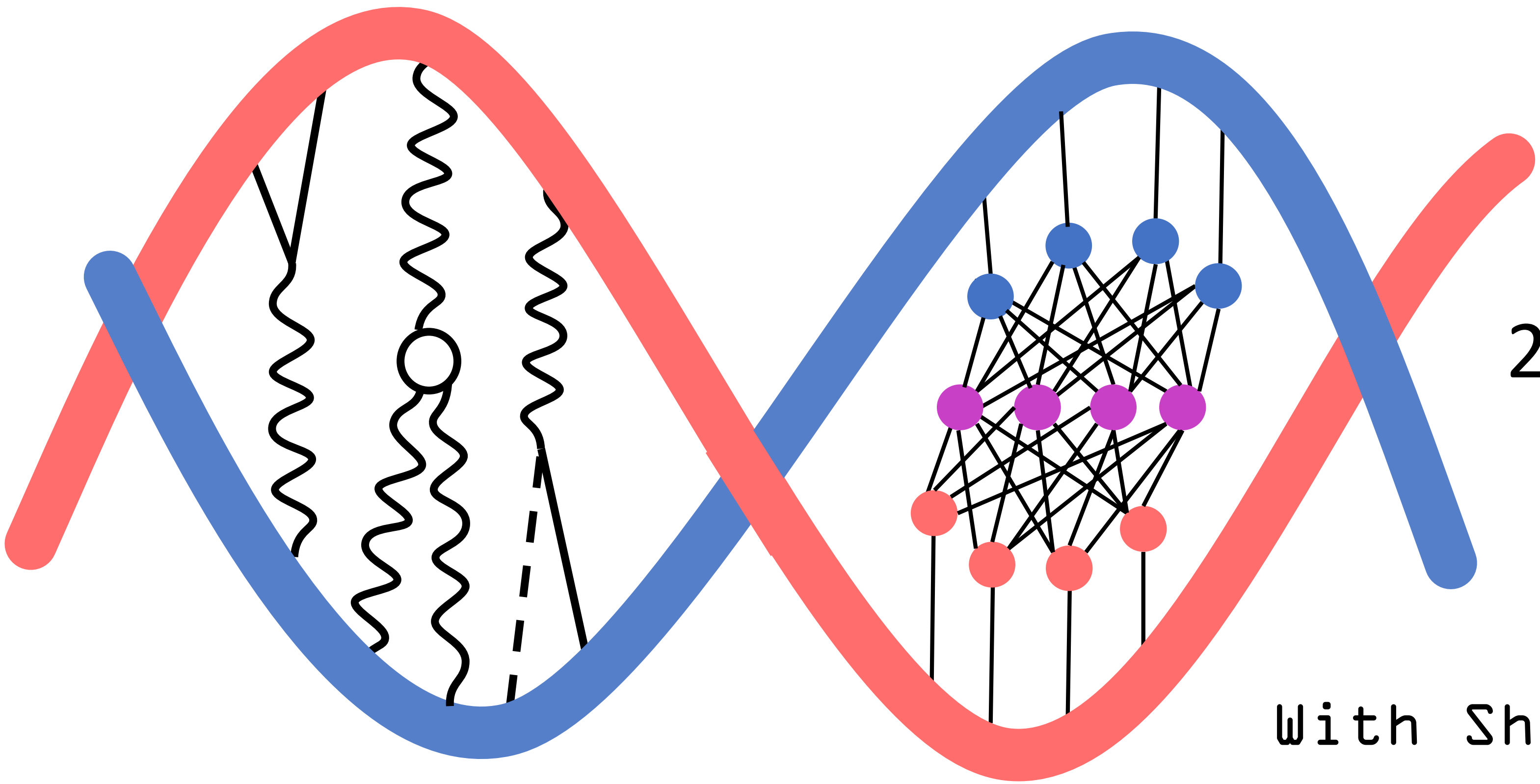
FROM OBSERVABLE TO EFT:
a Case Study of aTGC Measurement with Machine Learning



