FASTColor

Full-color **A**mplitude **S**urrogate **T**oolkit for QCD

[arXiv: **2509.07068**] Javier Mariño Villadamigo, Rikkert Frederix, Tilman Plehn, Timea Vitos, and Ramon Winterhalder

ML4Jets @ Vienna, September 16th 2026

SPONSORED BY THE

Federal Ministry of Education and Research

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

Institut für Theoretische Physik - University of Heidelberg

Setting the stage

Read the paper!

What we do

Use Machine-Learning to **accelerate event generation at Full-Color (FC) accuracy** in QCD

The result

Average **x1.5 speed-up** across different processes and multiplicities

How we do it

We **build on Leading-Color (LC) based, two-step unweighting** approach [2409.12128]

The Leading-Color approximation

Read the paper!

1. Introduce the $\boldsymbol{h}$**elicity-dependent** squared matrix elements at FC accuracy and LC accuracy

$$|M(x,h)|^2 = \sum_{i,j} A_{i,h}(x)\, C_{ij}\, A^*_{j,h}(x) \qquad |M_{\mathrm{LC}}(x,h)|^2 = \sum_{i} C_{ii}\, |A_{i,h}(x)|^2$$

The Leading-Color approximation

Read the paper!

1. Introduce the ***h*elicity-dependent** squared matrix elements at FC accuracy and LC accuracy

$$|M(x,h)|^2 = \sum_{i,j} A_{i,h}(x)\, C_{ij}\, A^*_{j,h}(x) \qquad |M_{\mathrm{LC}}(x,h)|^2 = \sum_{i} C_{ii}\,|\,A_{i,h}(x)\,|^2$$

2. We also define the **color-dependent** (h-summed) LC matrix element and its relative h-contributions

$$|M^i_{\mathrm{LC}}(x)|^2 = \sum_{h} C_{ii}\,|\,A_{i,h}(x)\,|^2 \qquad p_i(h\,|\,x) = \frac{C_{ii}\,|\,A_{i,h}(x)\,|^2}{|M^i_{\mathrm{LC}}(x)|^2}$$

The Leading-Color approximation

Read the paper!

1. Introduce the $\boldsymbol{h}$**elicity-dependent** squared matrix elements at FC accuracy and LC accuracy

$$|M(x,h)|^2 = \sum_{i,j} A_{i,h}(x)\, C_{ij}\, A^*_{j,h}(x) \qquad\qquad |M_{\rm LC}(x,h)|^2 = \sum_i C_{ii}\, |\, A_{i,h}(x)\,|^2$$

2. We also define the **color-dependent** (h-summed) LC matrix element and its relative h-contributions

$$|M^i_{\rm LC}(x)|^2 = \sum_h C_{ii}\, |\, A_{i,h}(x)\,|^2 \qquad\qquad p_i(h\,|\,x) = \frac{C_{ii}\,|\,A_{i,h}(x)\,|^2}{|M^i_{\rm LC}(x)|^2}$$

3. The FC cross section can thus be approximated by:

$$\sigma_{\rm FC} \approx \sum_i \int {\rm d}\Phi\, |M^i_{\rm LC}(x)|^2 \left\langle \frac{|M(x,h)|^2}{|M_{\rm LC}(x,h)|^2} \right\rangle_{h\sim p_i(h|x)}$$

The Leading-Color approximation

Read the paper!

1. Introduce the ***h*elicity-dependent** squared matrix elements at FC accuracy and LC accuracy

$$|M(x,h)|^2 = \sum_{i,j} A_{i,h}(x)\, C_{ij}\, A^*_{j,h}(x) \qquad\qquad |M_{\mathrm{LC}}(x,h)|^2 = \sum_i C_{ii}\, |A_{i,h}(x)|^2$$

2. We also define the **color-dependent** (h-summed) LC matrix element and its relative h-contributions

$$|M^i_{\mathrm{LC}}(x)|^2 = \sum_h C_{ii}\, |A_{i,h}(x)|^2 \qquad\qquad p_i(h\,|\,x) = \frac{C_{ii}\, |A_{i,h}(x)|^2}{|M^i_{\mathrm{LC}}(x)|^2}$$

3. The FC cross section can thus be approximated by:

$$\sigma_{\mathrm{FC}} \approx \sum_i \int \mathrm{d}\Phi\, |M^i_{\mathrm{LC}}(x)|^2 \left\langle \boxed{\frac{|M(x,h)|^2}{|M_{\mathrm{LC}}(x,h)|^2}} \right\rangle_{h\sim p_i(h|x)}$$

$$r_{\mathrm{LC}\to\mathrm{FC}}(x,h)$$

LC-based unweighting

Read the paper!

generate PS point x

LC-based unweighting

Read the paper!

generate PS point x

↓

calculate LC event weight

$w^i_{\mathrm{LC}} = |M^i_{\mathrm{LC}}(x)|^2$

LC-based unweighting

Read the paper!

generate PS point x

$\downarrow$

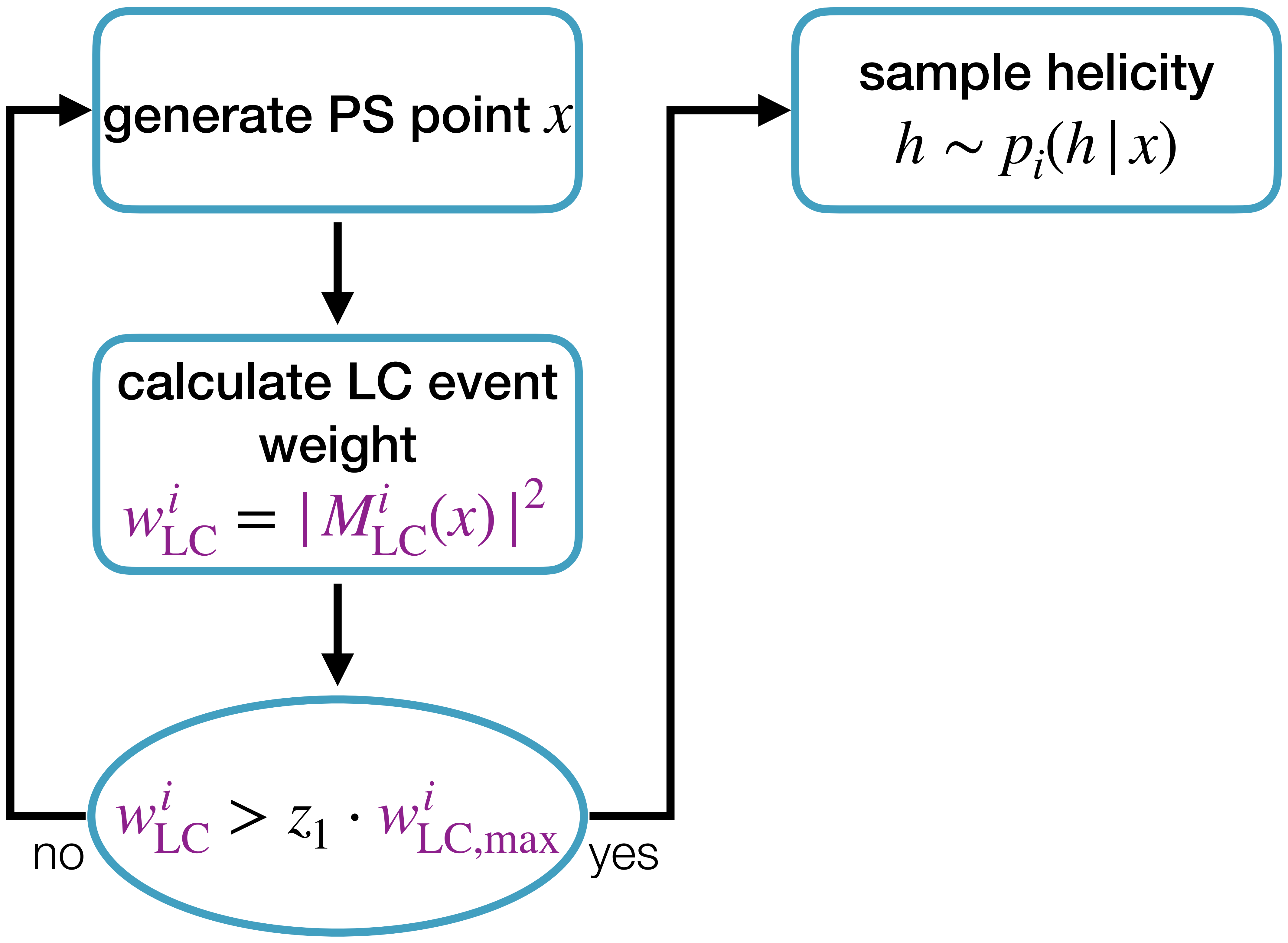
calculate LC event weight
$w^i_{\mathrm{LC}} = |M^i_{\mathrm{LC}}(x)|^2$

$\downarrow$

$w^i_{\mathrm{LC}} > z_1 \cdot w^i_{\mathrm{LC,max}}$

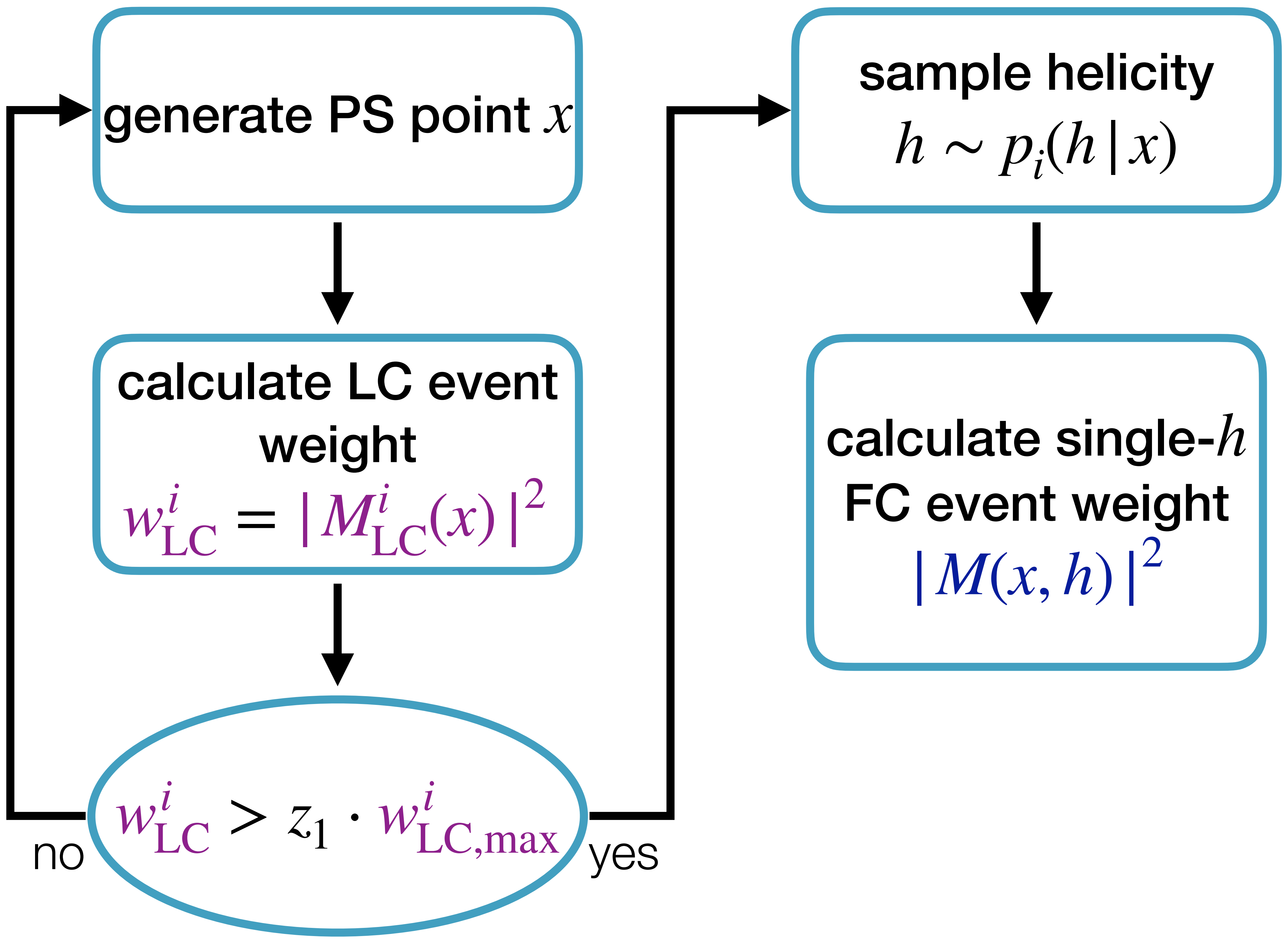
LC-based unweighting

Read the paper!

