Bad news: easy nice things are done already.

From optimal variables to machine learning and back: case of Z polarization and Higgs boson CP

Z. Was

- **(1)** We want to measure some quantities carrying good quality physics message
- **(2)** Good example is $\sin^2\theta_W^{effective}$ from τ polarization or angular distributions of leptons or Higgs boson parity.
- **(3)** Usually such quantities can be neither measured nor predicted
- **(4)** Of course if precision is required
- **(5)** In the following we will concentrate on how to “have cake and eat cake”
- **(6)** Message is: do not break rules, bend them
- **(7)** And later clear the mess: **sufficently** well.
- **(8)** Understand responsabilities and your domain. How to combine with others work.
- **(9)** Be ready for difficulties of: mathematics, physics, algorithms, computing, detectors

Optimal variables $\rightarrow$ Machine Learning $\rightarrow$ and back

1. Dream: clearly defined quantities easy to measure and interpret.
2. Often it starts like that, but later come complications:
   (a) experimental side: multi-dimensional, backgrounds, detector structure.
   (b) theory side: multi-dimensional due to jets and/or QED bremsstrahlung, quantum loops etc.
   (c) More serious quantum entaglement, mass $\rightarrow$
   (d) simplicity ? gone
3. BUT: Quantum Field Theory, the best to explain Nature (or New Physics).
4. ALSO: Acceleratos and detectors are marvels too.
5. We want details of Nature foundations.
6. ML is good for everything? If yes, what is the price.

Basic pictures $\longrightarrow$ faults $\longrightarrow$ recoveries

1. $$d\sigma = \Big(\sum_{\lambda_1\lambda_2} |\mathcal{M}^{prod}|^2\Big)\Big(\sum_{\lambda_1} |\mathcal{M}^{\tau^+}|^2\Big)\Big(\sum_{\lambda_2} |\mathcal{M}^{\tau^-}|^2\Big) wt_{spin} d\Omega_{prod}\, d\Omega_{\tau^+}\, d\Omega_{\tau^-}.$$

2. Formula above is correct, useless too. We work with:

3. $$d\sigma = \sum_{flav.} \int dx_1 dx_2 f(x_1,...) f(x_2,...) d\Omega_{prod}^{part.\ lev.}\, d\Omega_{\tau^+}\, d\Omega_{\tau^-}$$
$$\Big(\sum_{\lambda_1,\lambda_2} |\mathcal{M}^{prod}_{part.\ lev.}|^2\Big)\Big(\sum_{\lambda_1} |\mathcal{M}^{\tau^+}|^2\Big)\Big(\sum_{\lambda_2} |\mathcal{M}^{\tau^-}|^2\Big) wt_{spin}.$$

4. $$d\sigma_{Born}(x_1,x_2,\hat{s},\cos\theta) = \sum_{q_f,\bar{q}_f} [f^{q_f}(x_1,...) f^{\bar{q}_f}(x_2,...) d\sigma^{q_f\bar{q}_f}_{Born}(\hat{s},\cos\theta)$$
$$+ f^{\bar{q}_f}(x_1,...) f^{q_f}(x_2,...) d\sigma^{q_f\bar{q}_f}_{Born}(\hat{s},-\cos\theta)],$$

5. Nearly all is faulty. What can be corrected. What can be avoided.

Optimal variables $\rightarrow$ Machine Learning $\rightarrow$ and back

1. 50 years ago: measure single quantity
2. 30 years ago: look for most sensitive to physics, 1-dimensional distributions, hide some complexity in shapes
3. Prepare input for fits, with corrections breaking the picture parametrized also, or better estimate that they are negligible.
4. what does it mean for today when we attempt to use multi dimensional signatures with the help of e.g. Machine Learning techniques?
5. Let's go to practicalities of: CP-parity of Higgs, with τ decays and later τ polarization for $\sin^2\theta_W$ measurement.
6. Latter is more complicated because precision is higher.

Properties of matrix elements.

The Higgs boson's parity is imprinted in M.E.

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

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

is sensitive to the $\tau^\pm$ polarization vectors s^{τ^-} and s^{τ^+} (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 an 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: also from M.E.

- Higgs boson Yukawa coupling expressed with the help of the scalar–pseudo-scalar mixing angle ϕ

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

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

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

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

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

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

Reality to face...

- **Easy because:**
  - Binary classification: scalar/pseudoscalar or at most 1-parameter.
  - High precision is not required.
  - High statistics process $Z \to \tau^+\tau^-$ available for all kind of feasability detector/bacrounds etcstudies
- **Difficult because:**
  - Small samples
  - τ polarization vectors can not be measured directly ...
  - ... only through τ decays; many channels, many backgrounds.
  - Requires measurement of nearly overlapping $\pi^\pm$, π^0 and/or reconstruction of ν_τ momentum from kinematical constraints and decay vertex position is also required.

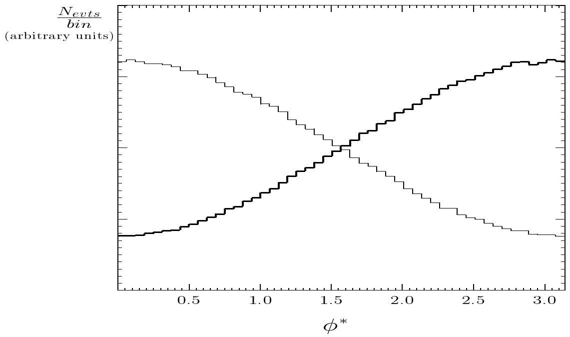
Tau-decay M.E. and attempts on observable.

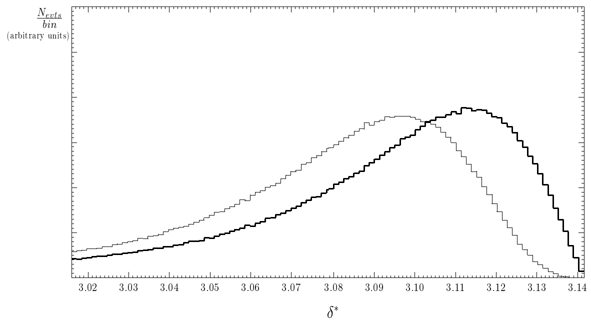
Transverse spin correlations through τ decays: ASYM or 1-DIM distr.

- *Case of $\tau \to \pi\nu_\tau$ decay,* $\mathcal{BR}(\tau \to \rho\nu_\tau) = 10\%$, *also M.E. expressed.*
- *First Measurable. Note that h^i is defined in τ r.f, and always $|\vec{h}| = 1$.*

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

$$h^i = p^i$$





- **Left:** *τ decay channel independent distribution of polarimetric h^i.* **Right:** *Acollinearity of π^+ and π^- is perfect. All sensitivity is transmitted to acollinearity distribution. Problem: one needs to reconstruct perfectly acollinearity of $\pi^+\pi^-$ in H rest-frame.* **Watch the scale** *Observable was realistic for Higgs of much lower mass !!!*