Lingfeng Li

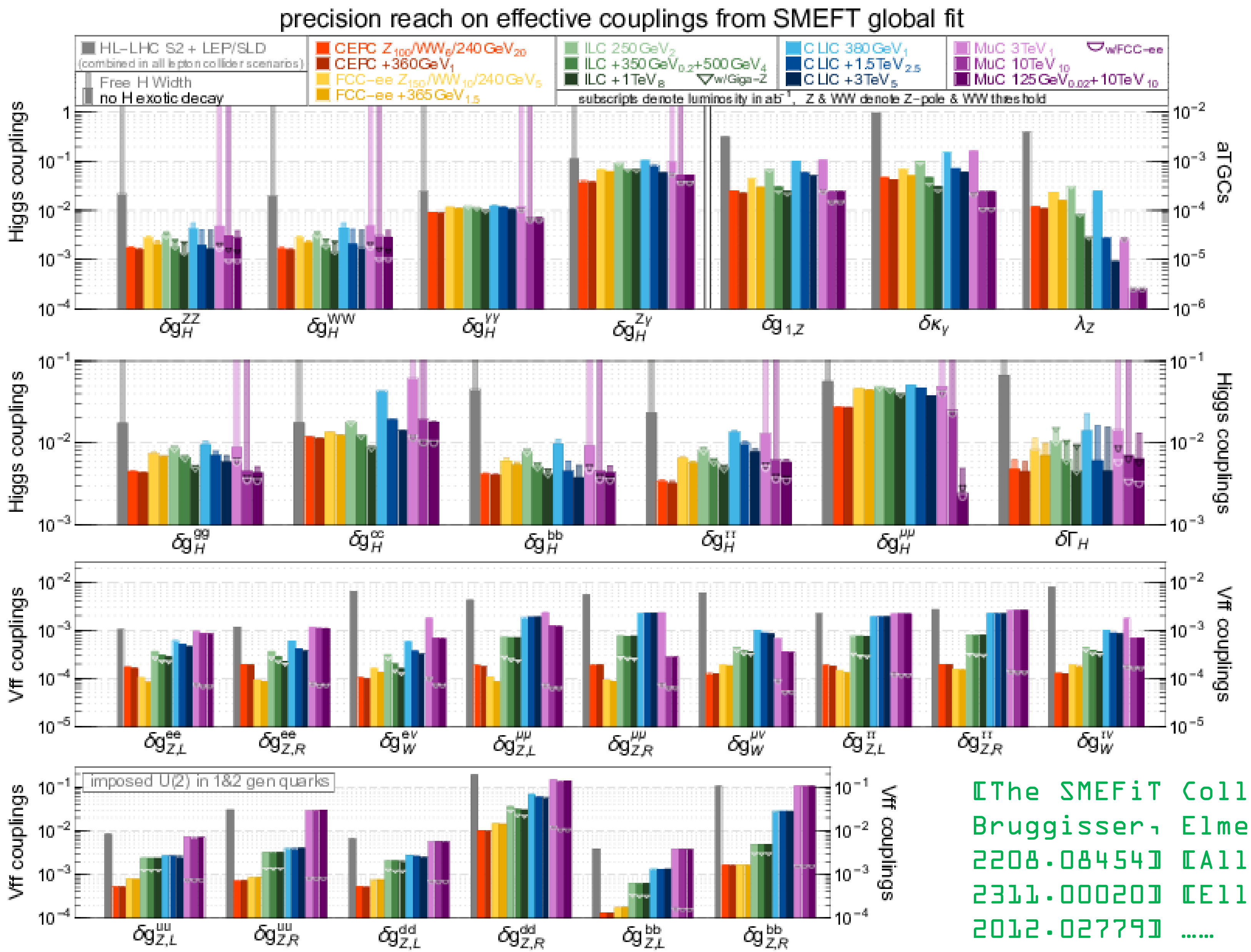
Brown. University

2nd US FCC workshop, MIT

Mar. 26, 2024

arXiv:2401.02474

With Shengdu Chai and Jiayin Gu



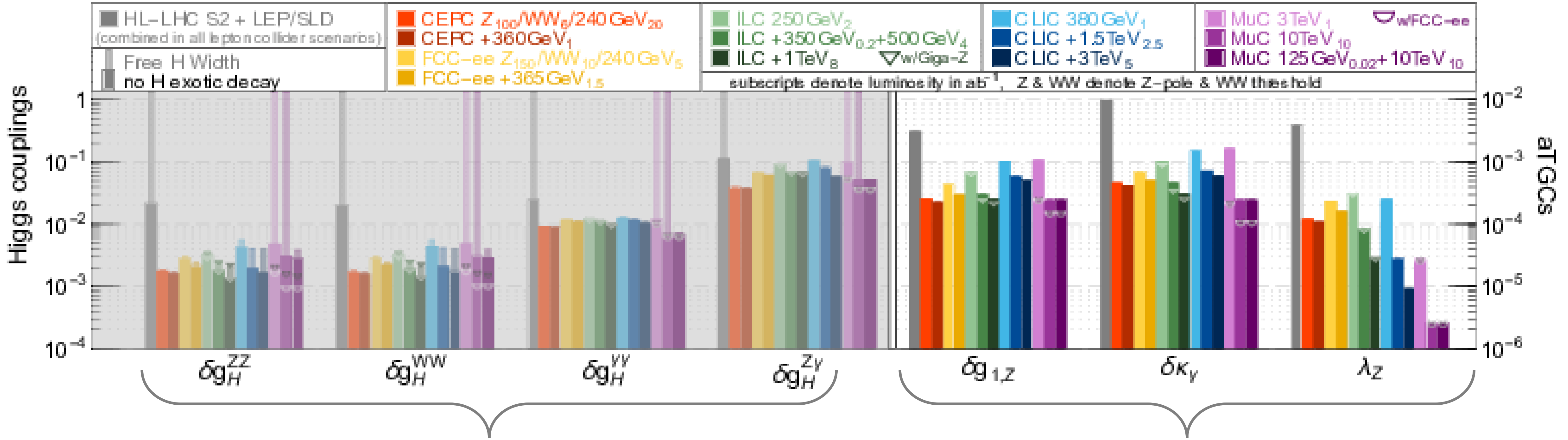
de Blas, Du, Grojean, Gu, Miralles, Peskin, Tian, Vos, Vryonidou, 2206.08326

Precision measurement of the SM is (one of) the main goal of FCC-ee

Efforts in FCC SMEFT fits:

[The SMEFiT Collaboration 2105.00006] [Brivio, Bruggisser, Elmer, Geoffray, Luchmann, Plehn, 2208.08454] [Allwicher, Cornella, Isidori, Stefanek, 2311.00020] [Ellis, Madigan, Mimasu, Sanz, You, 2012.02779] ……

“PROMISED” PRECISION OF ATGC



Higgs measurements. High precision mainly from Higgs BR

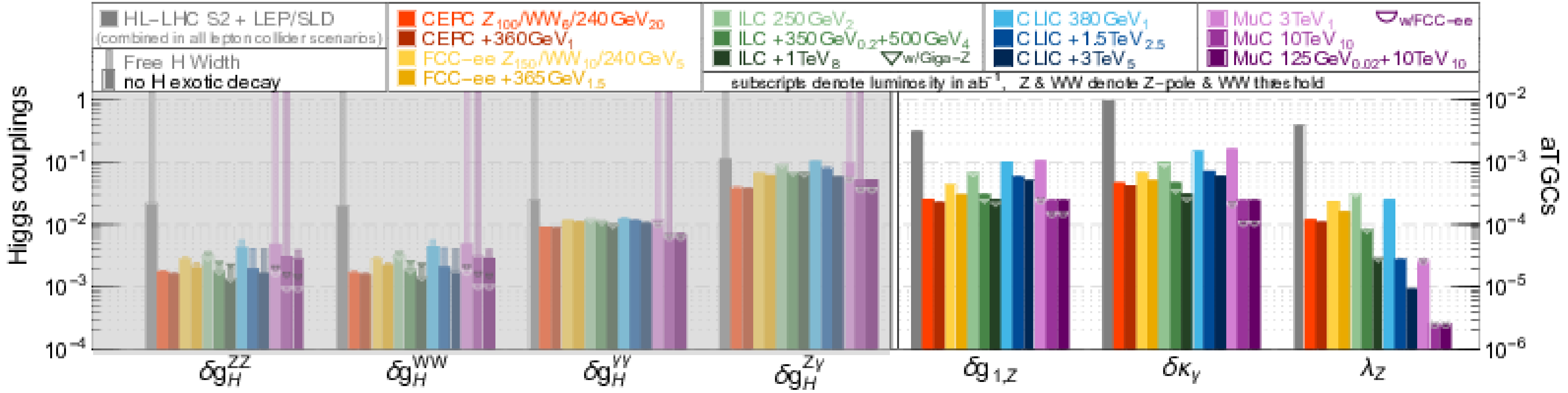
Anomalous Triple Gauge Couplings. From differential measurements of diboson events