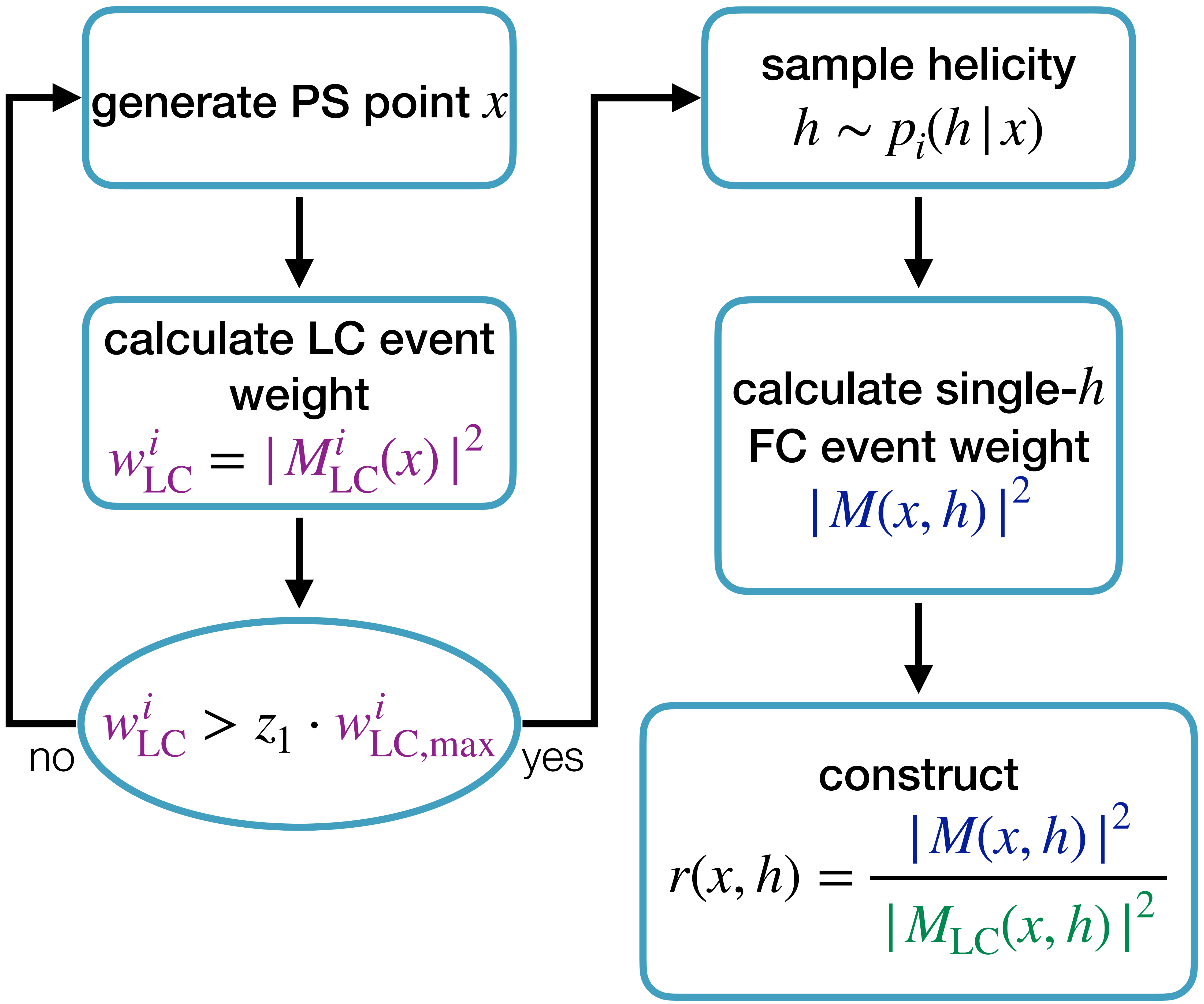


LC-based unweighting

Read the paper!

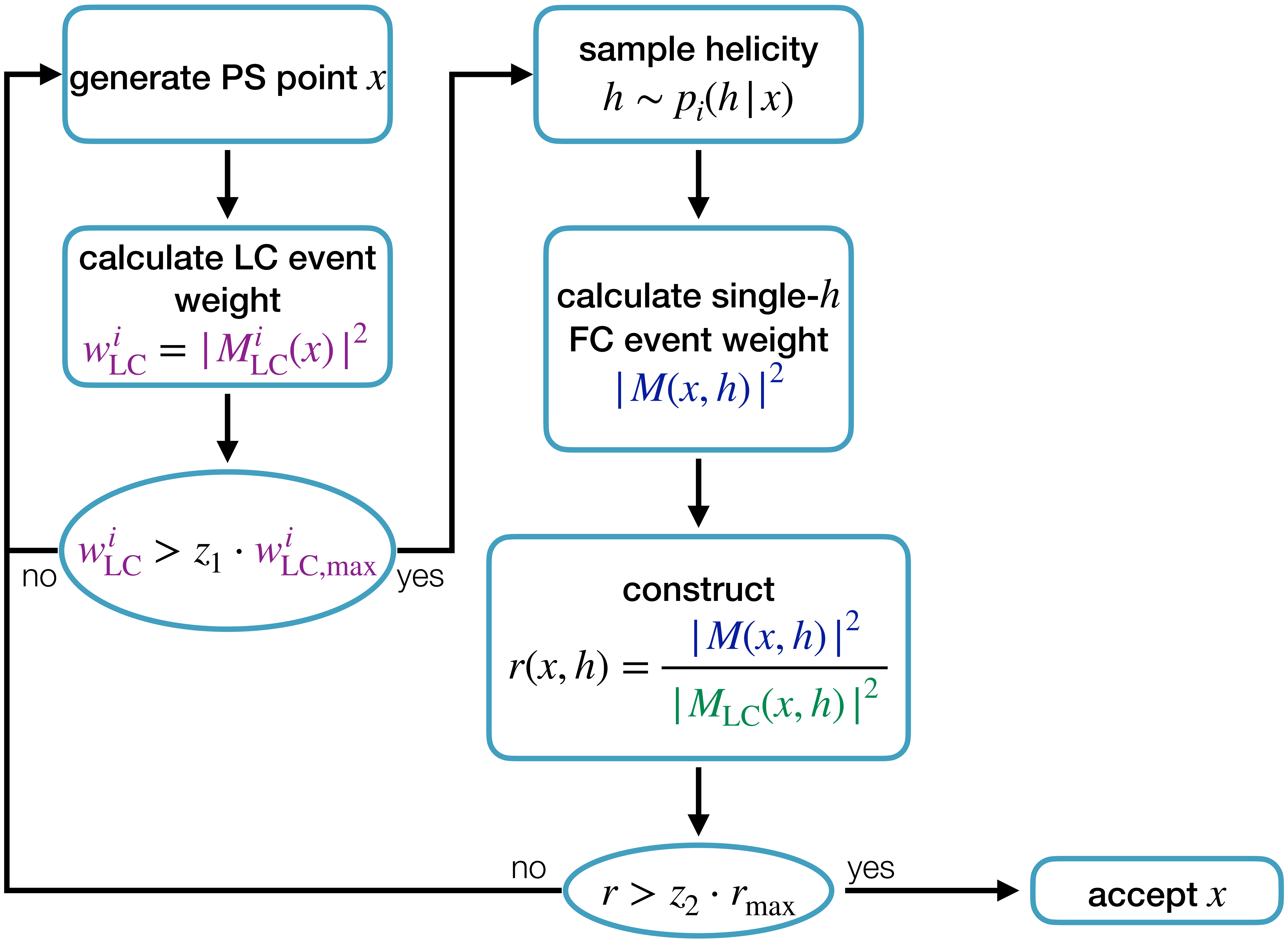


LC-based unweighting



Read the paper!

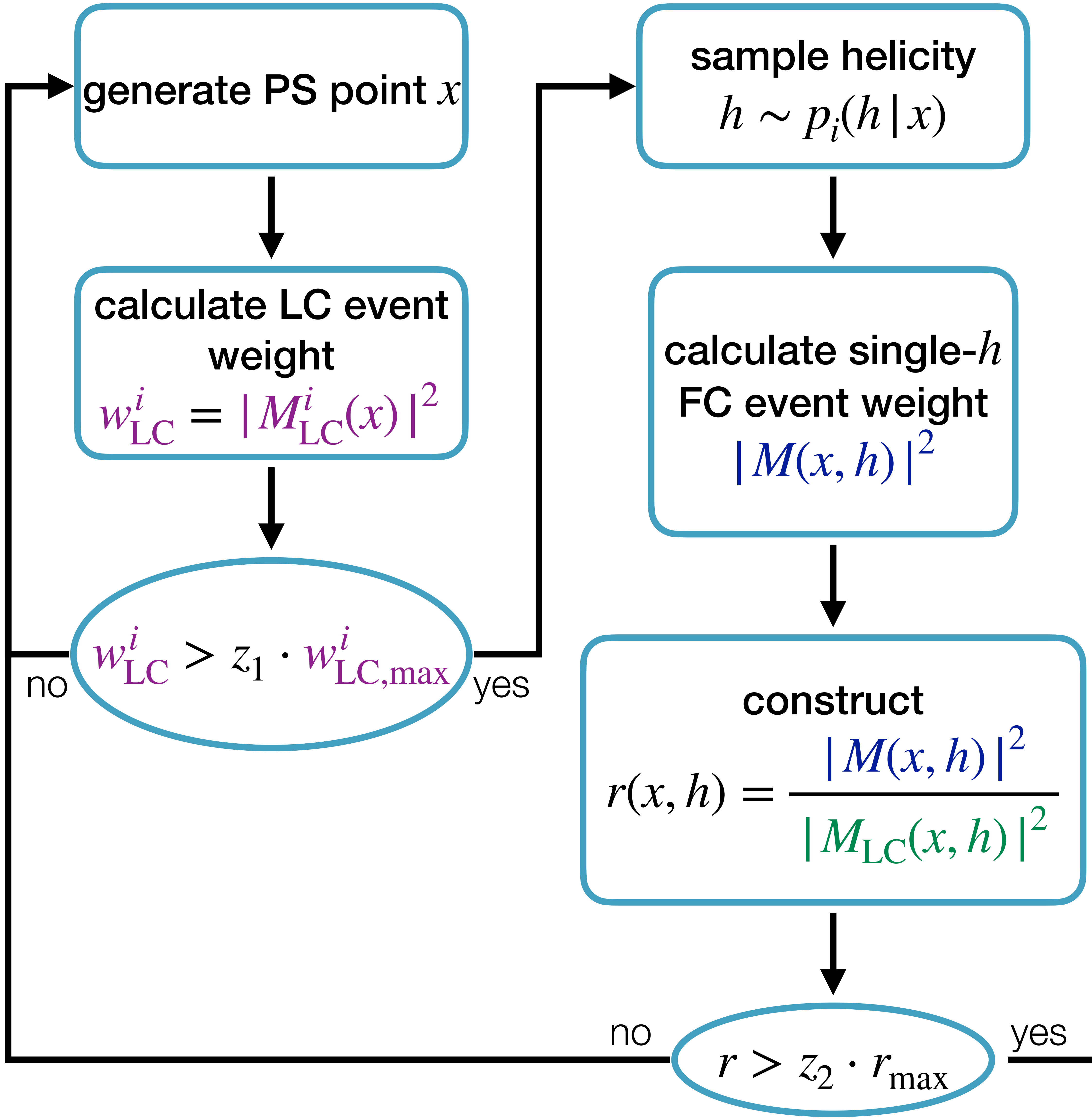
LC-based unweighting



Read the paper!

LC-based unweighting

Read the paper!



1. **generate LC events** for a single color flow i with weight w^i_{LC}
2. **assign a helicity h** with probability $p_i(h|x)$
3. **unweight to FC**

Only a single helicity sampled per event
but averaging for many PS points gives correct σ_{FC}

Surrogate-based unweighting

Read the paper!

Introduce **additional unweighting step** against ML surrogate

$$r_{\text{surr}}(x, h) \approx r(x, h)$$

Surrogate-based unweighting

Read the paper!

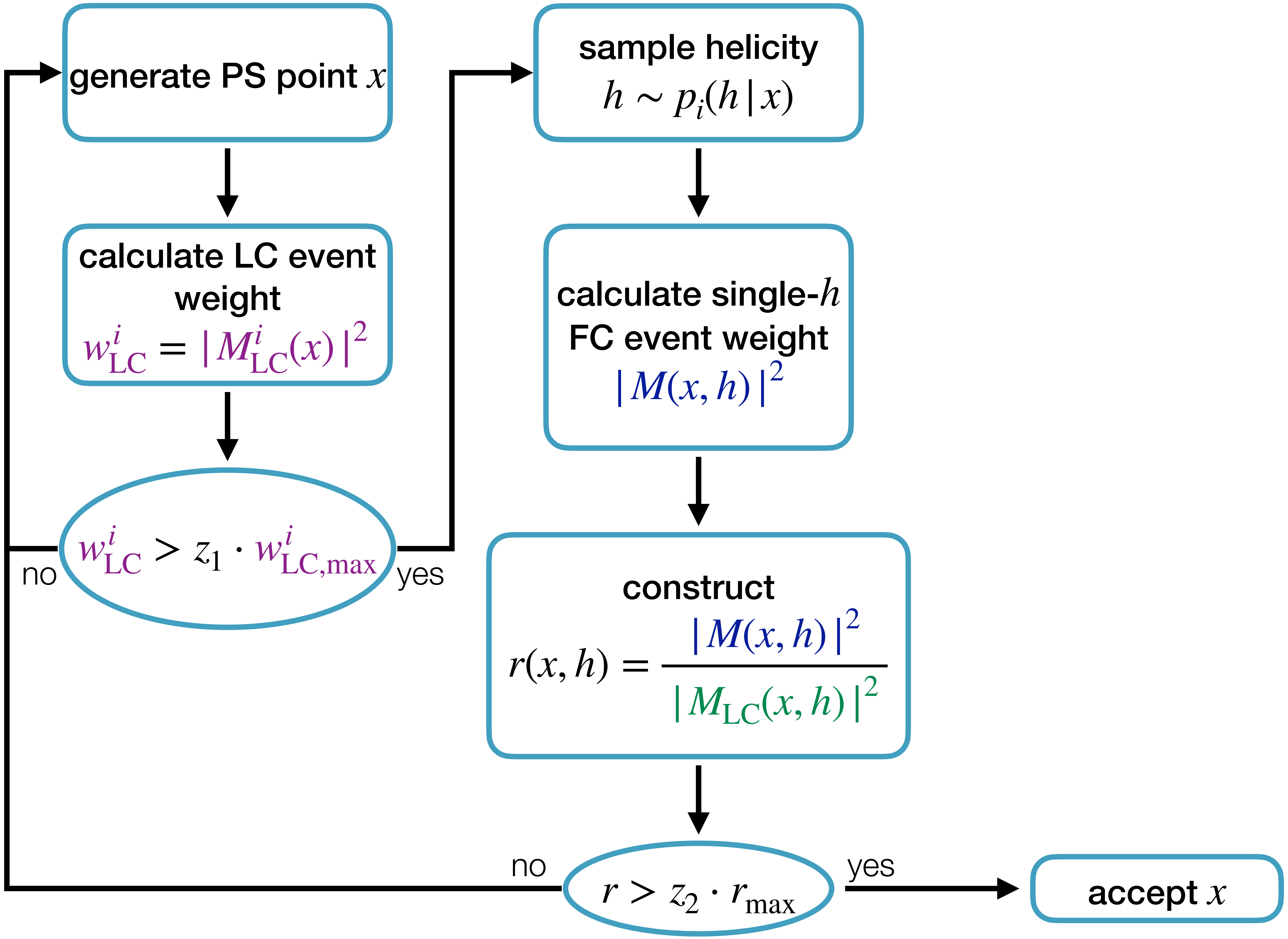
Introduce **additional unweighting step** against ML surrogate

$$r_{\text{surr}}(x,h) \approx r(x,h)$$

- **Faster to evaluate** than true $r_{\text{LC}\to\text{FC}}(x,h)$: e.g. $\mathcal{O}_{\text{FC}}(1)$ vs $\mathcal{O}_{\text{surr}}(10^{-5})$ s/event
- Much **better scaling** with particle multiplicity than $|M(x,h)|^2$ (and $|M_{\text{LC}}(x,h)|^2$)

