New Measurements with Archived e+e- Datasets from LEP Collider

Anthony Badea (UChicago)

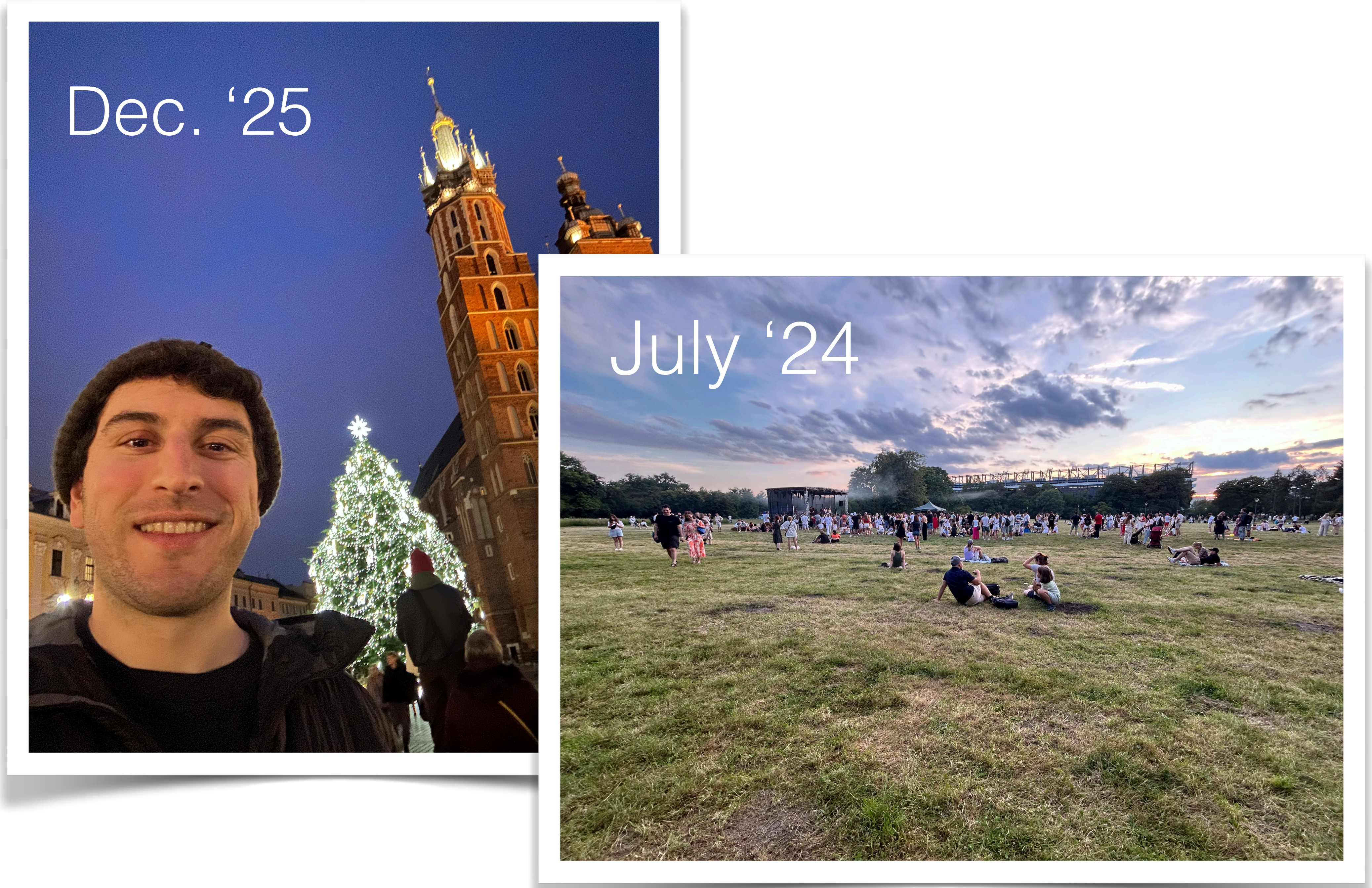
Seminar for AGH Krakow HEP Białasówka Seminar
December 12, 2025

Collaboration: e^+e^- Alliance, Austin Baty (UIC), Luke Lu, Jingyu Zhang, Luna Chen (Vanderbilt) Marcello Maggi (Bari), Ben Nachman (Stanford), Hannah Bossi, Yu-Chen Chen, Gian Michele Innocenti, Michael Peters, Tzu-An Sheng, Chris McGinn, Jesse Thaler, Yen-Jie Lee (MIT) AB work supported by Schmidt Sciences

Many Thanks to Iwona, Patrycja, All

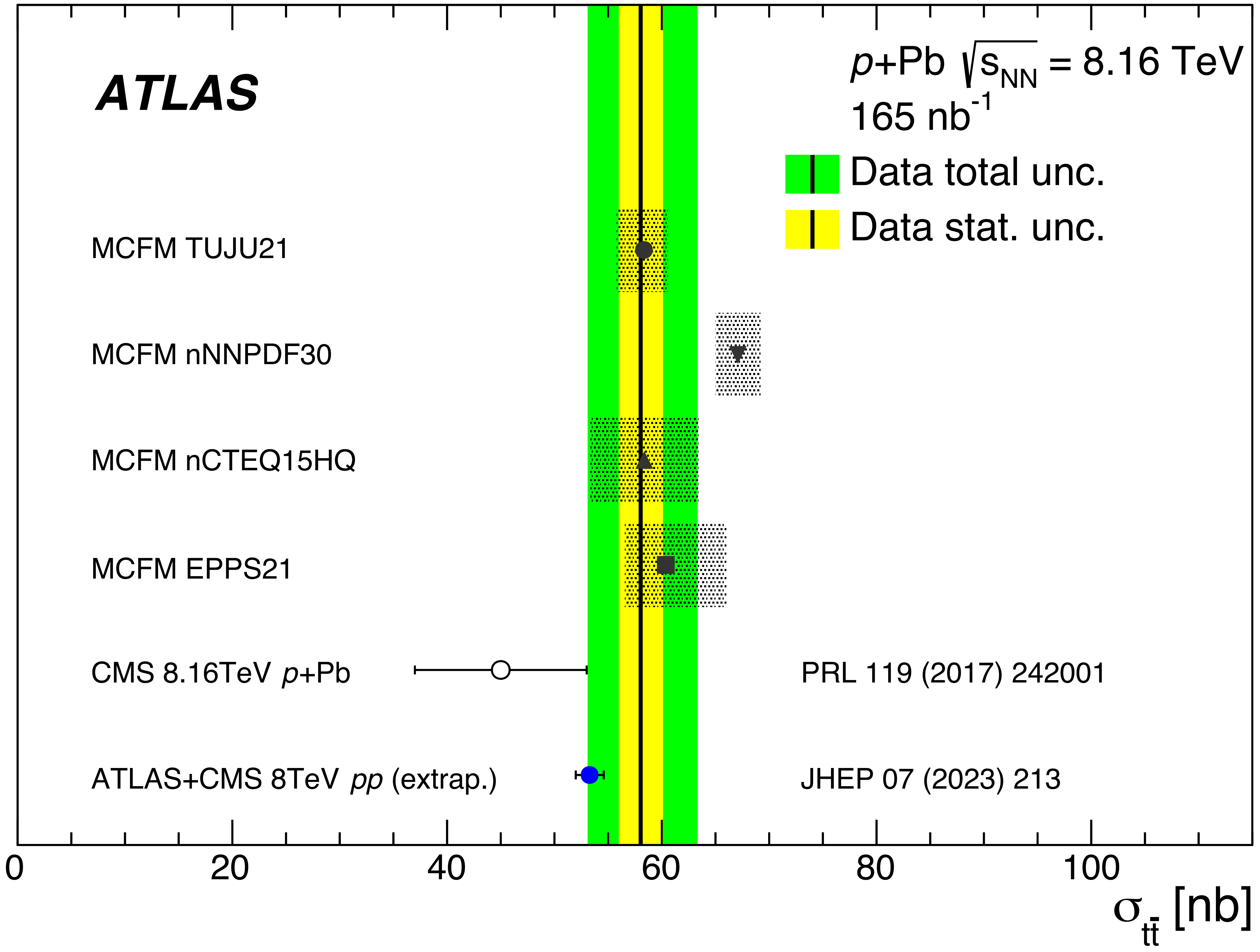
Congratulations Dr. Potepa!!

Great time visiting Krakow!!

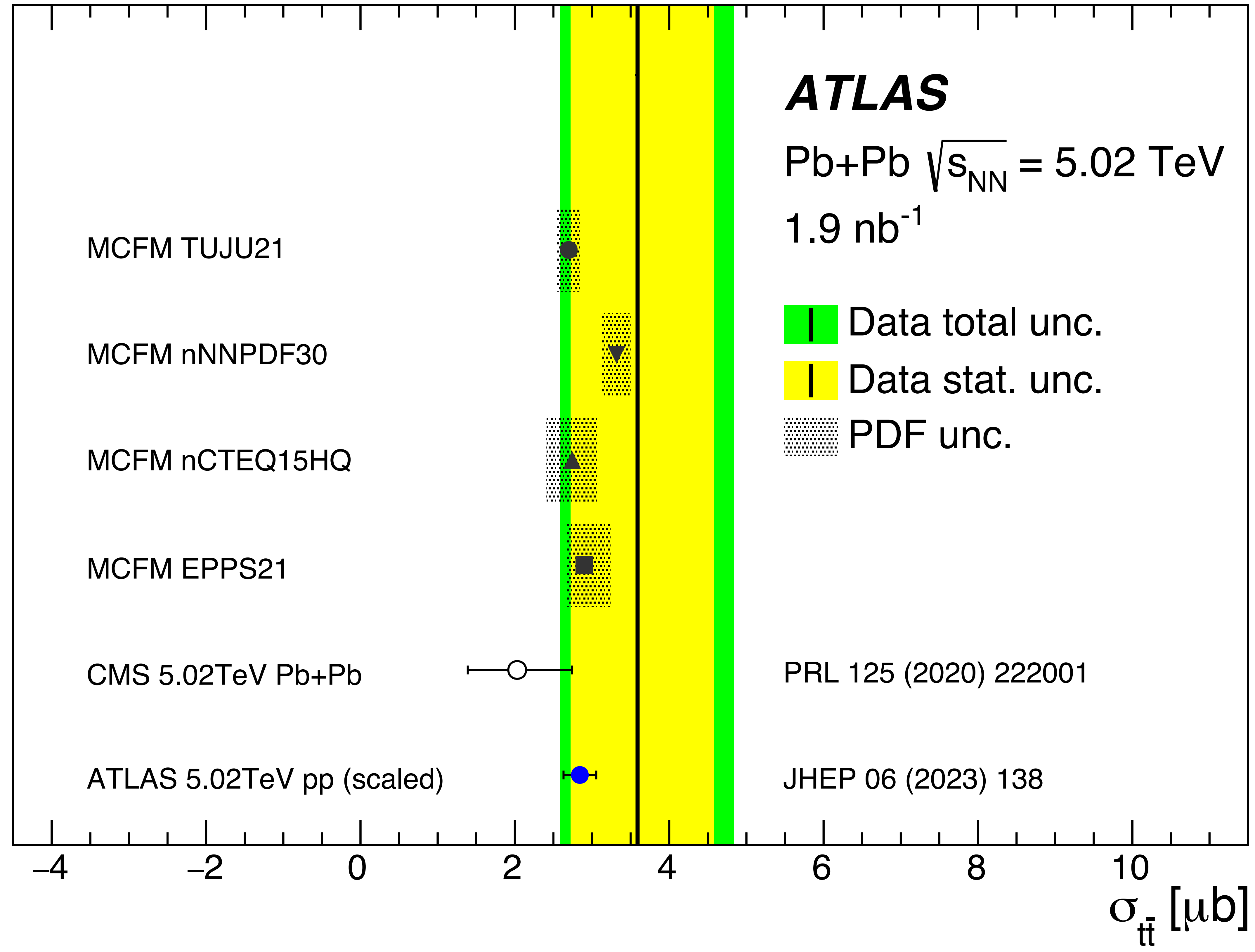


One Connection From Dr. Potepa's Talk

JHEP 11 (2024) 101

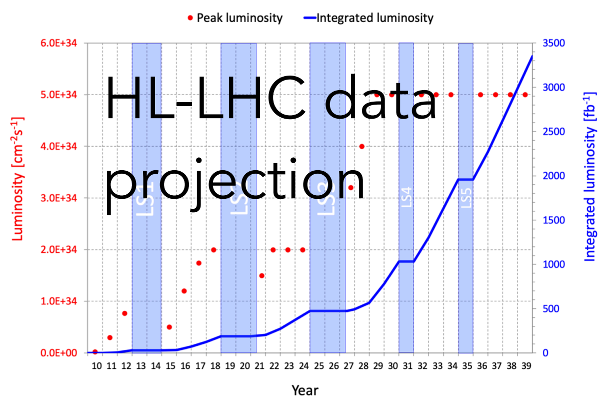


PRL 134 (2025) 142301



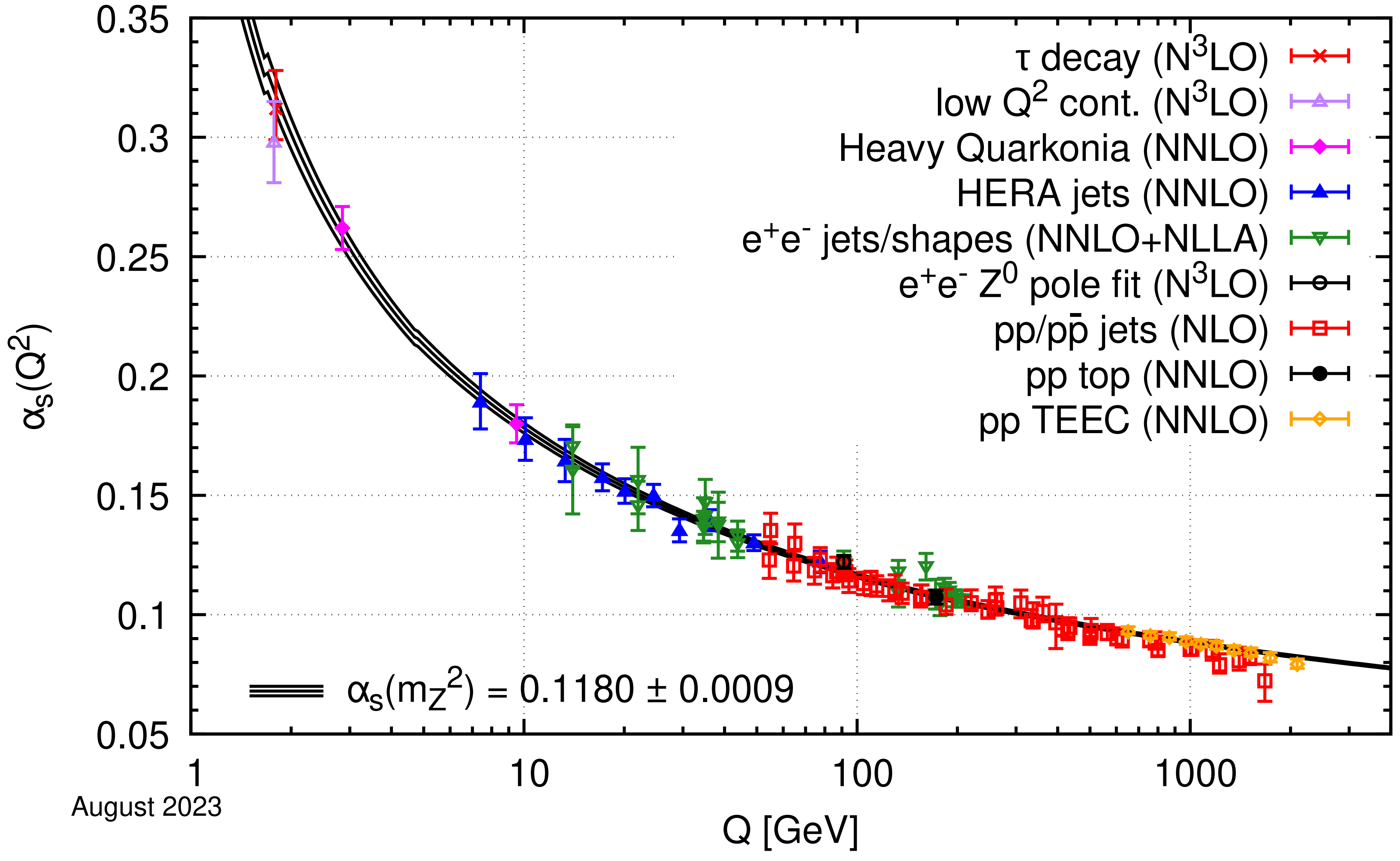
With high stats, systematics dominated driven by hadronic handle

Feature across frontiers of research in high energy, nuclear, astrophysics

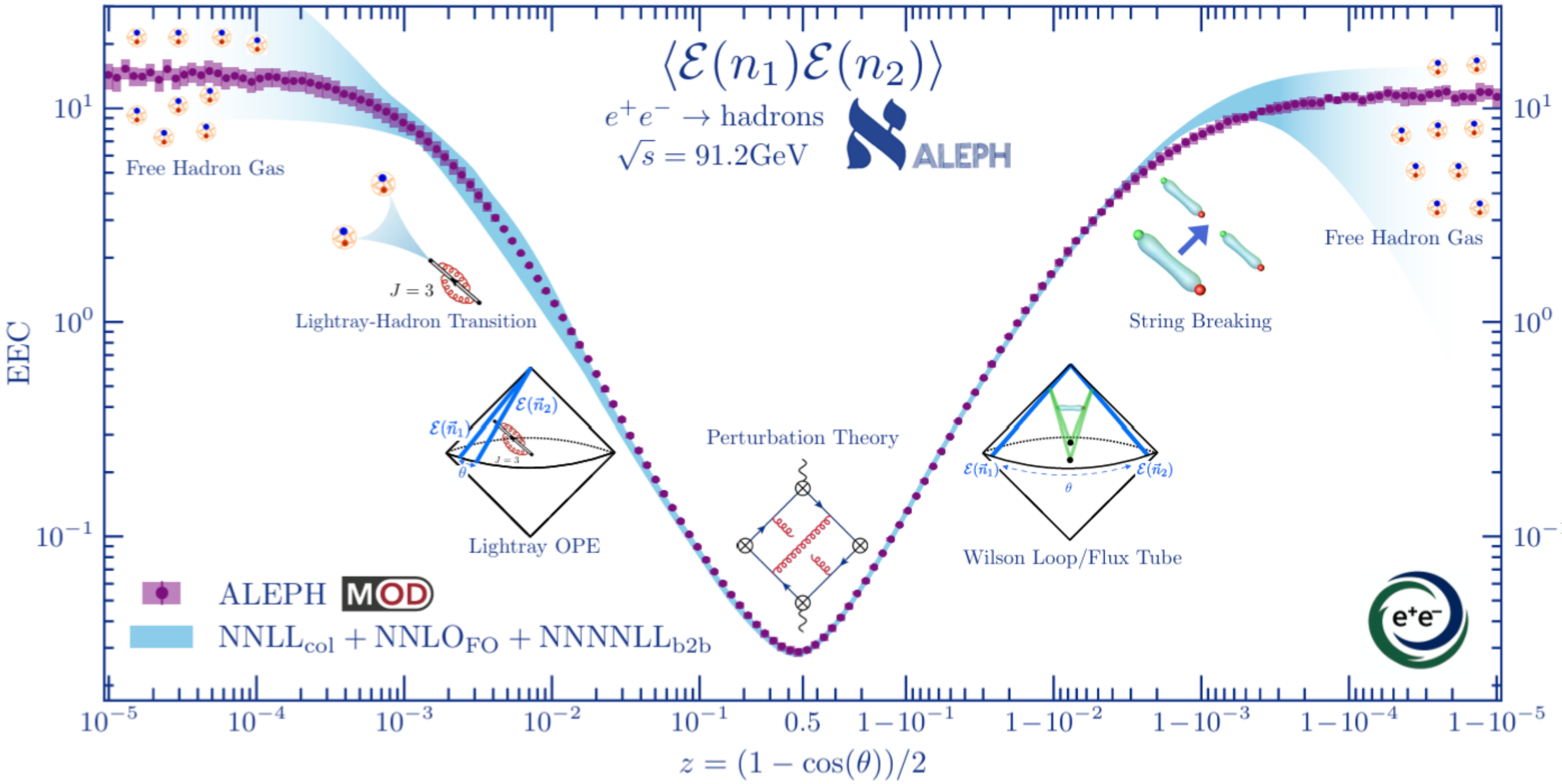


Quantum Chromodynamics

QCD Interesting Across Scales

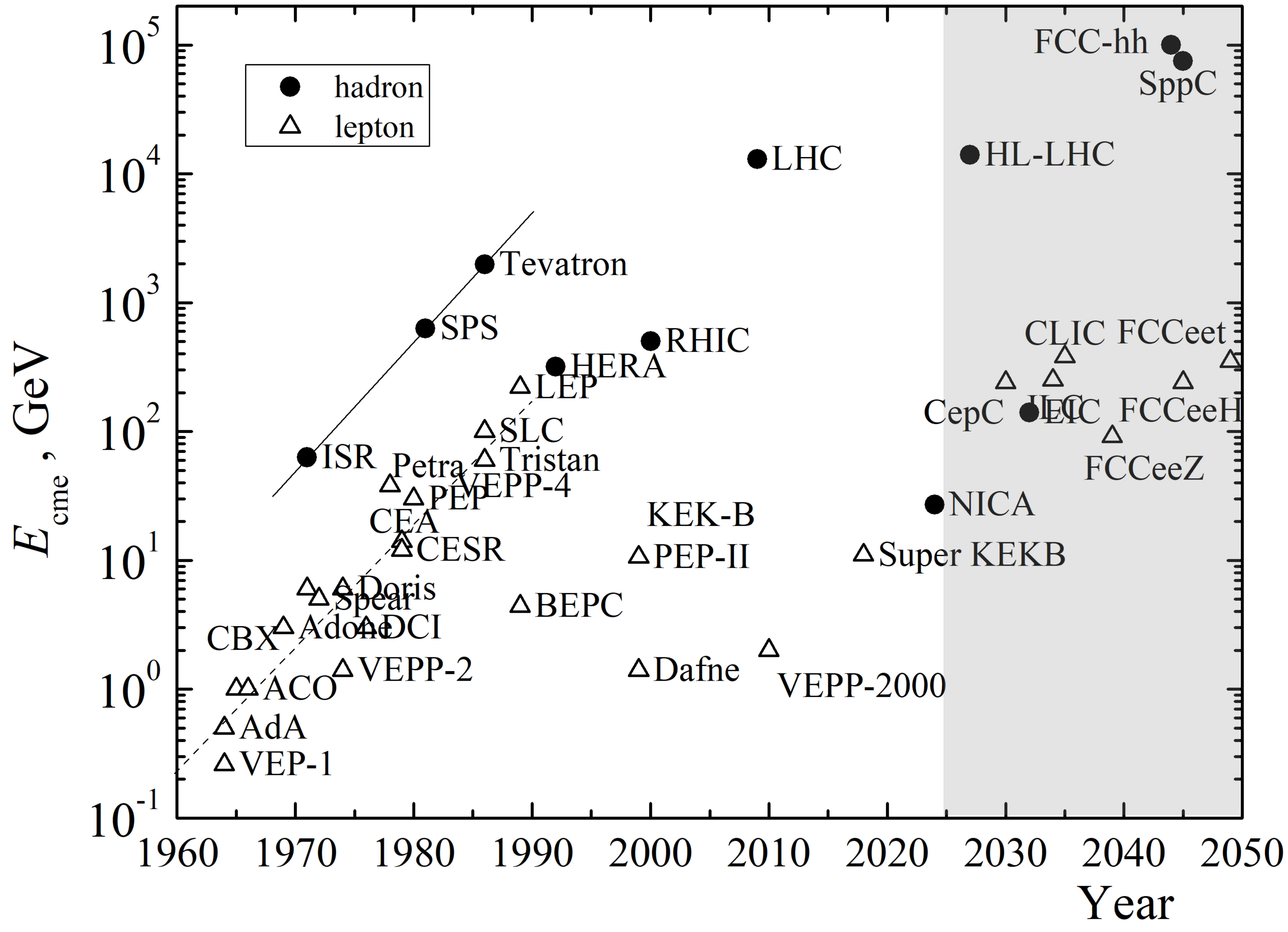


Perturbative (P) and Non-Perturbative (NP) QCD
New Physics in QCD Systems



- Quantum Chromodynamics (QCD), theory of strong force, central to Standard Model (SM)
- Study of QCD across wide energy scales, spanning P/NP regimes, of global importance
- Precision QCD enables discovery of new phenomena in strongly interacting systems

Landscape of Collider Data



Global operational high energy pp, pA, AA and low energy e^+e^- colliders

No current high energy e+e- machine but incredible foresight by previous experiments to archive those data

Window for progress to influence the crucial 2030's data taking

Electron-Positron Alliance's Reanalysis Efforts

Build community for curation, standardization, and reanalysis of archived/open e+e- data

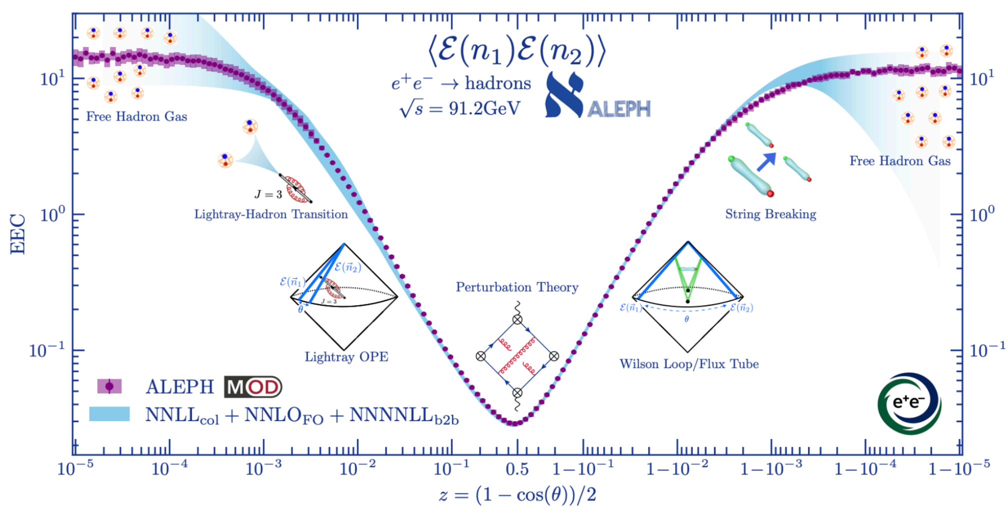
Enable **direct collaboration** between **experimentalists, analyzers and theorists**

Weekly group analysis meeting since 2017
10 LEP notes and publications on arXiv
2 publications with the Belle Collaboration
> 60 presentations in conferences / workshops

The Electron-Positron Alliance

About Papers Talks Links

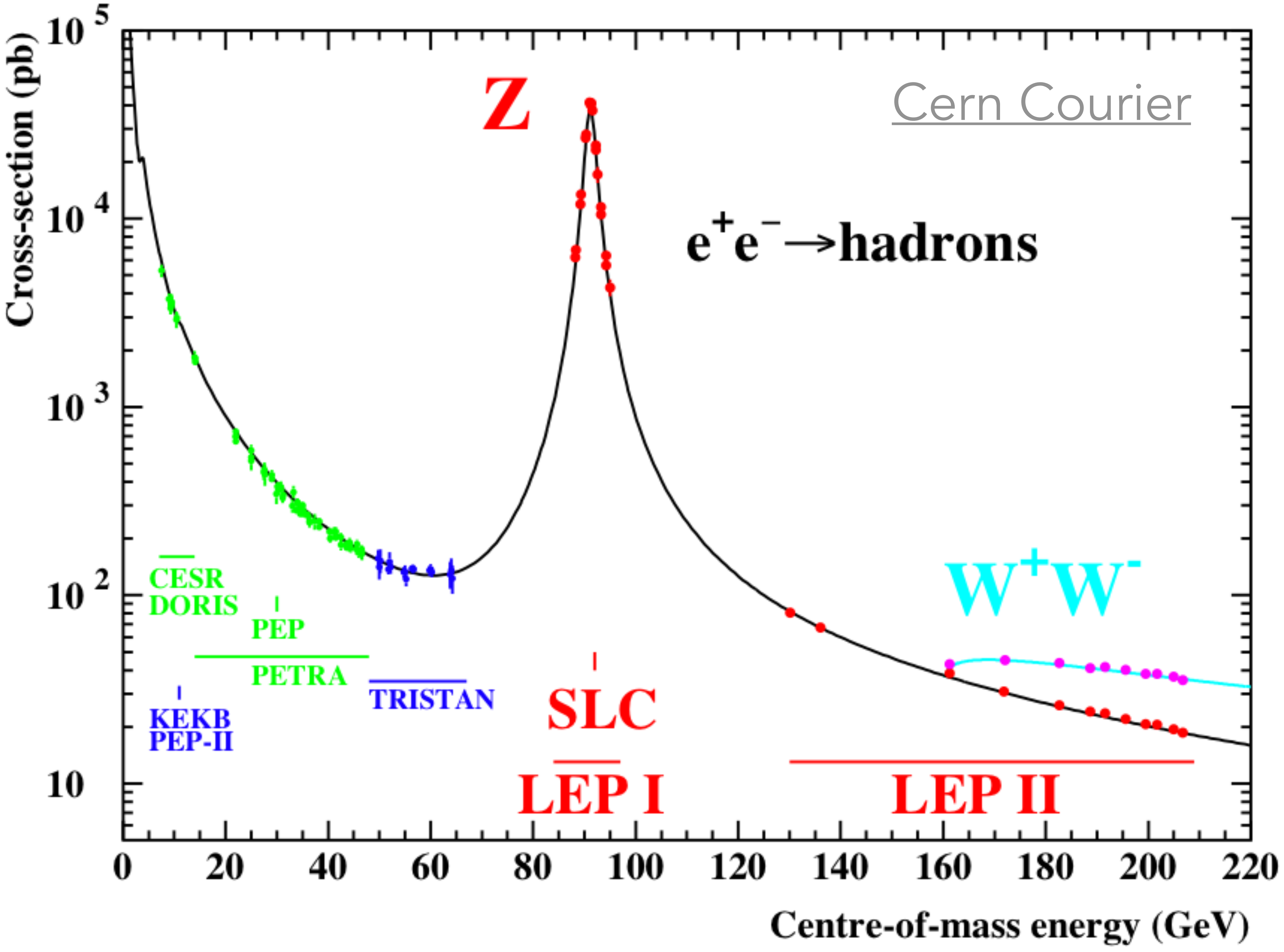
Maximize scientific impact of collected data

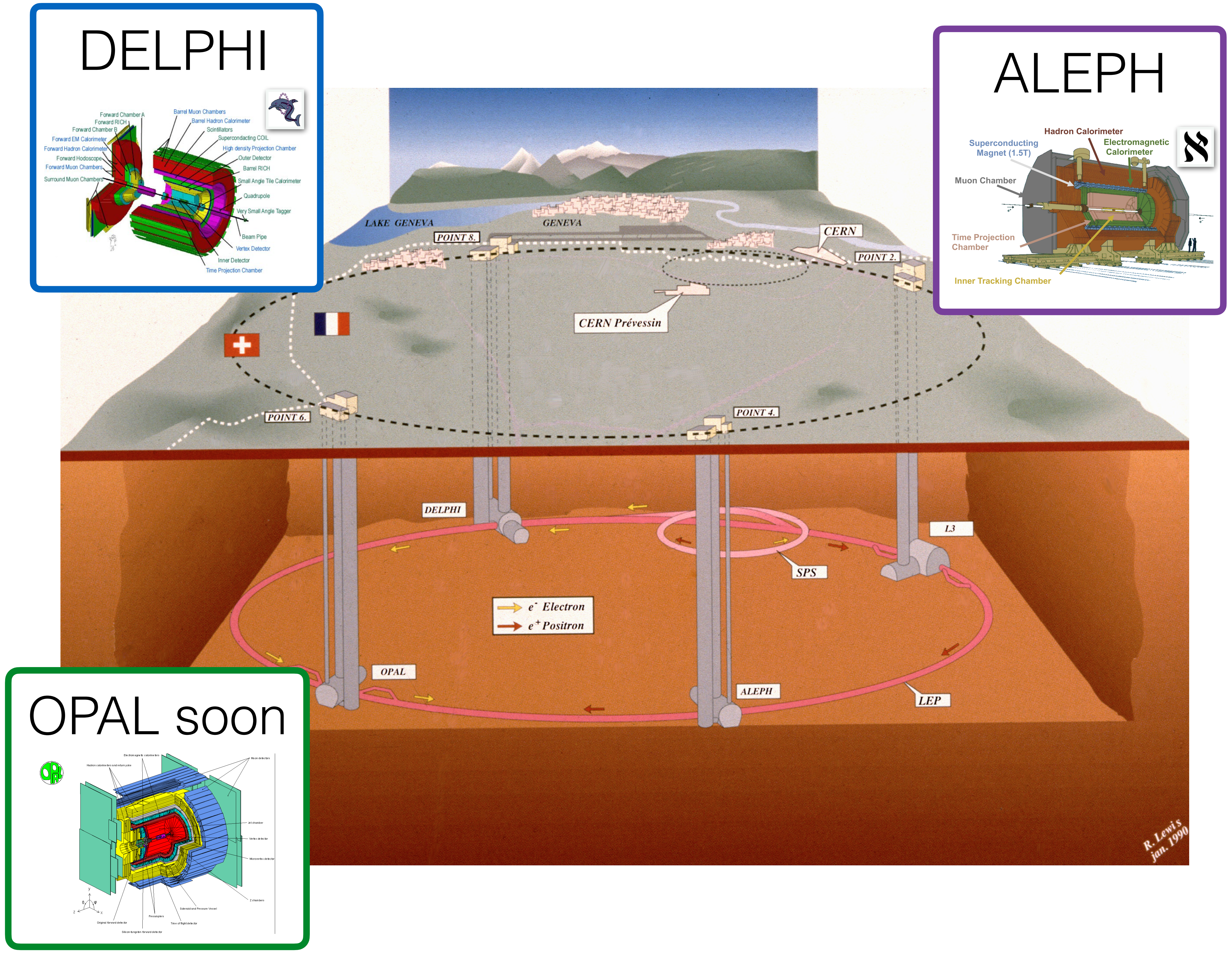


https://ee-alliance.org

Drawing from Luna Chen's and Yen-Jie Lee's talks

Large Electron-Positron (LEP) Datasets



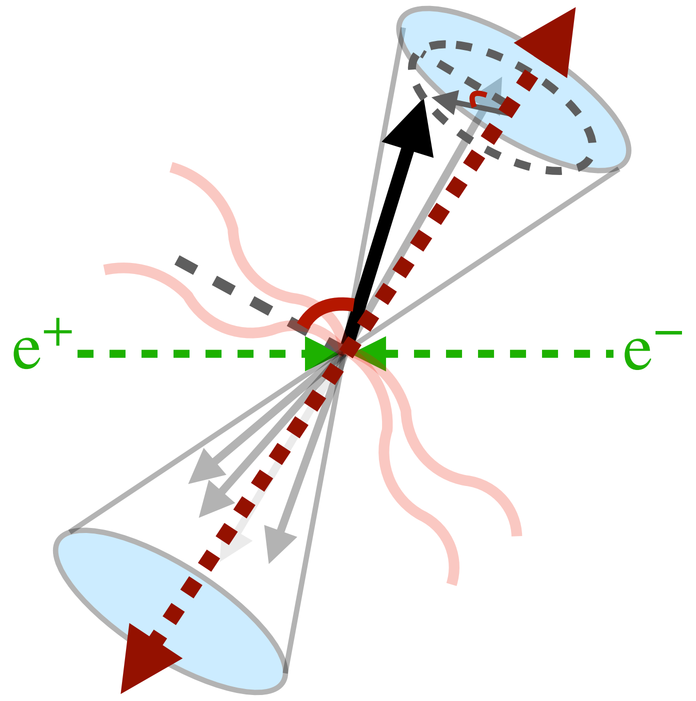


Active reanalysis of LEP1/2 data. Will discuss thrust and ongoing works

Outline

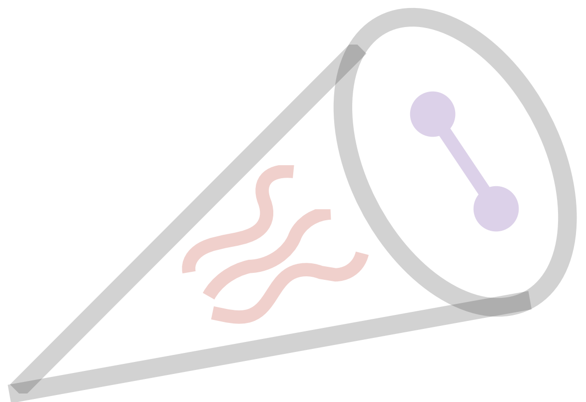
Physics Motivation (I)

Discovering new SM/BSM physics requires an unprecedented handle of QCD, motivating measurements across collision systems, regimes, and observables.



LEP Thrust Reanalyses (II)

Reanalysis of thrust in LEP archived data reveals new insights into α_S and P/NP QCD. Enables new studies of e+e- collisions with modern exp. and theory tools.



Ongoing Works and Outlook (III)

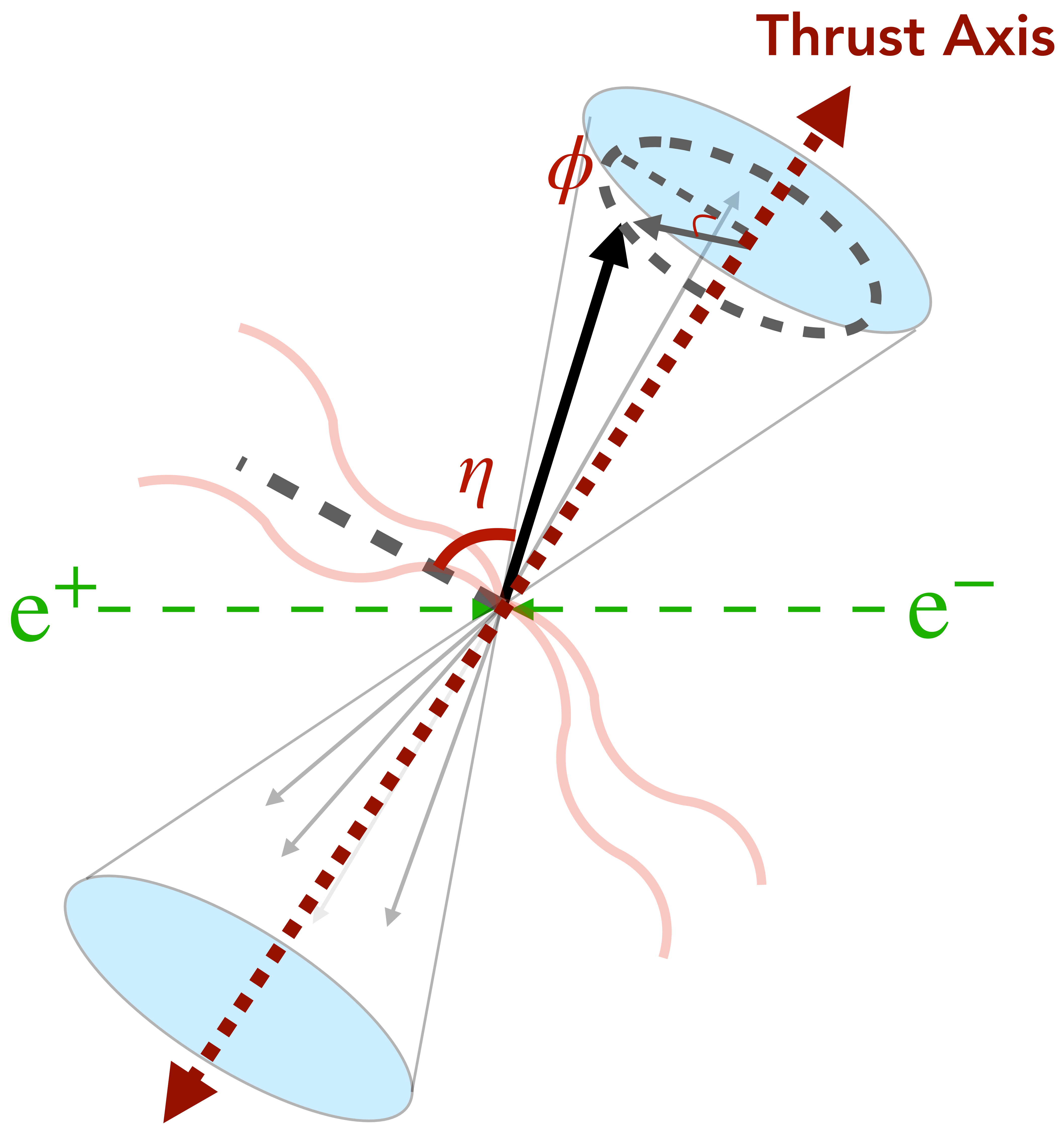
Continued investigations into thrust across relevant regimes. Complementary observables, studies across energies, application to constrain modern QCD.

Physics Motivation

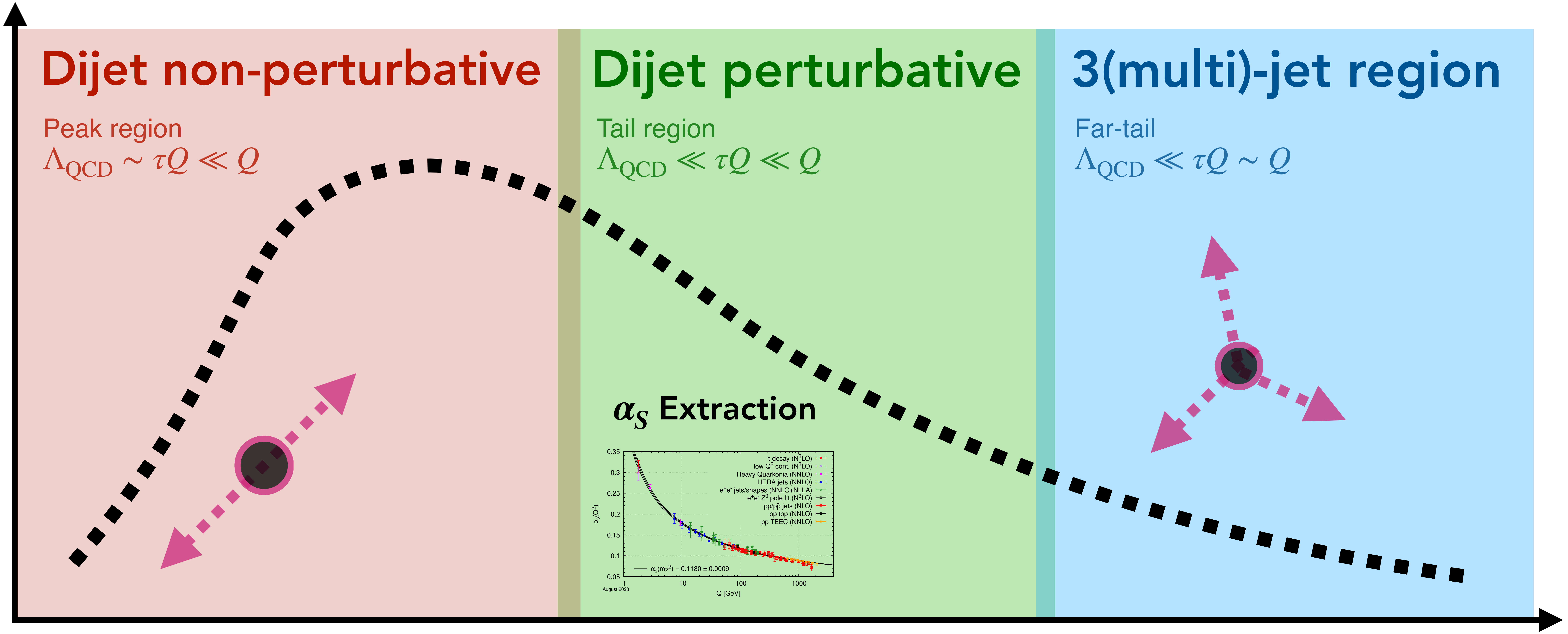
Thrust is an event shape observable that quantifies how dijet-like an event is, active use to extract α_s and constrain P/NP QCD

E. Farhi (1977): Phys. Rev. Lett. 39, 1587

$$T = \max_{\hat{n}} \frac{\sum_i |\vec{p}_i \cdot \hat{n}|}{\sum_i |\vec{p}_i|}$$



Interesting Thrust Regimes to Probe

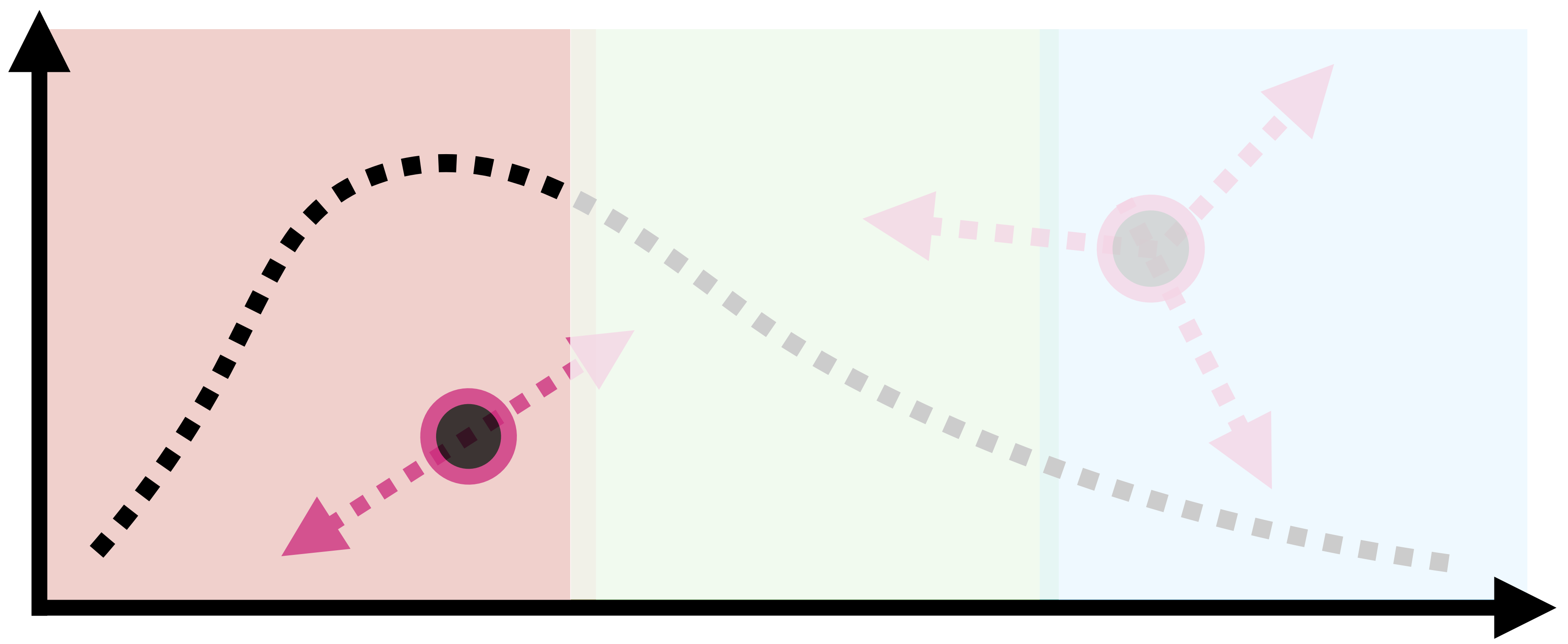


τ = 1-T

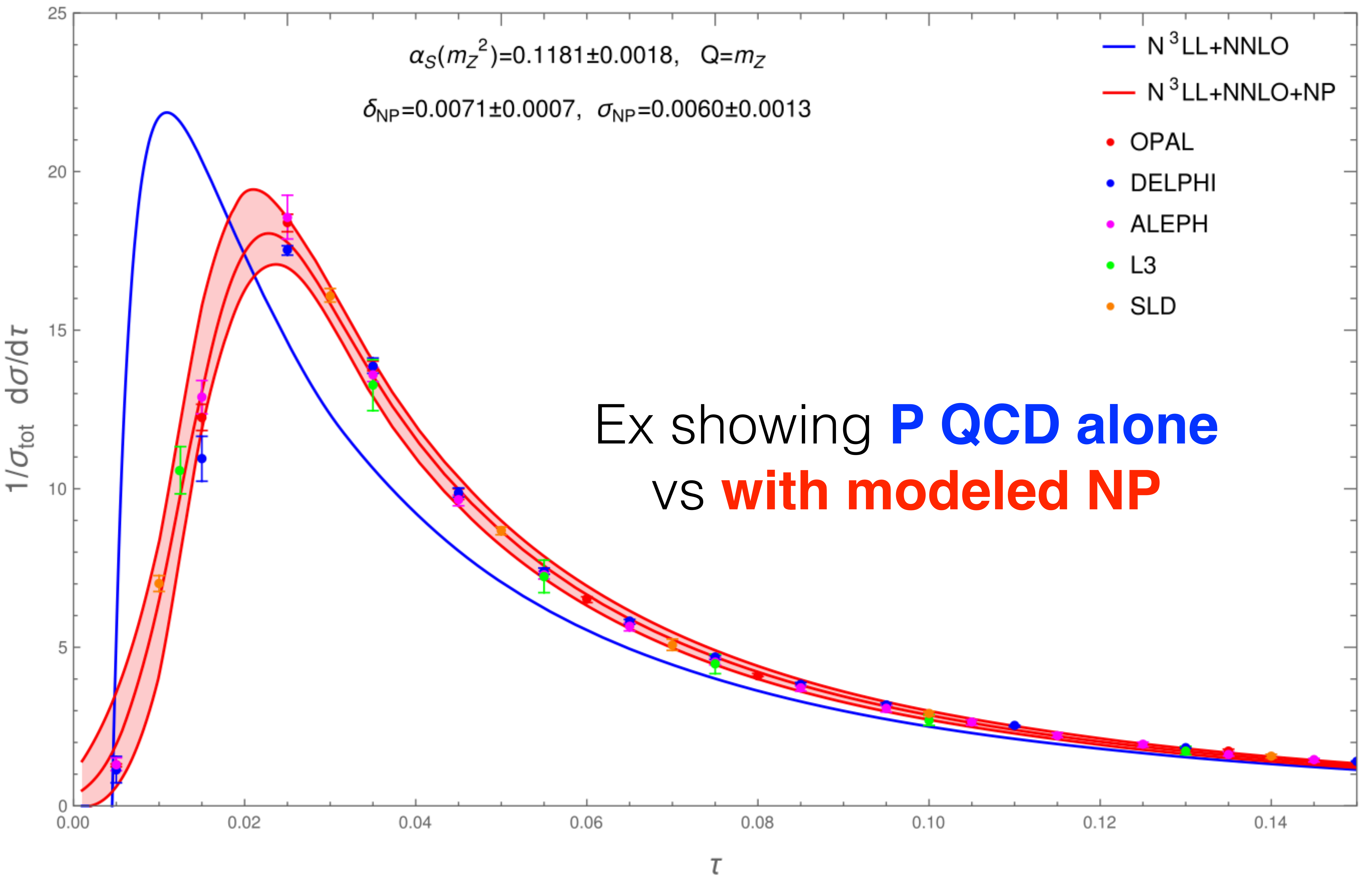
~3 interesting regimes

Modern Interest in Thrust Regime (I)

Dijet non-perturbative



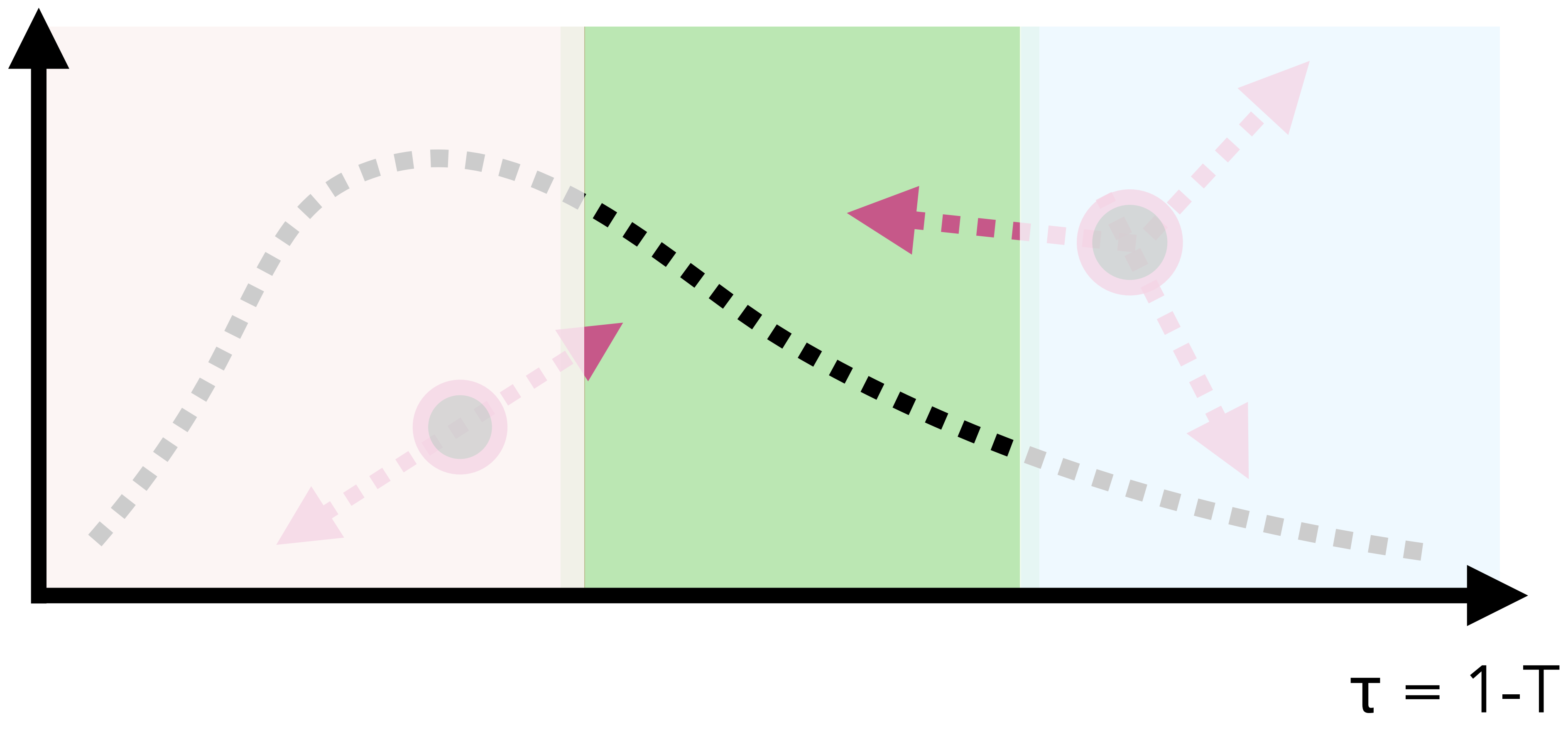
Large NP corrections applied using functional forms, large impact with limited data points



