Standard Model Effective Field Theory at Future Lepton Colliders

(with Machine Learning)

Jiayin Gu (顾嘉荫)

Fudan University

Interdisciplinary Center for Theoretical Study
Peng Huanwu Center for Fundamental Theory
USTC
Jun 15, 2023

What is particle physics?

+

=

Quantum Mechanics

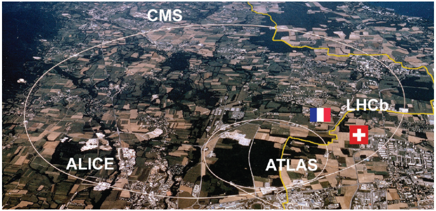
Special Relativity

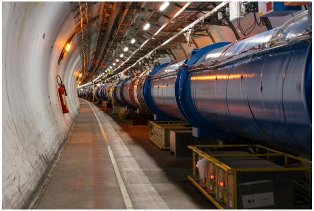
Quantum Field Theory

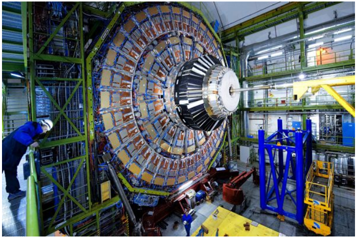
- **Quantum Field Theory** tells us:
  - Particles can be annihilated and created.
  - High energies $\Rightarrow$ heavy (new) particles.

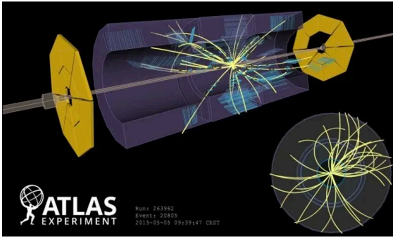
Jiayin Gu (顾嘉荫) Fudan University

particle physics ≈ collider physics





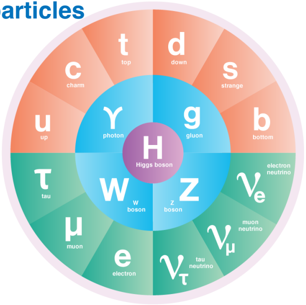




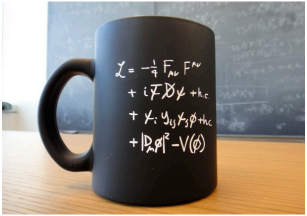
- **Build large colliders → go to high energy → discover new particles!**

The Standard Model

particles

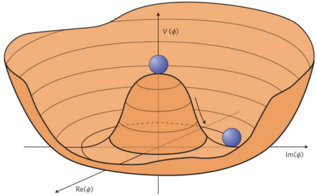


interactions



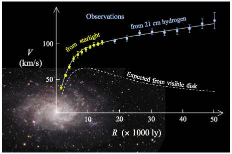
- **the Wheel**
  Only a few “elementary” particles.
- **the Mug**
  The Standard Model Lagrangian is simple!
- **the “Mexican Hat”**
  The Higgs Mechanism gives masses to the elementary particles.

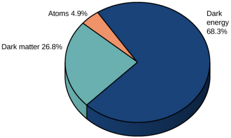
Higgs mechanism

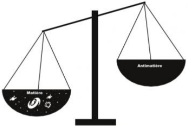


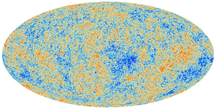
So many things we don't know...

- What is dark matter?
- What is dark energy?
- Why are there more matter than anti-matter?
- What is the origin of neutrino masses?
- What caused the inflation (if it happened)?
- Why is the electroweak scale so much smaller than the Planck scale?
- Why is the strong CP phase θ so small?
- Why is the CKM matrix somewhat close to 1?
- What is the theory of quantum gravity?
- ......









We need experiments to find the answers!

- ~~LHC will find Supersymmetry (or something else), which has a dark matter candidate and solves the Hierarchy problem!~~

- **Higgs and nothing else?**

LHC will definitely find new physics!

We need experiments to find the answers!

- ~~LHC will find Supersymmetry (or something else), which has a dark matter candidate and solves the Hierarchy problem!~~

- **Higgs and nothing else?**

LHC will definitely find new physics!

- **What's next?**
  - Build an even larger collider (~ 100 TeV)?
  - No guaranteed discovery!

We need experiments to find the answers!

- ~~LHC will find Supersymmetry (or something else), which has a dark matter candidate and solves the Hierarchy problem!~~

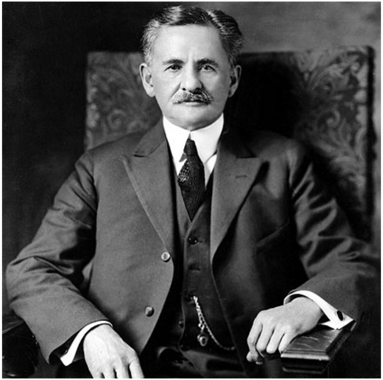
- **Higgs and nothing else?**
- **What's next?**
  - Build an even larger collider (~ 100 TeV)?
  - No guaranteed discovery!

- **Build large colliders** → go to high energy → discover new particles!

  ↘ **do precision measurements → discover new physics indirectly!**

  - **Higgs factory!** (HL-LHC, or a future lepton collider)
  - **S**tandard **M**odel **E**ffective **F**ield **T**heory (model independent approach)

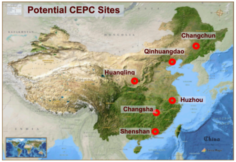
To summarize in one sentence...

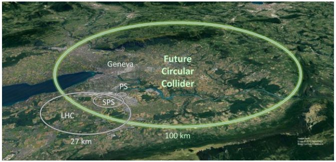


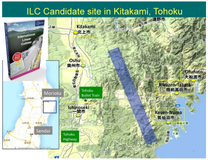
"Our future discoveries must be looked for in the sixth place of decimals."

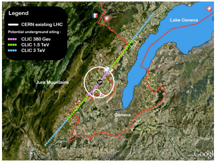
— Albert A. Michelson

Why lepton (e^+e^-) colliders?





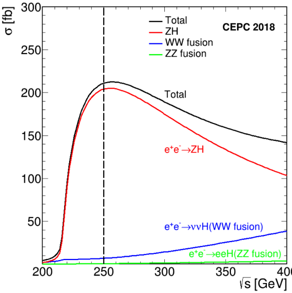


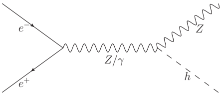


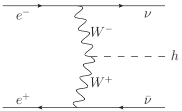
- **It's a Higgs (and Z, W, top) factory!**
  - Large statistics, clean environment
    $\Rightarrow$ **precision measurements!**
  - On the other hand, the LHC is designed to be a "discovery machine"...
- Circular vs. Linear
  - Circular: large luminosity, reuse the tunnel for a 100 TeV hadron collider.
  - Linear: high energy (up to a few TeVs), beam polarization.
- **Muon colliders?**

Why lepton (e^+e^-) colliders?

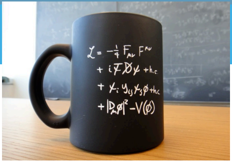
- Higgs
  - $e^+e^- \to hZ$
    cross section maximized at around 250 GeV
  - $e^+e^- \to \nu\bar{\nu}h$
    cross section increases with energy
  - $e^+e^- \to t\bar{t}h$,
    $e^+e^- \to Zhh$, $e^+e^- \to \nu\bar{\nu}hh$,
    ...
- and more
  - $e^+e^- \to Z \to f\bar{f}$
    Z-pole
  - $e^+e^- \to WW$
    WW threshold and above
  - $e^+e^- \to t\bar{t}$
    $t\bar{t}$ threshold and above







The Standard Model Effective Field Theory

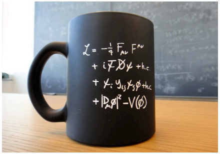


- $[\mathcal{L}_{\rm sm}] \leq 4$. Why?
  - ~~Bad things happen when we have non-renormalizable operators!~~
  - Everything is fine as long as we are happy with finite precision in perturbative calculation.
- **d=5:** $\frac{c}{\Lambda} LLHH \sim \frac{cv^2}{\Lambda} \nu\nu$, Majorana neutrino mass.
- Assuming Baryon and Lepton numbers are conserved,

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_j \frac{c_j^{(8)}}{\Lambda^4} \mathcal{O}_j^{(8)} + \cdots .$$

- If $\Lambda \gg v, E$, then **SM + dimension-6 operators** are sufficient to parameterize the physics around the electroweak scale.

