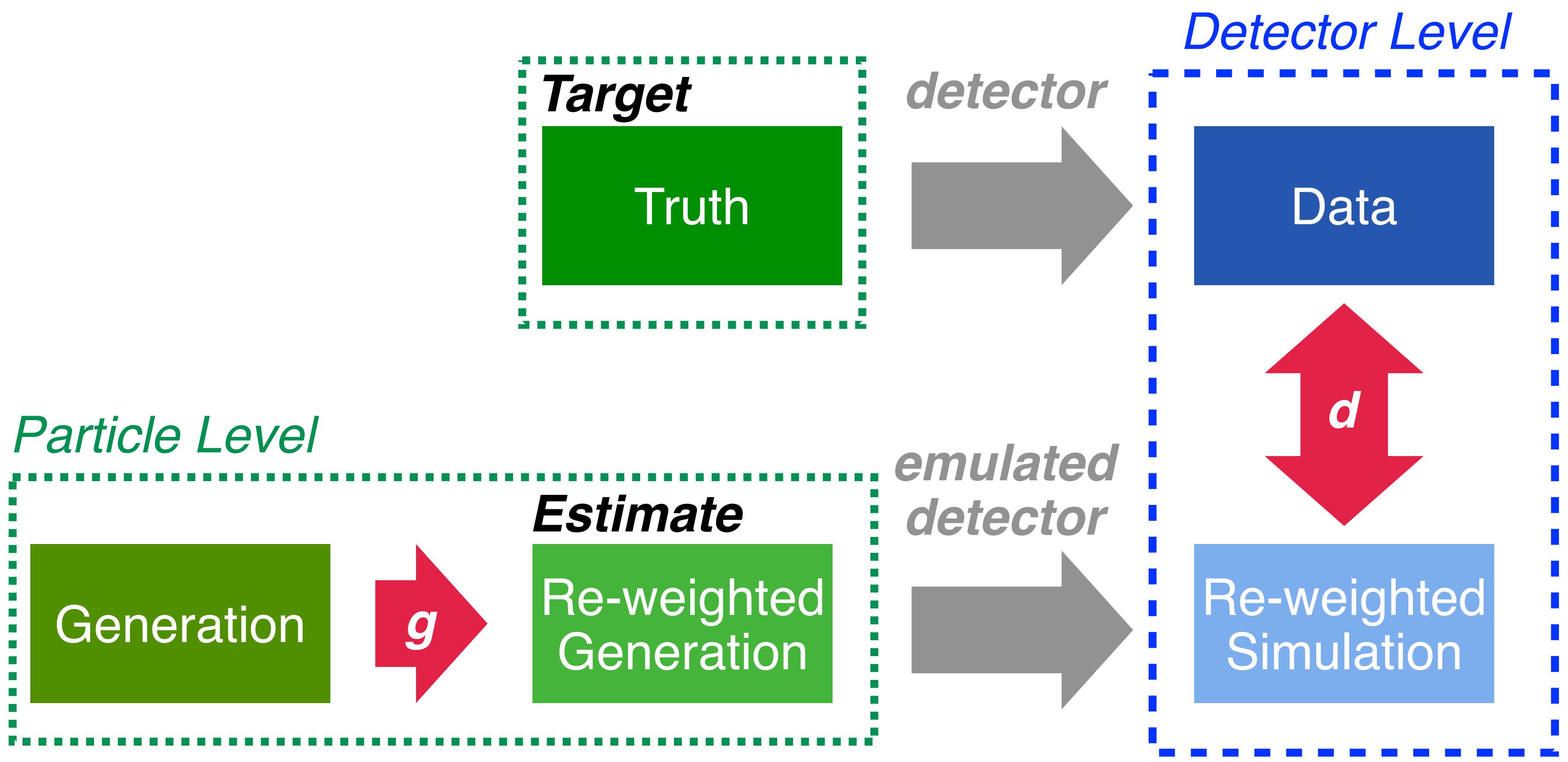
[Backes et al. 2212.08674](https://arxiv.org/abs/2212.08674)



[Andreassen at al. 1911.09107](https://arxiv.org/abs/1911.09107)



[Desai et al. 2407.11284](https://arxiv.org/abs/2407.11284)

