

# Lessons from Maschine Learning of Higgs CP measurements in $H \rightarrow \tau^+ \tau^-$ , $\tau^\pm \rightarrow (3\pi)^\pm \nu$ , for the TauSpinner and 'truth' simulation of leptons.

Z. Was\*,

\* Institute of Nuclear Physics, Polish Academy of Sciences Krakow

- **(1)** For me all started with 2001 publication on CP observability for  $H \rightarrow \tau^+ \tau^-$ ,  $\tau^\pm \rightarrow \pi^\pm \pi^0 \nu$ . Q: are there similar for other  $\tau$  decay modes? **A: not much progress despite several attempts.**
- **(2)** In the last two years, we could finally say **yes** for the  $\tau^\pm \rightarrow (3\pi)^\pm \nu$ .
- **(3)** Use of the Maschine Learning algorithm was instrumental to manage observable of 7 dimensional nature (at least). Our choice of particular software, [Y. LeCun, Y. Bengio, and G. Hinton, Nature 521 \(2015\), no. 7553 436](#) was because of advice/help we could get.
- **(4)** My perspective was what was needed from theoretical (experimental side– direction which need to be worked out in a great detail). Weighted events helpful.
- **(5)** I will present some results, but concentrate on the possible lesson for future applications **and for the design of simulation tools too.**

This work is partly supported by the grant of Narodowe Centrum Nauki 2016/23/B/ST2/03927

# Priorities:

- Theory first
- Experimental effects first
- Mathematics first\*
- Software development first\*

Everybody can choose his own  
*as the first,*

but it is natural that the others will do the same.

\* See for example <https://indico.cern.ch/event/673350/> <https://indico.cern.ch/event/687788/>

Only **transverse** spin correlations of  $\tau^+$  and  $\tau^-$  in  $H \rightarrow \tau^+\tau^-$  are **different for scalar and pseudoscalar Higgs**.

- The correlations can not be measured directly
- One need to measure distributions of  $\tau$  decay products
- Precisely their transverse (to  $\tau$  direction in Higgs boson rest frame) momenta
- Easiest to interpret is  $\tau^\pm \rightarrow \pi^\pm \nu$ , but it require indirect measurement of  $\nu_\tau$ .

M. Kramer, J. H. Kuhn, M. L. Stong and P. M. Zerwas, "Prospects of measuring the parity of Higgs particles," Z. Phys. C **64**, 21 (1994)

A. Rouge, "CP violation in a light Higgs boson decay from tau-spin correlations at a linear collider," Phys. Lett. B **619**, 43 (2005)

- The largest branching ratio (25 %) has  $\tau^\pm \rightarrow \pi^\pm \pi^0 \nu$  and we can look on transverse spin correlations of  $\rho^\pm \rightarrow \pi^\pm \pi^0$  decays.

There observable based on visible  $\tau$  decay products only, could be designed.

G. R. Bower, T. Pierzchala, Z. Was and M. Worek, "Measuring the Higgs boson's parity using tau  $\rightarrow$  rho nu," Phys. Lett. B **543** (2002) 227,

K. Desch, ZW, M. Worek, "Measuring the Higgs boson parity at a linear collider using the tau impact parameter and tau  $\rightarrow$  rho nu decay", Eur.Phys.J. C29 (2003) 491

## The Higgs boson's parity

- $H/A$  parity information can be extracted from the correlations between  $\tau^+$  and  $\tau^-$  spin components which are further reflected in correlations between the  $\tau$  decay products in the plane transverse to the  $\tau^+\tau^-$  axes.
- The decay probability

$$\Gamma(H/A \rightarrow \tau^+\tau^-) \sim 1 - s_{\parallel}^{\tau^+} s_{\parallel}^{\tau^-} \pm s_{\perp}^{\tau^+} s_{\perp}^{\tau^-}$$

is sensitive to the  $\tau^{\pm}$  polarization vectors  $s^{\tau^-}$  and  $s^{\tau^+}$  (defined in their respective rest frames). The symbols  $\parallel, \perp$  denote components parallel/transverse to the Higgs boson momentum as seen from the respective  $\tau^{\pm}$  rest frames.

- This idea and its practical refinements are universal: 'Higgs spin' is blind on Higgs origin. But it is not true for the background DY processes .

## *Phenomenology Of Mixed Parity Case*

- Higgs boson Yukawa coupling expressed with the help of the scalar–pseudoscalar mixing angle  $\phi$

$$\bar{\tau} N (\cos \phi + i \sin \phi \gamma_5) \tau$$

- *Decay probability for the mixed scalar–pseudoscalar case*

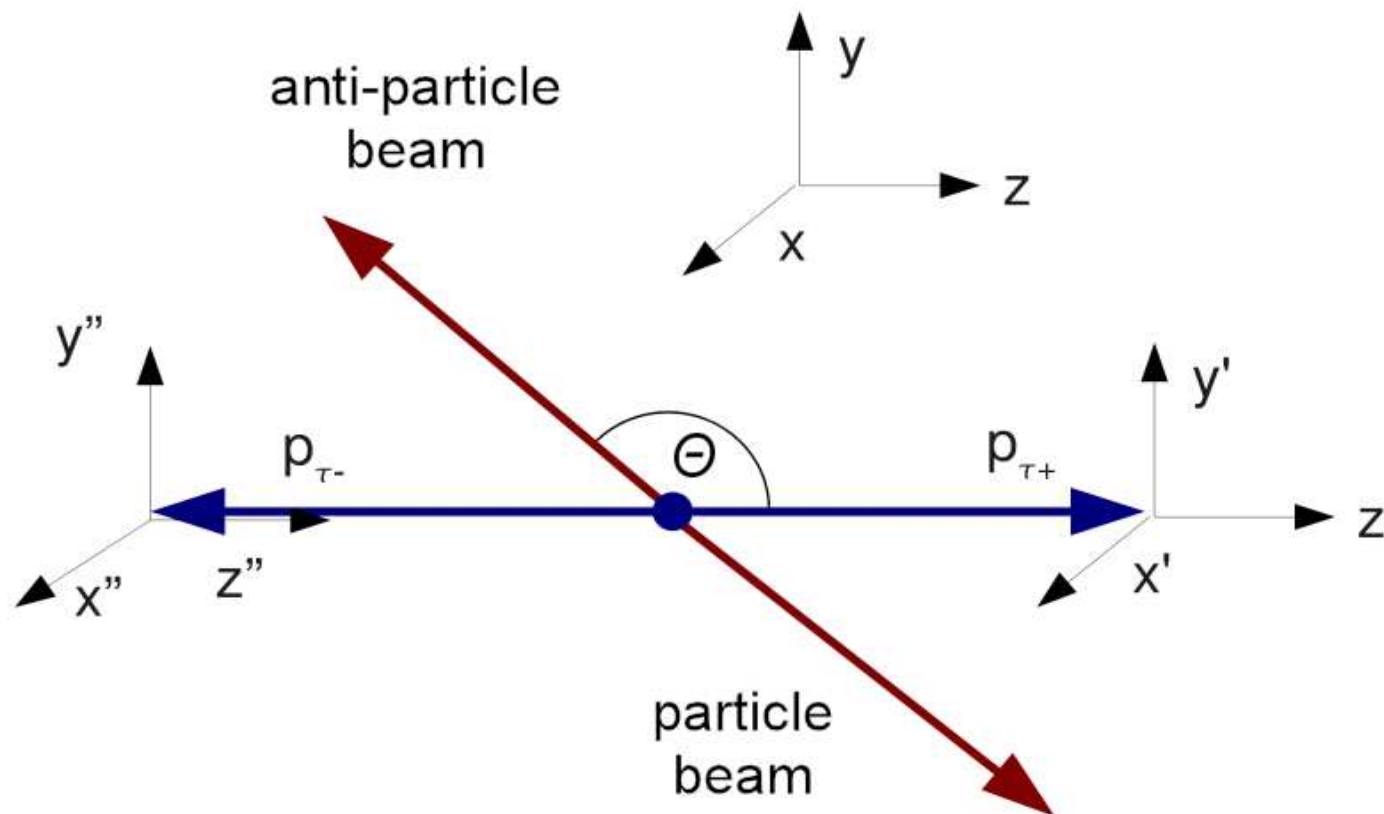
$$\Gamma(h_{mix} \rightarrow \tau^+ \tau^-) \sim 1 - s_{\parallel}^{\tau^+} s_{\parallel}^{\tau^-} + s_{\perp}^{\tau^+} R(2\phi) s_{\perp}^{\tau^-}$$