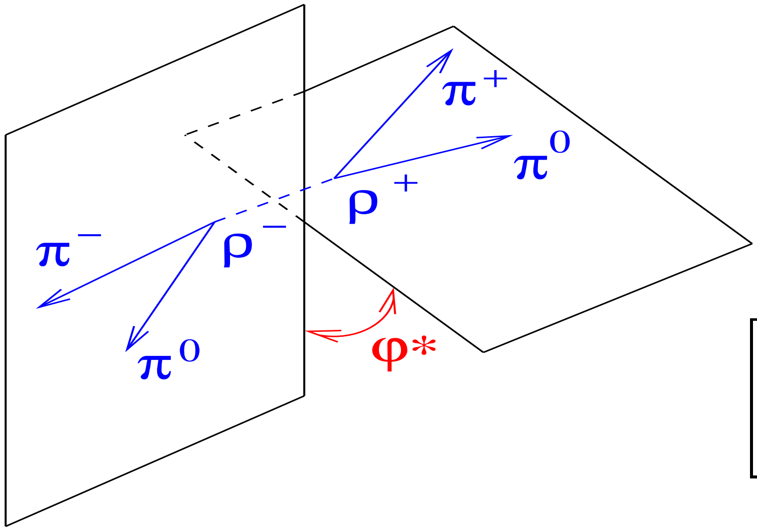
Transverse spin correlations through τ decays

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

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

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

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



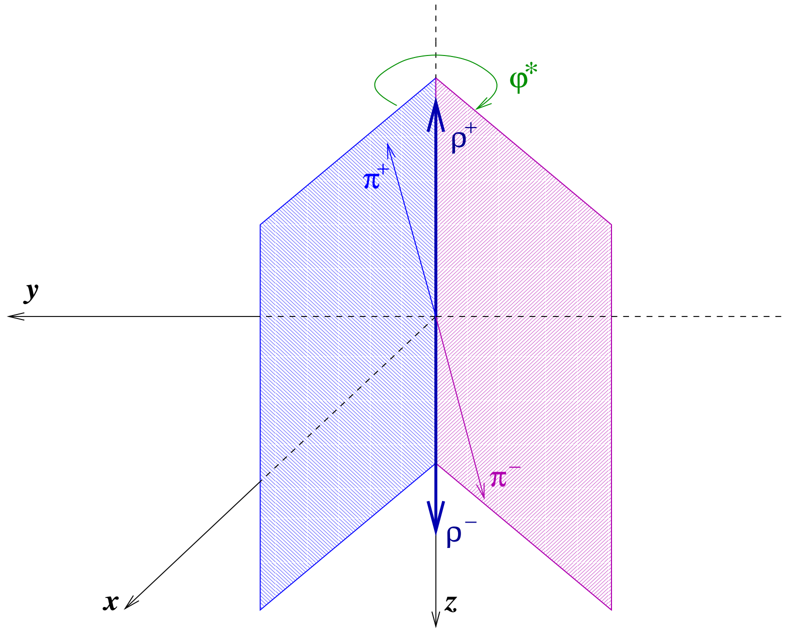
$$y_1 y_2 > 0\,; \qquad y_1 y_2 < 0 \text{ (in } \tau^\pm \text{ r.f.'s)}$$

$$y_1 = \frac{E_{\pi^+} - E_{\pi^0}}{E_{\pi^+} + E_{\pi^0}}\,; \qquad y_2 = \frac{E_{\pi^-} - E_{\pi^0}}{E_{\pi^-} + E_{\pi^0}}.$$

Observable of visible products in $H \to \tau^+\tau^- \to \pi^+\pi^0 \;\; \pi^-\pi$

Optimal Observable Mixed Scalar—Pseudoscalar Case

- For mixing angle ϕ, transverse component of τ^+ spin polarization vector is correlated with the one of τ^- rotated by angle 2ϕ.
- 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: $\mathbf{n}_\pm = \mathbf{p}_{\pi^\pm} \times \mathbf{p}_{\pi^0}$

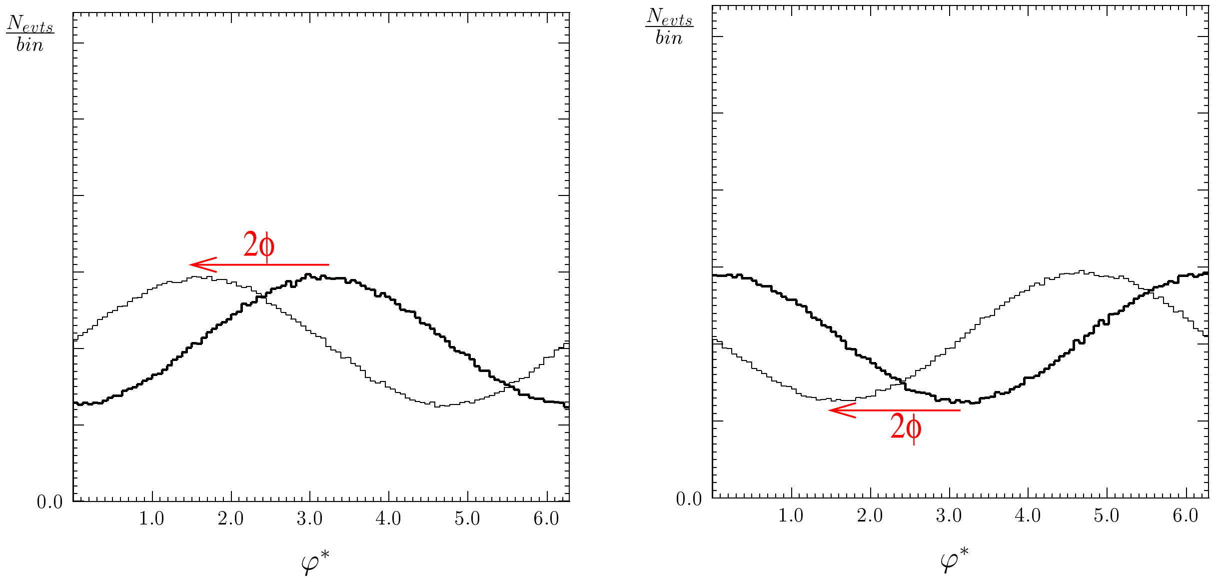
Find the sign of $\mathbf{p}_{\pi^-} \cdot \mathbf{n}_+$

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

Otherwise $2\pi - \varphi^*$

Observable of visible products in $H \to \tau^+\tau^- \to \pi^+\pi^0\ \pi^-\pi$

Old attempts, at the end 1-dim plot 'easy' to understand



- *Only events where the signs of y1 and y2 are the same whether calculated using the method without or with the help of the τ 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 $1ab^{-1}$ and 350 GeV CMS.

On dimensionality

Observable was constructed from 3 dimension distribution.