The Standard Model Effective Field Theory



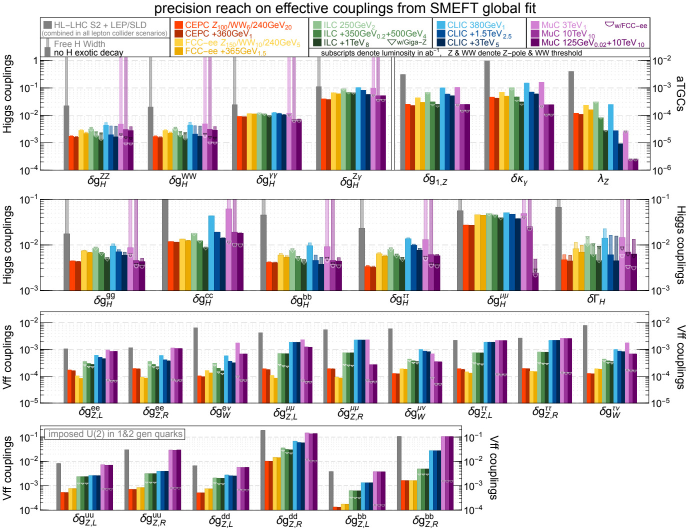
X^3		φ^6 and $\varphi^4 D^2$		$\psi^2\varphi^3$	
Q_G	$f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	Q_φ	$(\varphi^\dagger\varphi)^3$	$Q_{e\varphi}$	$(\varphi^\dagger\varphi)(\bar l_p e_r \varphi)$
$Q_{\widetilde G}$	$f^{ABC} \widetilde G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi\Box}$	$(\varphi^\dagger\varphi)\Box(\varphi^\dagger\varphi)$	$Q_{u\varphi}$	$(\varphi^\dagger\varphi)(\bar q_p u_r \widetilde\varphi)$
Q_W	$\varepsilon^{IJK} W_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi D}$	$\left(\varphi^\dagger D^\mu\varphi\right)^\star \left(\varphi^\dagger D_\mu\varphi\right)$	$Q_{d\varphi}$	$(\varphi^\dagger\varphi)(\bar q_p d_r \varphi)$
$Q_{\widetilde W}$	$\varepsilon^{IJK} \widetilde W_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$				
$X^2\varphi^2$		$\psi^2 X\varphi$		$\psi^2\varphi^2 D$	
$Q_{\varphi G}$	$\varphi^\dagger\varphi\, G^A_{\mu\nu} G^{A\mu\nu}$	Q_{eW}	$(\bar l_p \sigma^{\mu\nu} e_r)\tau^I \varphi W^I_{\mu\nu}$	$Q_{\varphi l}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar l_p \gamma^\mu l_r)$
$Q_{\varphi\widetilde G}$	$\varphi^\dagger\varphi\, \widetilde G^A_{\mu\nu} G^{A\mu\nu}$	Q_{eB}	$(\bar l_p \sigma^{\mu\nu} e_r)\varphi B_{\mu\nu}$	$Q_{\varphi l}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi)(\bar l_p \tau^I \gamma^\mu l_r)$
$Q_{\varphi W}$	$\varphi^\dagger\varphi\, W^I_{\mu\nu} W^{I\mu\nu}$	Q_{uG}	$(\bar q_p \sigma^{\mu\nu} T^A u_r)\widetilde\varphi\, G^A_{\mu\nu}$	$Q_{\varphi e}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar e_p \gamma^\mu e_r)$
$Q_{\varphi\widetilde W}$	$\varphi^\dagger\varphi\, \widetilde W^I_{\mu\nu} W^{I\mu\nu}$	Q_{uW}	$(\bar q_p \sigma^{\mu\nu} u_r)\tau^I \widetilde\varphi\, W^I_{\mu\nu}$	$Q_{\varphi q}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar q_p \gamma^\mu q_r)$
$Q_{\varphi B}$	$\varphi^\dagger\varphi\, B_{\mu\nu} B^{\mu\nu}$	Q_{uB}	$(\bar q_p \sigma^{\mu\nu} u_r)\widetilde\varphi\, B_{\mu\nu}$	$Q_{\varphi q}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi)(\bar q_p \tau^I \gamma^\mu q_r)$
$Q_{\varphi\widetilde B}$	$\varphi^\dagger\varphi\, \widetilde B_{\mu\nu} B^{\mu\nu}$	Q_{dG}	$(\bar q_p \sigma^{\mu\nu} T^A d_r)\varphi\, G^A_{\mu\nu}$	$Q_{\varphi u}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar u_p \gamma^\mu u_r)$
$Q_{\varphi WB}$	$\varphi^\dagger\tau^I\varphi\, W^I_{\mu\nu} B^{\mu\nu}$	Q_{dW}	$(\bar q_p \sigma^{\mu\nu} d_r)\tau^I \varphi\, W^I_{\mu\nu}$	$Q_{\varphi d}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar d_p \gamma^\mu d_r)$
$Q_{\varphi\widetilde WB}$	$\varphi^\dagger\tau^I\varphi\, \widetilde W^I_{\mu\nu} B^{\mu\nu}$	Q_{dB}	$(\bar q_p \sigma^{\mu\nu} d_r)\varphi\, B_{\mu\nu}$	$Q_{\varphi ud}$	$i(\widetilde\varphi^\dagger D_\mu\varphi)(\bar u_p \gamma^\mu d_r)$

$(\bar LL)(\bar LL)$		$(\bar RR)(\bar RR)$		$(\bar LL)(\bar RR)$	
Q_{ll}	$(\bar l_p \gamma_\mu l_r)(\bar l_s \gamma^\mu l_t)$	Q_{ee}	$(\bar e_p \gamma_\mu e_r)(\bar e_s \gamma^\mu e_t)$	Q_{le}	$(\bar l_p \gamma_\mu l_r)(\bar e_s \gamma^\mu e_t)$
$Q_{qq}^{(1)}$	$(\bar q_p \gamma_\mu q_r)(\bar q_s \gamma^\mu q_t)$	Q_{uu}	$(\bar u_p \gamma_\mu u_r)(\bar u_s \gamma^\mu u_t)$	Q_{lu}	$(\bar l_p \gamma_\mu l_r)(\bar u_s \gamma^\mu u_t)$
$Q_{qq}^{(3)}$	$(\bar q_p \gamma_\mu \tau^I q_r)(\bar q_s \gamma^\mu \tau^I q_t)$	Q_{dd}	$(\bar d_p \gamma_\mu d_r)(\bar d_s \gamma^\mu d_t)$	Q_{ld}	$(\bar l_p \gamma_\mu l_r)(\bar d_s \gamma^\mu d_t)$
$Q_{lq}^{(1)}$	$(\bar l_p \gamma_\mu l_r)(\bar q_s \gamma^\mu q_t)$	Q_{eu}	$(\bar e_p \gamma_\mu e_r)(\bar u_s \gamma^\mu u_t)$	Q_{qe}	$(\bar q_p \gamma_\mu q_r)(\bar e_s \gamma^\mu e_t)$
$Q_{lq}^{(3)}$	$(\bar l_p \gamma_\mu \tau^I l_r)(\bar q_s \gamma^\mu \tau^I q_t)$	Q_{ed}	$(\bar e_p \gamma_\mu e_r)(\bar d_s \gamma^\mu d_t)$	$Q_{qu}^{(1)}$	$(\bar q_p \gamma_\mu q_r)(\bar u_s \gamma^\mu u_t)$
		$Q_{ud}^{(1)}$	$(\bar u_p \gamma_\mu u_r)(\bar d_s \gamma^\mu d_t)$	$Q_{qu}^{(8)}$	$(\bar q_p \gamma_\mu T^A q_r)(\bar u_s \gamma^\mu T^A u_t)$
		$Q_{ud}^{(8)}$	$(\bar u_p \gamma_\mu T^A u_r)(\bar d_s \gamma^\mu T^A d_t)$	$Q_{qd}^{(1)}$	$(\bar q_p \gamma_\mu q_r)(\bar d_s \gamma^\mu d_t)$
				$Q_{qd}^{(8)}$	$(\bar q_p \gamma_\mu T^A q_r)(\bar d_s \gamma^\mu T^A d_t)$
$(\bar LR)(\bar RL)$ and $(\bar LR)(\bar LR)$		B-violating			
Q_{ledq}	$(\bar l_p^j e_r)(\bar d_s q_t^j)$	Q_{duq}	$\varepsilon^{\alpha\beta\gamma}\varepsilon_{jk}\left[(d_p^\alpha)^T C u_r^\beta\right]\left[(q_s^{\gamma j})^T C l_t^k\right]$		
$Q_{quqd}^{(1)}$	$(\bar q_p^j u_r)\varepsilon_{jk}(\bar q_s^k d_t)$	Q_{qqu}	$\varepsilon^{\alpha\beta\gamma}\varepsilon_{jk}\left[(q_p^{\alpha j})^T C q_r^{\beta k}\right]\left[(u_s^\gamma)^T C e_t\right]$		
$Q_{quqd}^{(8)}$	$(\bar q_p^j T^A u_r)\varepsilon_{jk}(\bar q_s^k T^A d_t)$	$Q_{qqq}^{(1)}$	$\varepsilon^{\alpha\beta\gamma}\varepsilon_{jk}\varepsilon_{mn}\left[(q_p^{\alpha j})^T C q_r^{\beta k}\right]\left[(q_s^{\gamma m})^T C l_t^n\right]$		
$Q_{lequ}^{(1)}$	$(\bar l_p^j e_r)\varepsilon_{jk}(\bar q_s^k u_t)$	$Q_{qqq}^{(3)}$	$\varepsilon^{\alpha\beta\gamma}(\tau^I\varepsilon)_{jk}(\tau^I\varepsilon)_{mn}\left[(q_p^{\alpha j})^T C q_r^{\beta k}\right]\left[(q_s^{\gamma m})^T C l_t^n\right]$		
$Q_{lequ}^{(3)}$	$(\bar l_p^j \sigma_{\mu\nu} e_r)\varepsilon_{jk}(\bar q_s^k \sigma^{\mu\nu} u_t)$	Q_{duu}	$\varepsilon^{\alpha\beta\gamma}\left[(d_p^\alpha)^T C u_r^\beta\right]\left[(u_s^\gamma)^T C e_t\right]$		