- *$R(2\phi)$  – operator for the rotation by angle  $2\phi$  around the  $\parallel$  direction.*

$$R_{11} = R_{22} = \cos 2\phi \quad R_{12} = -R_{21} = \sin 2\phi$$

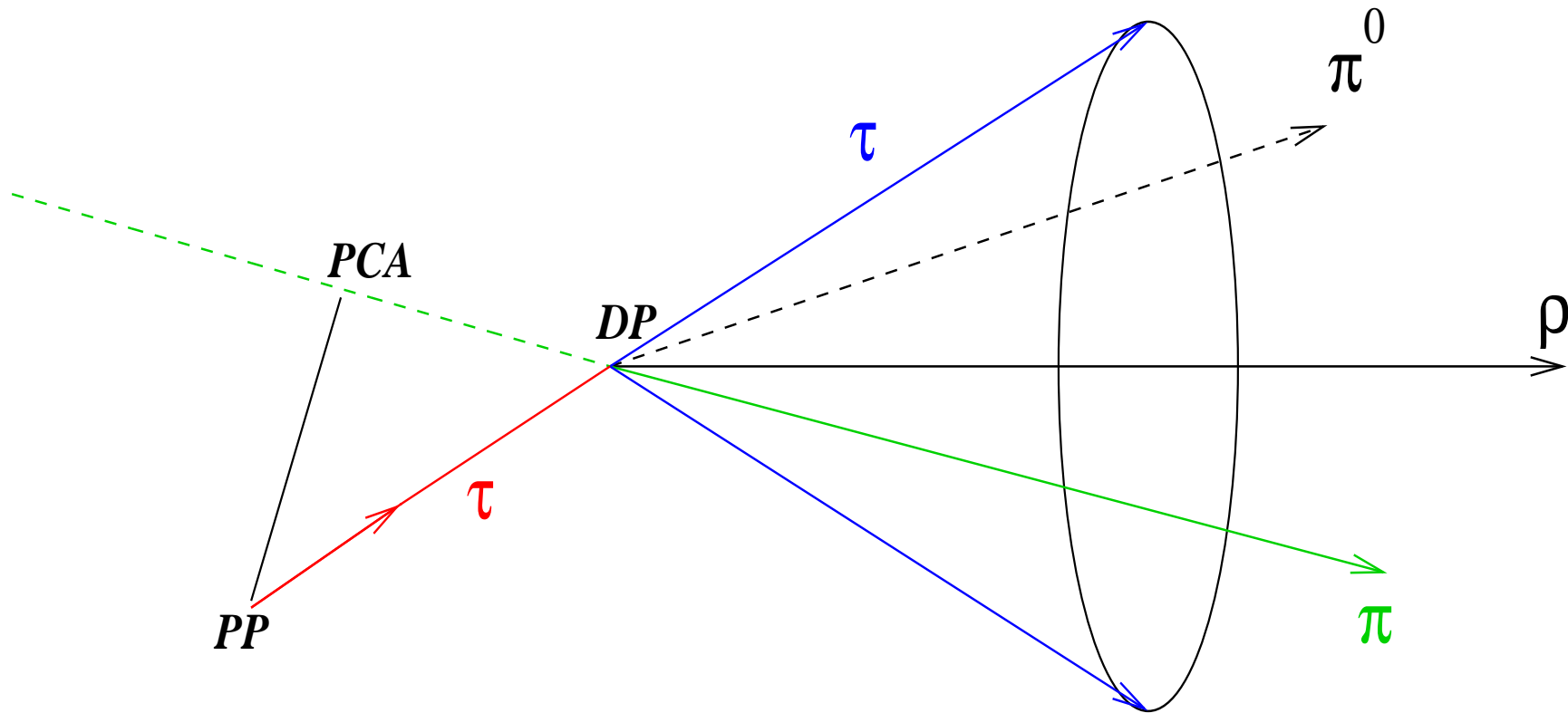
- *Pure scalar case is reproduced for  $\phi = 0$ .*
- *For  $\phi = \pi/2$  we reproduce the pure pseudoscalar case.*

*What does it mean  $\tau$  pair?*



Coordinate systems of H (or  $Z/\gamma^*$ ),  $\tau^+$  and  $\tau^-$  rest frames

*What does it mean  $\tau$  lepton in the detector?*



Measurable  $\pi^\pm$ ,  $\pi^0$ ,

$\nu_\tau$  through impact parameter and reconstruction (if at all).

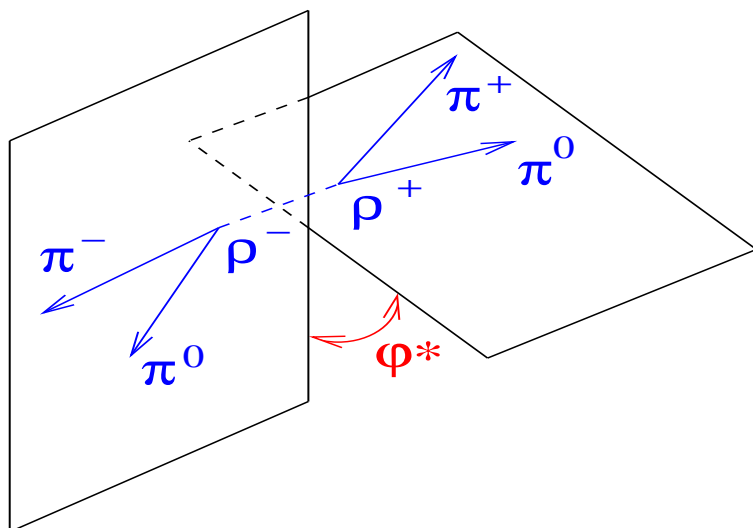
## Transverse spin correlations through $\tau$ decays

- Case of  $\tau \rightarrow \rho\nu_\tau$  decay,  $\mathcal{BR}(\tau \rightarrow \rho\nu_\tau) = 25\%$
- Polarimeter vector  $h^i$  is (where  $q$  for  $\pi^\pm - \pi^0$  and  $N$  for  $\nu_\tau$  four momenta).

$$h^i = \mathcal{N} \left( 2(q \cdot N)q^i - q^2 N^i \right)$$

$$q \cdot N = (E_{\pi^\pm} - E_{\pi^0})m_\tau$$

- Acoplanarity of  $\rho^+$  and  $\rho^-$  decay prod. (in  $\rho^+ \rho^-$  r.f.) and events separation.



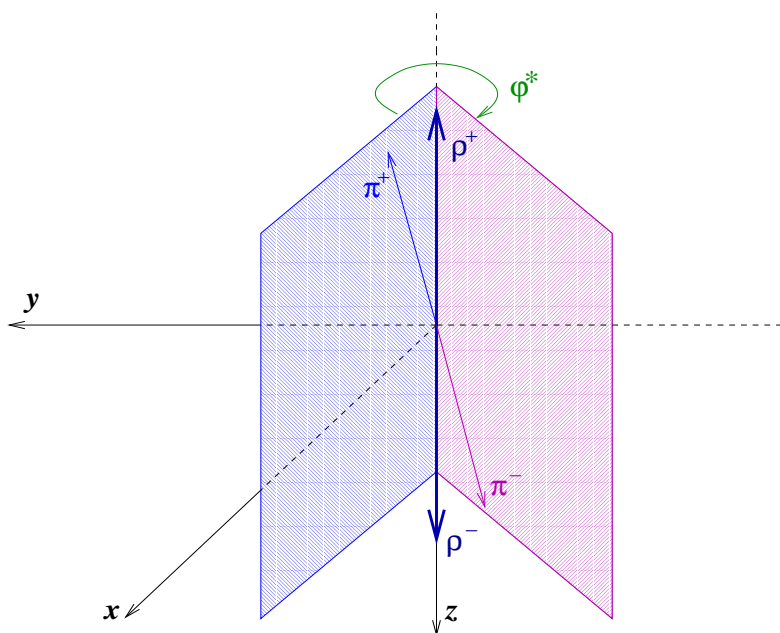
$$y_1 y_2 > 0; \quad y_1 y_2 < 0 \text{ (in } \tau^\pm \text{ r.f.'s)}$$

$$y_1 = \frac{E_{\pi^+} - E_{\pi^0}}{E_{\pi^+} + E_{\pi^0}}; \quad y_2 = \frac{E_{\pi^-} - E_{\pi^0}}{E_{\pi^-} + E_{\pi^0}}$$



## Observable For Mixed Scalar–Pseudoscalar Case

- For mixing angle  $\phi$ , transverse component of  $\tau^+$  spin polarization vector is correlated with the one of  $\tau^-$  rotated by angle  $2\phi$ .
- Acoplanarity  $0 < \varphi^* < 2\pi$  is of physical interest, not just  $\arccos \mathbf{n}_- \cdot \mathbf{n}_+$ .
- Distinguish between the two cases  $0 < \varphi^* < \pi$  and  $2\pi - \varphi^*$
- If no separation made the parity effect would wash itself out.



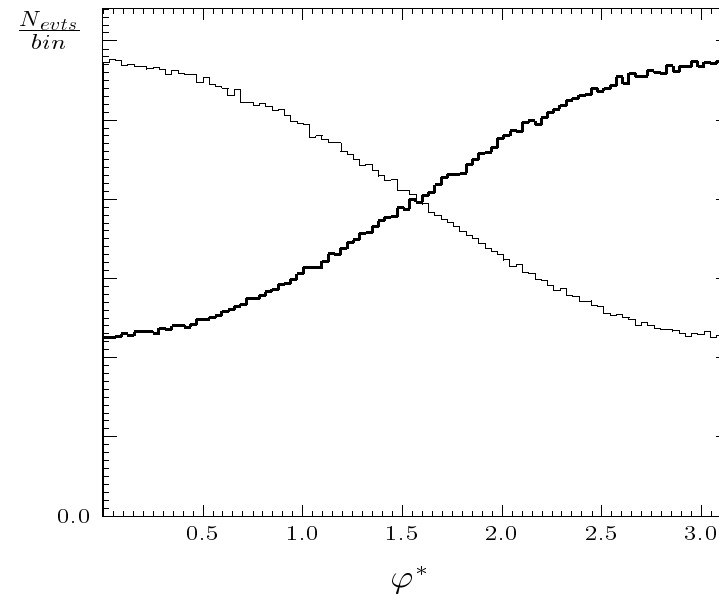
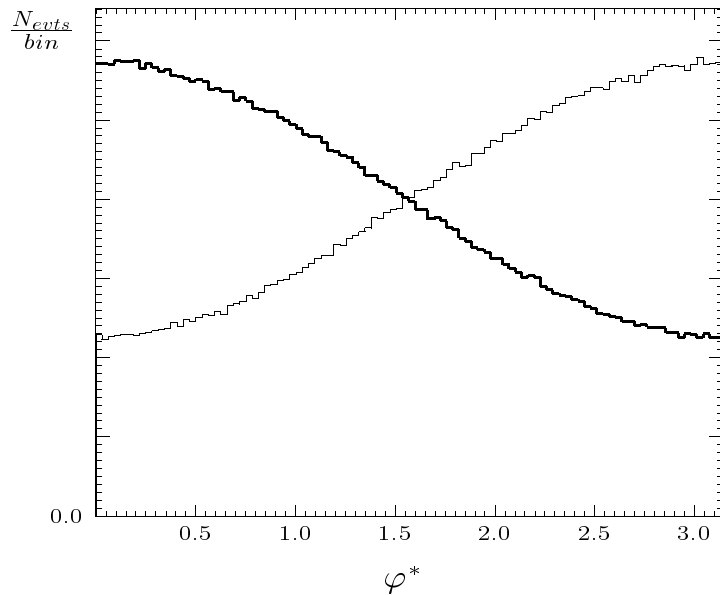
Normal to planes:  $\mathbf{n}_{\pm} = \mathbf{p}_{\pi^{\pm}} \times \mathbf{p}_{\pi^0}$