Issue raised in the ECFA report

J. Maestre et al., 2401.07564

Lingfeng Li | 2401.02474

"PROMISED" PRECISION OF ATGC



High Ecm advantageous, Higgs factory mode instead of WW mode

$$\mathcal{L}_{\rm TGC} = ie(W^+_{\mu\nu}W^{-\mu} - W^-_{\mu\nu}W^{+\mu})A^\nu + ie(1+\delta\kappa_\gamma)A^{\mu\nu}W^+_\mu W^-_\nu$$
$$+ igc_w\left[(1+\delta g_{1Z})(W^+_{\mu\nu}W^{-\mu} - W^-_{\mu\nu}W^{+\mu})Z^\nu + \left(1+\delta g_{1Z} - \frac{s_w^2}{c_w^2}\delta\kappa_\gamma\right)Z^{\mu\nu}W^+_\mu W^-_\nu\right]$$
$$+ \frac{ig\lambda_Z}{m_W^2}\left(s_w W^{+\nu}_\mu W^{-\rho}_\nu A^\mu_\rho + c_w W^{+\nu}_\mu W^{-\rho}_\nu Z^\mu_\rho\right),$$

Lingfeng Li | 2401.02474

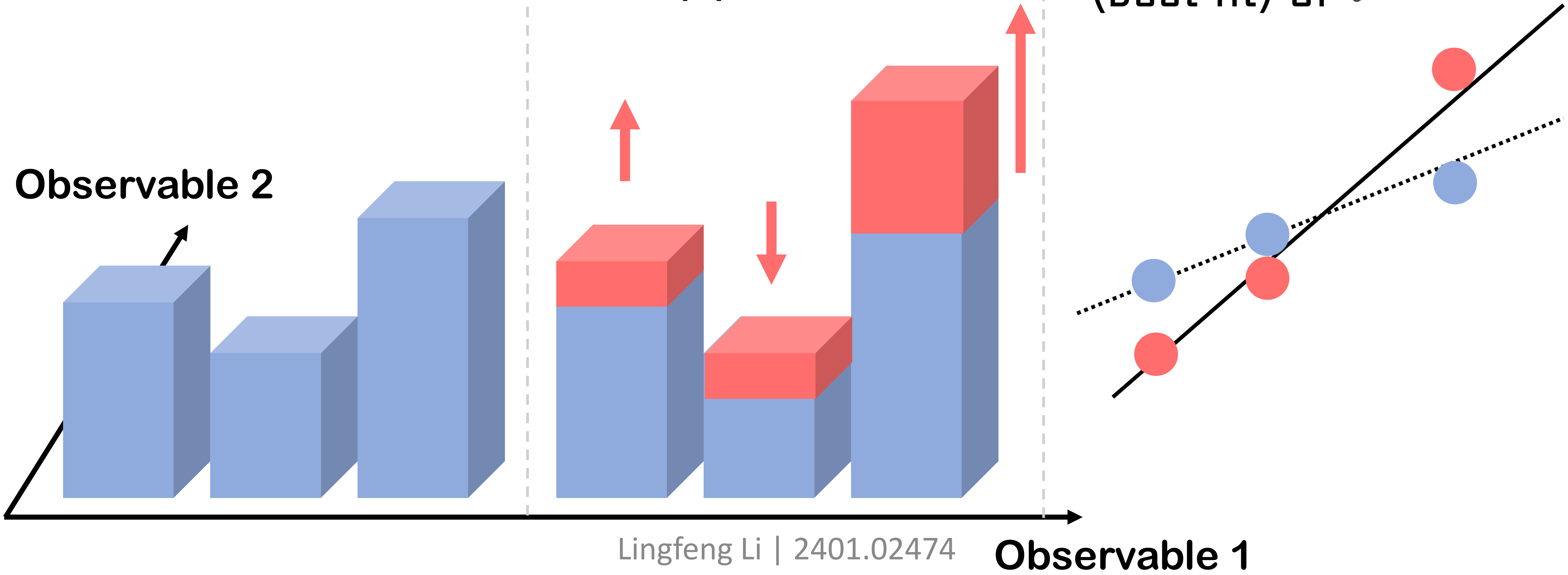
Similar stories in $|V_{cb}|$ from W decays 【In prep.】