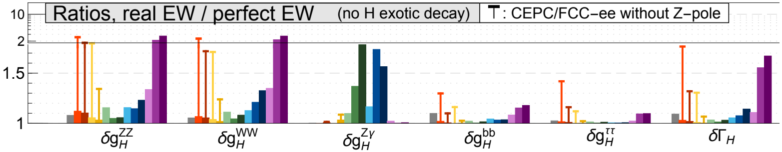
- Write down all possible (non-redundant) dimension-6 operators ...
- 59 operators (76 parameters) for 1 generation, or 2499 parameters for 3 generations. [arXiv:1008.4884] Grzadkowski, Iskrzyński, Misiak, Rosiek, [arXiv:1312.2014] Alonso, Jenkins, Manohar, Trott.
- A **full global fit** with all measurements to all operator coefficients?
  - We usually only need to deal with a subset of them, *e.g.* $\sim 20\text{-}30$ parameters for **Higgs and electroweak** measurements.
- Do a global fit and present the results with some fancy bar plots!

Higgs + EW, Results from the Snowmass 2021 (2022) study

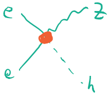
[2206.08326] de Blas, Du, Grojean, JG, Miralles, Peskin, Tian, Vos, Vryonidou



Impacts of (lack of) the Z-pole run



- Without good *Z*-pole measurements, the *eeZh* contact interaction may have a significant impact on the Higgs coupling determination.
- Current (LEP) *Z*-pole measurements are not good enough for CEPC/FCC-ee Higgs measurements!
  - **A future Z-pole run is important!**



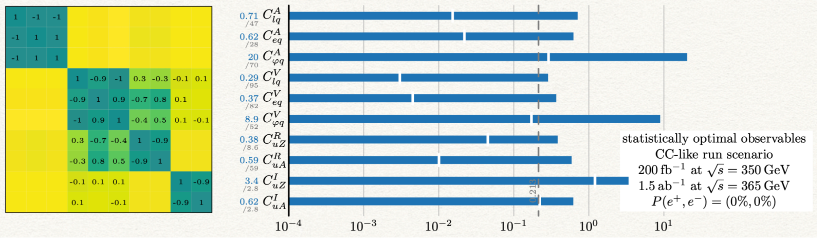
- Linear colliders suffer less from the lack of a Z-pole run. **(Win Win!)**

Probing Top operators with $e^- e^+ \to t\bar{t}$ [arXiv:1807.02121] Durieux, Perelló, Vos, Zhang

$$
\begin{aligned}
O^1_{\varphi q} &\equiv \tfrac{y_t^2}{2}\ \bar{q}\gamma^\mu q\ \ \varphi^\dagger i\overleftrightarrow{D}_\mu\varphi, & O_{uG} &\equiv y_t g_s\ \bar{q}T^A\sigma^{\mu\nu}u\ \ \epsilon\varphi^* G^A_{\mu\nu},\\
O^3_{\varphi q} &\equiv \tfrac{y_t^2}{2}\ \bar{q}\tau^I\gamma^\mu q\ \ \varphi^\dagger i\overleftrightarrow{D}^I_\mu\varphi, & O_{uW} &\equiv y_t g_W\ \bar{q}\tau^I\sigma^{\mu\nu}u\ \ \epsilon\varphi^* W^I_{\mu\nu},\\
O_{\varphi u} &\equiv \tfrac{y_t^2}{2}\ \bar{u}\gamma^\mu u\ \ \varphi^\dagger i\overleftrightarrow{D}_\mu\varphi, & O_{dW} &\equiv y_t g_W\ \bar{q}\tau^I\sigma^{\mu\nu}d\ \ \epsilon\varphi^* W^I_{\mu\nu},\\
O_{\varphi ud} &\equiv \tfrac{y_t^2}{2}\ \bar{u}\gamma^\mu d\ \ \varphi^T\epsilon\, iD_\mu\varphi, & O_{uB} &\equiv y_t g_Y\ \bar{q}\sigma^{\mu\nu}u\ \ \epsilon\varphi^* B_{\mu\nu},
\end{aligned}
$$

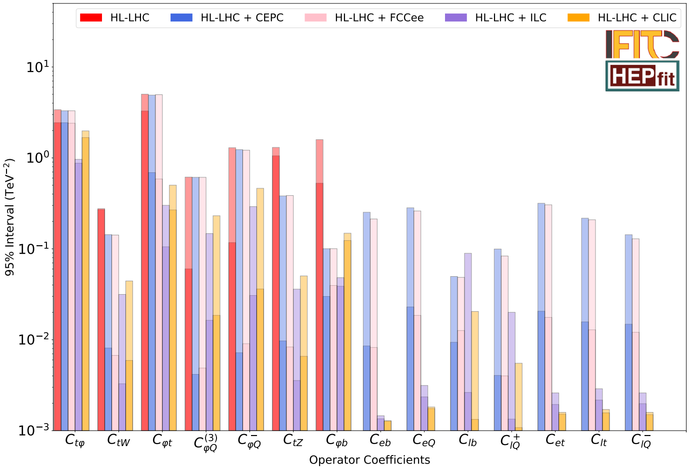
$$
\begin{aligned}
O^1_{lq} &\equiv \tfrac{1}{2}\ \bar{q}\gamma_\mu q\ \ \bar{l}\gamma^\mu l,\\
O^3_{lq} &\equiv \tfrac{1}{2}\ \bar{q}\tau^I\gamma_\mu q\ \ \bar{l}\tau^I\gamma^\mu l,\\
O_{lu} &\equiv \tfrac{1}{2}\ \bar{u}\gamma_\mu u\ \ \bar{l}\gamma^\mu l,\\
O_{eq} &\equiv \tfrac{1}{2}\ \bar{q}\gamma_\mu q\ \ \bar{e}\gamma^\mu e,\\
O_{eu} &\equiv \tfrac{1}{2}\ \bar{u}\gamma_\mu u\ \ \bar{e}\gamma^\mu e,
\end{aligned}
$$

- Also need to include **top dipole** interactions and *eett* contact interactions!
- Hard to resolve the top couplings from $4f$ interactions with just the 365 GeV run.
  - Can't really separate $e^+e^- \to Z/\gamma \to t\bar{t}$ from $e^+e^- \to Z' \to t\bar{t}$.
  - Is that a big deal?