Find the sign of  $\mathbf{p}_{\pi^-} \cdot \mathbf{n}_+$

Negative  $0 < \varphi^* < \pi$

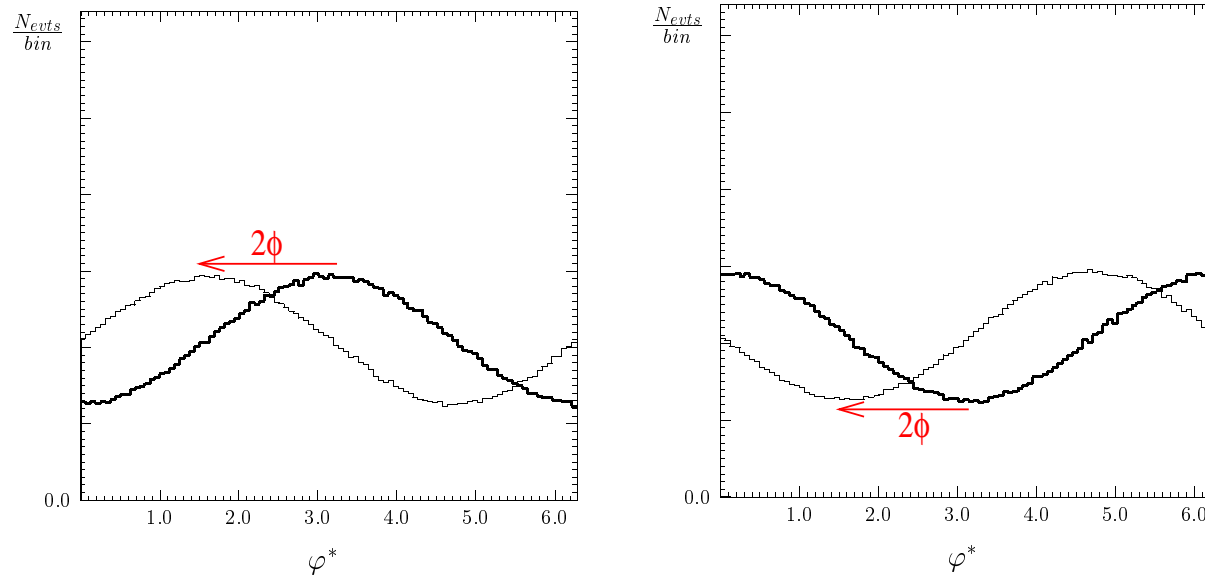
Otherwise  $2\pi - \varphi^*$

## Results Without Smearing



- The  $\rho^+ \rho^-$  decay products' acoplanarity distribution without any smearing .
- Selection  $y_1 y_2 > 0$  is used in the left plot,  $y_1 y_2 < 0$  is used for the right plot.
- Thick line denote the case of the scalar Higgs and thin lines the pseudoscalar.
- Complete spin correlations of  $h \rightarrow \tau^+ \tau^-$ ,  $\tau^\pm \rightarrow \rho^\pm \nu$ ,  $\rho^\pm \rightarrow \pi^\pm \pi^0$  incl.

## Old attempts on detector and Impact Parameter



- Only events where the signs of  $y_1$  and  $y_2$  are the same whether calculated using the method without or with the help of the  $\tau$  impact parameter.
- Tesla-like set-up SIMDET used, K. Desch, A. Imhof, ZW, M. Worek, Phys.Lett. B579 (2004) 157.
- The thick line corresponds to a scalar Higgs boson, the thin line to a mixed one.

Precision on  $\phi \sim 6^\circ$ , for  $1\text{ab}^{-1}$  and 350 GeV CMS.

Q: Why ML and not e.g. TMVA?

A: Manpower: in the not so distant past, and now.

I do not plan to cover this topic.

Anyway used by us solutions are outdated they are of nearly wo years ago...

Q: what is ML?

A: I will not cover this topic as well.

I will go to the results

Note:

- From the  $\pi^-$ ,  $\pi^0$  we can define a plane for acoplanarity
- From the  $\pi^-$ ,  $\pi^-$ ,  $\pi^+$  we can define four such planes.
- Each plane bring its own  $y_i$  variable to avoid cancellations due to properties of  $\tau$  decay ME:  $\langle \cos \theta \rangle = 0$ .

Features/variables	$\rho^\pm - \rho^\mp$ $\rho^\pm \rightarrow \pi^0 \pi^\pm$	$a_1^\pm - \rho^\mp$ $a_1^\pm \rightarrow \rho^0 \pi^\mp, \rho^0 \rightarrow \pi^+ \pi^-$ $\rho^\mp \rightarrow \pi^0 \pi^\mp$	$a_1^\pm - a_1^\mp$ $a_1^\pm \rightarrow \rho^0 \pi^\pm, \rho^0 \rightarrow \pi^+ \pi^-$
True classification	0.782	0.782	0.782
$\varphi_{i,k}^*$	0.500	0.500	0.500
$\varphi_{i,k}^*$ and $y_i, y_k$	0.624	0.569	0.536
<b>4-vectors</b>	0.638	0.590	0.557
$\varphi_{i,k}^*$ , 4-vectors	0.638	0.594	0.573
$\varphi_{i,k}^*, y_i, y_k$ and $m_i^2, m_k^2$	0.626	0.578	0.548
$\varphi_{i,k}^*, y_i, y_k, m_i^2, m_k^2$ and 4-vectors	0.639	0.596	0.573

Table 1: Average probability  $p_i$  that a model predicts correctly event  $x_i$  to be of a type  $A$  (scalar), with training being performed for separation between type  $A$  and  $B$  (pseudo-scalar). **Looks beautiful ML do all for us. Really?** Let's check details.

Features				Ideal $\pm$ (stat)	Smeared $\pm$ (stat) $\pm$ (syst)
$\phi^*$	4-vec	$y_i$	$m_i$		
$a_1 - \rho$ Decays					
✓	✓	✓	✓	$0.6035 \pm 0.0005$	$0.5923 \pm 0.0005 \pm 0.0002$
✓	✓	✓	-	$0.5965 \pm 0.0005$	$0.5889 \pm 0.0005 \pm 0.0002$
✓	✓	-	✓	$0.6037 \pm 0.0005$	$0.5933 \pm 0.0005 \pm 0.0003$
-	✓	-	-	$0.5971 \pm 0.0005$	$0.5892 \pm 0.0005 \pm 0.0002$
✓	✓	-	-	$0.5971 \pm 0.0005$	$0.5893 \pm 0.0005 \pm 0.0002$
✓	-	✓	✓	$0.5927 \pm 0.0005$	$0.5847 \pm 0.0005 \pm 0.0002$
✓	-	✓	-	$0.5819 \pm 0.0005$	$0.5746 \pm 0.0005 \pm 0.0002$
$a_1 - a_1$ Decays					
✓	✓	✓	✓	$0.5669 \pm 0.0004$	$0.5657 \pm 0.0004 \pm 0.0001$
✓	✓	✓	-	$0.5596 \pm 0.0004$	$0.5599 \pm 0.0004 \pm 0.0001$
✓	✓	-	✓	$0.5677 \pm 0.0004$	$0.5661 \pm 0.0004 \pm 0.0001$
-	✓	-	-	$0.5654 \pm 0.0004$	$0.5641 \pm 0.0004 \pm 0.0001$
✓	✓	-	-	$0.5623 \pm 0.0004$	$0.5615 \pm 0.0004 \pm 0.0001$
✓	-	✓	✓	$0.5469 \pm 0.0004$	$0.5466 \pm 0.0004 \pm 0.0001$
✓	-	✓	-	$0.5369 \pm 0.0004$	$0.5374 \pm 0.0004 \pm 0.0001$

**Table 2:** AUC for NN trained to separate scalar and pseudoscalar hypotheses with combinations of input features marked with a ✓. Results in the column labelled "Ideal" are from NNs trained with ideal MC (particle-level simulation). The results in column labelled "Smeared" are from NNs trained with smeared MC.