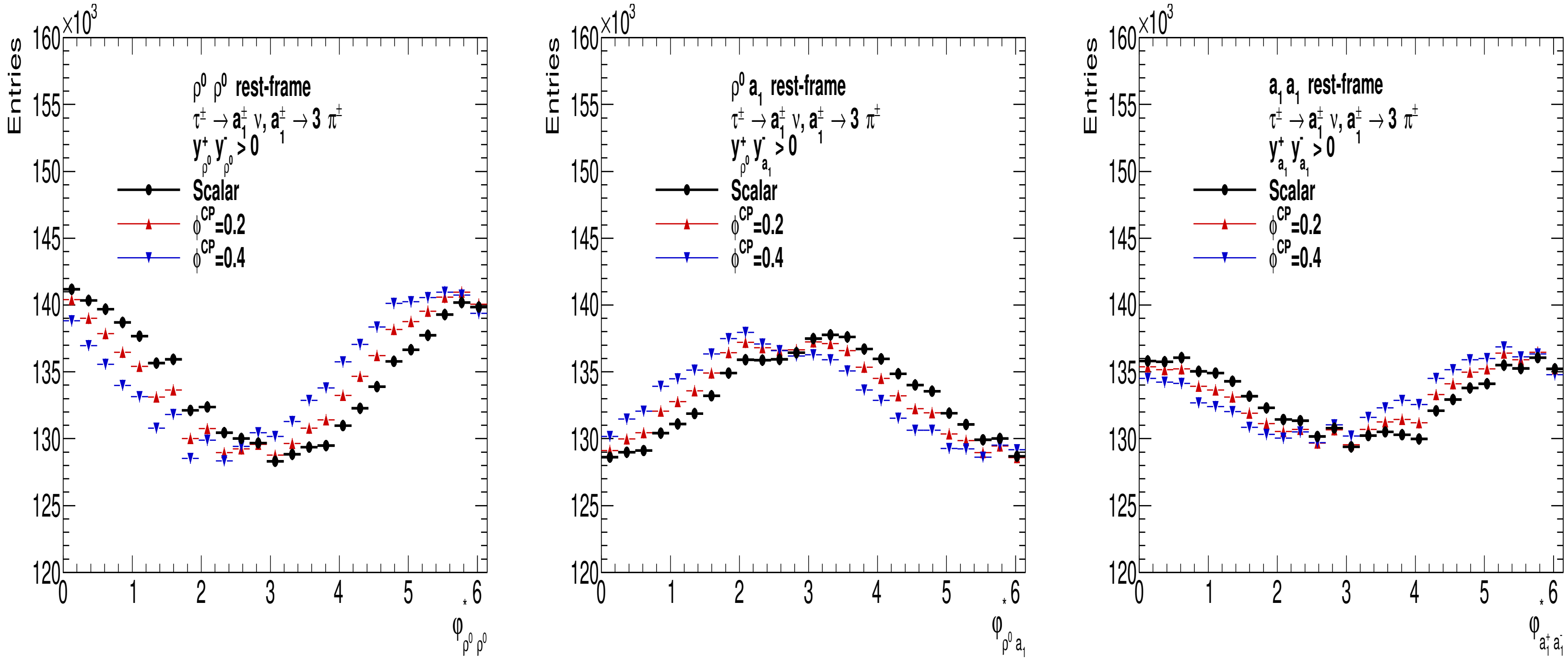
Even if plots were 1 dimensional

To progress:

- $\mathcal{BR}(\tau \to \rho\nu_\tau) = 25\%$, that mean 6$\%$ of $H \to \tau\tau$. Why not use other decay modes? They all have (in principle) the same sensitivity to spin: J. H. Kuhn, Phys. Rev. D52 (1995) 3128, but in practice ν_τ is not observable and:
- from the π^-, π^0 we can define one plane for acoplanarity,
- from the π^-, π^-, π^+ we can define four such planes.
- Each plane bring its own y_i variable to avoid cancellations due to properties of τ decay ME.

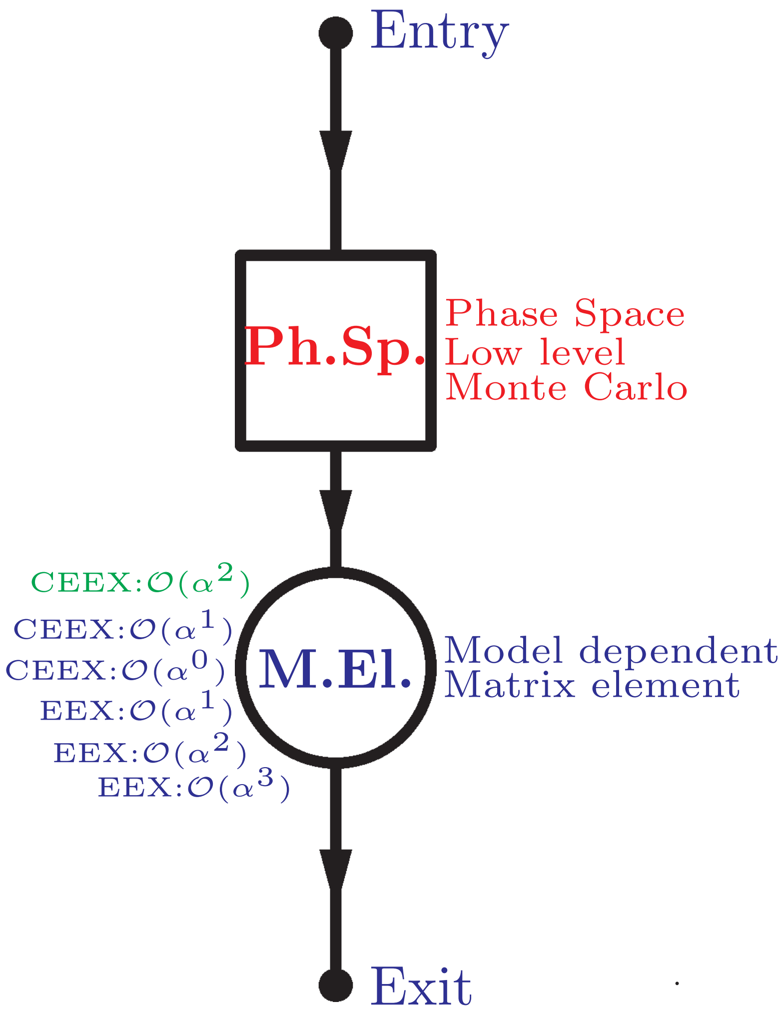
On dimensionality

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 ϕ^{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. Physics model depends on 1 parameter only ϕ^{CP} mixing scalar pseudo-scalar angle, which brings linear shift. I remained frustrated for 15 years, how to digest...



ML: my first solution.

Textbook principle “matrix element × full phase space” useful



- Phase-space Monte Carlo module producing “raw events”.
- Library of models for provides input for “model weight”
- **The scalar from pseudo-scalar distinguished by M.E. weight attributed to each event**
- Ratios define probability that event could be scalar or pseudoscalar Higgs.
- Convenient for ML training sample.

ML: my first solution.

- I can not present ML technology.
- I will flash some results only.
- Essentially probabilities that Network will identify event to be scalar, when it was scalar.
- 0.5 means random choce. 1.0 would mean certainty.
- Anything in-between was something useful.
- To get classification I had to:
- boost events to rest frame of all visible objects combined,
- rotate all to set τ^+ primary decay resonance along z axis.

Features/variables	$\rho^\pm - \rho^\mp$ $\rho^\pm \to \pi^0\,\pi^\pm$	$a_1^\pm - \rho^\mp$ $a_1^\pm \to \rho^0\pi^\mp,\ \rho^0 \to \pi^+\pi^-$ $\rho^\mp \to \pi^0\,\pi^\mp$	$a_1^\pm - a_1^\mp$ $a_1^\pm \to \rho^0\pi^\pm,$ $\rho^0 \to \pi^+\pi^-$
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
4-vectors	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).

$\varphi^*_{i,k}$ and y_i: expert variables In rest frame of all visible, aligned along z. Essential for measure of event distance.

With detector smearing: Phys.Rev. D96 (2017) 073002

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

Table 2: AUC for NN to separate scalar and pseudo-scalar hypotheses. Inputs with a ✓used. Results in column “Ideal" - from NNs trained/used with particle-level simulation, in column “Smeared" - from NNs trained/used with smearing. NN trained on smeared samples, for used on exact samples give similar results as “Ideal".

$H \to \tau\tau; \tau \to 2(3)\pi$ channels and ML

1. We have played with input, and we have observed:
2. Our precious expert variables were not necessary from some point
3. But seemingly trivial overall boosts and rotations were indispensable
4. Only some time later we understood why: network required help to separate longitudinal from transverse degrees of freedom.
5. There was not problem that some variables were then systematically big or small. Such properties were easy for NN to understand. Re scaling was in the system.
6. It does not need to be always like that. It will be application domain dependent.
7. My training case was in a sense easy, we could get help from ME. calculation and adjust variable set accordingly.
8. Only some time after finishing work and after some studies of literature I understood this ML contexts.