DIFFERENTIAL MEASUREMENT

Cut SR into several bins

Counts change with theory parameter θ

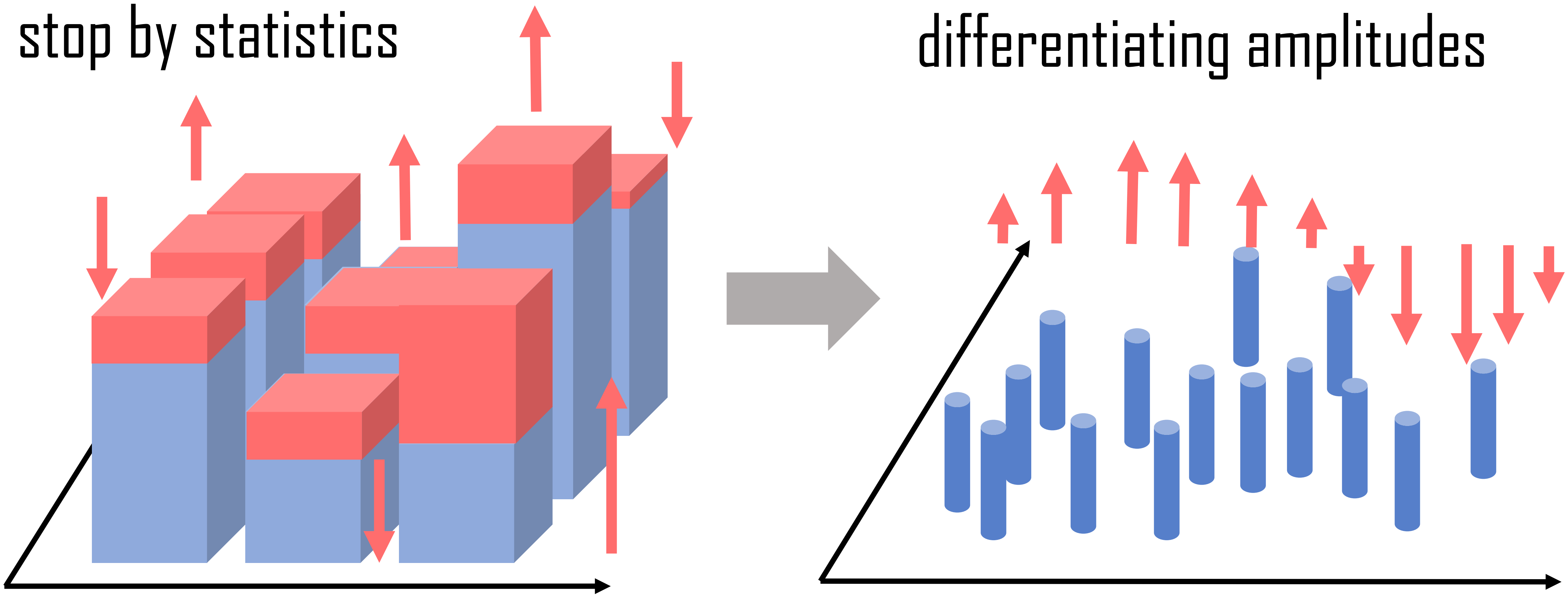
Find the max likelihood (best fit) of θ



FROM BINS TO OPTIMAL OBSERVABLES

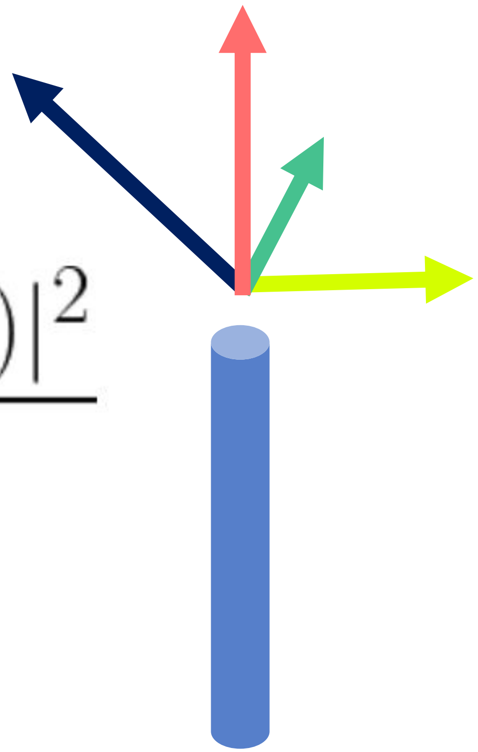
Define more and more bins,
stop by statistics

Dependence of each event by
differentiating amplitudes



FROM BINS TO OPTIMAL OBSERVABLES

Promote the dependence into event-by-event vectors



$$\vec{\alpha}_i(\mathbf{x}) \propto \frac{\partial |\mathcal{A}(\mathbf{x})|^2}{\partial \theta_i}$$

Can be achieved during MC:

J. Brehmer, F. Kling, I. Espejo, K. Cranmer, 1907.10621

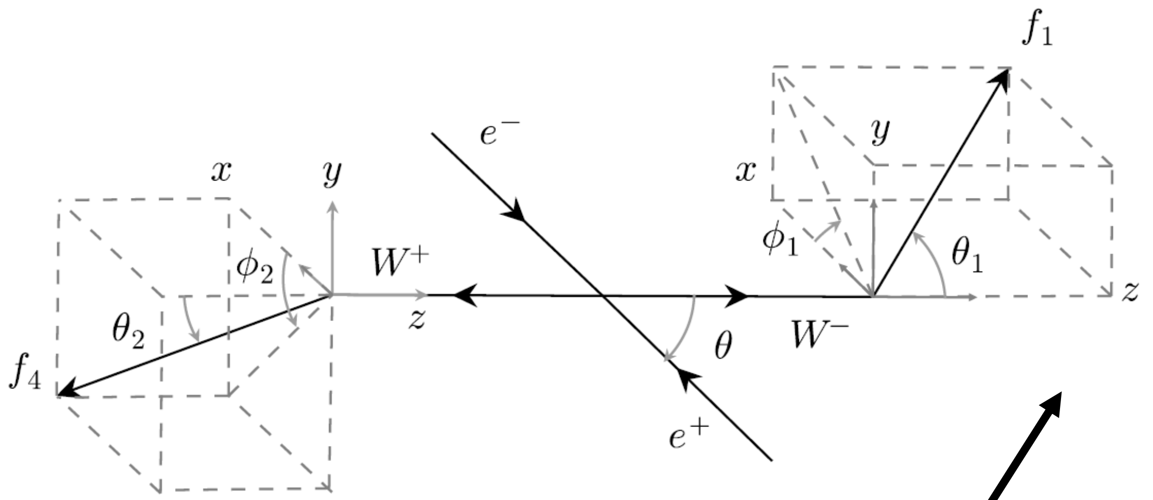
The Optimal Observable (OO)

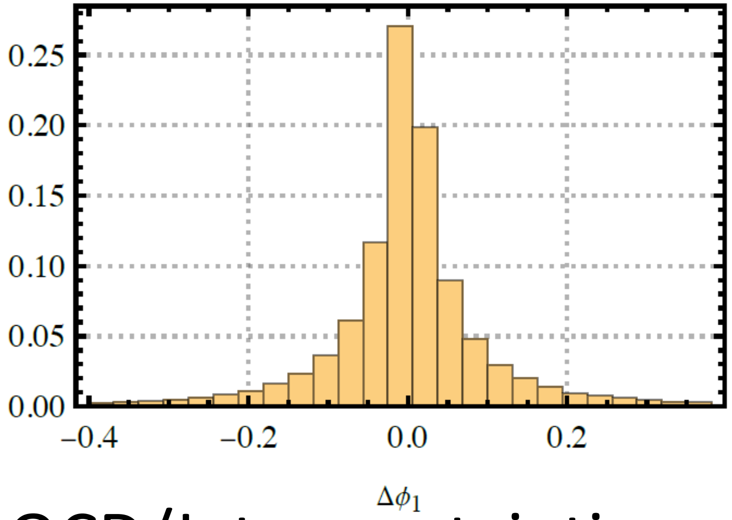
M. Diehl and O. Nachtmann, Z.Phys.C 62 (1994) 397-412

Expectation & limits on θ are extracted from the likelihood

$$\mathcal{L} \propto \sum_{i \in \text{events}} (1 + \vec{\alpha} \cdot \vec{\theta})$$