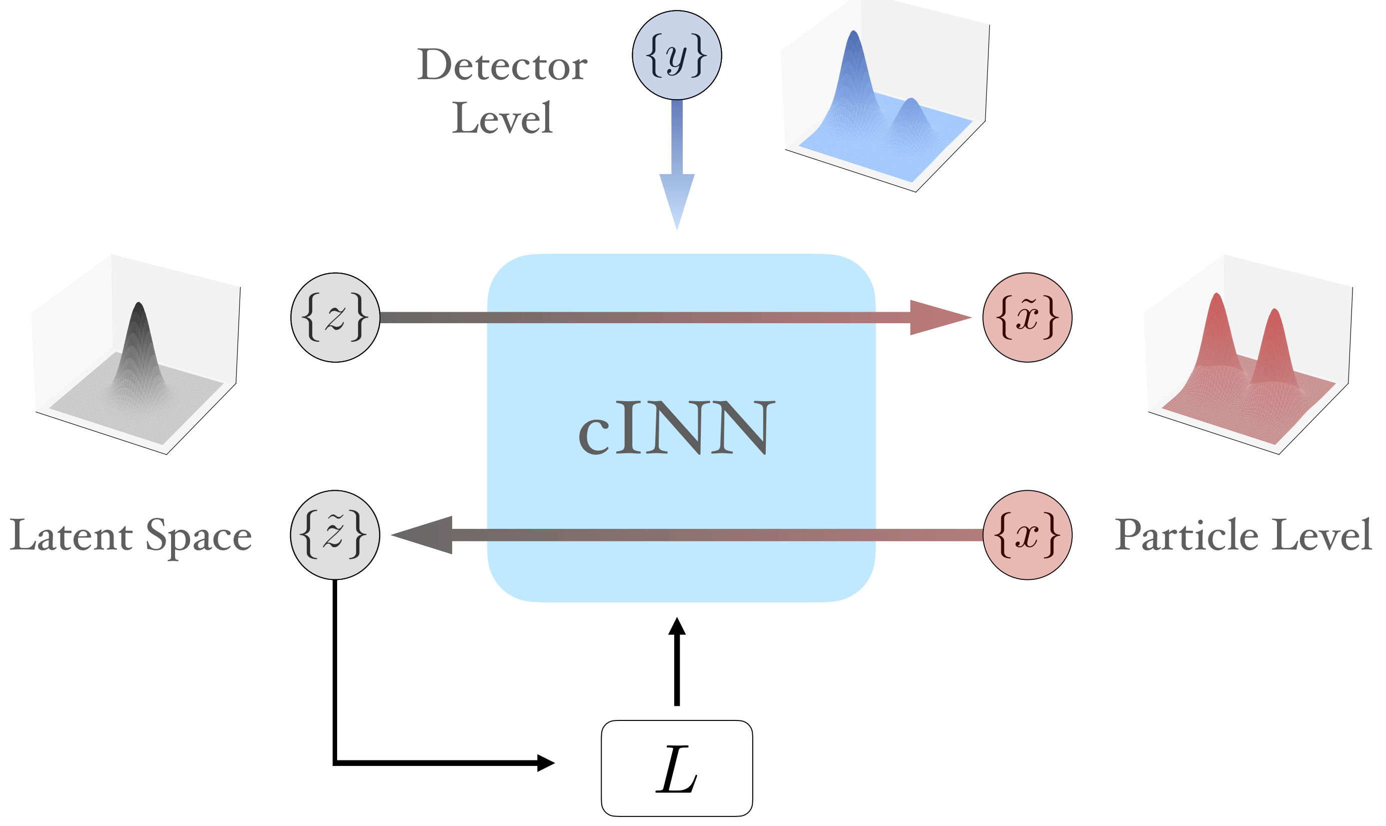


Adversarial

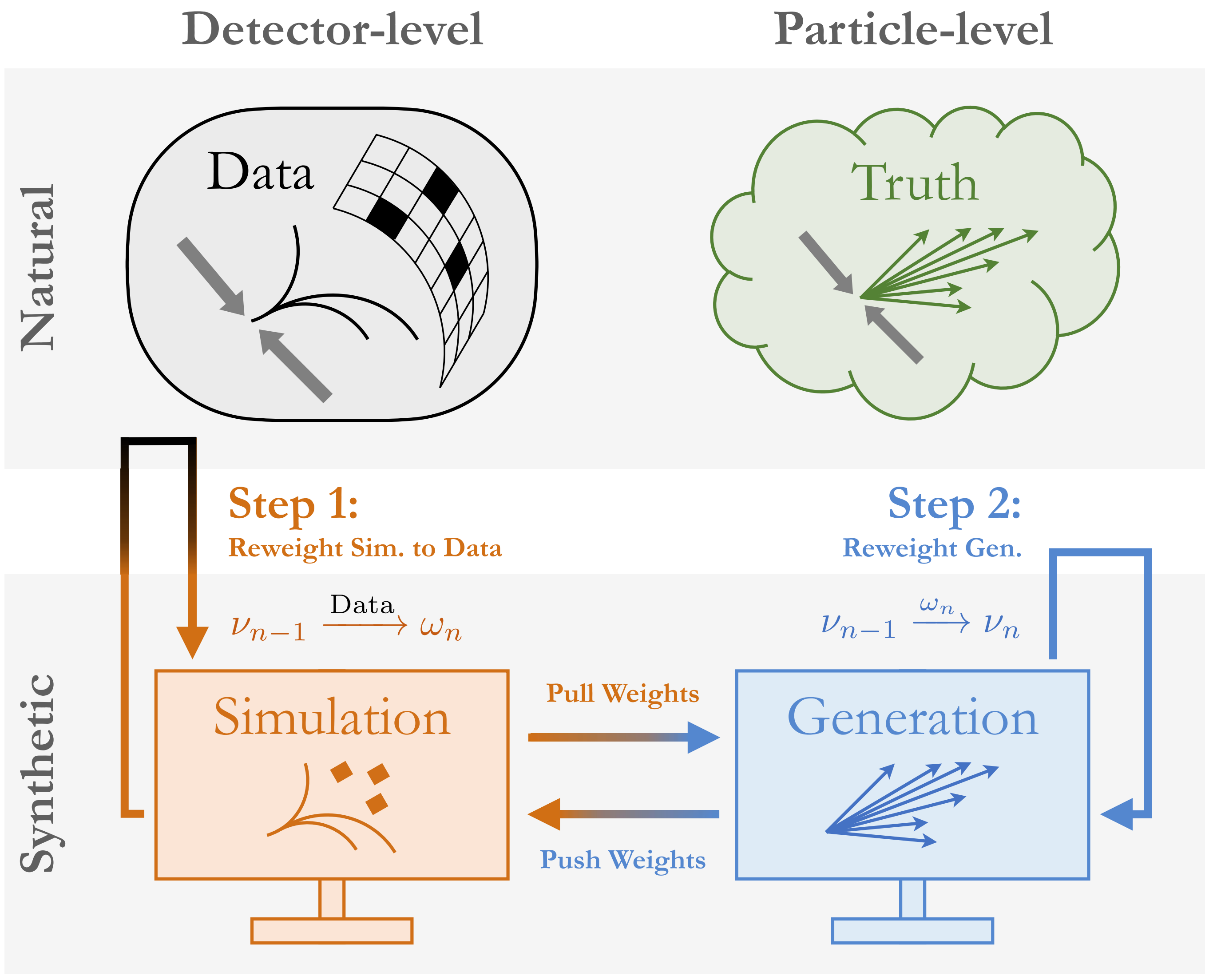
ML methods: Iterations, Adversaries, Surrogates