Features/variables	$\rho^\pm - \rho^\mp$ $\rho^\pm \rightarrow \pi^0 \pi^\pm$	$a_1^\pm - \rho^\mp$ $a_1^\pm \rightarrow \rho^0 \pi^\pm, \rho^0 \rightarrow \pi^+ \pi^-$ $\rho^\mp \rightarrow \pi^0 \pi^\mp$	$a_1^\pm - a_1^\mp$ $a_1^\pm \rightarrow \rho^0 \pi^\pm,$ $\rho^0 \rightarrow \pi^+ \pi^-$
$\varphi_{i,k}^*$	1	4	16
$\varphi_{i,k}^*$ and $y_i, y_k$	3	9	24
$\varphi_{i,k}^*$ , 4-vectors	25	36	64
$\varphi_{i,k}^*, y_i, y_k$ and $m_i, m_k$	5	13	30
$\varphi_{i,k}^*, y_i, y_k, m_i, m_k$ and 4-vectors	29	45	78

Table 3: Dimensionality of the features which may be used in each discussed configuration of the decay modes. Note that in principle  $y_i^\pm, y_k^\mp$  may be calculated in the rest frame of the resonance pair used to define  $\varphi_{i,k}^*$  planes, but in practice, choice of the frames is of no numerically significant effect. We do not distinguish such variants.



*New attempt on detector and Impact Parameter*

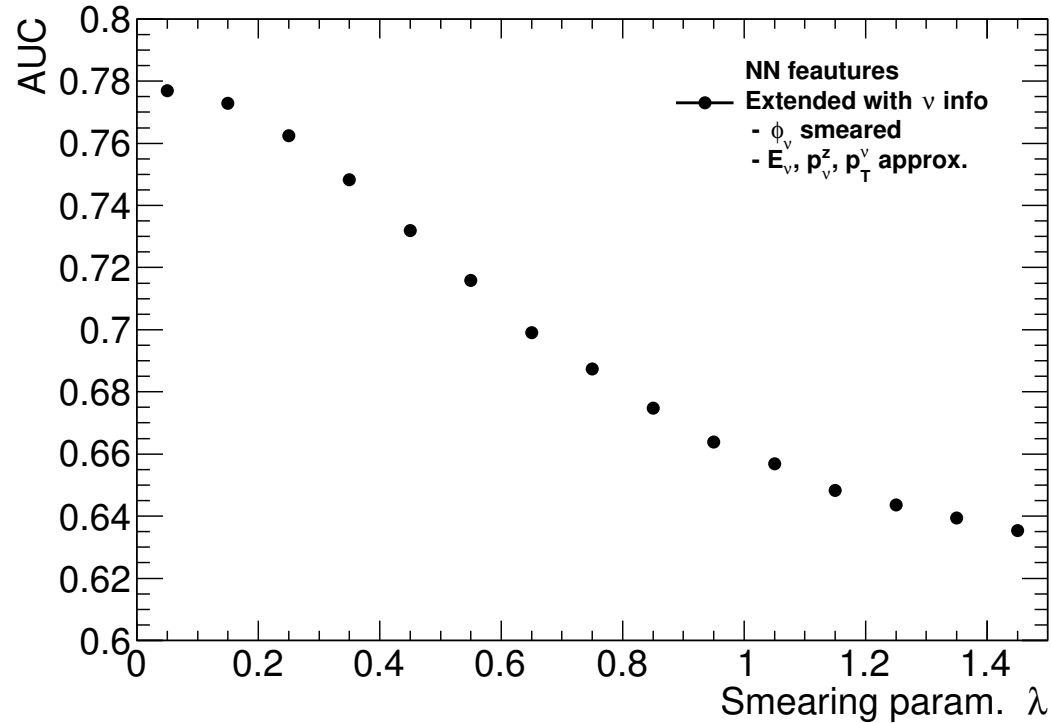


Figure 1: Plot of AUC score for training and validation sample of  $\tau\tau$  pairs with one  $\tau$  decaying  $\tau^\pm \rightarrow a_1^\pm \nu \rightarrow 3\pi^\pm \nu$ , second  $\tau$  decaying  $\tau^\pm \rightarrow \rho^\pm \nu$ , as a function of smearing parameter  $\lambda$ , for orientation angle of neutrinos directions around  $\tau$  hadronic decay products. The rest of kinematic reconstructed from mass or missing  $p_T$  constraints.

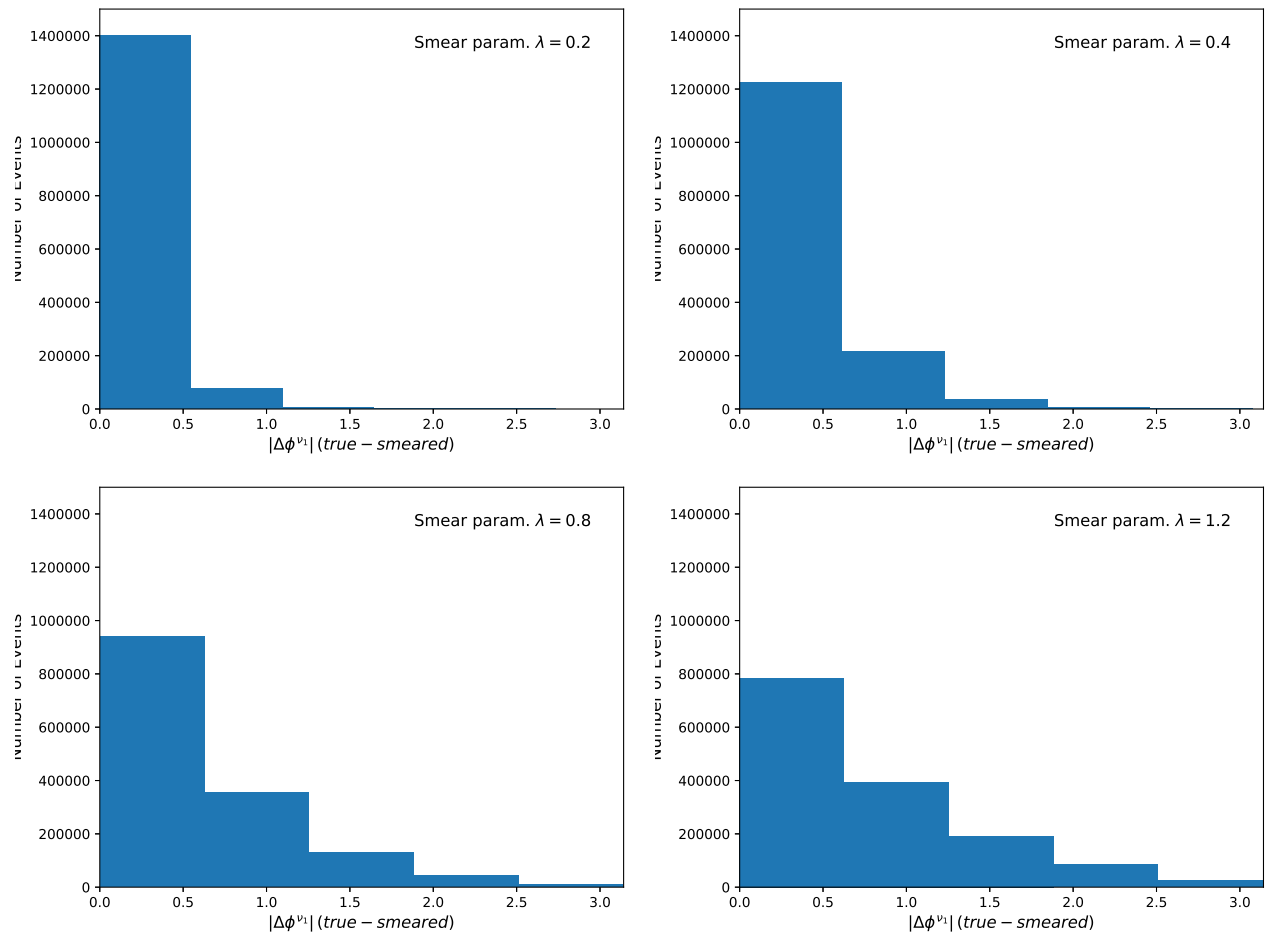


Figure 2: The  $|\Delta\phi_{\nu_1}|$  (true - smeared) plot for smearing parameter  $\lambda=0.2, 0.4, 0.8, 1.2$ . Note that moderate smearing mean no loss of precision, there is 'critical' range of values, where sensitivity from neutrinos disappear.

- Does it sound like the end of physics?
- I have recalled some results demonstrating sensitivity and its robustness.
- However:
  - (i) **Are ML results without hidden conditions?**
  - (ii) **Are ML results trustworthy?**
- (i) We had to boost all four momenta into rest-frame of all visible H decay products combined.
- (i) We had to align events along the  $z$  axis
- (i) **Otherwise no sensitivity at all!**
- (ii) I will recall definitions and results for classical one dimensional plots.  
For the understanding and trust.
- But first: **we are not alone with the concerns.**



- Result depend on model assumptions. Models inspired with results ...  
**Fitting setup → biases.**
- Our algorithms are far less elaborate than human eye/brain.
- That may look worrisome.

- Biases in art, Giuseppe Arcimboldo (1572 - 1593).



**Figure 3:** Artificial Neural Networks have spurred remarkable recent progress in image classification and speech recognition. But even though these are very useful tools based on well-known mathematical methods, we actually understand surprisingly little of why certain models work and others don't.

From <http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html>

Pattern recognition is an active field and deep concern and not only for us.