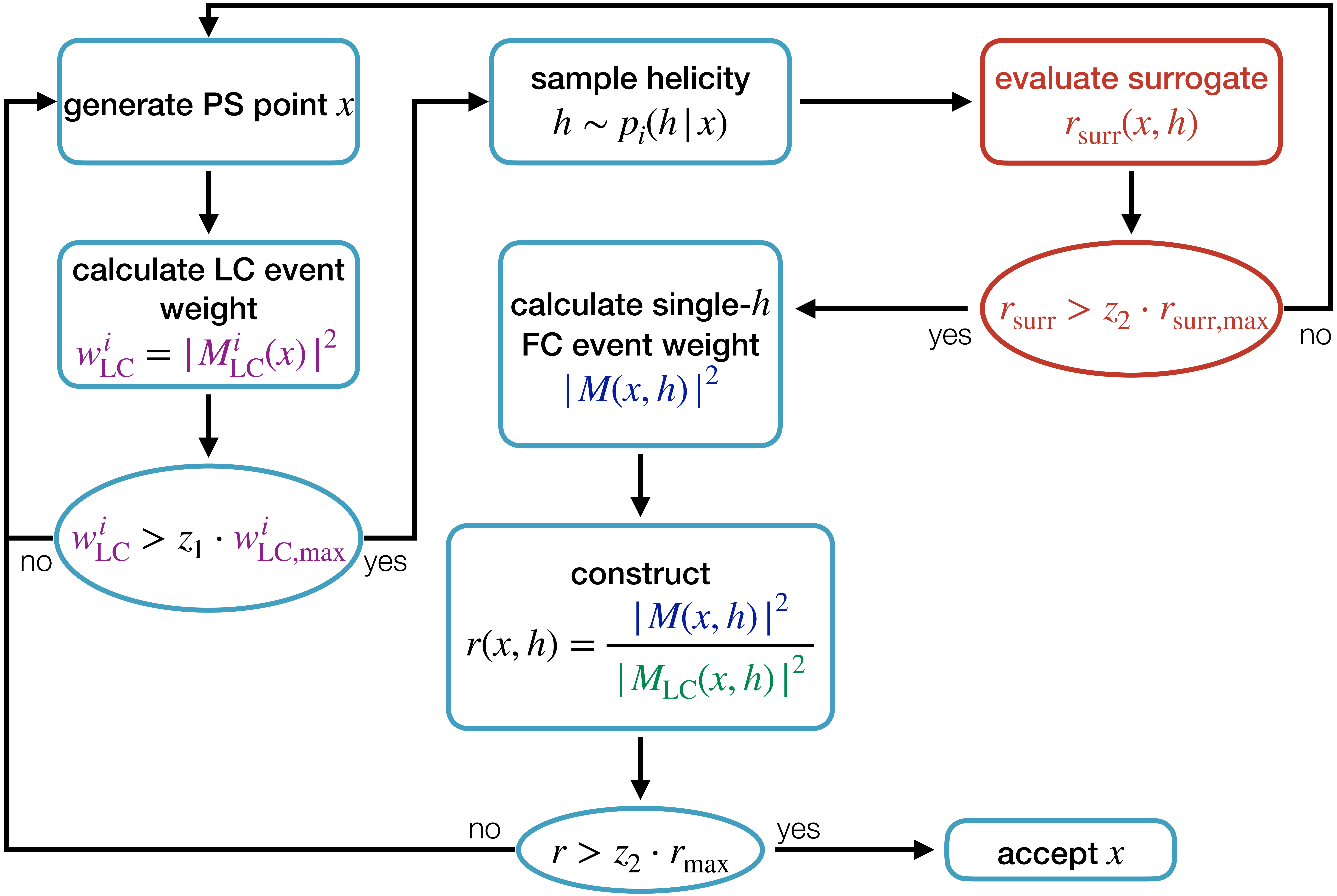
Surrogate-based unweighting

Read the paper!



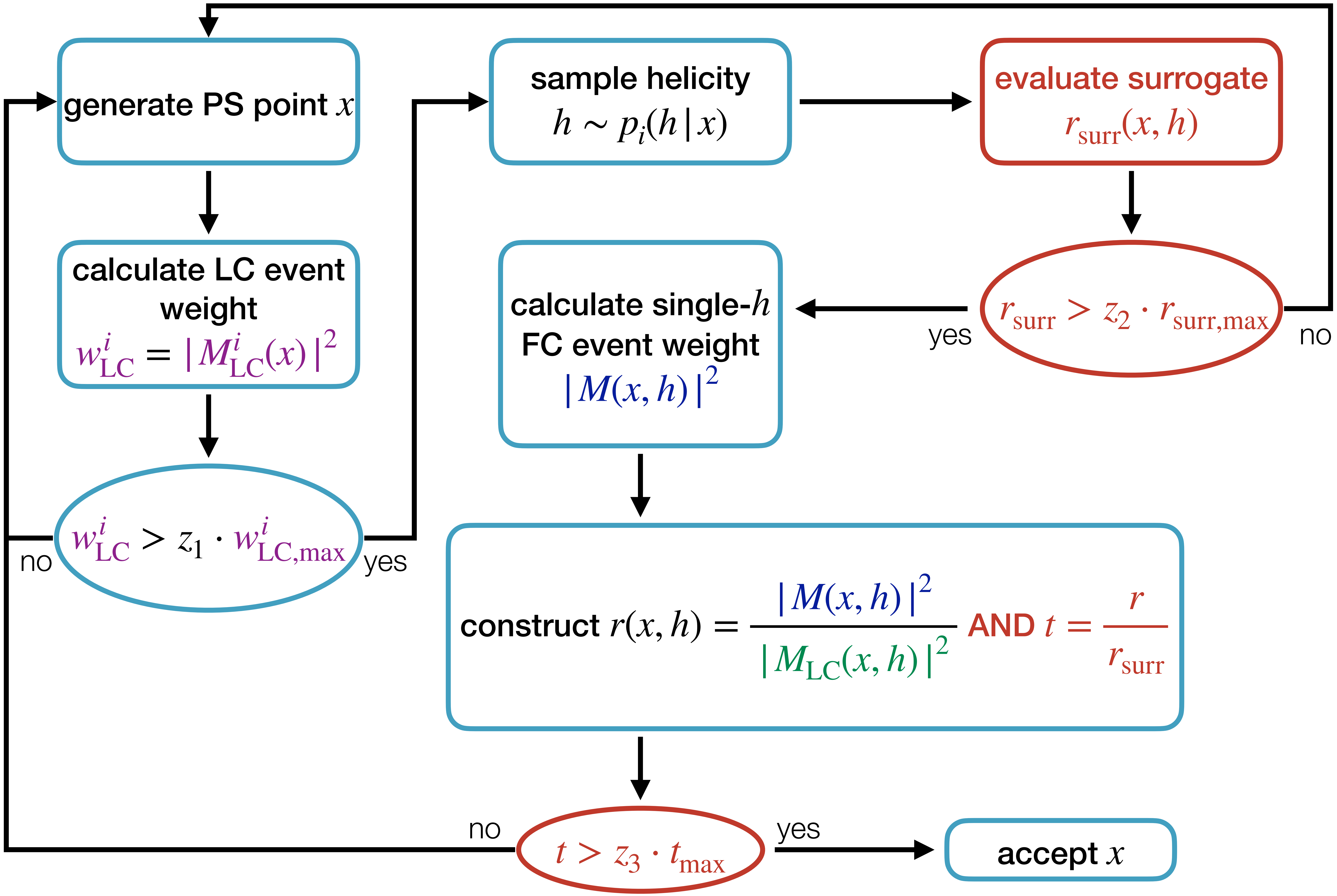
Surrogate-based unweighting

Read the paper!



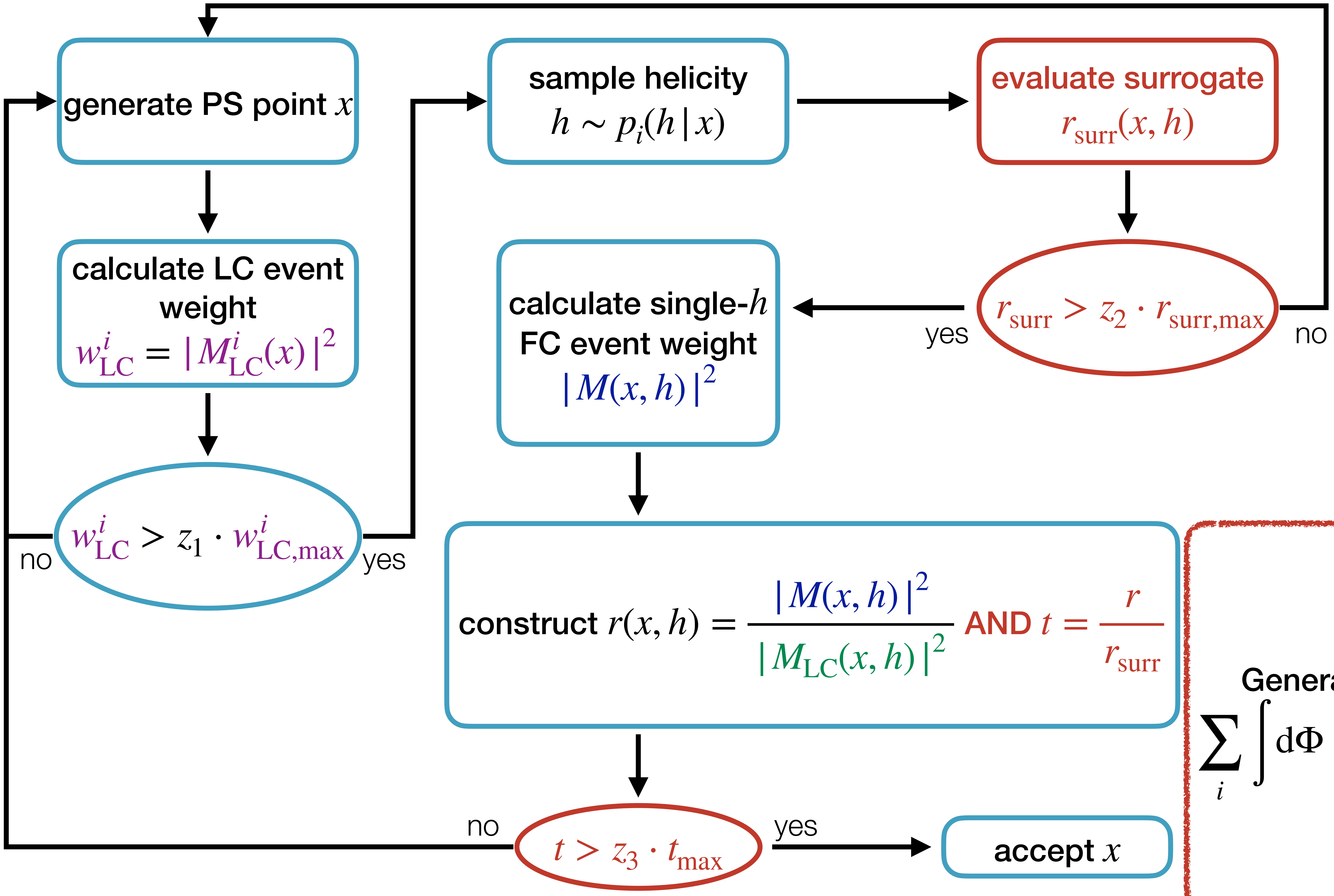
Surrogate-based unweighting

Read the paper!



Surrogate-based unweighting

Read the paper!



Also unbiased!

Generated events follow FC density

$$\sum_i \int \mathrm{d}\Phi \, |M^i_{\mathrm{LC}}(x)|^2 \left\langle r_{\mathrm{surr}}(x,h) \cdot t \right\rangle_{h \sim p_i(h|x)}$$

$$\approx \sigma_{\mathrm{FC}}$$

Surrogate-based unweighting

Read the paper!

How to quantify

gains in performance?

Surrogate-based unweighting

Read the paper!

How to quantify gains in performance?

Effective gain factor:

$$f^{\text{eff}} \equiv \frac{T_{\text{LC}}}{T_{\text{surr}}}$$

- T takes into account evaluation time, efficiency, and statistical power of the generated sample

$$T^{(M-\text{step})} = N \sum_{i=1}^{M} \frac{\langle t_i \rangle}{\prod_{j=i}^{M} \epsilon_j}$$

Datasets

Read the paper!