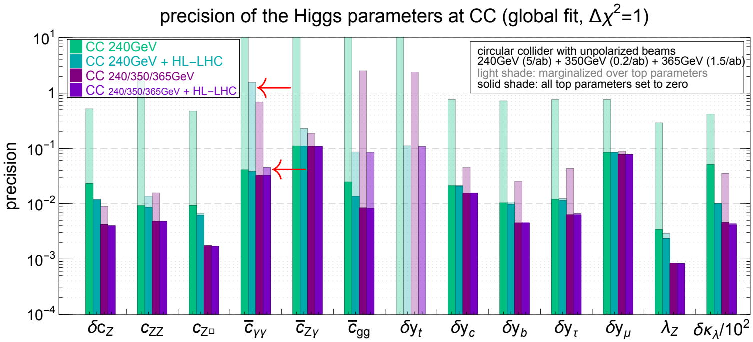
Results from the recent snowmass study

[2206.08326] de Blas, Du, Grojean, JG, Miralles, Peskin, Tian, Vos, Vryonidou

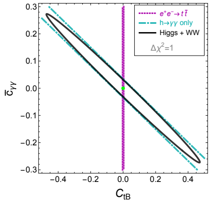


Top operators in loops (Higgs processes)

[1809.03520] G. Durieux, JG, E. Vryonidou, C. Zhang



- $O_{tB} = (\bar{Q}\sigma^{\mu\nu} t)\, \tilde{\varphi} B_{\mu\nu} + h.c.$ is not very well constrained at the LHC, and it generates dipole interactions that contributes to the $h\gamma\gamma$ vertex.
- Deviations in $h\gamma\gamma$ coupling $\Rightarrow$ run at ~ 365 GeV to confirm?



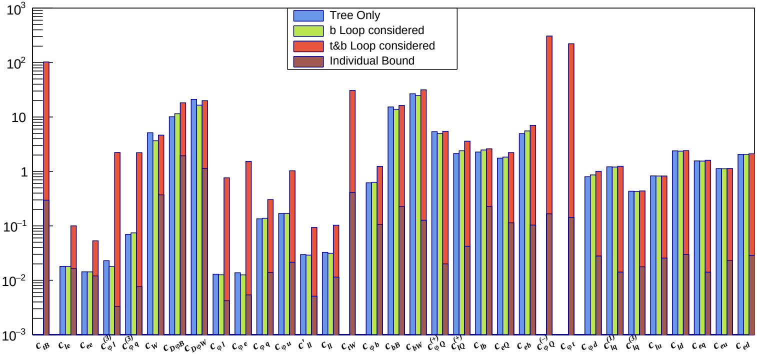
Top operators in loops (current EW processes)

[2205.05655] Y. Liu, Y. Wang, C. Zhang, L. Zhang, JG

	Experiment	Observables
Low Energy	CHARM/CDHS/ CCFR/NuTeV/ APV/QWEAK/ PVDIS	Effective Couplings
Z-pole	LEP/SLC	Total decay width Γ_Z Hadronic cross-section σ_{had} Ratio of decay width R_f Forward-Backward Asymmetry A^f_{FB} Polarized Asymmetry A_f
W-pole	LHC/Tevatron/ LEP/SLC	Total decay width Γ_W W branching ratios $Br(W \to l\nu_l)$ Mass of W Boson M_W
$ee \to qq$	LEP/TRISTAN	Hadronic cross-section σ_{had} Ratio of cross-section R_f Forward-Backward Asymmetry for b/c A^f_{FB}
$ee \to ll$	LEP	cross-section σ_f Forward-Backward Asymmetry A^f_{FB} Differential cross-section $\frac{d\sigma_f}{dcos\theta}$
$ee \to WW$	LEP	cross-section σ_{WW} Differential cross-section $\frac{d\sigma_{WW}}{dcos\theta}$

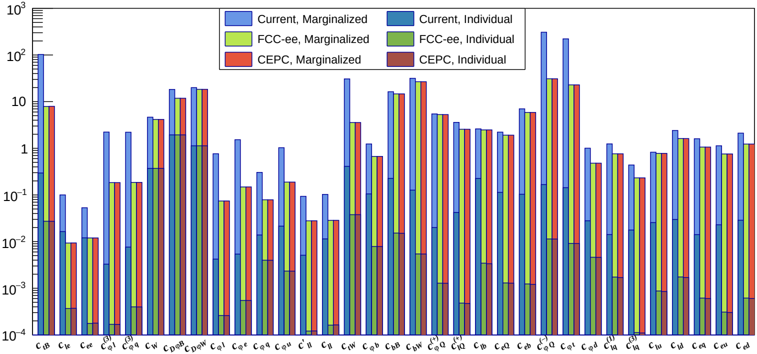
- Top operators (1-loop) + EW operators (tree, including bottom dipole operators)
- $e^+e^- \to f\bar{f}$ at different energies, $e^+e^- \to WW$.

Top operators in loops (current EW processes)



- Good sensitivities, but too many parameters for a global fit...

Top operators in loops (future EW processes)



- Good sensitivities, but too many parameters for a global fit...
- It shows the importance of directly measuring $e^+e^- \to t\bar{t}$.

Machine learning in SMEFT analyses

Machine learning is not physics!

past

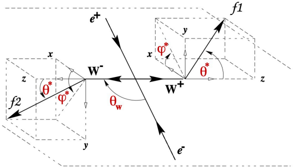
now

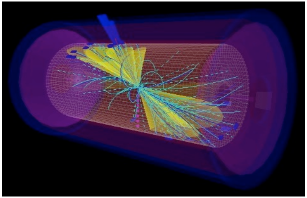
- Current work with Shengdu Chai (柴声都), Lingfeng Li (李凌风) on $e^+e^- \to WW$.
- Current work with Yifan Fei (费昳帆), Tong Shen (沈同) and Kerun Yu (余柯润) on $e^+e^- \to t\bar{t}$.

Why Machine learning?

- In many cases, the new physics contributions are sensitive to the differential distributions.
  - $e^+e^- \to WW \to 4f \quad \Rightarrow \quad$ 5 angles
  - $e^+e^- \to t\bar{t} \to bW^+\bar{b}W^- \to 6f$
    $\Rightarrow$ 9 angles
  - How to extract information from the differential distribution?
  - If we have the full knowledge of $\frac{d\sigma}{d\Omega} \Rightarrow$ matrix-element method, optimal observables...

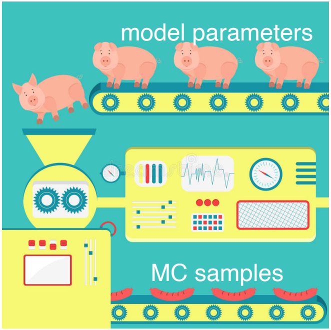
- The ideal $\frac{d\sigma}{d\Omega}$ we can calculate is not the $\frac{d\sigma}{d\Omega}$ that we actually measure!
  - detector acceptance, measurement uncertainties, ISR/beamstrahlung ...
  - In practice we only have **MC samples**, not analytic expressions, for $\frac{d\sigma}{d\Omega}$.



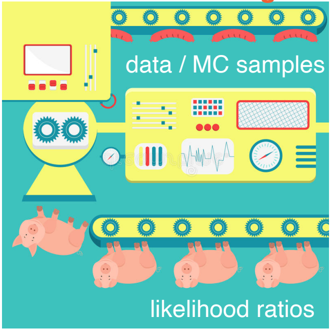


The “inverse problem”

Monte Carlo simulation



Machine Learning



- **Forward:** From model parameters we can calculate the ideal $\frac{d\sigma}{d\Omega}$, simulate complicated effects and produce MC samples.
- **Inverse:** From data / MC samples, how do we know the model parameters?
- With **Neural Network** we can (in principle) reconstruct $\frac{d\sigma}{d\Omega}$ (or likelihood ratios) from MC samples.