Already from J. H. Kuhn, Phys. Rev. D52 (1995) 3128 is was clear that having all information on τ decays should give ideal 0.782 classification from all τ decay modes.

Missing neutrinos was a challenge, the following approaches can be listed:

- K. Desch, Z. Was and M. Worek, “Measuring the Higgs boson parity at a linear collider using the tau impact parameter and tau —> rho nu decay,” Eur. Phys. J. C **29** (2003) 491
- A. Rouge, “CP violation in a light Higgs boson decay from tau-spin correlations at a linear collider,” Phys. Lett. B **619**, 43 (2005)
- S. Berge, W. Bernreuther and S. Kirchner, “Prospects of constraining the Higgs boson’s CP nature in the tau decay channel at the LHC,” Phys. Rev. D **92**, 096012 (2015)

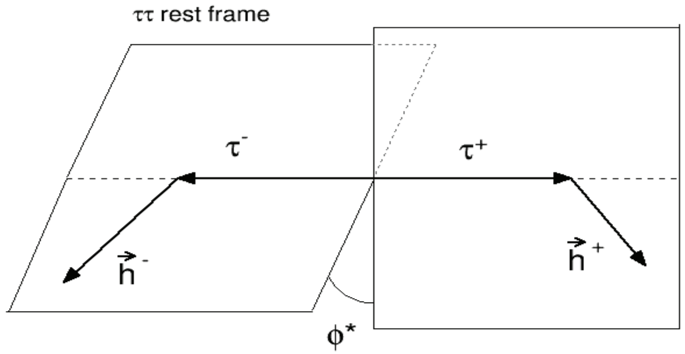
Is there anything that ML or new efforts can bring?

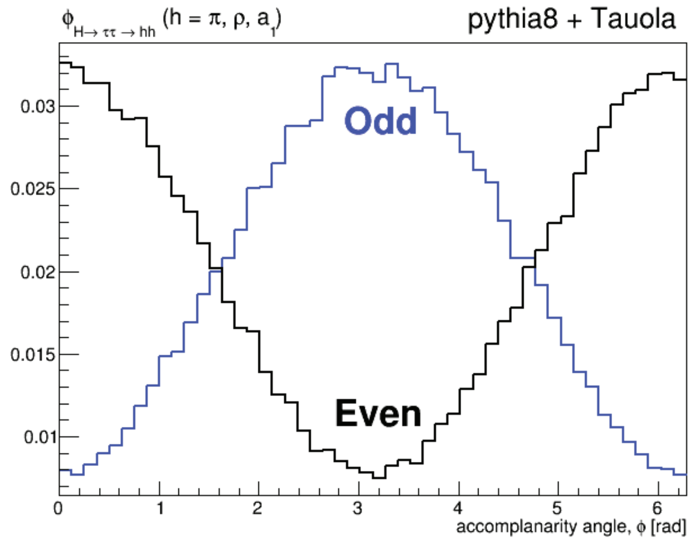
Our struggling with Impact parameter 15 years later

- A key for the improvement is to reconstruct ν_τ momenta: 6 variables.
- We have 3 good constraints (partly correlated) and two less precise:
  - $m_{\tau^\pm}$
  - m_H
  - $p_{T\ mis.}^{x,y}$
- We are missing two angles, which need to be obtained from impact parameter measurement.
- On the other hand, attempt to obtain optimal variable may be complicated by all kind of smearing. Kinematical constraints may be difficult.
- ML may be helpful? $\to$ arXiv:1812.08140
- Or, may be use ML to control h^i? rather than complete Higgs decay chain?
  Higgs CP is "easy" because matrix element weight straightforward to calculate.
- Then at the end we have 1-dim optimal variable, decay mode independent.

Applications of TauSpinner: Transverse spin correlation in H→ ττ at LHC (Full kinematic analysis)

- Alternatively one may try to estimate the invisible part of the polarimetric vectors of both τ leptons
- The estimation of the Higgs r.f. does not need to be excellent. We need at least some information on neutrinos momenta
- Accomplanarity observable can be build using direction of τ’s in H r.f. and polarimetric vectors
- Ideally carries the full analyzing power and irrespective of the τ decay channel





Factorizing effective Born from hadronic events

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.

Introduction to part 2 and 3 – $Z \to \tau\tau$ for $\sin^2\theta_W$

1. By far more complicated, expected precision on $\sin^2\theta_W$ 20 $\cdot 10^{-5}$ or better.
2. The energy and angle dependent $\sin^2\theta_W^{effective}$ must be used (electroweak loops).
3. Because we need to control spectra of τ decay products (ratios of $E_{products}/E_\tau$) we need to control QED final state bremsstrahlung
4. Because we do not collide quarks, we need to control PDFs and ...
5. ... intrinsic partons p_T and ...
6. ... QCD matrix element for jets and corresponding loop corrections and ...
7. ... underlying event.
8. Control may mean that we construct observables independent from ambiguities.
9. Monte Carlos may be helpful. Note that factorizations due to properties also due to lifetime of Z W H are helpful.
10. Note that the best measured at LHC are directions of leptons

properties of Effective Born

Let us start with the lowest order coupling constants (without EW corrections) of the Z boson to fermions, $\sin\theta_W^2 = s_W^2 = 1 - m_W^2/m_Z^2$ (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

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

properties of Effective Born

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

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

Z-boson and photon propagators read respectively as

$$\chi_\gamma(s) = 1 \quad (4)$$

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

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 ...

properties of Effective Born

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

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

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:

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

properties of Effective Born

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_τ 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 $e^+ e^- \to \mu^+ \mu^- \ (\gamma)$

$$s = 2p_+ \cdot p_-, \quad t = 2p_+ \cdot q_+, \quad u = 2p_+ \cdot q_-$$

$$s' = 2q_+ \cdot q_-, \quad t' = 2p_- \cdot q_-, \quad u' = 2p_- \cdot q_+$$

$$\sigma_{\mathrm{hard}} = \int \mathrm{d}\tau (X_{\mathrm{i}} + X_{\mathrm{f}} + X_{\mathrm{int}}),$$

The explicit forms of the three terms in σ_{hard} read:

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

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

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

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

2

Extending definition of Mustraal frame

- We extended this frame to pp -> $l^+ l^-$ 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 (θ_1, ϕ_1), (θ_2, ϕ_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

Effective Born and jets ..

- 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 θ, ϕ, 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:

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

Numerical results, Born recovery?

We will use samples of events generated with the `MadGraph` matrix element Monte Carlo for Drell-Yan production of τ-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).

Moments method

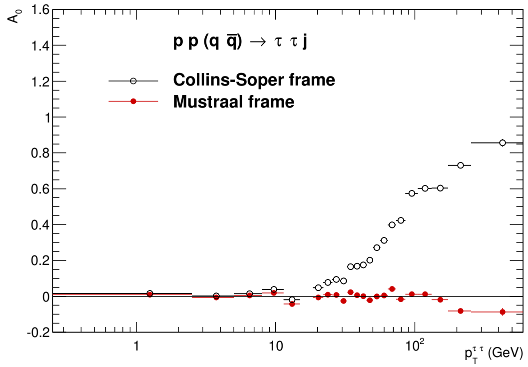
- We use the Monte Carlo sample of $Z \to \ell^\pm \ell^\mp$ events and extract angular coefficients of Eq. (8) using moments methods [Mirkes:1994]. The moment of a polynomial $P_i(\cos\theta, \phi)$, integrated over a specific range of p_T, Y defines:

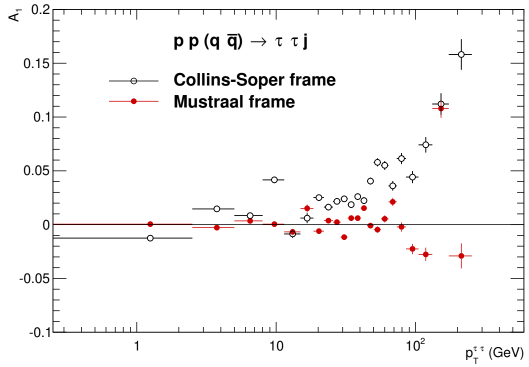
$$\langle P_i(\cos\theta,\phi)\rangle = \frac{\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi\, P_i(\cos\theta,\phi) d\sigma(\cos\theta,\phi)}{\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi\, d\sigma(\cos\theta,\phi)}. \quad (9)$$

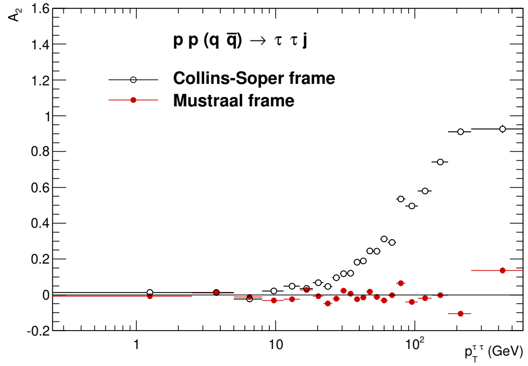
- Owing to the orthogonality of the spherical polynomials of Eq. (8), the weighted average of the angular distributions with respect to any specific polynomial, Eq. (9), isolates its corresponding coefficient, averaged over some phase-space region.

- We obtain:

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







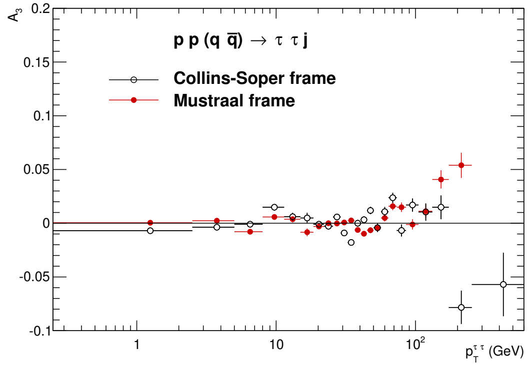


Figure 1: The A_i coefficients of Eq. (8)) calculated in Collins-Soper (black) and in `Mustraal` (red) frames for $pp(q\bar{q}) \to \tau\tau j$ process generated with `MadGraph`. From **Eur.Phys.J. C76 (2016) 473** . Tree level ME+ collinear pdf's used for analyzed sample.

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

Moments method

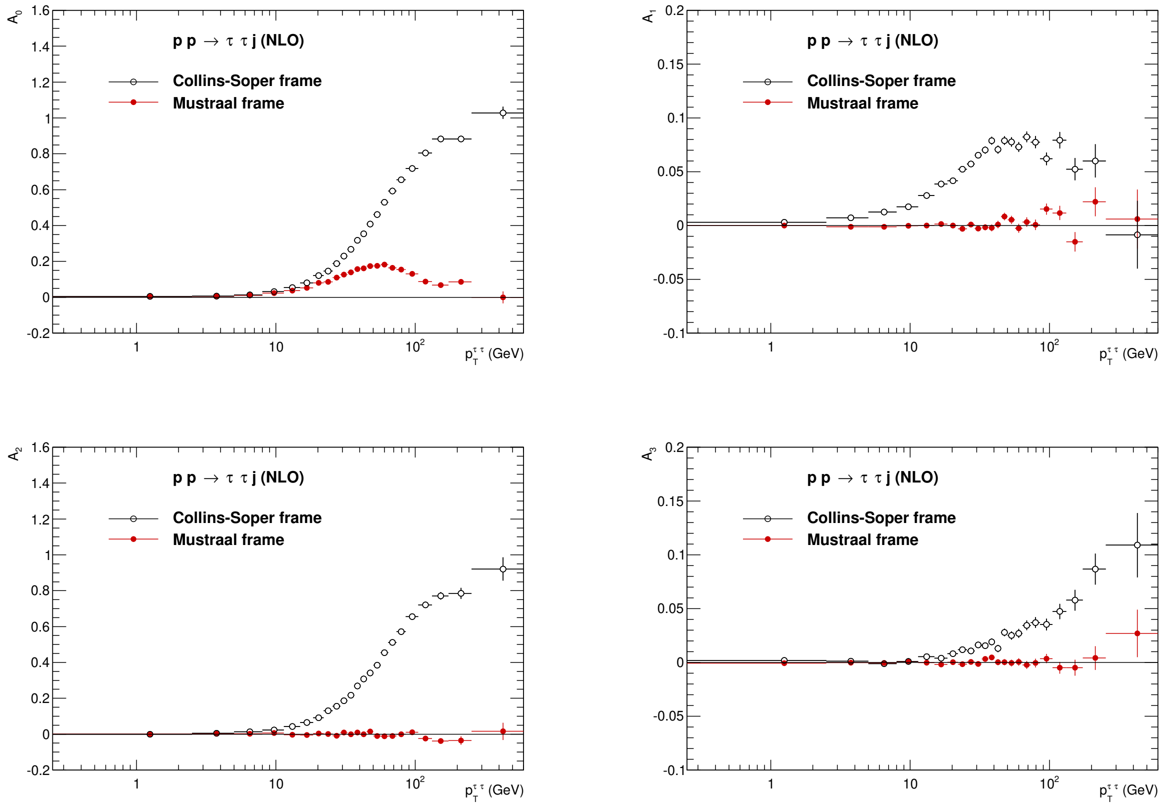


Figure 2: The A_i coefficients of Eq. (8)) calculated in Collins-Soper (black) and in `Mustraal` (red) frames for $pp \to \tau\tau j$ (NLO) process generated with `Powheg+MiNLO`. From **Eur.Phys.J. C76 (2016) 473**.

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.