We consider processes of the form $ab \to 12\ldots n$ for varying multiplicities $n \in \{4,5,6,7\}$

All-gluons processes:

$$gg \to ng$$

One-quark-line processes:

$$gg \to d\bar{d} + (n-2)g \qquad d\bar{d} \to ng$$

Two-quark-lines processes:

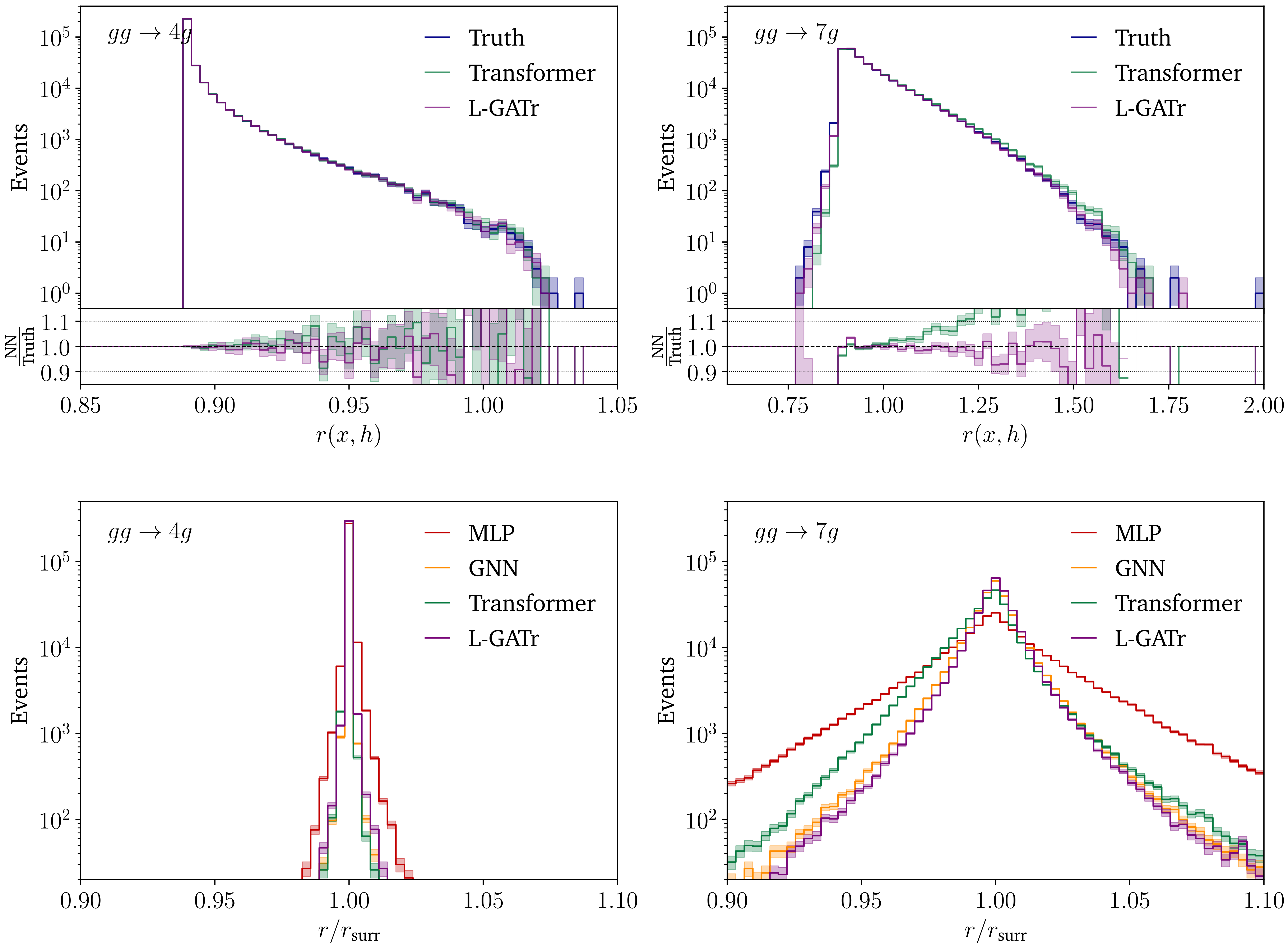
$$d\bar{d} \to u\bar{u} + (n-2)g \qquad gg \to d\bar{d}u\bar{u} + (n-4)g$$

[CO1, CO2]

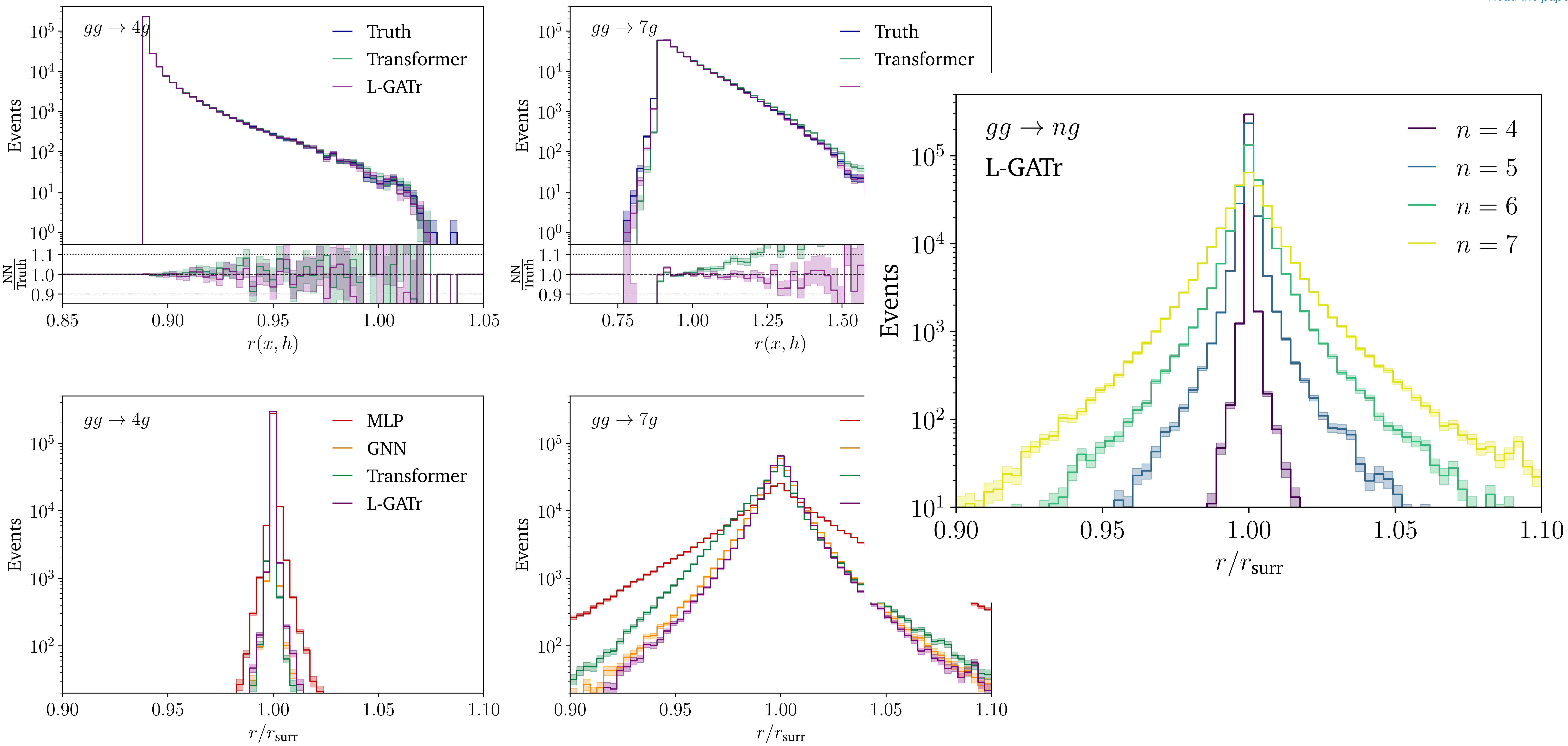
- We **regress r and compute f^{eff} for different NNs**

Results: $gg \to \{4{,}7\}g$

Read the paper!

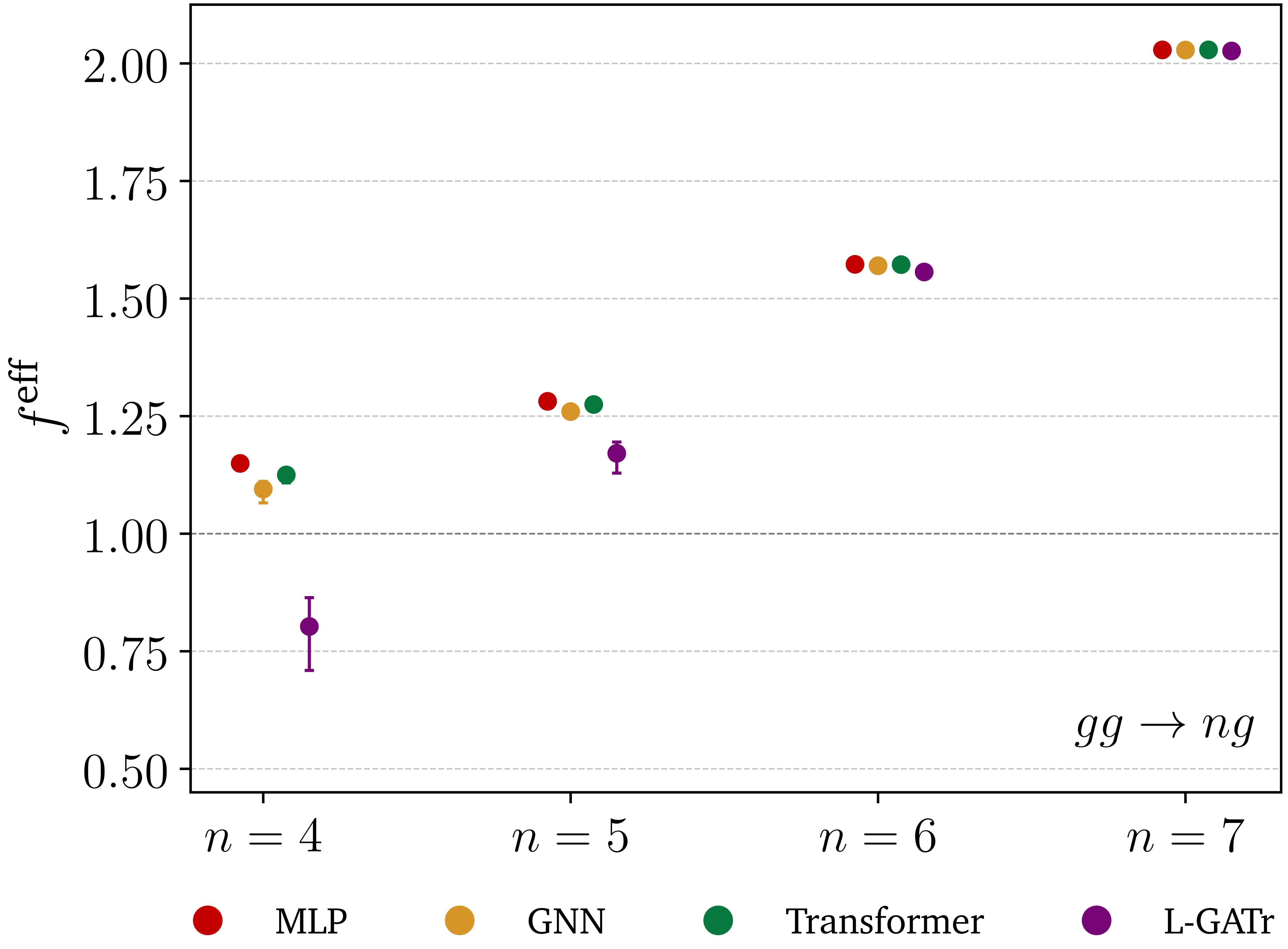


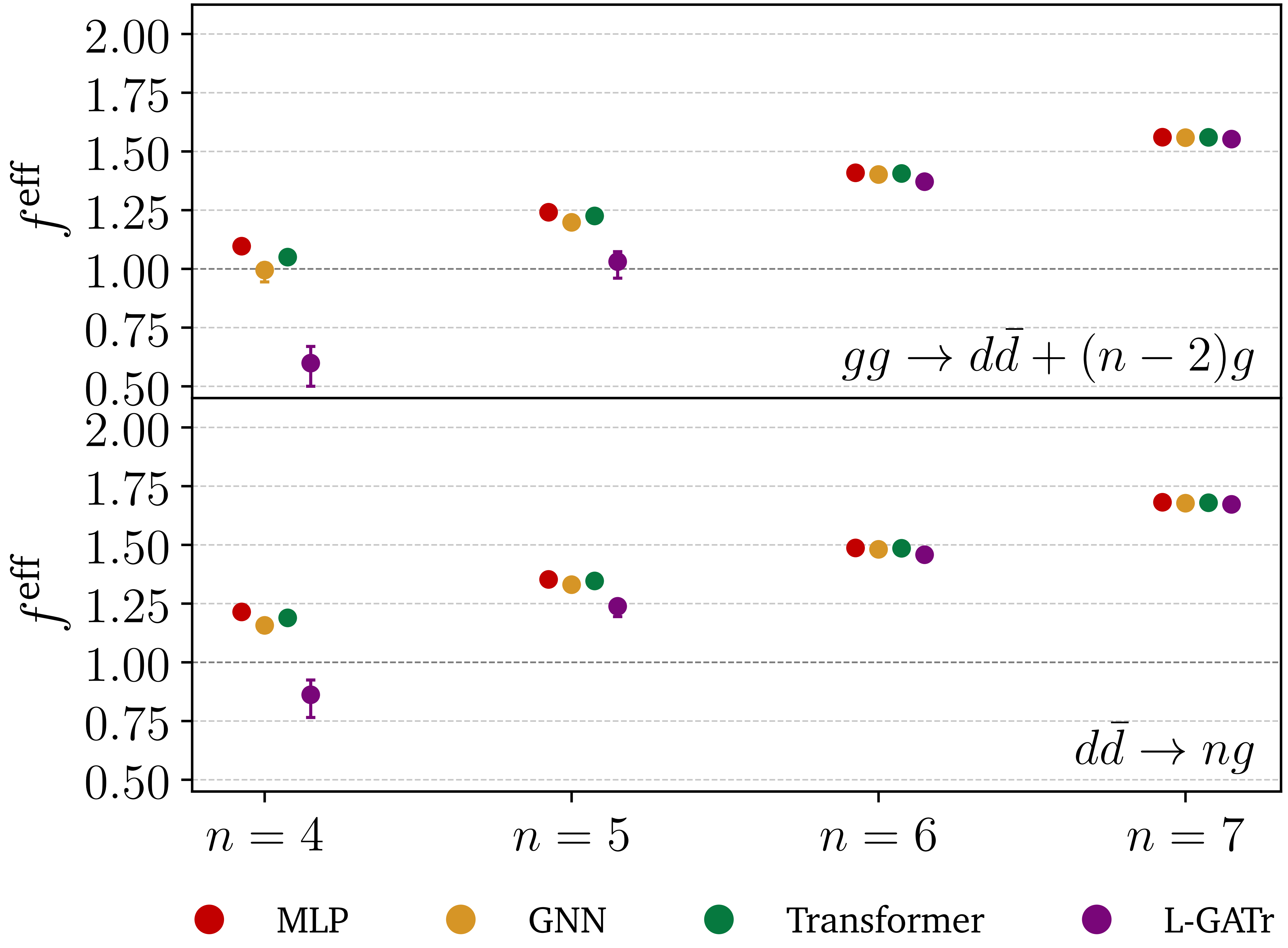
Results: $gg \to \{4,7\}g$



Read the paper!