SYSTEMATIC EFFECTS

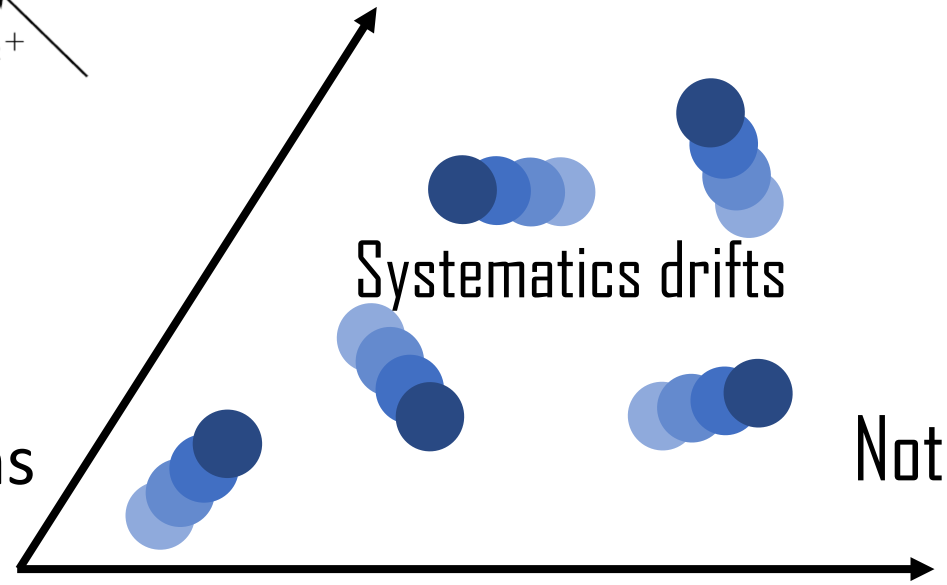




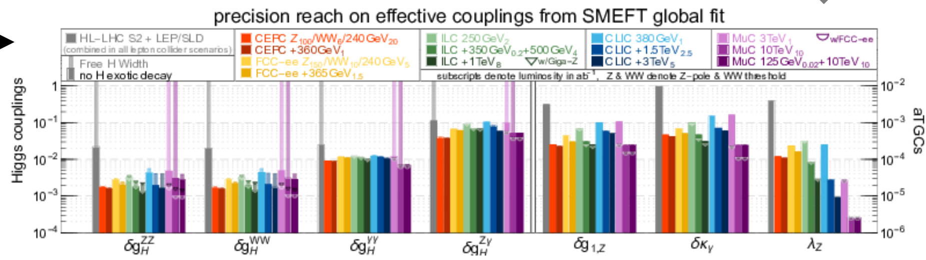
- ❖QCD/Jet uncertainties
- ❖Neutrino info only inferred (non-linear)
- ❖Beam ISR
- ❖Detector resolution

Leading channel: WW semileptonic decays

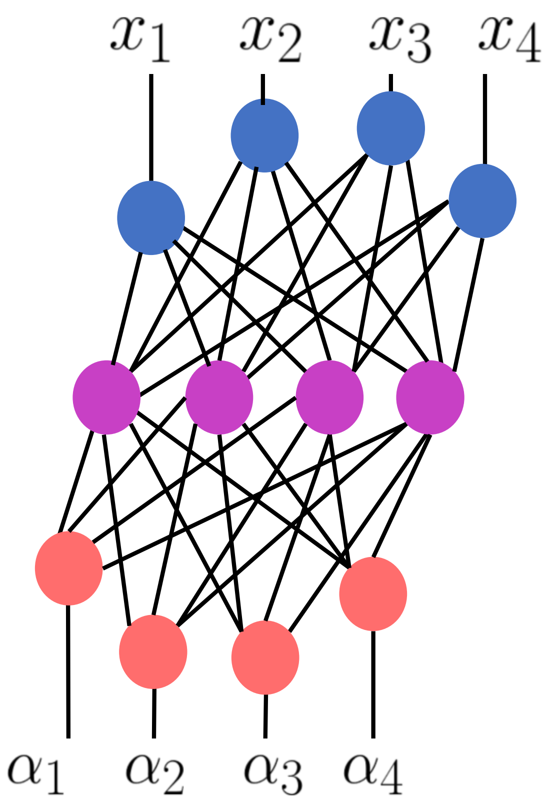
- ✓Discernable W bosons
- ✓Large BR (~50%)
- ✓Good reconstruction efficiency



Not yet included in projections



INFERENCE FROM NEURON NETWORKS



From classification to regression

SALLY (Score approximates likelihood locally) algorithm, the Optimal Observable extension in ML J. Brehmer, K. Cranmer, G. Louppe and J. Pavez 1805.00013 , 1805.00020

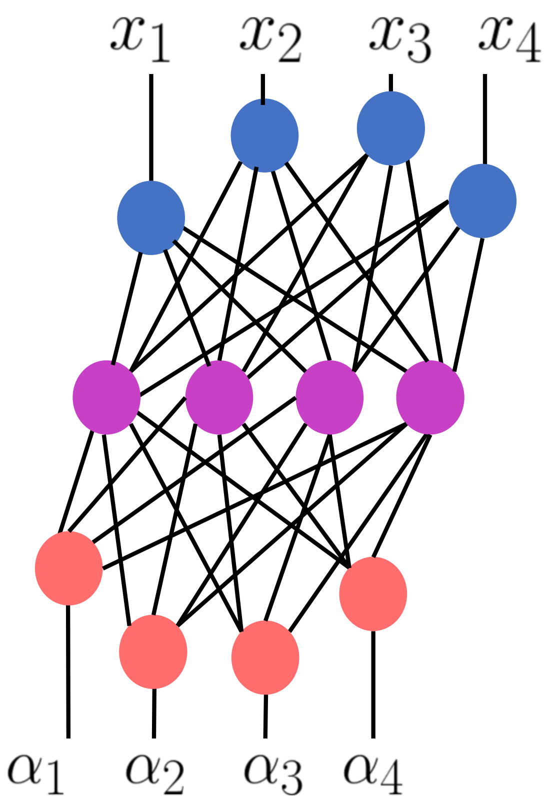
$$L = \sum |\hat{\alpha}_i(x) - \alpha_i|^2$$

Loss function(al) Inferenced from observables Truth value

Other "likelihood free" (binless) algorithms also apply. Not as good for very small θ range

See also [Chen, Glioti, Panico, Wulzer, 2007.10356, [Ambrosio, Hoeve, Madigan, Rojo, Sanz, 2211.02058]