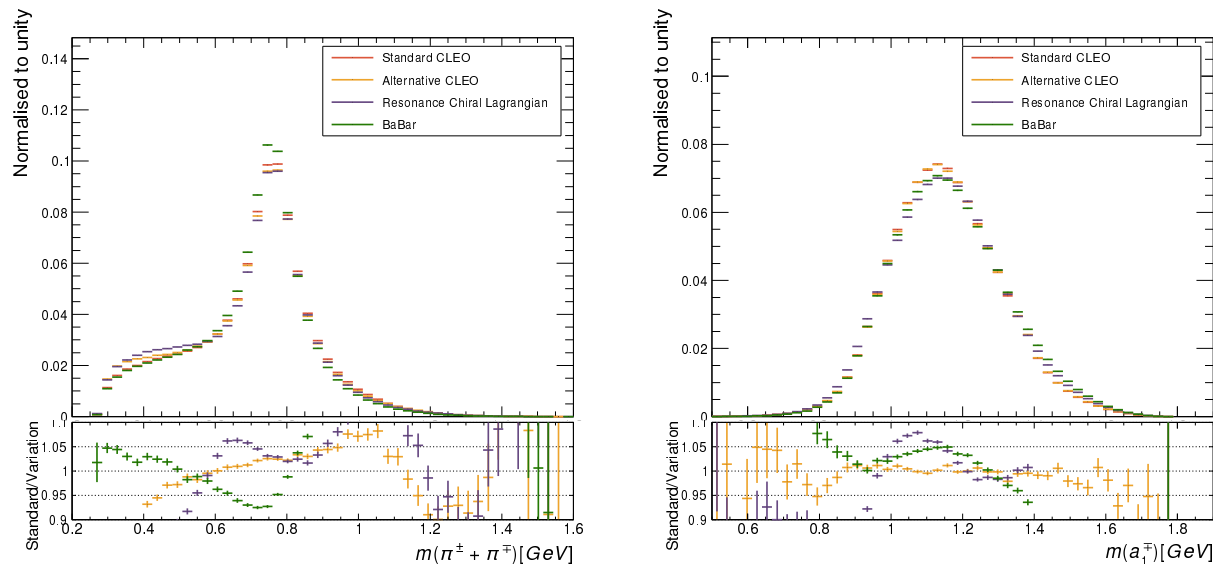
# Scalar or Pseudoscalar? Human mind

22

The acoplanarity distributions I have presented for  $\rho^+ \rho^-$  on slides 10-11, were very close to optimal observables, very useful for interpretation.

They can be useful, even if they are just an element for evaluation of systematic error and to gain confidence that ML results are not too good (to be true).

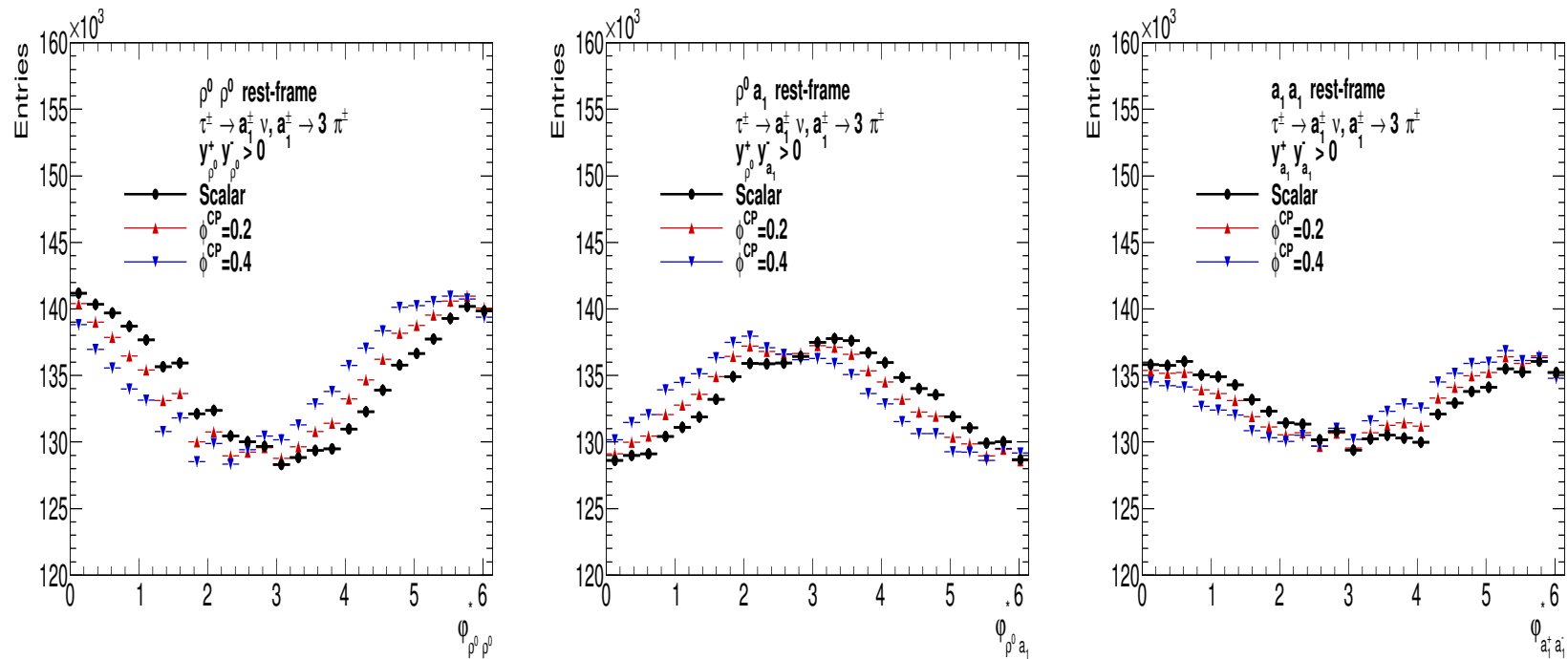
Four models of hadronic currents and control distributions give consistent results, also if different models used for training and analysis for CP classification all went OK.



**from  $\rho^\pm$  to  $a_1^\pm$  case.**

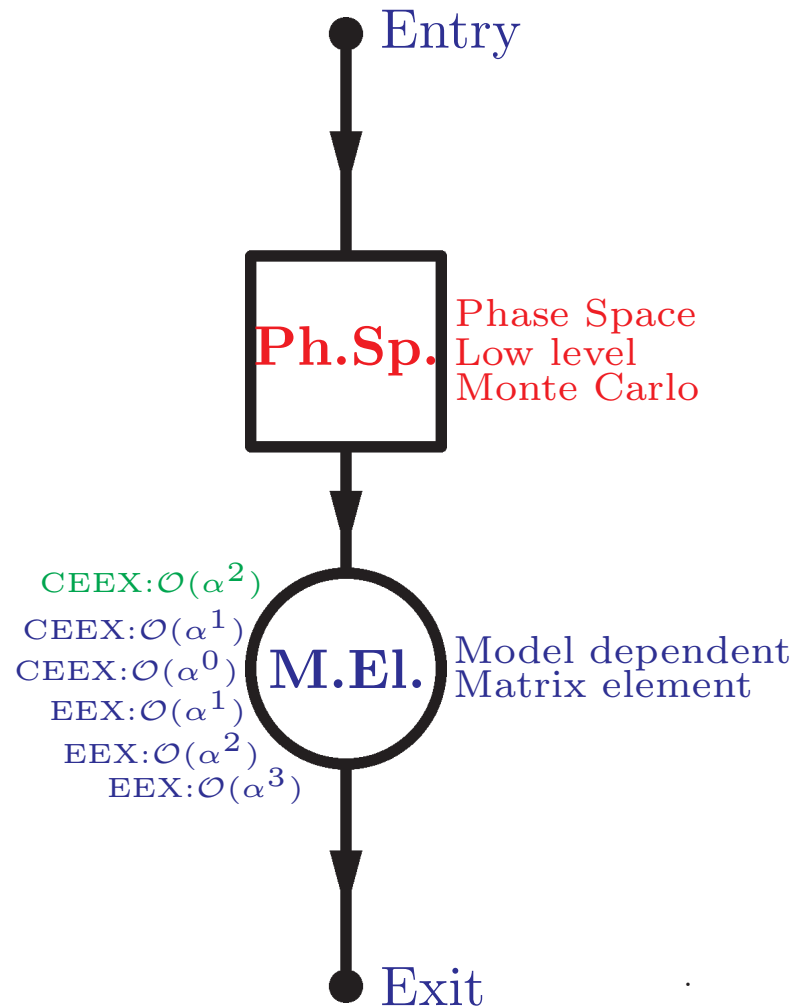
1. In case of  $\tau \rightarrow \rho\nu$  there was one decay plane to define and sign of CP sensitive sinusoid was dependent on sign of  $y_+y_-$ .
2. In case of  $\tau \rightarrow a_1\nu$  four planes can be defined. Two for  $a_1 \rightarrow \pi\rho^0$  and another two for  $\rho^0 \rightarrow \pi^+\pi^-$  decays.
3. We end up with 4 (or 16) angular distributions and number of  $y_i$  like variables.
4. That means many sub categories to define sample ...
5. All distributions are correlated.
6. Methods of Machine Learning useful, to evaluate sensitivity of mult-dimensional signatures
  - $\rightarrow$  may be more sensitivity can be obtained?
  - $\rightarrow$  may be even more sensitivity than possible?

Acoplanarity angles of oriented half decay planes:  $\varphi_{\rho^0 \rho^0}^*$  (left),  $\varphi_{a_1 \rho^0}^*$  (middle) and  $\varphi_{a_1 a_1}^*$  (right), for events grouped by the sign of  $y_{\rho^0}^+ y_{\rho^0}^-$ ,  $y_{a_1}^+ y_{\rho^0}^-$  and  $y_{a_1}^+ y_{a_1}^-$  respectively. Three CP mixing angles  $\phi^{CP} = 0.0$  (scalar), 0.2 and 0.4. Note scale, effect on individual plot is so much smaller now. But up to **16 plots like that** have to be measured, correlations understood. But physics model depends on 1 parameter only and effect of  $\phi^{CP}$ , the Higgs mixing scalar pseudoscalar angle, is always a linear shift.