Results: gain factors

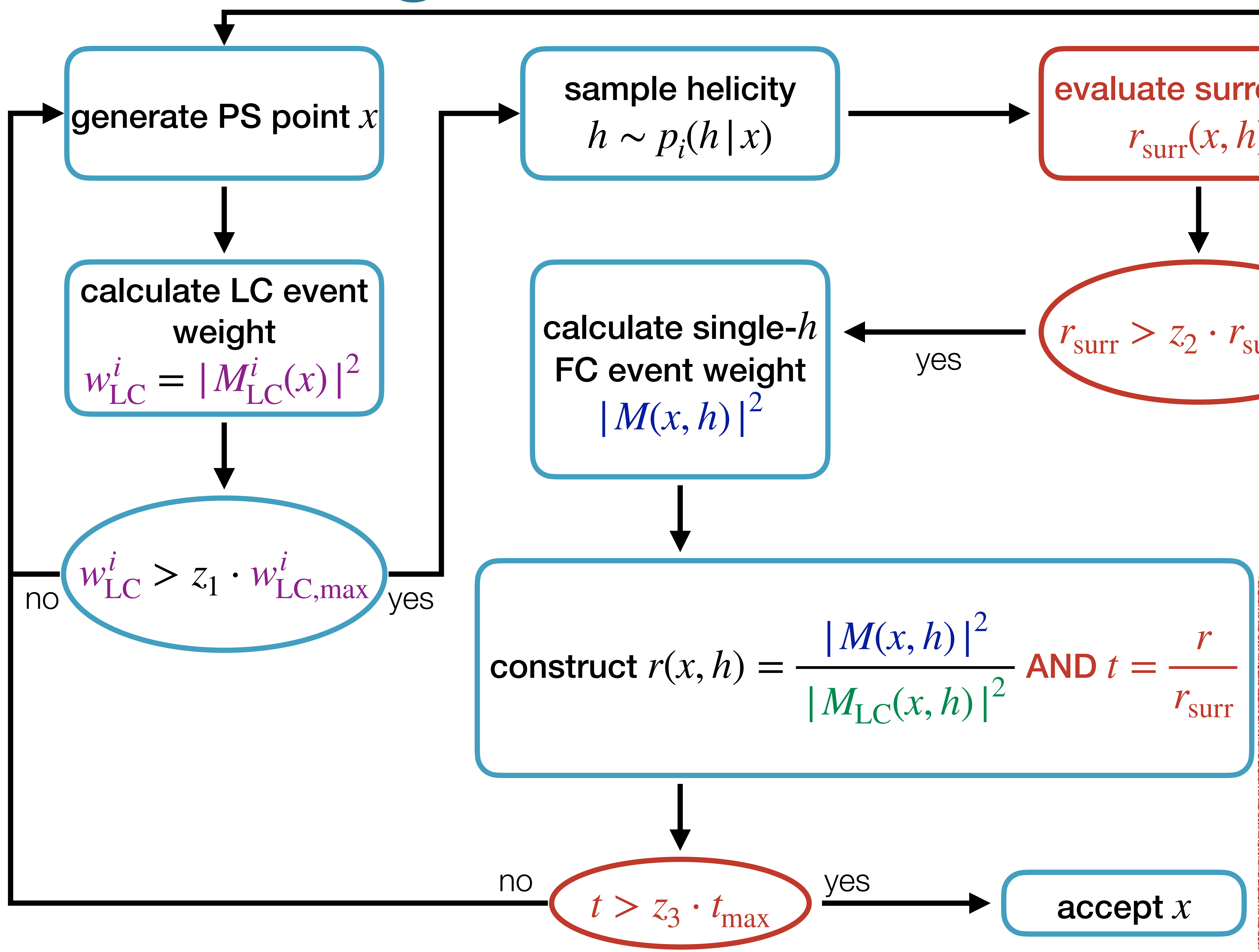




Gains across all processes, from x1.1 to x2 in the all-gluons $n = 7$ case

Results: gain factors eliminating FC unw.

Read the paper!



Also unbiased!

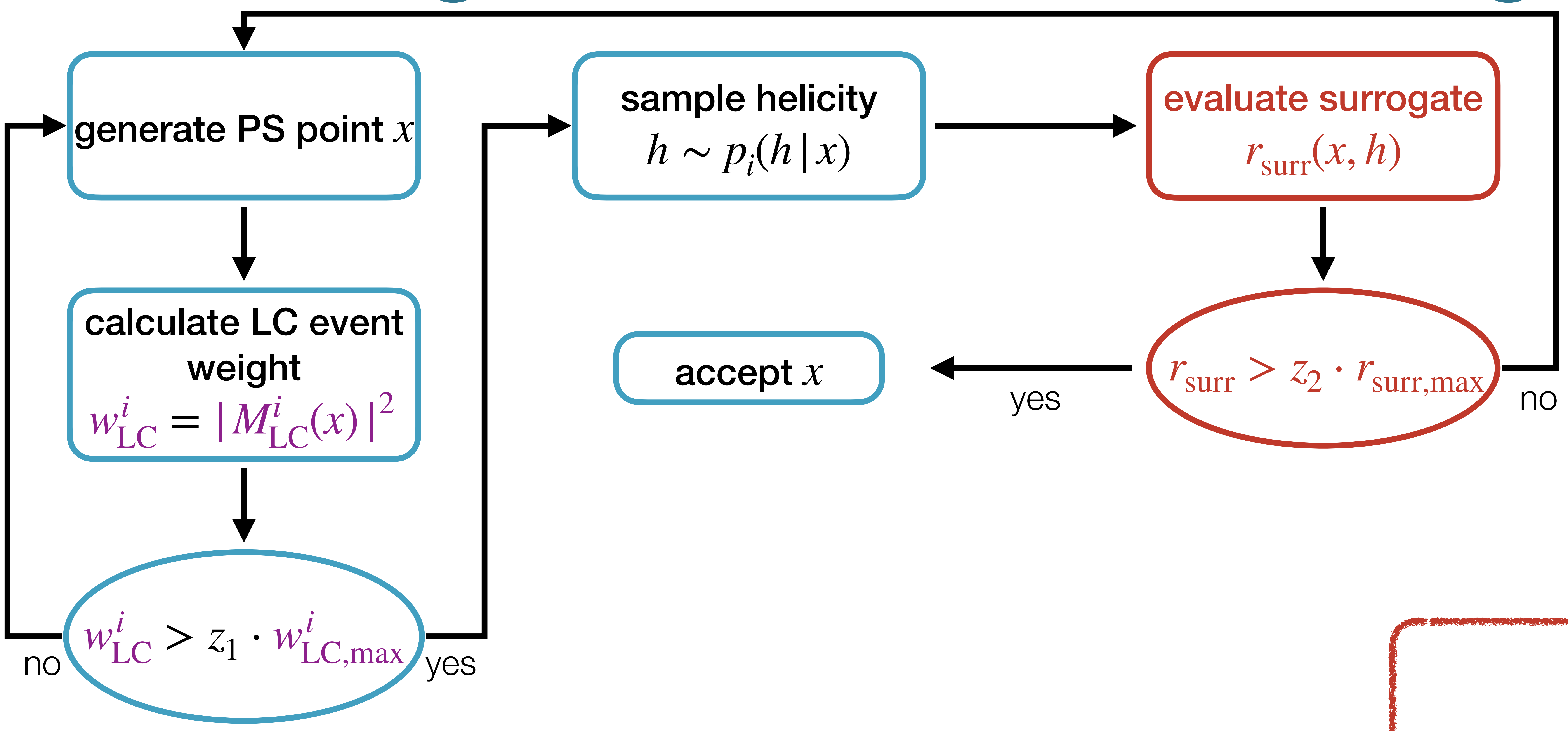
Generated events follow FC density

$$\sum_i \int \mathrm{d}\Phi\, |M^i_{\text{LC}}(x)|^2 \left\langle r_{\text{surr}}(x,h) \cdot t \right\rangle_{h \sim p_i(h|x)}$$

$$\approx \sigma_{\text{FC}}$$

Results: gain factors eliminating FC unw.

Read the paper!



Also unbiased!

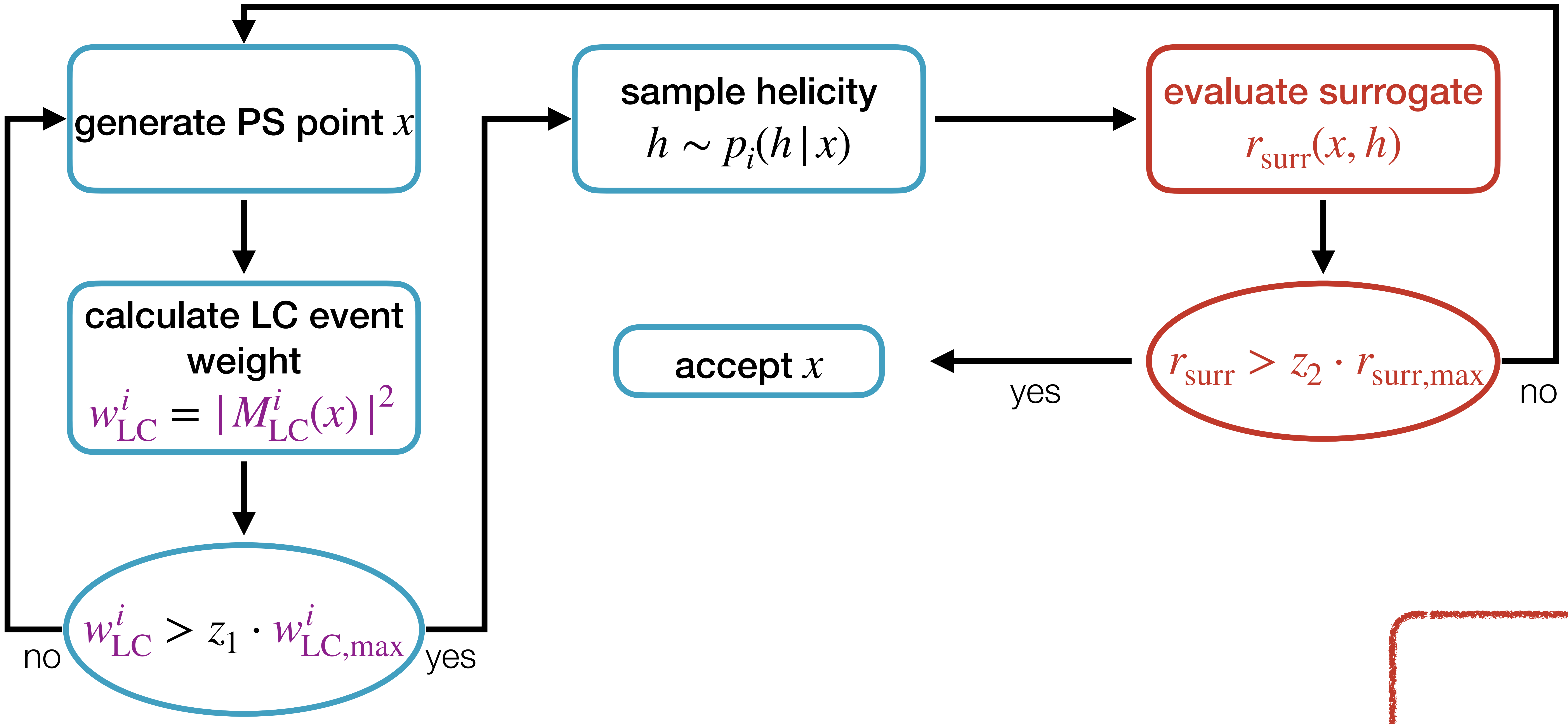
Generated events follow FC density

$$\sum_i \int \mathrm{d}\Phi \, |M^i_{\mathrm{LC}}(x)|^2 \left\langle r_{\mathrm{surr}}(x,h) \cdot t \right\rangle_{h \sim p_i(h|x)}$$

$$\approx \sigma_{\mathrm{FC}}$$

Results: gain factors eliminating FC unw.

Read the paper!



Also unbiased!