A rough sketch

- We have a theory (SMEFT) that gives a differential cross section $\frac{d\sigma}{d\Omega}$ which is a function of the parameters of interest $\mathbf{c}$ (Wilson coefficients).
  - For simplicity, let's ignore the total rate and focus on $\frac{1}{\sigma}\frac{d\sigma}{d\Omega} \equiv p(\mathbf{x}|\mathbf{c})$, *i.e.* it's a probability density function of the observables $\mathbf{x}$.
  - Define the likelihood function $\mathcal{L}(\mathbf{c}|\mathbf{x}) \equiv p(\mathbf{x}|\mathbf{c})$. For a sample of N events, maximizing the joint likelihood $\prod_{i=1}^{N} \mathcal{L}(\mathbf{c}|\mathbf{x}_i)$ (or the log likelihood) gives the best estimator for $\mathbf{c}$. (matrix-element method)

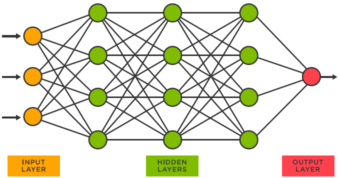
- Suppose we have two equal-size samples $\{\mathbf{x}_{i,\mathbf{c}_0}\} \sim p(\mathbf{x}|\mathbf{c}_0)$ and $\{\mathbf{x}_{i,\mathbf{c}_1}\} \sim p(\mathbf{x}|\mathbf{c}_1)$, one could define the cross-entropy loss function(al)

$$L(\hat{s}) = -\sum_{i=1}^{N} \log \hat{s}(\mathbf{x}_{i,\mathbf{c}_1}) - \sum_{i=1}^{N} \log\left(1 - \hat{s}(\mathbf{x}_{i,\mathbf{c}_0})\right) ,$$

which is minimized by the optimal decision function

$$s(\mathbf{x}|\mathbf{c}_0, \mathbf{c}_1) = \frac{p(\mathbf{x}|\mathbf{c}_1)}{p(\mathbf{x}|\mathbf{c}_0) + p(\mathbf{x}|\mathbf{c}_1)} .$$

A rough sketch



- From neural network we can construct a function $\hat{s}(\mathbf{x})$. By minimizing $L(\hat{s})$ with respect to $\hat{s}(\mathbf{x})$ we can obtain an estimator for the likelihood ratio

$$\hat{r}(\mathbf{x}|\mathbf{c}_0,\mathbf{c}_1) = \frac{1-\hat{s}(\mathbf{x}|\mathbf{c}_0,\mathbf{c}_1)}{\hat{s}(\mathbf{x}|\mathbf{c}_0,\mathbf{c}_1)} = \frac{\hat{p}(\mathbf{x}|\mathbf{c}_0)}{\hat{p}(\mathbf{x}|\mathbf{c}_1)},$$

  which is the same as the true likelihood ratio in the ideal limit (large sample, perfect training).

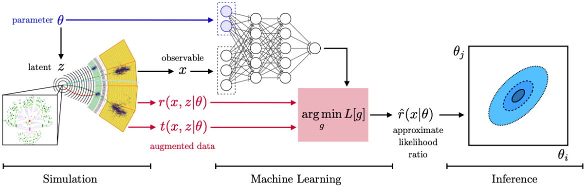
- There are many other ways to construct a loss function(al)....

- With additional assumptions on how $\frac{d\sigma}{d\Omega}$ depends on $\mathbf{c}$ (*i.e.*, a quadratic relation), we only need to train a finite number of times to know how the likelihood ratio depend on $\mathbf{c}$.

Particle physics structure

- One could make use of latent variable "z" (the parton level analytic result for $\frac{d\sigma}{d\Omega}$) to increase the performance of ML.

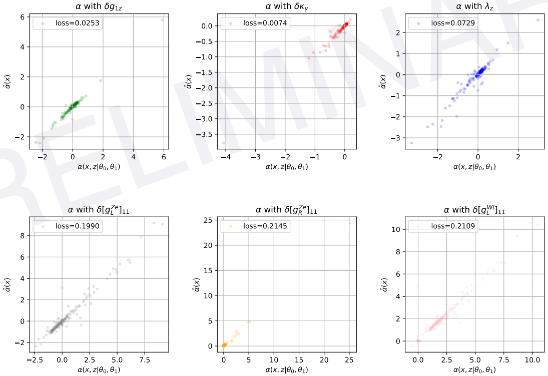
  [1805.00013, 1805.00020] Brehmer, Cranmer, Louppe, Pavez



- Assuming linear dependences $\frac{d\sigma}{d\Omega} = S_0 + \sum_i S_{1,i}\, c_i$, there is a method called SALLY (Score approximates likelihood locally).
  - In this case, for each parameter we only need to train once to obtain $\alpha_i \equiv \frac{S_{1,i}}{S_0}$. (It is basically the ML version of Optimal Observables.)
  - We can calculate the "ideal" $\alpha(z)$ which will help us train the actual $\alpha(x)$.

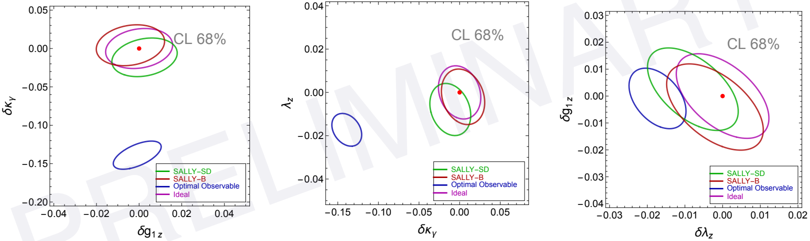
$$L[\hat{\alpha}(x)] = \sum_{x_i, z_i \sim \mathrm{SM}} |\alpha(z_i) - \hat{\alpha}(x_i)|^2 .$$

Machine Learning in $e^+e^- \to WW$ (preliminary results, Shengdu Chai, JG, Lingfeng Li)



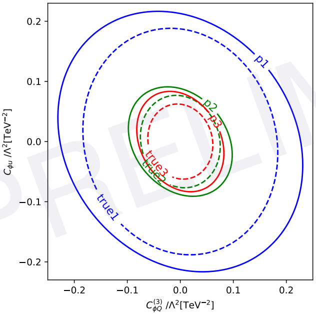
- Semileptonic channel, MadGraph/Pythia/Delphes (CEPC detector card), with *ZZ* backgrounds.

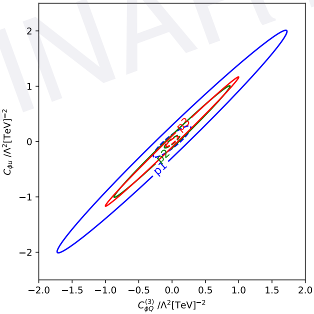
Machine Learning in $e^+e^- \to WW$ (preliminary results, Shengdu Chai, JG, Lingfeng Li)



- 3-aTGC fit, scaled to 10^4 events.
  - OO+classifier: hybrid method that uses a classifier to discriminate background.
- Naively applying truth-level optimal observables could lead to a large bias!
- It's easier for machine learning to take care of systematics!

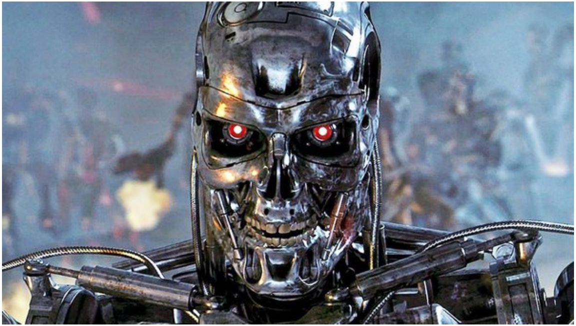
Machine Learning in $e^+e^- \to t\bar{t}$ (very preliminary results, Yifan Fei, JG, Tong Shen, Kerun Yu)





- $e^+e^- \to t\bar{t}$, 3 different channels (no background yet)
- **Left:** $\sqrt{s}$ =1TeV, **Right:** $\sqrt{s}$ =360 GeV

Machine learning



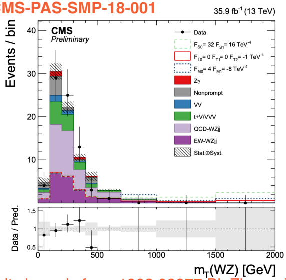
- **When will Machine take over?**
  - Before or after a future lepton collider is built?