$W \to l\nu_l$ production at LHC, 1609.02536

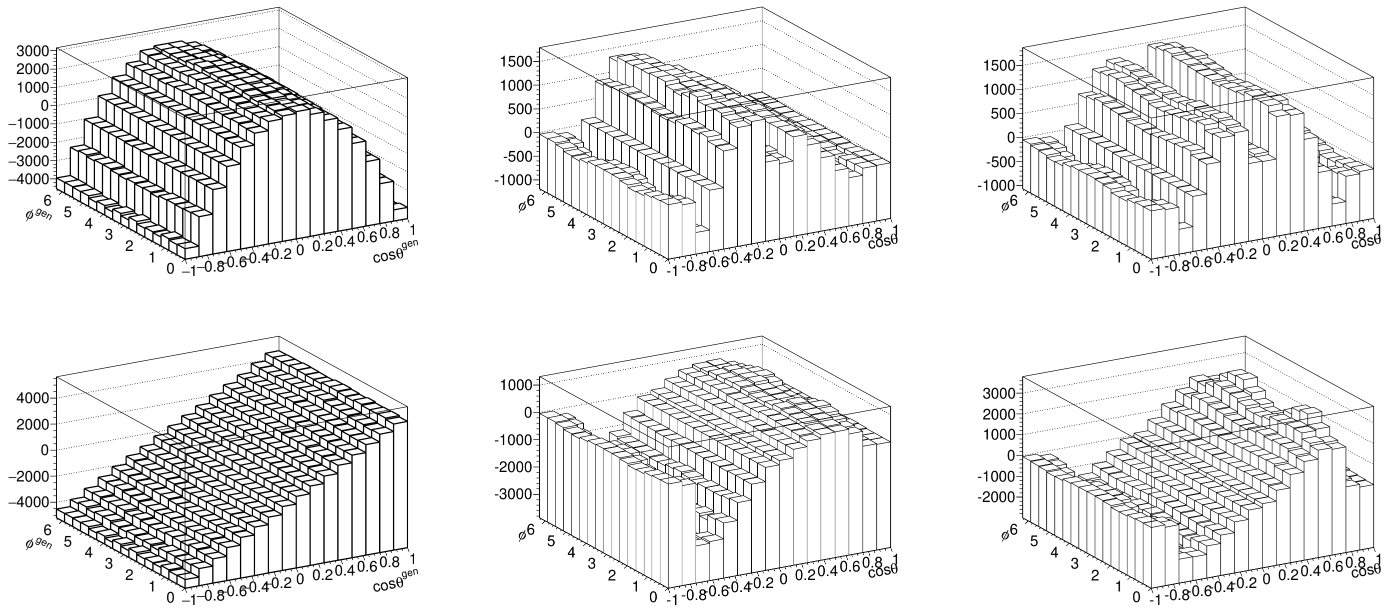


Figure 3: Analytical shape of the polynomial P_0 (top) and P_4 (bottom) in the full phase-space (left) and templates for polynomials after reconstructing p_Z^{ν} and fiducial selection for: W^- (middle) and W^+ (right). Original spherical harmonics of second order for $W \to l\nu$ decay angles are strongly deformed, but can be measured even for W. For the benefit of initial state hadronic interaction.

$W \to l\nu_l$ production at LHC, 1609.02536

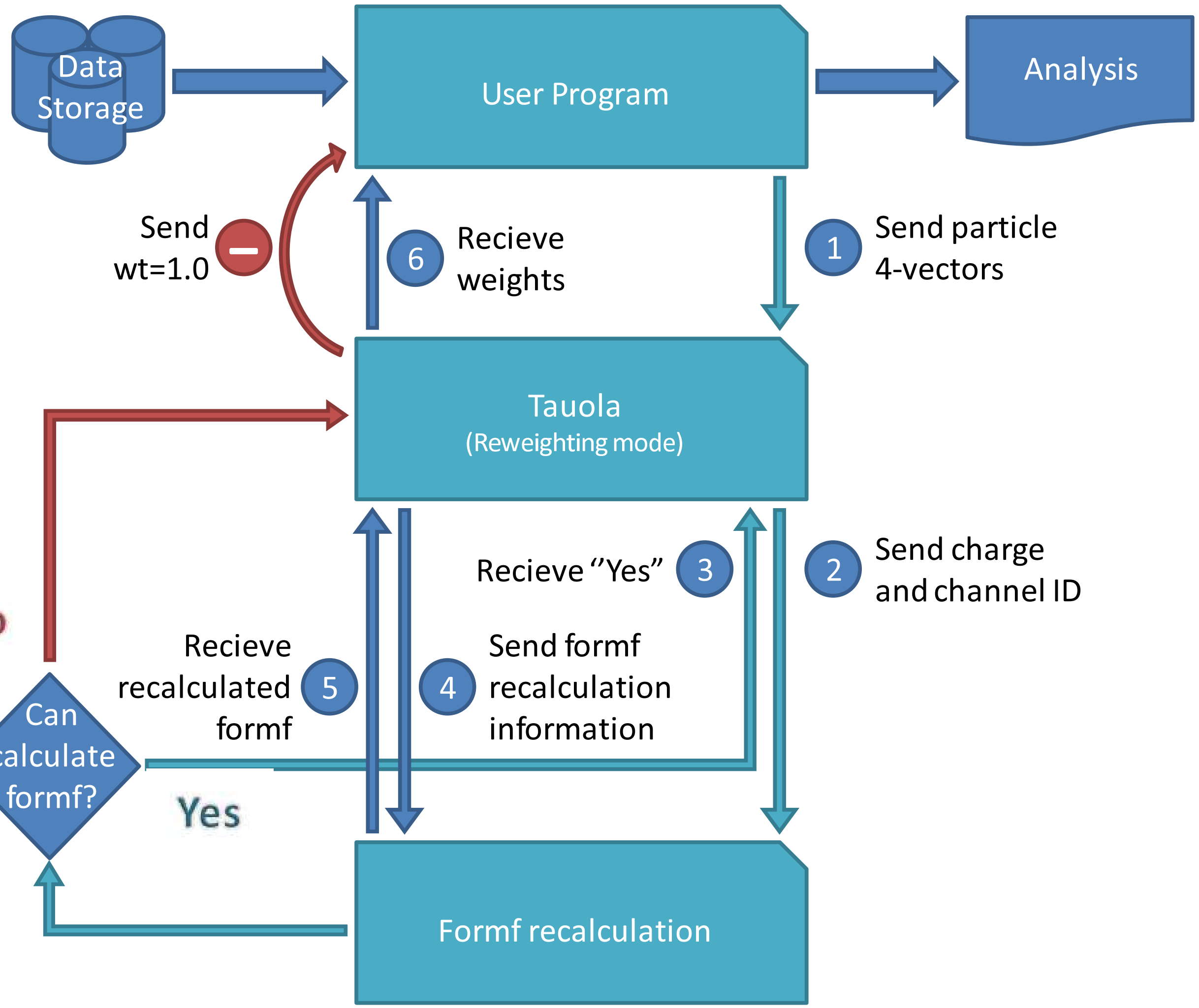


Figure 4: Flow chart for communication when already stored events are modified with the weights. Useful at LHC and at low energy applications as well.

Applications of TauSpinner: Longitudinal τ polarization in Z→ ττ at LHC

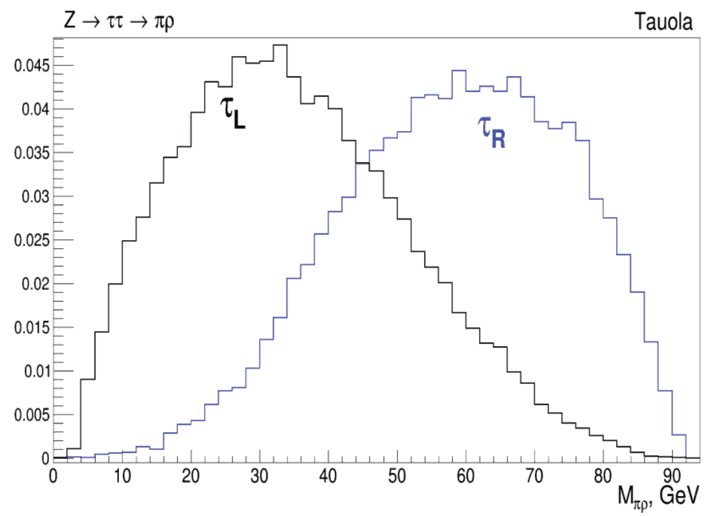
Extensive measurements of tau polarization have been performed at LEP.