Textbook principle “matrix element  $\times$  full phase space” must be obeyed



- Universal Phase-space Monte Carlo simulator is a separate module producing “raw events” (including importance sampling for possible intermediate resonances)
- Library of several types of models for signal/background provides input for “model weight” which is another independent module
- This is exactly like in case of KORALZ or KKMC of LEP time
- This applies to all information used for ML input too!
- There is plenty of room for biases.
- **Responsability of physics Monte Carlos  $\rightarrow$  all the following slides**

**Formalism for  $\tau^+\tau^-$**  : helps to separate production and decay

- Because narrow  $\tau$  width approximation can be obviously used for phase space, cross-section for the process  $f\bar{f} \rightarrow \tau^+\tau^-Y; \tau^+ \rightarrow X^+\bar{\nu}; \tau^- \rightarrow \nu\nu$  reads:

$$d\sigma = \sum_{spin} |\mathcal{M}|^2 d\Omega = \sum_{spin} |\mathcal{M}|^2 d\Omega_{prod} d\Omega_{\tau^+} d\Omega_{\tau^-}$$

- This formalism is fine, but because of over 20  $\tau$  decay channels we have over 400 distinct processes. Also picture of production and decay are mixed.
- Below only  $\tau$  spin indices are explicitly written:

$$\mathcal{M} = \sum_{\lambda_1 \lambda_2 = 1}^2 \mathcal{M}_{\lambda_1 \lambda_2}^{prod} \mathcal{M}_{\lambda_1}^{\tau^+} \mathcal{M}_{\lambda_2}^{\tau^-}$$

- Cross section can be re-written into **core formula of spin algorithms**

$$d\sigma = \left( \sum_{spin} |\mathcal{M}^{prod}|^2 \right) \left( \sum_{spin} |\mathcal{M}^{\tau^+}|^2 \right) \left( \sum_{spin} |\mathcal{M}^{\tau^-}|^2 \right) wt d\Omega_{prod} d\Omega_{\tau^+} d\Omega_{\tau^-}$$

## General formalism for semileptonic decays

- Matrix element used in TAUOLA for semileptonic decay

$$\tau(P, s) \rightarrow \nu_\tau(N) X$$

$$\mathcal{M} = \frac{G}{\sqrt{2}} \bar{u}(N) \gamma^\mu (v + a\gamma_5) u(P) J_\mu$$

- $J_\mu$  the current depends on the momenta of all hadrons

$$|\mathcal{M}|^2 = G^2 \frac{v^2 + a^2}{2} (\omega + H_\mu s^\mu)$$

$$\omega = P^\mu (\Pi_\mu - \gamma_{va} \Pi_\mu^5), \quad h_\mu = H_\mu / \omega$$

$$H_\mu = \frac{1}{M} (M^2 \delta_\mu^\nu - P_\mu P^\nu) (\Pi_\nu^5 - \gamma_{va} \Pi_\nu)$$

$$\Pi_\mu = 2[(J^* \cdot N) J_\mu + (J \cdot N) J_\mu^* - (J^* \cdot J) N_\mu]$$

$$\Pi^{5\mu} = 2 \operatorname{Im} \epsilon^{\mu\nu\rho\sigma} J_\nu^* J_\rho N_\sigma$$

$$\gamma_{va} = -\frac{2va}{v^2 + a^2}$$

$$\hat{\omega} = 2 \frac{v^2 - a^2}{v^2 + a^2} m_\nu M (J^* \cdot J)$$

$$\hat{H}^\mu = -2 \frac{v^2 - a^2}{v^2 + a^2} m_\nu \operatorname{Im} \epsilon^{\mu\nu\rho\sigma} J_\nu^* J_\rho P_\sigma$$

## *Higgs Boson Parity*

- *Again scalar/pseudoscalar decay probability in formalism of Kramer et al.*

$$\Gamma(H/A^0 \rightarrow \tau^+ \tau^-) \sim 1 - s_{\parallel}^{\tau^+} s_{\parallel}^{\tau^-} \pm s_{\perp}^{\tau^+} s_{\perp}^{\tau^-}$$

- $s^{\tau}$  is the  $\tau$  polarization vectors.
- $\parallel / \perp$  denote components parallel / transverse to the Higgs boson momentum.
- The spin weight  $0 < wt < 4$  is given by the following formula

$$wt = \frac{1}{4} \left( 1 + \sum_{ij=1}^3 R_{ij} h^i h^j \right)$$

$$R_{33} = -1, R(2\phi) \text{ defined earlier and for the general case}$$

## Results relevant for fitting and for $\tau$ leptons.

### • Applications

- E. Barberio, B. Le, E. Richter-Was, Z. Was, D. Zanzi and J. Zaremba, “Deep learning approach to the Higgs boson CP measurement in  $H \rightarrow \tau\tau$  decay and associated systematics,” Phys. Rev. D **96**, 073002 (2017)
- *W production at LHC: lepton angular distributions and reference frames for probing hard QCD*, E. Richter-Was and Z. Was, arXiv:1609.02536
- *Separating electroweak and strong interactions in Drell–Yan processes at LHC: leptons angular distributions and reference frames*, E. Richter-Was and Z. Was, Eur.Phys.J. C76 (2016) 473

### • Tools:

- “*TauSpinner Program for Studies on Spin Effect in tau Production at the LHC*”, Z. Czyczula, T. Przedzinski and Z. Was, Eur. Phys. J. C **72**, 1988 (2012)
- “*. Production of tau lepton pairs with high  $p_T$  jets at the LHC and the TauSpinner reweighting algorithm*”, J. Kalinowski, W. Kotlarski, E. Richter-Was and Z. Was, arXiv:1604.00964

1. We have demonstrated that ML techniques can be useful to distinguish in statistically controllable way between hypotheses of Higgs coupling to tau being CP even, CP-odd or even CP-mix  
*for the observables which are massively multi-dimensional.*

2. I have pointed issues of mis-interpretations known in the industry.

3. It is known in High energy physics too, e.g. in the domain of jets:

<https://indico.cern.ch/event/667334/>

*Advanced Machine Learning for Classification, Regression, and Generation in Jet Physics* ,

Ben Nachman (LBL) CERN Nov 15 2017

4. **LESSON: it is important to separate those degrees of freedom which can be controlled, from those where more effort is still needed.**

5. In case of signal, that is Higgs production and decay, it is easy: Higgs is narrow and its spin is zero, production is well separated from decay.

6. Problem may come from background. Note that Drell Yan is in comparison huge and intermediate Z state is broader and carries spin.

Let us start with the lowest order coupling constants (without EW corrections) of the  $Z$  boson to fermions,  $\sin^2 \theta_W^2 = s_W^2 = 1 - m_W^2/m_Z^2$  (on-shell scheme) and  $T_3^f$  denotes third component of the isospin.

The vector  $v_e, v_f$  and axial  $a_e, a_f$  couplings for leptons and quarks are defined with the formulas below:

$$\begin{aligned}
 v_e &= (2 \cdot T_3^e - 4 \cdot q_e \cdot s_W^2) / \Delta \\
 v_f &= (2 \cdot T_3^f - 4 \cdot q_f \cdot s_W^2) / \Delta \\
 a_e &= (2 \cdot T_3^e) / \Delta \\
 a_f &= (2 \cdot T_3^f) / \Delta
 \end{aligned} \tag{1}$$

where

$$\Delta = \sqrt{16 \cdot s_W^2 \cdot (1 - s_W^2)} \tag{2}$$

With this notation, matrix element for the  $q\bar{q} \rightarrow Z/\gamma^* \rightarrow l^+l^-$ ,  $ME_{Born}$ , can be written as:

$$\begin{aligned}
 ME_{Born} &= [\bar{u}\gamma^\mu v g_{\mu\nu} \bar{v}\gamma^\nu u] \cdot (q_e \cdot q_f) \cdot \frac{\chi_\gamma(s)}{s} \\
 &+ [\bar{u}\gamma^\mu v g_{\mu\nu} \bar{\nu}\gamma^\nu u \cdot (v_e \cdot v_f) + \bar{u}\gamma^\mu v g_{\mu\nu} \bar{\nu}\gamma^\nu \gamma^5 u \cdot (v_e \cdot a_f) \\
 &+ \bar{u}\gamma^\mu \gamma^5 v g_{\mu\nu} \bar{\nu}\gamma^\nu u \cdot (a_e \cdot v_f) + \bar{u}\gamma^\mu \gamma^5 v g_{\mu\nu} \bar{\nu}\gamma^\nu \gamma^5 u \cdot (a_e \cdot a_f)] \cdot \frac{\chi_Z(s)}{s}
 \end{aligned} \tag{3}$$

$Z$ -boson and photon propagators read respectively as

$$\chi_\gamma(s) = 1 \tag{4}$$

$$\chi_Z(s) = \frac{G_\mu \dot{M}_Z^2}{\sqrt{2} \cdot 8\pi \cdot \alpha_{QED}(0)} \cdot \Delta^2 \cdot \frac{s}{s - M_Z^2 + i \cdot \Gamma_Z \cdot M_Z} \tag{5}$$

At the peak of resonance  $|\chi_Z(s)| \times (v_e \cdot v_f) > (q_e \cdot q_f)$  and as a consequence, angular distribution asymmetries of leptons are proportional to

$v_e = (2 \cdot T_3^e - 4 \cdot q_e \cdot s_W^2)$ . This gives good sensitivity for  $s_W^2$  measurement.