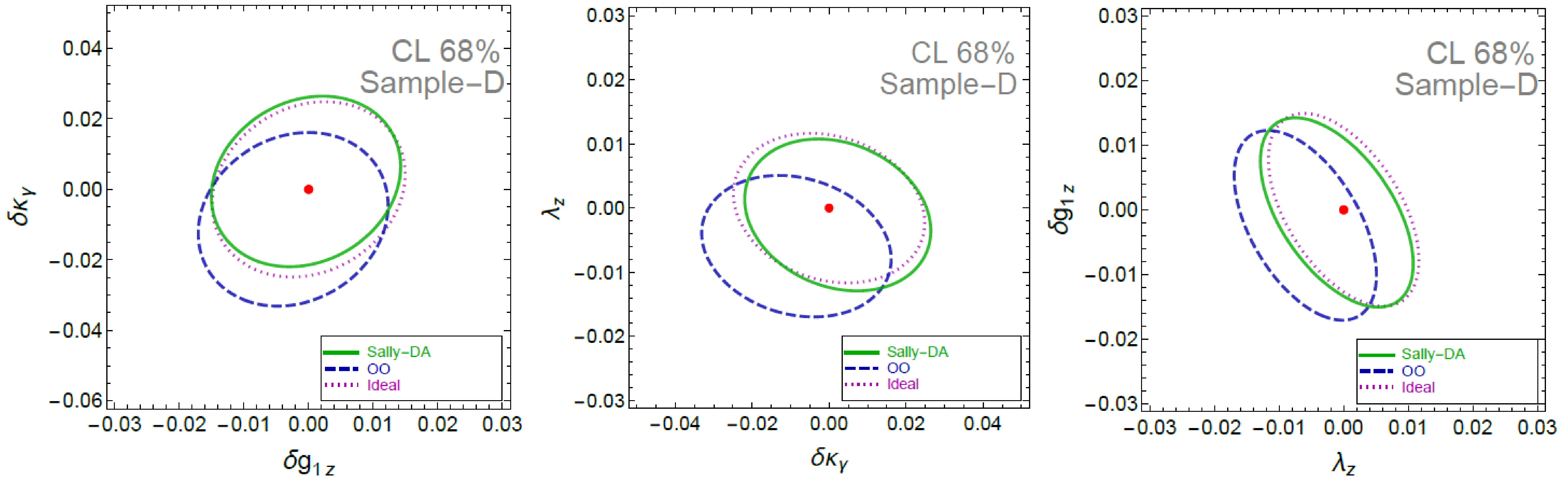
INFERENCE FROM NEURON NETWORKS



Learnt post-parton stages

- Only basic network structure and inputs
- Fully connected (vanilla) network, 9 layers with 200 node each
- Inputs: clustered jet momenta, lepton momentum and a few derived values
- 21 input dimensions
- 1/10 events are 4 fermion background (semileptonic ZZ events with unidentified lepton)

AGAINST ISR & FINITE RESOLUTION

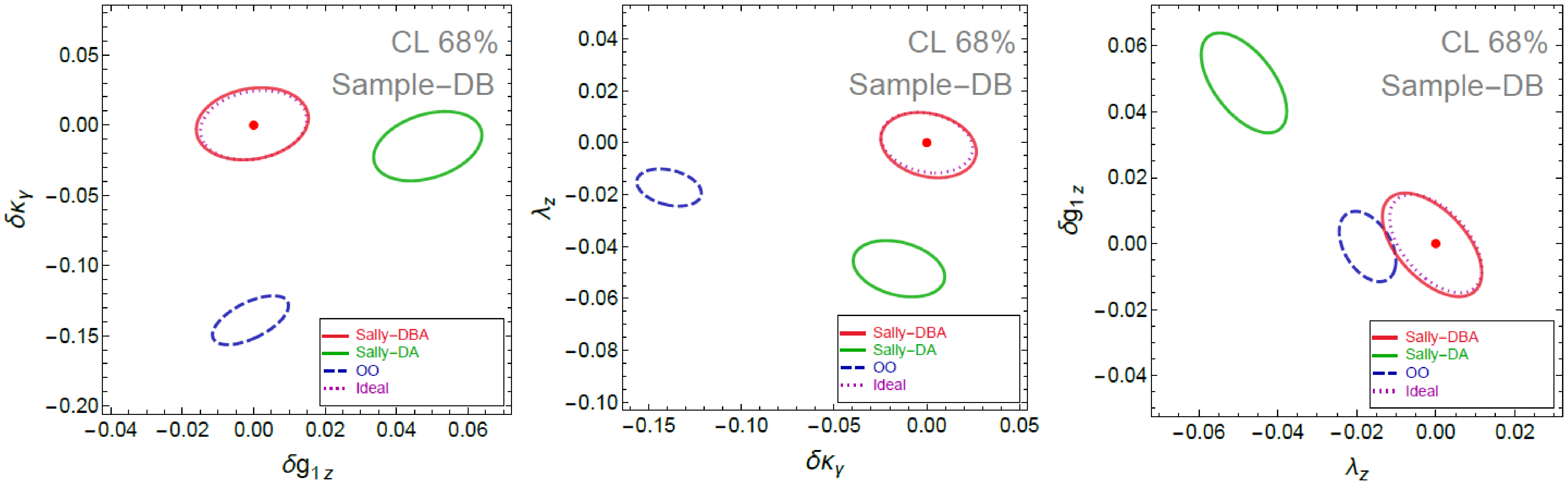


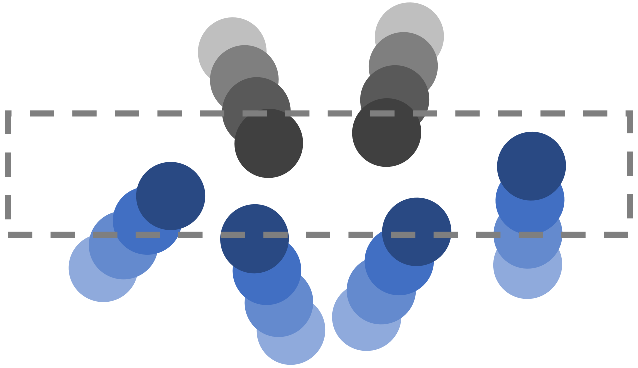
The ideal case, or OO at the truth level
Maximal information use

The OO that knows nothing about ISR and detector

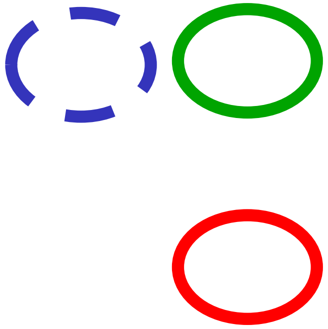
The NN w/ SALLY algorithm that removes systematics