Phys. Rev. Lett. 134, 251904, 0809.3326

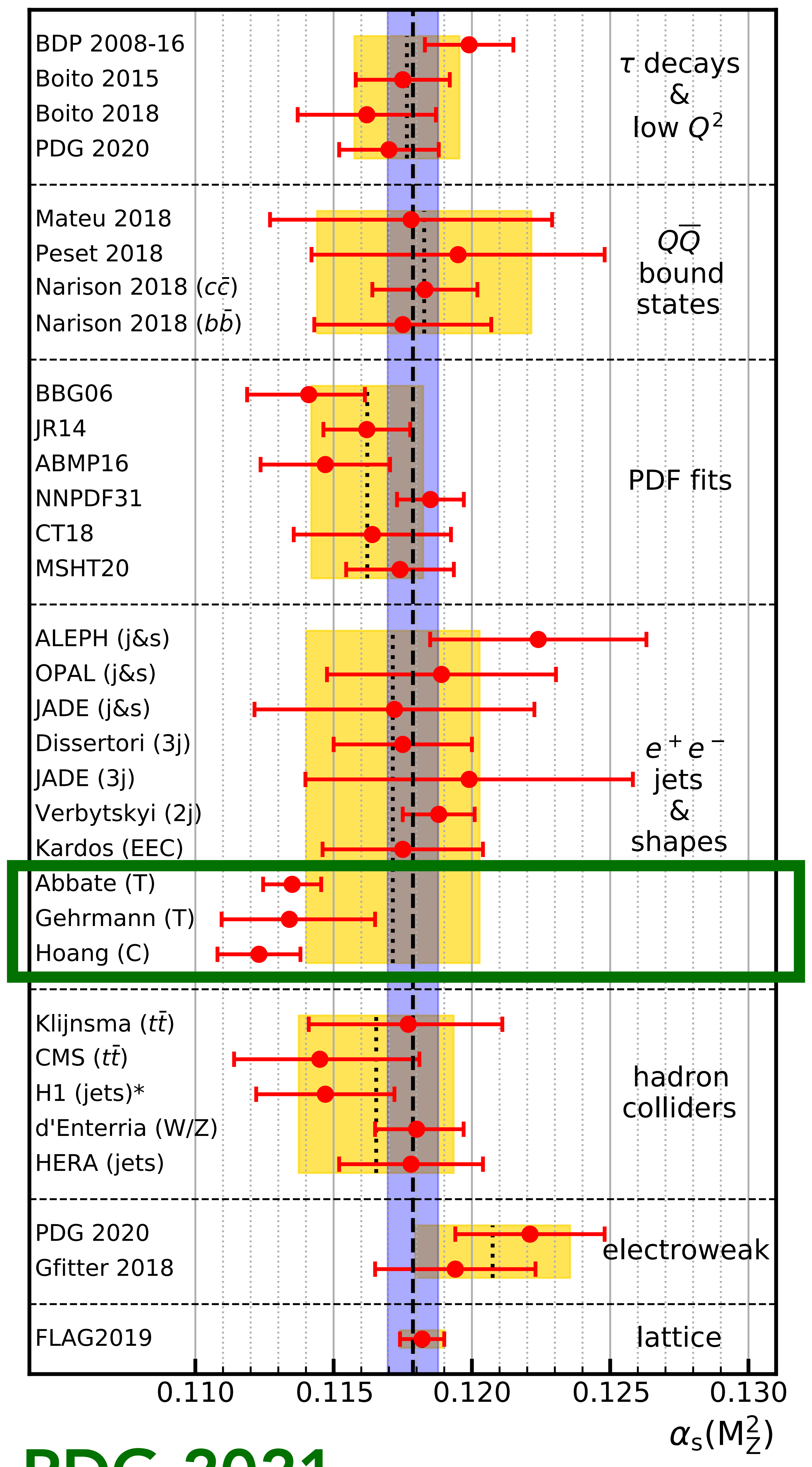
Modern Interest in Thrust Regime (II)

Dijet perturbative

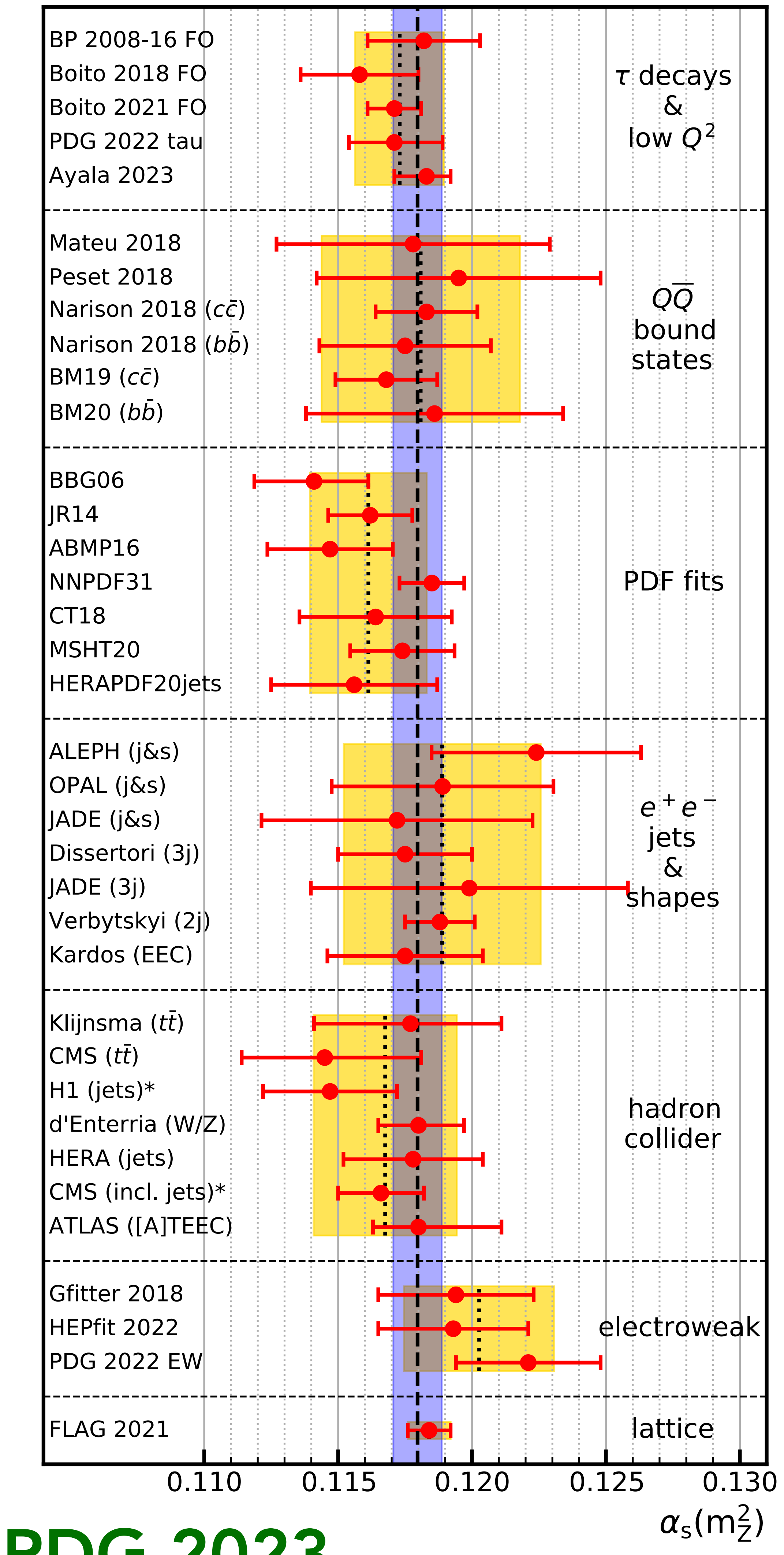


Large differences in $\alpha_S(m_Z)$ between e+e- event shapes (e.g. Thrust) extractions as well as with extractions from other observables (PDG world):

JHEP 07 (2025) 249 (τ resum.):	0.1136 ± 0.0012
PRL 134 (2025) 251904 (L resum.):	0.1181 ± 0.0018
PDG 2023 World Average:	0.1180 ± 0.0009



PDG 2021

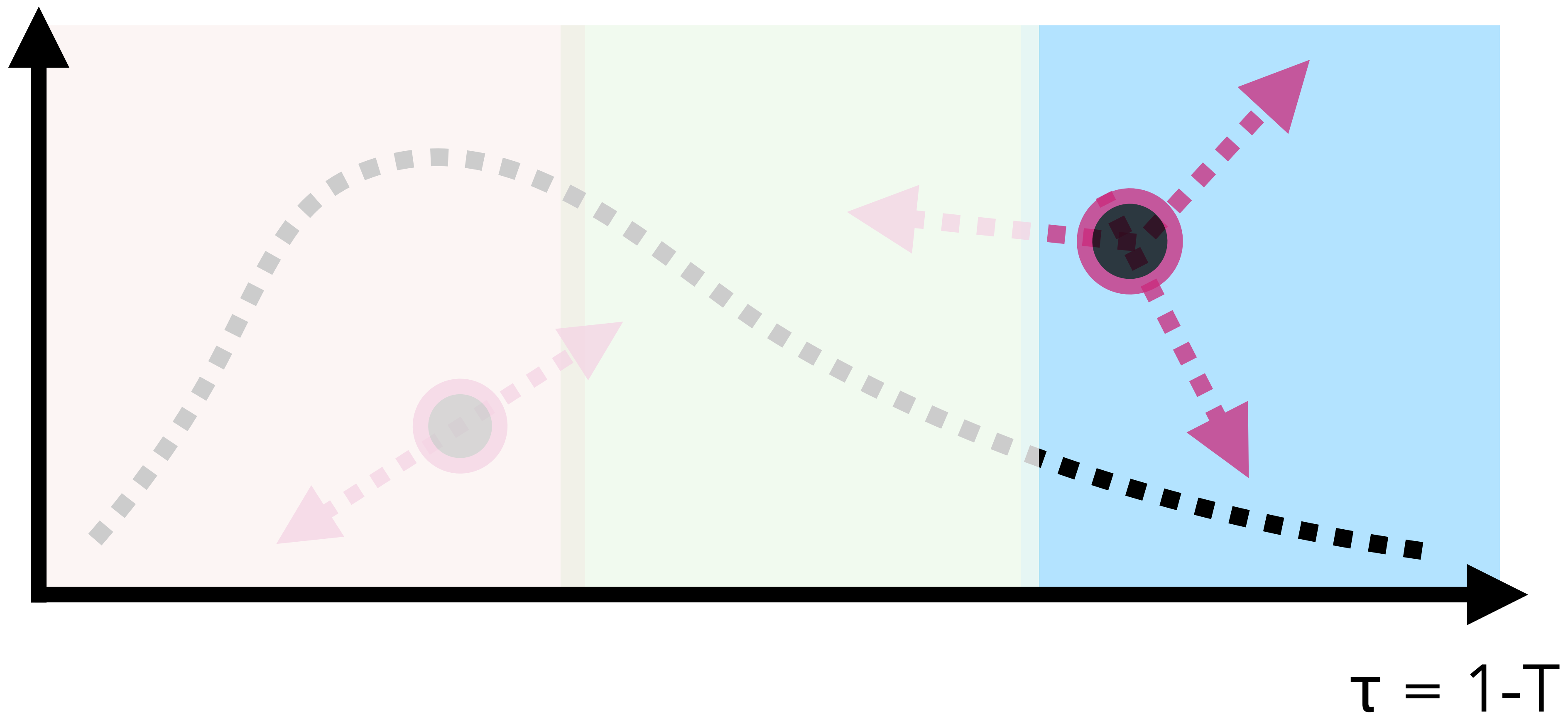


PDG 2023

see e.g. CERN QCD Seminar 02/10/2025 G. Vita, 1006.3080, 2412.15164, Slide 8 of P. Skands

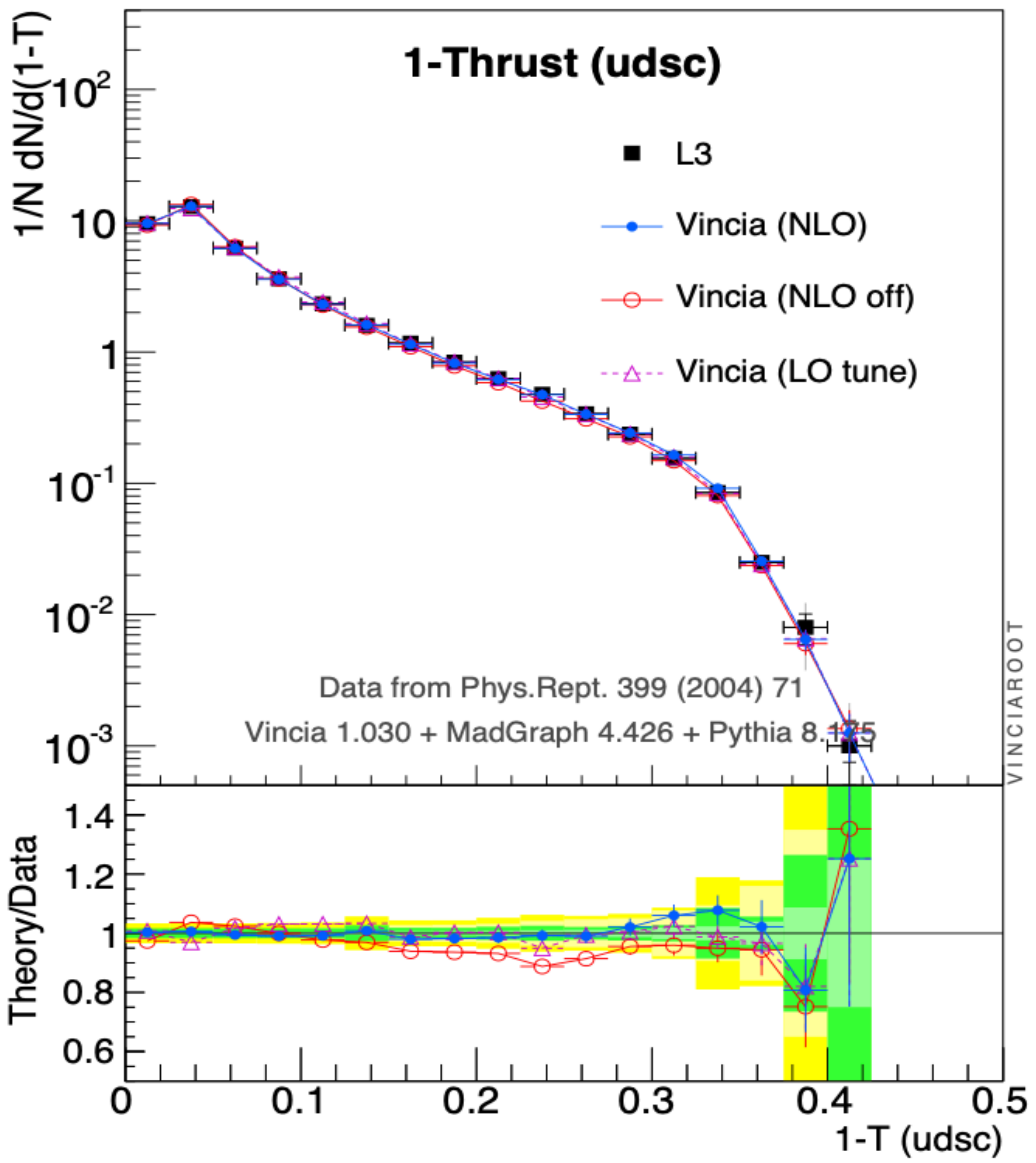
Modern Interest in Thrust Regime (III)

3(multi)-jet region

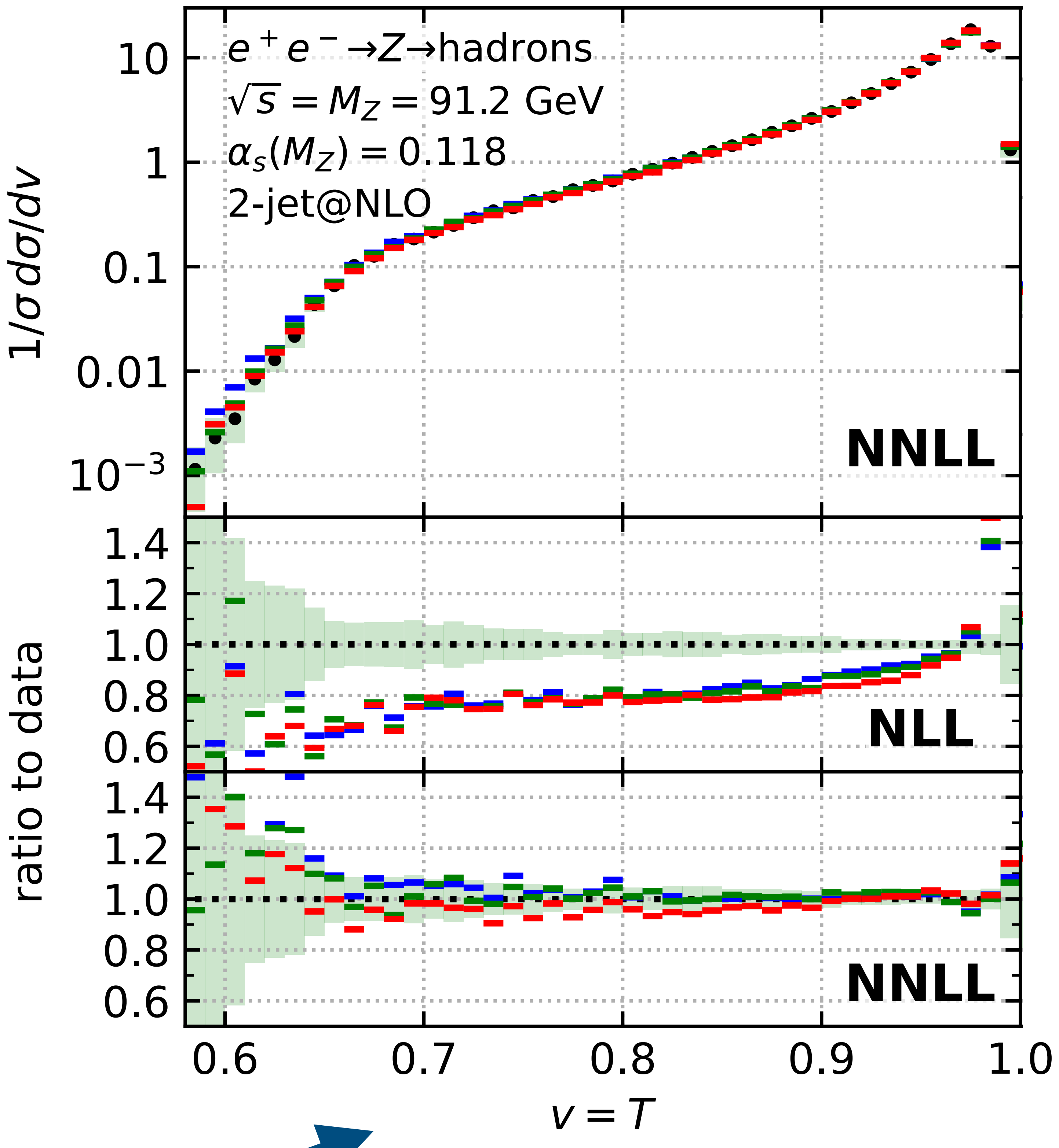


Validate/tune modern P QCD predictions and parton shower Monte Carlo (MC) simulations

Pythia Tuning



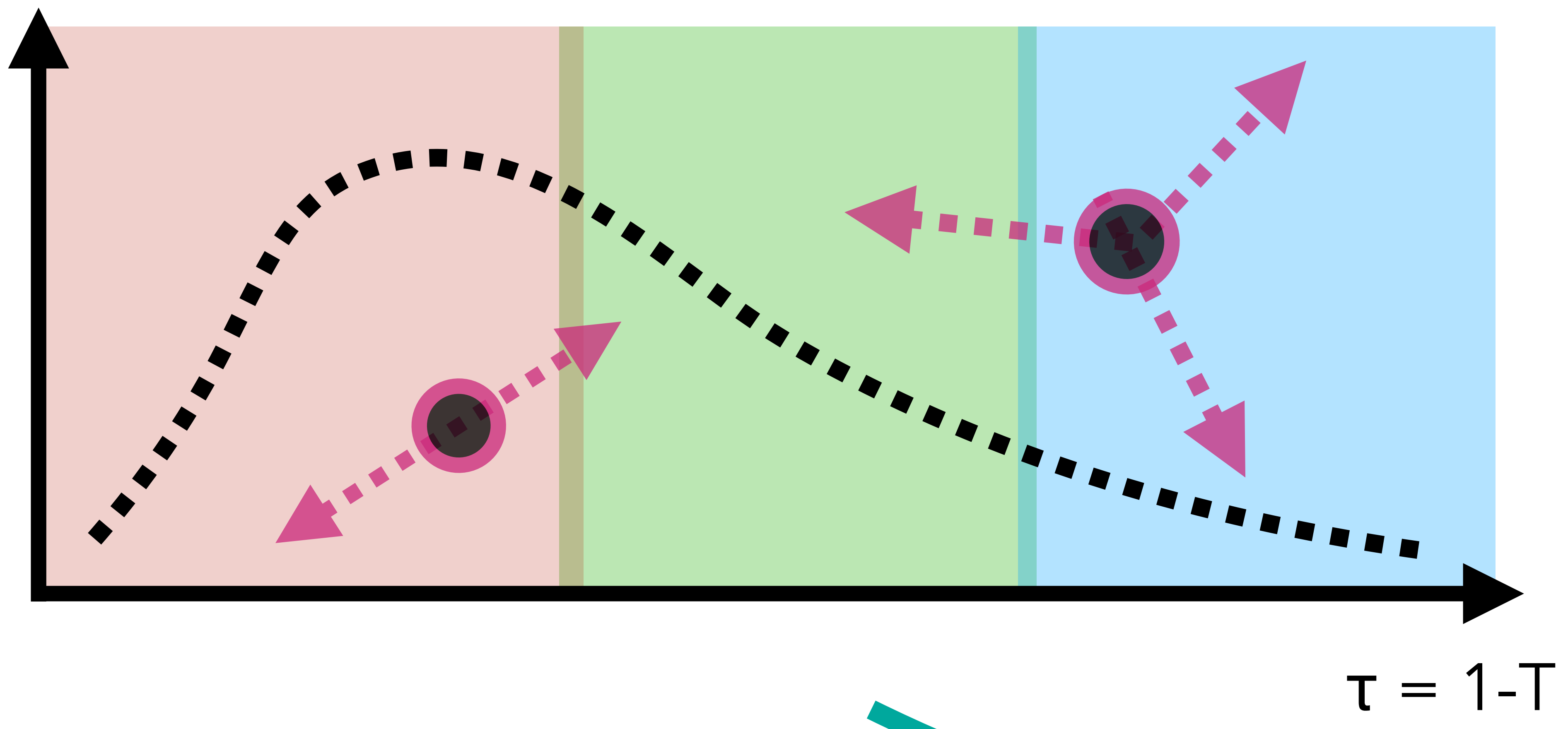
PanScales NNLL Showers

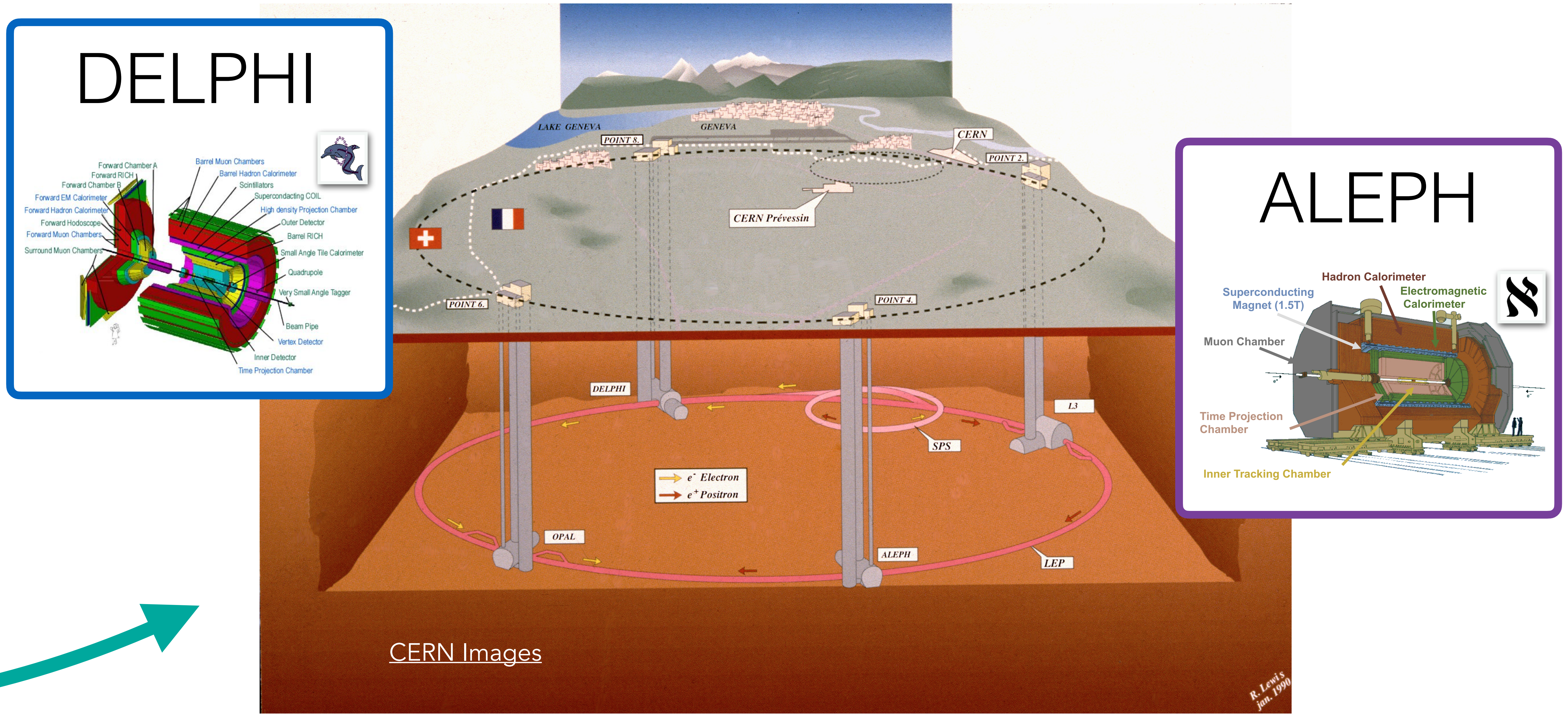


Event Generator Tuning (P. Skands), 1303.4974
PRL 134 (2025) 1, 011901

Physics Questions to be Addressed

Probe across P/NP Regimes

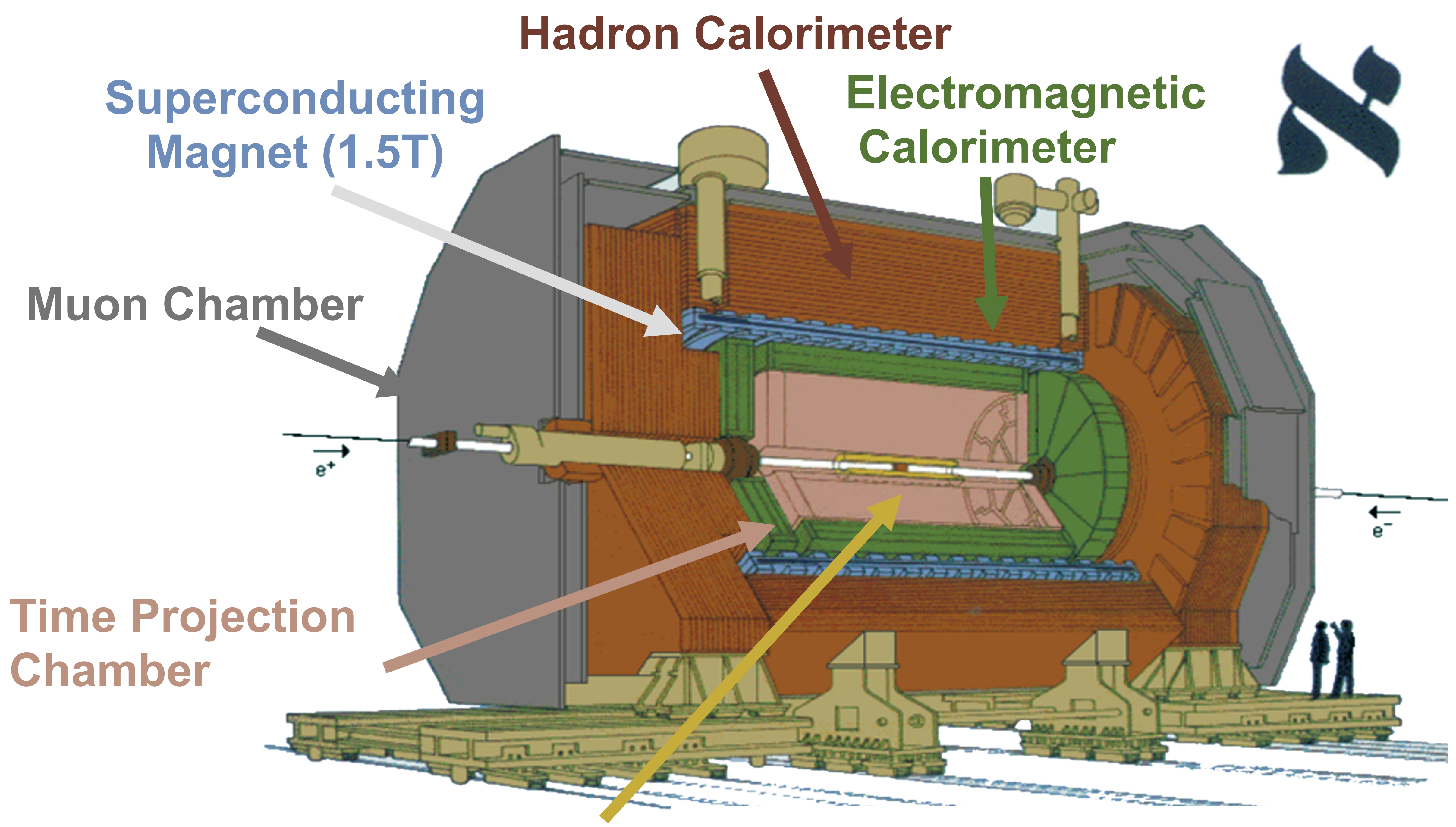




LEP reanalyses!

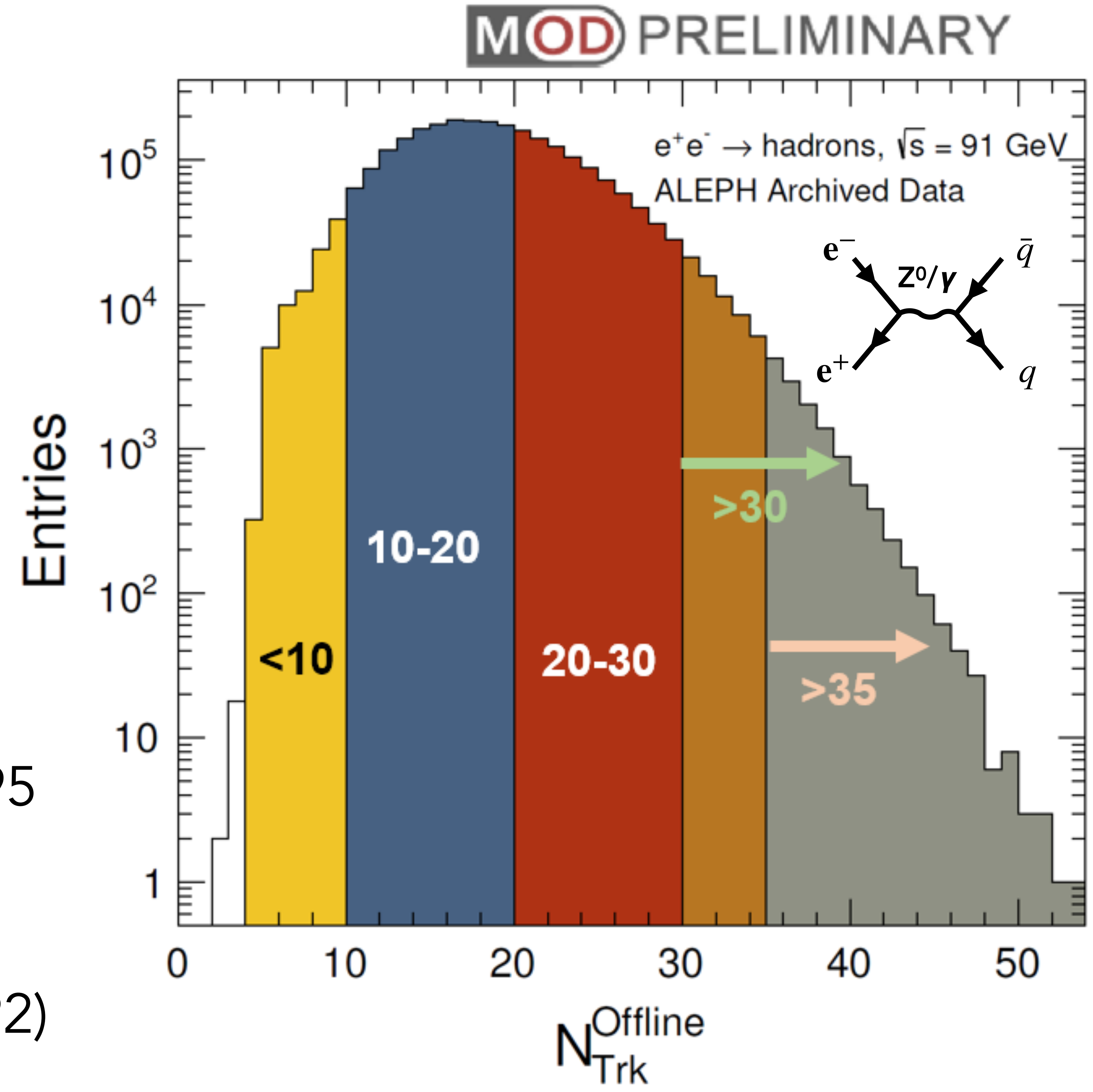
- What causes the α_S discrepancies in e^+e^- event shape extractions?
- What is the scale of the NP effects at high Q^2? Can the NP shape functions be constrained beyond first moment?
- How well does P QCD predict the final state particle spectrum?

ALEPH Experiment



Apparatus for LEP PHysics (ALEPH)

Charged particle multiplicity

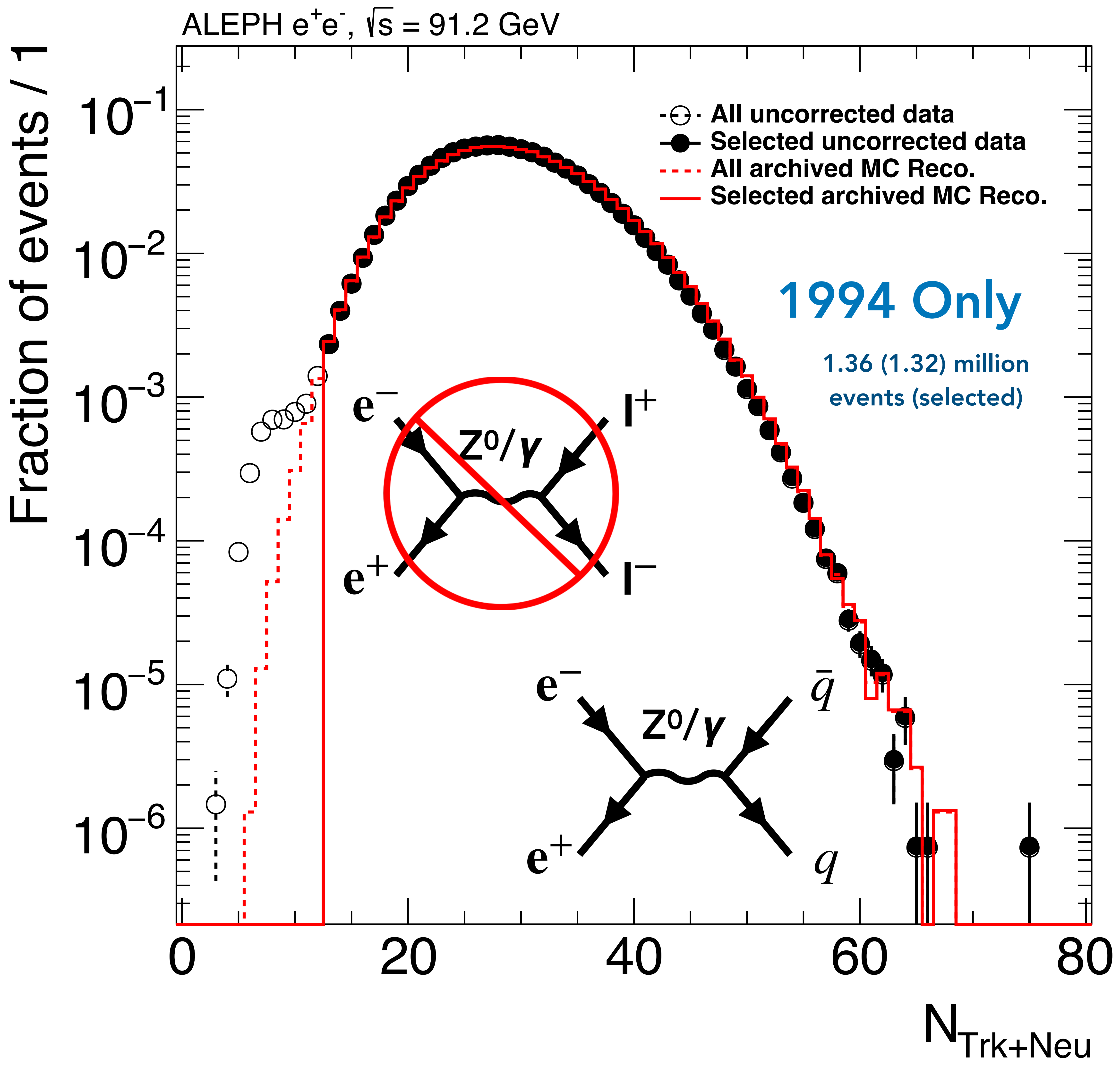


- LEP1 e^+e^- data at Z pole (91 GeV) taken between 1992-1995
- Approximately 2.5 million hadronic events were recorded
- Raw txt files converted to modern ntuples. Access to low level detector quantities. Archived MC (LEP1 1994, full LEP2)

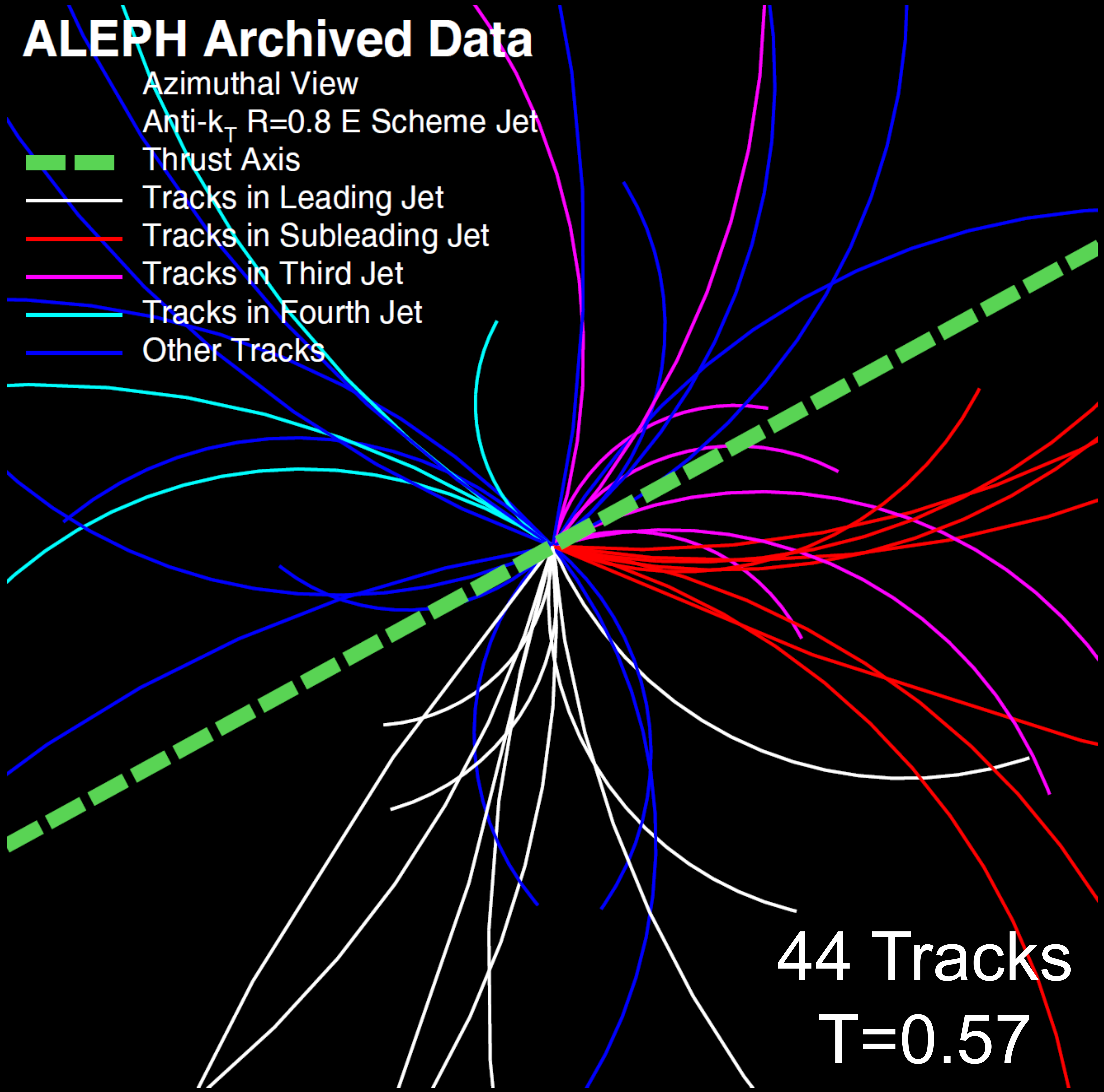
Hadronic Event Selection

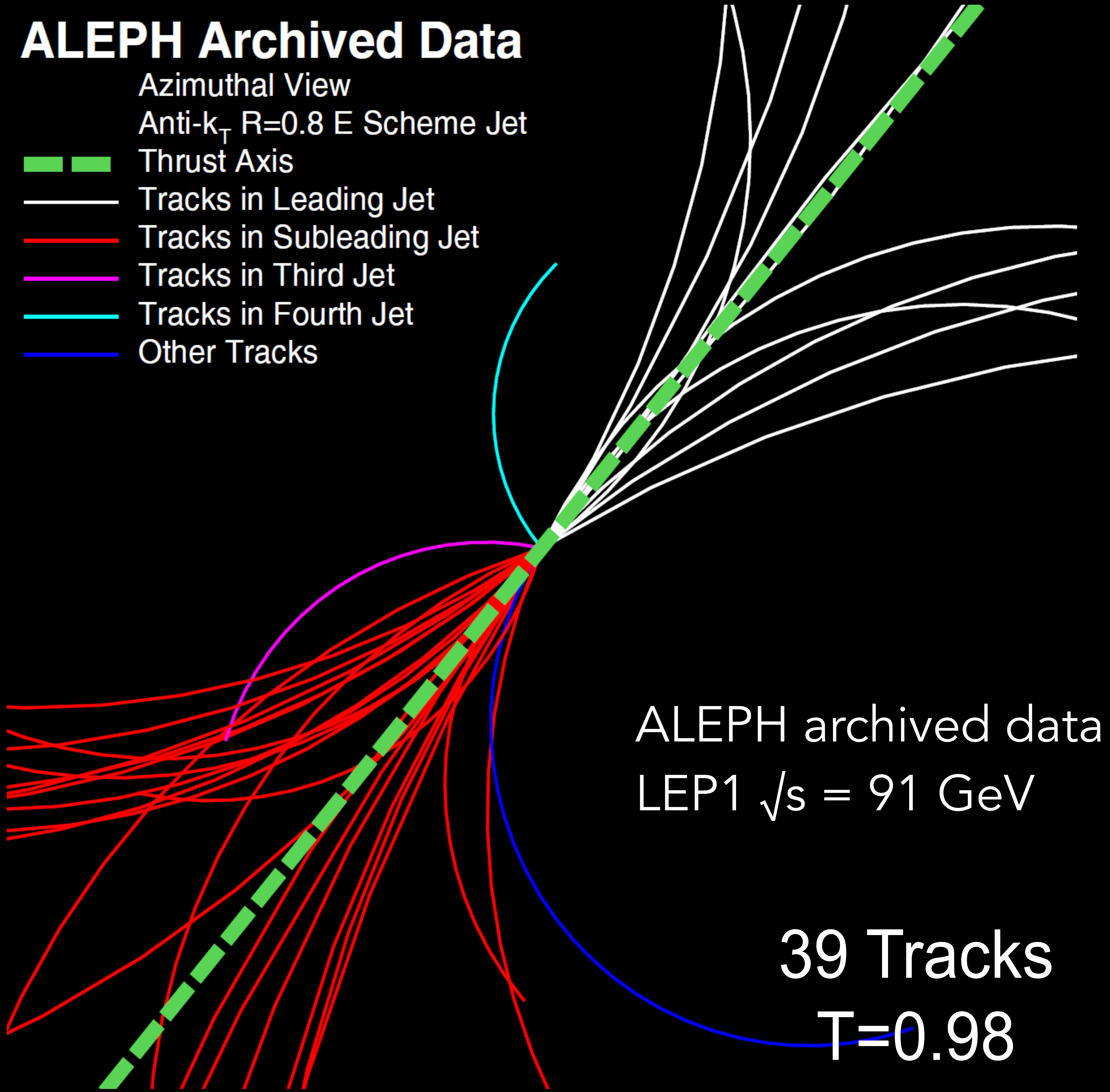
Same selection criteria from ALEPH 2004 QCD paper

- **Track Selections:**
  - Number of TPC hits for a charged tracks ≥ 4
  - $|d0| \leq 2$ cm
  - $|z0| \leq 10$ cm
  - $|\cos\theta| \leq 0.94$ (corresponding to $|\eta| \leq 1.74$)
  - $pT \geq 0.2$ GeV (pT with respect to beam axis)
  - $\chi^2/ndf < 1000$
- **Neutral Hadron Selections:**
  - ECAL/HCAL objects
  - $E \geq 0.4$ GeV
  - $|\cos\theta| \leq 0.98$
- **Event Selections:**
  - Number of good charged particles ≥ 5 (including charged hadrons and leptons)
  - Number of good charged+neutral ≥ 13
  - $E_{charged} \geq 15$ GeV
  - $|\cos\theta_{sphericity}| \leq 0.82$



Low and High Thrust e^+e^- Events in LEP1 Data





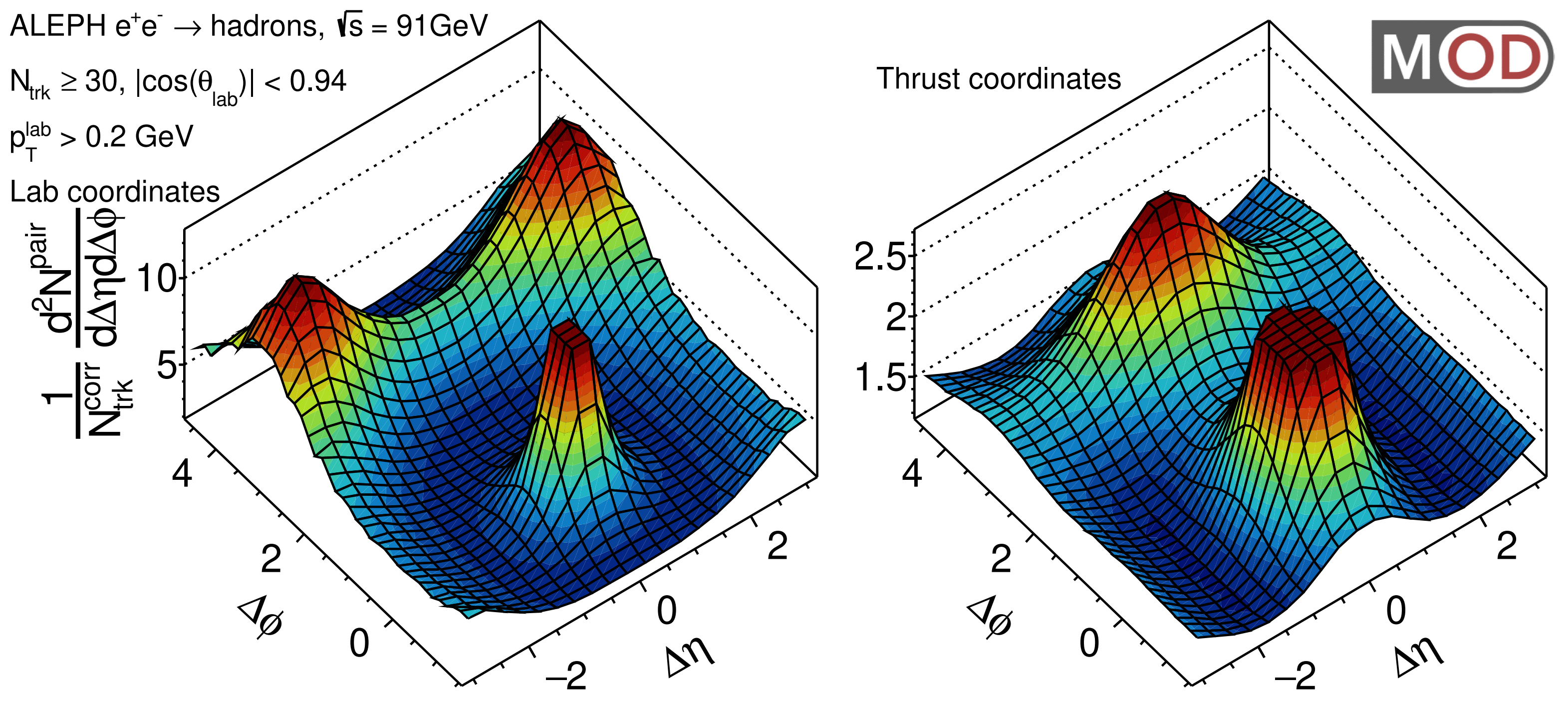
ALEPH archived data

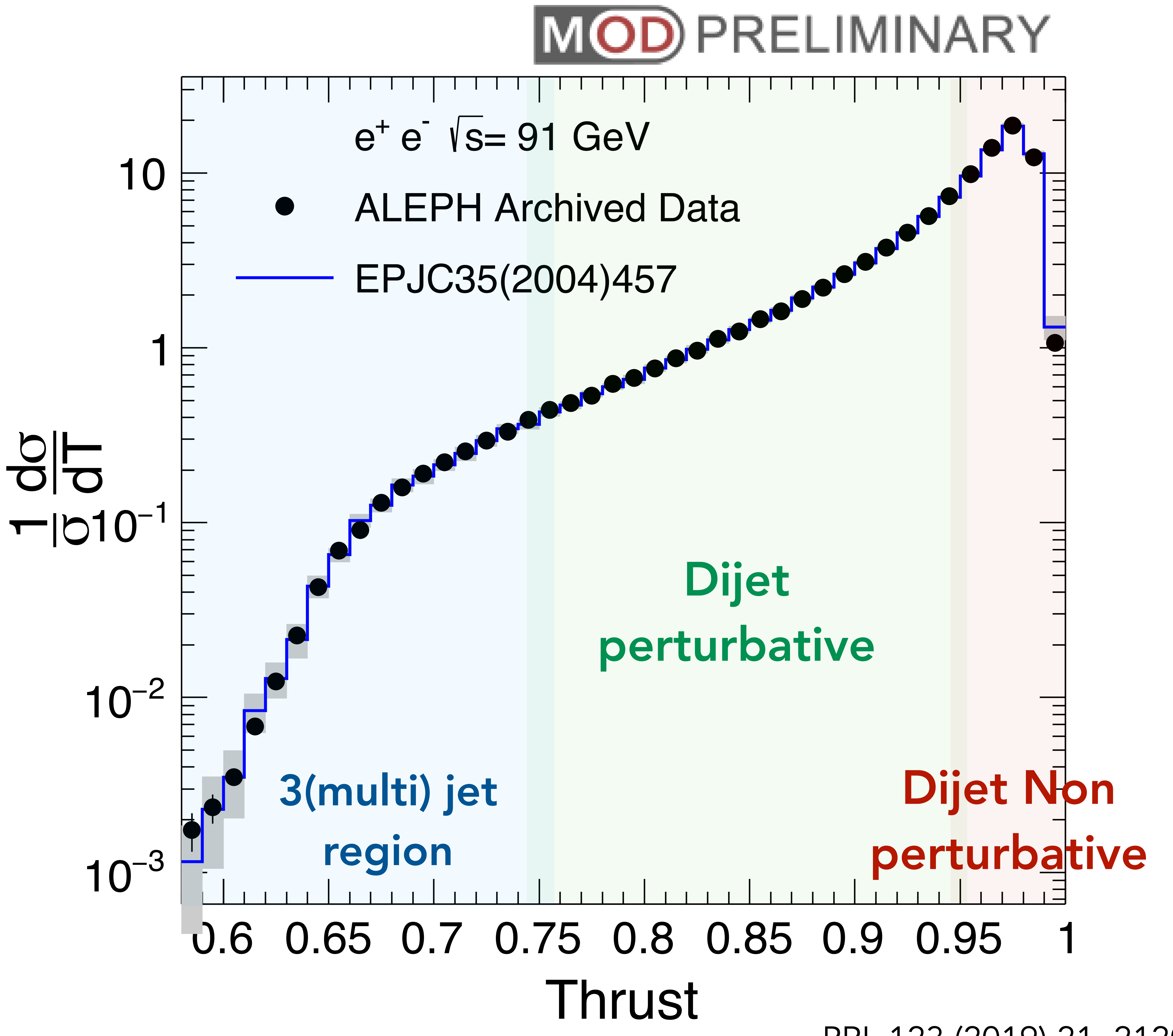
LEP1 $\sqrt{s}$ = 91 GeV

~2019 Thrust in e+e- ALEPH Archived Data

First thrust study to validate dataset and use for two particle correlation

Corrected with iterative bayesian unfolding (IBU), stat only uncertainty, compared with previous ALEPH result



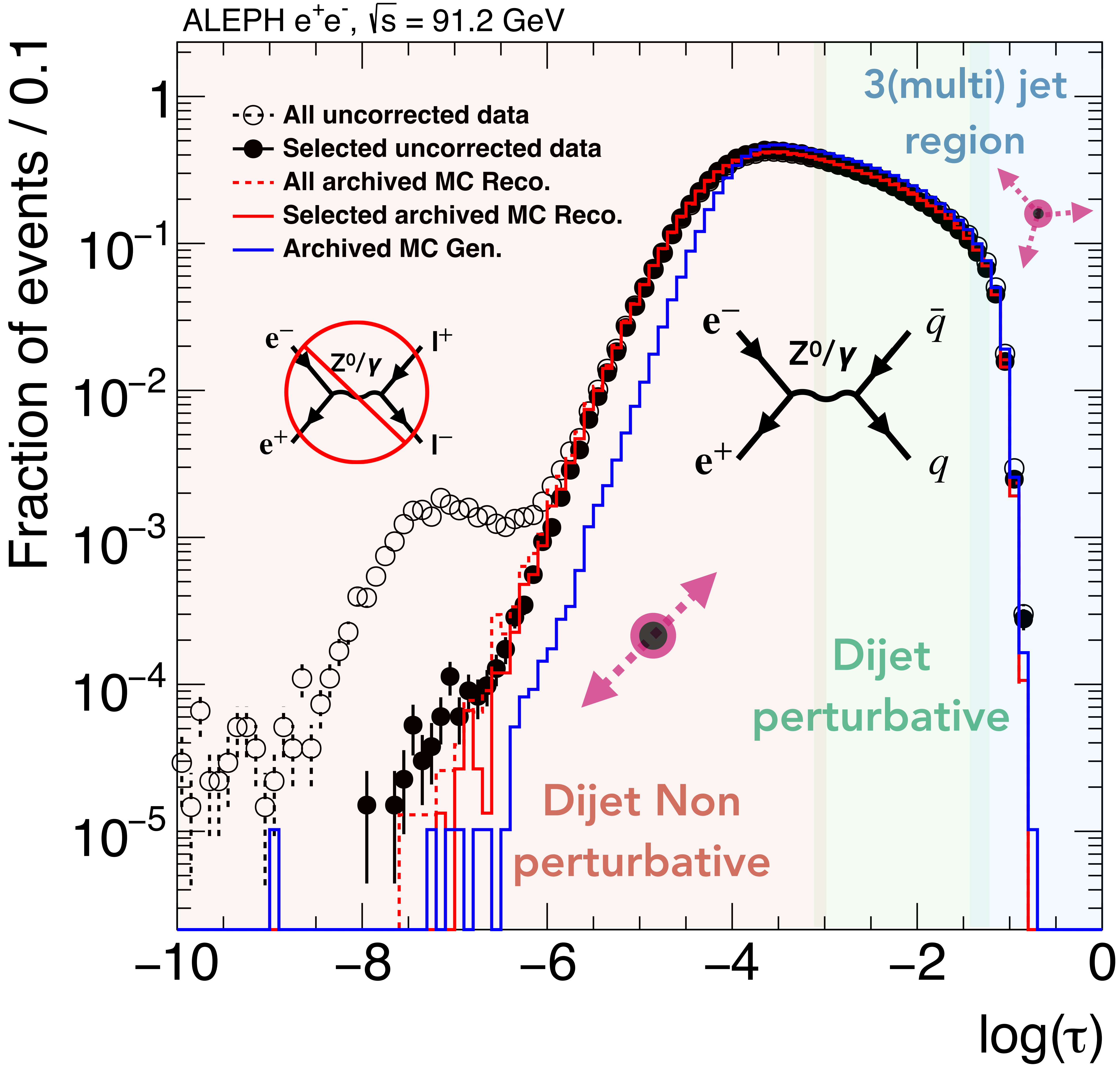


PRL 123 (2019) 21, 212002
Badea MIT B.S. Thesis (2019)