Probing dimension-8 operators?

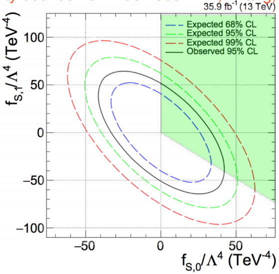
- The dimension-8 contribution has a large energy enhancement ($\sim E^4/\Lambda^4$)!
- It is difficult for LHC to probe these bounds.
  - Low statistics in the high energy bins.
  - Example: Vector boson scattering.
  - $\Lambda \lesssim \sqrt{s}$, the EFT expansion breaks down!
- Can we separate the dim-8 and dim-6 effects?
  - Precision measurements at several different $\sqrt{s}$?

    (A **very** high energy lepton collider?)
  - Or find some special process where dim-8 gives the leading new physics contribution?

CMS-PAS-SMP-18-001



positivity bounds from 1902.08977 Bi, Zhang, Zhou

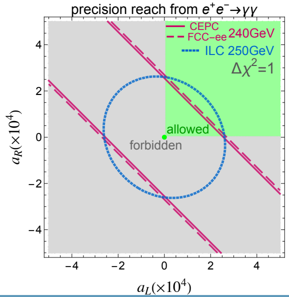


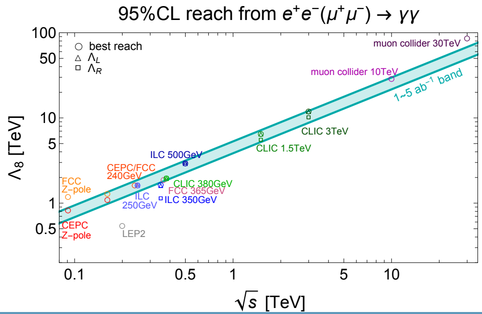
The diphoton channel [arXiv:2011.03055] Phys.Rev.Lett. 129, 011805, JG, Lian-Tao Wang, Cen Zhang

- $e^+e^- \to \gamma\gamma$ (or $\mu^+\mu^- \to \gamma\gamma$), SM, non-resonant.
- Leading order contribution: **dimension-8 contact interaction**.
  ($f^+f^- \to \bar{e}_L e_L$ or $e_R \bar{e}_R$)

$$\mathcal{A}(f^+ f^- \gamma^+ \gamma^-)_{\mathrm{SM+d8}} = 2e^2 \frac{\langle 24 \rangle^2}{\langle 13 \rangle \langle 23 \rangle} + \frac{a}{v^4} [13][23] \langle 24 \rangle^2 \,.$$

- Can probe dim-8 operators (and their positivity bounds) at a **Higgs factory** (~ 240 GeV)!





Conclusion

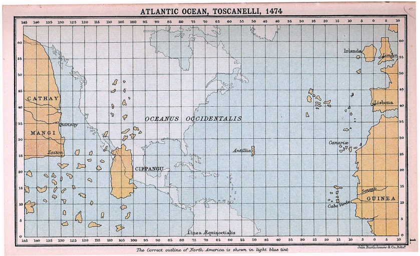
- **We have no idea what is the new physics beyond the Standard Model.**

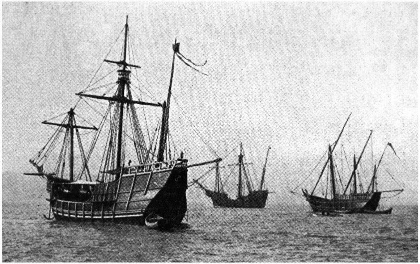
- **One important direction to move forward is to do precision measurements of the Standard Model processes.**
  - A future lepton collider is an ideal machine for that.
  - SMEFT is a good theory framework (but is not everything).
  - Expanding the theory framework?
    - Loop contributions, dimension-8 operators, HEFT ...

- **Machine learning is** (likely to be) **the future!**

A lesson from Christopher Columbus (哥伦布发现美洲大陆)

- **You need to have a theory.**
  - The earth is round, India is in the east...
- **Your theory can be wrong!**
  - Columbus did not find India, but found America instead...
- **You need to ask money from the government!**
  - Columbus convinced the monarchs of Spain to sponsor him.
- **Will we discover the new world?**

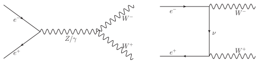


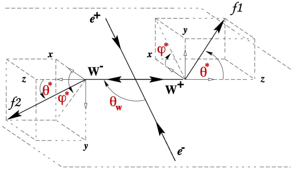


backup slides

$e^+e^- \to WW$ with Optimal Observables

- TGCs (and additional EFT parameters) are sensitive to the differential distributions!
  - One could do a fit to the binned distributions of all angles.
  - Not the most efficient way of extracting information.
  - Correlations among angles are sometimes ignored.



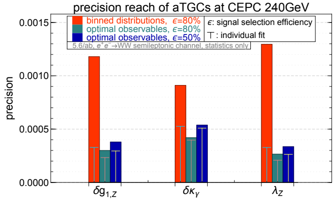


- What are optimal observables?

  (See *e.g.* Z.Phys. C62 (1994) 397-412 Diehl & Nachtmann)
  - In the limit of large statistics (everything is Gaussian) and small parameters (linear contribution dominates), the best possible reaches can be derived analytically!

$$\frac{d\sigma}{d\Omega} = S_0 + \sum_i S_{1,i}\, g_i\,, \qquad c_{ij}^{-1} = \int d\Omega \frac{S_{1,i} S_{1,j}}{S_0} \cdot \mathcal{L}\,,$$

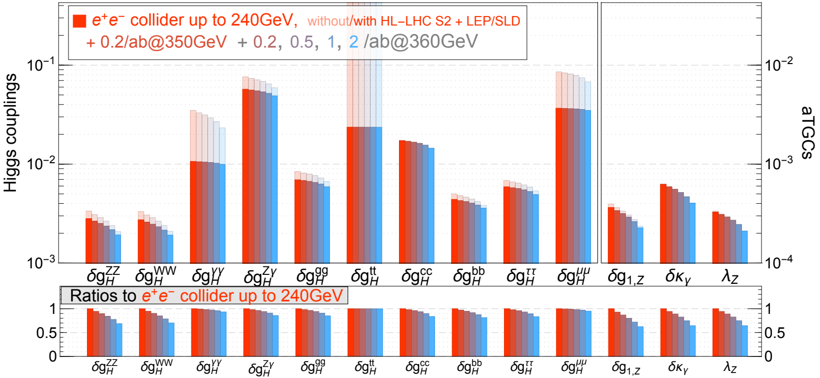
- The optimal observables are given by $\mathcal{O}_i = \frac{S_{1,i}}{S_0}$, and are functions of the 5 angles.



[arXiv:1907.04311] de Blas, Durieux, Grojean, JG, Paul

Impact of a 350/360 GeV run

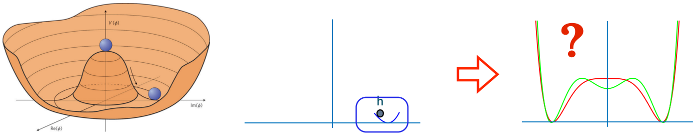
precision reach of the full EFT fit (effective couplings and aTGCs)



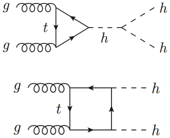
- $5.6\,\mathrm{ab}^{-1}$ at 240 GeV assumed.
- Measurements at 350/360 GeV provides additional handles on the anomalous couplings (*e.g.* $hZ^\mu Z_\mu$ *vs.* $hZ^{\mu\nu}Z_{\mu\nu}$).
- Also improves the measurements of $e^+e^- \to WW$ (aTGCs).

Higgs self-coupling

- We know very little about the Higgs potential!

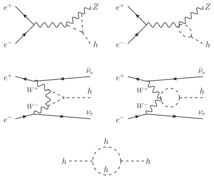


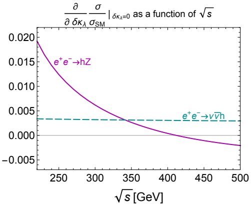
- To know more about the Higgs potential, we need to measure the Higgs self-couplings (**hhh** and **hhhh** couplings).
- The $(H^\dagger H)^3$ operator can modify the Higgs self-couplings.
- Probing the *hhh* coupling at Hadron colliders.
  - $gg \to hh$
  - $\lesssim 50\%$ at HL-LHC.
  - $\lesssim 5\%$ at a 100 TeV collider.