INJECTING BACKGROUND NOISE





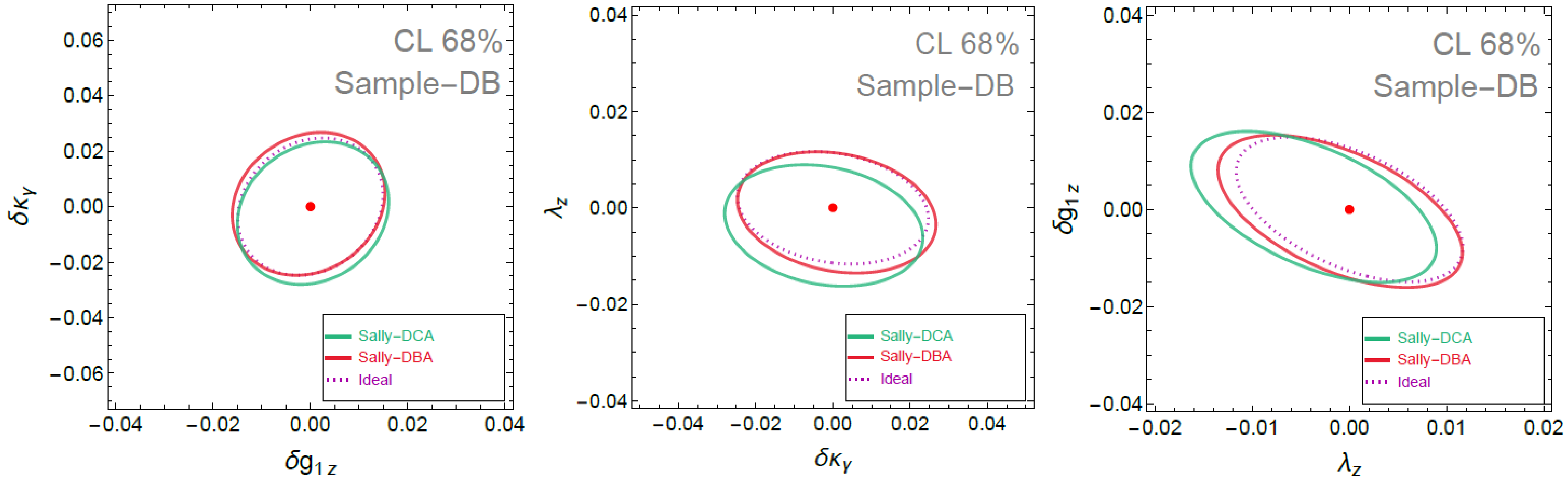
Backgrounds injected due to missing lepton in detector (~10% total)



Never seen backgrounds before

NN ensemble trained with backgrounds

TWO WAYS AGAINST BACKGROUND



Training the SALLY algorithm with background events

One single network system

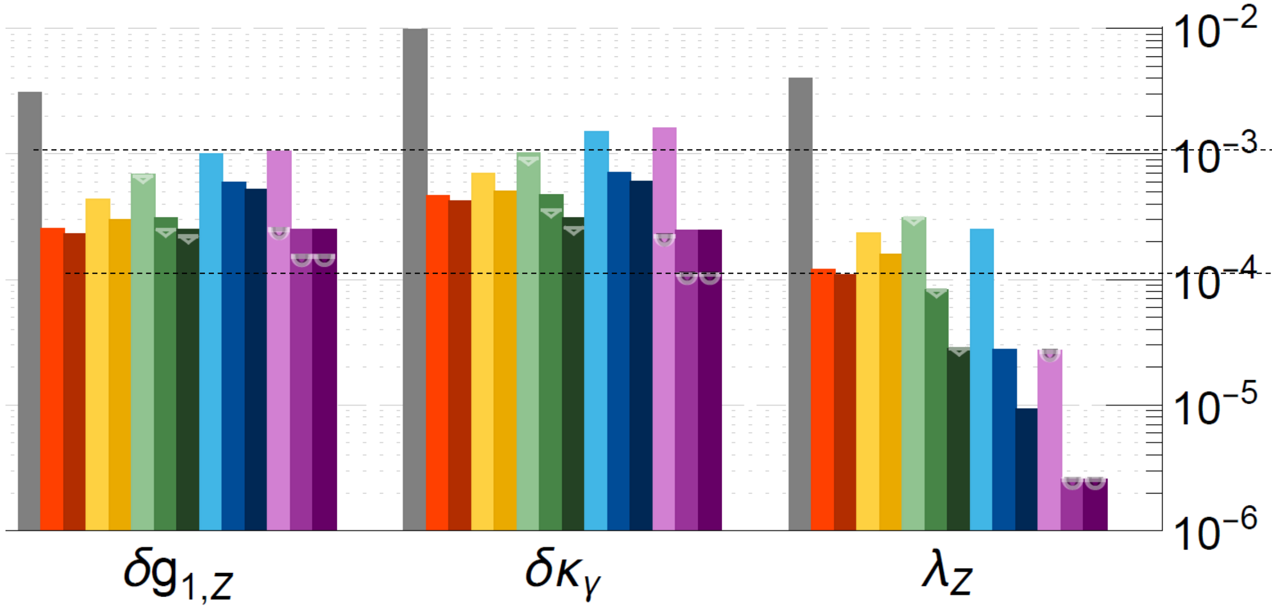
Ideal Training

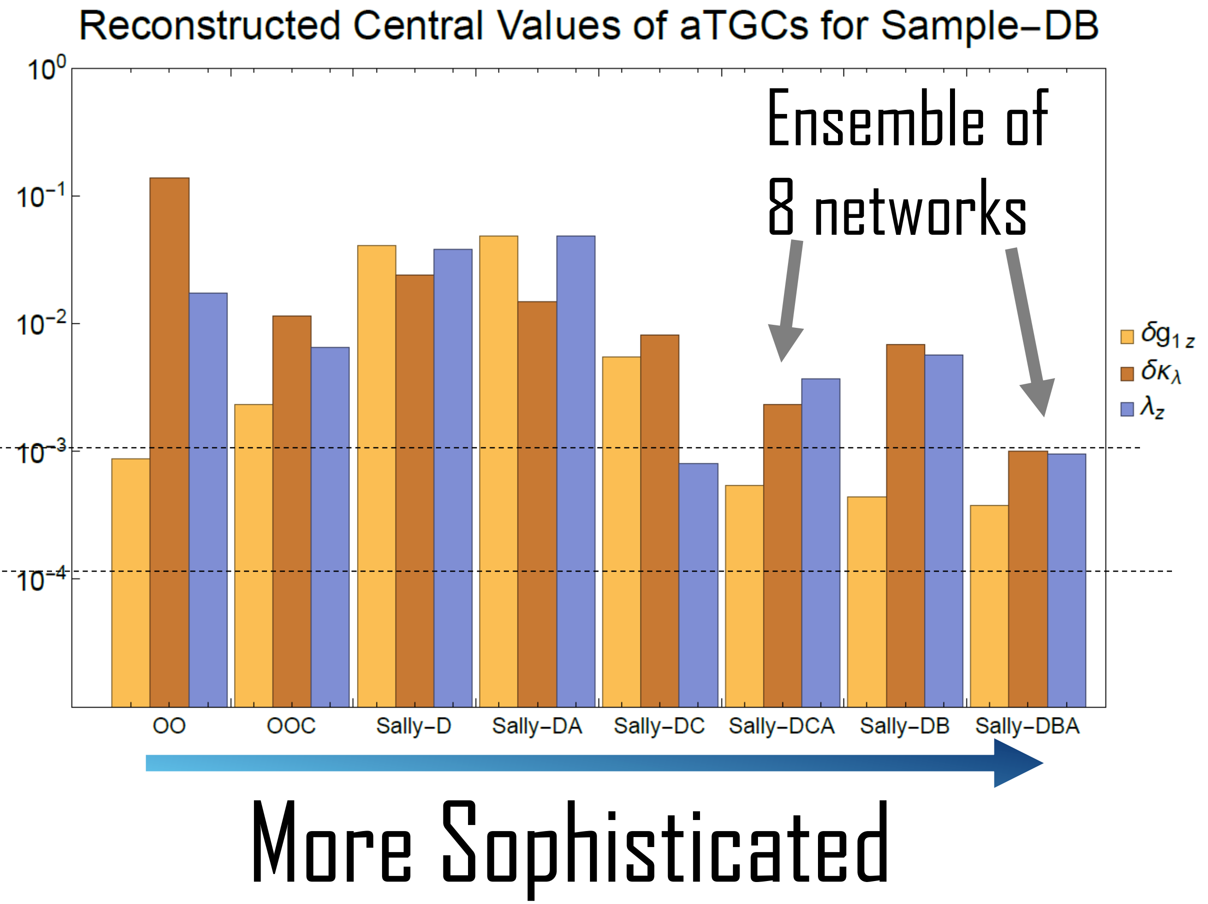
Excluding the background by a classifier.