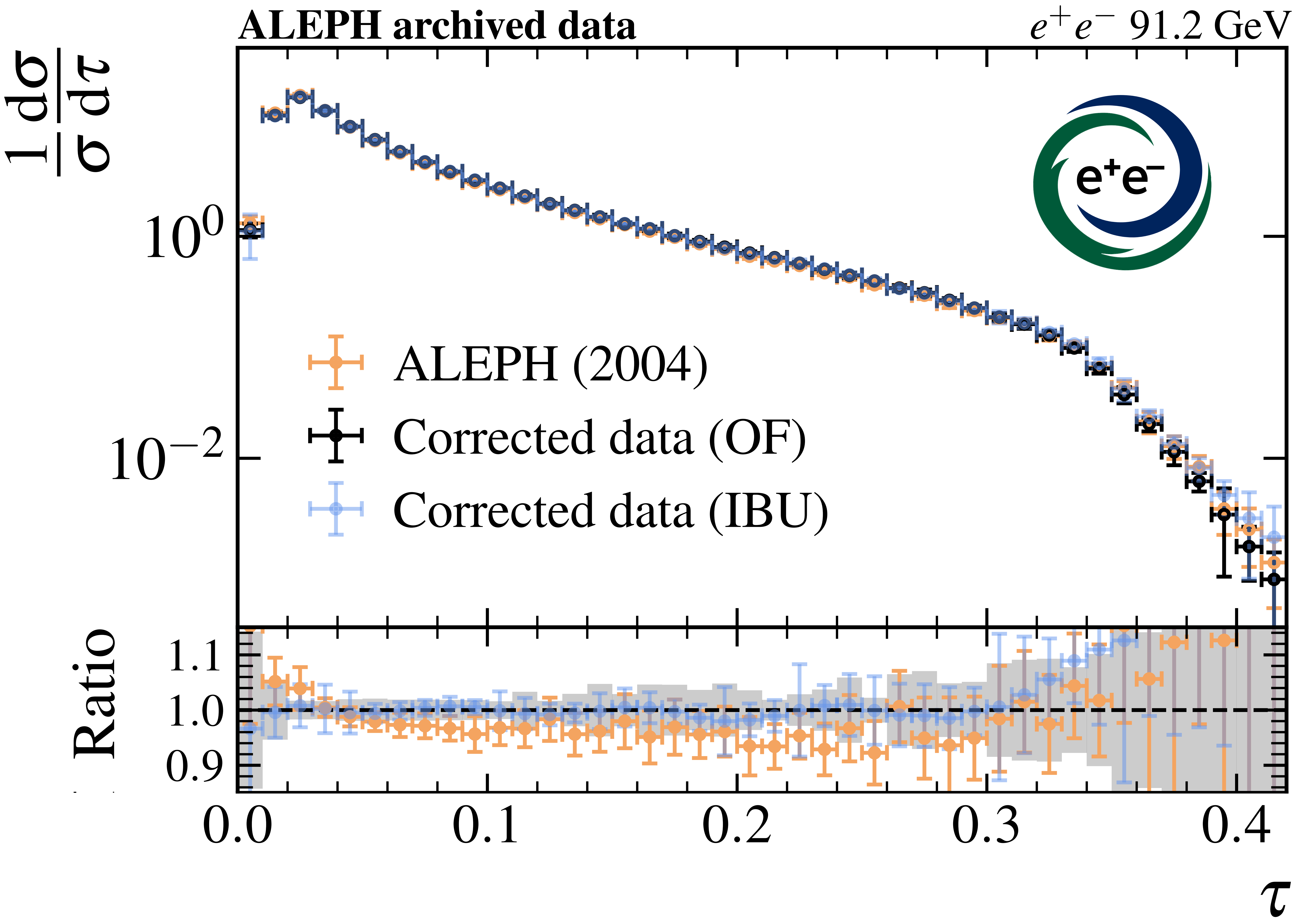
2025 Unbinned Measurement of Thrust

Analysis Summary

- Corrections for hadronic event selection, tracking efficiency, and EM ISR/FSR effects
- OmniFold unbinned unfolding from detector to particle level
- Systematics on event selections, charged track selections, neutral particle energy and efficiency, theory MC prior



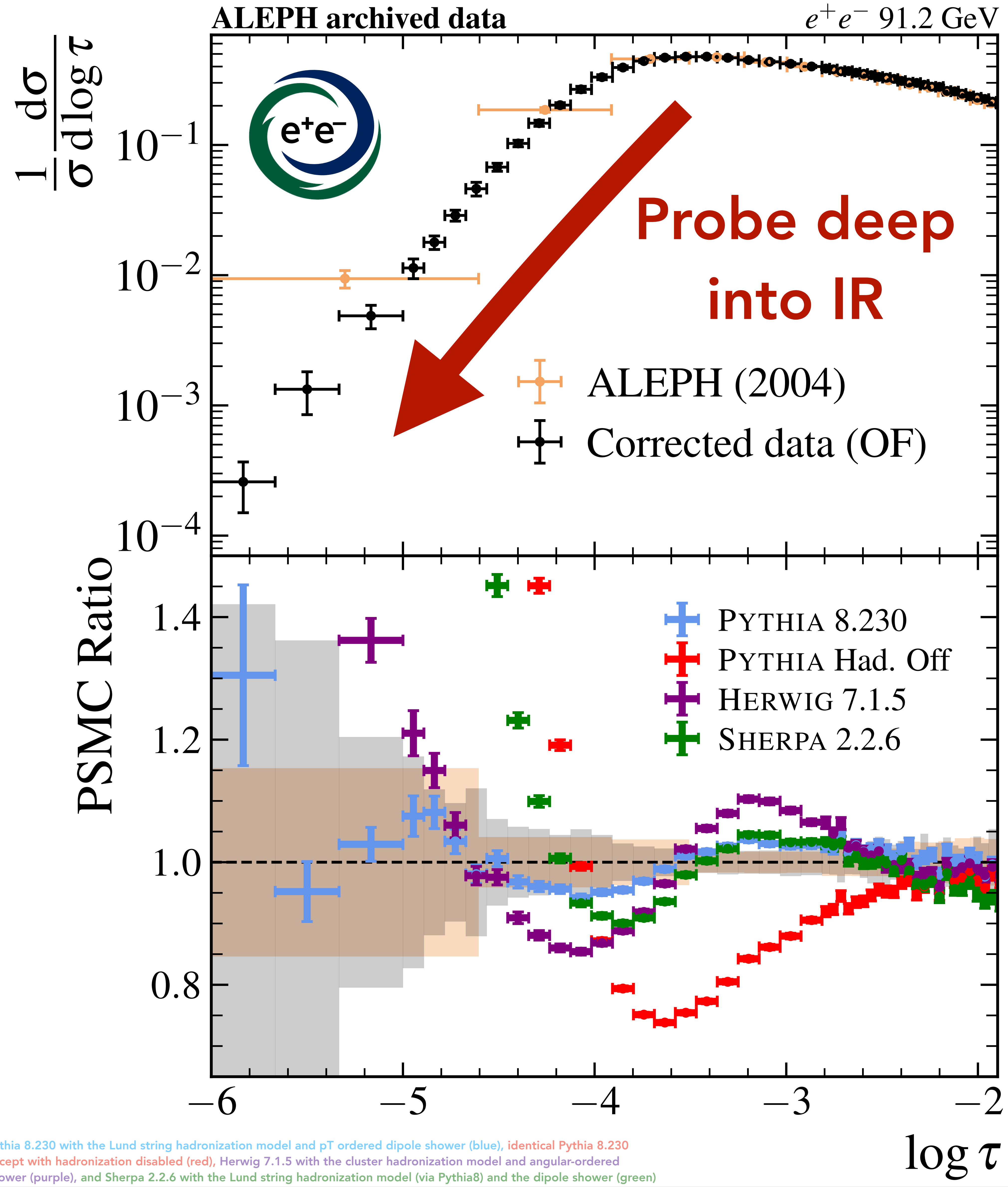
Fully Corrected Thrust in ALEPH Bins



Good agreement between OF/IBU result

Differences between ALEPH and reanalysis??

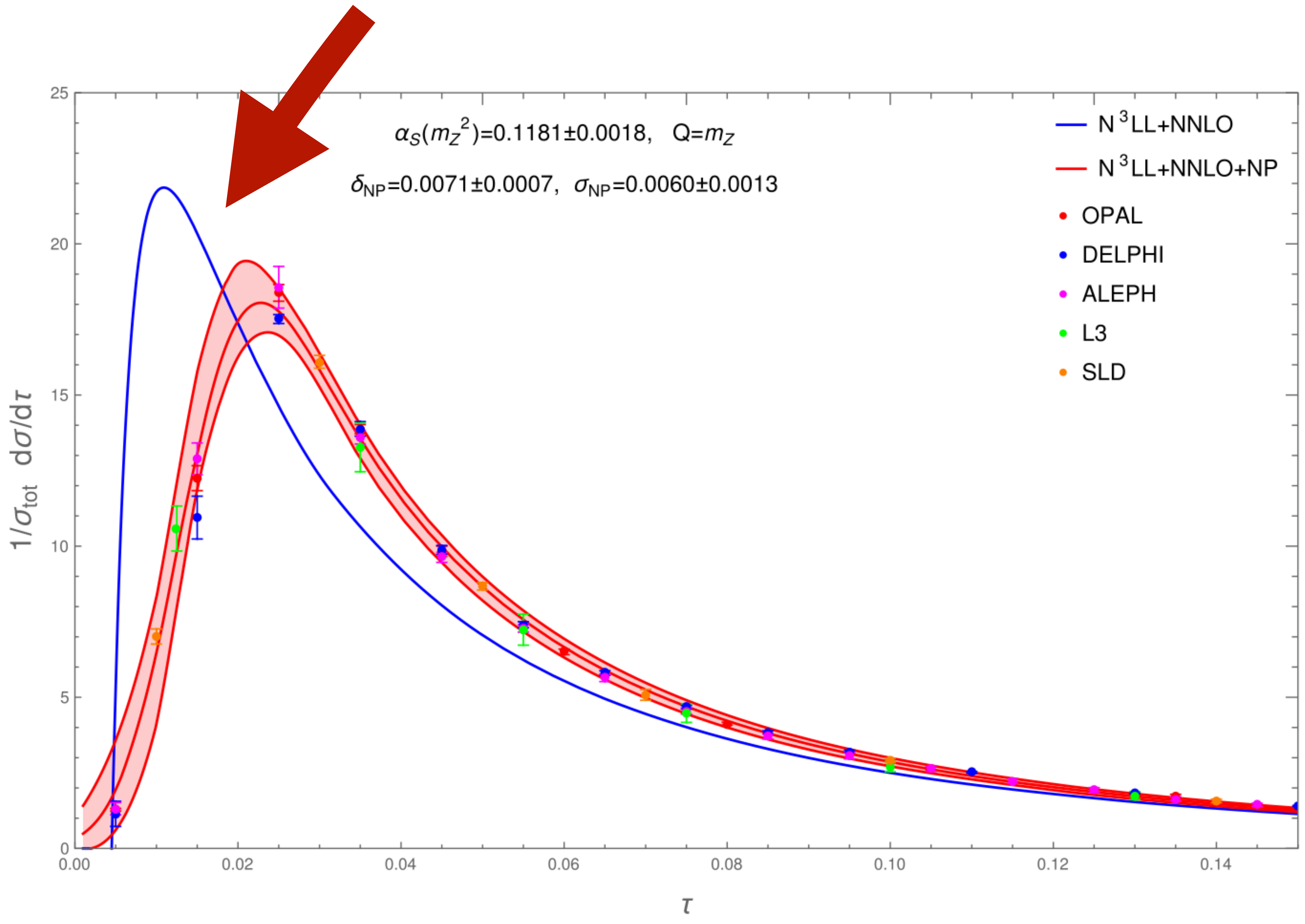
Fully Corrected Thrust Deep Into the IR



Pythia 8.230 with the Lund string hadronization model and pT ordered dipole shower (blue), identical Pythia 8.230 except with hadronization disabled (red), Herwig 7.1.5 with the cluster hadronization model and angular-ordered shower (purple), and Sherpa 2.2.6 with the Lund string hadronization model (via Pythia8) and the dipole shower (green)

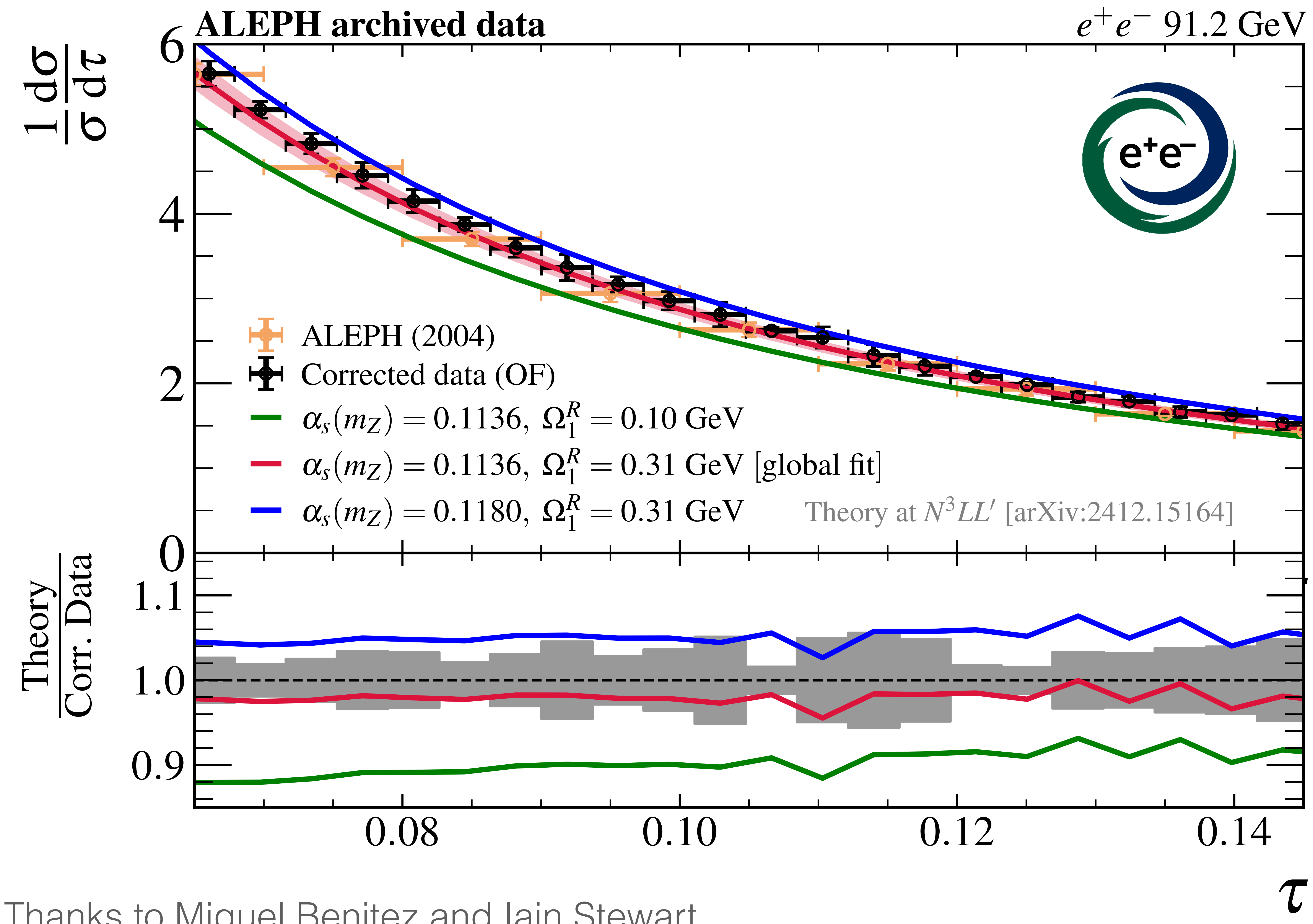
Able to probe deep into the infrared where NP QCD effect become significant

(observable $\tau \sim \Lambda_{QCD}/\sqrt{s}$)



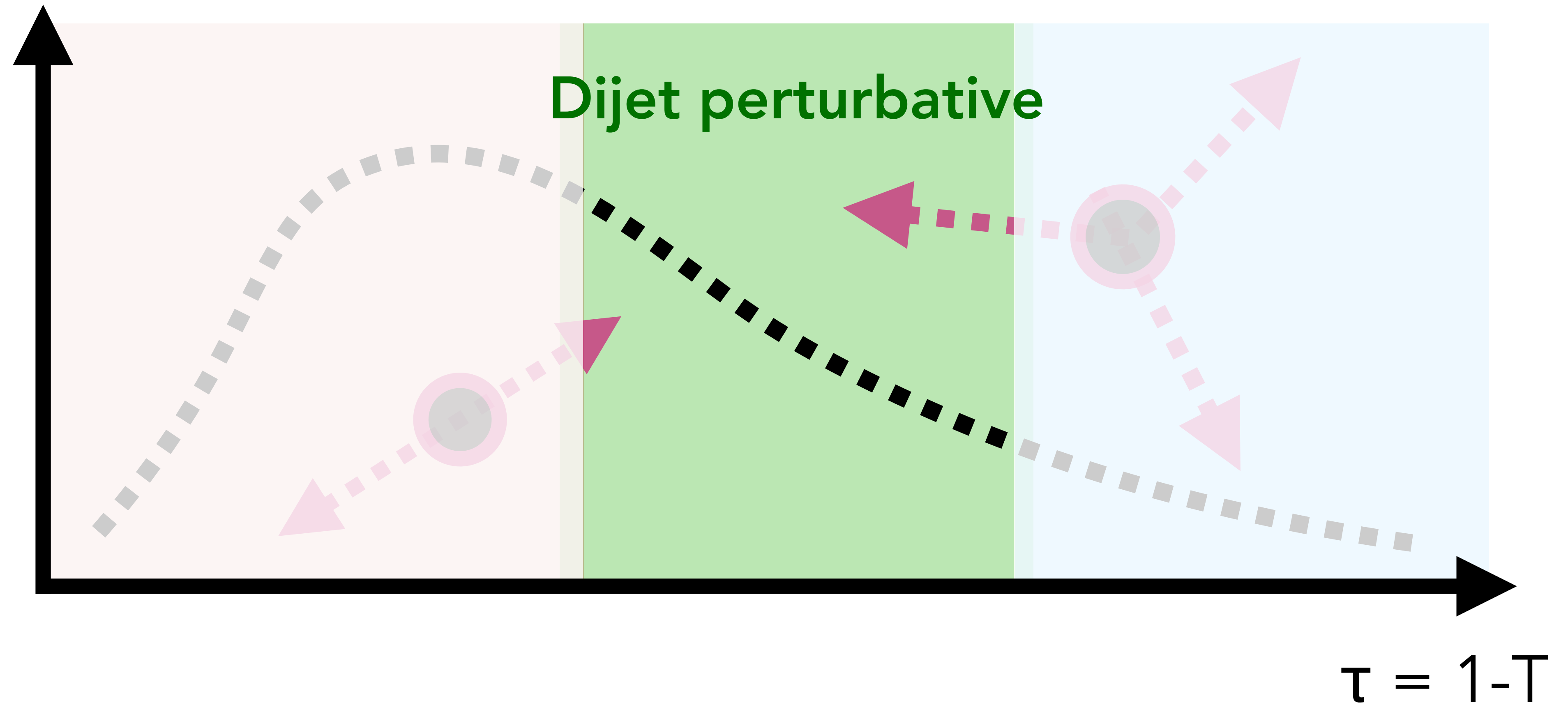
Phys. Rev. Lett. 134, 251904, 0809.3326

Very Interesting Theory Comparison!



Comparison with the state of the art theoretical extraction yielding lower α_S

Potential agreement with theory global fit $\alpha_S(m_Z)$ = 0.1136 (excluding this result) and world average for fixed Ω_1^R = 0.31 GeV (NP shape function first moment)



Thanks to Miguel Benitez and Iain Stewart

On Determining $\alpha_S(m_Z)$ from Dijets in e+e- Thrust, 2412.15164

Investigating the Thrust Shift

(1) Thrust Observable Calculation:

Modern exact vs historic heuristic calculation
Selections result in no change in spectrum from varied calculation

(2) Unfolding and Resolution Effect Correction:

IBU vs. Omnifold vs Bin-by-bin
Consistent results covered by unfolding uncertainties

(3) Charged only vs Charged + Neutral vs. adding MET:

Performed the systematics check and the difference are quoted as part of systematics

(4) Charged particle p_T threshold:

Performed the systematics checks. If we apply higher charged particle pT threshold than 0.2 GeV (used in ALEPH QCD paper) could **narrow the raw spectra** (as expected)

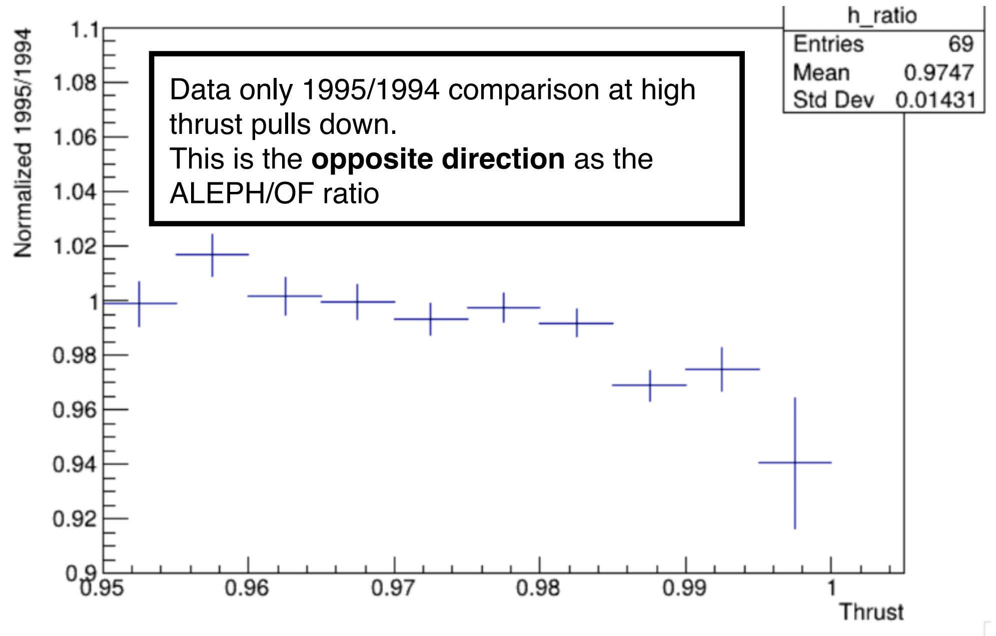
(5) More simulated MC samples:

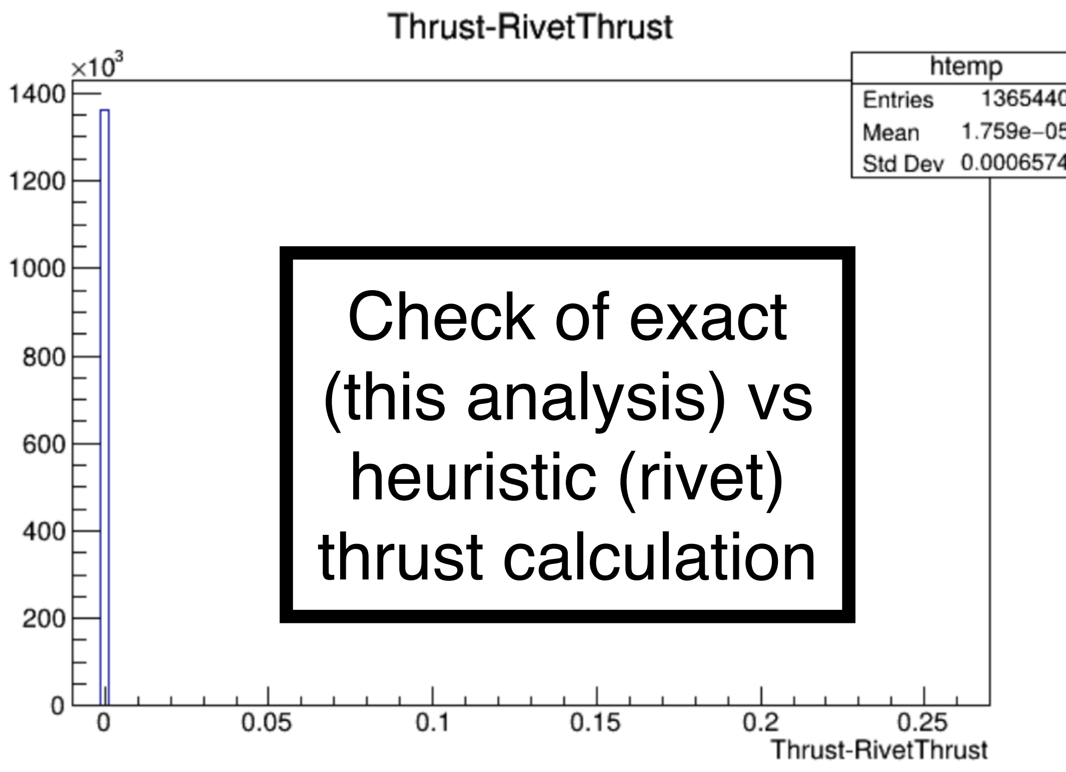
Currently we have analyzed archived 1994 MC available to us.
In the near future, we will get the galeph simulation to work for further studies.

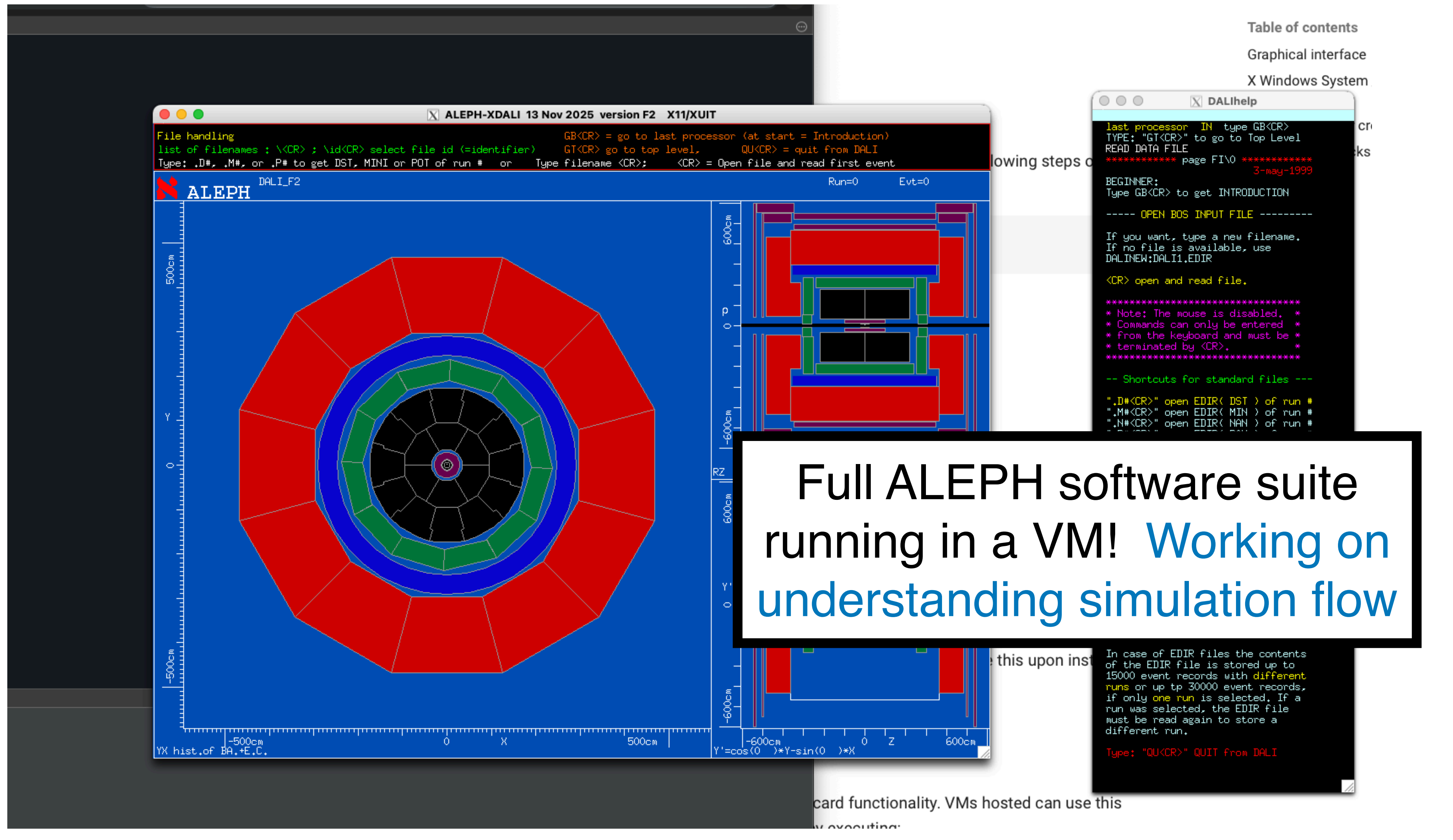
(6) Cross-experiment comparison:

Comparison with DELPHI measurement!

Comments and suggestions are welcome!!

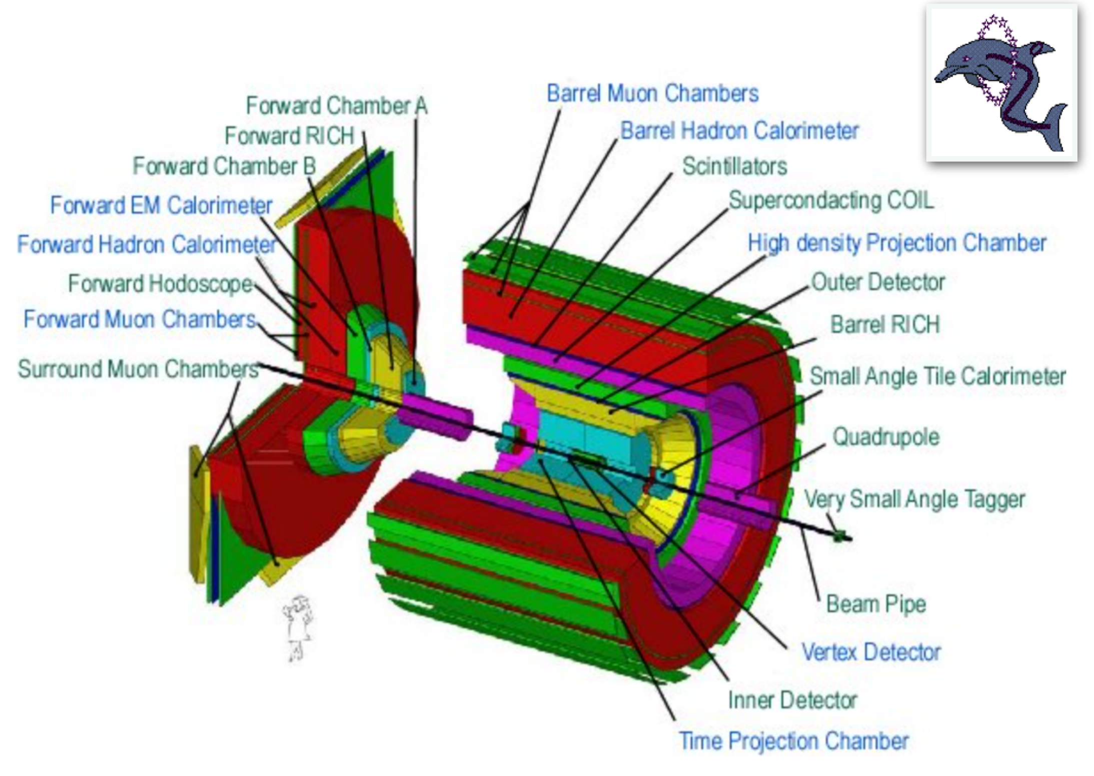


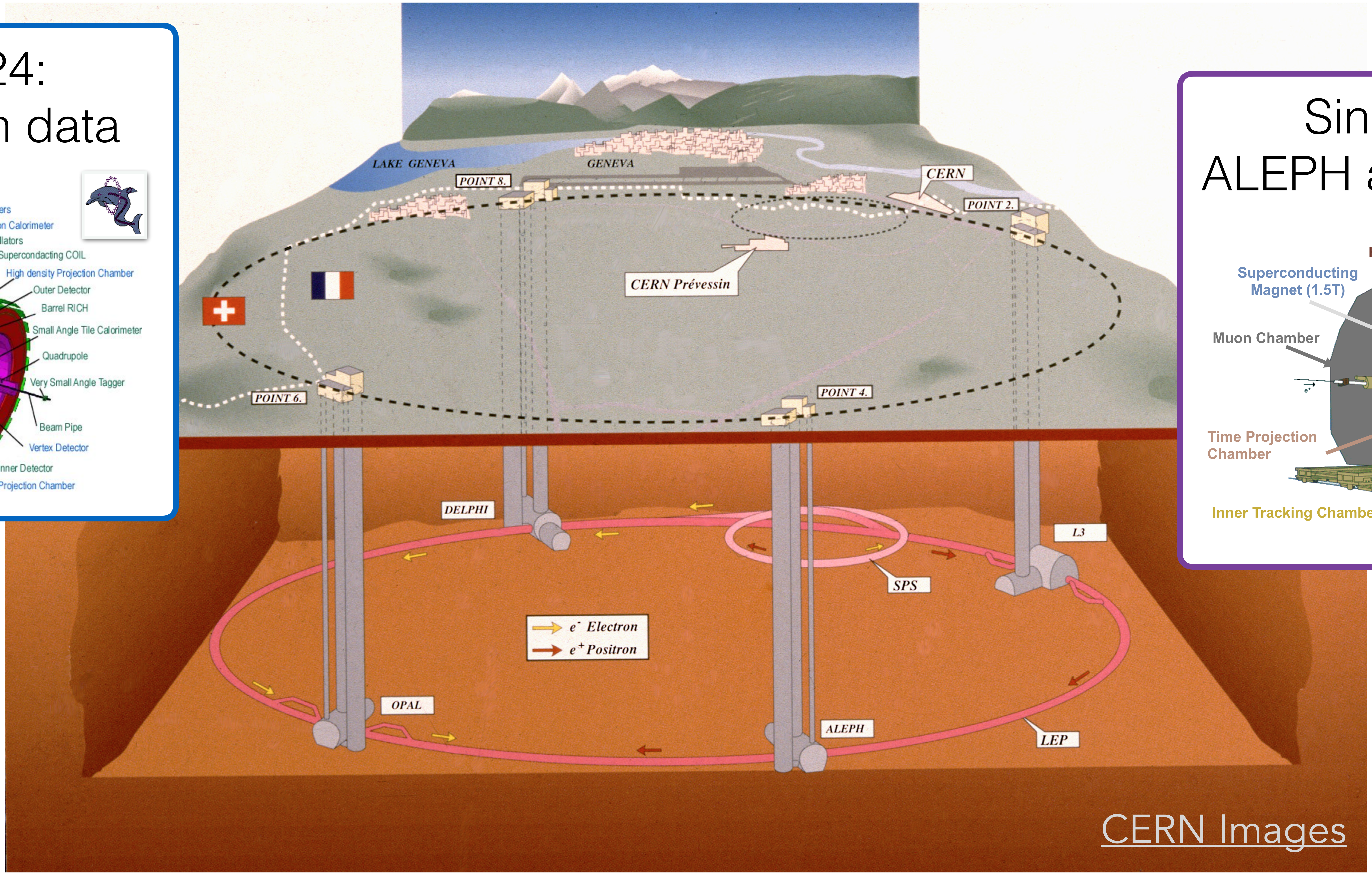




Cross-Experiment Comparison

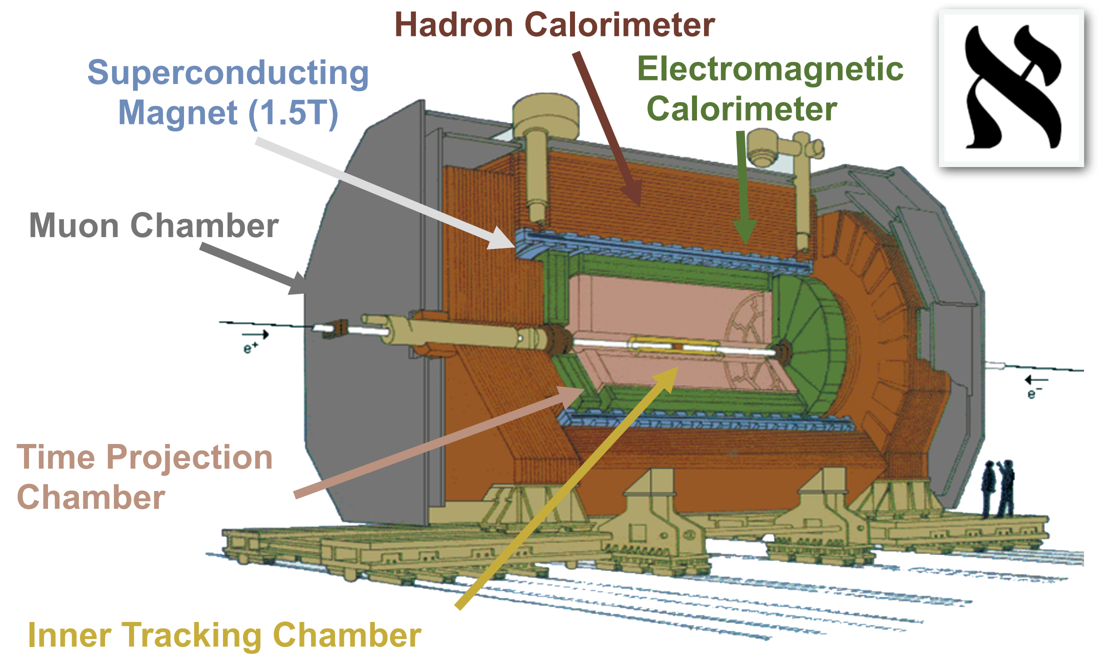
Since 2024:
DELPHI open data

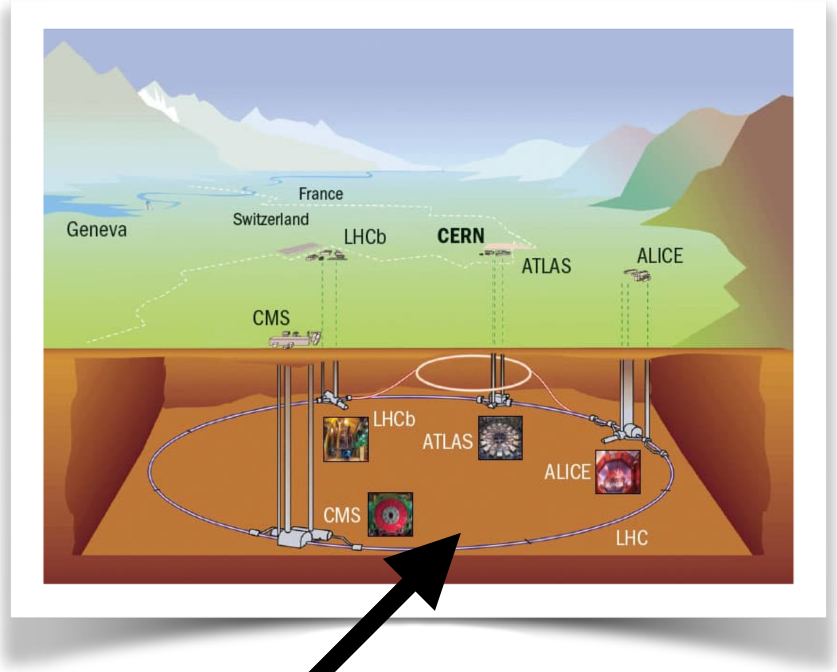




[CERN Images]

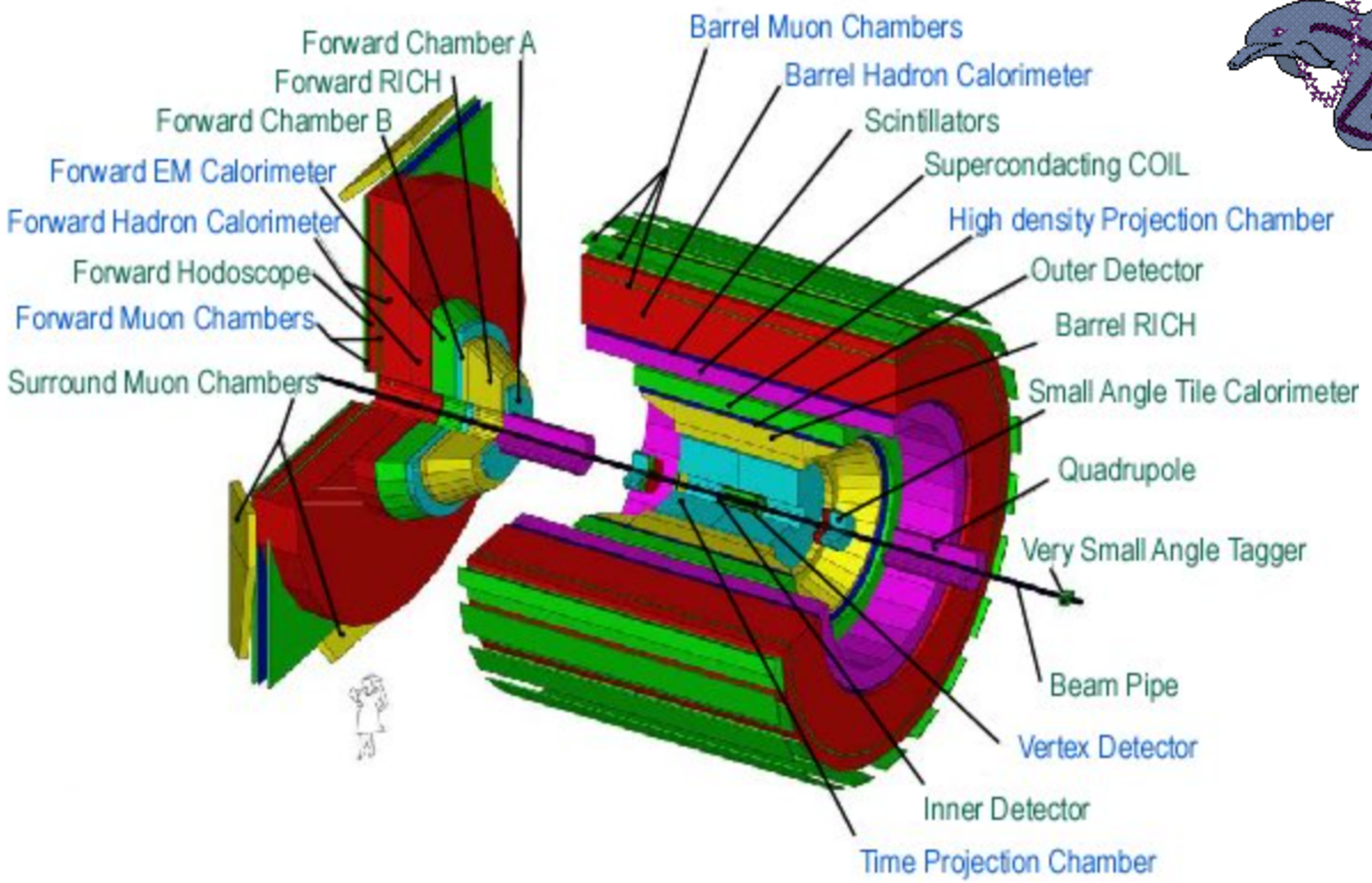
Since 2018:
ALEPH archived data





Can we achieve a LEP archived data cross-experiment comparison? (e.g. LHC style ATLAS vs CMS)

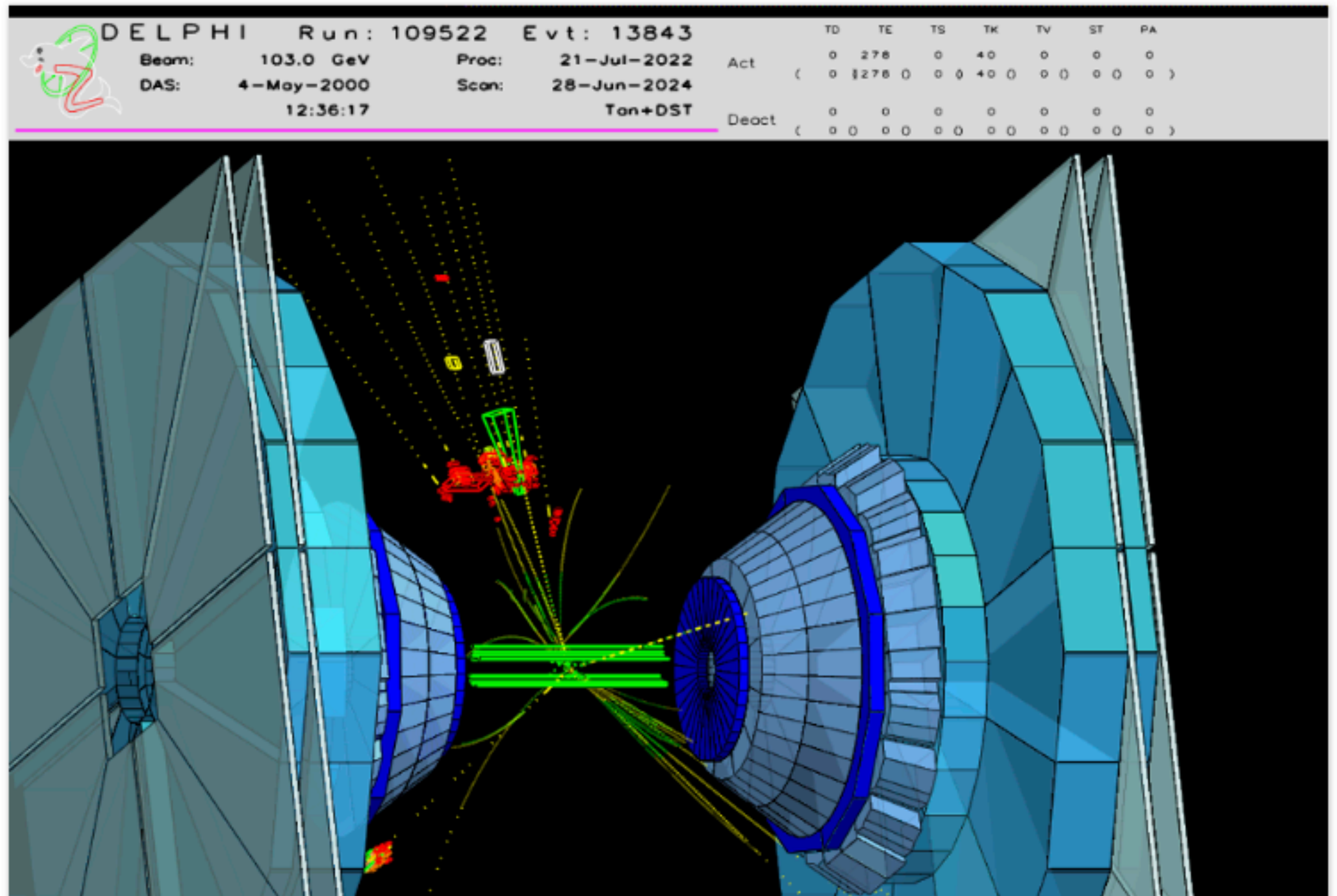
DELPHI Experiment Open Data



Digital archaeology: new LEP data now available to all

Retrieving and preserving access to data from experiments that ran in the 90s is a complex task carried out by passionate experts in the IT department

15 JULY, 2024 | By Antonella Del Rosso



Jingyu Zhang (Vanderbilt)

Luna Chen (Vanderbilt)

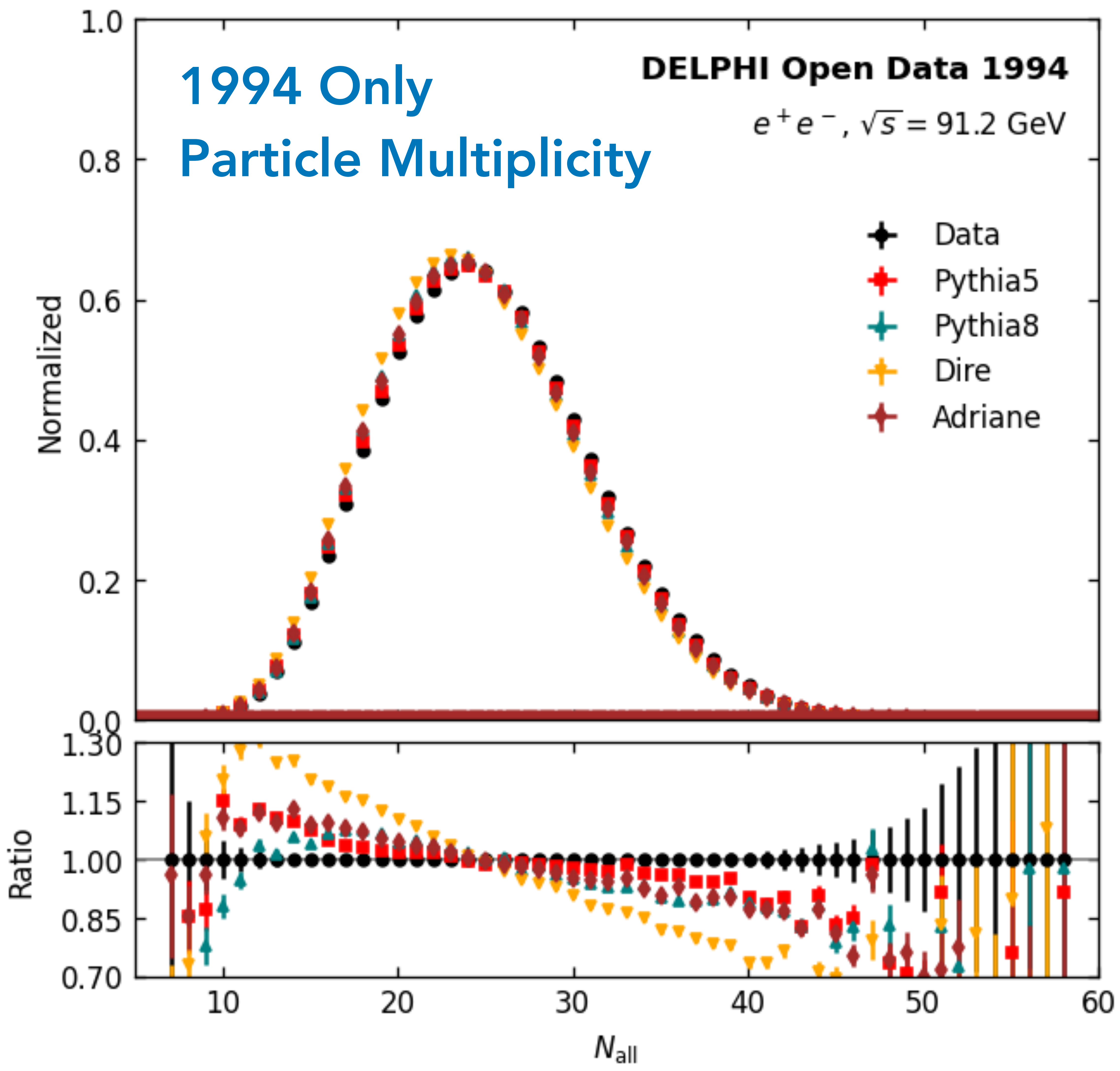
- LEP1 e^+e^- data at Z pole (91 GeV) taken between 1992-1995
- Open data released late 2024, converted to ntuples. Access to all simulation software, enabling new sim with modern MC

Hadronic Event Selection

Particle and Event Level Selections

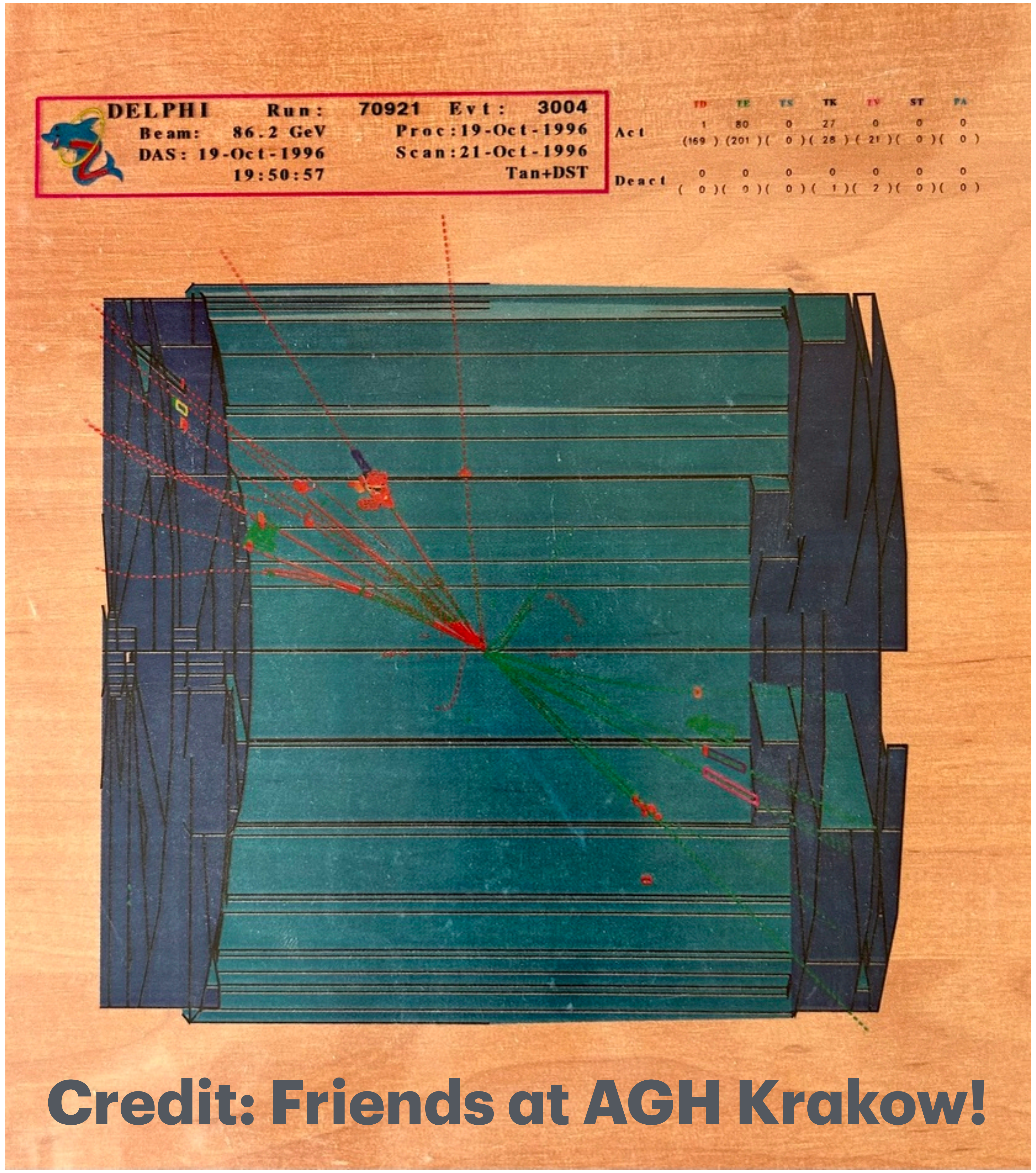
neutral particle selection	$E \geq 0.5\,\text{GeV}$ $20° \leq \theta \leq 160°$
charged particle selection	$0.4\,\text{GeV} \leq p \leq 100\,\text{GeV}$ $\Delta p/p \leq 1.0$ measured track length ≥ 30 cm distance to I.P. in $r\phi$ plane ≤ 4 cm distance to I.P. in $z \leq 10$ cm $20° < \theta < 160°$
Standard event selection	$N_{\text{ch}} \geq 7$ $30° \leq \theta_{\text{Thrust}} \leq 150°$ $E_{\text{tot}} \geq 0.50 E_{\text{cm}}$

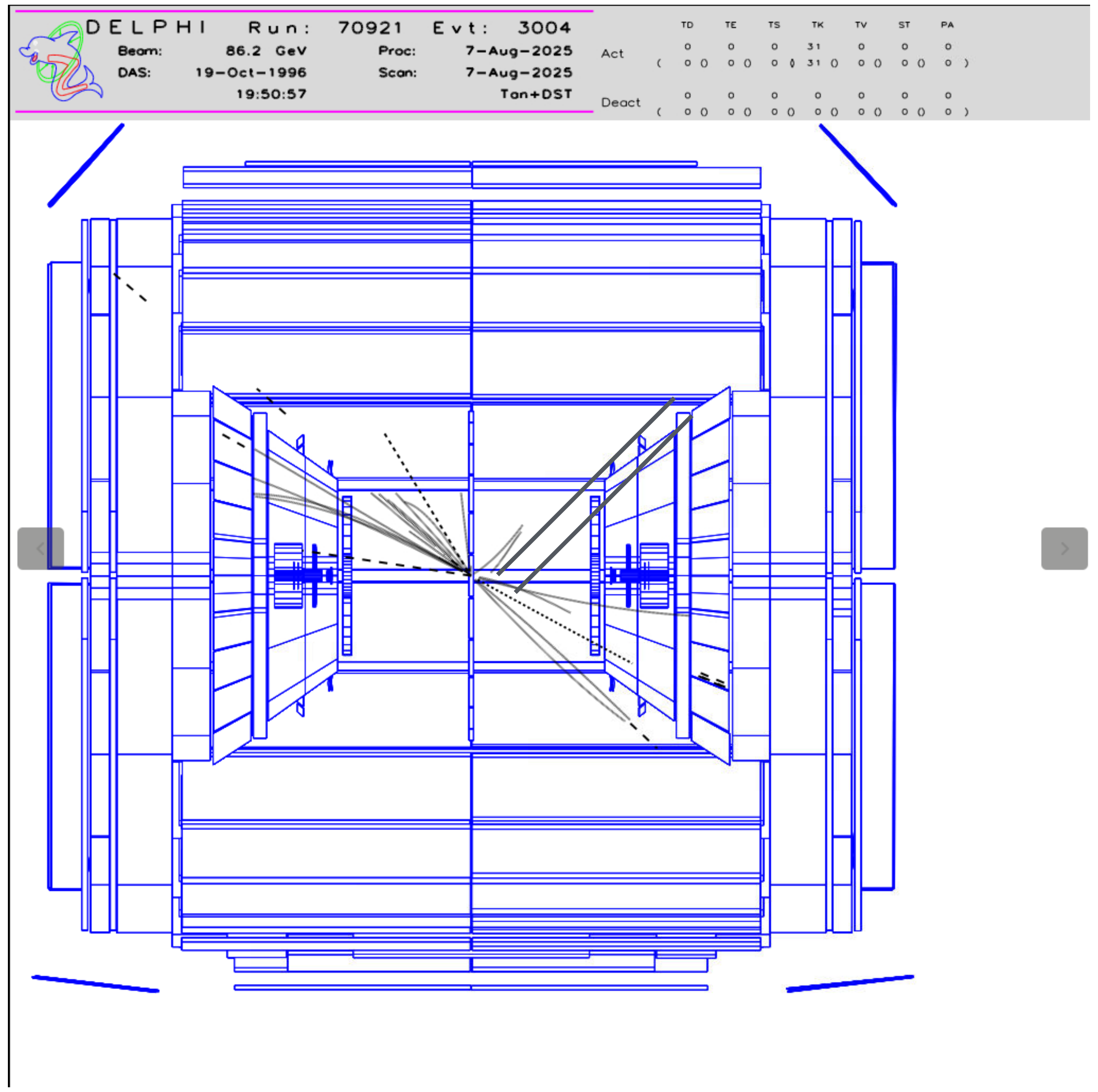
- Pythia 5.7 and ARIADNE (open data with DELPHI tune)
- Pythia 8.3 and Dire (**new MC + DELSIM**)!!