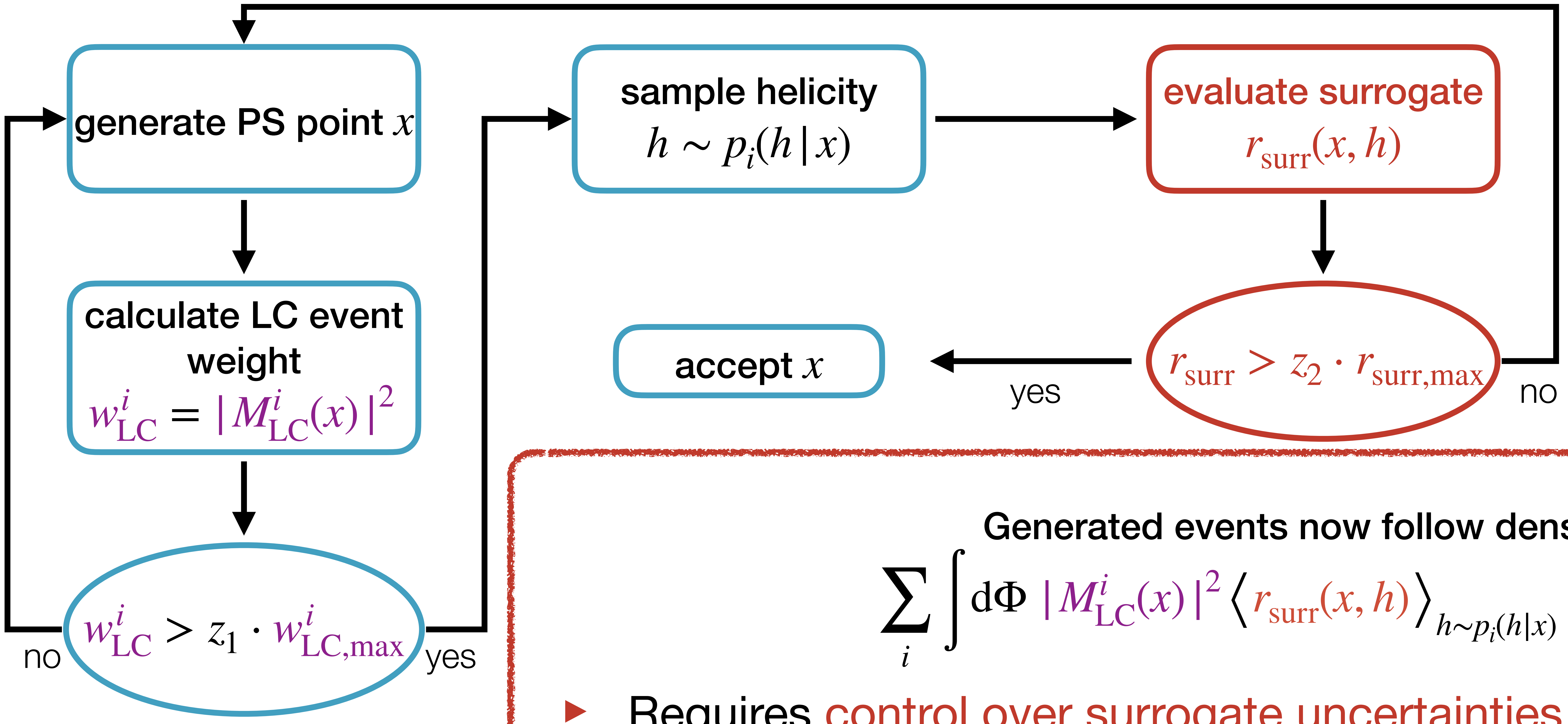
~~Generated ev... ...ow FC density~~

$$\sum_i \int \mathrm{d}\Phi\ |M^i_{\mathrm{LC}}(x \ldots \ldots r_{\mathrm{surr}}(x,h) \cdot t\rangle_{h\sim p_i(h|x)}$$

$$\approx \sigma_{\mathrm{FC}}$$

Results: gain factors eliminating FC unw.

Read the paper!



Generated events now follow density

$$\sum_i \int \mathrm{d}\Phi \, |M^i_{\mathrm{LC}}(x)|^2 \left\langle r_{\mathrm{surr}}(x, h) \right\rangle_{h \sim p_i(h|x)} \overset{\equiv}{\approx} \sigma_{\mathrm{surr,FC}}$$

- Requires control over surrogate uncertainties
- But very feasible!

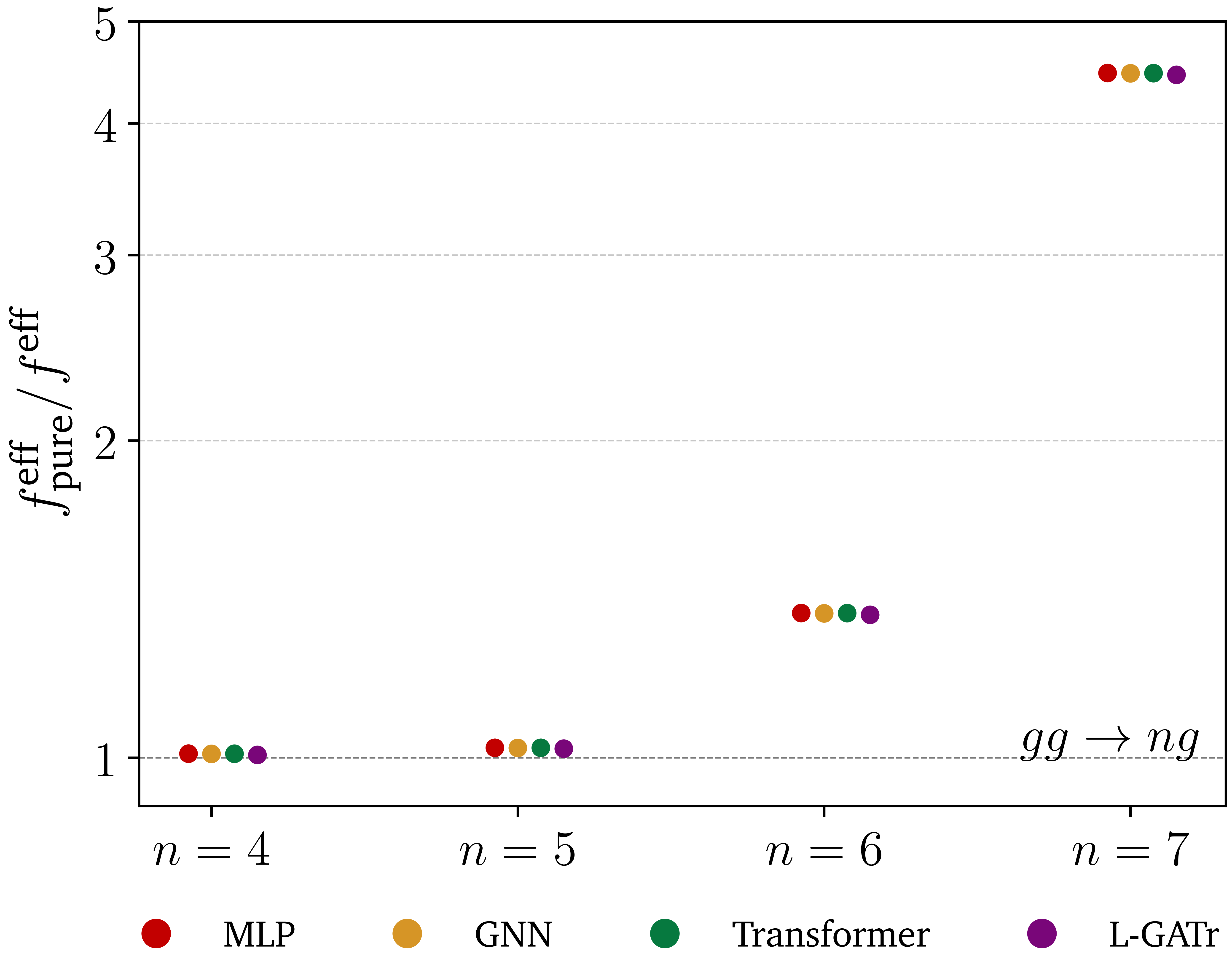
Results: gain factors eliminating FC unw.

Read the paper!

Check out e.g. Suprio's / Lorenz's talks on Monday

Results: gain factors eliminating FC unw.

Read the paper!



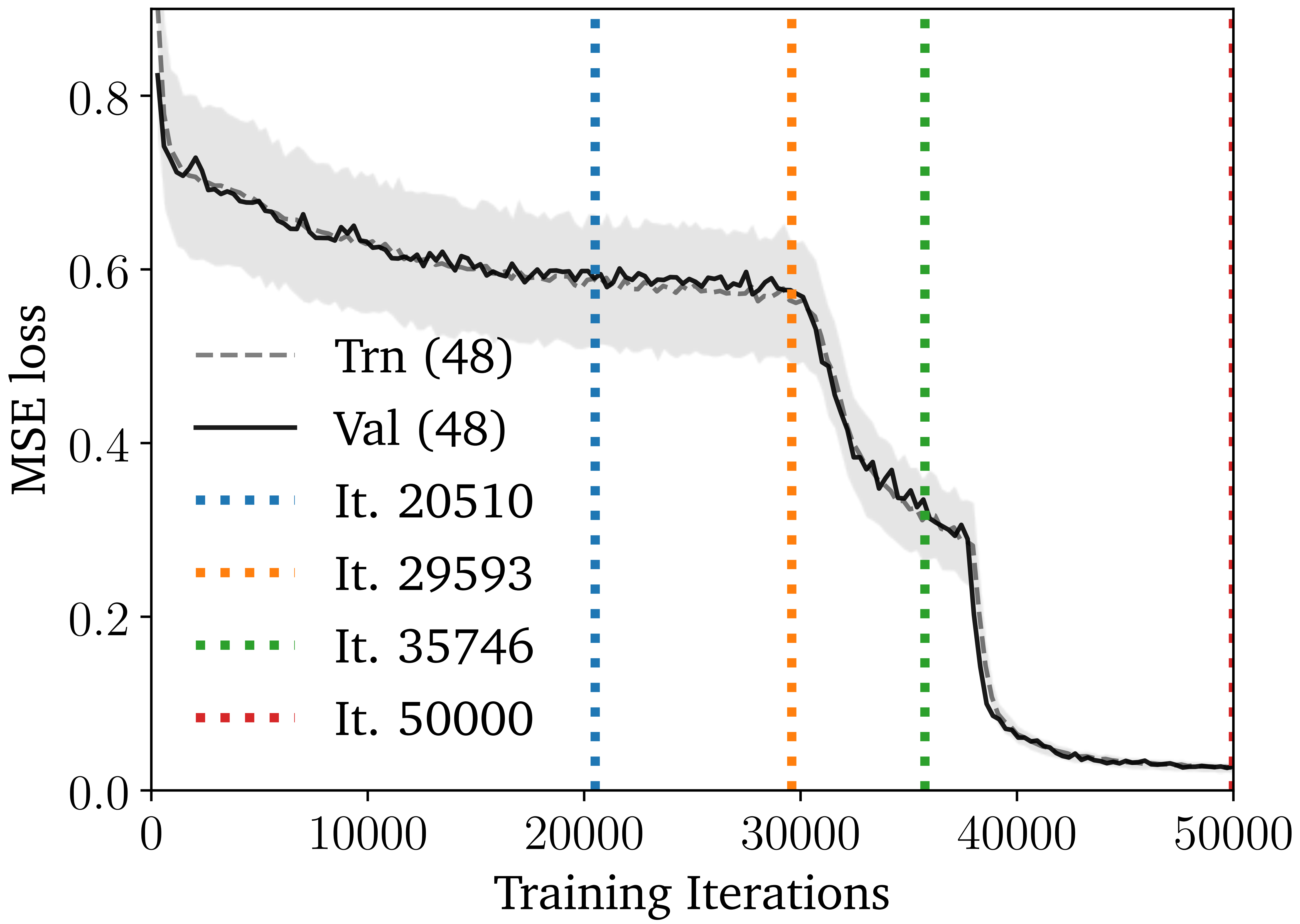
Large improvements with respect to the three-step method

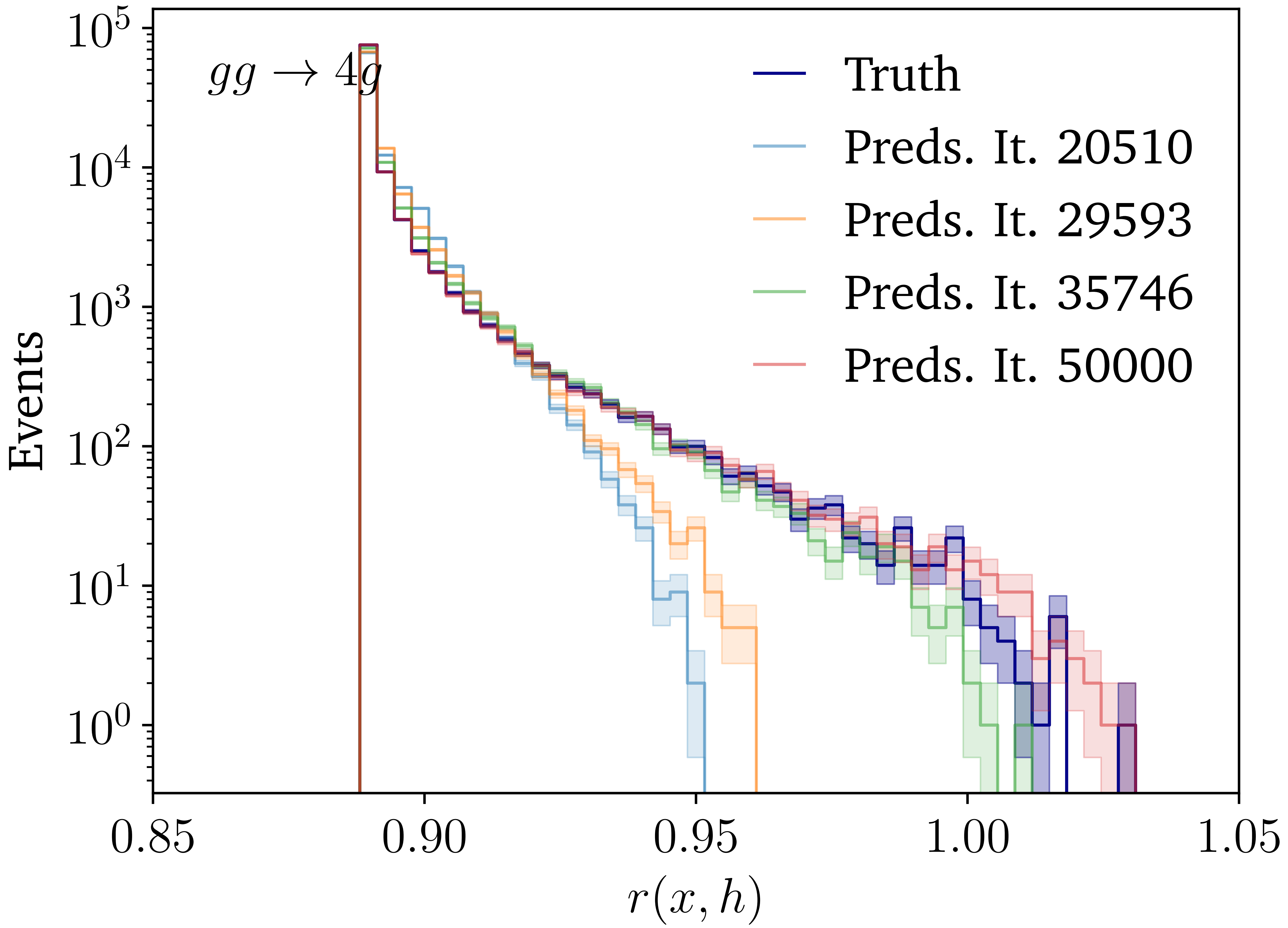
$$f^{\text{eff}}_{\text{pure}} \Big/ f^{\text{eff}} \approx 4.5$$