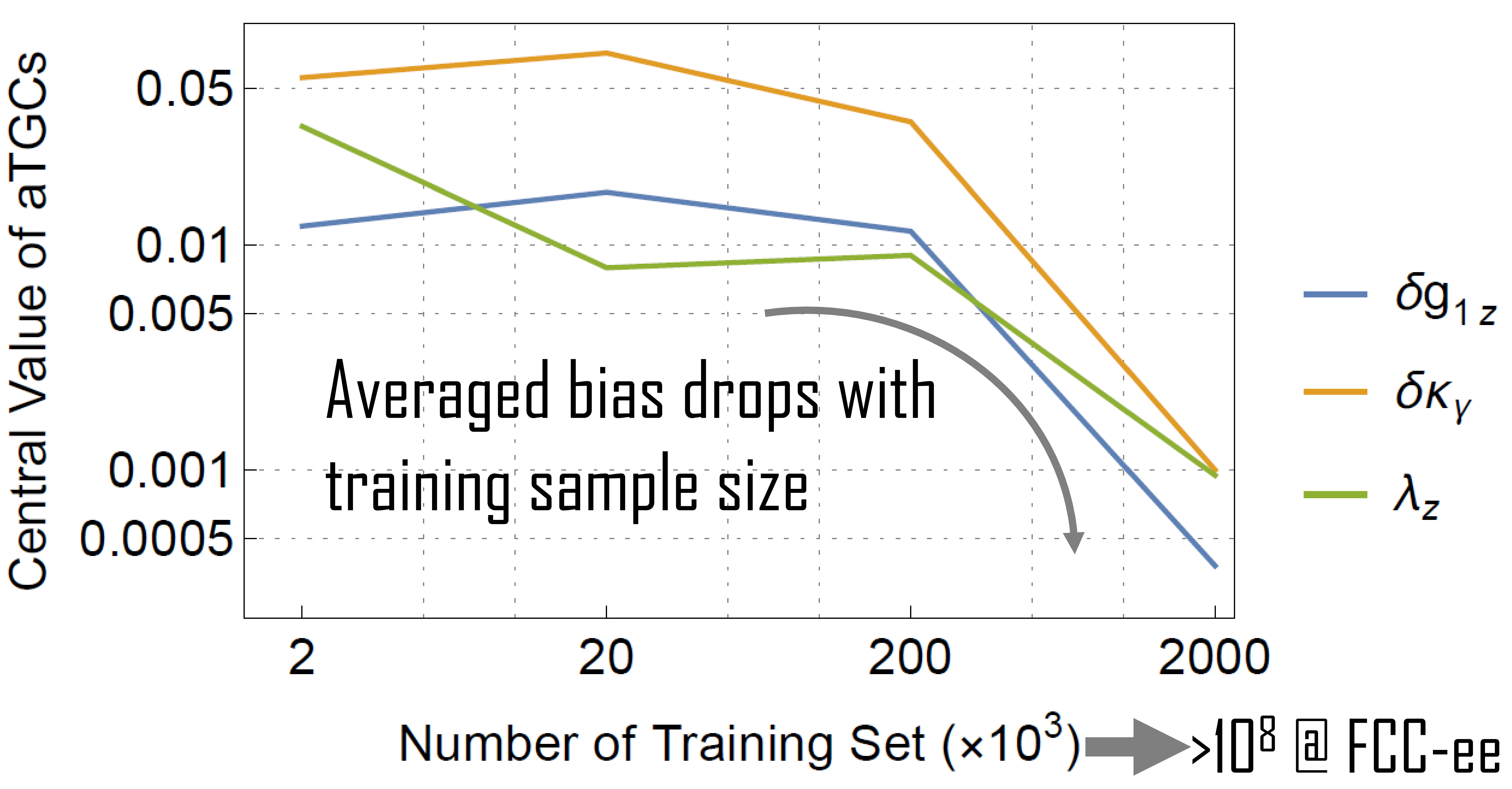
Work separated in two networks

STABILIZE PERFORMANCE

Training sample size: $2 \cdot 10^6$, much smaller than 10^8 expected at Higgs factories

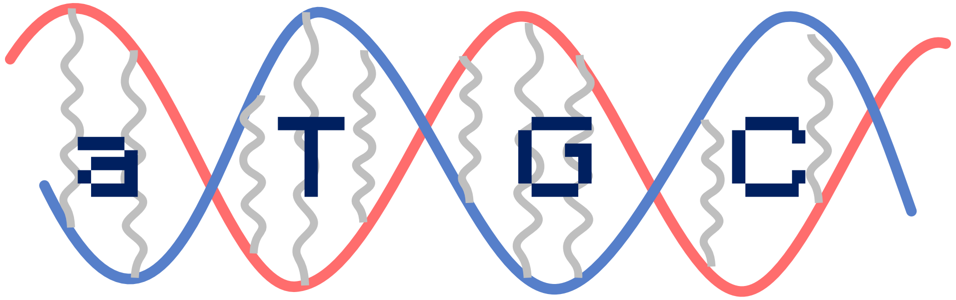





Central Value of aTGCs
0.05
0.01
0.005
0.001
0.0005
2
20
200
2000
Number of Training Set (×10³)
>10⁸ @ FCC-ee
Averaged bias drops with training sample size
δg1 z
δκγ
λz

SUMMARY

- ❑SM precision tests relies on differential measurements (especially aTGC)
- ❑Introduced ML to work against systematic injections
- ❑Binless SALLY algorithm
- ❑Potential to control uncertainties with a small training set



STILL MANY TO DO...

- ❑Apply to other problem at FCC
- ❑More advanced network structure
- ❑ML self-calibration of systematics
- ❑......

THANK YOU