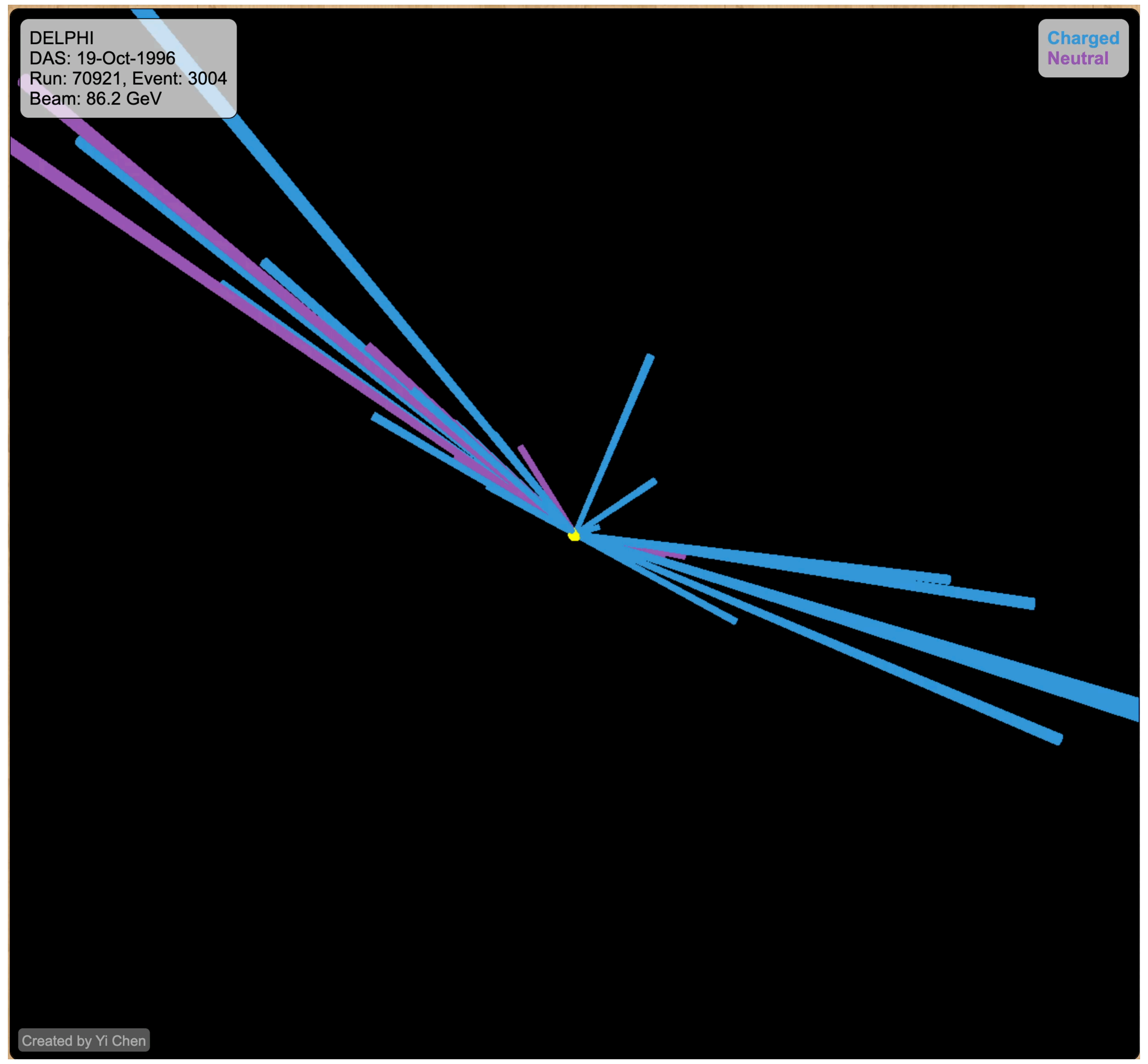


Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

The Same Event Display Over 30 Years







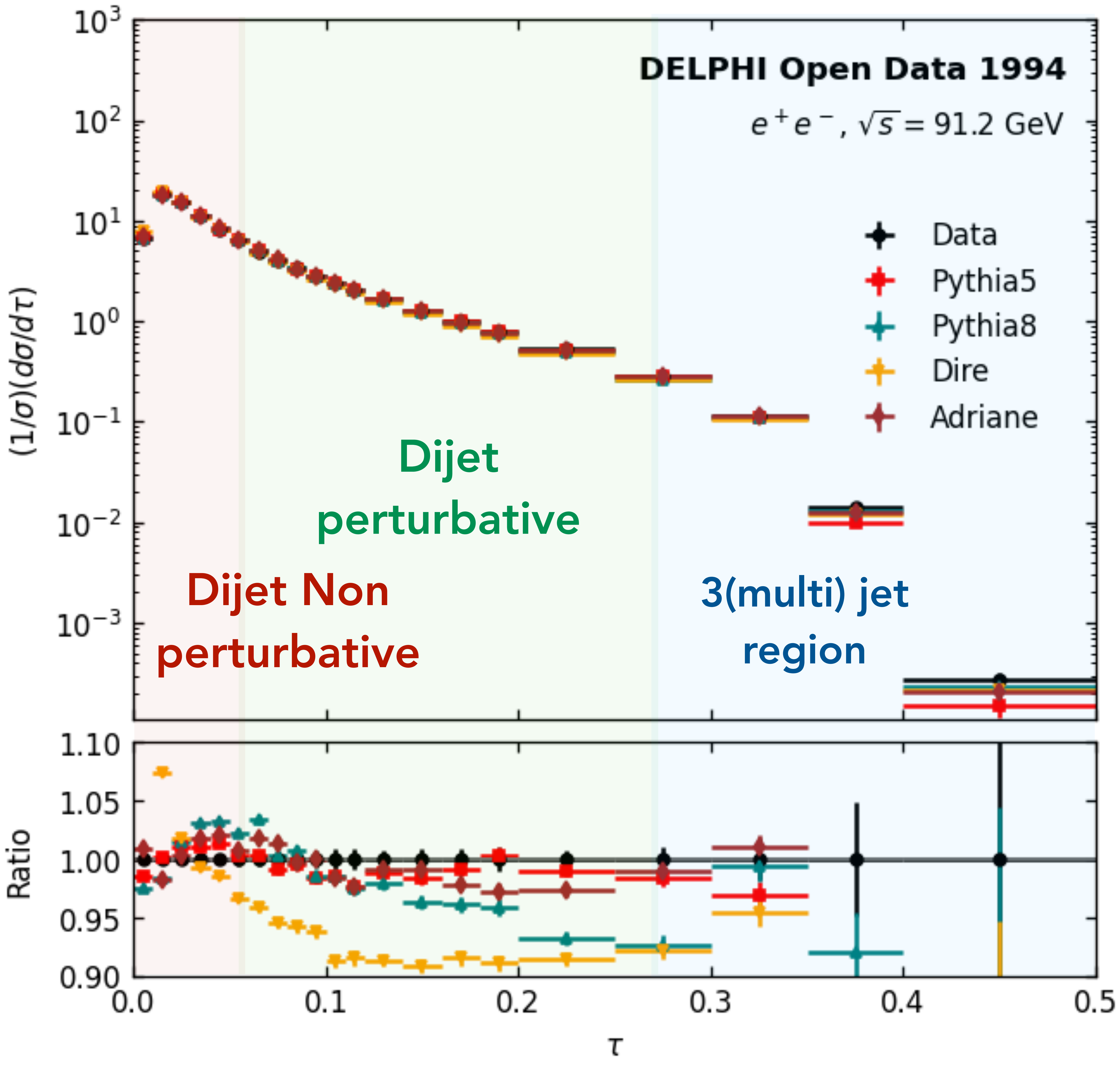
Event from Oct. 19 1996 traced from physical print out → original software → modern event displays

Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

2025 Measurement of Thrust

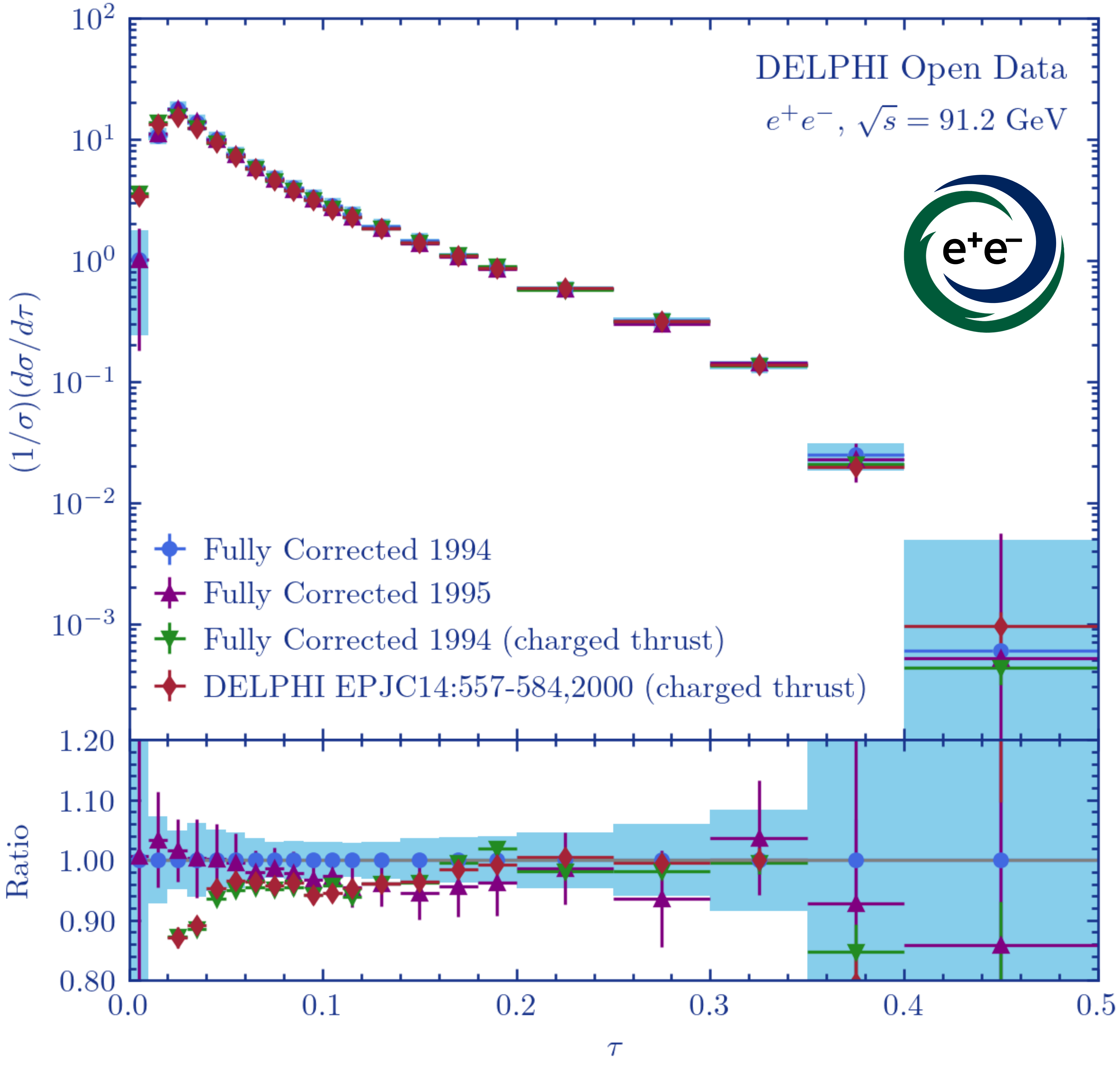
Analysis Summary

- Corrections for hadronic event selection, tracking efficiency
- IBU unfolding from detector to particle level
- Fiducial to full phase space and EM ISR/FSR effects
- Systematics on event selections, charged track selections, neutral particle energy and efficiency, theory MC prior



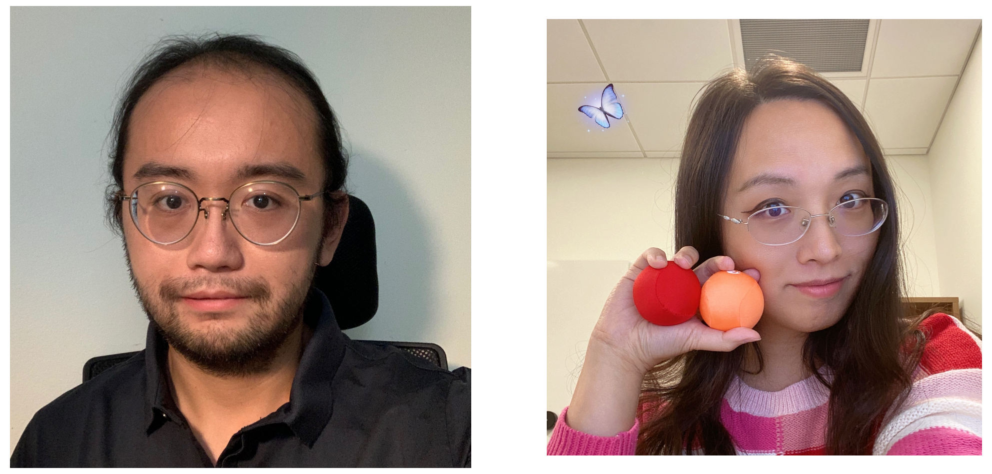
Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

Fully Corrected Thrust Spectrum



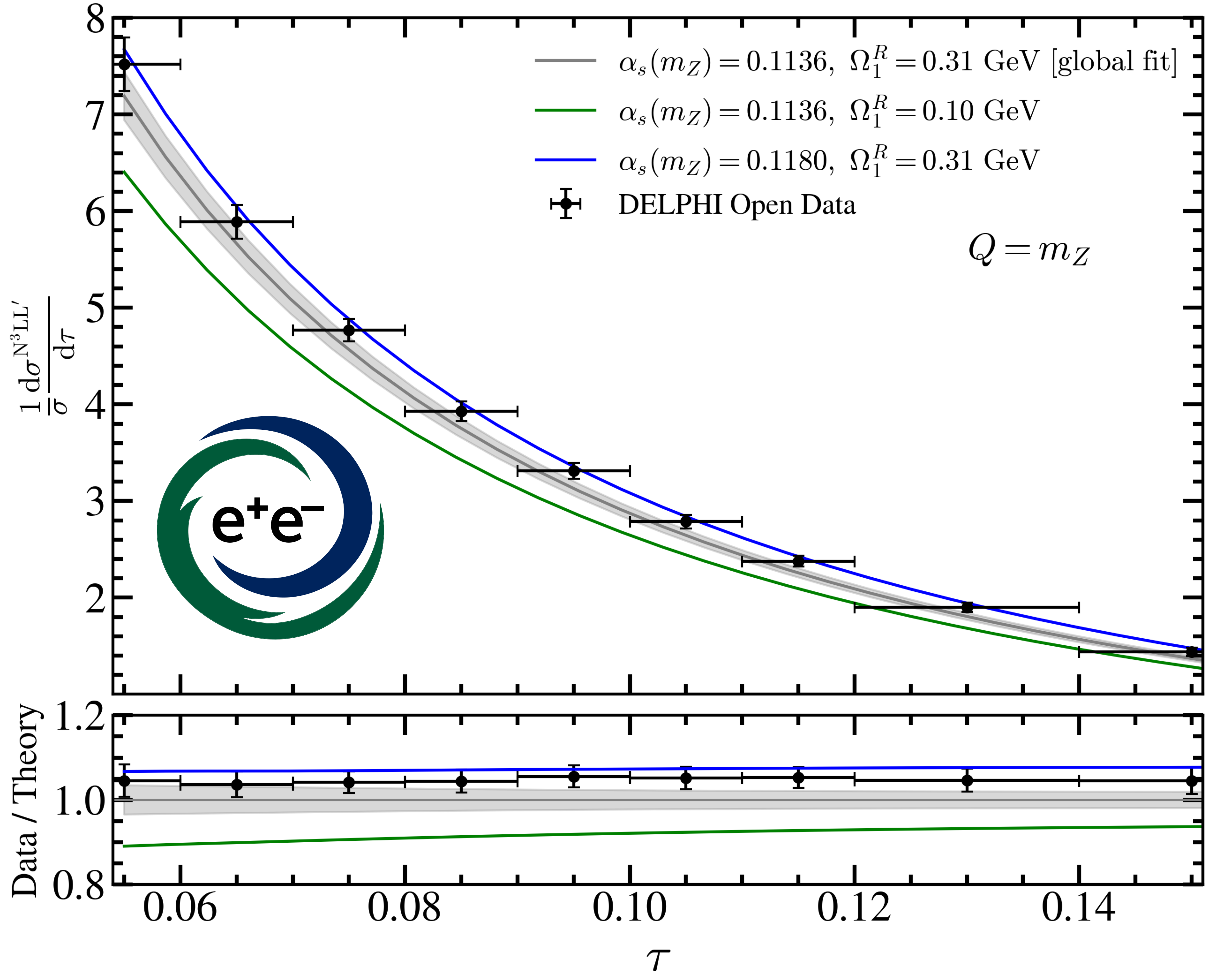
Previous DELPHI all particle thrust measurements results not on HepData

For charged particle only measurement, cross-check (no systematics) agrees well with previous DELPHI result using 1994



Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

Theory Comparison

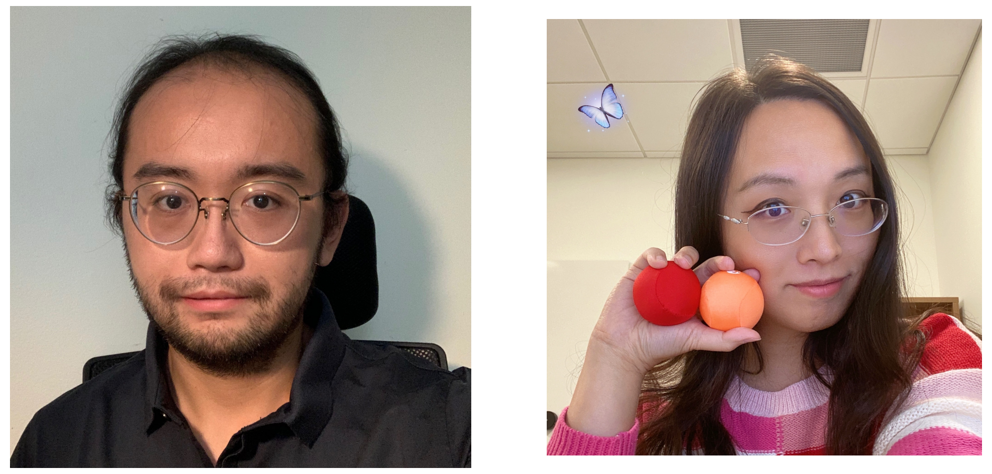


Combined 1994/1995 by luminosity

Possible agreement also with an α_S shifted up in comparison to theory global fit (excluding these analyses)!

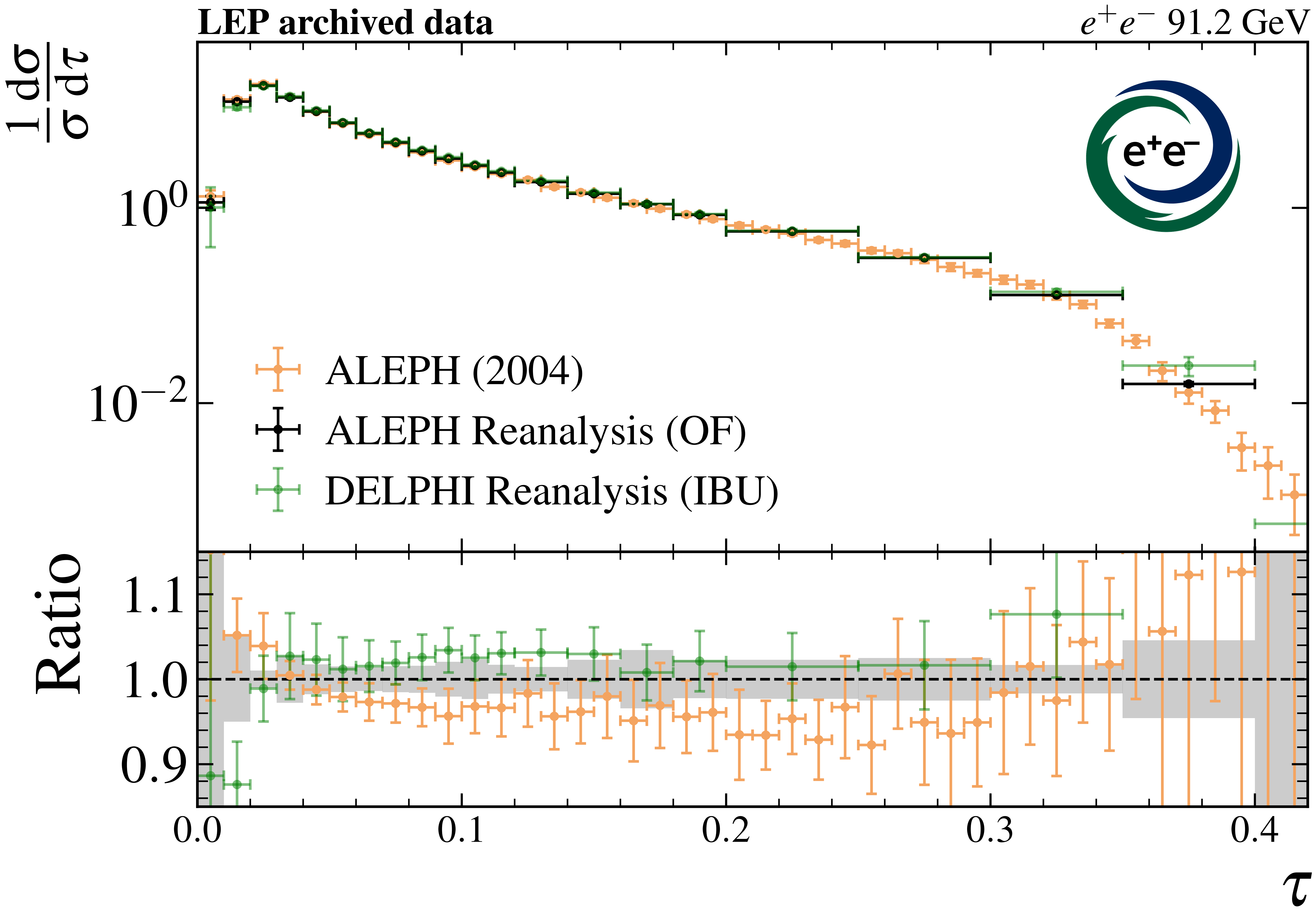
How do the ALEPH and DELPHI measurements compare??

Thanks to Miguel Benitez and Iain Stewart, 2412.15164



Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

ALEPH vs DELPHI Reanalysis



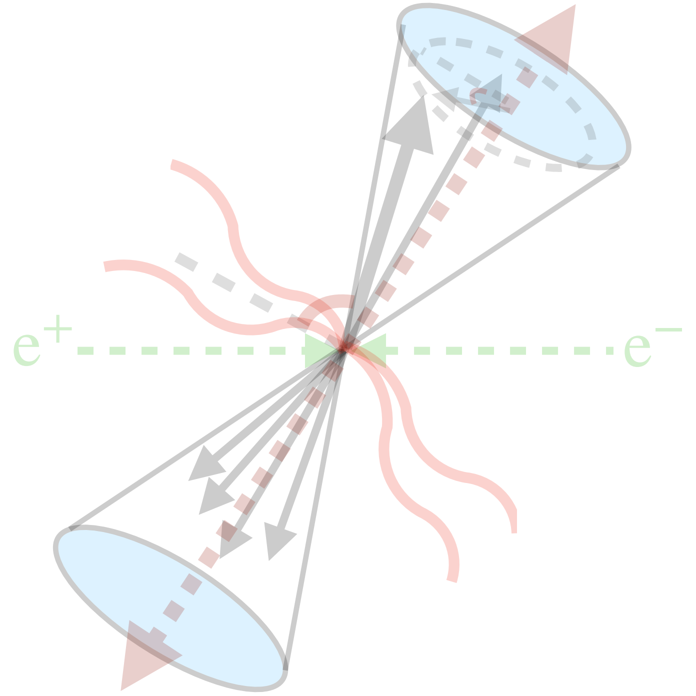
- Reanalyses of ALEPH/DELPHI archived data agree
- DELPHI shifted towards even broader τ spectrum
- Motivates new theory fits for α_S and Ω_1^R

First open e+e- data equivalent of modern day LHC comparisons (e.g. ATLAS vs CMS)! Opens door to many future directions across those datasets

Outline

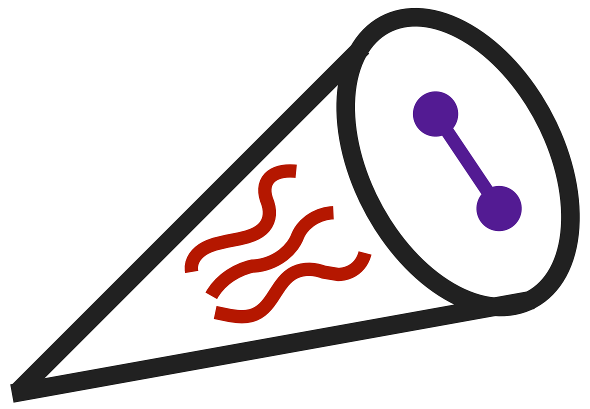
Physics Motivation (I)

Discovering new SM/BSM physics requires an unprecedented handle of QCD, motivating measurements across collision systems, regimes, and observables.



LEP Thrust Reanalyses (II)

Reanalysis of thrust in LEP archived data reveals new insights into α_S and P/NP QCD. Enables new studies of e+e- collisions with modern exp. and theory tools.



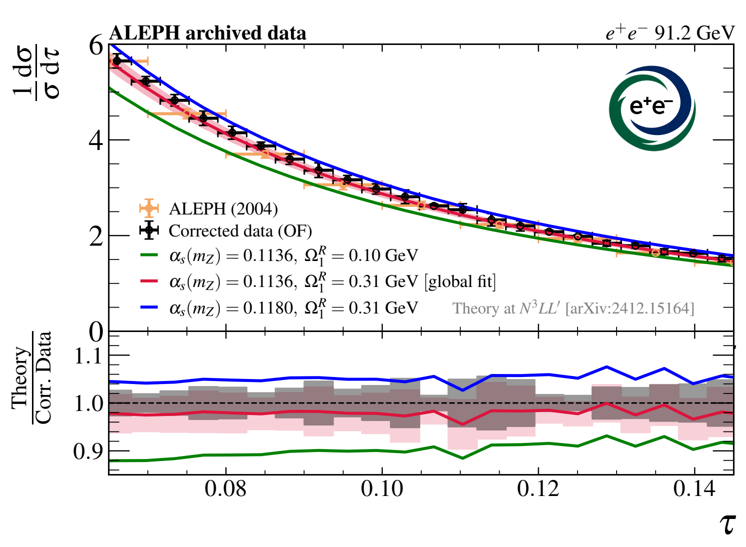
Ongoing Works and Outlook (III)

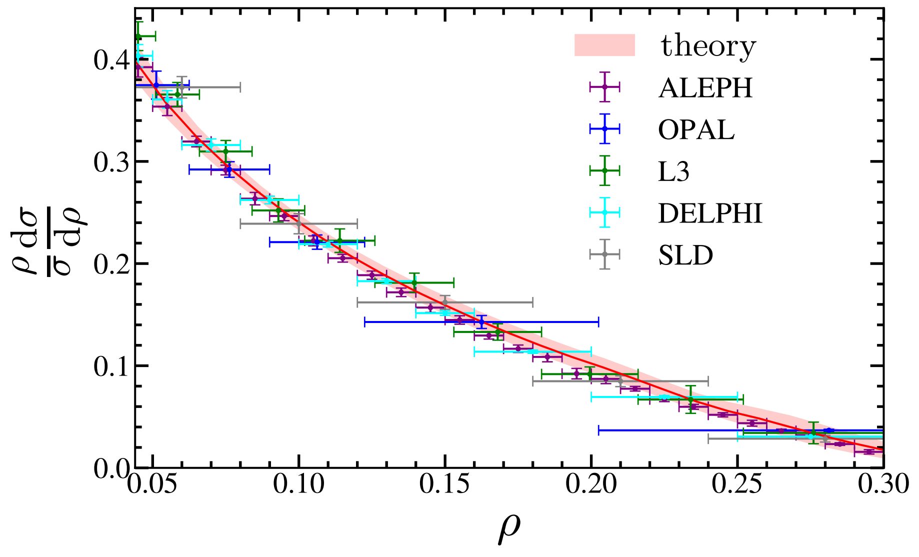
Continued investigations into thrust across relevant regimes. Complementary observables, studies across energies, application to constrain modern QCD.

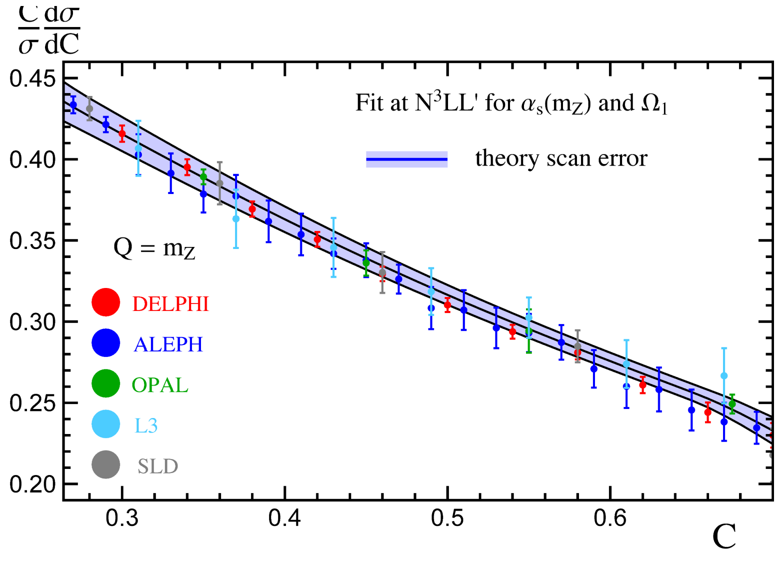
Natural Extension: 1D Thrust to Multi-Dimensional

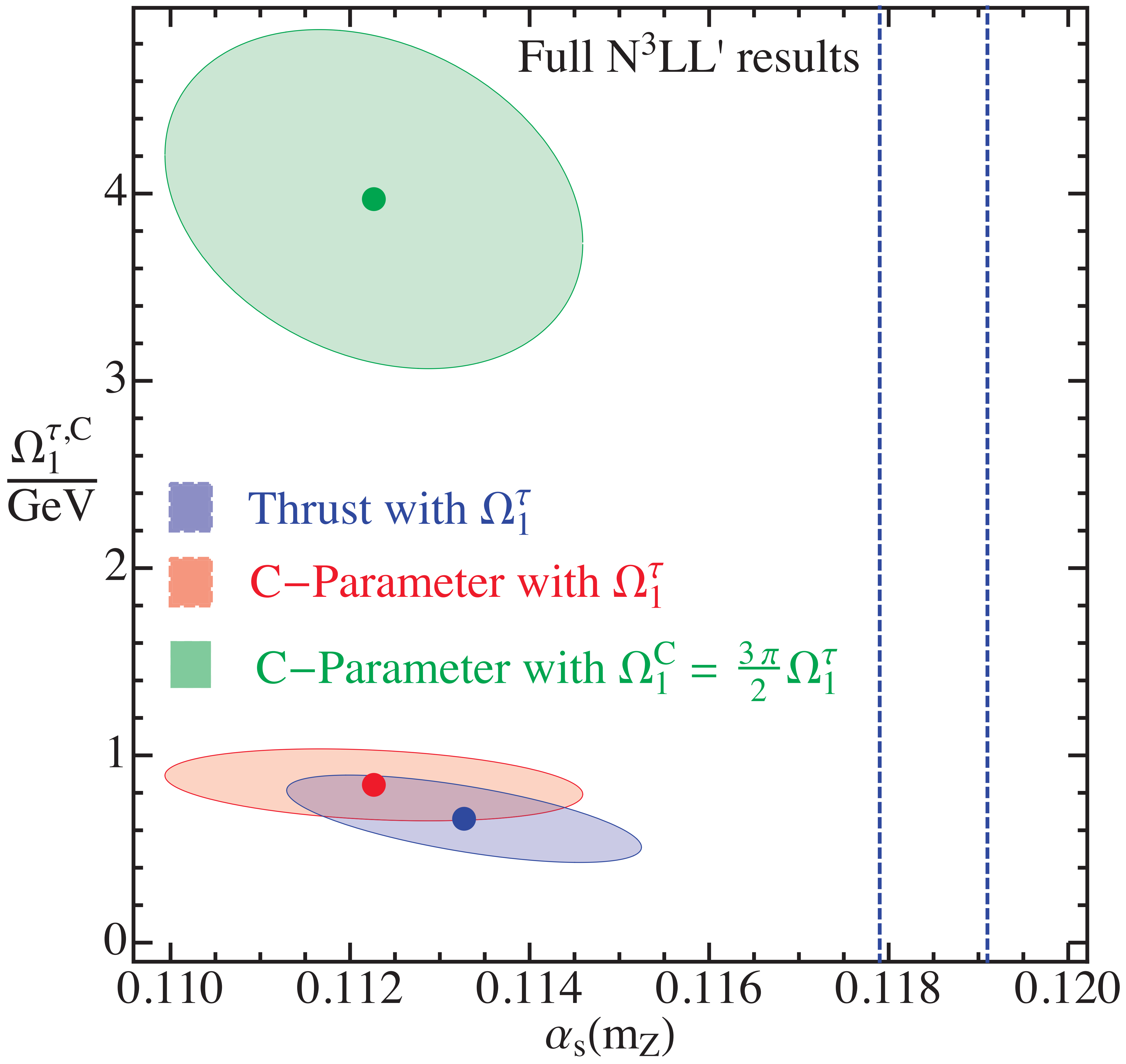
Major goal: multi-dimensional measurement of complementary event shape observables

Ex. Thrust, C-parameter, Heavy Jet Mass
- Leverage unbinned unfolding machinery
- Test universality of NP shape functions
- Test resummation effects on α_S
- Constrain across P/NP regimes for MC







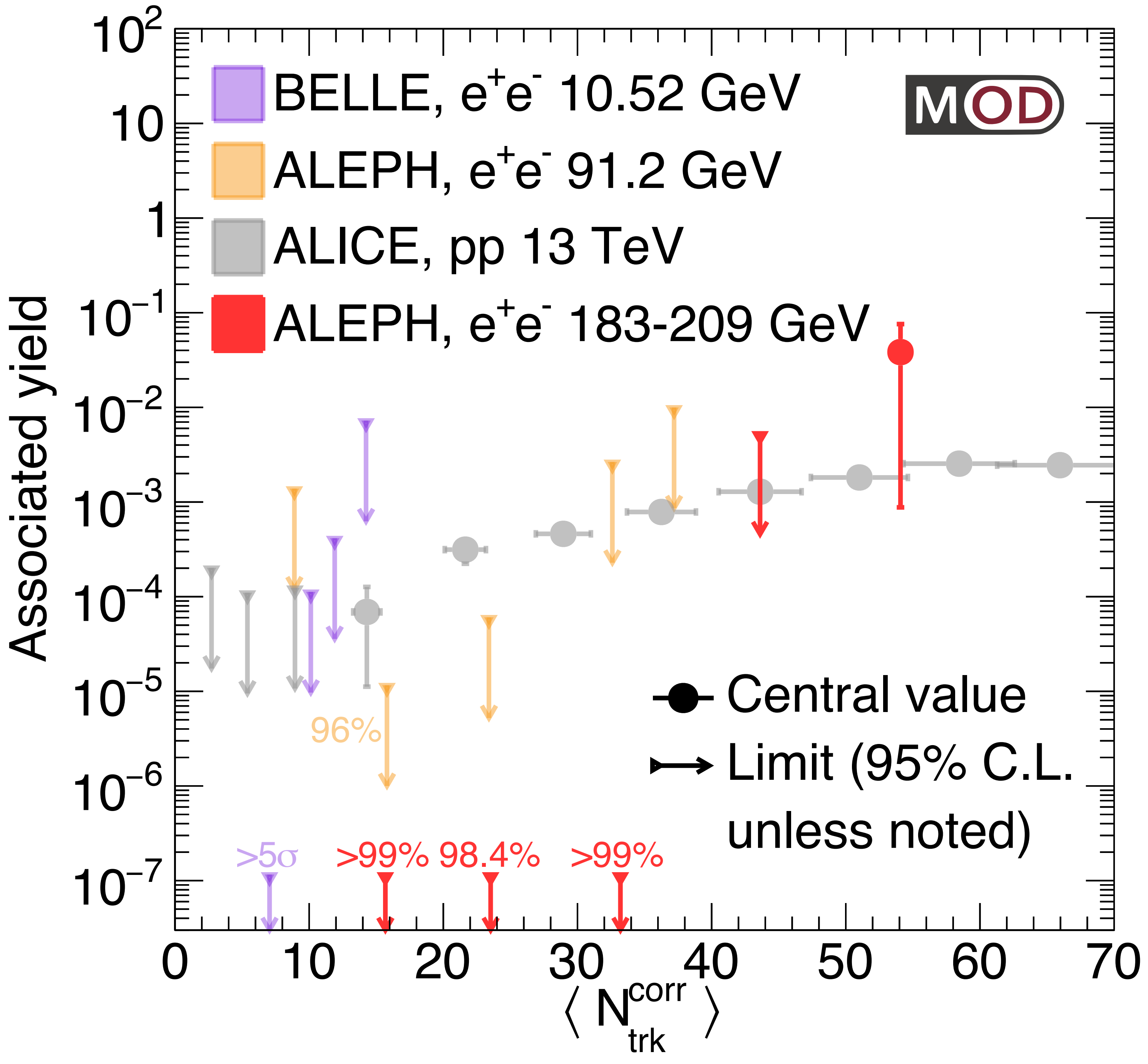


PRD 91 (2015), 094018, JHEP 07 (2025) 249, arXiv:2502.12253

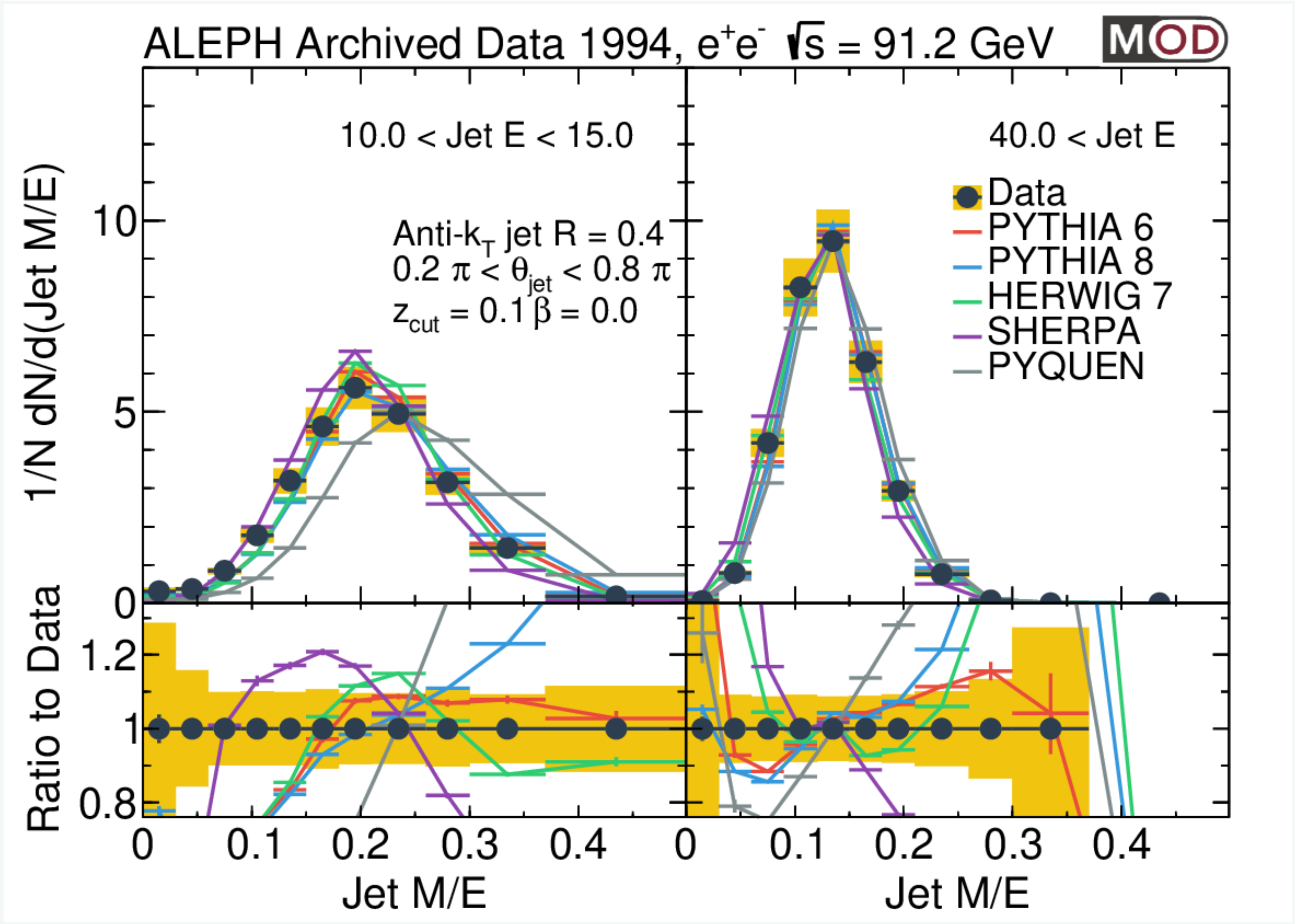
Many Other Exciting Areas of Research

Many other interesting observables, great for theory/exp. collaboration!
Involving many great young scientists!

Two-Particle Correlations

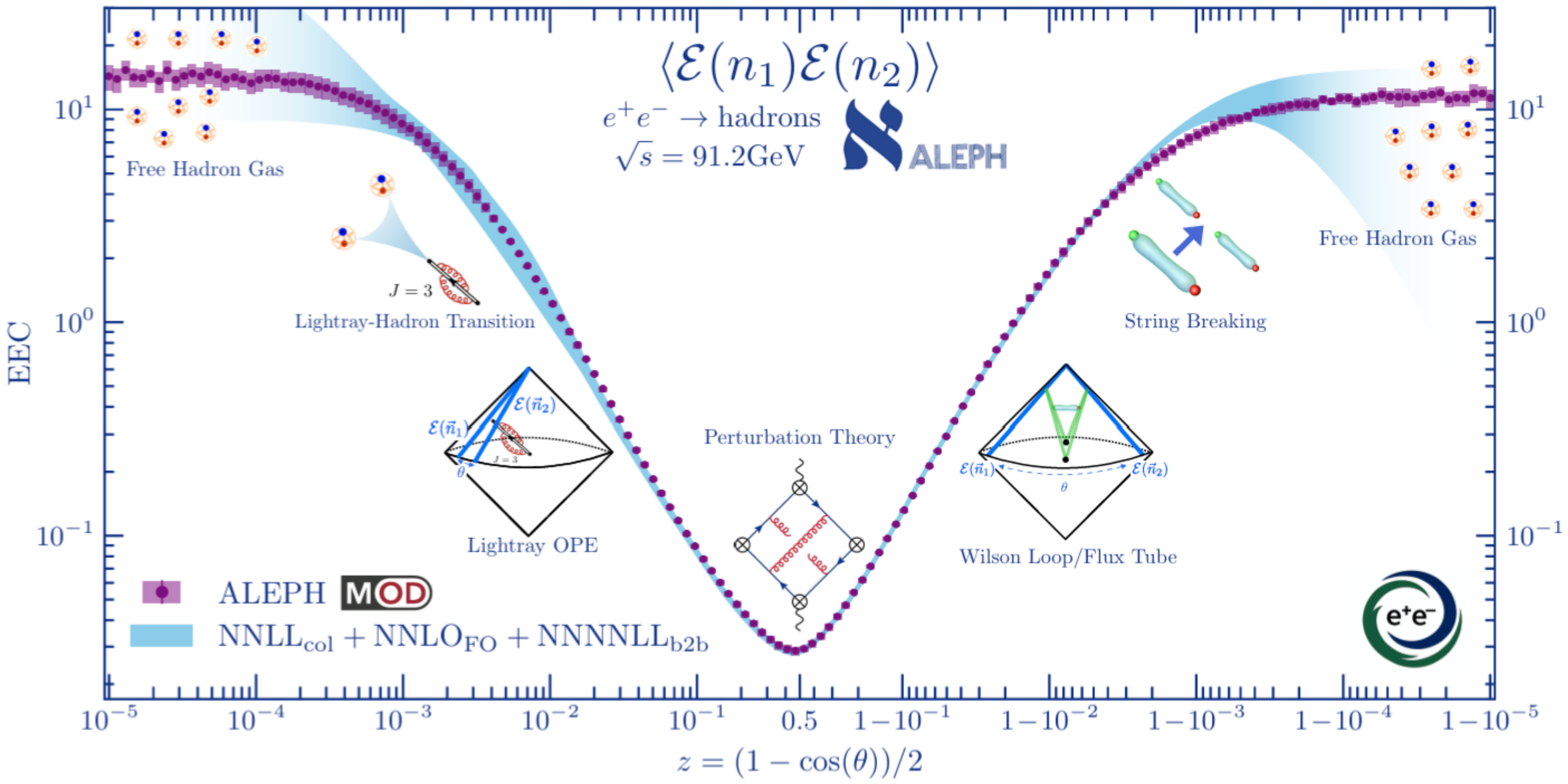


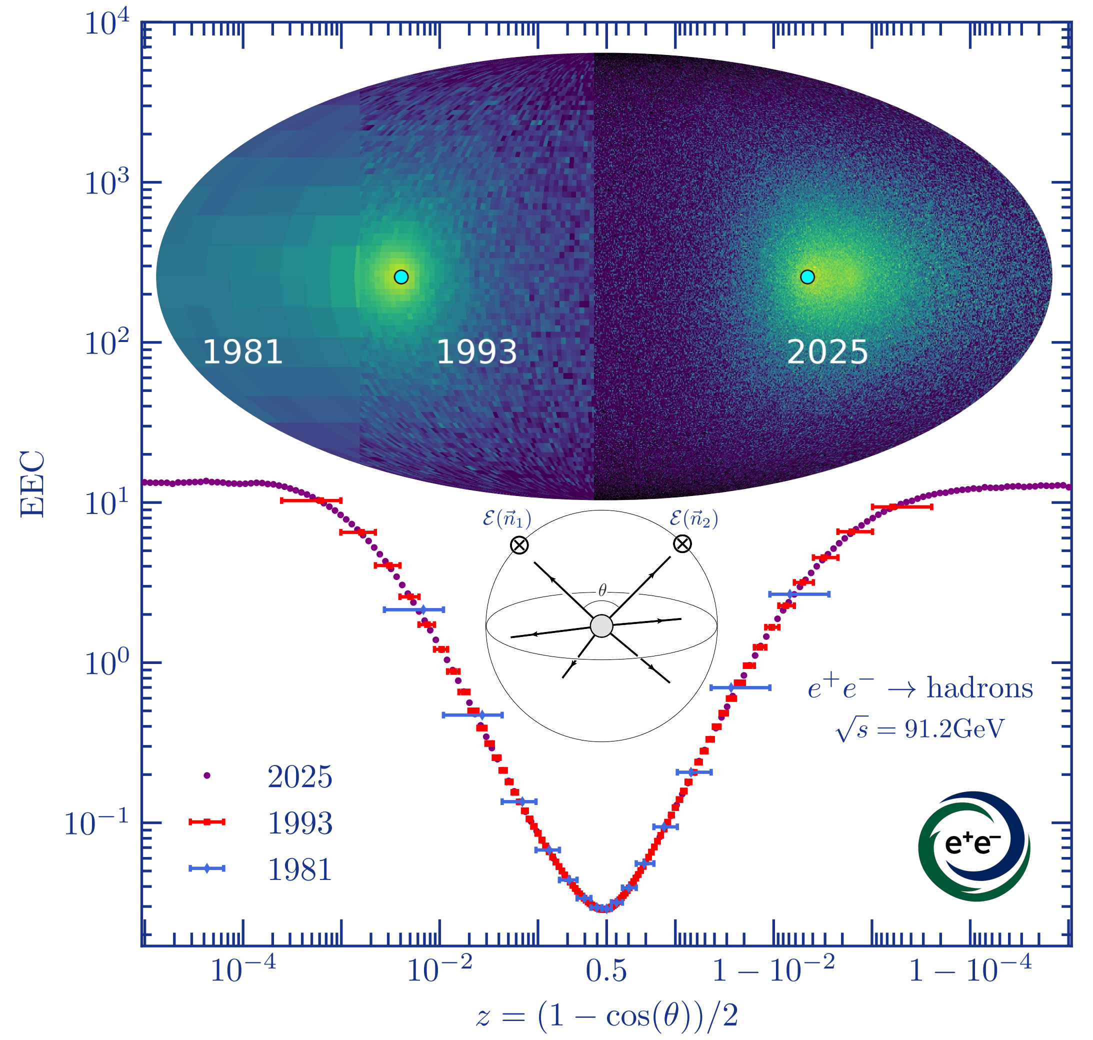
Jet Physics

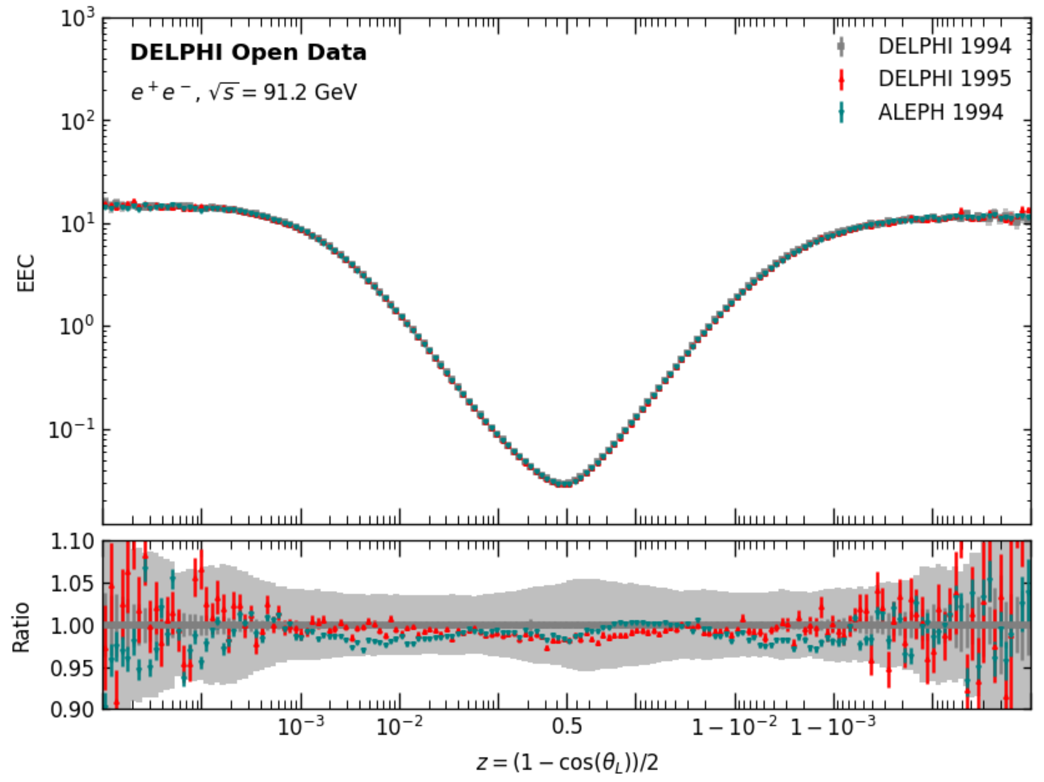


Yu-Chen "Janice" Chen, Tzu-An Sheng
Hannah Bossi, Jingyu Zhang, Chris McGinn,
Austin Baty, Luna (Yi) Chen, Yen-Jie Lee

One Highlight: Track Energy-Energy Correlator







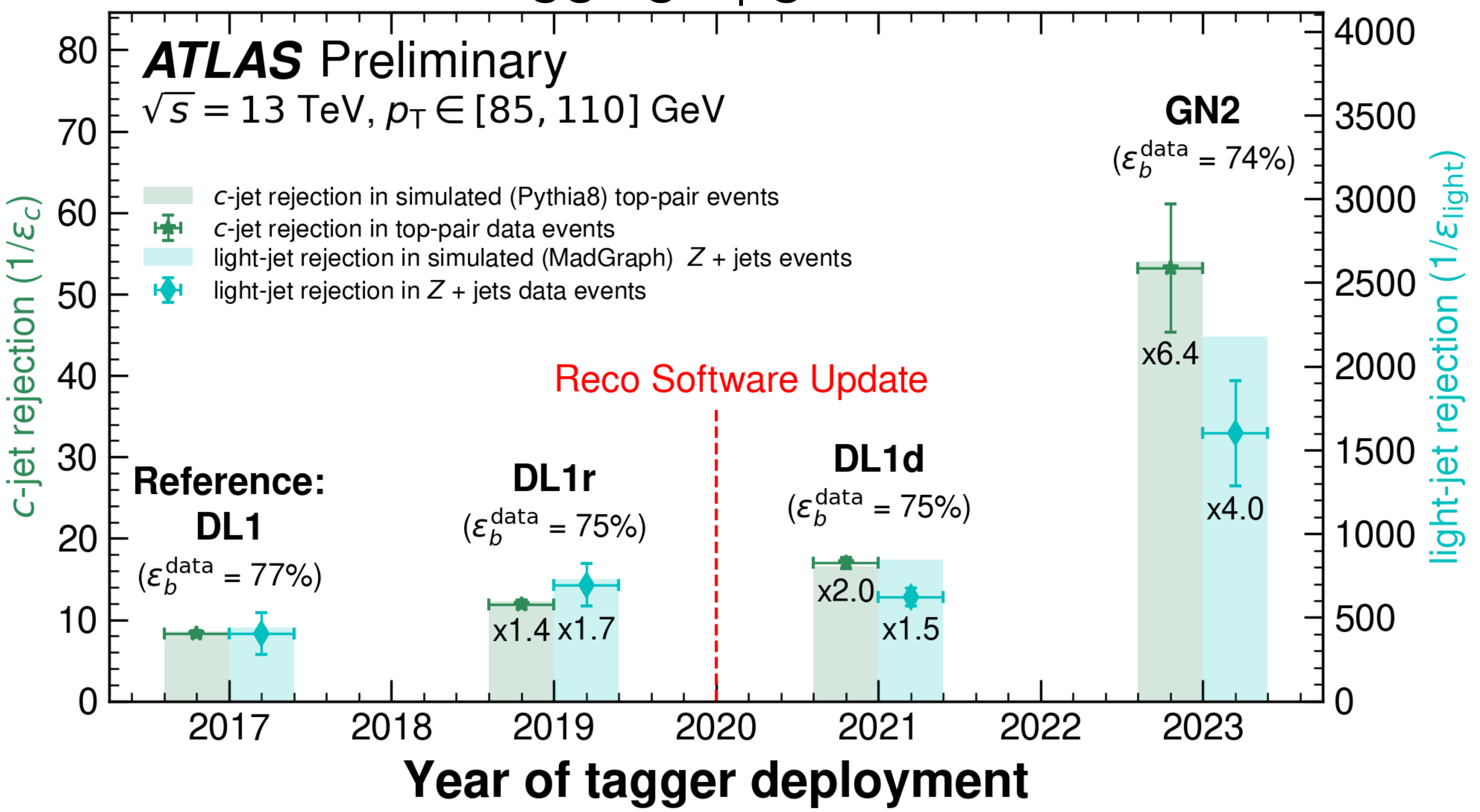
Th: Max Jaarsma, Yibei Li, Ian Moult, Wouter Waalewijn, HuaXing Zhu
Exp: Hannah Bossi, Jingyu Zhang, Luna (Yi) Chen, Yen-Jie Lee, e+e- All.