Iterative

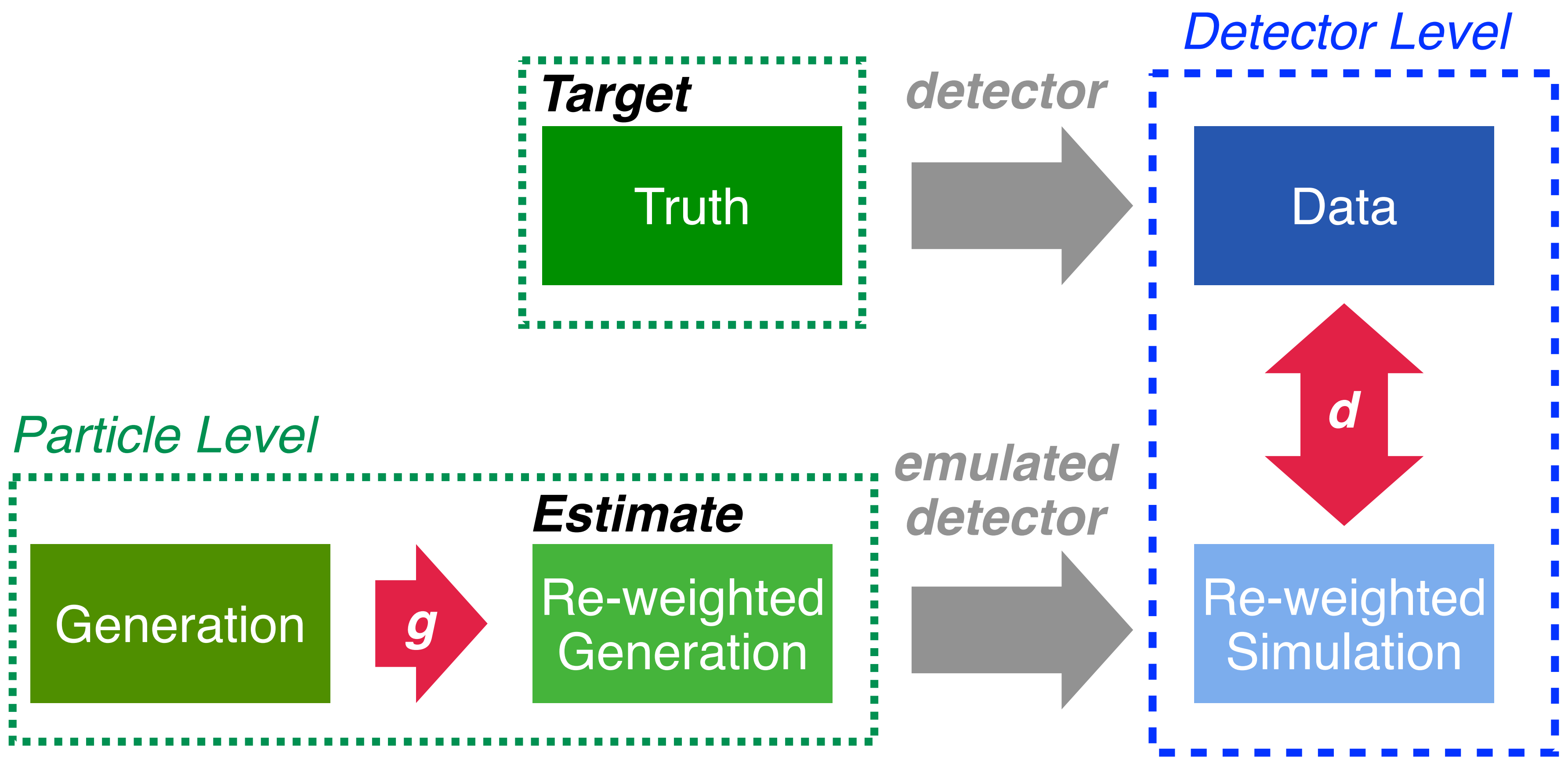
Backes et al. 2212.08674



Andreassen at al. 1911.09107

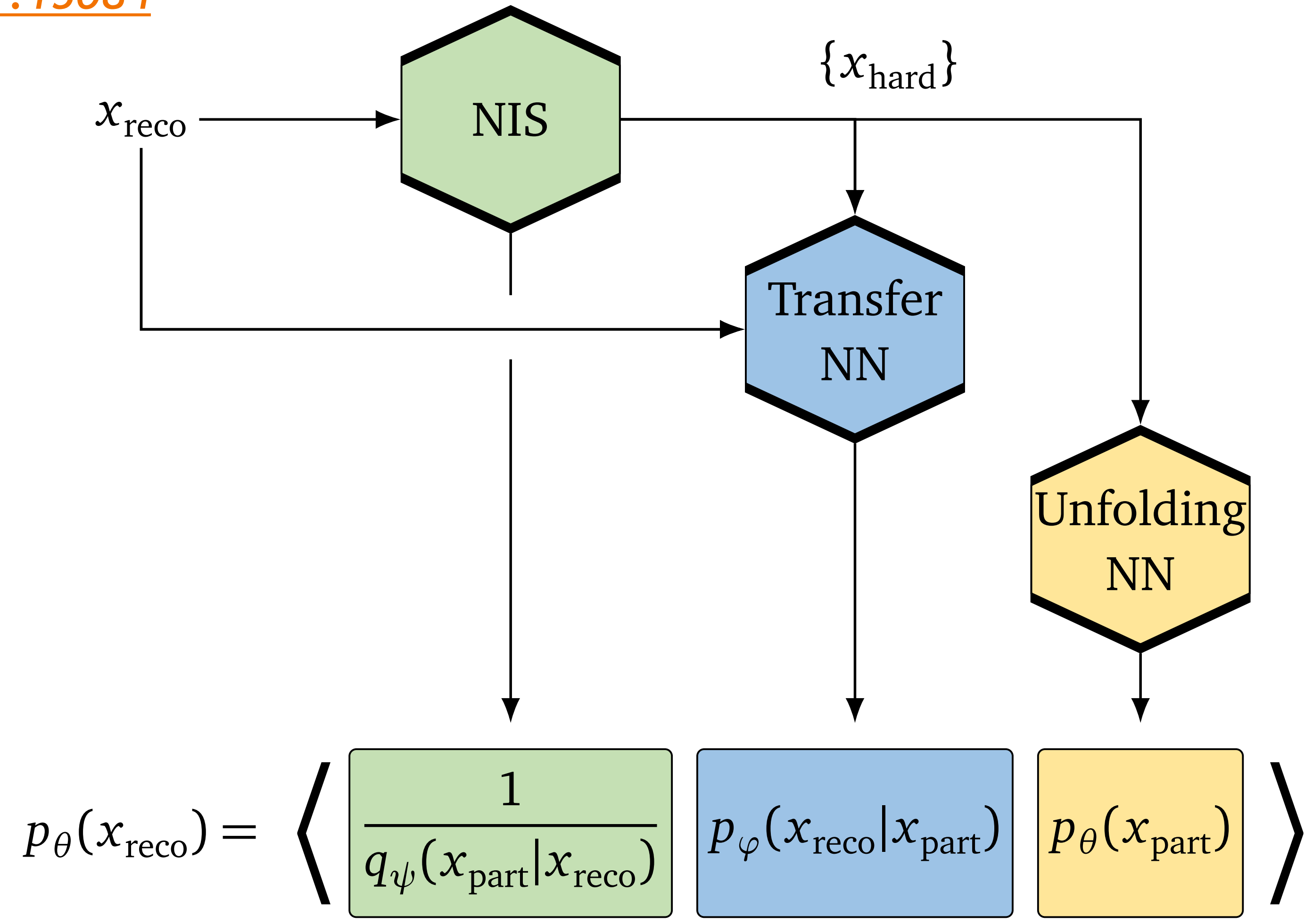


Desai et al. 2407.11284



Adversarial

Butter et al. 2507.15084



Surrogate-based

AUSSIE

- We introduce: **Adversary-free Unfolding SanS Iteration or Emulation (AUSSIE)**

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

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

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 $R(x)$

$$\underbrace{\frac{p_{\text{data}}(x)}{p_{\text{sim}}(x)}}_{R(x)} = \int dz\, p_{\text{sim}}(z|x)\, \underbrace{\frac{p_{\text{unfold}}(z)}{p_{\text{sim}}(z)}}_{\overline{R}(z)}$$

AUSSIE Step 2

- *First, understand **OmniFold**'s second step:*

$$\mathcal{L}_{\text{MSE}} = \left\langle \left(R(x) - \overline{R}_\varphi(z) \right)^2 \right\rangle_{p_{\text{sim}}(x,z)}$$

$$\Big\downarrow \; \frac{\delta \mathcal{L}_{\text{MSE}}}{\delta \overline{R}_\varphi(z)} = 0$$

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

! Inverted !

AUSSIE Step 2

- First, understand **OmniFold**'s second step:

$$\mathcal{L}_{\mathrm{MSE}} = \left\langle \left(R(x) - \overline{R}_\varphi(z) \right)^2 \right\rangle_{p_{\mathrm{sim}}(x,z)}$$

$$\frac{\delta \mathcal{L}_{\mathrm{MSE}}}{\delta \overline{R}_\varphi(z)} = 0$$

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

! Inverted !