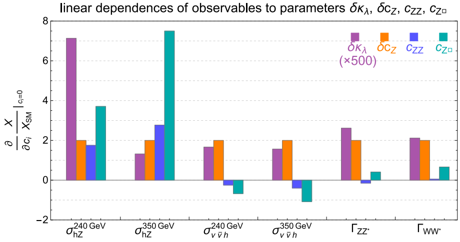


Triple Higgs coupling at one-loop order

[arXiv:1711.03978] Di Vita, Durieux, Grojean, JG, Liu, Panico, Riembau, Vantalon

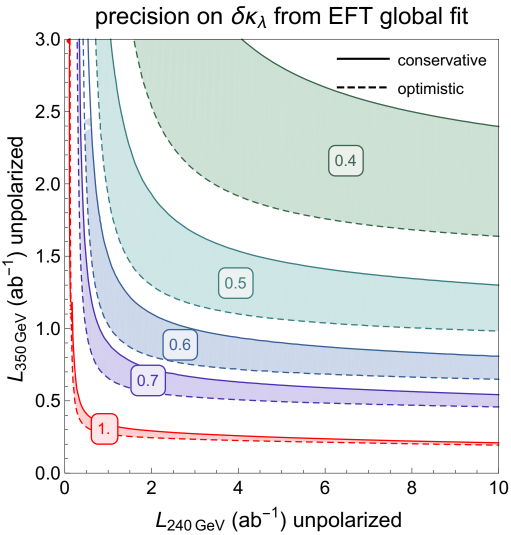






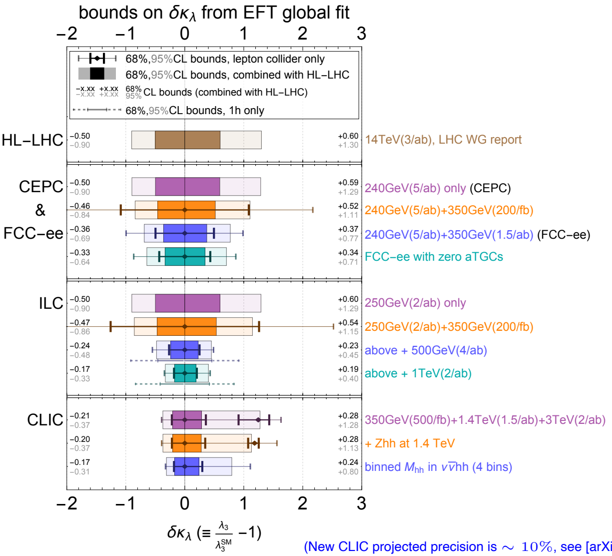
- $\kappa_\lambda \equiv \frac{\lambda_{hhh}}{\lambda_{hhh}^{SM}}$,
  $\delta\kappa_\lambda \equiv \kappa_\lambda - 1 = c_6 - \frac{3}{2}c_H$,
  with $\mathcal{L} \supset -\frac{c_6\,\lambda}{v^2}(H^\dagger H)^3$.
- One loop corrections to all Higgs couplings (production and decay).
- 240 GeV: hZ near threshold (more sensitive to $\delta\kappa_\lambda$)
- at 350-365 GeV:
  - WW fusion
  - hZ at a different energy
- $h \to WW^* / ZZ^*$ also have some discriminating power (but turned out to be not enough).

Triple Higgs coupling from EFT global fits



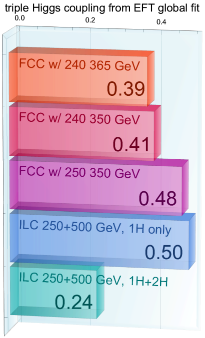
- Runs at two different energies (240 GeV and 350/365 GeV) are needed to obtain good constraints on the triple Higgs coupling in a global fit!

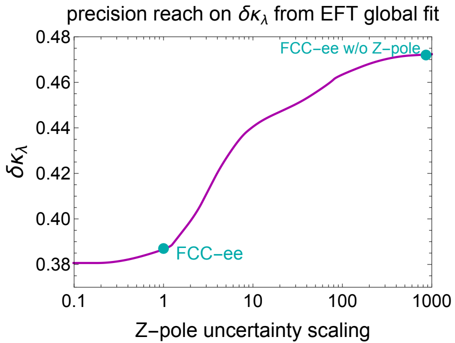
Triple Higgs coupling from global fits [arXiv:1711.03978]



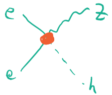
(New CLIC projected precision is $\sim 10\%$, see [arXiv:1901.05897] Roloff *et al.*)

Updates on the triple Higgs coupling determination from EFT global fits



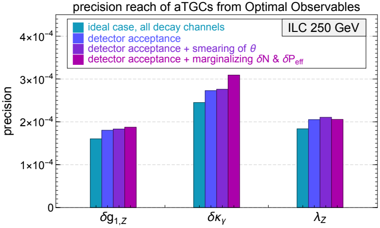


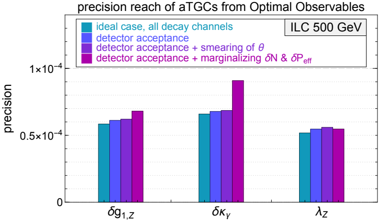
- 240, 365 GeV are better than 250, 350 GeV.
- Impacts of Z-pole measurements are not negligible.
  ($eeZ(h)$ contact interaction enters $e^+e^- \to hZ$.)



Updates on the *WW* analysis with Optimal Observables

- How well can we do it in practice?
  - detector acceptance, measurement uncertainties, ...
- What we have done (current work for the snowmass study)
  - detector acceptance ($|\cos\theta| < 0.9$ for jets, < 0.95 for leptons)
  - some smearing (production polar angle only, $\Delta = 0.1$)
  - ILC: marginalizing over total rate (δN) and effective beam polarization (δP_{eff})
- Constructing full EFT likelihood and feed it to the global fit. (For illustration, only showing the 3-aTGC fit results here.)
- Further verifications (by experimentalists) are needed.