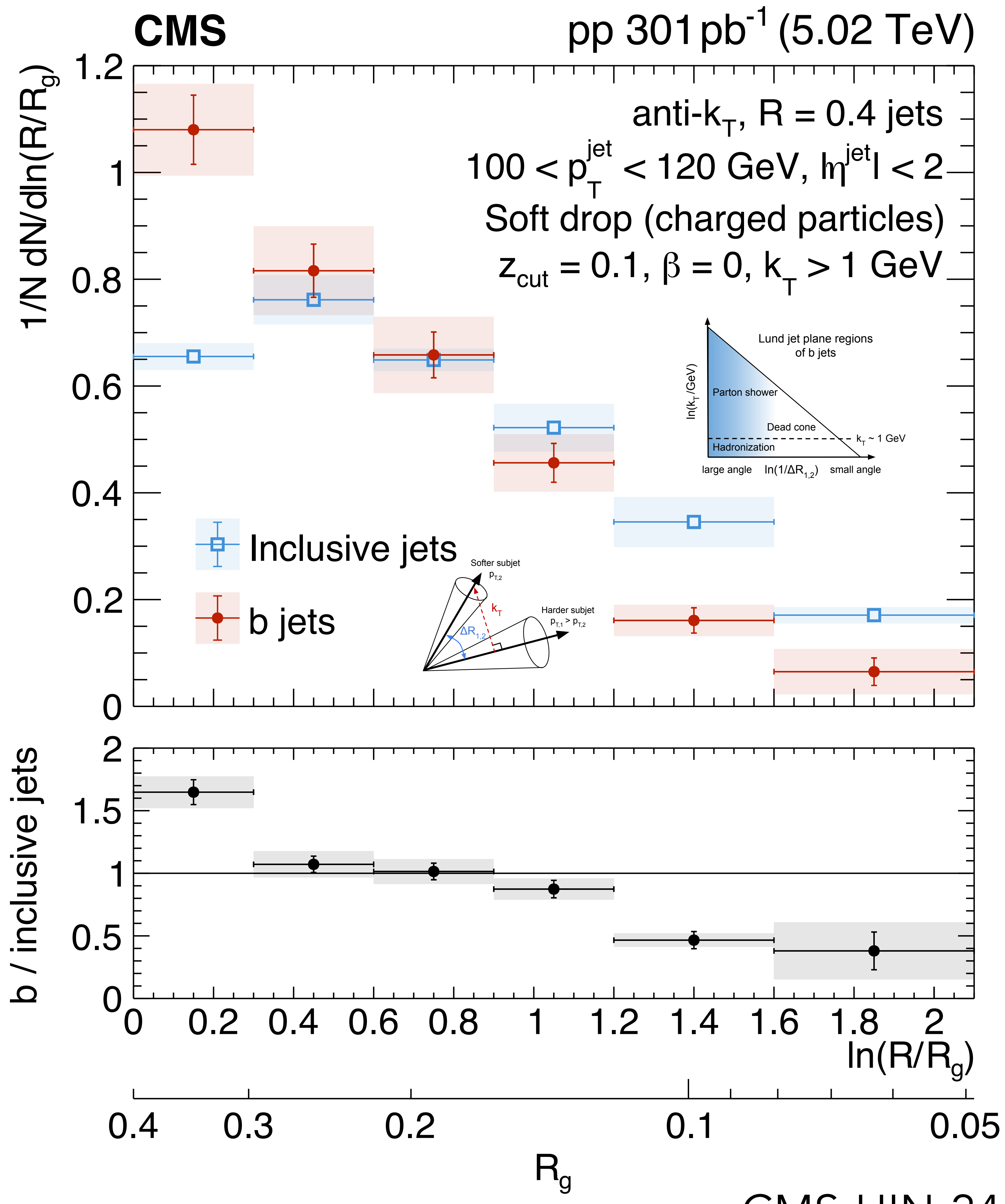
Probe HF Quark Mass Effects

Apply modern HF tagging (ML) to the LEP data to constrain mass effects on parton evolution

Requires work on data curation and algorithms

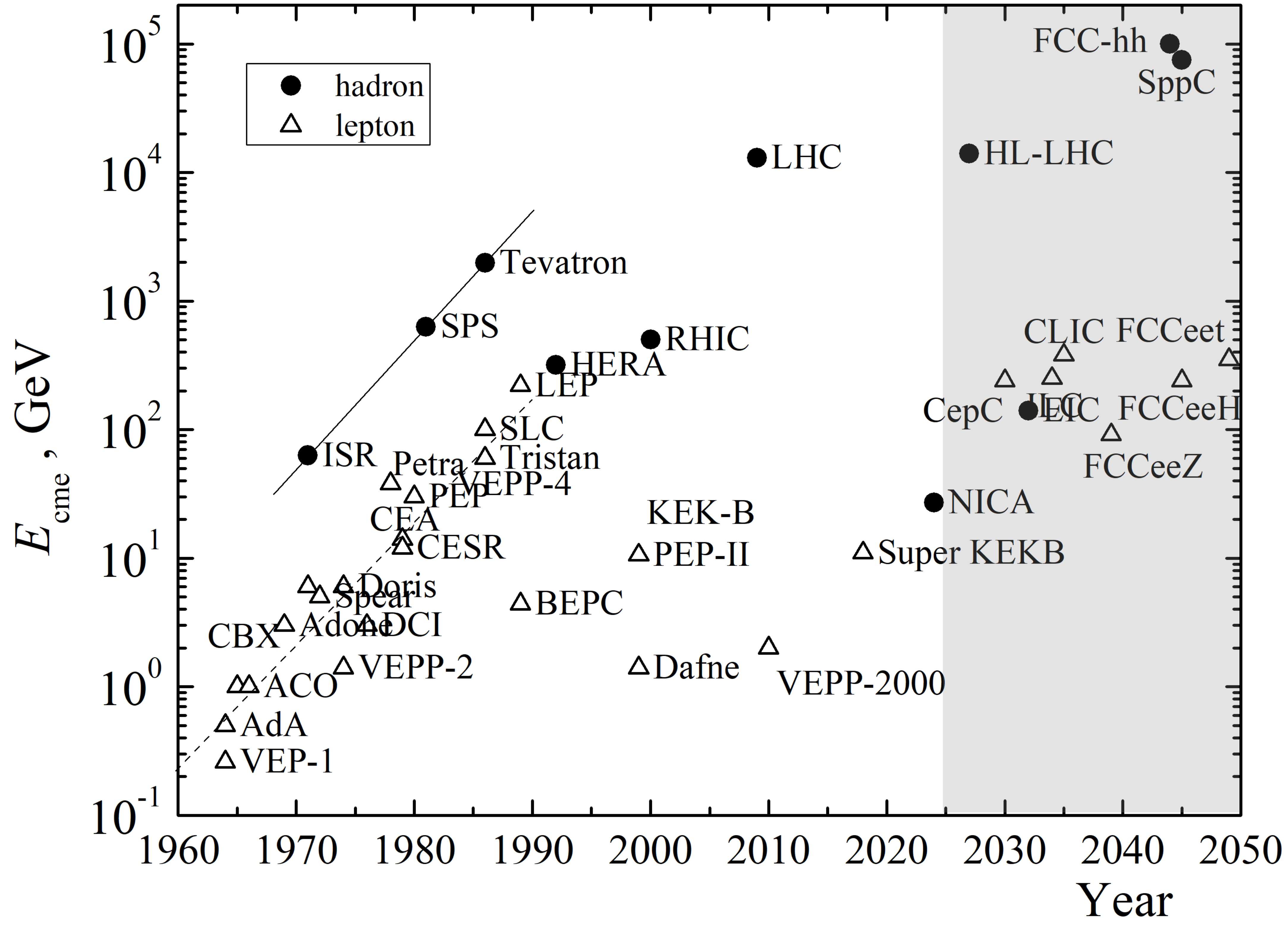
Ex. ATLAS HF tagging Upgraded with AI





CMS-HIN-24-005

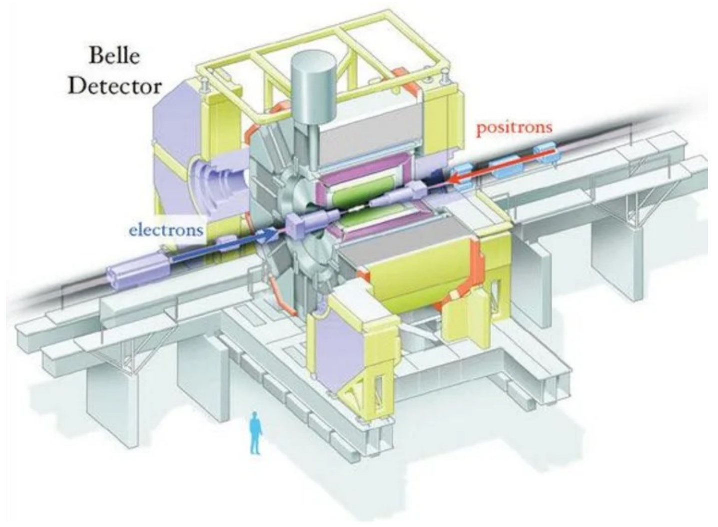
Birth of a Virtual e+e- Experiment Across Energies

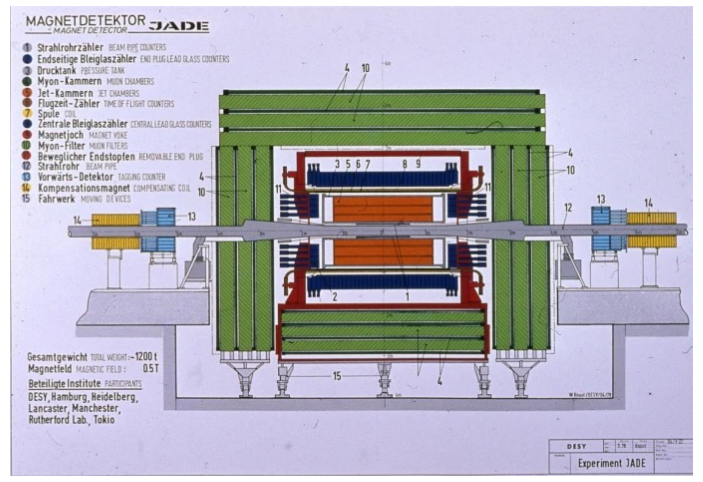


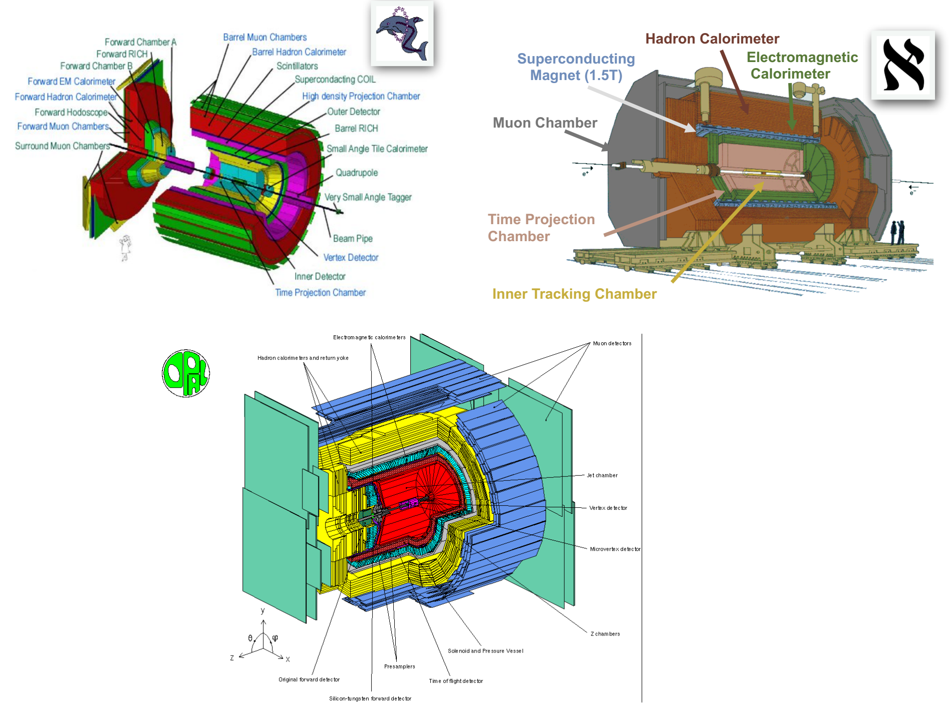
LEP: ALEPH, DELPHI, OPAL

KEK: BELLE 1

PETRA: JADE





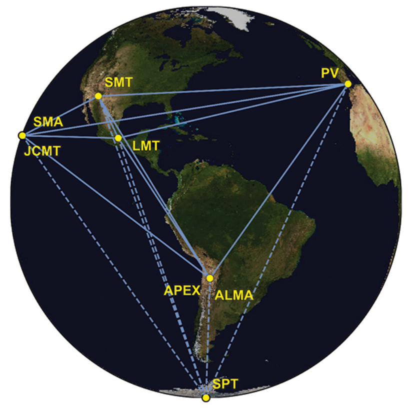


← 11 GeV (1999-2010)	14 - 45 GeV (1979-1986)	(←) 91 - 200 GeV (1989-2000)

→ $\sqrt{s}$

Possible to study from few to 200 GeV with e+e- archive data

Requires technical data curation to enable the reanalysis

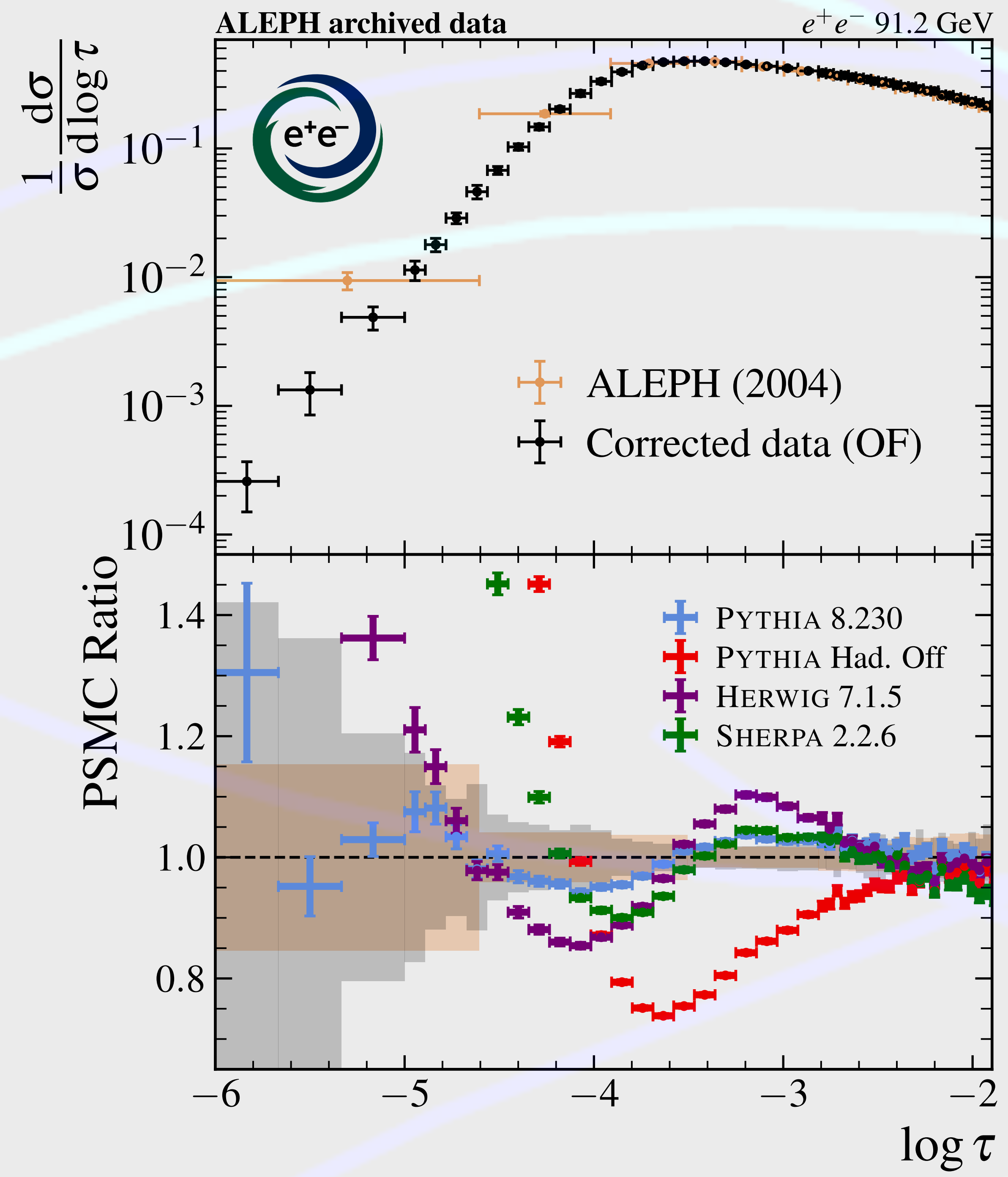


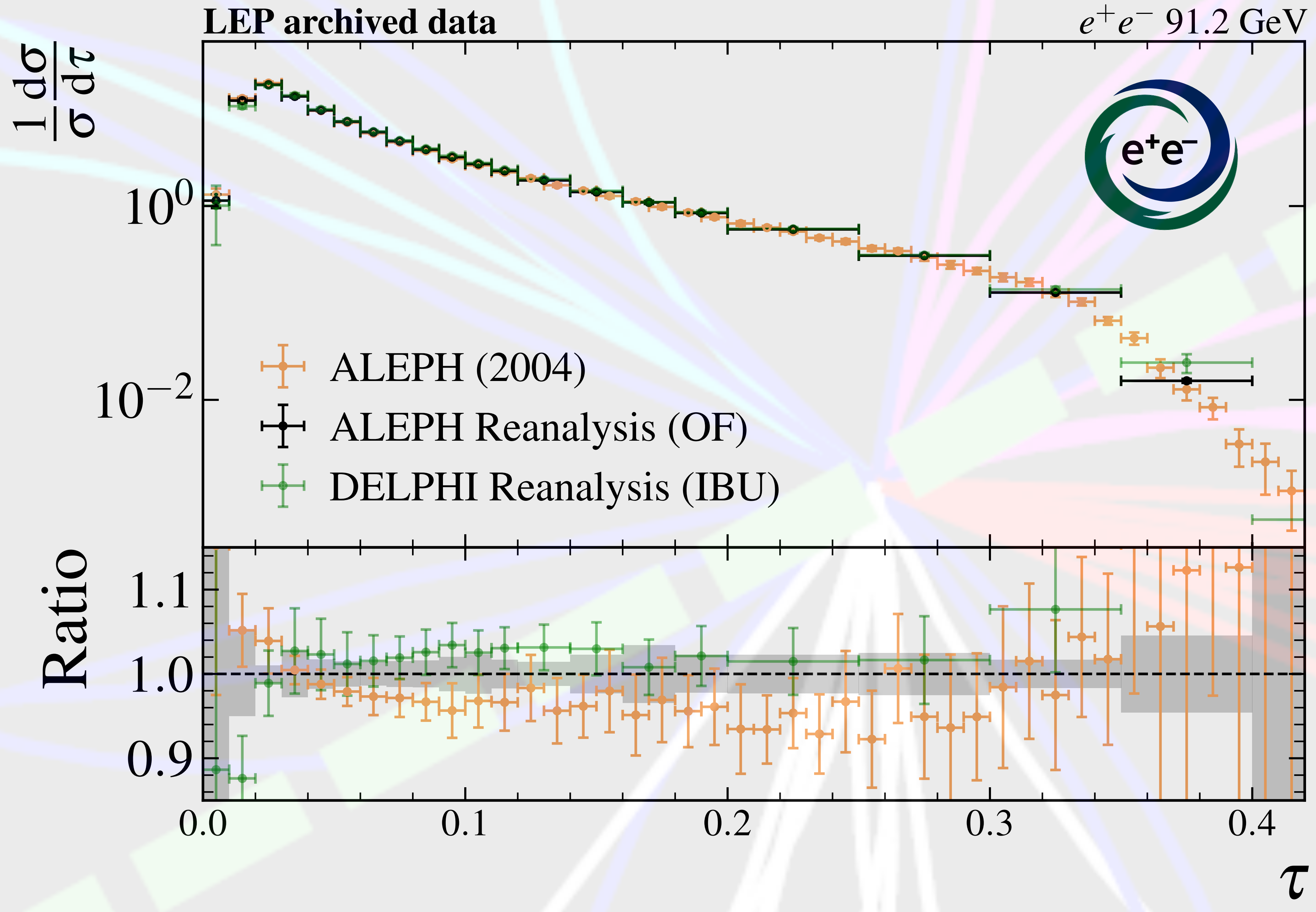
~ Like the *Virtual* Event Horizon Telescope?

Summary

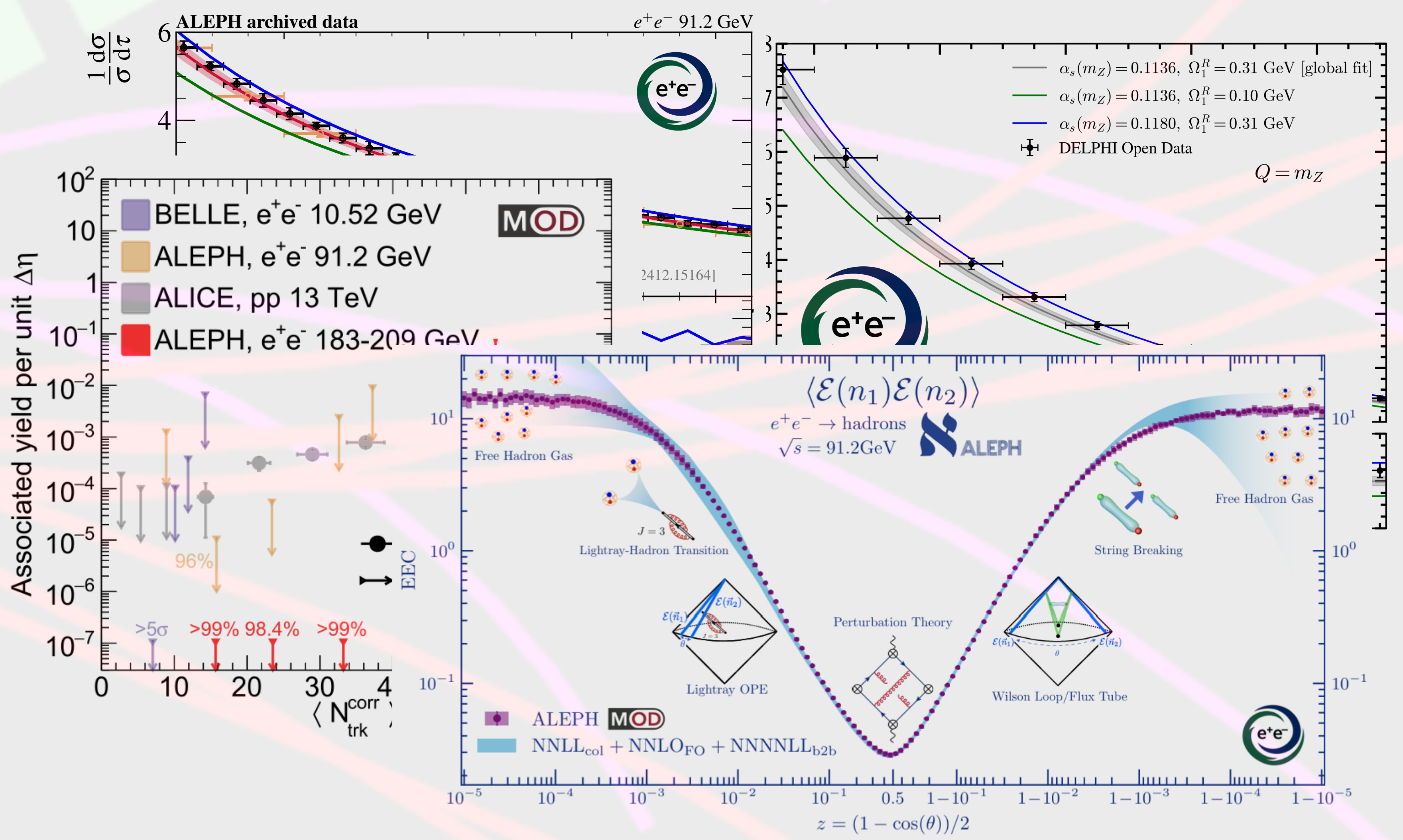
Active community to curate and reanalyze archived e+e- datasets

- New insights with modern experimental and theory tools
- Establishing a virtual e+e- experiment program across energies
- Connections to future facilities (HL-LHC, FCC, EIC, …)





Many exciting directions!

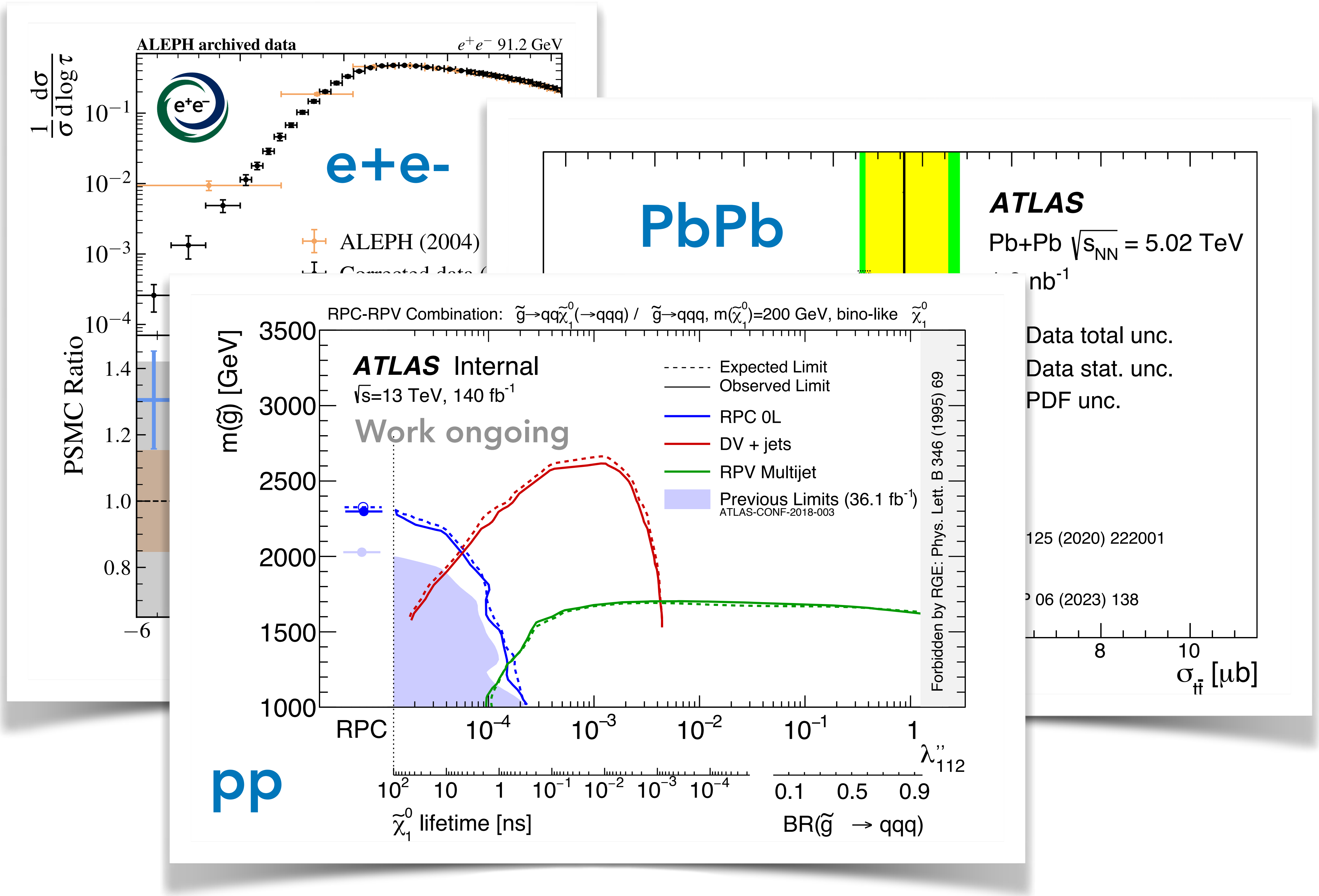


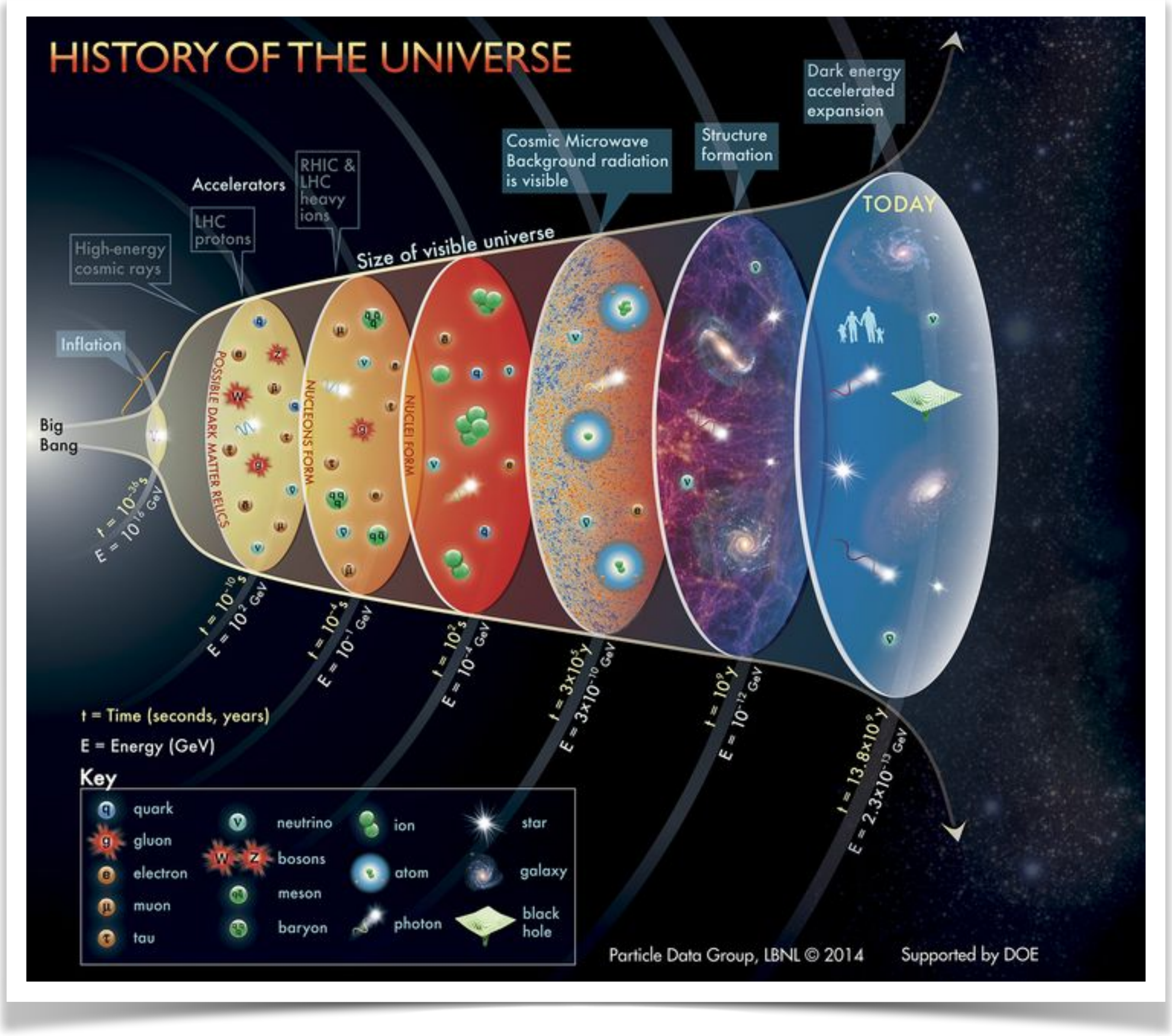
BACKUP

Misc.

Examples Across My Research

Hadronic processes key to understanding the baryonic evolution of universe from early to current times



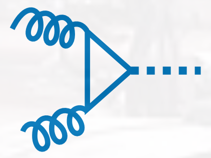


e.g. partonic hadronization, in-medium behavior, and cosmological baryon asymmetry of early universe

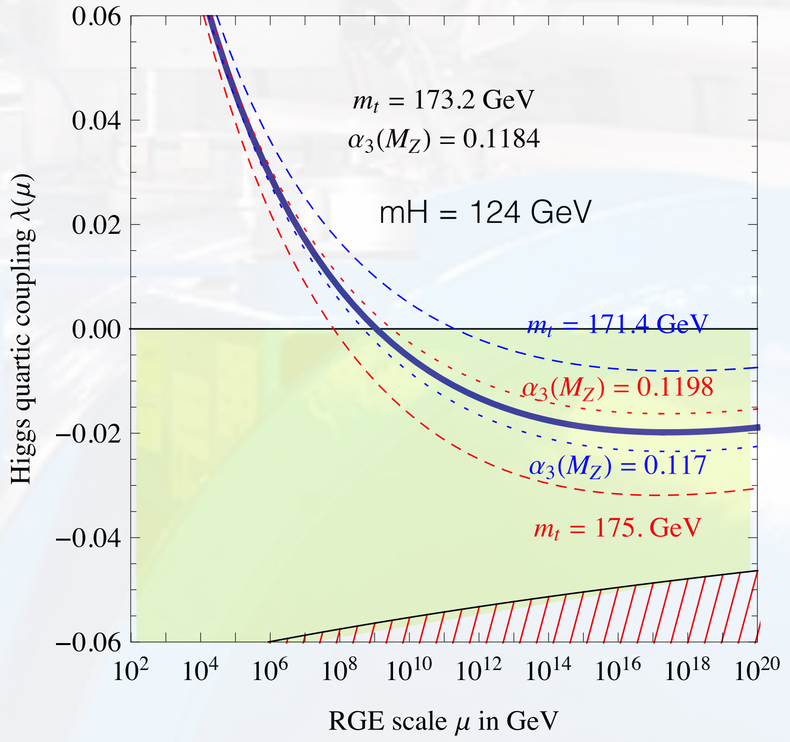
Examples Across the Community

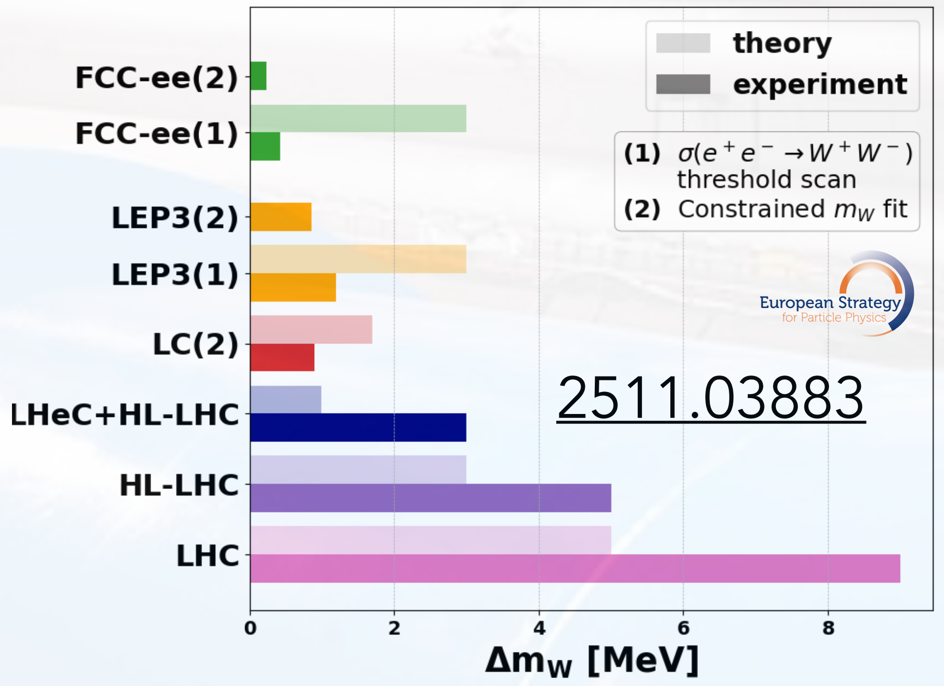
Precise handle on hadronic processes is key to global goal of discovering new physics across HEP, nuclear, astrophysics, cosmology, AI/ML, future programs …

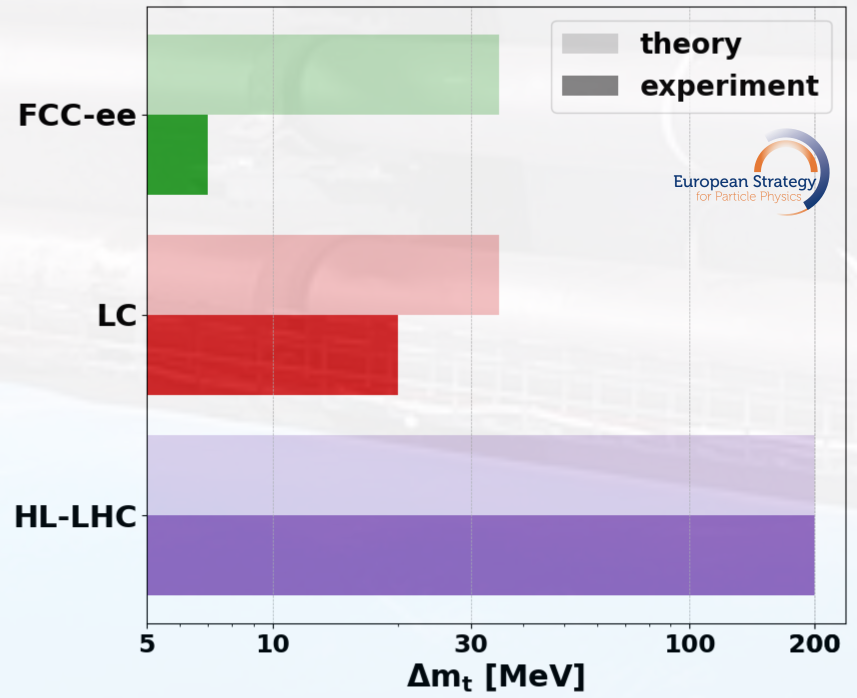
ggF Higgs cross sections, 1602.00695, 1112.3022

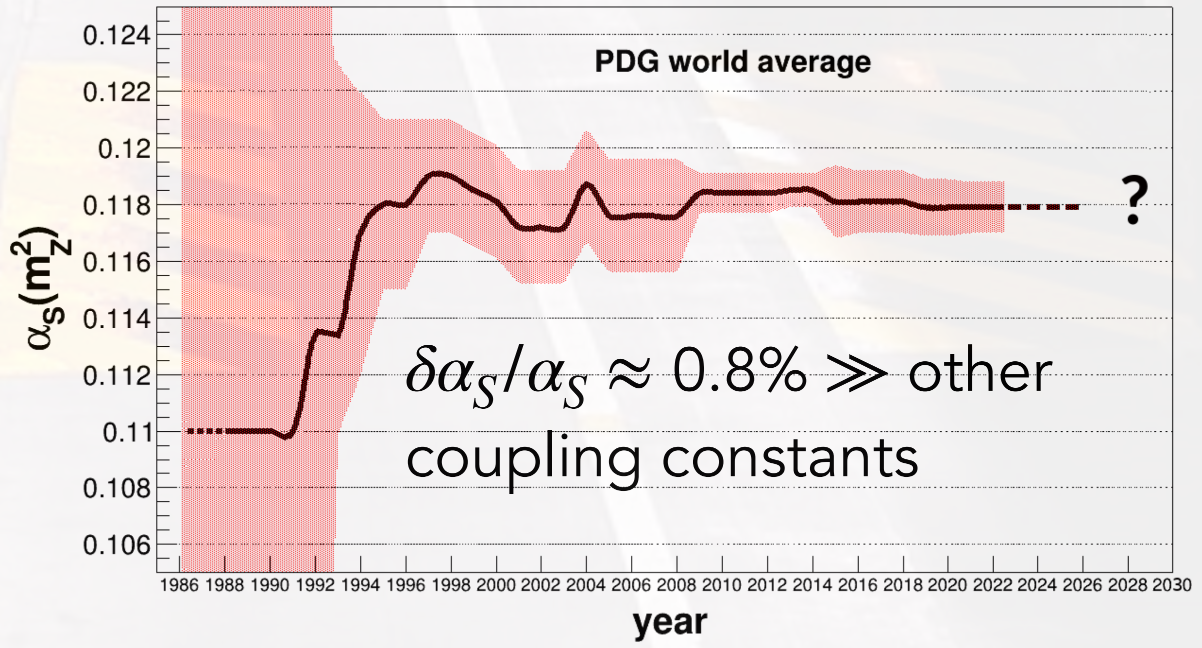


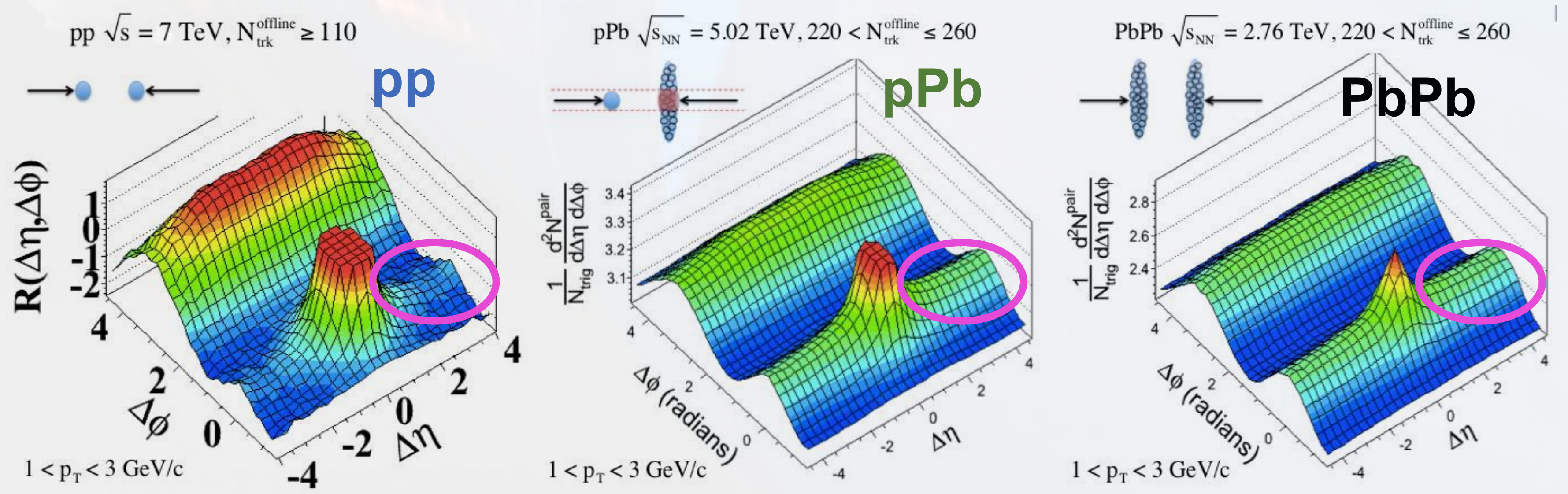
E_{CM}	σ	δ(theory)	δ(PDF)	$\delta(\alpha_s)$
2 TeV	1.10 pb	$^{+0.04\text{pb}}_{-0.09\text{pb}}\left(^{+4.06\%}_{-7.88\%}\right)$	$\pm$ 0.03 pb ($\pm$ 3.17%)	$^{+0.04\text{pb}}_{-0.04\text{pb}}\left(^{+3.36\%}_{-3.69\%}\right)$
7 TeV	16.85 pb	$^{+0.74\text{pb}}_{-1.17\text{pb}}\left(^{+4.41\%}_{-6.96\%}\right)$	$\pm$ 0.32 pb ($\pm$ 1.89%)	$^{+0.45\text{pb}}_{-0.45\text{pb}}\left(^{+2.67\%}_{-2.66\%}\right)$
8 TeV	21.42 pb	$^{+0.95\text{pb}}_{-1.48\text{pb}}\left(^{+4.43\%}_{-6.90\%}\right)$	$\pm$ 0.40 pb ($\pm$ 1.87%)	$^{+0.57\text{pb}}_{-0.56\text{pb}}\left(^{+2.65\%}_{-2.62\%}\right)$
13 TeV	48.58 pb	$^{+2.22\text{pb}}_{-3.27\text{pb}}\left(^{+4.56\%}_{-6.72\%}\right)$	$\pm$ 0.90 pb ($\pm$ 1.86%)	$^{+1.27\text{pb}}_{-1.25\text{pb}}\left(^{+2.61\%}_{-2.58\%}\right)$
14 TeV	54.67 pb	$^{+2.51\text{ pb}}_{-3.67\text{ pb}}\left(^{+4.58\%}_{-6.71\%}\right)$	$\pm$1.02 pb ($\pm$ 1.86%)	$^{+1.43\text{pb}}_{-1.41\text{pb}}\left(^{+2.61\%}_{-2.59\%}\right)$

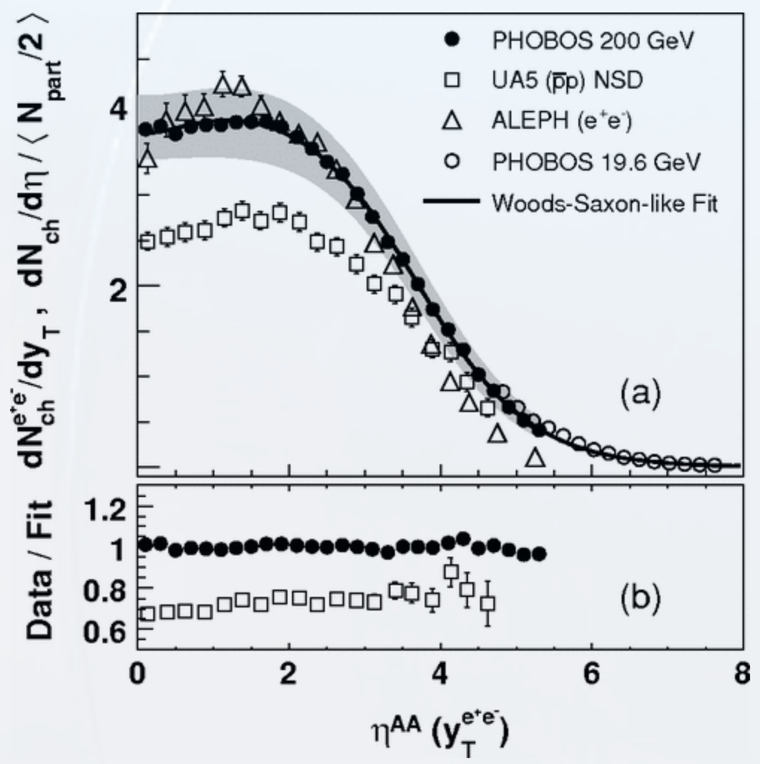


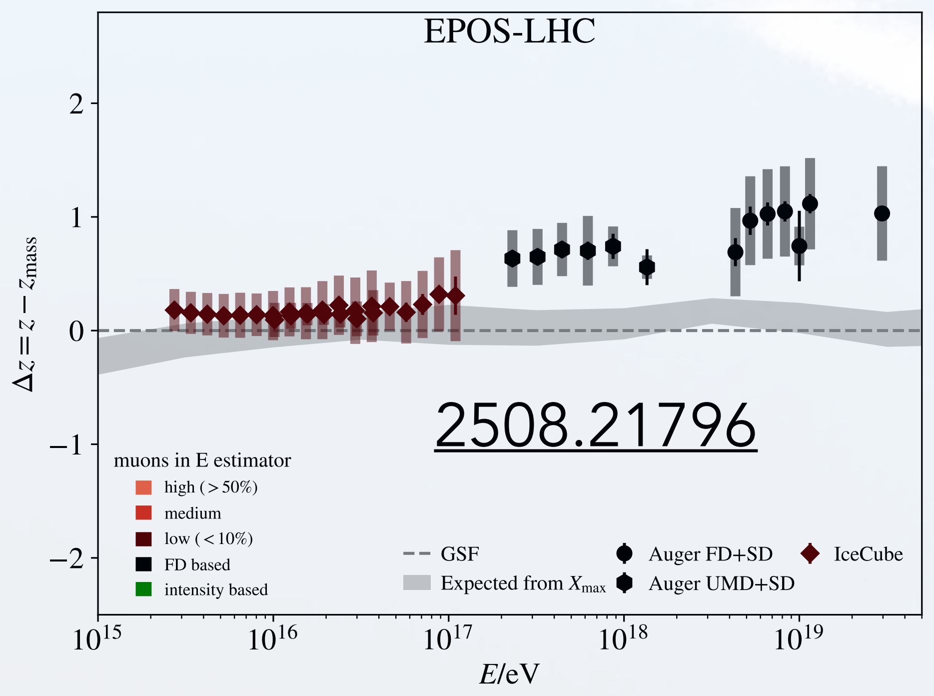




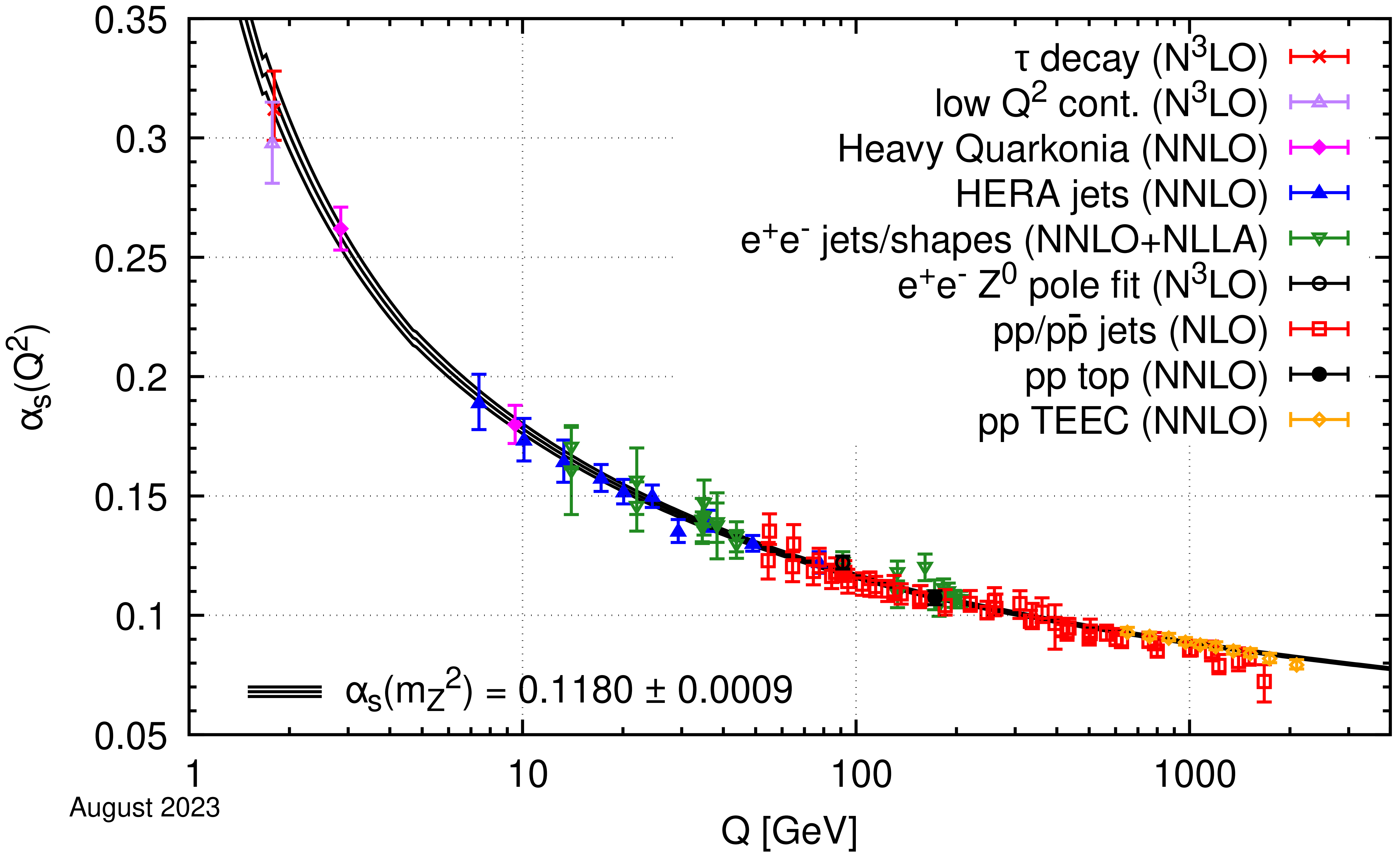




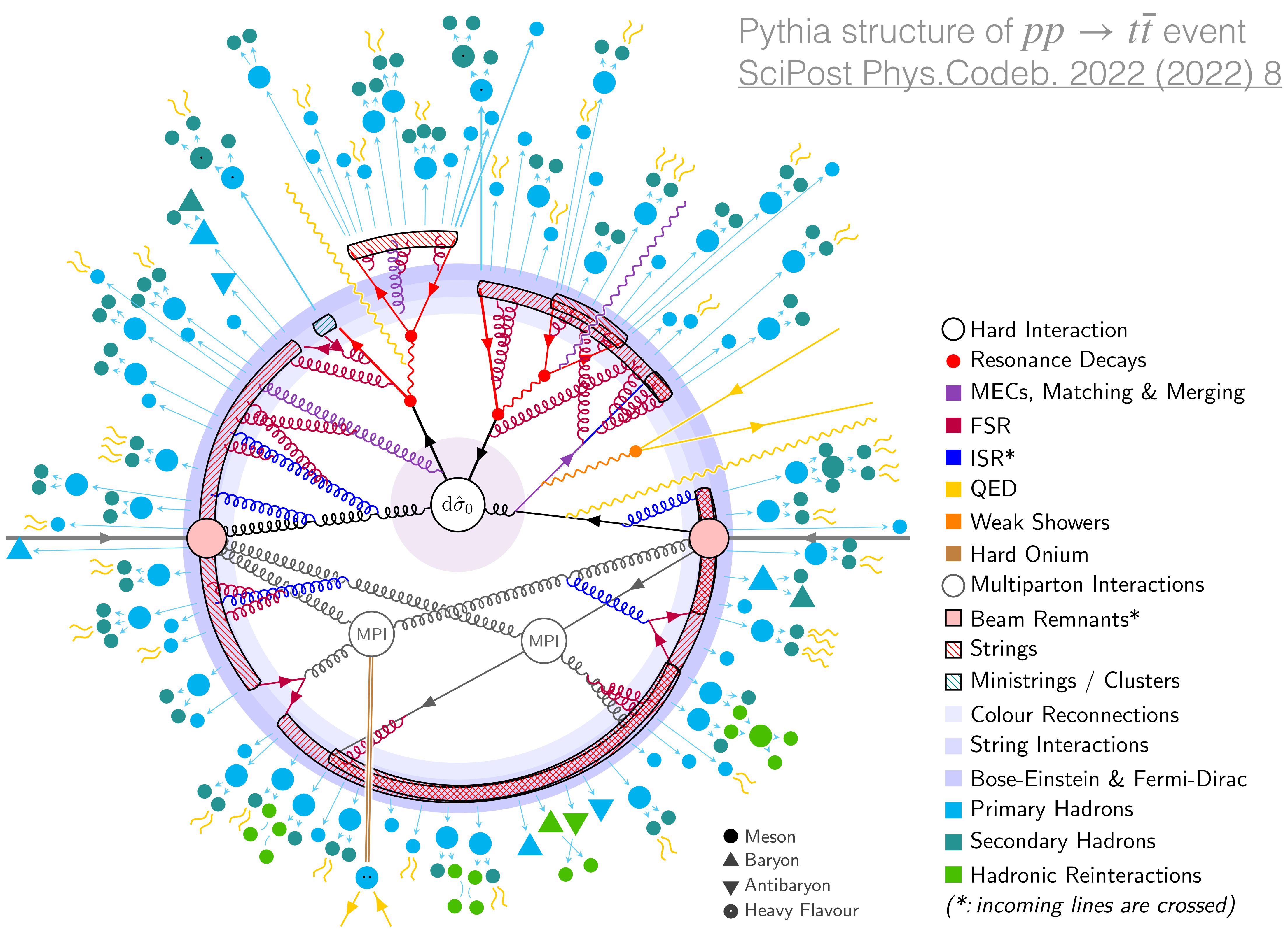




Must Achieve Precise Handle of QCD



Pythia structure of $pp \to t\bar{t}$ event
SciPost Phys.Codeb. 2022 (2022) 8



Unprecedented handle on Quantum Chromodynamics (QCD): Strong Coupling Constant (α_S), Perturbative (P) and Non-Perturbative (NP) regimes, Heavy flavour (HF) quark mass effects, …

Interesting History

PHYSICAL REVIEW LETTERS

VOLUME 39 | 14 NOVEMBER 1977 | NUMBER 20

Simple Quantum-Chromodynamics Prediction of Jet Structure in e^+e^- Annihilation

Howard Georgi and Marie Machacek
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 September 1977)

We propose a simple alternative to the sphericity as a measure of jet structure in e^+e^- annihilation. Our variable has the property that it can be reliably calculated in perturbation theory in quantum chromodynamics (QCD) for large Q^2. It is not sensitive to the details of quark and gluon decay into color-singlet hadrons. We discuss the nonperturbative effects which are important at moderate Q^2.