- In **AUSSIE**, consider the mirrored problem:

$$\frac{\delta \mathcal{L}_{\mathrm{MSE}}}{\delta R(x)} = 0$$

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

Solution

AUSSIE Step 2

- First, understand **OmniFold**'s second step:

$$\mathcal{L}_{\text{MSE}} = \left\langle \left(R(x) - \overline{R}_\varphi(z) \right)^2 \right\rangle_{p_{\text{sim}}(x,z)}$$

$$\frac{\delta \mathcal{L}_{\text{MSE}}}{\delta \overline{R}_\varphi(z)} = 0$$

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

! Inverted !

- In **AUSSIE**, consider the mirrored problem:

$$\boxed{\frac{\delta \mathcal{L}_{\text{MSE}}}{\delta R(x)} = 0}$$

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

Solution

AUSSIE Step 2: Implementations

Target

$$\boxed{\frac{\delta \mathcal{L}_{\text{MSE}}}{\delta R(x)} = 0}$$

1. **Explicit:** Minimise a norm of $\dfrac{\delta \mathcal{L}_{\text{MSE}}}{\delta R(x)} = \left(R(x) - \int dz\, p_{\text{sim}}(z|x)\, \overline{R}(z) \right) p_{\text{sim}}(x)$

AUSSIE Step 2: Implementations

Target

$$\boxed{\frac{\delta\mathcal{L}_{\text{MSE}}}{\delta R(x)} = 0}$$

1. Explicit: Minimise a norm of $\frac{\delta\mathcal{L}_{\text{MSE}}}{\delta R(x)} = \left(R(x) - \int dz\, p_{\text{sim}}(z|x)\,\overline{R}(z)\right) p_{\text{sim}}(x)$

$$\mathcal{L}_{\text{AUSSIE}} = \left\| \frac{\delta\mathcal{L}_{\text{MSE}}}{\delta R(x)} \right\|_K^2$$