Above and below resonance we are sensitive to lepton and quark charge instead ...



Born cross-section, for  $q\bar{q} \rightarrow Z/\gamma^* \rightarrow \ell^+\ell^-$  can be expressed as:

$$\frac{d\sigma_{Born}^{q\bar{q}}}{d\cos\theta}(s, \cos\theta, p) = (1+\cos^2\theta)F_0(s) + 2\cos\theta F_1(s) - p[(1+\cos^2\theta)F_2(s) + 2\cos\theta F_3(s)] \quad (6)$$

$p$  polarization of the outgoing leptons. The  $\cos\theta$  of angle between incoming quark and outgoing lepton in the rest frame of outgoing leptons. All rely on second order spherical harmonics. Also with transverse spin. Form-factors read:

$$\begin{aligned} F_0(s) &= \frac{\pi\alpha^2}{2s} [q_f^2 q_\ell^2 \cdot \chi_\gamma^2(s) + 2 \cdot \chi_\gamma(s) \text{Re}\chi_Z(s) q_f q_\ell v_f v_\ell + |\chi_Z^2(s)|^2 (v_f^2 + a_f^2)(v_\ell^2 + a_\ell^2)], \\ F_1(s) &= \frac{\pi\alpha^2}{2s} [2\chi_\gamma(s) \text{Re}\chi(s) q_f q_\ell v_f v_\ell + |\chi^2(s)|^2 2v_f a_f 2v_\ell a_\ell], \\ F_2(s) &= \frac{\pi\alpha^2}{2s} [2\chi_\gamma(s) \text{Re}\chi(s) q_f q_\ell v_f v_\ell + |\chi^2(s)|^2 (v_f^2 + a_f^2) 2v_\ell a_\ell], \\ F_3(s) &= \frac{\pi\alpha^2}{2s} [2\chi_\gamma(s) \text{Re}\chi(s) q_f q_\ell v_f v_\ell + |\chi^2(s)|^2 (v_f^2 + a_f^2) 2v_\ell a_\ell], \end{aligned} \quad (7)$$

## *Why is it of interest?*

1. Condition:  $s_W^2 = 1 - m_W^2/m_Z^2$  is important for some gauge cancellations, in case of multileg processes, but at the same time bring inconsistencies with measurements:
2. either  $m_W$  must be off by many experimental errors
3. or electroweak observables such as  $A_{FB}$  or  $P_\tau$  by 50 % of their measurable values.
4. Nonetheless such on mass shell scheme is used by many programs of importance for QCD phenomenology.
5. Technical solutions using calculation of correcting weights are of interest.
6. **BY-PRODUCT: separate leptonic degrees of freedom from the hadronic ones.**

## Mustraal frame

[18] F. A. Berends, R. Kleiss, and S. Jadach, *Comput. Phys. Commun.* **29** (1983) 185–200.

**Mustraal: Monte Carlo for  $e^+ e^- \rightarrow \mu^+ \mu^- (\gamma)$**

$$\begin{aligned} s &= 2p_+ \cdot p_-, & t &= 2p_+ \cdot q_+, & u &= 2p_+ \cdot q_- \\ s' &= 2q_+ \cdot q_-, & t' &= 2p_- \cdot q_-, & u' &= 2p_- \cdot q_+ \end{aligned}$$

$$\sigma_{\text{hard}} = \int d\tau (X_i + X_f + X_{\text{int}}),$$

The explicit forms of the three terms in  $\sigma_{\text{hard}}$  read:

$$X_i = \frac{Q^2 \alpha}{4\pi^2 s} \frac{1 - \Delta}{k_+ k_-} s'^2 \left[ \frac{d\sigma^B}{d\Omega}(s', t, u) + \frac{d\sigma^B}{d\Omega}(s', t', u') \right], \quad (3.4)$$

$$X_f = \frac{Q'^2 \alpha}{4\pi^2 s} \frac{1 - \Delta'}{k'_+ k'_-} s^2 \left[ \frac{d\sigma^B}{d\Omega}(s, t, u') + \frac{d\sigma^B}{d\Omega}(s, t', u) \right], \quad (3.5)$$

$$\begin{aligned} X_{\text{int}} &= \frac{QQ'\alpha}{4\pi^2 s} W \frac{\alpha^2}{2ss'} \left[ (u^2 + u'^2 + t^2 + t'^2) \tilde{f}(s, s') + \frac{1}{2}(u^2 + u'^2 - t^2 - t'^2) \tilde{g}(s, s') \right] \\ &+ \frac{QQ'\alpha^3}{4\pi^2 s} \frac{(s - s') M \Gamma}{k_+ k_- k'_+ k'_-} \epsilon_{\mu\nu\rho\sigma} p_+^\mu p_-^\nu q_+^\rho q_-^\sigma \left[ \tilde{E}(s, s')(t^2 - t'^2) + \tilde{F}(s, s')(u^2 - u'^2) \right], \end{aligned} \quad (3.6)$$

Resulting optimal frame used to minimise higher order corrections from initial state radiation in  $e^+e^- \rightarrow Z/\gamma^* \rightarrow \mu \mu$  for algorithms of genuine EW corrections implementation in LEP time Monte Carlo's like Koral Z.

## Extending definition of Mustraal frame

- We extended this frame to  $pp \rightarrow l^+ l^- j (j)$  case
  - reconstruct  $x_1, x_2$  of incoming partons from final state kinematics (information on jets used)
  - assume the quark is following  $x_1$  direction (equivalent to what done in CS frame)
  - calculate  $(\theta_1, \phi_1), (\theta_2, \phi_2)$  of two Born's, weight with probability calculated not using couplings

$$wt_1 = \frac{E_{p1}^2(1 + \cos \theta_1^2)}{E_{p1}^2(1 + \cos \theta_1^2) + E_{p2}^2(1 + \cos \theta_2^2)}, \quad wt_2 = \frac{E_{p2}^2(1 + \cos \theta_2^2)}{E_{p1}^2(1 + \cos \theta_1^2) + E_{p2}^2(1 + \cos \theta_2^2)}$$

3

- We can see that distribution is a stochastic sum of Born-like distributions with coefficients which are *positive thus like probabilities*. **But it is only QED!**

*What are the Limitations and Perspectives for case of QCD jets:*

- E. Mirkes and J. Ohnemus, “Angular distributions of Drell-Yan lepton pairs at the Tevatron: Order  $\alpha - s^2$  corrections and Monte Carlo studies,” PRD **51** (1995) 4891
- R. Kleiss, “Inherent Limitations in the Effective Beam Technique for Algorithmic Solutions to Radiative Corrections,” Nucl. Phys. B **347**, 67 (1990).

If jets are present definition of angles  $\theta, \phi$ , of effective Born becomes an **issue**. However, only  $\alpha_s^2 \sim 0.01$  corrections to spherical harmonics independently of the choice of reference frame,  $p_T$  transverse momentum of  $\tau\tau$ -pair,  $Y$  rapidity:

$$\frac{d\sigma}{dp_T^2 dY d\cos\theta d\phi} = \frac{3}{16\pi} \frac{d\sigma^{U+L}}{dp_T^2 dY} [(1 + \cos^2\theta) + 1/2 A_0(1 - 3\cos^2\theta) + A_1 \sin(2\theta) \cos\phi + 1/2 A_2 \sin^2\theta \cos(2\phi) + A_3 \sin\theta \cos\phi + A_4 \cos\theta + A_5 \sin^2\theta \sin(2\phi) + A_6 \sin(2\theta) \sin\phi + A_7 \sin\theta \sin\phi] \quad (8)$$

We will use samples of events generated with the MadGraph matrix element Monte Carlo for Drell-Yan production of  $\tau$ -lepton pairs, with  $m_{\tau\tau} = 80 - 100$  GeV and 13 TeV  $pp$  collisions. Lowest order spin amplitudes are used in this program for the parton level process. For the EW scheme we have used default initialisation of MadGraph, with on-shell definition of  $\sin^2 \theta_W = 1 - m_W^2/m_Z^2 = 0.2222$ , which determines value of the axial coupling for leptons and quarks to the Z-boson. The incoming partons are distributed accordingly to PDFs (using CTEQ6L1 PDFs).