Farhi[9] has independently come to many of the same conclusions. We are grateful to him and to S. Weinberg and H. D. Politzer for valuable discussions. One of us (H.G.) wishes to thank the staff of the Aspen Center for Physics for their hospitality in Aspen and to acknowledge useful conversations at the Center with T. Appelquist, A. De Rújula (who invented the name *spherocity*), and G. Ross.

Phys. Rev. Lett. 39, 1237

PHYSICAL REVIEW LETTERS

VOLUME 39 | 19 DECEMBER 1977 | NUMBER 25

Quantum Chromodynamics Test for Jets

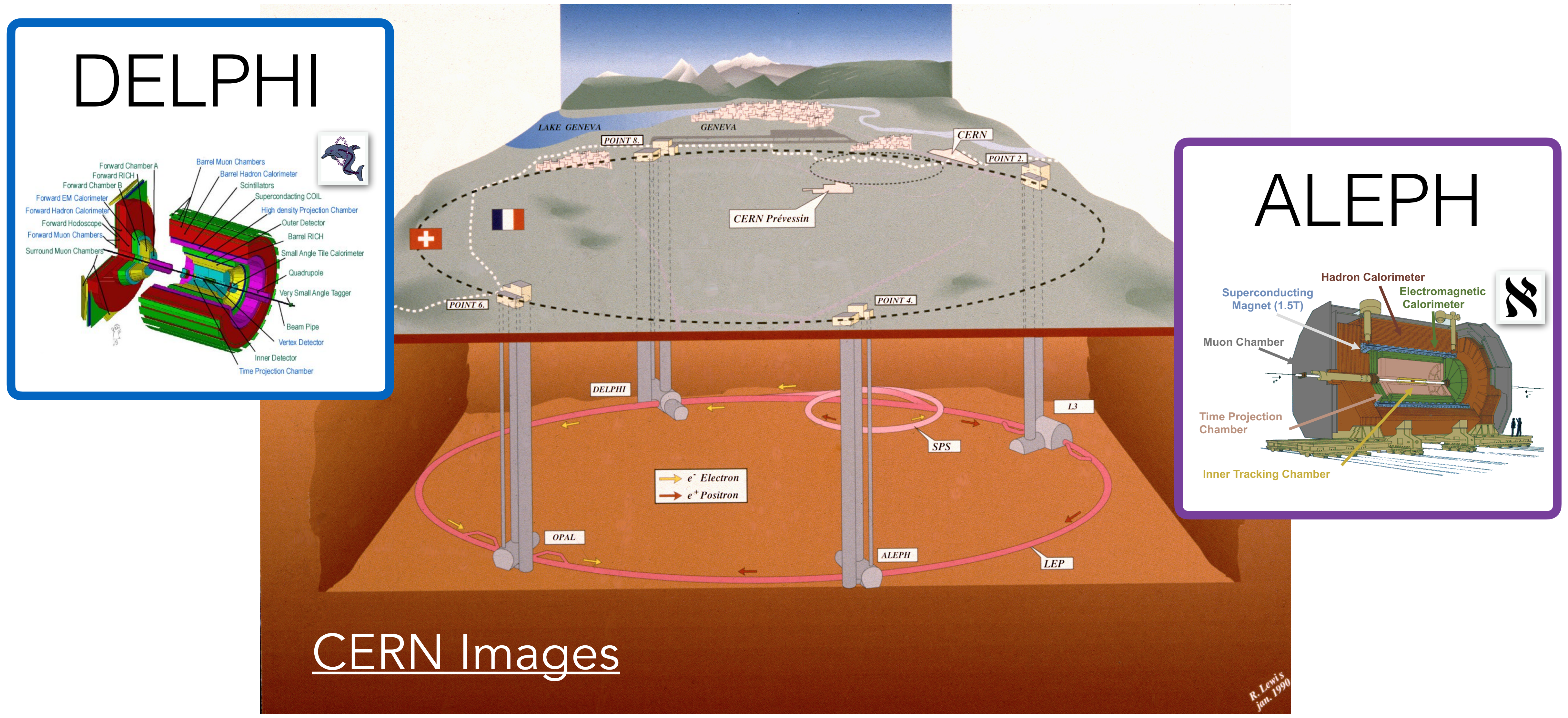
Edward Farhi
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 September 1977)

A new quantity, the maximum of the directed momentum, is proposed which measures jetlikeness in e^+e^- annihilation. This quantity is computed in the quark-gluon model by using renormalization-group improved perturbation theory.

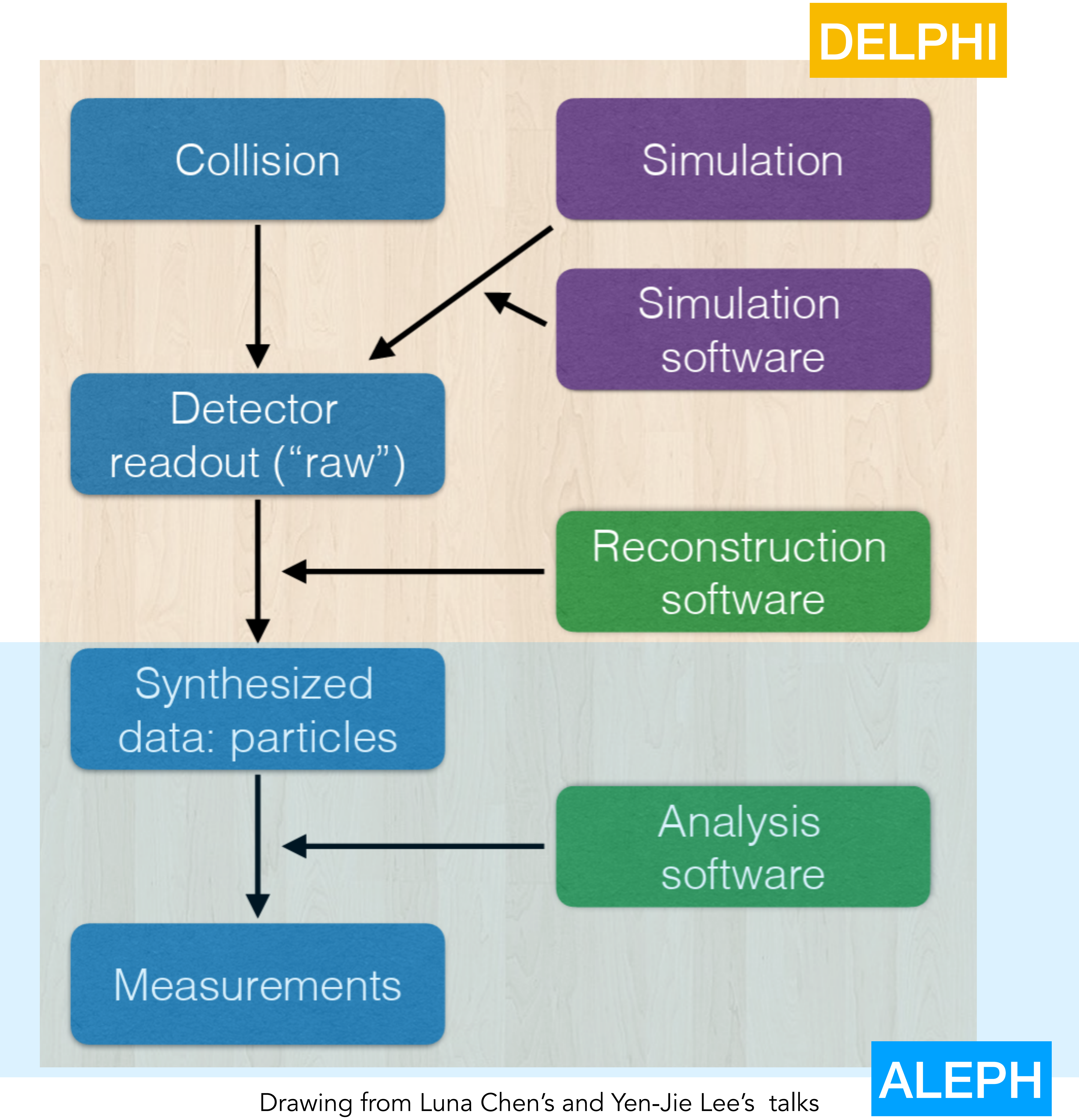
I would like to thank Orlando Alvarez and Steven Weinberg for their help. Howard Georgi and Marie Machacek[9] have independently done work similar to this and I would also like to thank them for valuable discussions. This research was supported in part by the National Science Foundation under Grant No. PHY75-20427.

Phys. Rev. Lett. 39, 1587

Summary: Reanalyses of e+e- LEP Archived Data



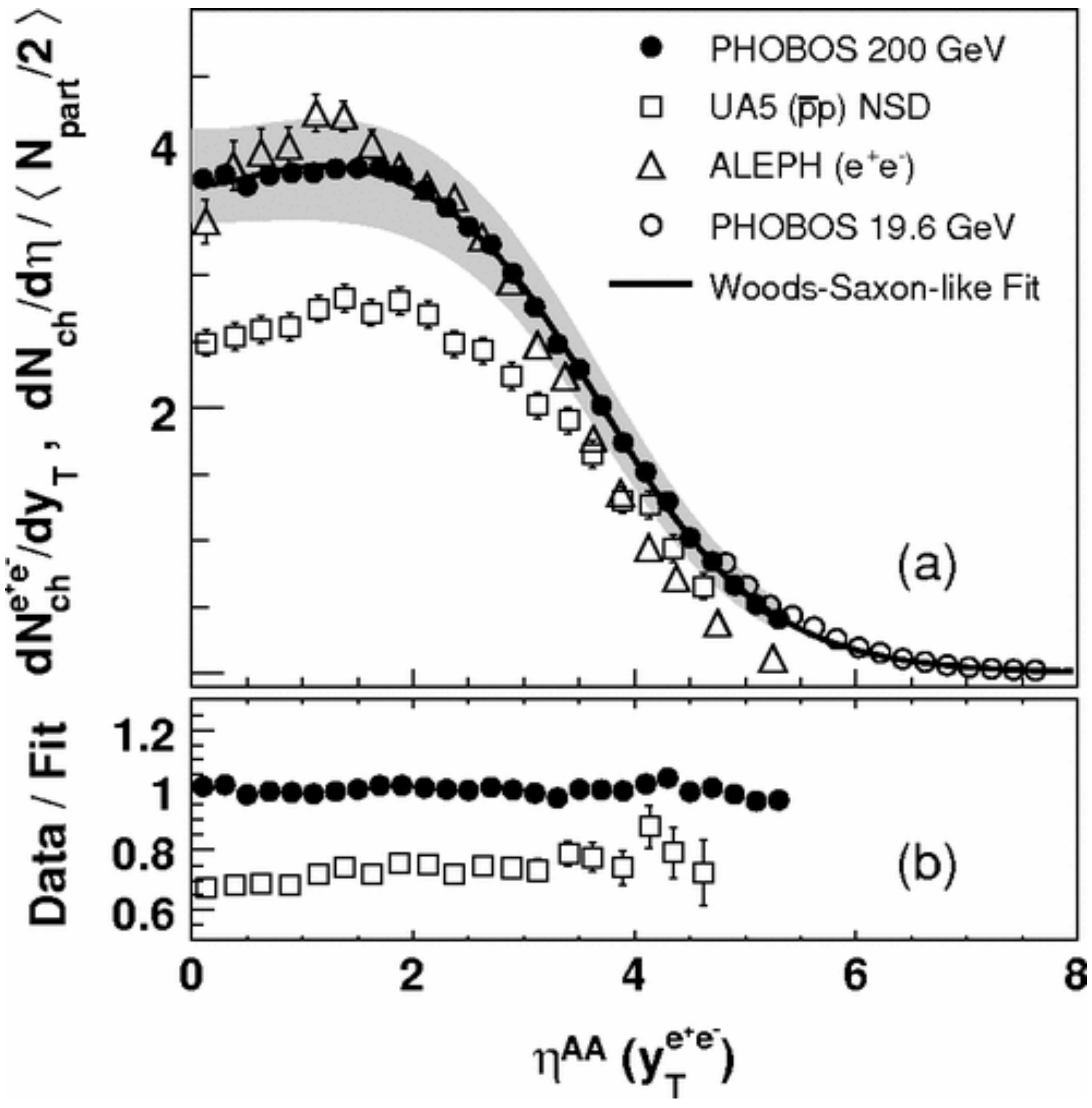
	Access to low level detector quantities	Ability to run full simulations with modern MC
ALEPH	✅	⏸️
DELPHI	✅	✅
OPAL …	😄	😄



Drawing from Luna Chen's and Yen-Jie Lee's talks

Particle Production Across Systems

Centrality and energy dependence of charged-particle multiplicities in heavy ion collisions in the context of elementary reactions, PhysRevC.74.021902



Sphericity Tensor and Derived Observables

Linearized sphericity tensor

The eigenvalues of this tensor have a direct geometrical interpretation. Given the normalisation of the M_{xyz} tensor its eigenvalues (λ_i) are defined such that $\sum_i \lambda_i = 1$. When two of them are zero, the third one must be equal to one. In this case, the final state consists of two back-to-back jets. If there are three jets with the momenta lying on the same plane, one of the eigenvalues will be equal to 0. Instead, if the spread of the momenta in the final state is close to spherical, the eigenvalues will have similar values between each other, close to 1/3.

$$M_{xyz} = \frac{1}{\sum_i |\vec{p_i}|} \sum_i \frac{1}{|\vec{p_i}|} \begin{pmatrix} p_{x,i}^2 & p_{x,i}p_{y,i} & p_{x,i}p_{z,i} \\ p_{y,i}p_{x,i} & p_{y,i}^2 & p_{y,i}p_{z,i} \\ p_{z,i}p_{x,i} & p_{z,i}p_{y,i} & p_{z,i}^2 \end{pmatrix}.$$

Define ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$

Sphericity	$S = 3/2(\lambda_2 + \lambda_3)$
Aplanarity	$A = 3/2 \cdot \lambda_3$
C-parameter	$C = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)$
D-parameter	$D = 27(\lambda_1\lambda_2\lambda_3)$

Using eigenvalues can define useful geometric quantities

Many Ongoing Efforts and Opportunities

- Standard model measurements
  - Jets (e.g. radius-dependent spectra, substructure), energy-energy correlator, thrust, heavy-jet mass, …
- Heavy-ion collision focus
  - Strangeness enhancement, collectivity small system, jet substructure correlation, …
- Electron-ion collider references
  - Jet properties (e.g. charge, in-jet anisotropy), …

Luna (Yi) Chen's talk

Many Ongoing Efforts and Opportunities

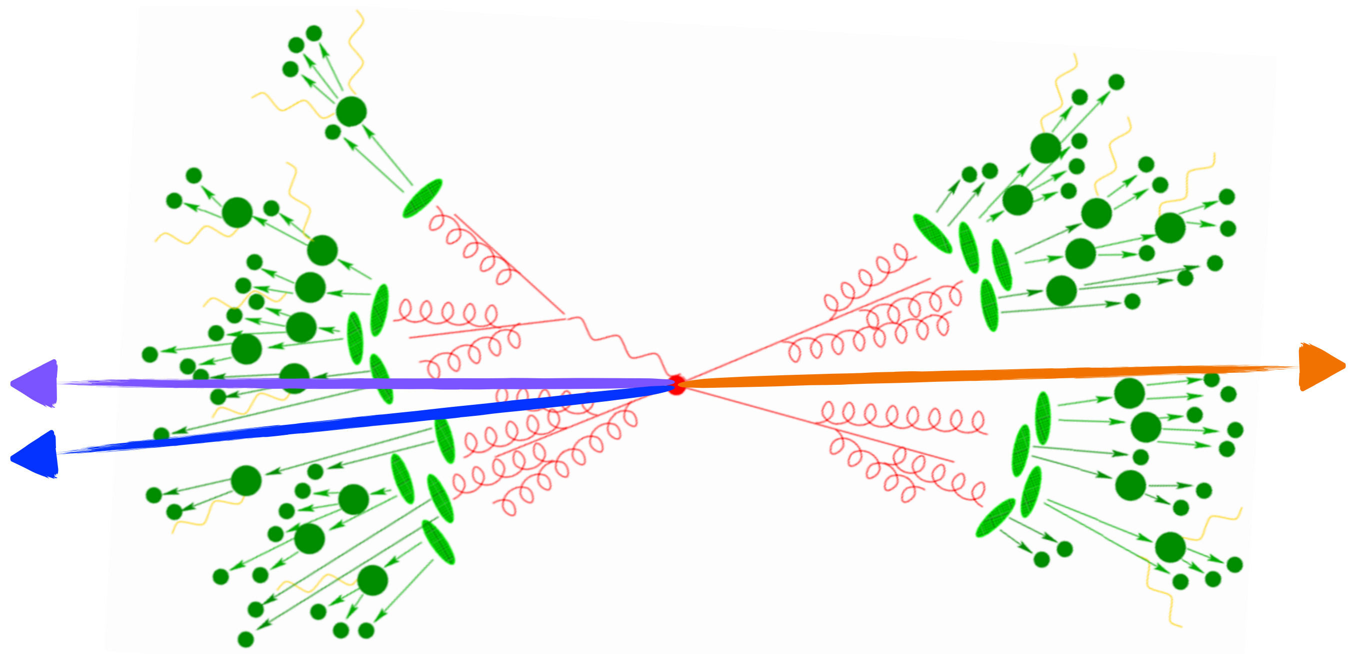
- Standard model measurements
  - Jets (e.g. radius-dependent spectra, substructure), energy-energy correlator, thrust, heavy-jet mass, …
- Heavy-ion collision focus
  - Strangeness enhancement, collectivity small system, jet substructure correlation, …
- Electron-ion collider references
  - Jet properties (e.g. charge, in-jet anisotropy), …

Physics Motivation

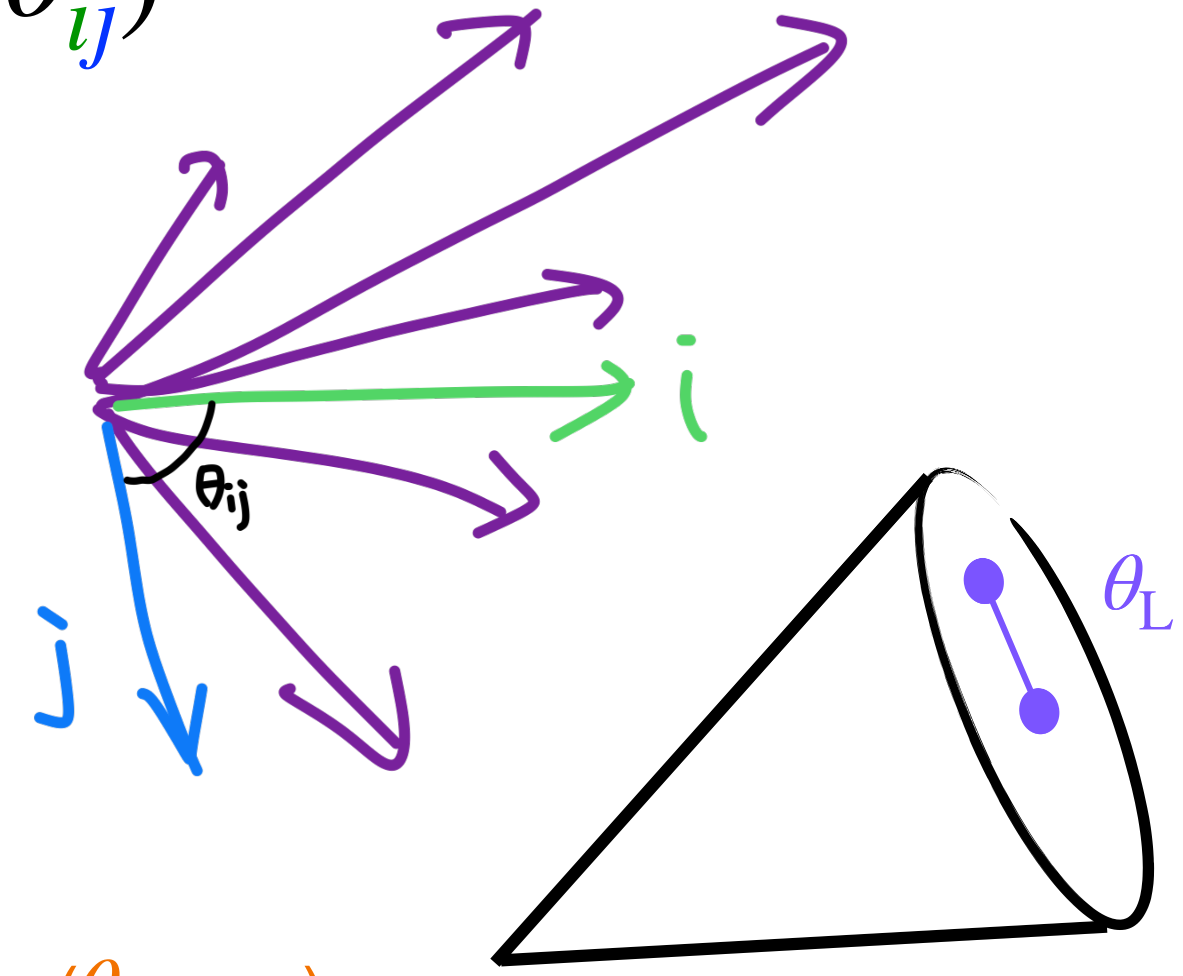
Emerging observable across systems:

Energy-Energy Correlator (EEC)

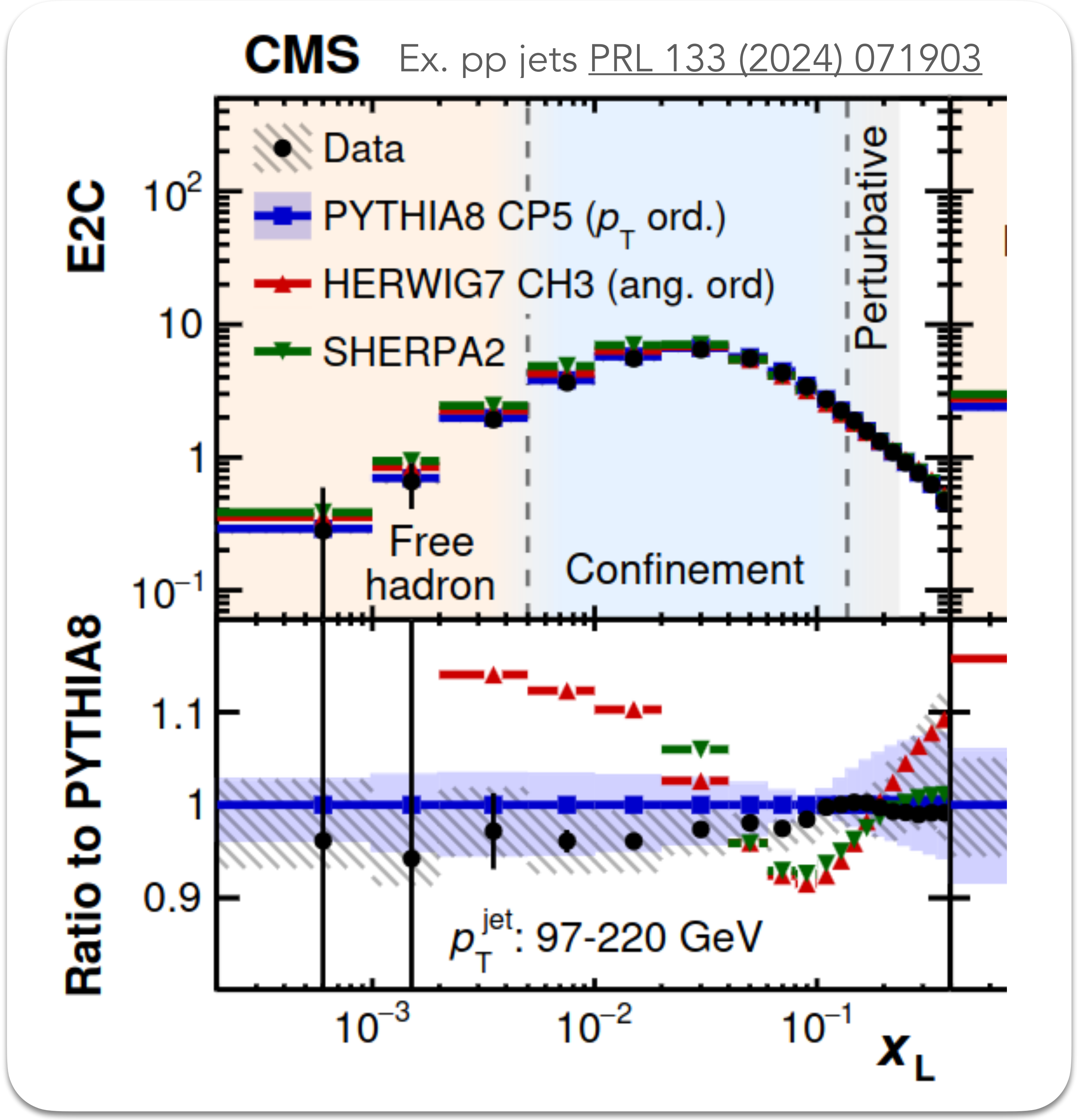
$$\mathrm{EEC}(\theta) \equiv \sum_{i>j} \int d\sigma \frac{E_i E_j}{Q^2} \delta(\theta_L - \theta_{ij})$$



Collinear ($\theta_L \sim 0$) Back-to-back or Sudakov ($\theta_L \sim \pi$)



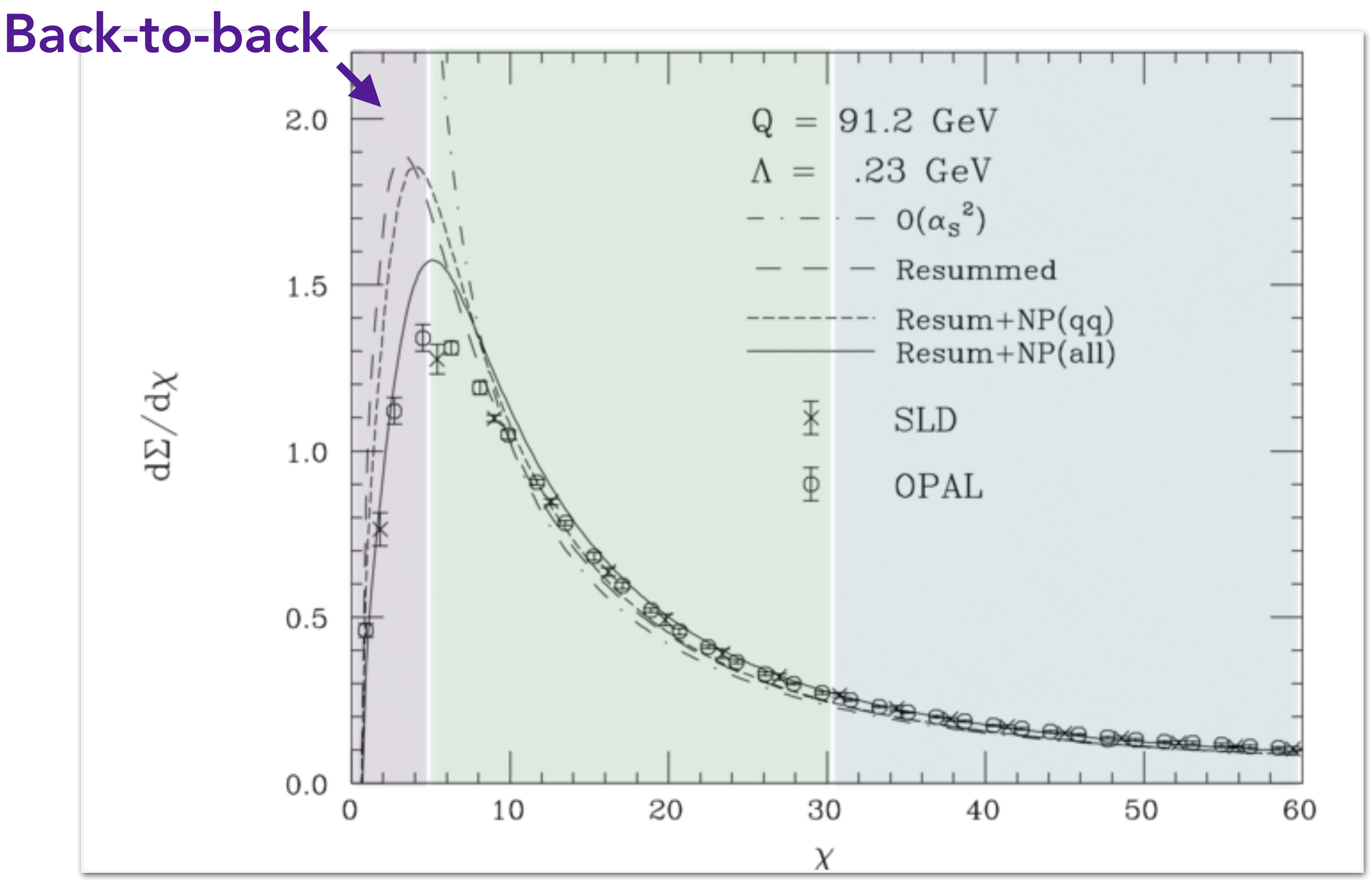
Probe phase transitions, confinement mechanism, parton shower dynamics, …

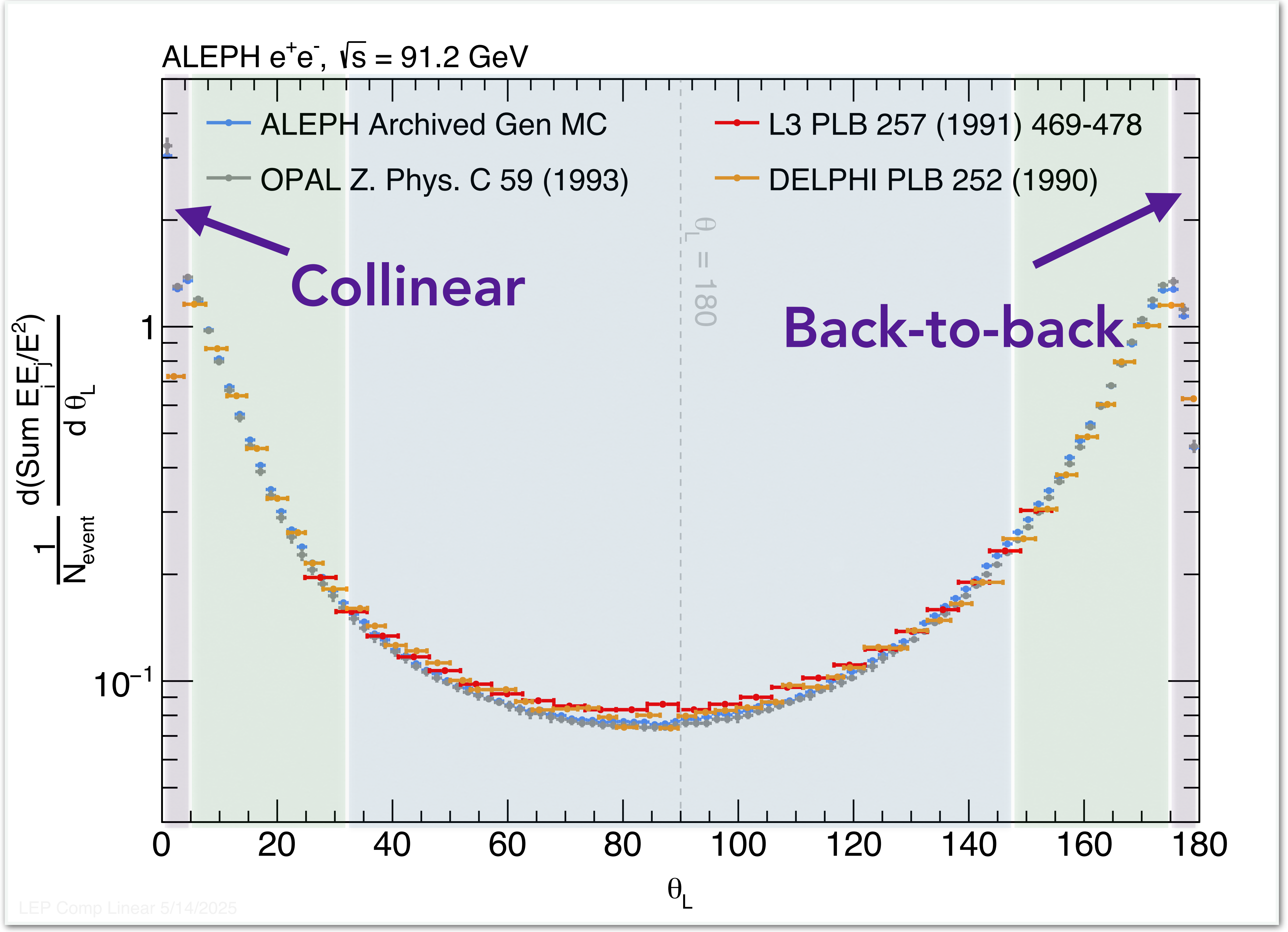


Talks by: Hannah Bossi (MIT) and Luna (Yi) Chen (Vanderbilt)
Energy Correlations in e+e- Annihilations: Testing QCD, PRL 41, 1585 (1978)
Onset of direct Exp-Th collaboration: https://indico.cern.ch/event/1388246/

Th: Max Jaarsma, Yibei Li, Ian Moult, Wouter Waalewijn, HuaXing Zhu
Exp: Hannah Bossi, Jingyu Zhang, Luna (Yi) Chen, Yen-Jie Lee, e+e- All.

Previous Measurements

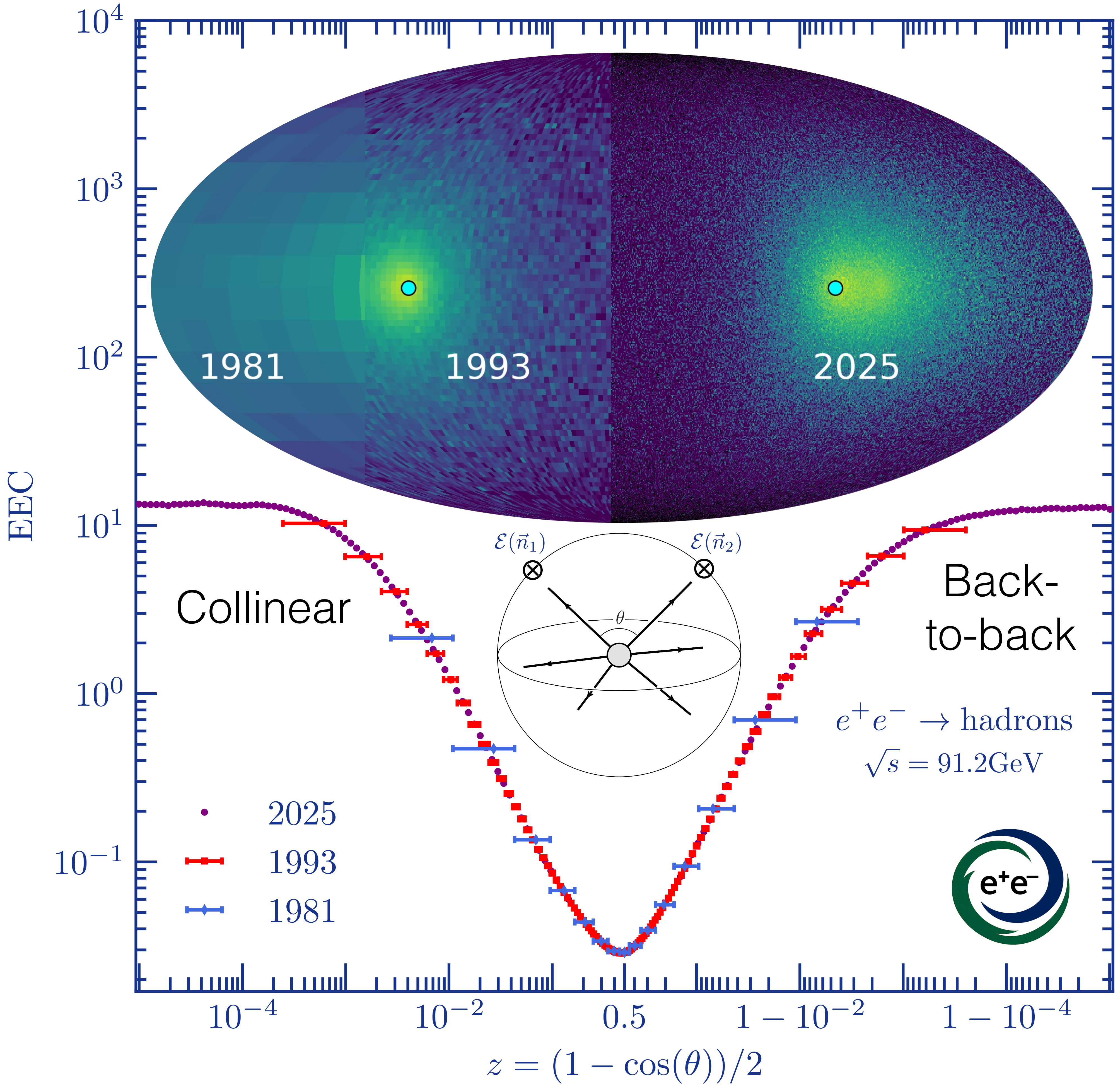




Different Focus: many small bins in intermediate angle regions but few in small angle or back-to-back (different focus also previously for thrust) → **Chance for reanalysis**

Th: Max Jaarsma, Yibei Li, Ian Moult, Wouter Waalewijn, HuaXing Zhu
Exp: Hannah Bossi, Jingyu Zhang, Luna (Yi) Chen, Yen-Jie Lee, e+e- All.

Charged Particle EEC Reanalysis



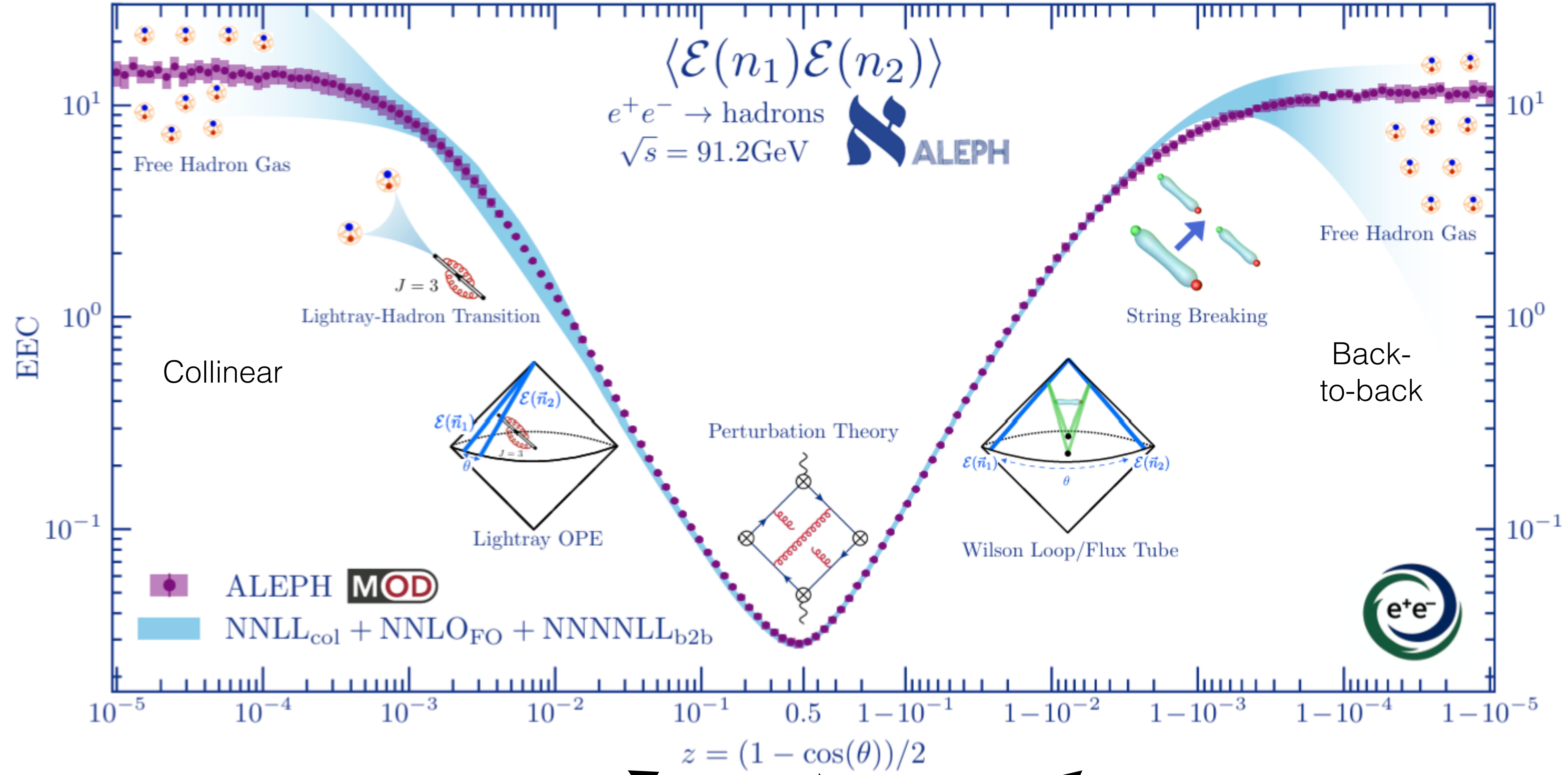
Leverage track function formalism and high charged particle angular precision

Reparameterize to $z = (1 - \cos\theta)/2$
Extending to small and large angles

Improved "resolution" to fine structure imprinted on EEC in different regimes

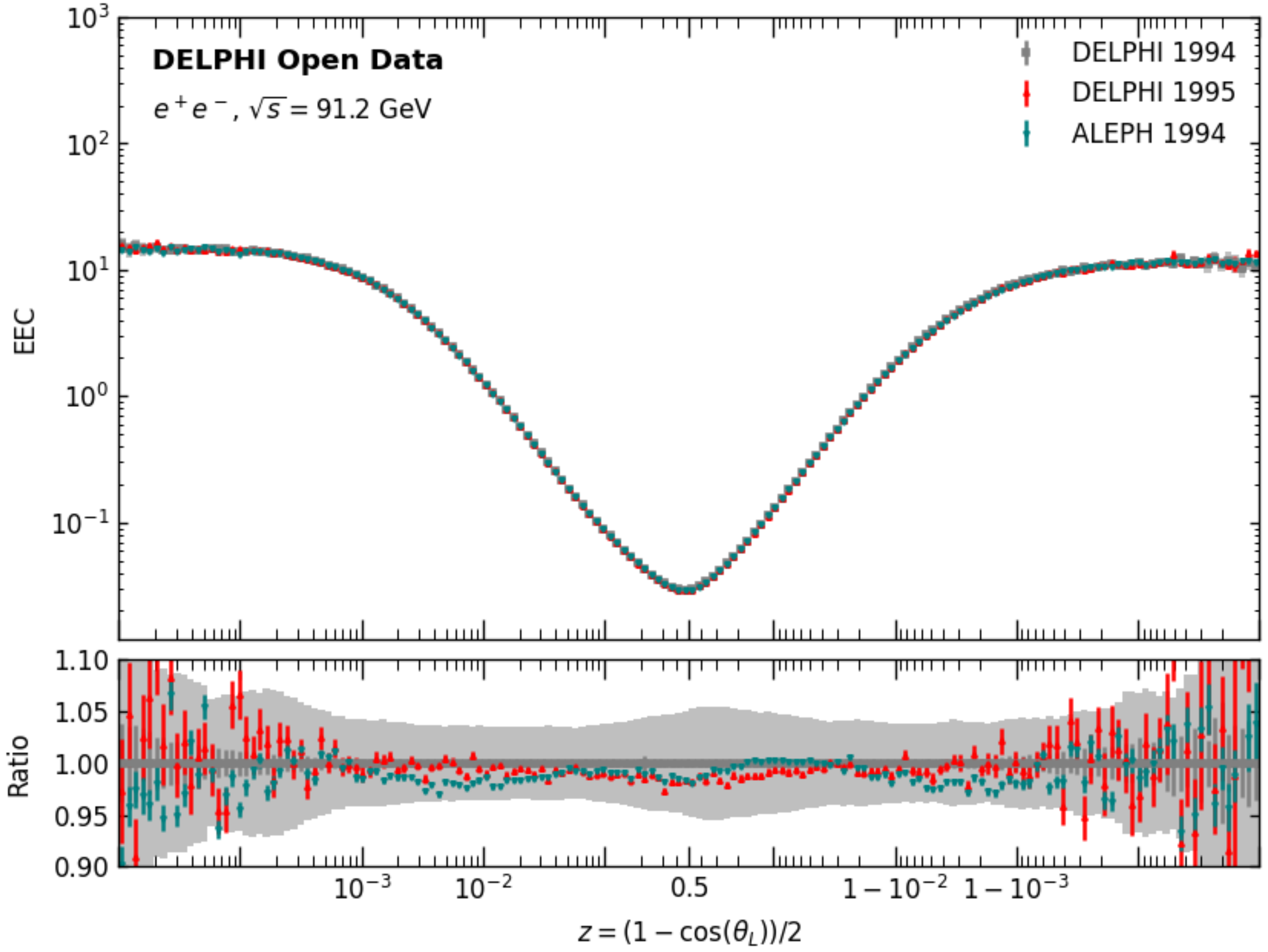
Th: Max Jaarsma, Yibei Li, Ian Moult, Wouter Waalewijn, HuaXing Zhu
Exp: Hannah Bossi, Jingyu Zhang, Luna (Yi) Chen, Yen-Jie Lee, e+e- All.

Charged Particle EEC and State-of-art Theory



Comparison to NNLL+NNLO+N^{4}LL calculation

Cross Experiment Comparison and Future



Different analyses, detectors, calibration, years. Compatible between ALEPH archived data and DELPHI open data!

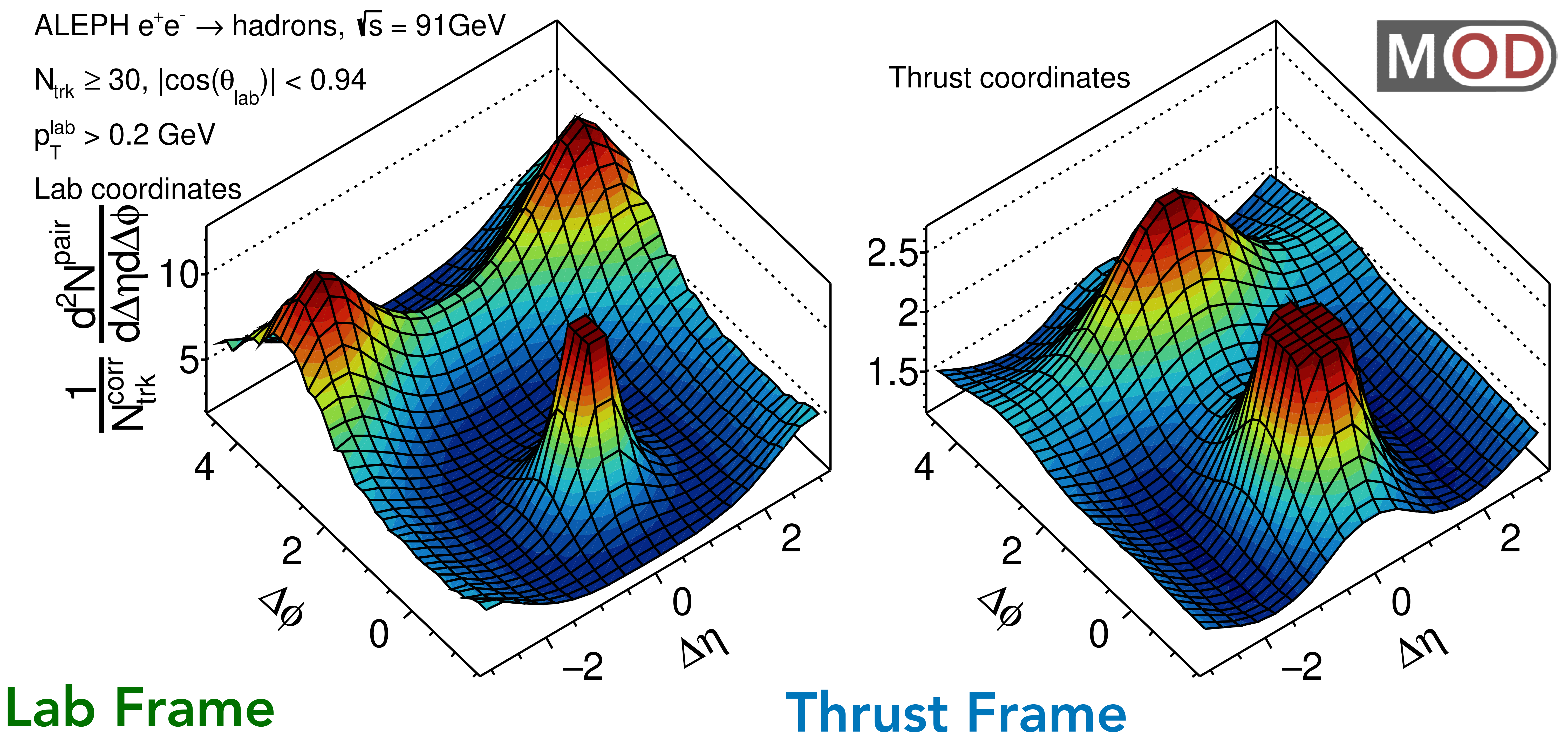
Many exciting possibilities from here (e.g. LEP2, HF tagging)!!