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

Positive-definite kernel

- *Linear kernel norm is tractable*
- *Fast*
- *Effective in **low dim** (~20)*

AUSSIE Step 2: Implementations

Target

$$\boxed{\frac{\delta \mathcal{L}_{\mathrm{MSE}}}{\delta R(x)} = 0}$$

1. Explicit: Minimise a norm of $\frac{\delta \mathcal{L}_{\mathrm{MSE}}}{\delta R(x)} = \left(R(x) - \int dz\, p_{\mathrm{sim}}(z|x)\, \bar{R}(z) \right) p_{\mathrm{sim}}(x)$

$$\mathcal{L}_{\mathrm{AUSSIE}} = \left\| \frac{\delta \mathcal{L}_{\mathrm{MSE}}}{\delta R(x)} \right\|_K^2$$

$$= \left\langle \left(R(x) - \bar{R}_\varphi(z) \right) \underbrace{K(x,x')}_{\text{Positive-definite kernel}} \left(R(x') - \bar{R}_\varphi(z') \right) \right\rangle_{p_{\mathrm{sim}}(x,z) p_{\mathrm{sim}}(x',z')}$$

- *Linear kernel norm is tractable*
- *Fast*
- *Effective in **low dim** (~20)*

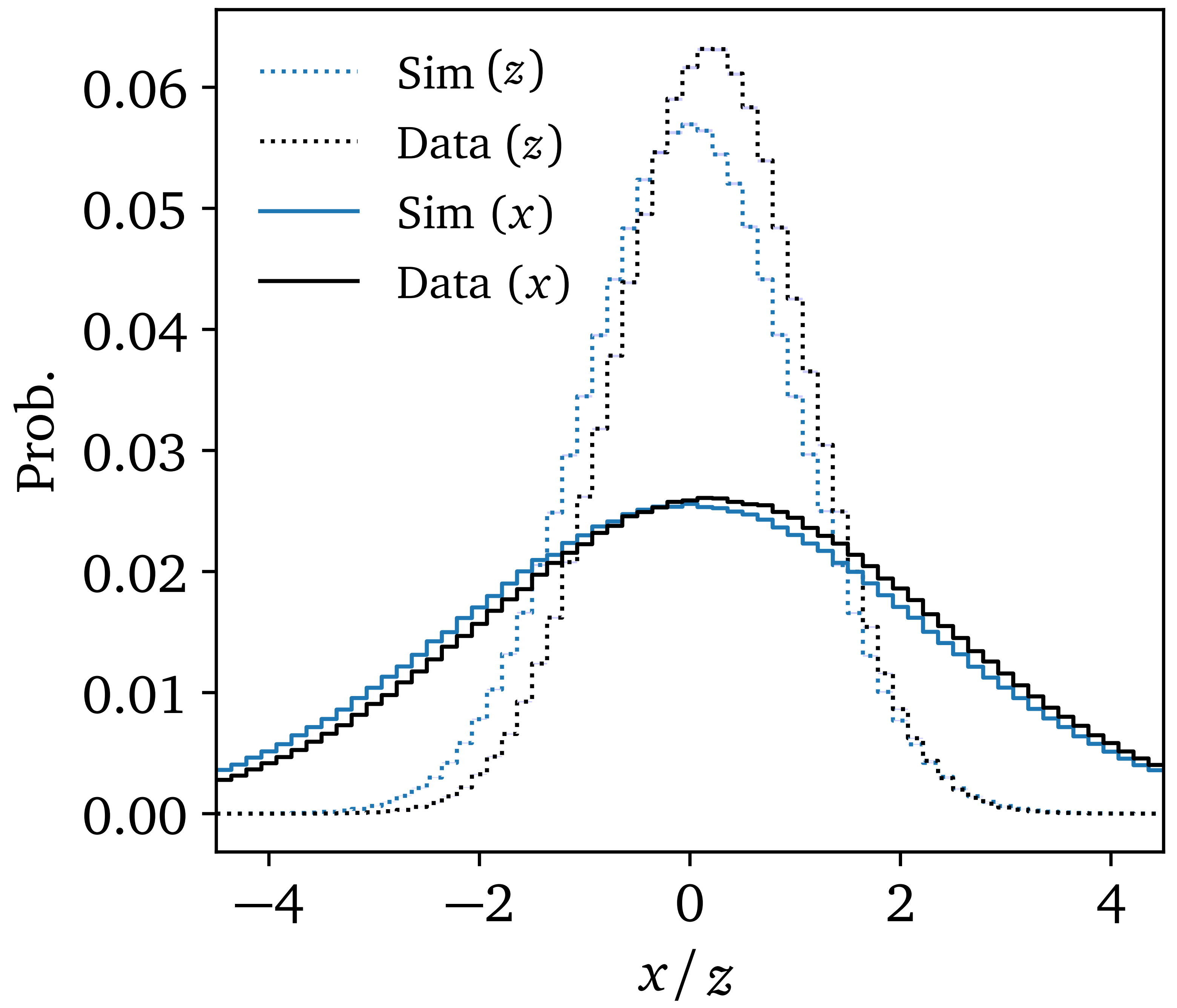
2. Implicit: Minimise parameter gradient norm of $R_\theta(x)$

$$\mathcal{L}_{\mathrm{AUSSIE}} = \left\| \nabla_\theta \mathcal{L}_{\mathrm{MSE}} \right\|^2$$

- *Leverages sensitivity and symmetries of $R_\theta(x)$*
- *Applies to **high-dim** point clouds*
- *Equivalent to **1.** choosing K as Neural Tangent Kernel (NTK)*