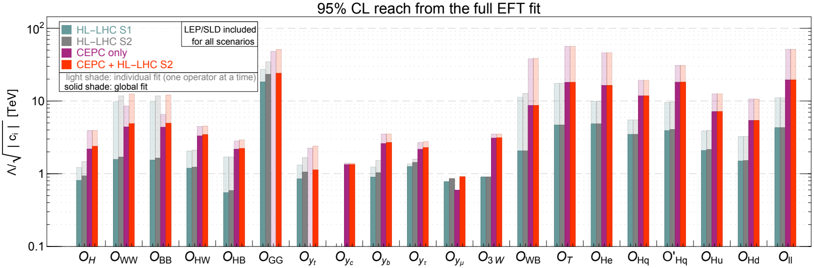


D6 operators

$\mathcal{O}_H = \frac{1}{2}(\partial_\mu \lvert H^2 \rvert)^2$	$\mathcal{O}_{GG} = g_s^2 \lvert H\rvert^2 G^A_{\mu\nu} G^{A,\mu\nu}$
$\mathcal{O}_{WW} = g^2 \lvert H\rvert^2 W^a_{\mu\nu} W^{a,\mu\nu}$	$\mathcal{O}_{y_u} = y_u \lvert H\rvert^2 \bar{q}_L \tilde{H} u_R + \text{h.c.} \quad (u \to t, c)$
$\mathcal{O}_{BB} = g'^2 \lvert H\rvert^2 B_{\mu\nu} B^{\mu\nu}$	$\mathcal{O}_{y_d} = y_d \lvert H\rvert^2 \bar{q}_L H d_R + \text{h.c.} \quad (d \to b)$
$\mathcal{O}_{HW} = ig(D^\mu H)^\dagger \sigma^a (D^\nu H) W^a_{\mu\nu}$	$\mathcal{O}_{y_e} = y_e \lvert H\rvert^2 \bar{l}_L H e_R + \text{h.c.} \quad (e \to \tau, \mu)$
$\mathcal{O}_{HB} = ig'(D^\mu H)^\dagger (D^\nu H) B_{\mu\nu}$	$\mathcal{O}_{3W} = \frac{1}{3!} g \epsilon_{abc} W^{a\,\nu}_{\mu} W^b_{\nu\rho} W^{c\,\rho\mu}$
$\mathcal{O}_W = \frac{ig}{2}(H^\dagger \sigma^a \overleftrightarrow{D}_\mu H) D^\nu W^a_{\mu\nu}$	$\mathcal{O}_B = \frac{ig'}{2}(H^\dagger \overleftrightarrow{D}_\mu H) \partial^\nu B_{\mu\nu}$
$\mathcal{O}_{WB} = gg' H^\dagger \sigma^a H W^a_{\mu\nu} B^{\mu\nu}$	$\mathcal{O}_{H\ell} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{\ell}_L \gamma^\mu \ell_L$
$\mathcal{O}_T = \frac{1}{2}(H^\dagger \overleftrightarrow{D}_\mu H)^2$	$\mathcal{O}'_{H\ell} = iH^\dagger \sigma^a \overleftrightarrow{D}_\mu H \bar{\ell}_L \sigma^a \gamma^\mu \ell_L$
$\mathcal{O}_{\ell\ell} = (\bar{\ell}_L \gamma^\mu \ell_L)(\bar{\ell}_L \gamma_\mu \ell_L)$	$\mathcal{O}_{He} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{e}_R \gamma^\mu e_R$
$\mathcal{O}_{Hq} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{q}_L \gamma^\mu q_L$	$\mathcal{O}_{Hu} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{u}_R \gamma^\mu u_R$
$\mathcal{O}'_{Hq} = iH^\dagger \sigma^a \overleftrightarrow{D}_\mu H \bar{q}_L \sigma^a \gamma^\mu q_L$	$\mathcal{O}_{Hd} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{d}_R \gamma^\mu d_R$

- SILH' basis (eliminate $\mathcal{O}_{WW}$, $\mathcal{O}_{WB}$, $\mathcal{O}_{H\ell}$ and $\mathcal{O}'_{H\ell}$)
- Modified-SILH' basis (eliminate $\mathcal{O}_W$, $\mathcal{O}_B$, $\mathcal{O}_{H\ell}$ and $\mathcal{O}'_{H\ell}$)
- Warsaw basis (eliminate $\mathcal{O}_W$, $\mathcal{O}_B$, $\mathcal{O}_{HW}$ and $\mathcal{O}_{HB}$)

Reach on the scale of new physics

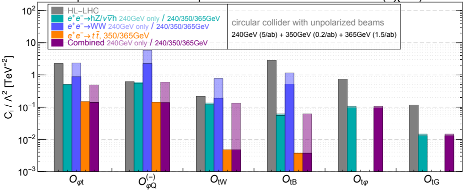


- Reach on the scale of new physics Λ.
- Note: reach depends on the couplings c_i!

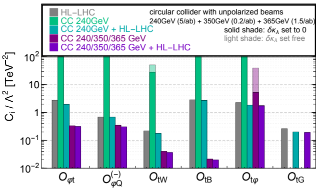
Top operators in loops

[arXiv:1809.03520] G. Durieux, JG, E. Vryonidou, C. Zhang

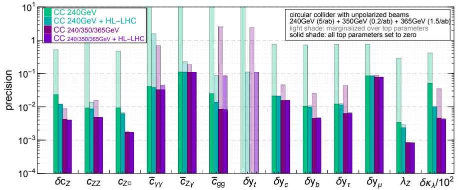
precision of individual operator coefficients at LHC and CC ($\Delta\chi^2$=1)



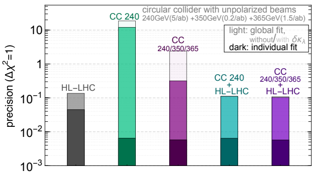
precision of top operator coefficients (global fit, $\Delta\chi^2$=1)



precision of the Higgs parameters at CC (global fit, $\Delta\chi^2$=1)



precision reach on the top Yukawa coupling

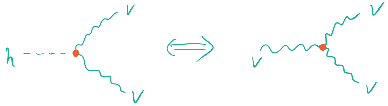


- Higgs precision measurements have sensitivity to the top operators in the loops.
  - But it is challenging to discriminate many parameters in a global fit!
- HL-LHC helps, but a 360 or 365 GeV run is better.
- Indirect bounds on the top Yukawa coupling.

You can't really separate Higgs from the EW gauge bosons!

- $\mathcal{O}_{H\ell} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{\ell}_L \gamma^\mu \ell_L$,
  $\mathcal{O}'_{H\ell} = iH^\dagger \sigma^a \overleftrightarrow{D}_\mu H \bar{\ell}_L \sigma^a \gamma^\mu \ell_L$,
  $\mathcal{O}_{He} = iH^\dagger \overleftrightarrow{D}_\mu H \bar{e}_R \gamma^\mu e_R$
  (or the ones with quarks)
  - modifies gauge couplings of fermions,
  - also generates *hVff* type contact interaction.

- $\mathcal{O}_{HW} = ig(D^\mu H)^\dagger \sigma^a (D^\nu H) W^a_{\mu\nu}$,
  $\mathcal{O}_{HB} = ig'(D^\mu H)^\dagger (D^\nu H) B_{\mu\nu}$
  - generate **aTGCs** $\delta g_{1,Z}$ and $\delta \kappa_\gamma$,
  - also generates *HVV* **anomalous couplings** such as $hZ_\mu \partial_\nu Z^{\mu\nu}$.



You also have to measure the Higgs!

- Some operators can only be probed with the **Higgs particle**.
- $|H|^2 W_{\mu\nu} W^{\mu\nu}$ and $|H|^2 B_{\mu\nu} B^{\mu\nu}$
  - $H \to v/\sqrt{2}$, corrections to gauge couplings?
  - Can be absorbed by field redefinition! This applies to any operators in the form $|H|^2 \mathcal{O}_{\rm SM}$.

$$c_{\rm SM}\mathcal{O}_{\rm SM} \quad vs. \quad c_{\rm SM}\mathcal{O}_{\rm SM} + \frac{c}{\Lambda^2}|H|^2\mathcal{O}_{\rm SM}$$
$$= (c_{\rm SM} + \frac{c\,v^2}{2\,\Lambda^2})\mathcal{O}_{\rm SM} + \text{terms with } h$$
$$= c'_{\rm SM}\mathcal{O}_{\rm SM} + \text{terms with } h$$

  - probed by measurements of the $h\gamma\gamma$ and $hZ\gamma$ couplings, or the hWW and hZZ anomalous couplings.
  - or Higgs in the loop (different story...)
- Yukawa couplings, Higgs self couplings, ...

Why lepton colliders?

- EFT is good for lepton colliders.
  - A systematic parameterization of Higgs (and other) couplings.
- Lepton colliders are also good for EFT!
  - High precision $\Rightarrow$ $E \ll \Lambda$
    Ideal for EFT studies!
  - LHC is built for discovery, but ....

Why lepton colliders?

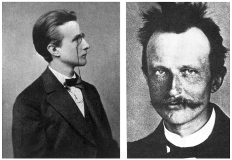
- EFT is good for lepton colliders.
  - A systematic parameterization of Higgs (and other) couplings.
- Lepton colliders are also good for EFT!
  - High precision $\Rightarrow$ $E \ll \Lambda$
    Ideal for EFT studies!
  - LHC is built for discovery, but ....
- **Energy** *vs.* **Precision**
  - Poor measurements at the high energy tails lead to problems in the interpretation of EFT...

A lesson from history

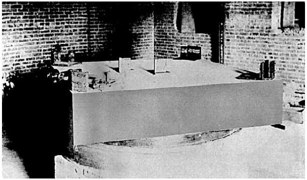
- In 1875, a young Max Planck was told by his advisor Philipp von Jolly not to study physics, since there was nothing left to be discovered.
  - **Planck did not listen.**

Max Planck:

Before quantum physics: After quantum physics:



- In 1887, Michelson and Morley tried to find ether, the postulated medium for the propagation of light that was widely believed to exist.
  - **They didn't find it.**



- **"Our future discoveries must be looked for in the sixth place of decimals." — Albert A. Michelson**

Conclusion

Waiting for a future lepton collider to be built...