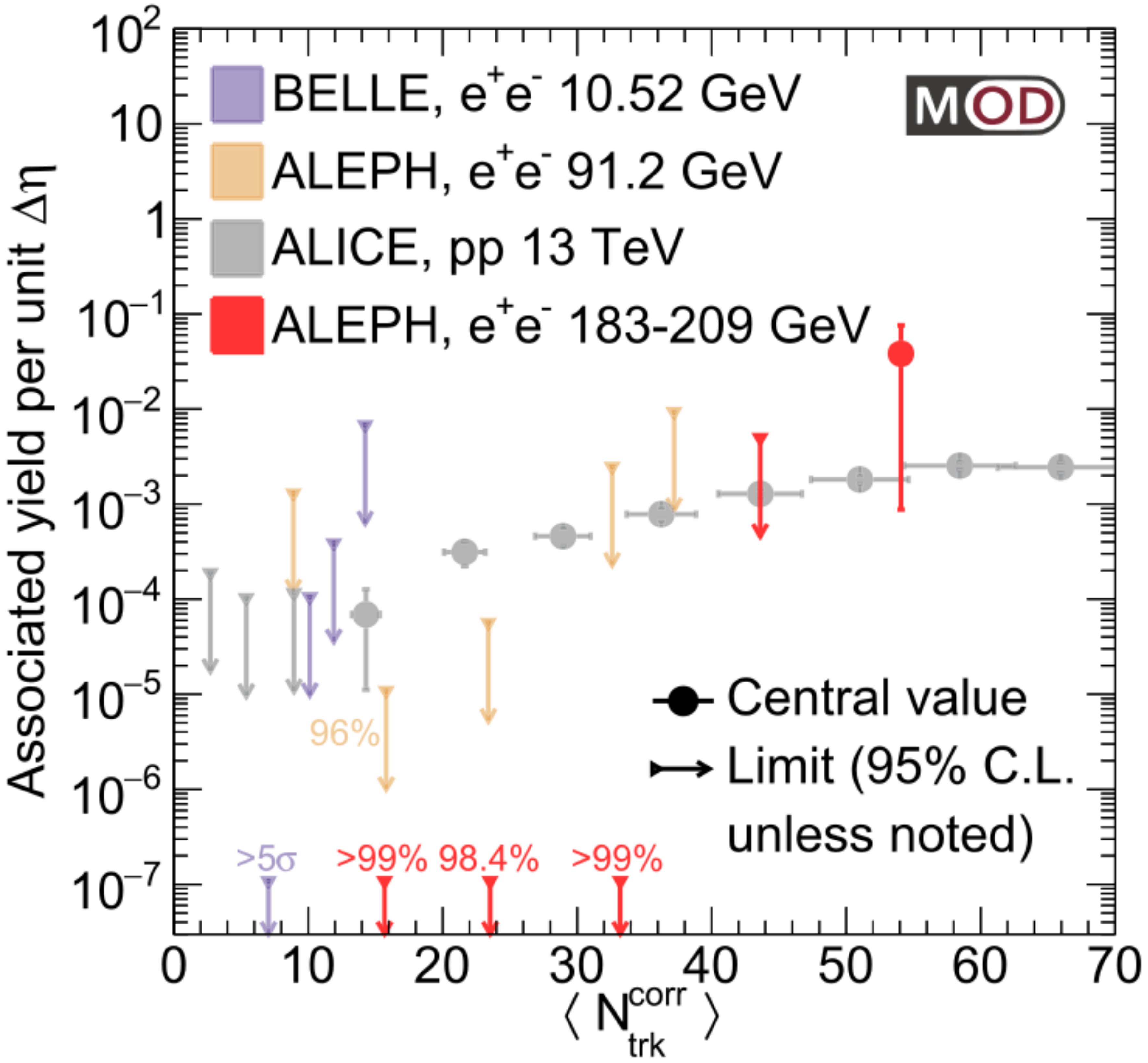
Th: Max Jaarsma, Yibei Li, Ian Moult, Wouter Waalewijn, HuaXing Zhu
Exp: Hannah Bossi, Jingyu Zhang, Luna (Yi) Chen, Yen-Jie Lee, e+e- All.

Many Ongoing Efforts and Opportunities

- Standard model measurements
  - Jets (e.g. radius-dependent spectra, substructure), energy-energy correlator, thrust, heavy-jet mass, …
- Heavy-ion collision focus
  - Strangeness enhancement, collectivity small system jet substructure correlation, …
- Electron-ion collider references
  - Jet properties (e.g. charge, in-jet anisotropy), …

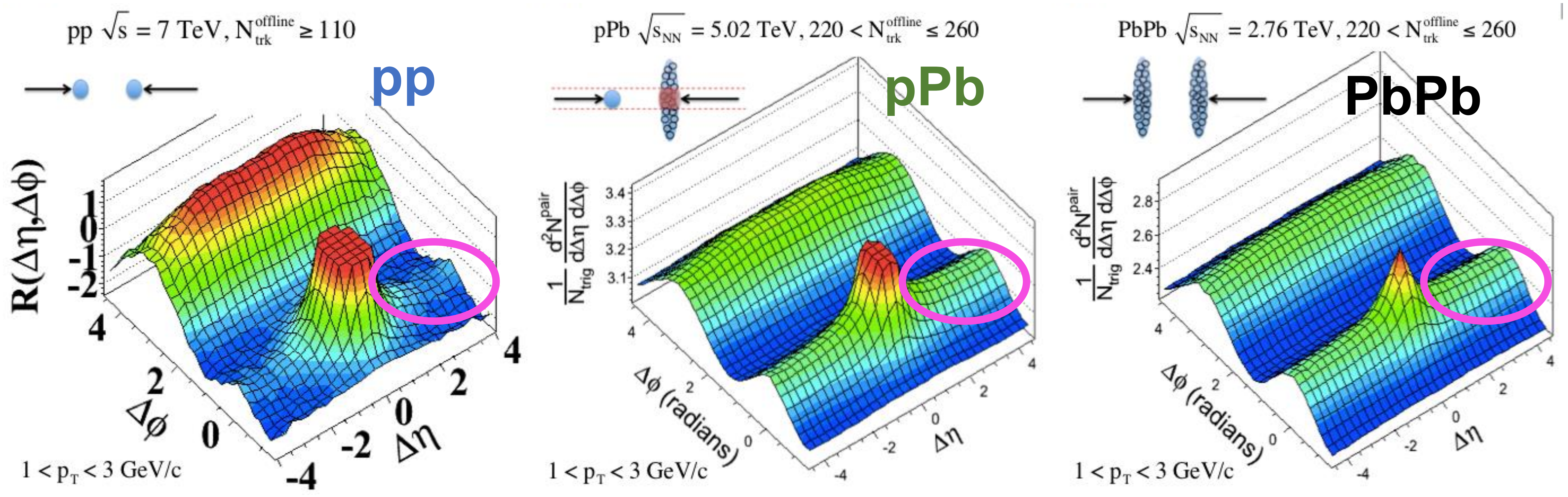
Thrust in Two-Particle Correlation Analyses



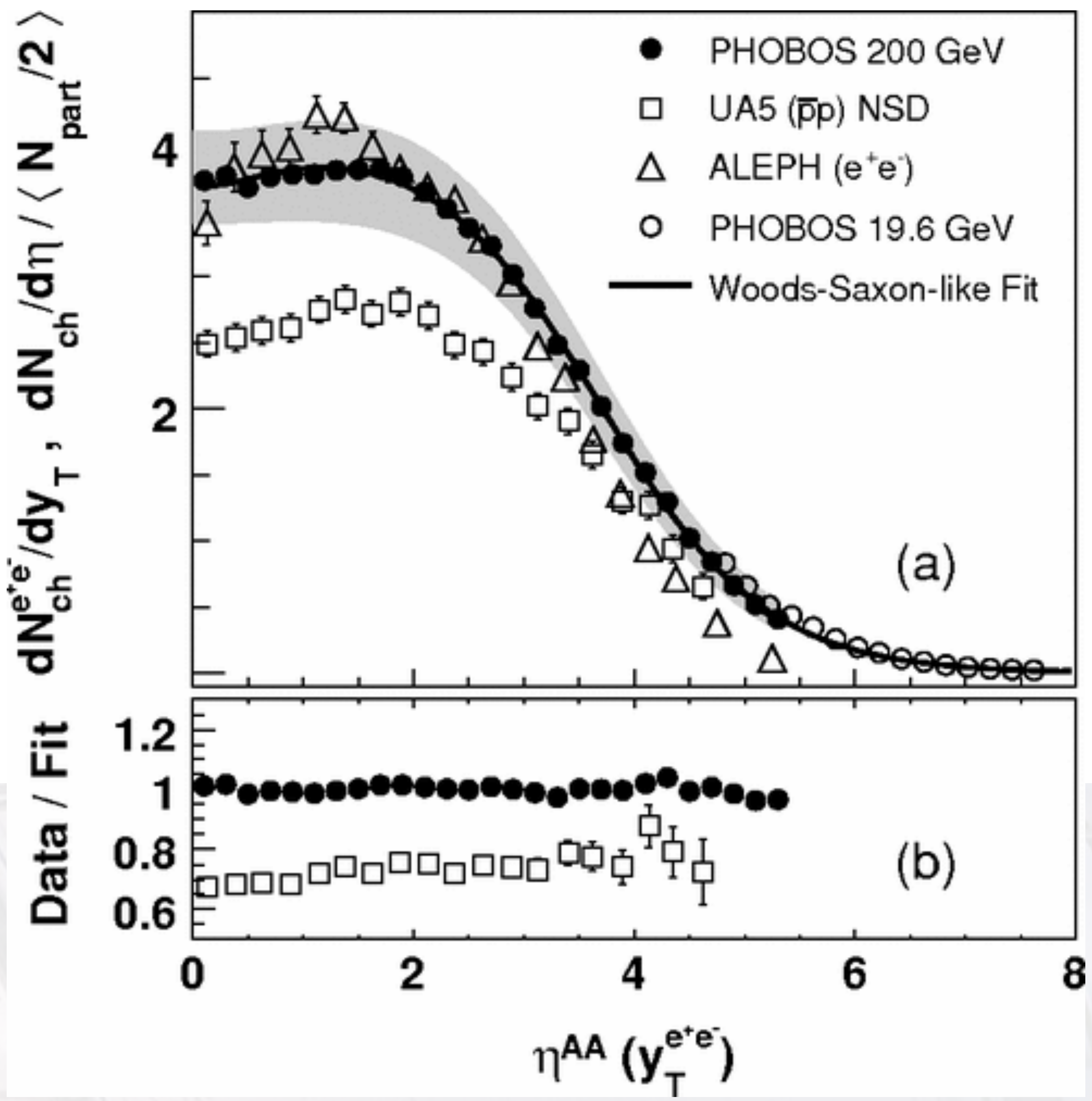


- Use thrust to align our 2PC measurements with outgoing color string, connecting small and large systems. Feel free to ask questions about these measurements!
- Now dedicated thrust measurement to simultaneously probe the P/NP regimes

Physics Motivation



Inspired by Yen-Jie Lee's, Yu-Chen "Janice" Chen's, Luna (ii) Chen's Talks



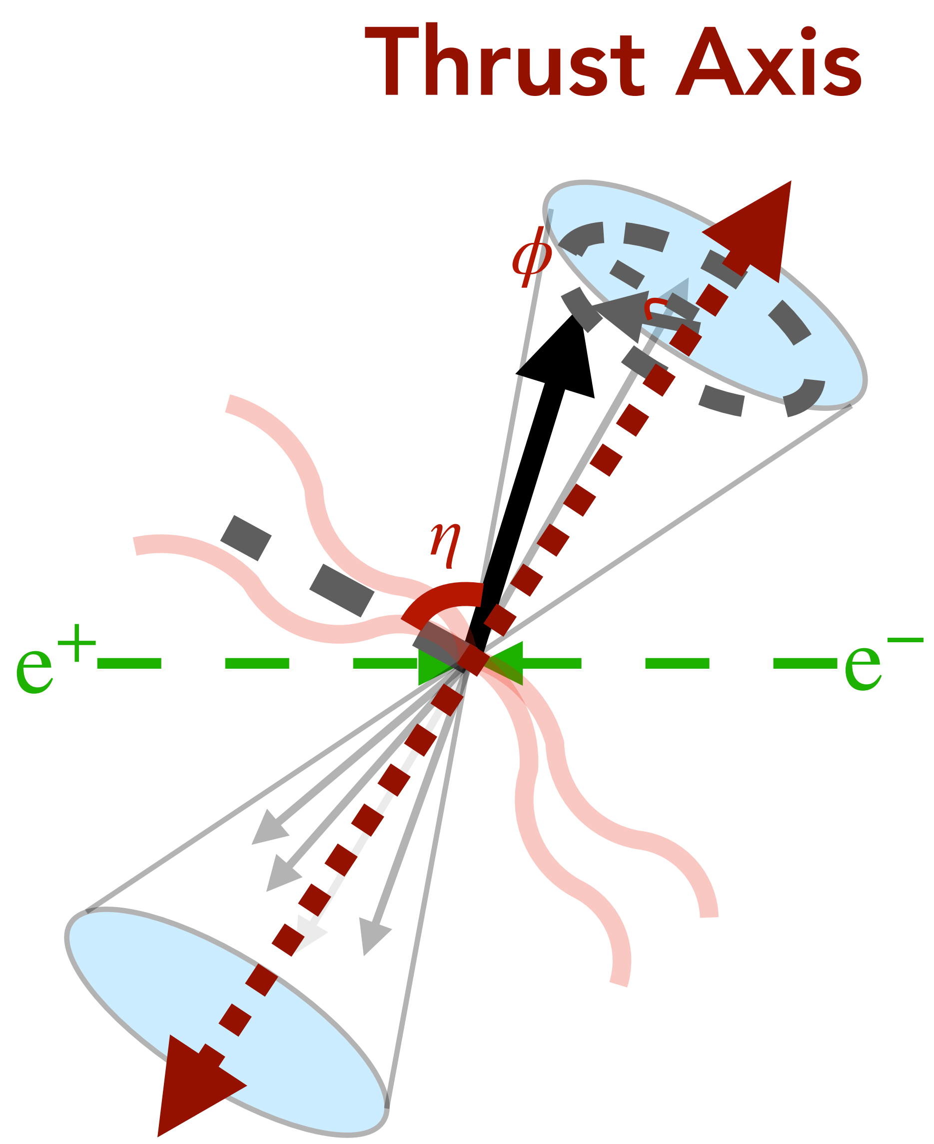
- **First unexpected discovery at LHC**: Ridge in high multiplicity pp

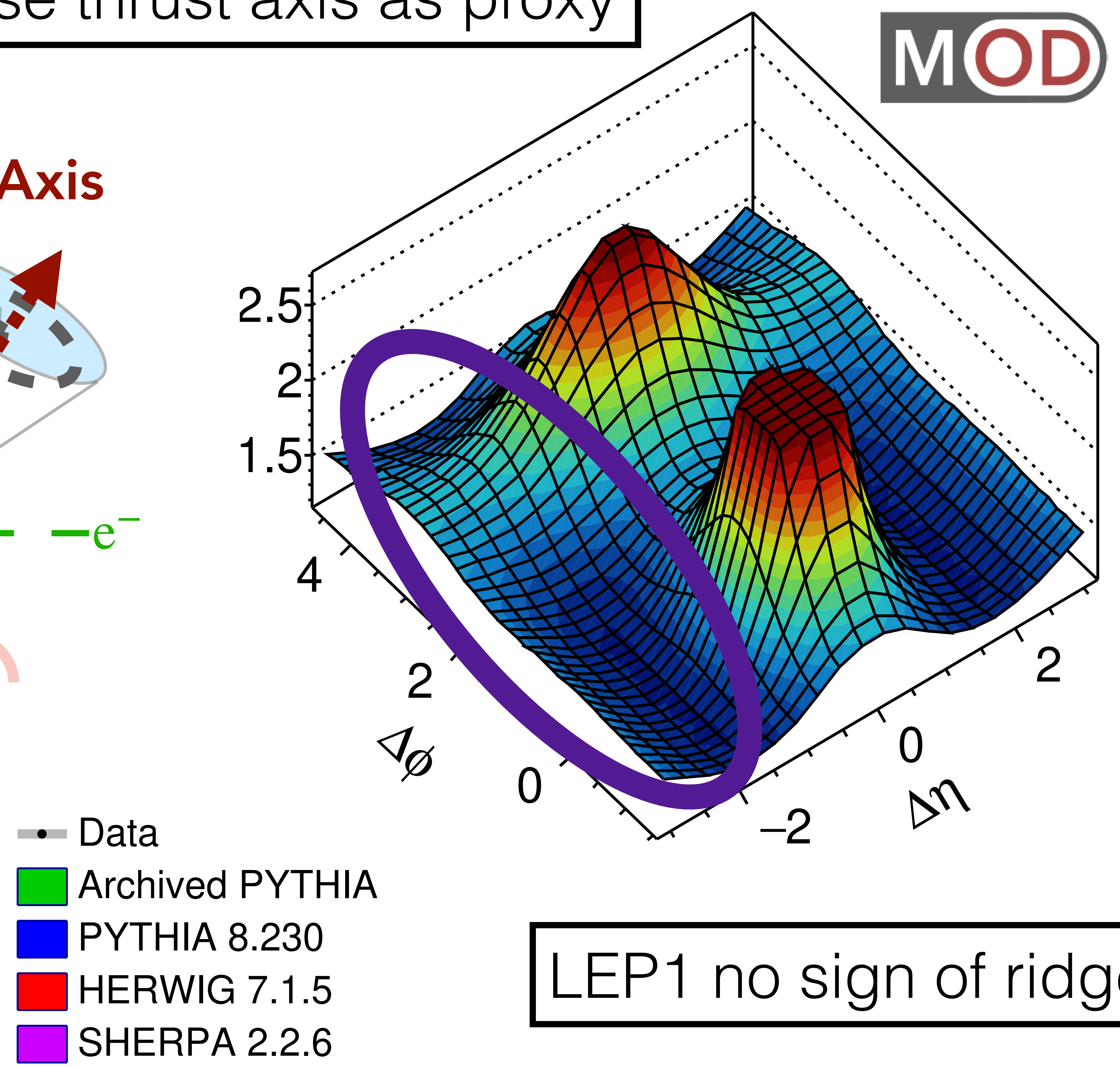
  "Ridge" refers to the long-range (large-$\Delta\eta$) near-side (small-$\Delta\phi$) correlation signal $\frac{1}{N_{\text{trig}}}\frac{d^2N^{\text{pair}}}{d\Delta\eta d\Delta\phi}$

- Origin may not necessarily be hydrodynamics from Quark Gluon Plasma (QGP), e.g. in HI collisions
- Discovery motivates the questions: What are the minimal conditions for QGP formation? How is particle production similar across systems? **What happens in the smallest collision system of point like particles?**

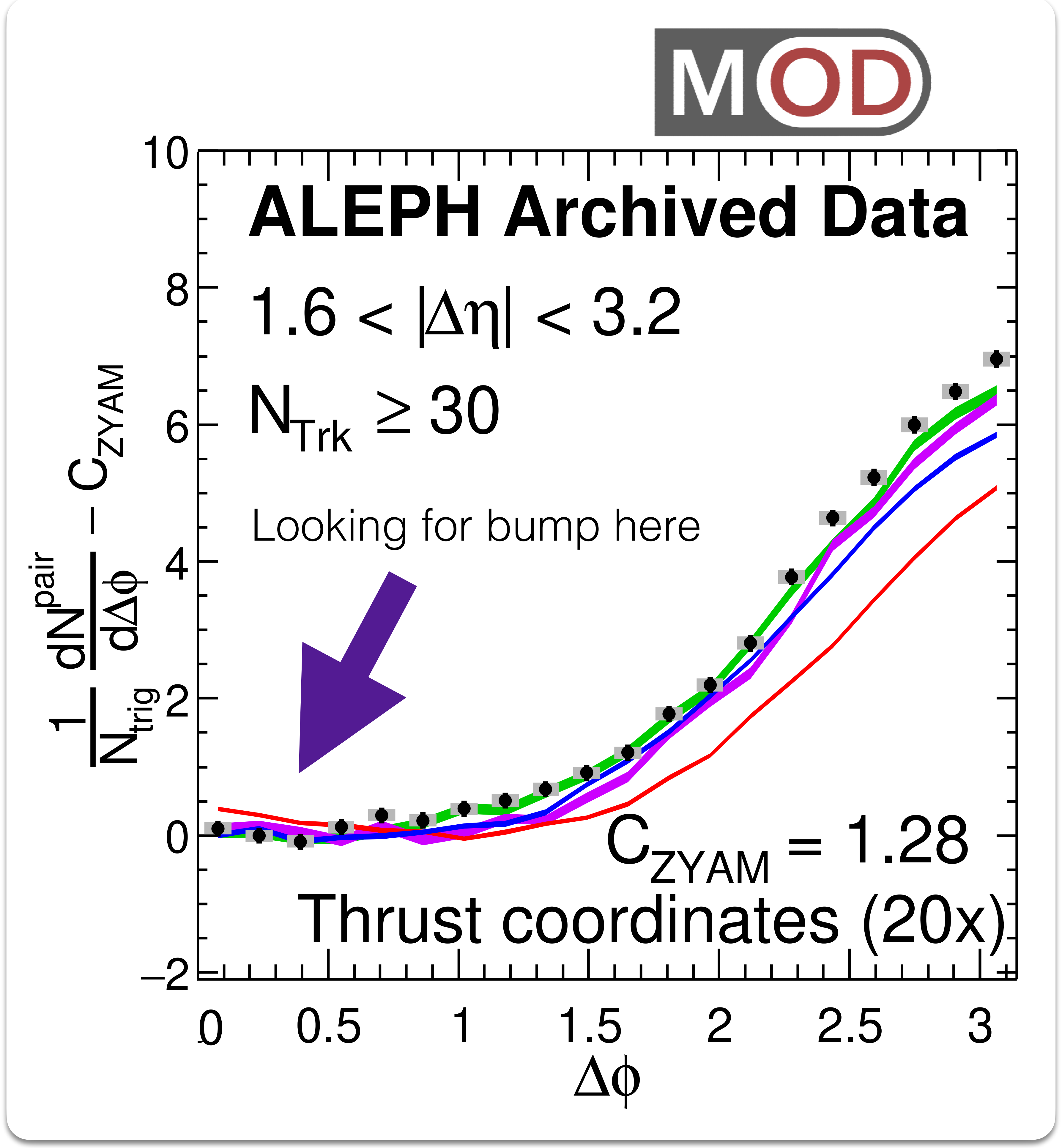
LEP 1 Charged Particle Two-Particle Correlation

Search for signatures transverse to color string, use thrust axis as proxy





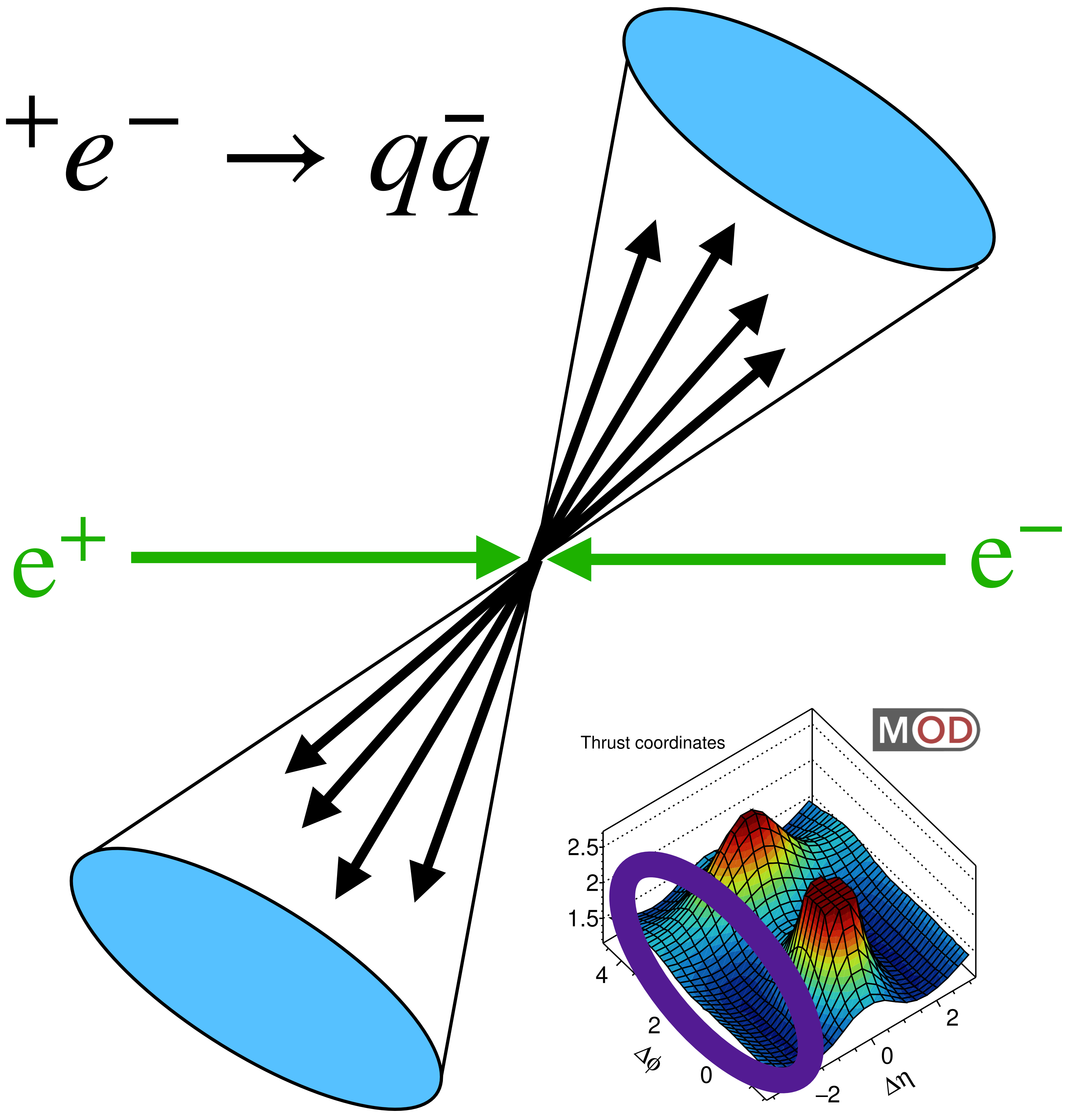
LEP1 no sign of ridge



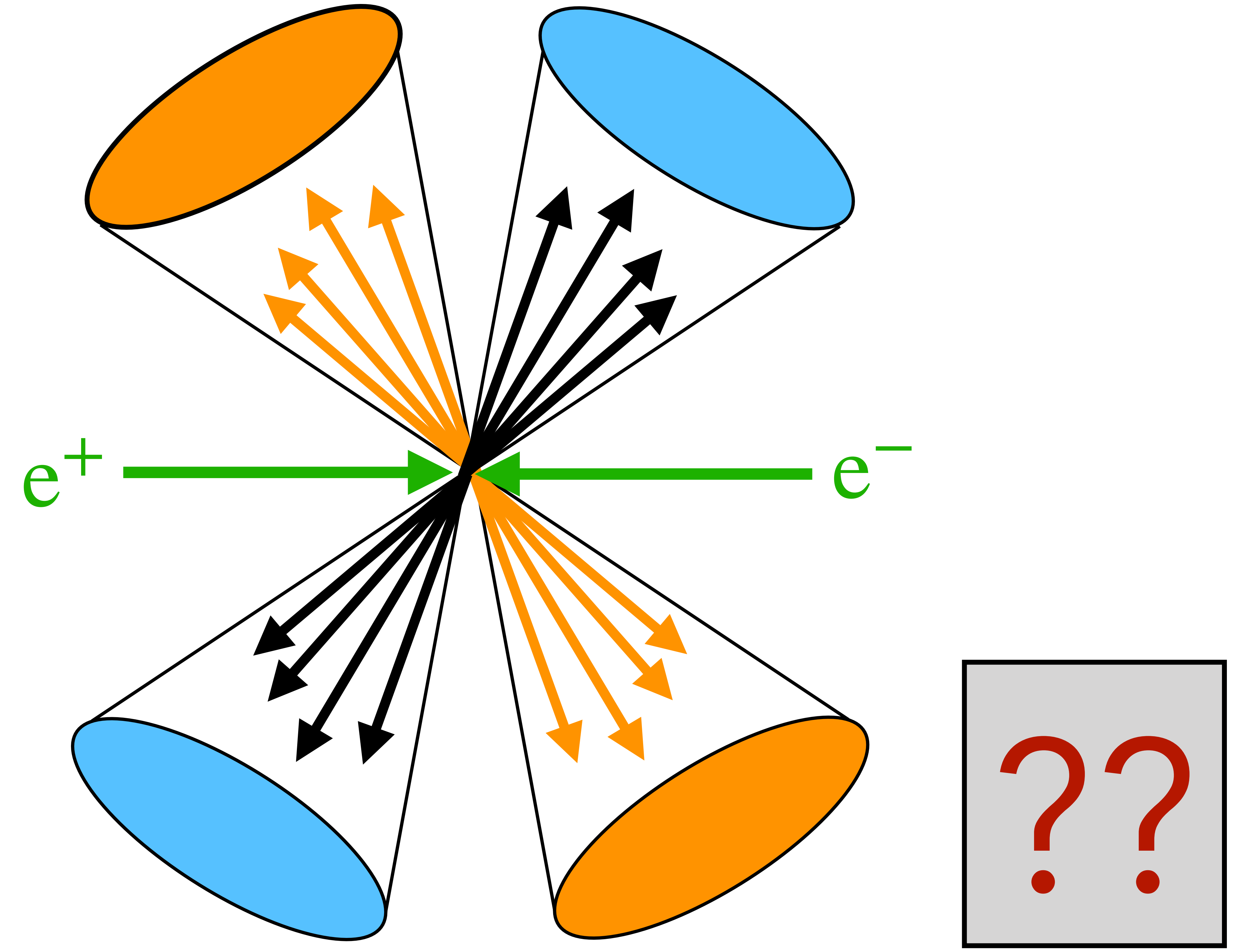
A.Badea MIT Undergraduate Thesis, PRL. 123, 212002 (2019)

What if We Overlap Two Color Strings?

$e^+e^- \rightarrow q\bar{q}$



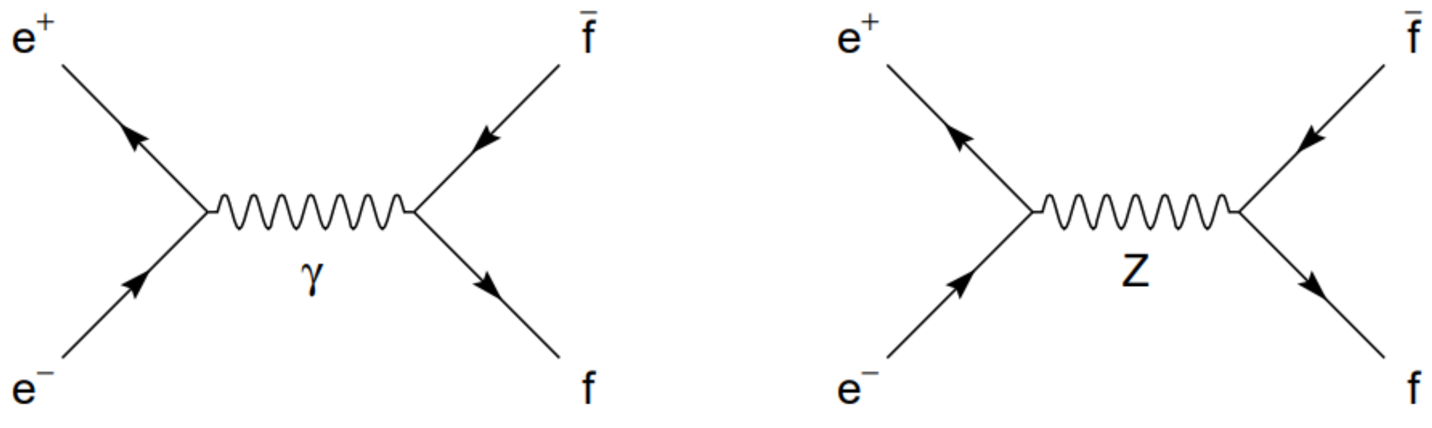
$e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}q\bar{q}$



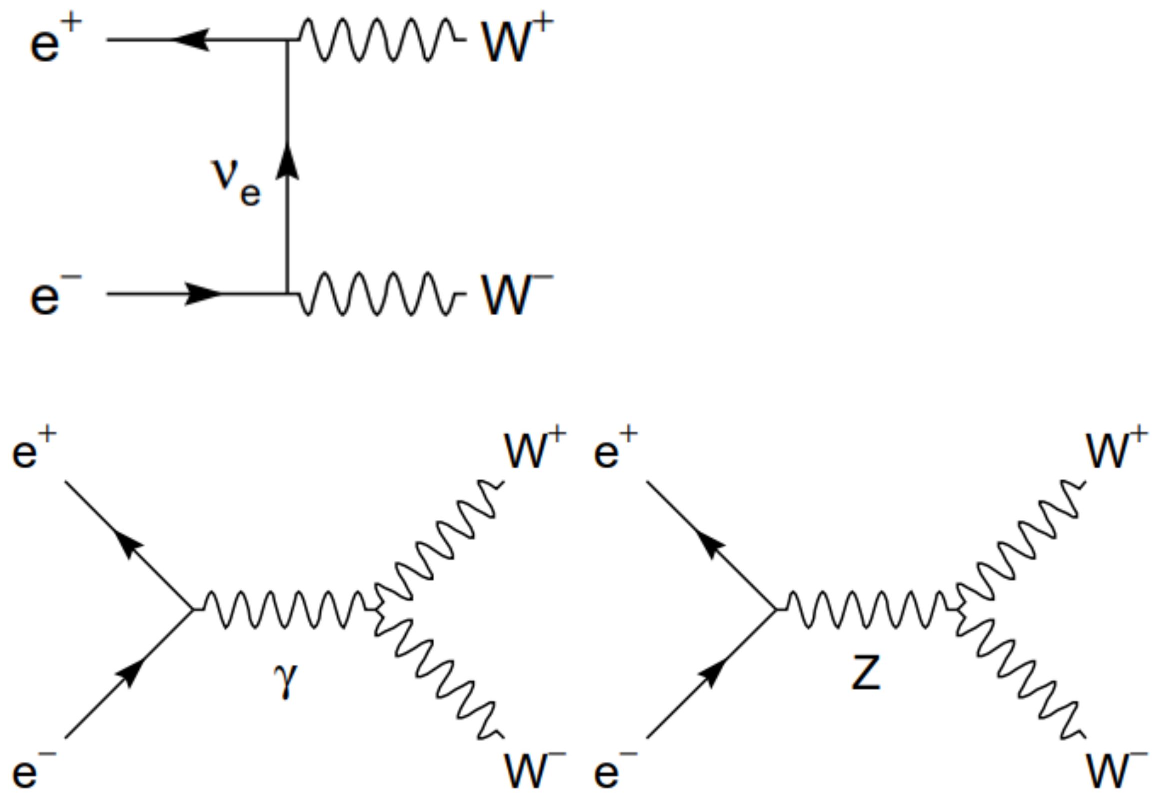
- $e^+e^- \rightarrow q\bar{q}$ single color-string fragmenting in vacuum (DGLAP evolution), hadronizing into charged particles
- $e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}q\bar{q}$ possible (partonic) interactions between two color-strings. Phase space for higher multiplicity production

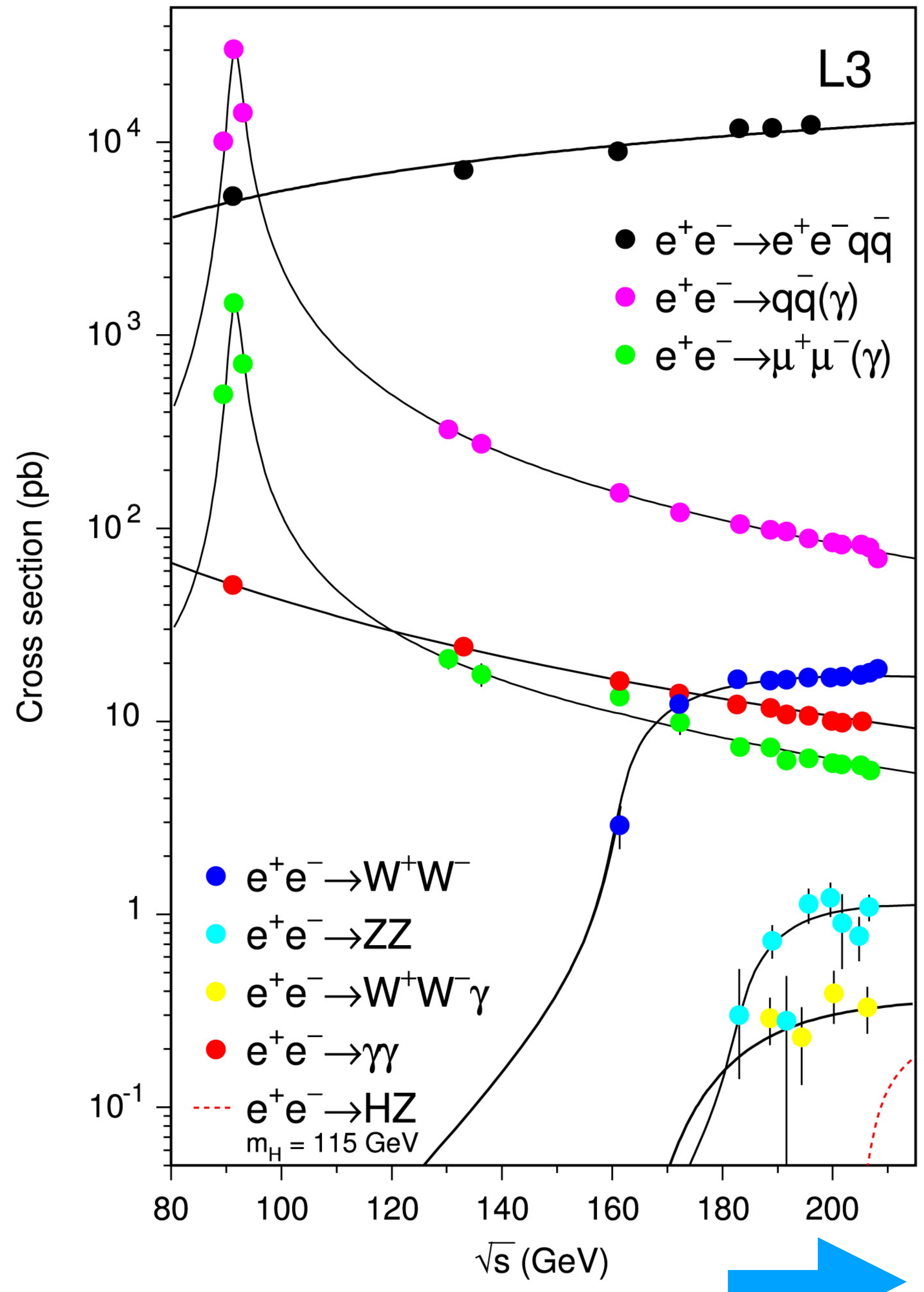
High Multiplicity Reach in LEP 2

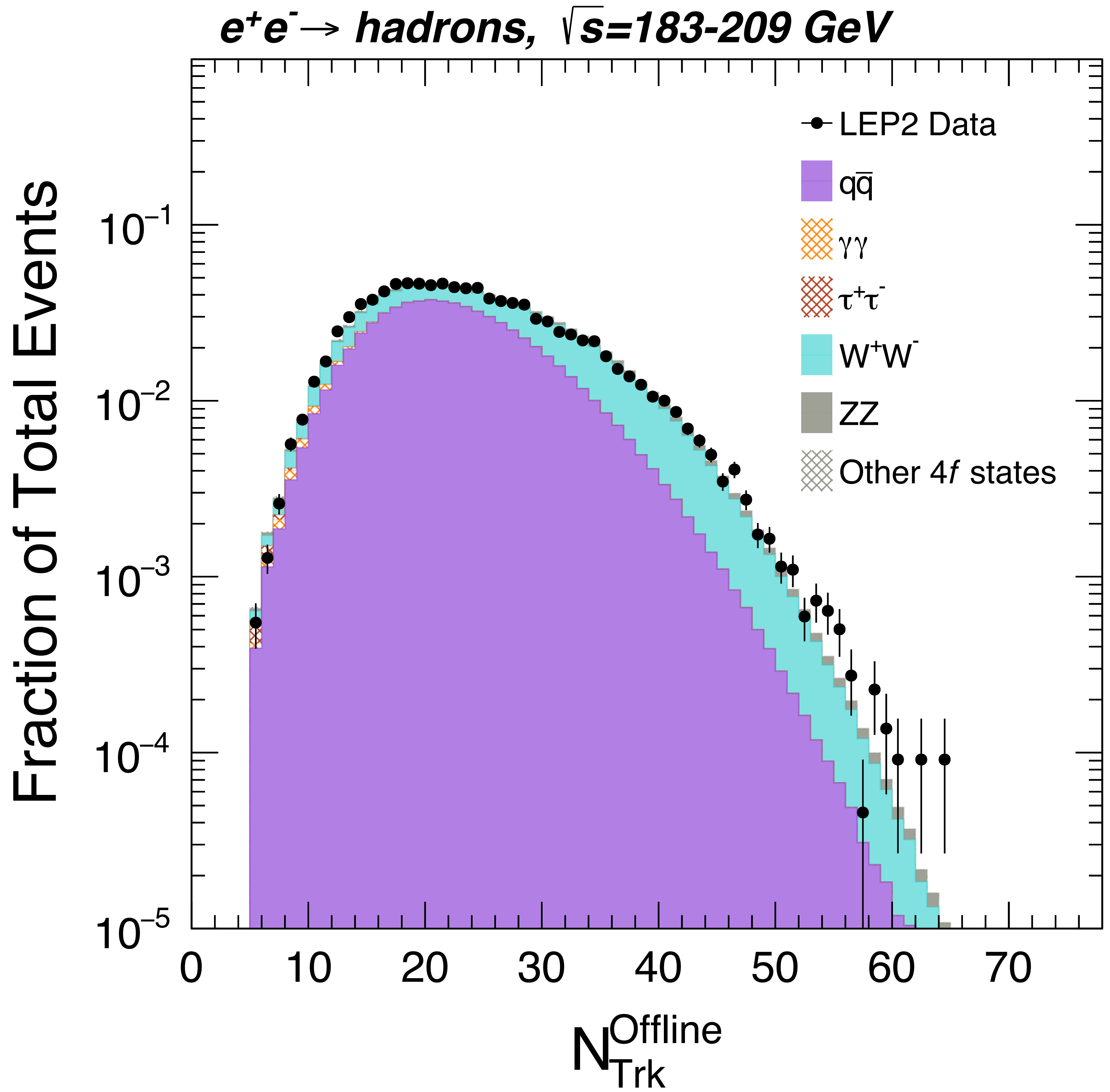
● $e^+e^- \to q\bar{q}(\gamma)$



● $e^+e^- \to W^+W^-$



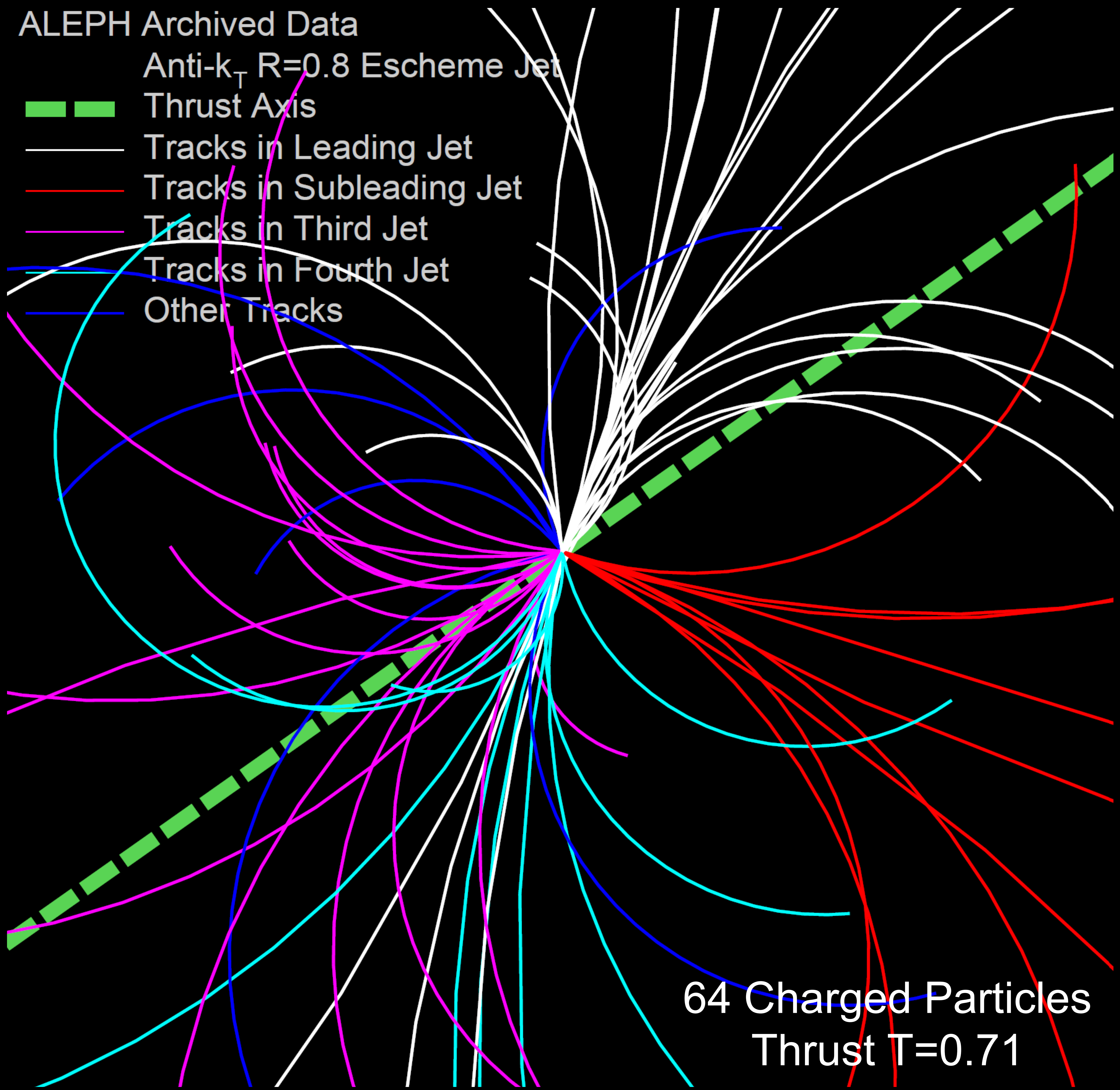


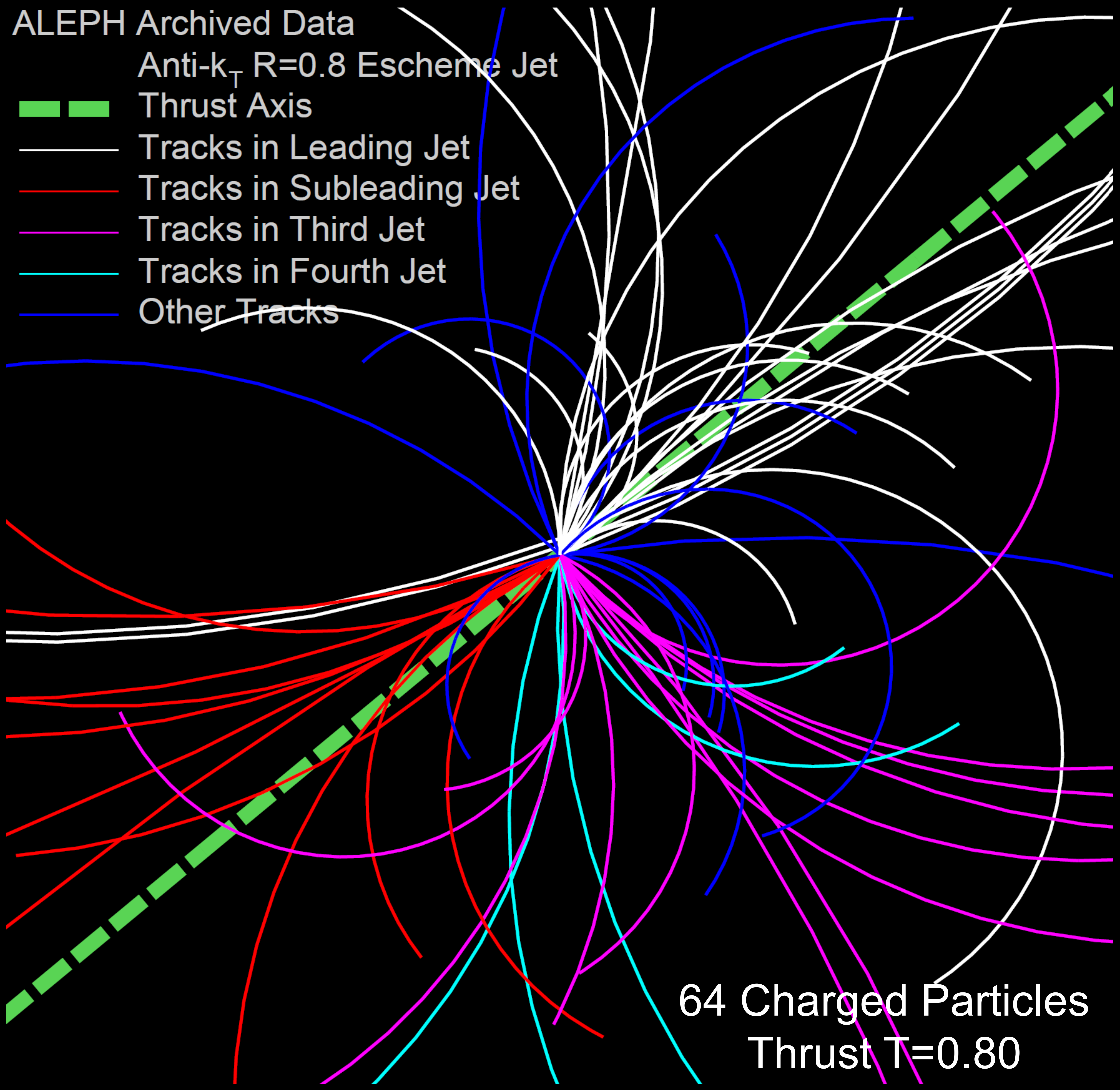


- **LEP2 energies** gives access to different physics processes
- **W+W- contribution** significant at high multiplicity ($N_{\mathrm{Trk}}^{\mathrm{Offline}}$> 40)

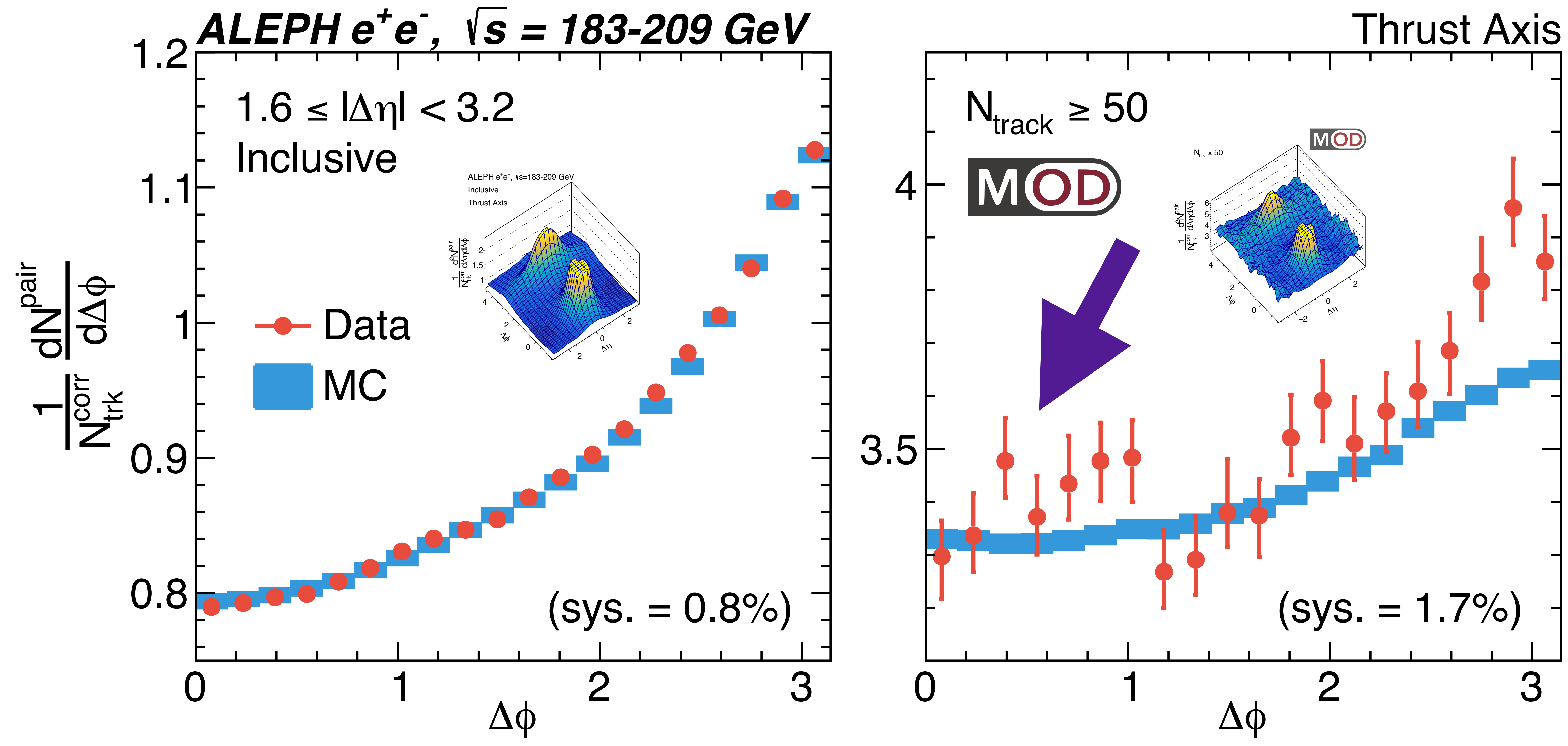
Yu-Chen "Janice" Chen, Tzu-An Sheng, Yen-Jie Lee (MIT)