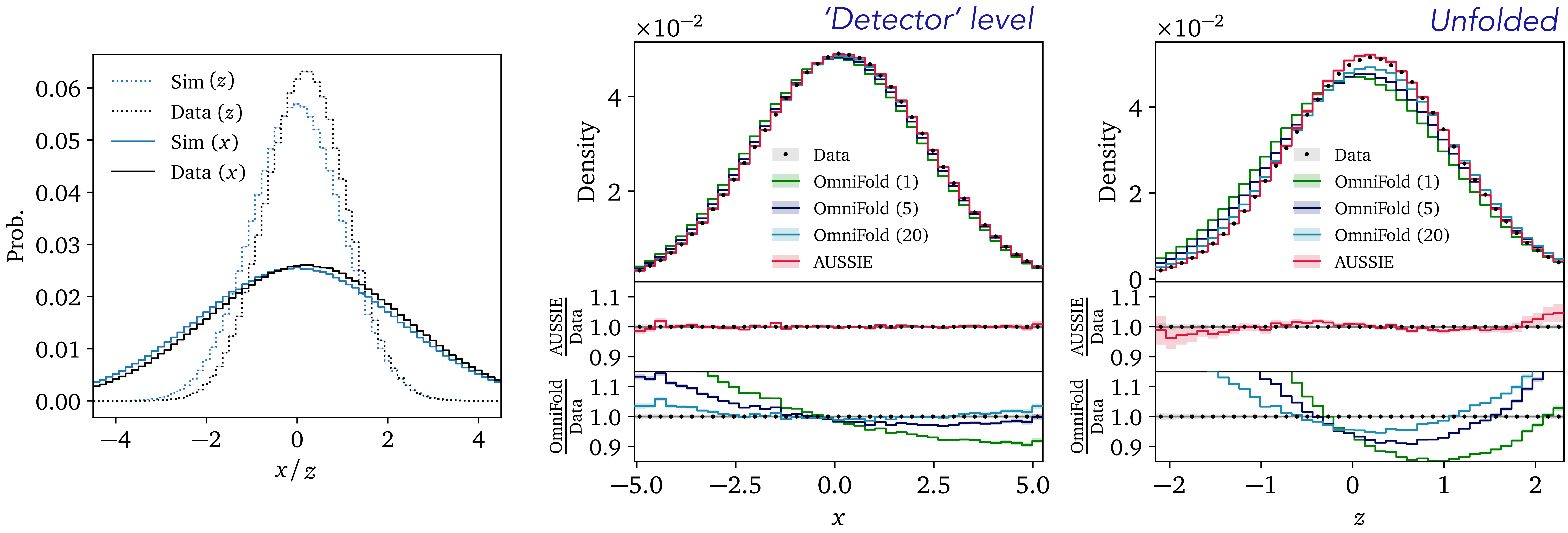
Gaussian example

- Toy model with large smearing exemplifies advantage of AUSSIE



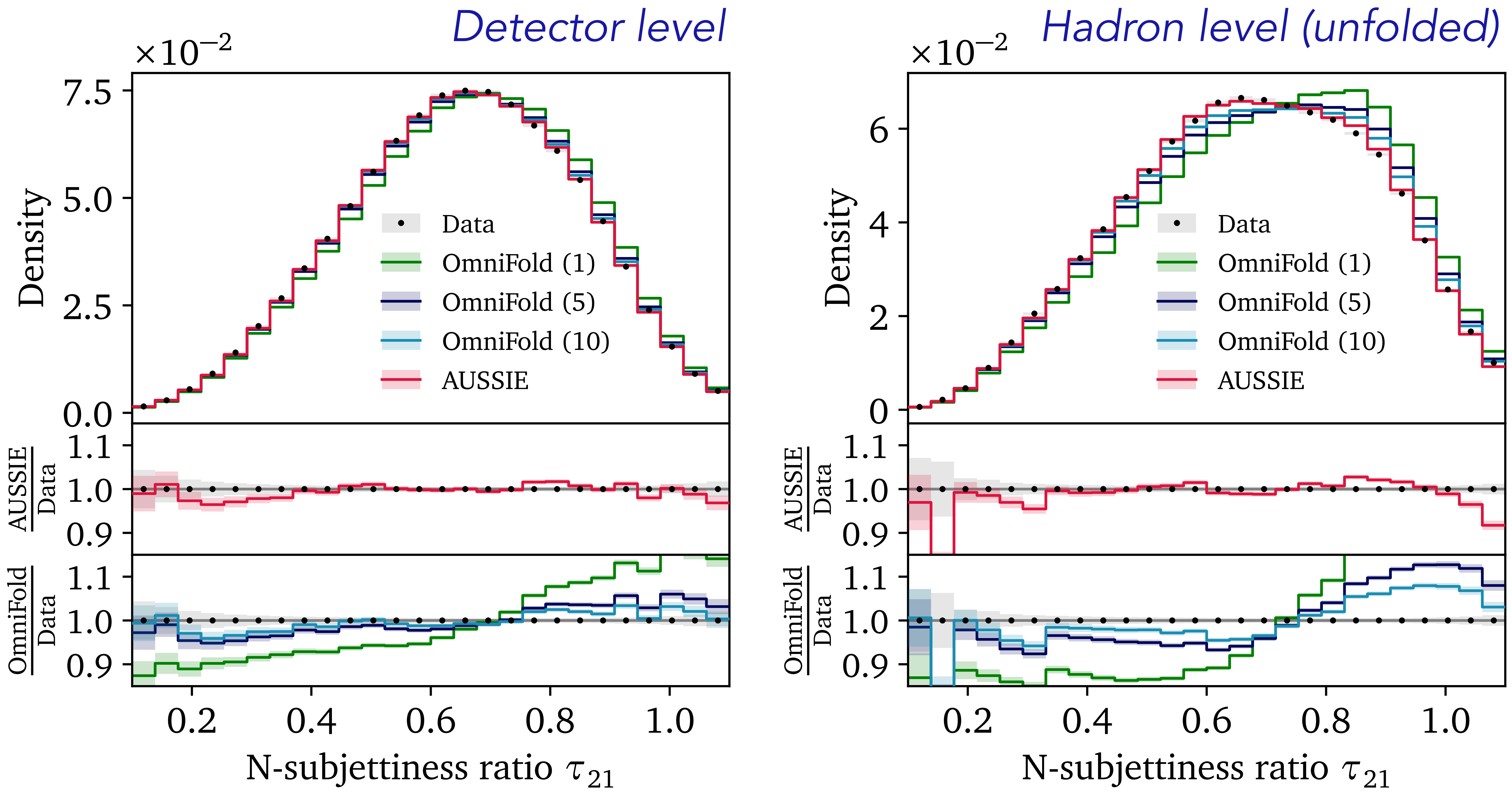
Gaussian example

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

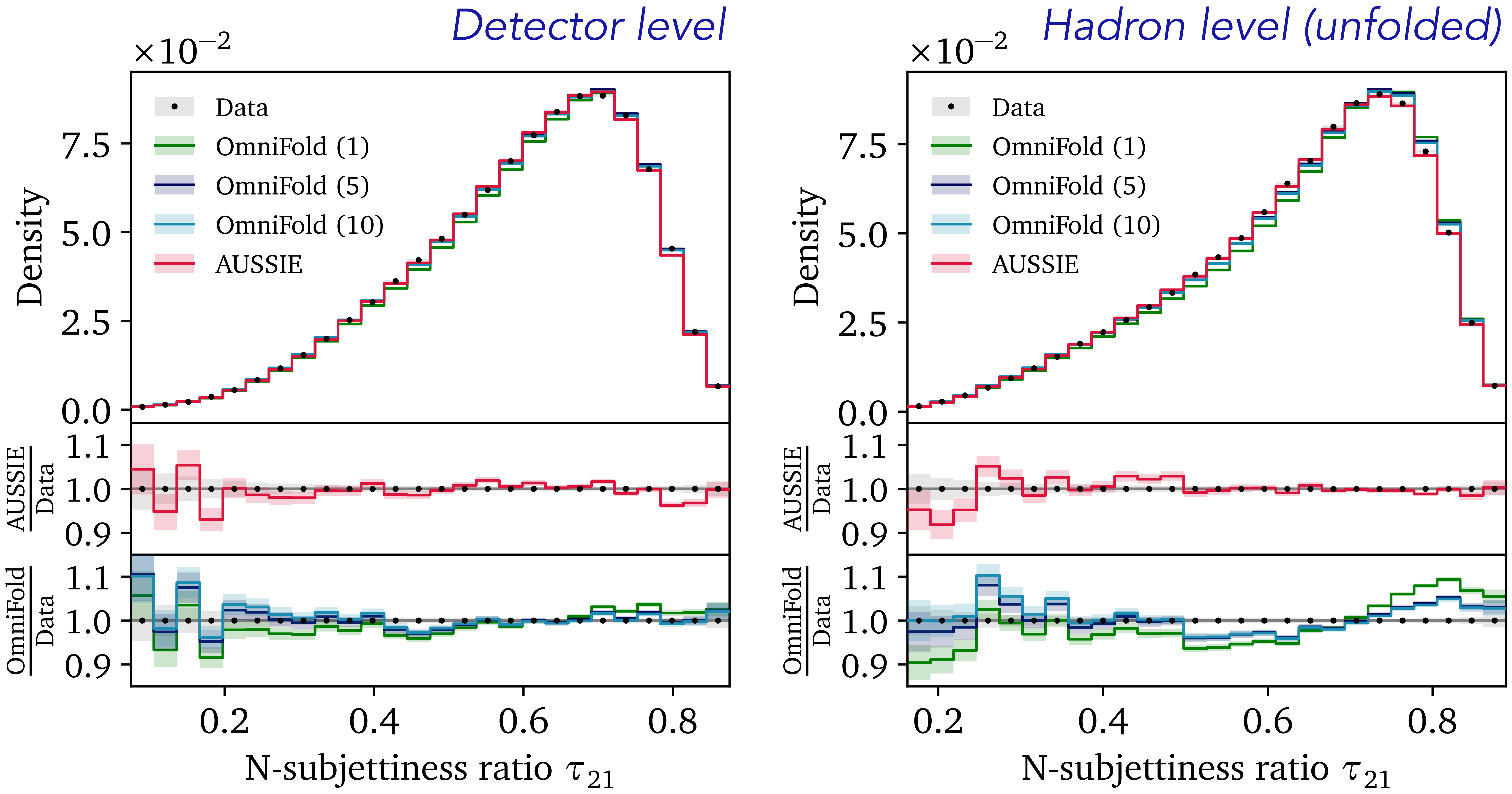


Jet substructure (high-level)

- ***Classic benchmark dataset:*** *6D unfolding problem*
- $R_\theta(x)$ *and* $\overline{R}_\varphi(z)$ *small MLPs*
- *Unit Gaussian kernel after preprocessing (no tuning)*
- *AUSSIE matches while OmniFold struggles*



Jet substructure (constituent-level)



- ***Full phase space unfolding:*** *particle cloud up to N=150*
- *L-GATr* $R_\theta(x)$ *and* $\overline{R}_\varphi(z)$
- *Adaptive kernel from* ***Neural Tangent Kernel***
- *AUSSIE mitigates residual bias toward simulated dataset*