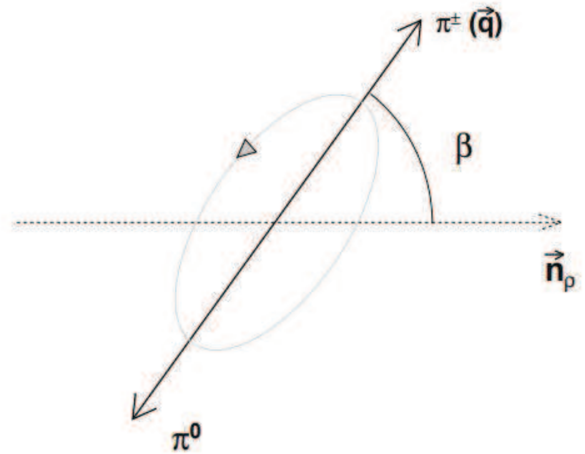
At LHC one may apply the visible analysis:

- Acollinearity and energy-energy correlation between decay products of both τ's

  R. Alemany, et.al., Nucl.Phys. B379 (1992) 3-23

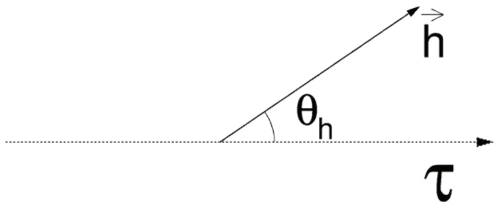
- Angles between measured decay products in τ→ρν and τ→a_1ν decays, for example angle β in τ→ρν decay.





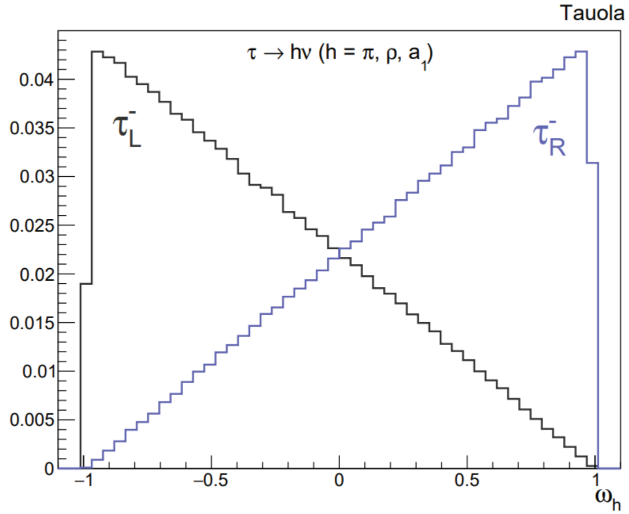
Applications of TauSpinner: Longitudinal τ polarization in Z→ ττ at LHC (Full kinematic analysis)

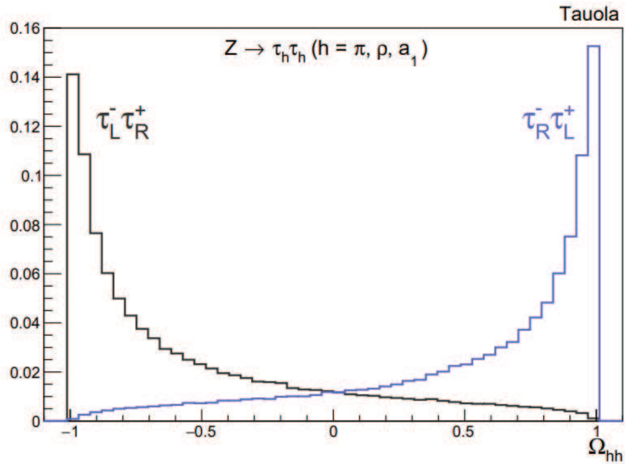
- A competitive precision to LEP can be achieved analysing all τ decays that can be identified at LHC
- Similarly to H→ ττ one may try to maximise the analyzing power reconstructing the event kinematic and the full polarimetric vector
- The observable is an angle between τ direction and the polarimetric vector



Corresponds to the optimal observable $\omega = cos\theta_h$
M. Davier, et.al., Phys.Lett. B306 (1993) 411-417

- ≈100% anti-correlation of τ leptons spins allows to further gain the analysing power





Conclusions and outlook

- My attempt was to demonstrate how phenomenology of some processes at LHC can be prepared.
- Higgs CP was 'easy' because it is scalar and production does not affect its decay
- Template methods for τ helicities. Can helicity attribution be trusted
- Template methods for lepton pair production using angular distribution of lepton directions in lepton pair frame as guiding feature.

Conclusions and outlook

Figure 5: 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.

Example of papers from huge ML domain

1. Impressive talks https://indico.cern.ch/event/673350/
   https://indico.cern.ch/event/687788/
2. A lot of solutions available from https://root.cern.ch/tmva
3. Examples of applications (papers I read):
   (a) K. Fraser and M. D. Schwartz, arXiv:1803.08066
   (b) P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, "Jet Substructure Classification in High-Energy Physics with Deep Neural Networks," Phys. Rev. D **93** (2016) no.9, 094034
   (c) E. Bothmann and L. Del Debbio, "Reweighting a parton shower using a neural network: the final-state case," arXiv:1808.07802
   (d) D. Guest, K. Cranmer and D. Whiteson, "Deep Learning and its Application to LHC Physics," arXiv:1806.11484
   (e) . Baldi, P. Sadowski and D. Whiteson, Nature Commun. **5**, 4308 (2014)
4. I am only a user interested in how to prepare input for ML solution and in results.
5. I am worried about systematic errors of multi-dimensional distributions.
6. Big interest among students for projects of contact with industrial/other domain techniques is present. They should find the ways. **If** they want to use such tools they **need to understand mathematical foundations**.

News on the programs: PHOTOS, Tauspinner

1. `Photos` http://photospp.web.cern.ch/photospp/ Monte Carlo for bremsstrahlung in decays was enriched with emission of pairs, tests are published:

   S. Antropov, A. Arbuzov, R. Sadykov and Z. Was, "Extra lepton pair emission corrections to Drell-Yan processes in PHOTOS and SANC," Acta Phys. Polon. B **48** (2017) 1469

2. `TauSpinner` http://tauolapp.web.cern.ch/tauolapp/, algorithm for re-weighting τ production and decays was enriched. This program was used to obtain results presented in earlier parts of my talk. References:

a T. Przedzinski, E. Richter-Was and Z. Was, "Documentation of TauSpinner algorithms – program for simulating spin effects in tau-lepton production at LHC," arXiv:1802.05459.

b E. Richter-Was and Z. Was, "The TauSpinner approach for electroweak corrections in LHC Z to ll observable," arXiv:1808.08616

c E. Barberio, B. Le, E. Richter-Was, Z. Was, D. Zanzi and J. Zaremba, "Deep learning approach to the Higgs boson CP measurement in $H \to \tau\tau$ decay and associated systematic," Phys. Rev. D **96** (2017) no.7, 073002