- We use the Monte Carlo sample of  $Z \rightarrow \ell^\pm \ell^\mp$  events and extract angular coefficients of Eq. (8) using moments methods [Mirkes:1994]. The moment of a polynomial  $P_i(\cos \theta, \phi)$ , integrated over a specific range of  $p_T, Y$  defines:

$$\langle P_i(\cos \theta, \phi) \rangle = \frac{\int_{-1}^1 d \cos \theta \int_0^{2\pi} d\phi P_i(\cos \theta, \phi) d\sigma(\cos \theta, \phi)}{\int_{-1}^1 d \cos \theta \int_0^{2\pi} d\phi d\sigma(\cos \theta, \phi)}. \quad (9)$$

- Owing to the orthogonality of the spherical polynomials of Eq. (8), the weighted average of the angular distributions with respect to any specific polynomial, Eq. (9), isolates its corresponding coefficient, averaged over some phase-space region.
- We obtain:

$$\begin{aligned} \left\langle \frac{1}{2}(1 - 3 \cos^2 \theta) \right\rangle &= \frac{3}{20} (A_0 - \frac{2}{3}); & \langle \sin 2\theta \cos \phi \rangle &= \frac{1}{5} A_1; \\ \langle \sin^2 \theta \cos 2\phi \rangle &= \frac{1}{10} A_2; & \langle \sin \theta \cos \phi \rangle &= \frac{1}{4} A_3; \\ \langle \cos \theta \rangle &= \frac{1}{4} A_4; & \langle \sin^2 \theta \sin 2\phi \rangle &= \frac{1}{5} A_5; \\ \langle \sin 2\theta \sin \phi \rangle &= \frac{1}{5} A_6; & \langle \sin \theta \sin \phi \rangle &= \frac{1}{4} A_7. \end{aligned} \quad (10)$$

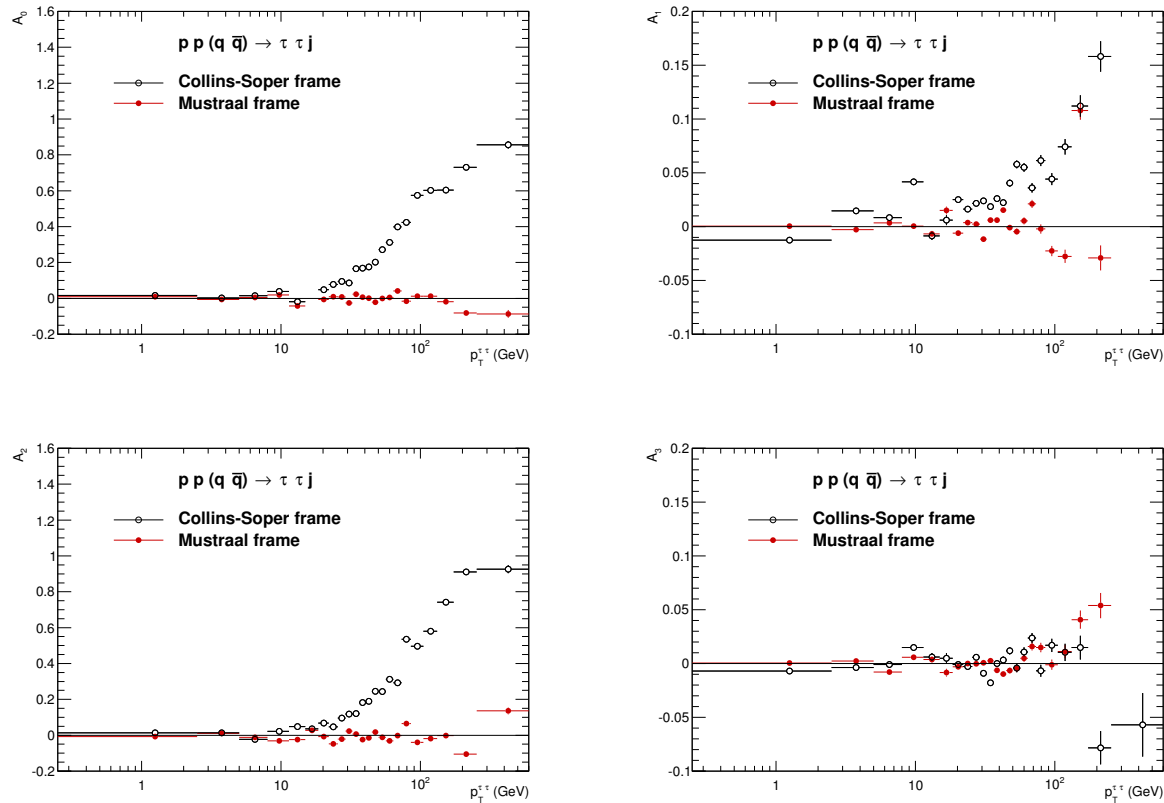


Figure 4: The  $A_i$  coefficients of Eq. (8) calculated in Collins-Soper (black) and in Mustraal (red) frames for  $pp(q\bar{q}) \rightarrow \tau\tau j$  process generated with MadGraph. From **Eur.Phys.J. C76 (2016) 473** . Tree level ME+ collinear pdf's used for analyzed sample.

**Mustraal frame works PERFECT. Note that our probabilities/weights were stripped from dependence on EW parameters. It could be not so, but IS SO**



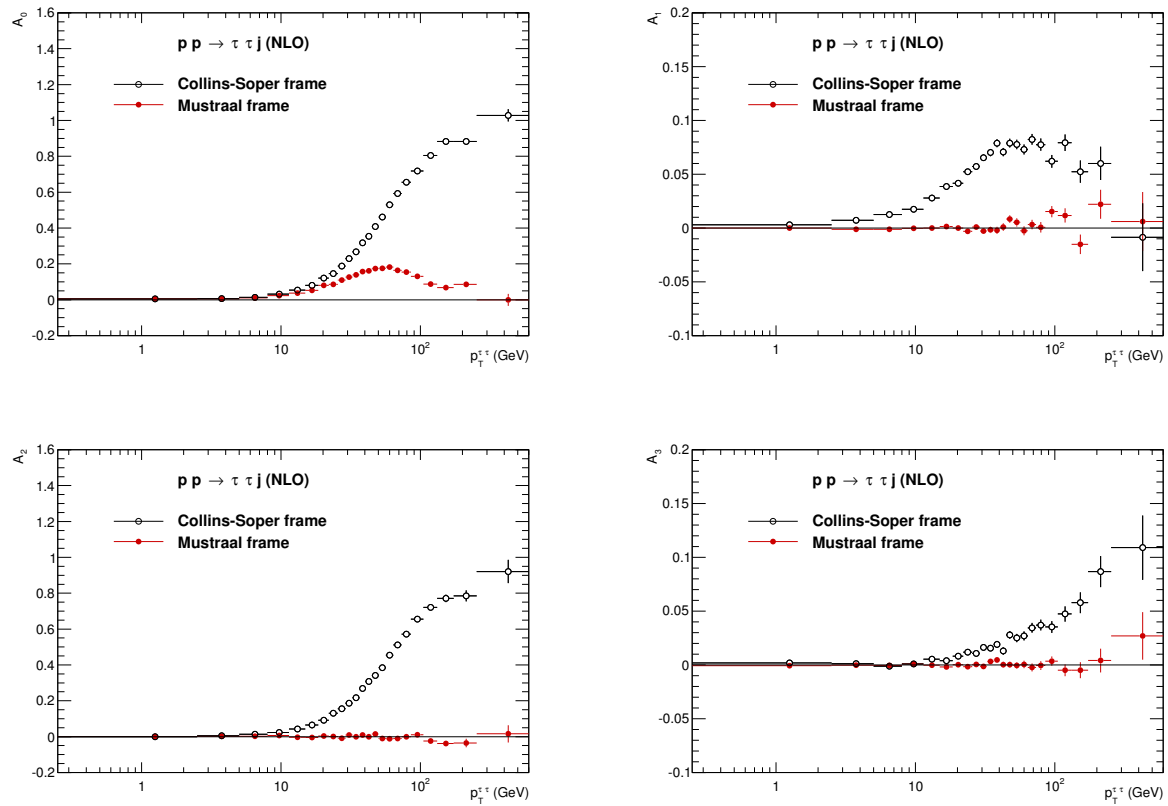


Figure 5: The  $A_i$  coefficients of Eq. (8)) calculated in Collins-Soper (black) and in Mustraal (red) frames for  $pp \rightarrow \tau\tau j$  (NLO) process generated with Powheg+MiNLO. From **Eur.Phys.J. C76 (2016) 473**.

Note that for complete QCD Z+1jet NLO plus MiNLO pattern remained!

- The choice of Mustraal frame is result of careful study of single photon (gluon) emission)
- In Ref of 1982 it was shown, that differential distribution is a sum of two born-like distributions convoluted with emission factors.
- This is a consequence of Lorentz group representation and that is why it generalizes to the case of double gluon or even double parton emissions.
- **Impact of jets on effective Born is like change of orientation of frames.**
- This observation is helpful to separate leptonic degrees of freedom from the ones of hadronic jets, where modelling could bring problems.

- Still another example, where multi-dimensionality was important for precision
- LEP times precision breakthrough: from 2-3 % on luminosity measurement down to 0.041 %.
- In principle it was just counting experiment.
- Once precision improved, nothing remained simple. Simulation became essential.
- Thanks to introduction into simulation of final states consisting of electron and soft collinear photons...
- **... one could identify that corresponding events was not a detector malfunction.**
- I can not find the appropriate (plastic) slide of that times.

- I have presented general principle of CP Higgs parity measurement in  $H \rightarrow \tau\tau$  decays
- I have demonstrated how computer algebraic methods or pattern recognition techniques (Machine Learning) was useful to manage observable of multi-dimensional nature.
- I have adressed the facorization properties which were helpful to design observables and later to control if NN response was not 'too good'.
- Starting from Higgs decay I have moved to discuss similar factorization properties od DY. They may be also offering a necessary benchmark path for the forthcoming applications.
- I recalled result of LEP sub permille precision level observable. There, details of multidimensional nature were of a great importance.
- **This can be only more true for the future applications. ML learning can enforce importance of multidimensional predictions and necessity of phase-space  $\times$  ME approach.**

- Case of Higgs parity is 'easy'. Dynamic of tau decays is known, Higgs production separates from decay
- case of its background is manageable: Dynamic of tau decays is still known, and  $Z\gamma^*$  production convolutes with decays through spherical harmonics of the second order.
- Case of jet physics is by far more complex, see e.g. afternoon session talks of SMP-J annual workshop, afternoon talks: <https://indico.cern.ch/event/684248/>
- Managing detector granularity to improve precision? I think it is good set of topics.