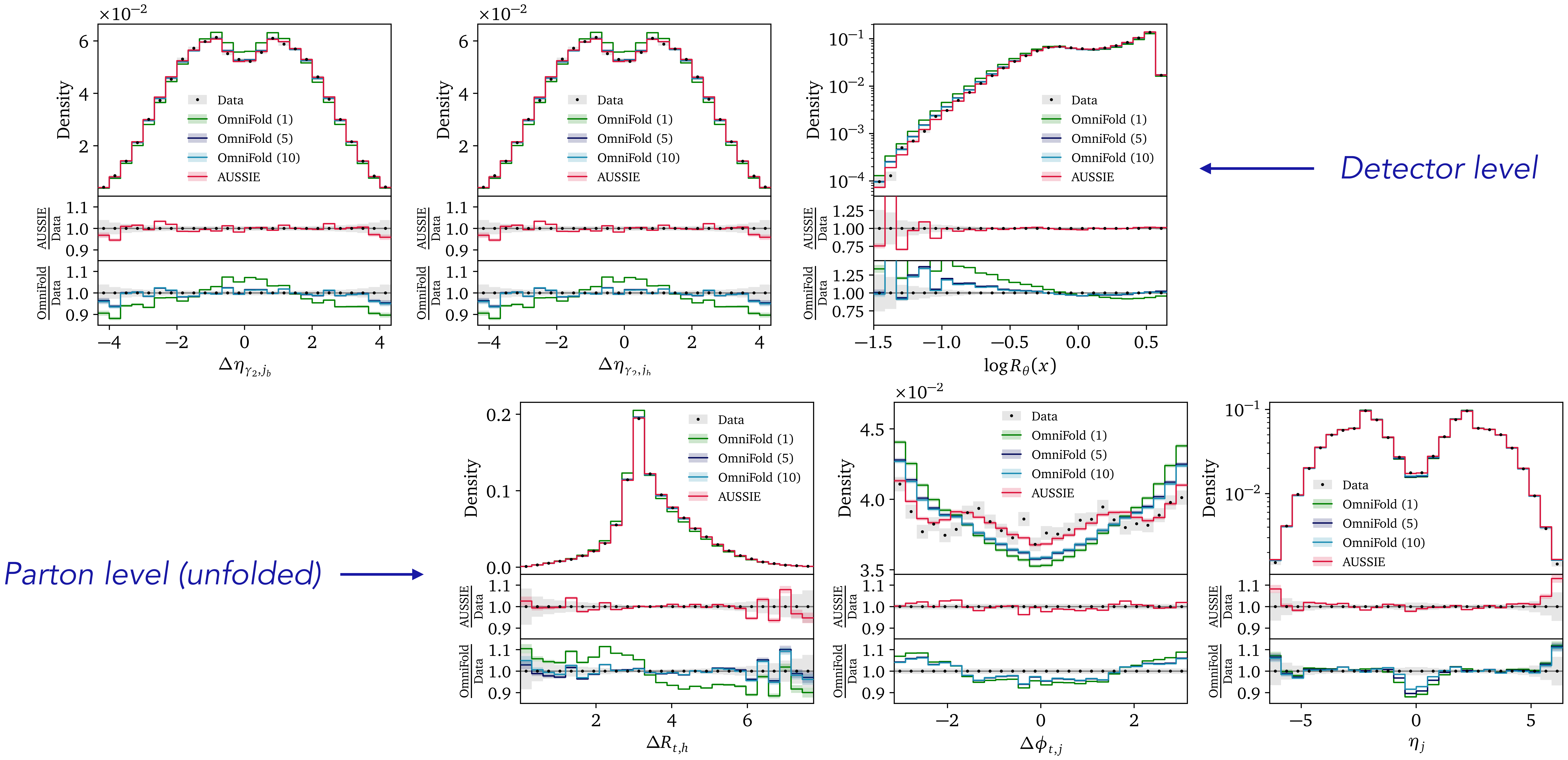
Summary

- ***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*

Read more:
Paper *arXiv:2602.24282*
Code *github.com/heidelberg-hepml/aussie*

Bonus: Parton-level unfolding



Bonus: Some detail

Regression loss

$$\mathcal{L}_{\text{MSE}}[\bar{R}_\varphi] = \left\langle \left(R(x) - \bar{R}_\varphi(z)\right)^2 \right\rangle_{p_{\text{sim}}(x,z)} \quad = \int dx\, dz\, p_{\text{sim}}(x,z) \left(R(x) - \bar{R}_\varphi(z)\right)^2$$

Functional derivative

$$\frac{\delta \mathcal{L}_{\text{MSE}}[R, \bar{R}_\varphi]}{\delta R(x)} = \left(R(x) - \int dz\; p_{\text{sim}}(z|x)\, \bar{R}_\varphi(z)\right) p_{\text{sim}}(x)$$

AUSSIE loss

$$\mathcal{L}_{\text{AUSSIE}}[\bar{R}_\varphi] = \left\| \frac{\delta \mathcal{L}_{\text{MSE}}[R, \bar{R}_\varphi]}{\delta R(x)} \right\|_K^2 \equiv \int dx\, dx' \left(\frac{\delta \mathcal{L}_{\text{MSE}}[R, \bar{R}_\varphi]}{\delta R(x)} \right) K(x, x') \left(\frac{\delta \mathcal{L}_{\text{MSE}}[R, \bar{R}_\varphi]}{\delta R(x')} \right)$$

$$= \left\langle \left(R(x) - \bar{R}_\varphi(z)\right) K(x, x') \left(R(x') - \bar{R}_\varphi(z')\right) \right\rangle_{p_{\text{sim}}(x,z)\, p_{\text{sim}}(x',z')}$$