@ $gg \to 7g$

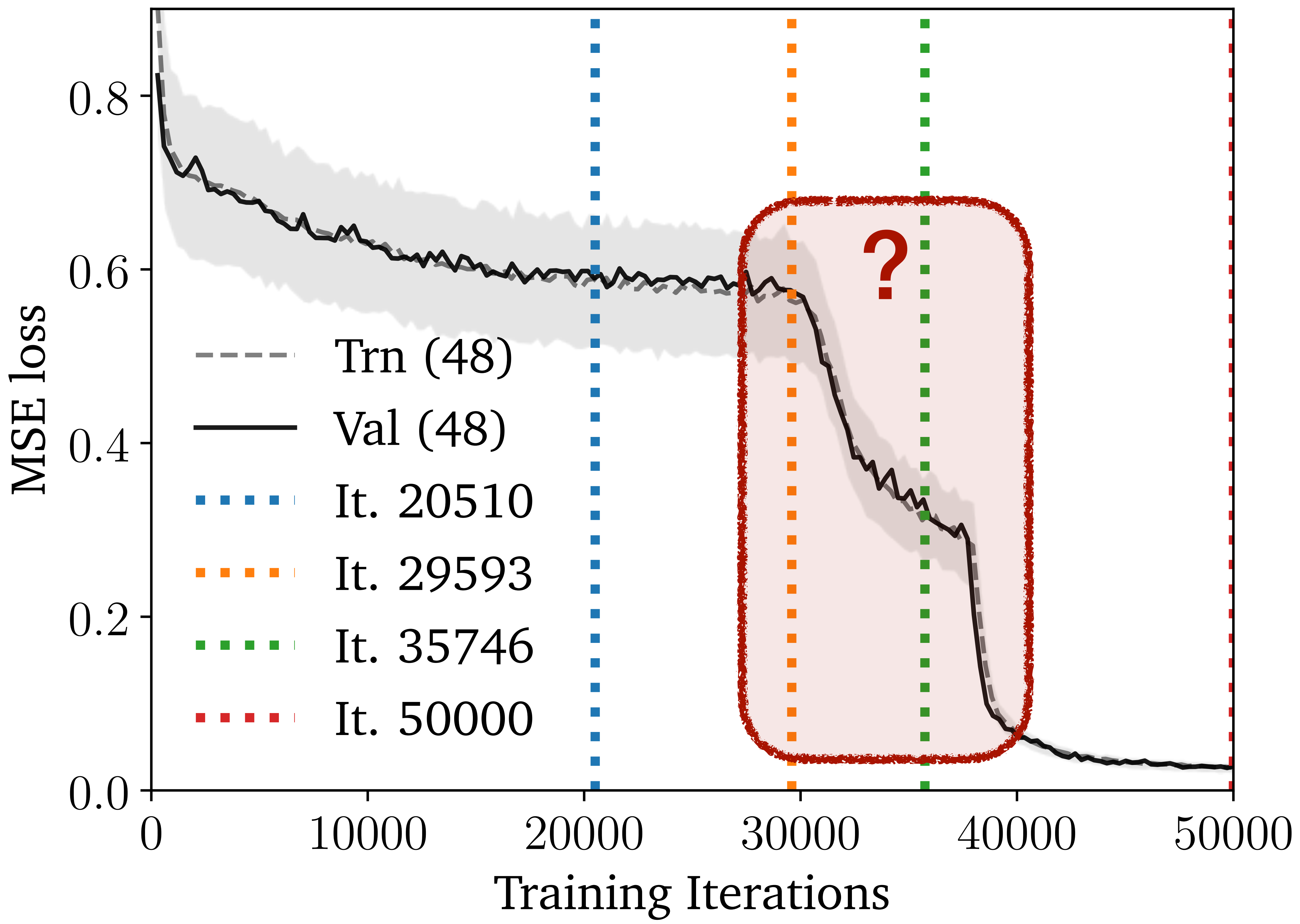
Symmetry learning

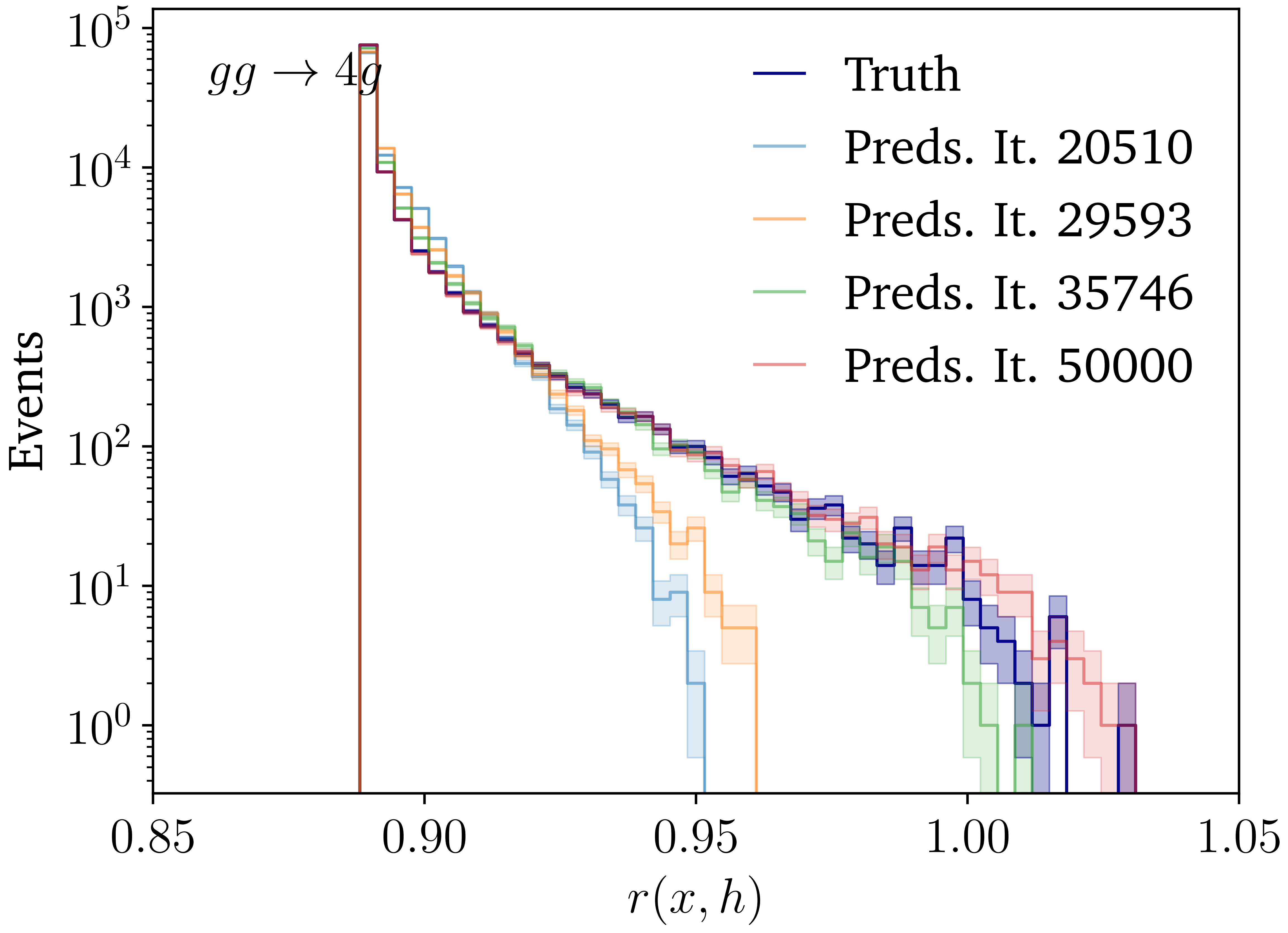
Read the paper!





Symmetry learning





Symmetry learning

Read the paper!

Consider a transformation g of the input

Symmetry learning

Read the paper!

Consider a transformation g of the input

During validation instances @ training time

Evaluate $r_{\text{surr}}(x, h)$ → Compute loss MSE^0

Symmetry learning

Read the paper!

Consider a transformation g of the input

During validation instances @ training time

Evaluate $r_{\mathrm{surr}}(x, h)$ → Compute loss MSE^0

Evaluate $r_{\mathrm{surr}}(g(x), h)$ → Compute loss MSE^g

Symmetry learning

Consider a transformation g of the input

During validation instances @ training time

Evaluate $r_{\text{surr}}(x, h)$ → Compute loss MSE^0

Evaluate $r_{\text{surr}}(g(x), h)$ → Compute loss MSE^g

$$g \in \{\text{SO}(2),\ \text{Boost},\ \text{SL}(4),\ \text{Shear}\}$$

Read the paper!

Symmetry learning

Read the paper!

Consider a transformation g of the input

During validation instances @ training time

Evaluate $r_{\text{surr}}(x, h)$ → Compute loss MSE^0

Evaluate $r_{\text{surr}}(g(x), h)$ → Compute loss MSE^g

$$g \in \{\boxed{\text{SO}(2),\ \text{Boost},}\ \text{SL}(4),\ \text{Shear}\}$$

Lorentz tfms

(symmetry present in data)

Symmetry learning

Read the paper!

Consider a transformation g of the input

During validation instances @ training time

Evaluate $r_{\text{surr}}(x, h)$ → Compute loss MSE^0

Evaluate $r_{\text{surr}}(g(x), h)$ → Compute loss MSE^g

$$g \in \{\boxed{\text{SO}(2),\ \text{Boost},}\ \boxed{\text{SL}(4),\ \text{Shear}}\}$$

Lorentz tfms

(symmetry present in data)

Random tfms

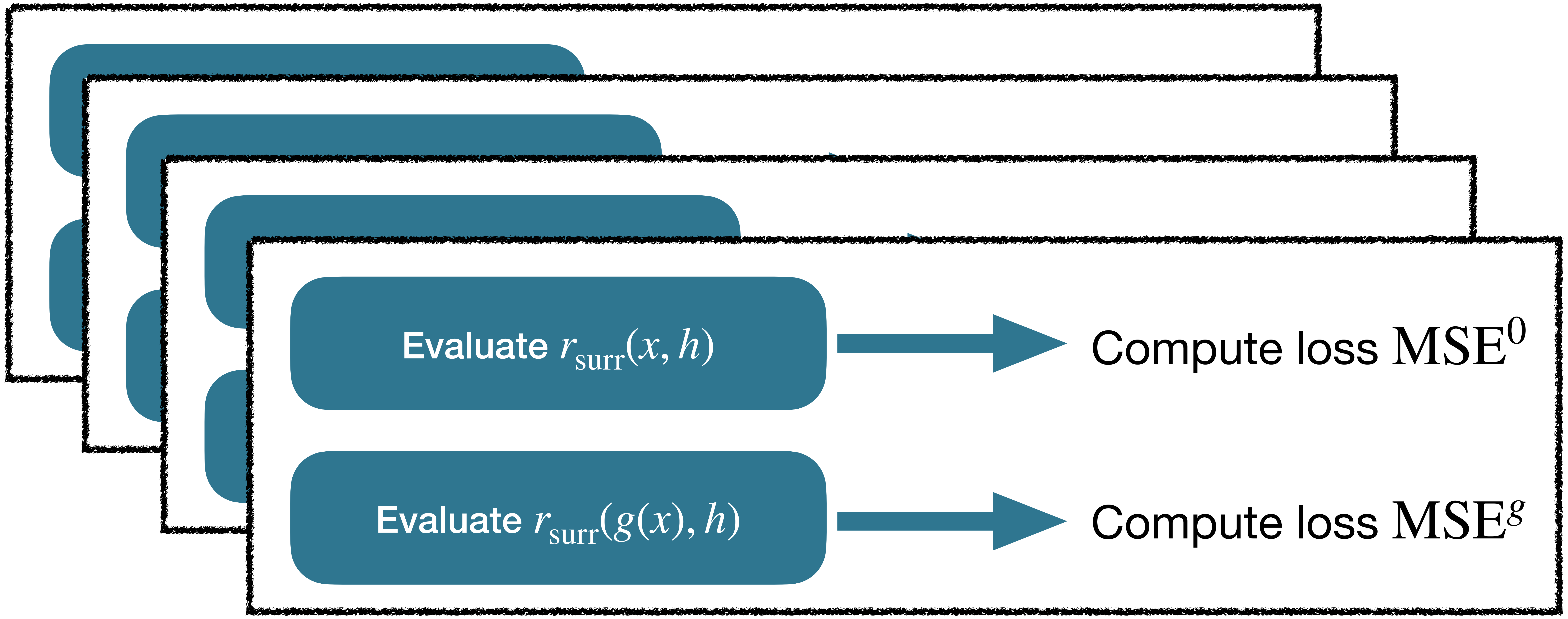
(not present)

Symmetry learning

Read the paper!