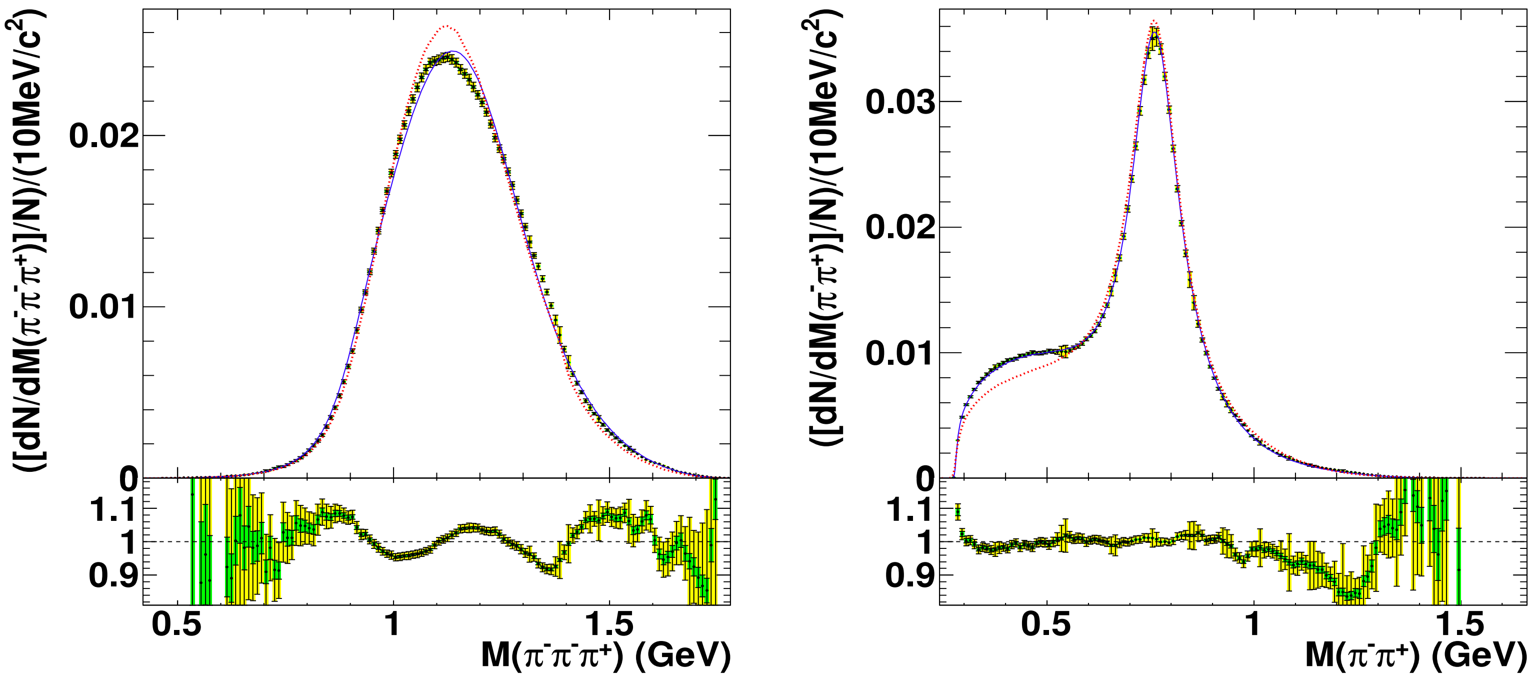
1. TAUOLA with new hadronic currents, up to 500 decay channels, which can be manipulated by user is published: M. Chrzaszcz, T. Przedzinski, Z. Was and J. Zaremba, Comput. Phys. Commun. **232**, 220 (2018) . For archivization purposes initialization of hadronic decay channels is compatible with defaults as BaBar was using.
2. Direction for work is essentially set. I have not received much feed-backs, but there were no objections too.
3. Program is prepared to be translated piece after piece into C++, or other language. Whenever a need will arrive.
4. **Theoretical uncertainty of the models can be $\frac{1}{N_C}$, $\frac{1}{N_C^2}$ or $\cdots$,** but experimental precision has to be assumed to be better than 0.001. That is a factor of 100 better.
5. Not much to add ...

- The precision of the longitudinal and transverse spin measurement in H→ $\tau\tau$ and Z→ $\tau\tau$ can be gained considering all possible τ decays
- The full kinematic analysis might help to maximize the analysing power in measurements of τ spin effects in H→ $\tau\tau$ and Z→ $\tau\tau$
- Several tools exist that can be used for estimation the invisible momentum in τ decays (using missing transverse energy, track impact parameters, τ decay vertex etc)
- In practice a reasonably good performance is expected only in channels where robust reconstruction of τ decay point is possible: Z/H→ $\tau\tau$ $\rightarrow a_1 + X$, $a_1 + a_1$ with a_1 decaying to three charged pions.
- The final choice of discriminant in each decay category can be concluded after all detector effects are studied and understood.

CMS work in progress…

New currents for $\tau \to 3\pi$ and $\tau \to 2\pi$ decays

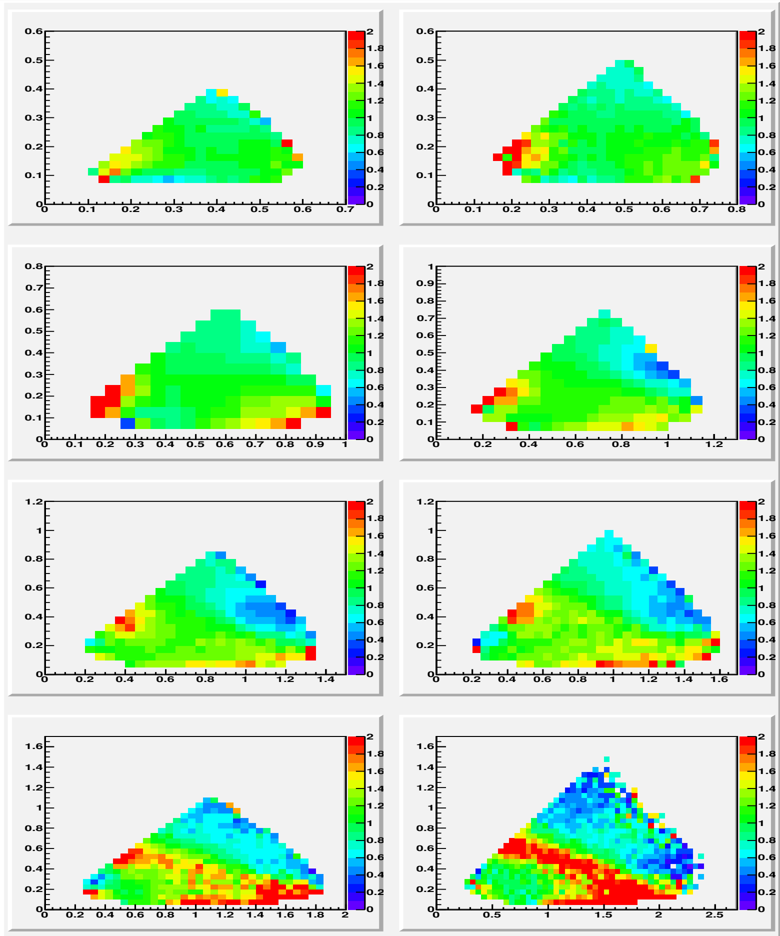
Currents from Resonance Chiral Lagrangian approach, fits to BaBar data. Experimental systematic errors considered. From: *Resonance Chiral Lagrangian Currents and Experimental Data for* $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$, I.M. Nugent, T. Przedzinski, P. Roig, O. Shekhovtsova, Z. Was, Phys. Rev. D 88, 093012 (2013). **Looks like a step of successful strategy. See the next slide for concern.**



From Z.W, Jakub Zaremba, Eur.Phys.J. C75 (2015) 566

Dalitz plot ratios (Red, blue differences ∼50% or more); **TAUOLA RChL** to **TAUOLA CLEO**,

$m^2(3\pi)$ in ranges: 0.36-0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69-1.96, 1.96-2.25, 2.25-3.24 GeV^2.



Conclusions and outlook

- Some examples of how new techniques for data analysis became useful for multi-dimensional distributions involving $H \to \tau\tau$ decays and Higgs CP-parity measurement.
- Useful for that properties of Monte Carlo simulation programs and properties of ML techniques were underlined.
- Recent developments for `Tauola` `TauSpiunner` and `Photos` programs were listed.
- Important properties of predictions for models used to describe $\tau \to 3\pi\nu$ decays were underlined.
- Fits involving multi-dimensional distributions are highly desirable.
- Essential for future developments will be thus control of systematic errors for such multi-dimensional distributions.
- If experimental data should have background subtracted, or if dominant backgrounds will be fitted simultaneously with the signature is technically less important.
- Question of manpower and training as well as motivation of involved people is very important. Competition for talent from other fields can not be ignored.

Reconstruction for precision observable at LEP

- 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.