Consider a transformation g of the input

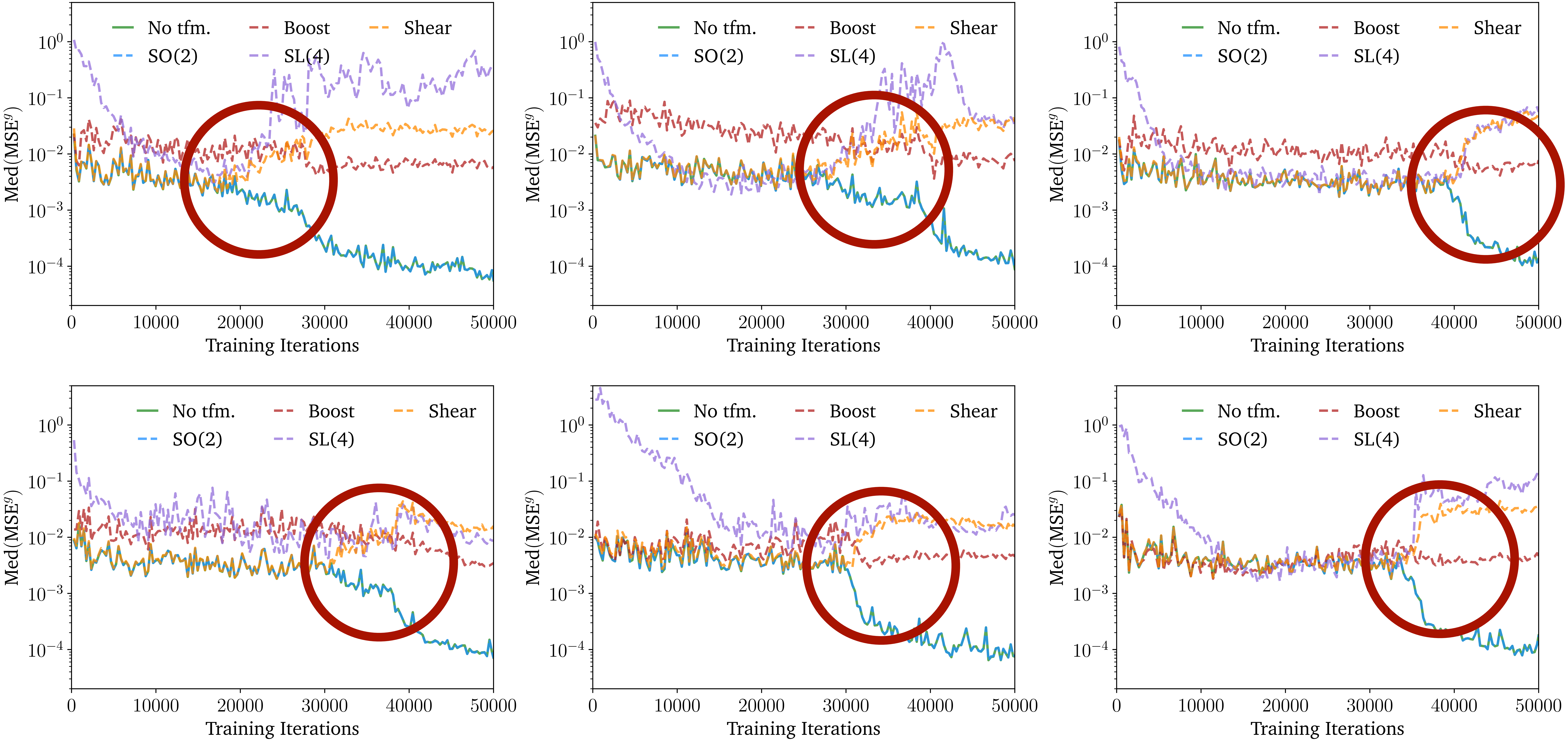
During validation instances @ training time



Train an ensemble of 120 networks

Symmetry learning

paper!



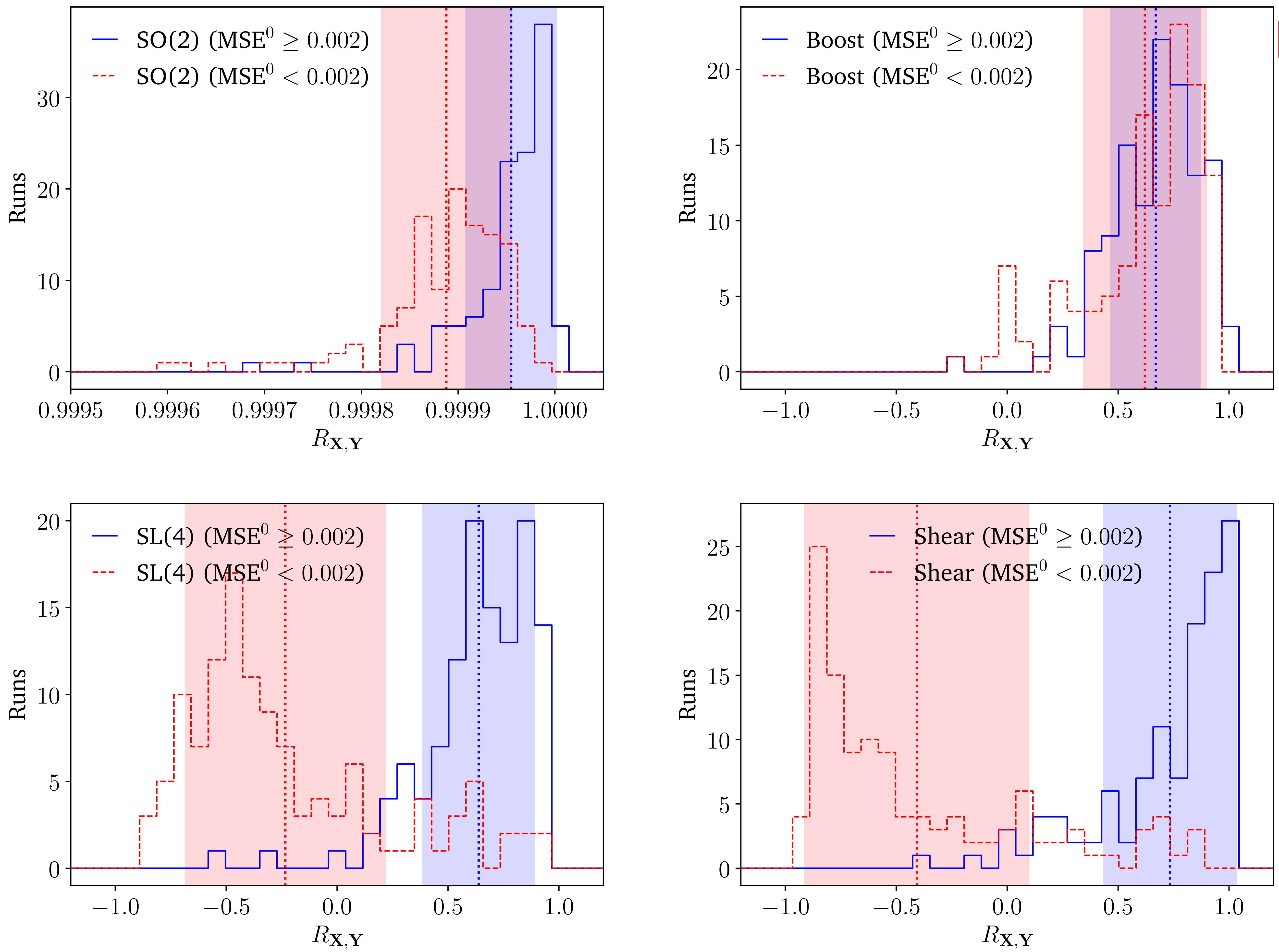
Symmetry learning

Read the paper!

$$R_{\mathbf{X},\mathbf{Y}}(g) = \frac{\mathrm{cov}(\mathbf{X},\mathbf{Y})}{\sqrt{\mathrm{var}(\mathbf{X})\mathrm{var}(\mathbf{Y})}}, \qquad \mathbf{X} = \mathrm{Med}(\mathrm{MSE}^0),\ \mathbf{Y} = \mathrm{Med}(\mathrm{MSE}^g)$$

Loss plateau

Loss collapse



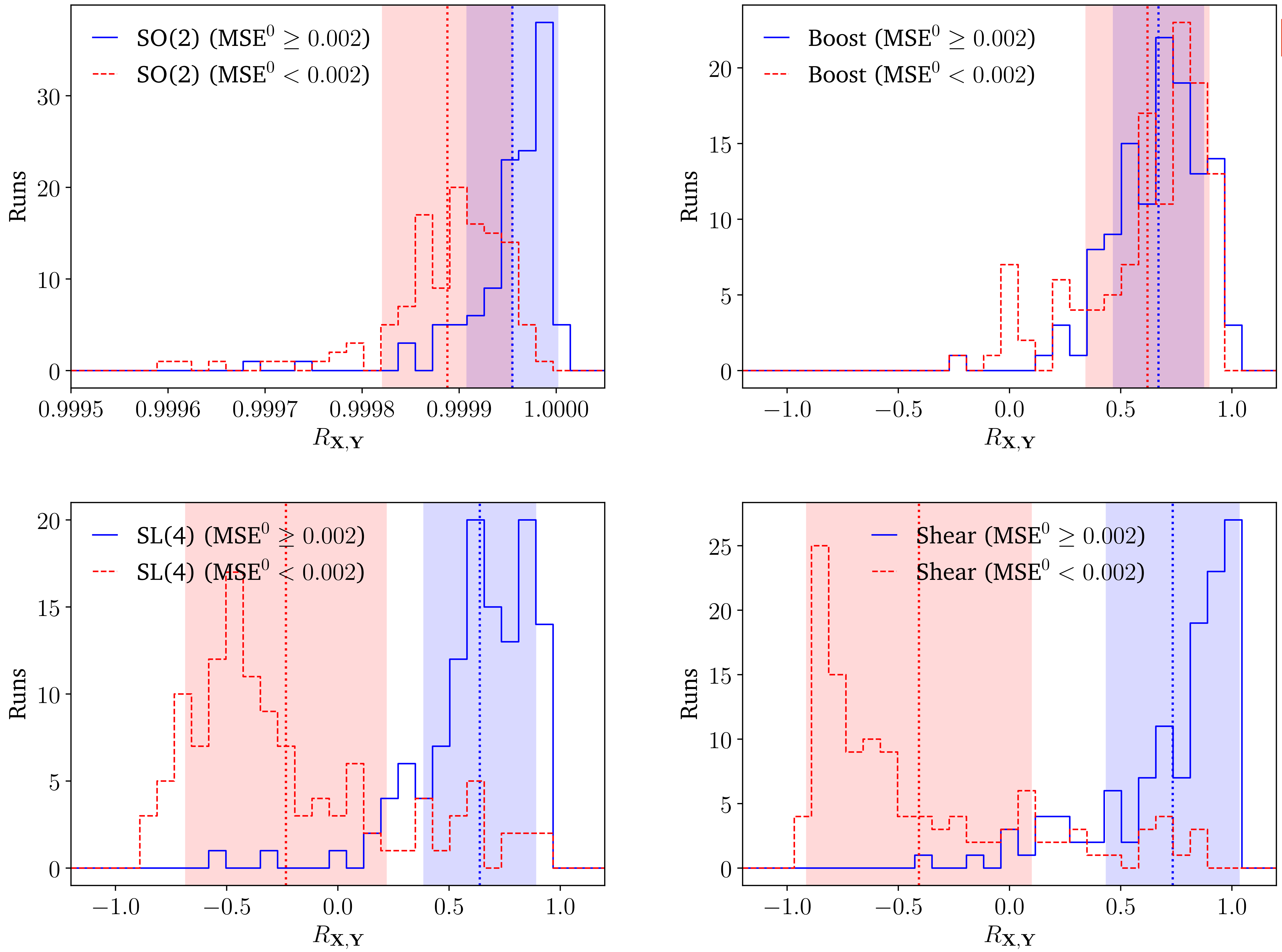
Symmetry learning

Read the paper!

$$R_{\mathbf{X},\mathbf{Y}}(g) = \frac{\mathrm{cov}(\mathbf{X},\mathbf{Y})}{\sqrt{\mathrm{var}(\mathbf{X})\mathrm{var}(\mathbf{Y})}}, \qquad \mathbf{X} = \mathrm{Med}(\mathrm{MSE}^0),\ \mathbf{Y} = \mathrm{Med}(\mathrm{MSE}^g)$$

Loss plateau

Loss collapse



Lorentz symmetries remain correlated with optimized loss

Non-Lorentz symmetries are discarded as training evolves

Summary

Read the paper!

- Achieved **x2 speed-ups** and **always** > 1
- Potentially **x9 speed-ups are possible** by controlling NN uncertainties
- **Integration with MC generators** can have large impact at HL-LHC
- Networks can **learn symmetries from underlying data**, enhancing generalization

Summary

Read the paper!

- Achieved **x2 speed-ups** and **always** > 1
- Potentially **x9 speed-ups are possible** by controlling NN uncertainties
- **Integration with MC generators** can have large impact at HL-LHC
- Networks can **learn symmetries from underlying data**, enhancing generalization

Thank you for your attention!