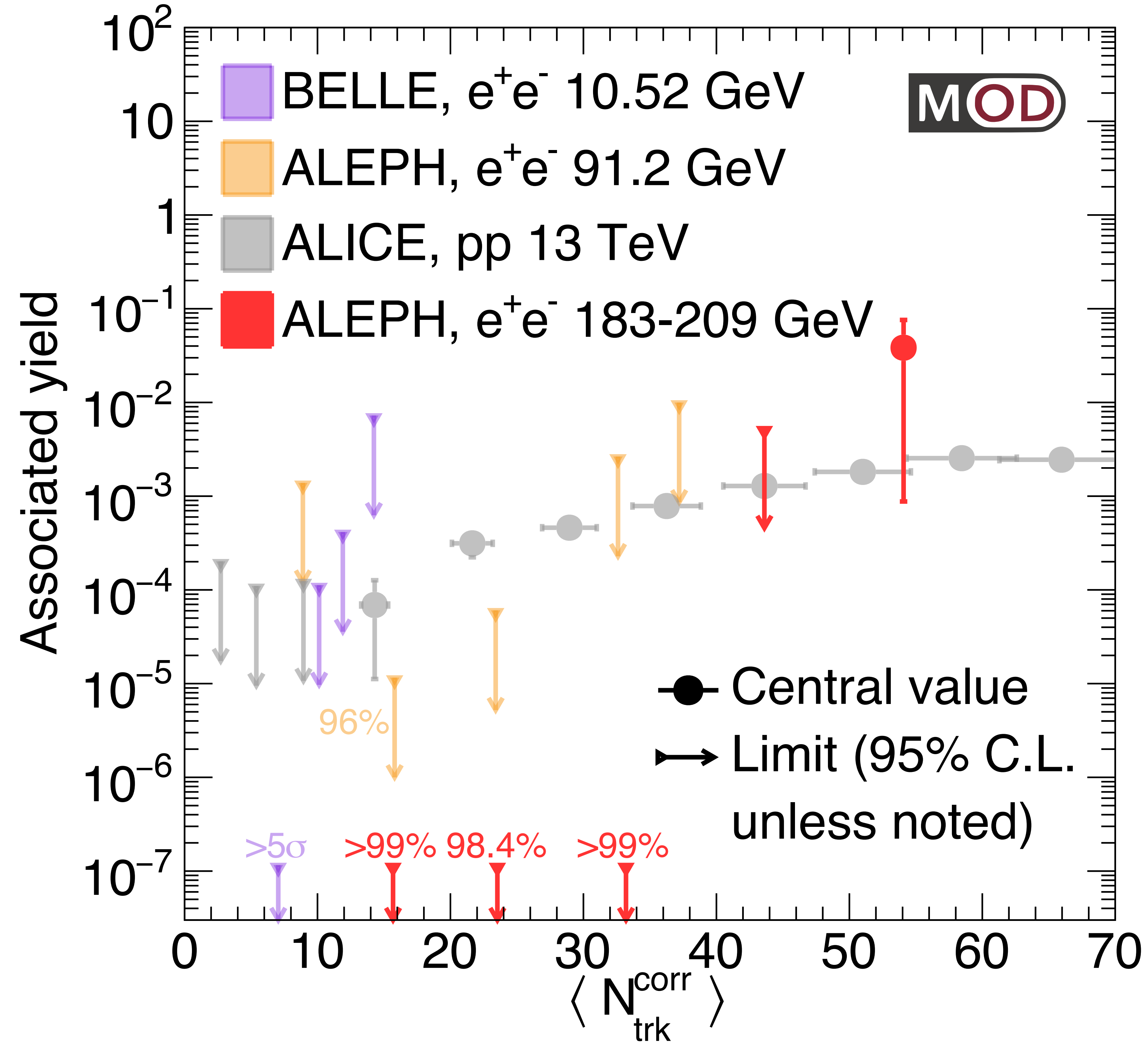
Highest Multiplicity LEP 2 Events





LEP 2 Charged Particle Two-Particle Correlation



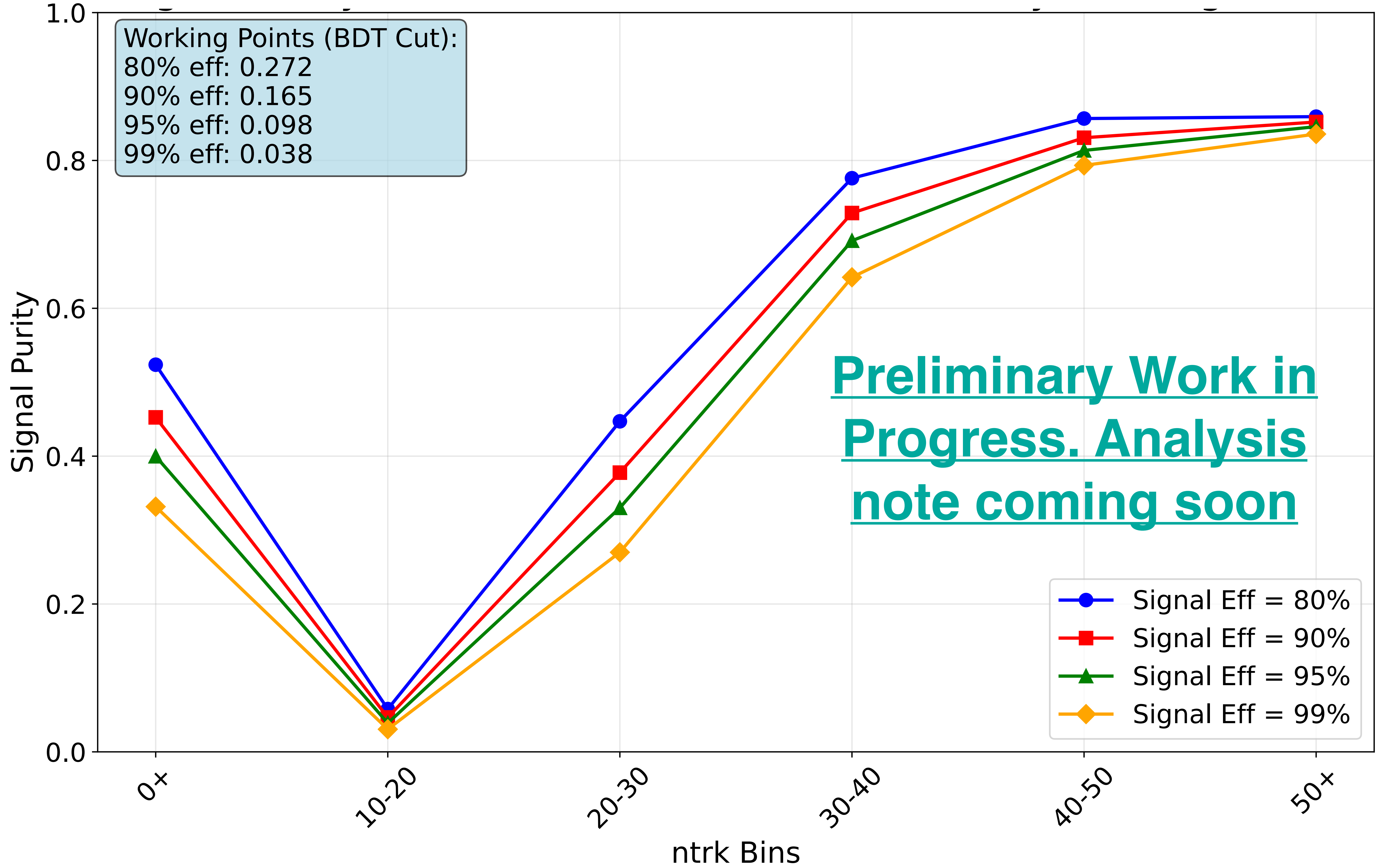


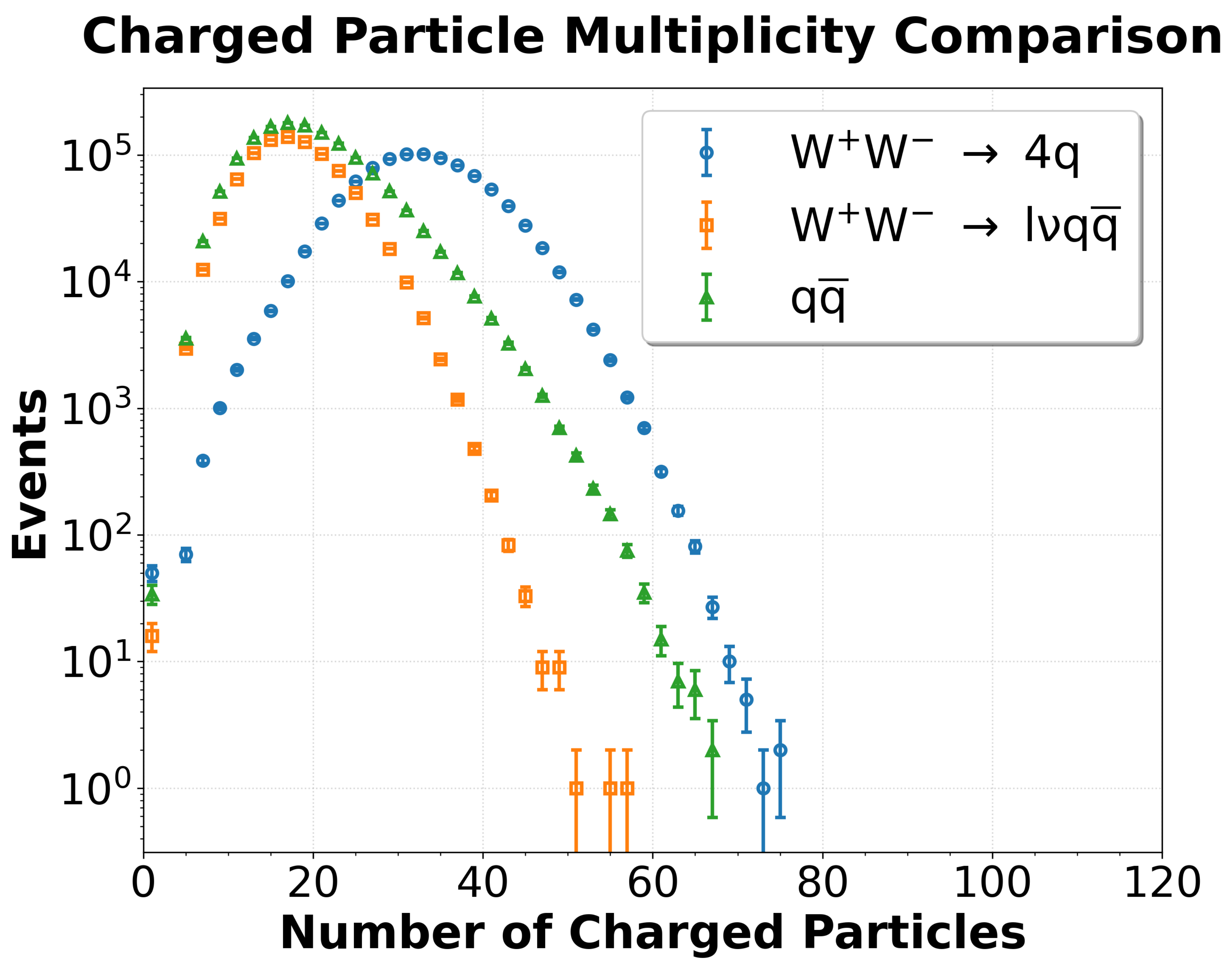
- Inclusive-multiplicity results show excellent agreement between archived **data** and **MC simulation**
- **Long-range near side correlation** signal appears at high multiplicity!
- Very tight upper limits with low multiplicity (NTrk < 40) (**LEP1**, **BELLE**, **LEP2**), for first time nonzero associated yield seen at high multiplicity (NTrk > 50)

Yu-Chen "Janice" Chen, Tzu-An Sheng, Yen-Jie Lee (MIT)

Future Works with Enhanced WW Purity

Boosted decision tree to enhance the WW two string vs one string (and other) component. These works open new door for LEP analyses





Yu-Chen “Janice” Chen, Tzu-An Sheng, Yen-Jie Lee (MIT)

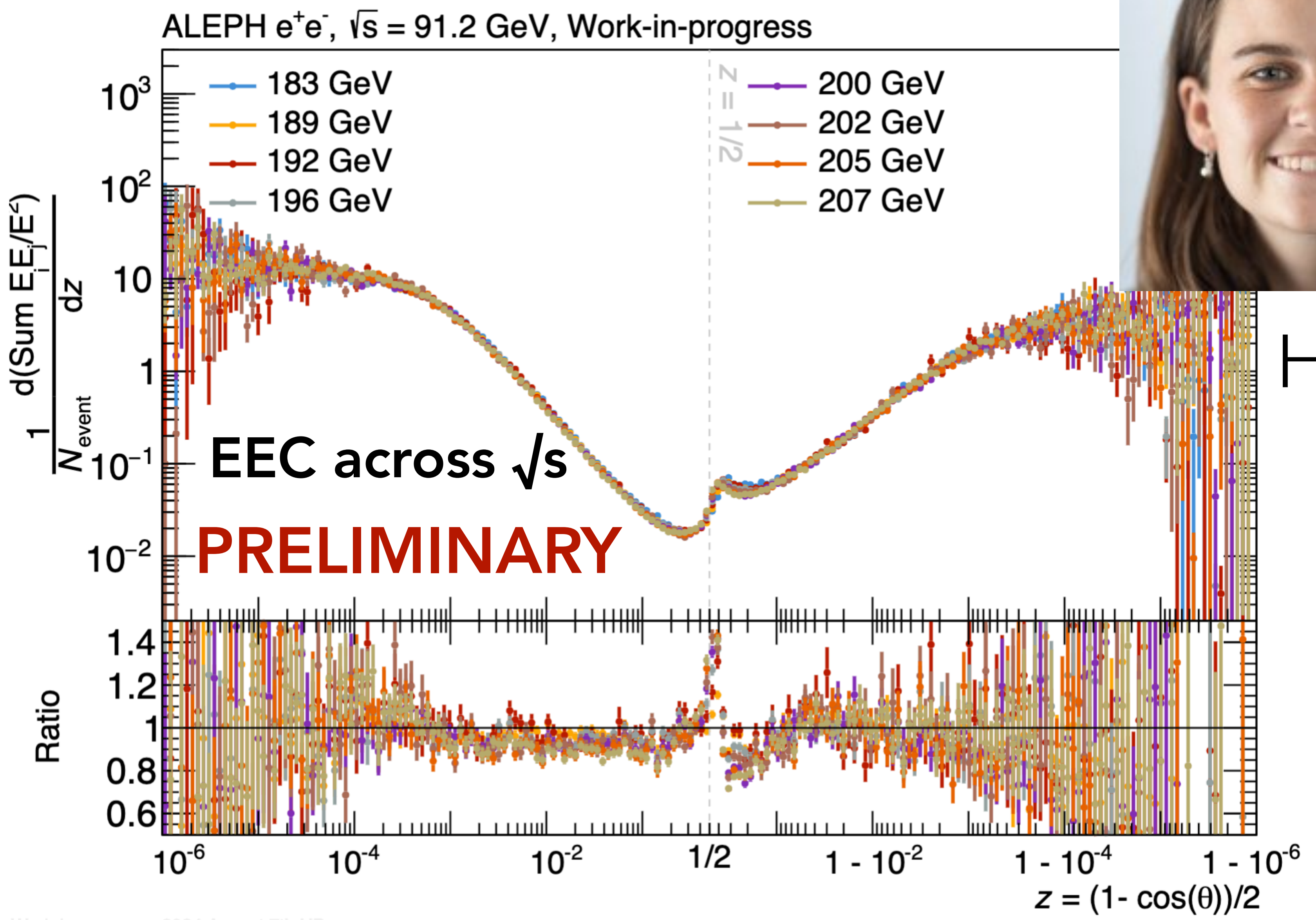
Ever Growing List of Exciting Physics Ideas

Non-exhaustive list of examples:

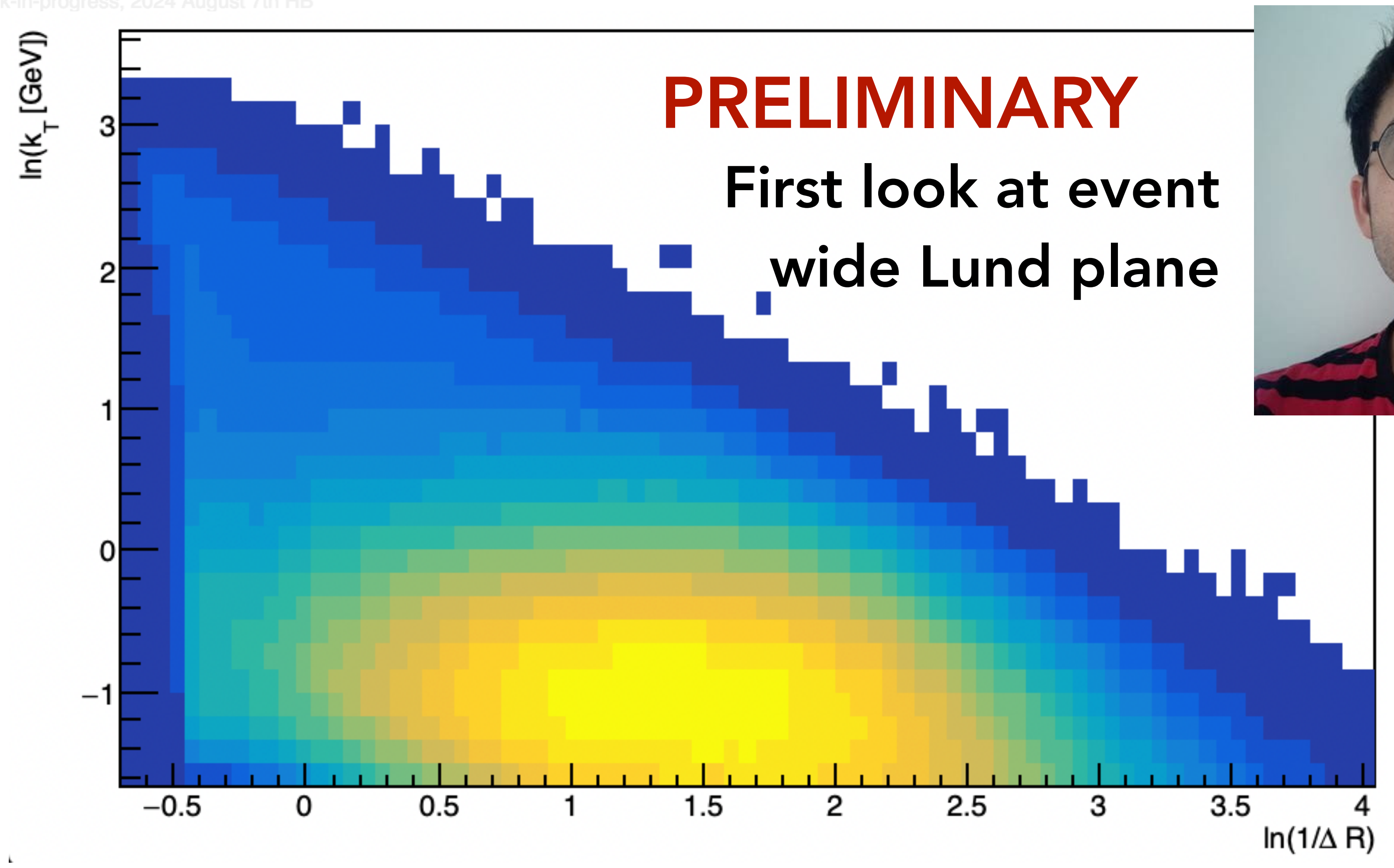
- All measurements across √s from 1-200 GeV
- Thrust with varied p vs E-scheme
- Unbinned multi-differential (e.g. τ, C, HJM)
- Unbinned full phase space
- Lund Plane Observables
- Modern HF tagging and observables
- anti-kT Jet measurements
- …

Excited for ideas beyond precision QCD:

- Revisit $Z \to b\bar{b}$ forward-backward asymmetry
- … 🤔😁🤨

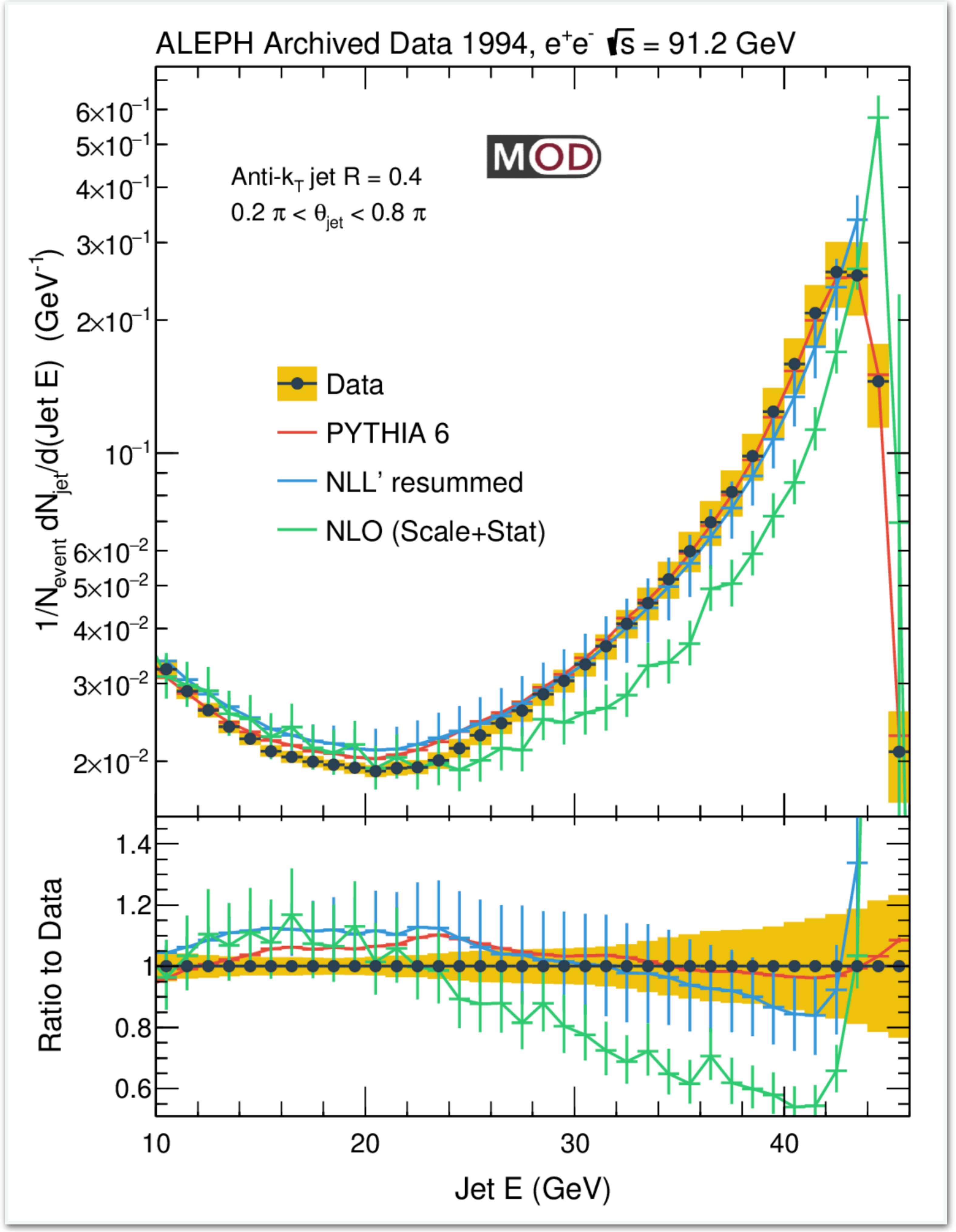


Hannah Bossi

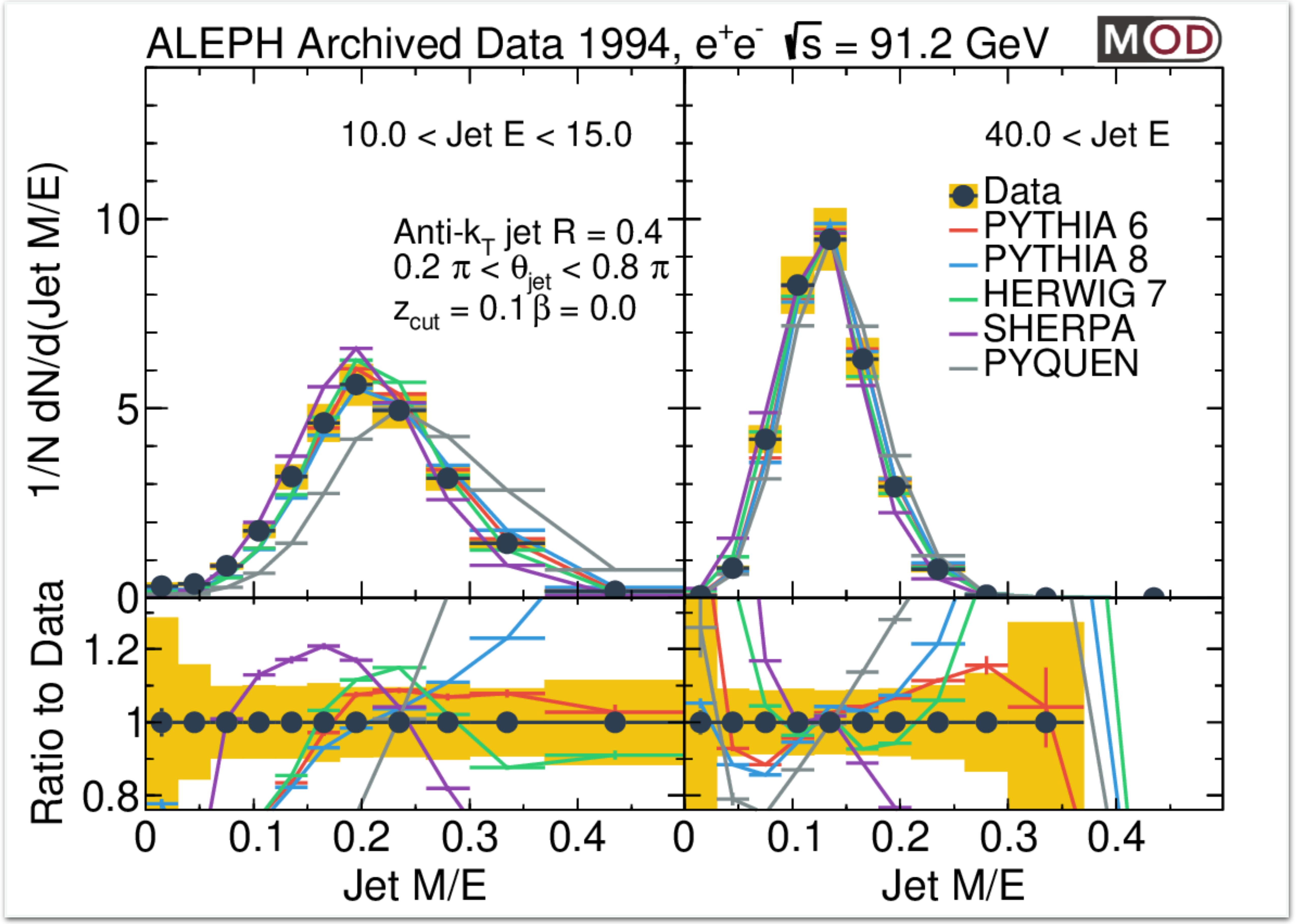


Cristian Barrera

Emerging Studies into Modern Anti-kT Jets



anti-k_T jet spectra



anti-k_T jet mass

Credit: Luna Chen (Vanderbilt) et al.

ALEPH Thrust Backup

Corrections to the Spectrum

(1) Hadronic event selection:

- High efficiency, but removal of hadronic efficient
- Removed fraction corrected for via unfolding

(2) Tracking efficiency:

- Efficiency to reconstruct charged particles accounted for in unfolding

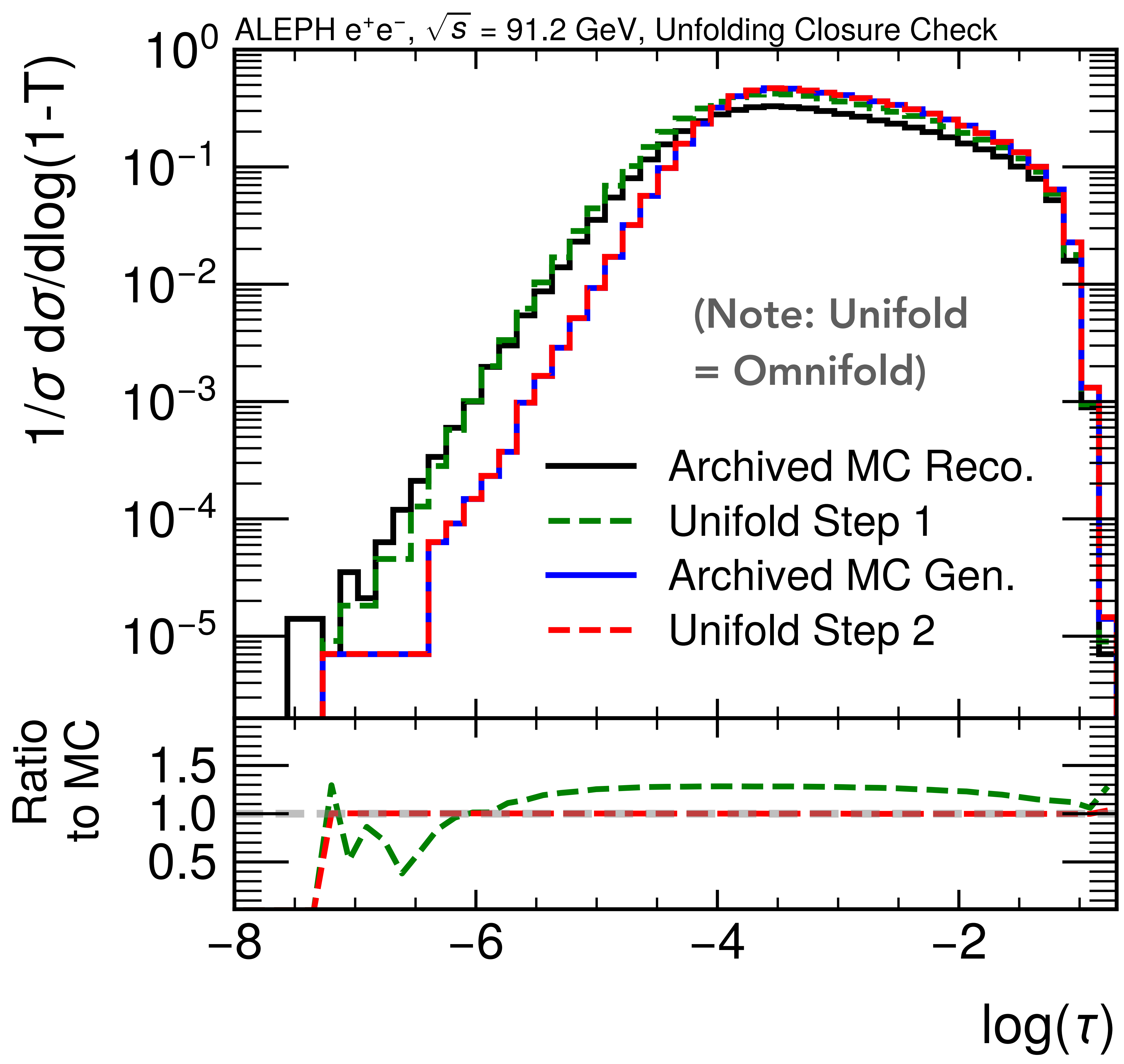
(3) EM radiation effects (ISR/FSR):

- e^+e^- from γ radiation → Trace gen/reco history for subsequent e^+e^- close in (θ,ϕ), Remove the e^+e^-
- Odd set of neutrals along beam pipe in MC → Likely radiated photons, Removed those particles

(4) Unbinned unfolding with OmniFold:

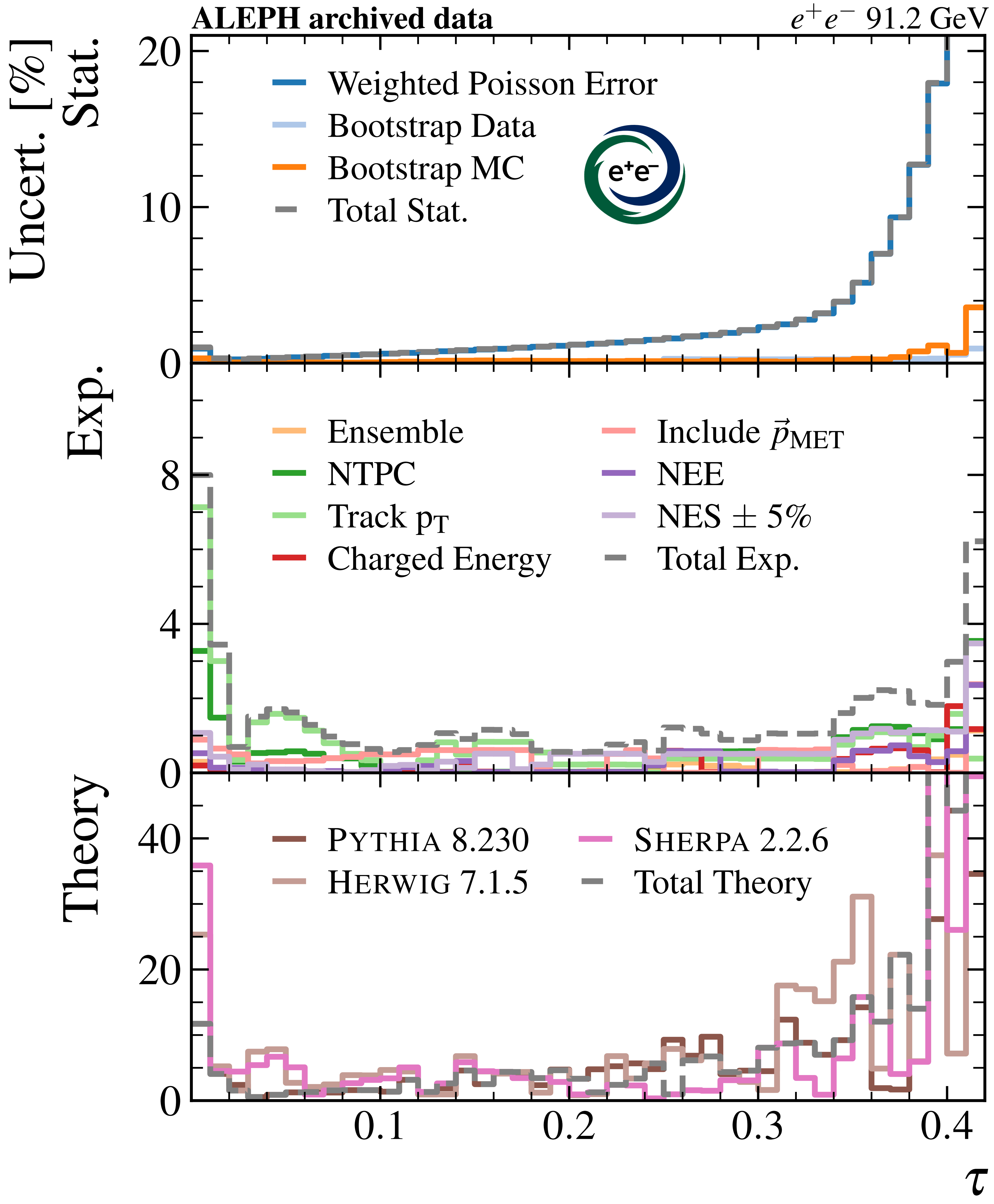
- Implicitly handle (1) - (3)
- Correct for detector effects

Successful closure check on MC



Uncertainties

- **Statistical uncertainty:**
  - Poisson uncertainty on final weights
  - Bootstrapping to assess impact on unfolding
- **Experimental uncertainty:**
  - Variations in the selections
  - Variation of neutral particle energy and efficiency
  - Ensembling from NN random initialization
- **Theory MC prior uncertainty:**
  - Max spread from variation of the MC prior in unfolding. Accomplished by reweighting at particle level archived Pythia 6 to modern MC's



Check Thrust Calculation Algorithm

We checked the difference between:

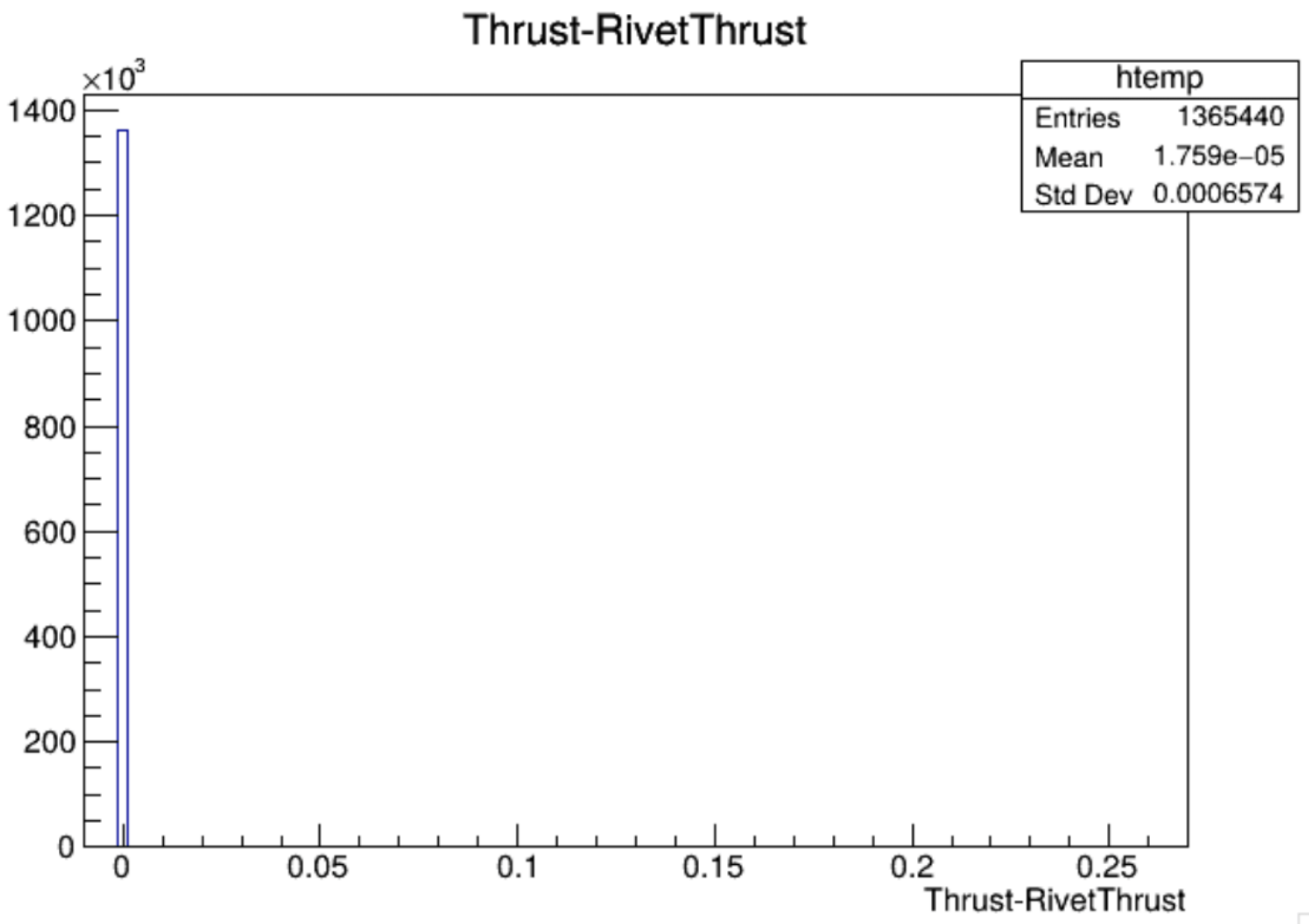
(1) ALEPH Rivet thrust algorithm [1]
(2) Exact thrust algorithm used in this paper

No significant difference between the two algorithms using archived data

Experimentally we tried to do the following variations:

(a) Include missing momentum vector in the experimental calculation
(b) Used only charged particle and unfold
in both cases, **we unfold to the same generator level definition (2)**.

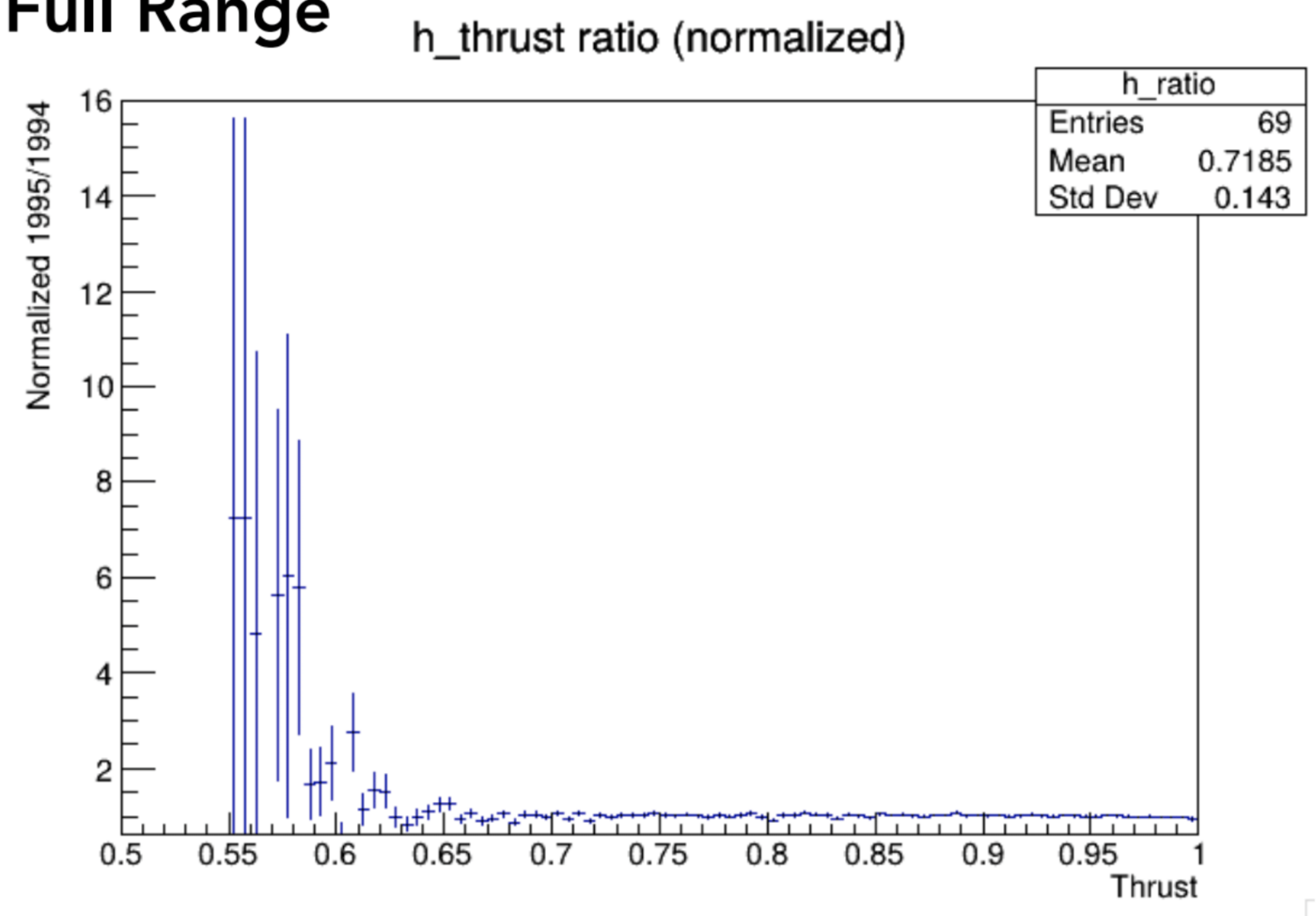
The difference is within quoted systematics.



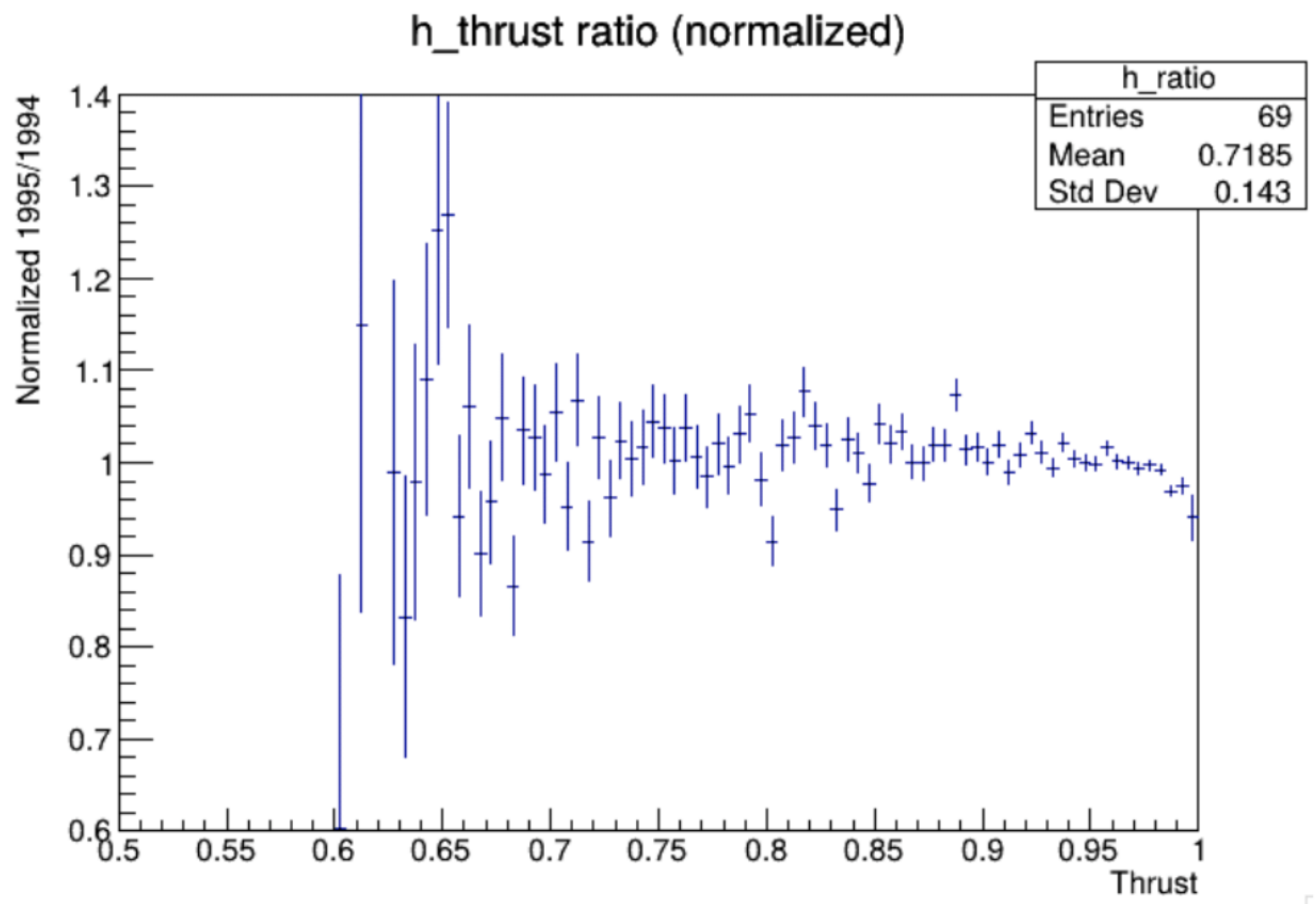
[1] Rivet Thrust code: https://rivet.hepforge.org/code/1.3.0/a00697.html

Thrust Raw Spectra in 1994 and 1995 Data

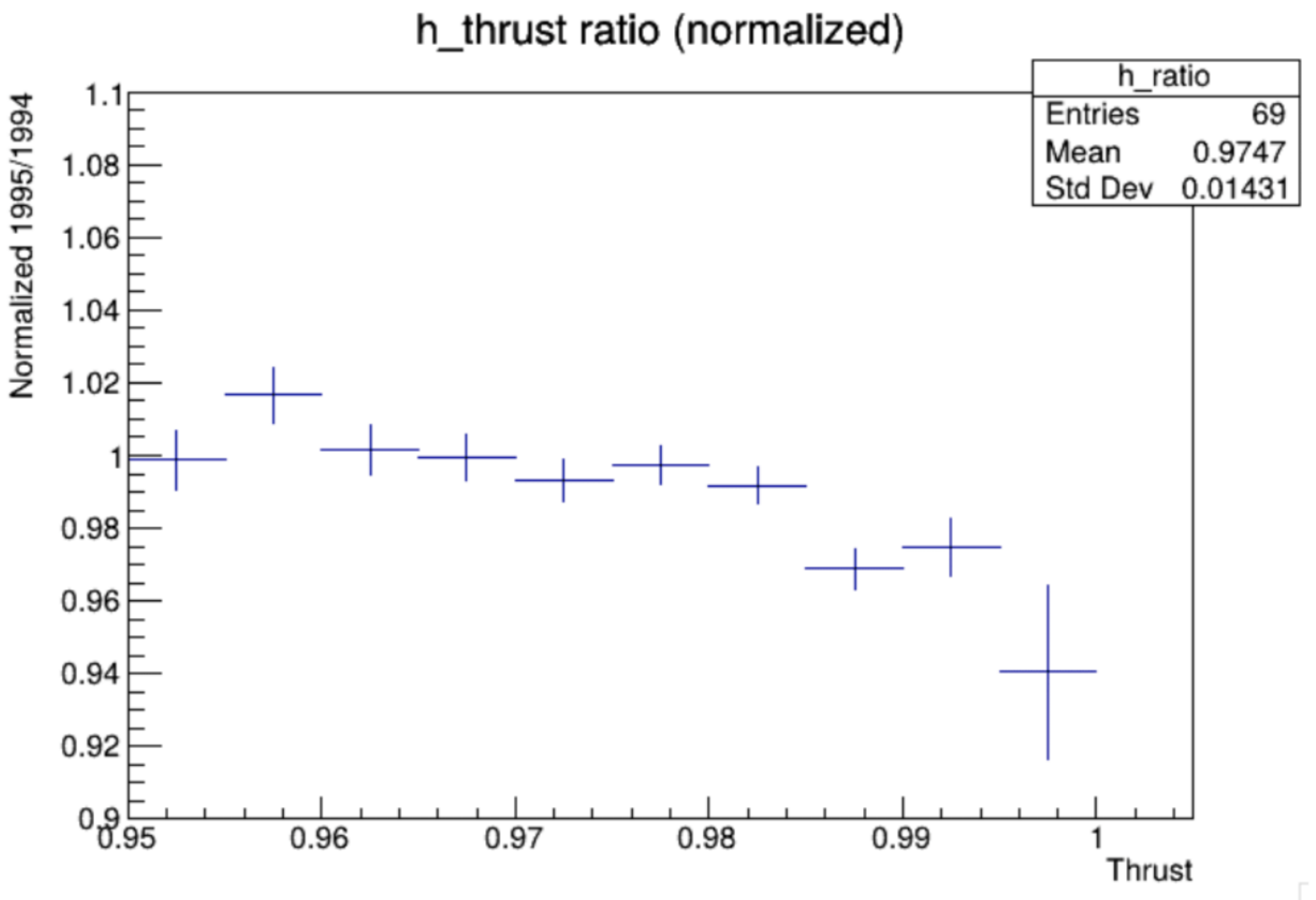
Full Range



Zoomed in around Y=1

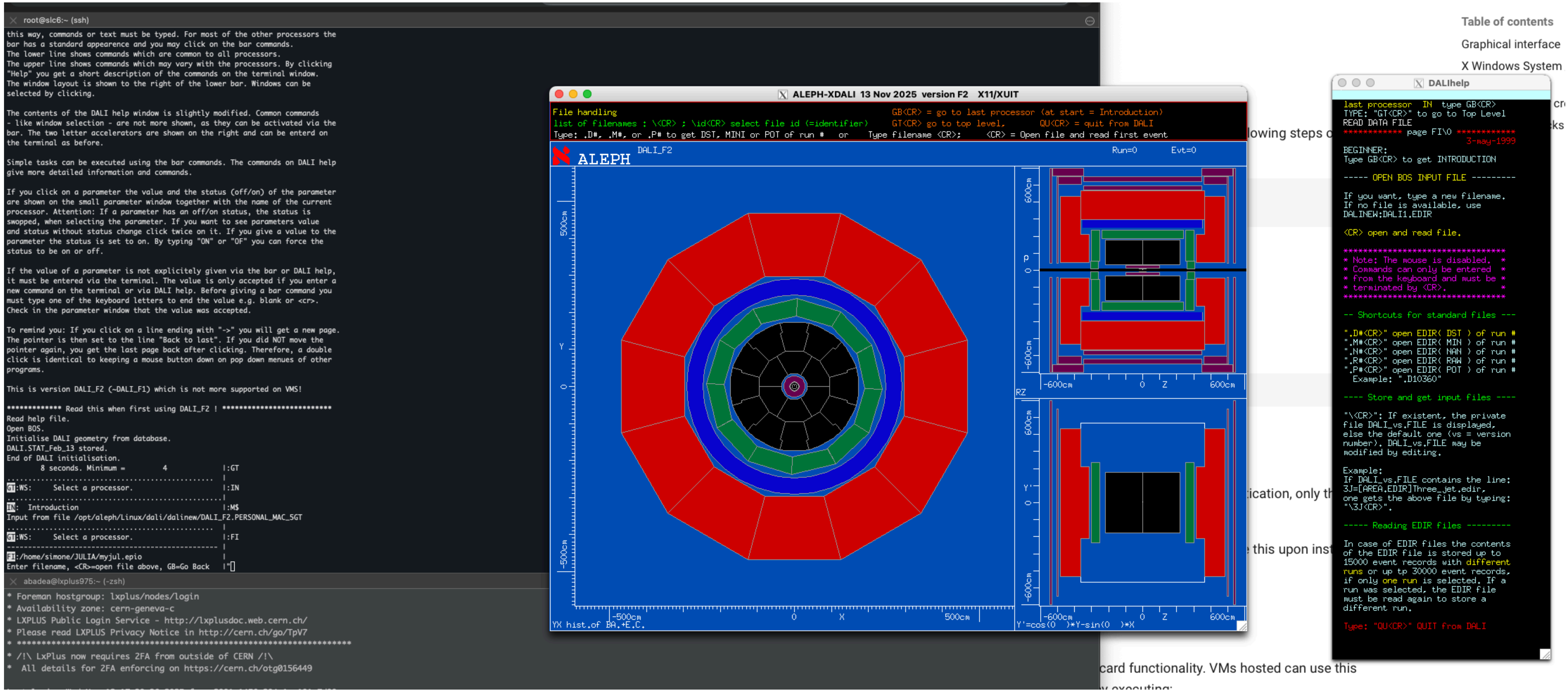


Zoomed in at T>0.95



Data only comparison at high thrust pulls down. This is the **opposite direction** as the ALEPH/OF ratio

Progress Towards More MC



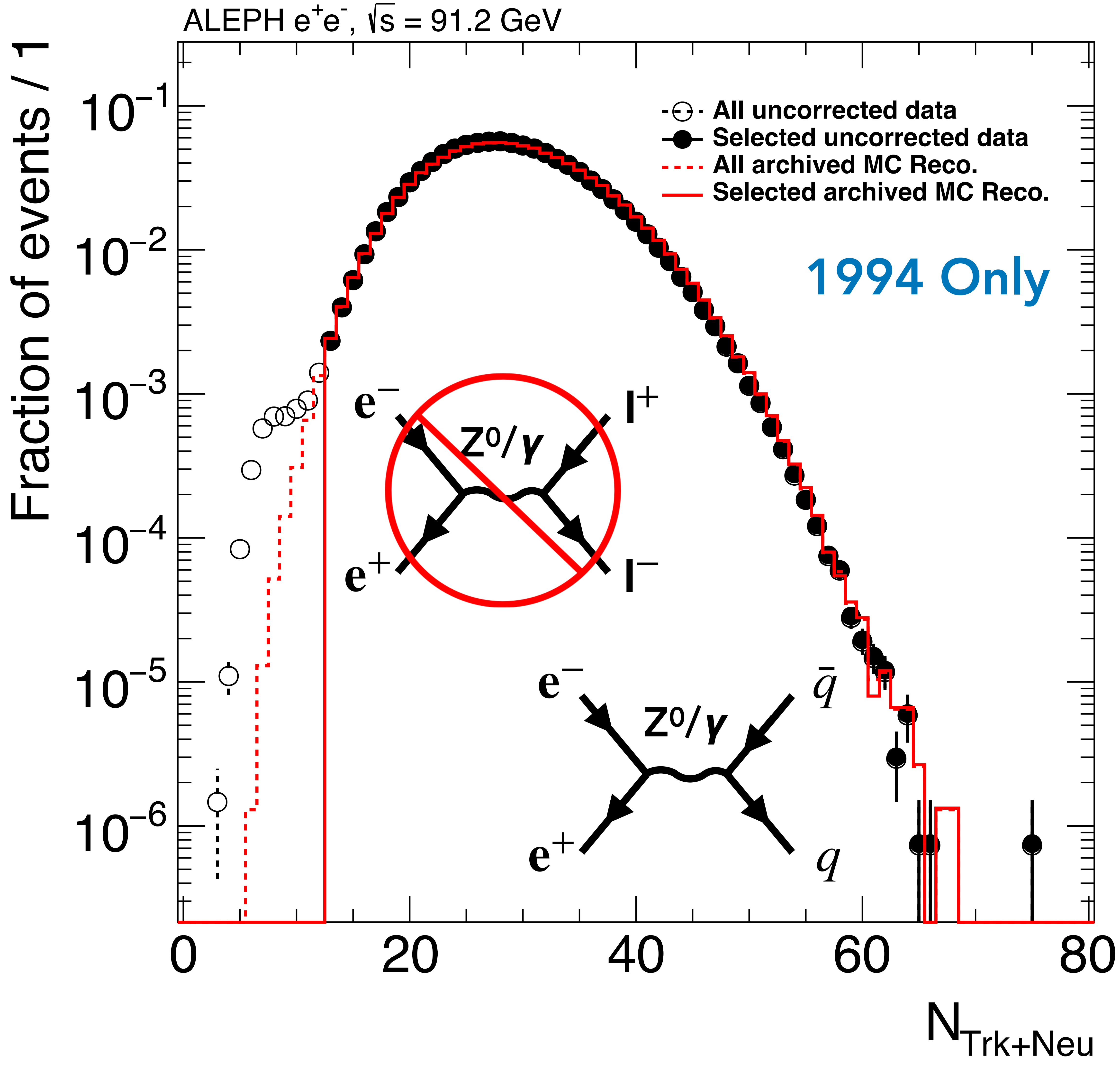
Full ALEPH software suite running in a VM! Working on understanding simulation flow

Hadronic Event Selection

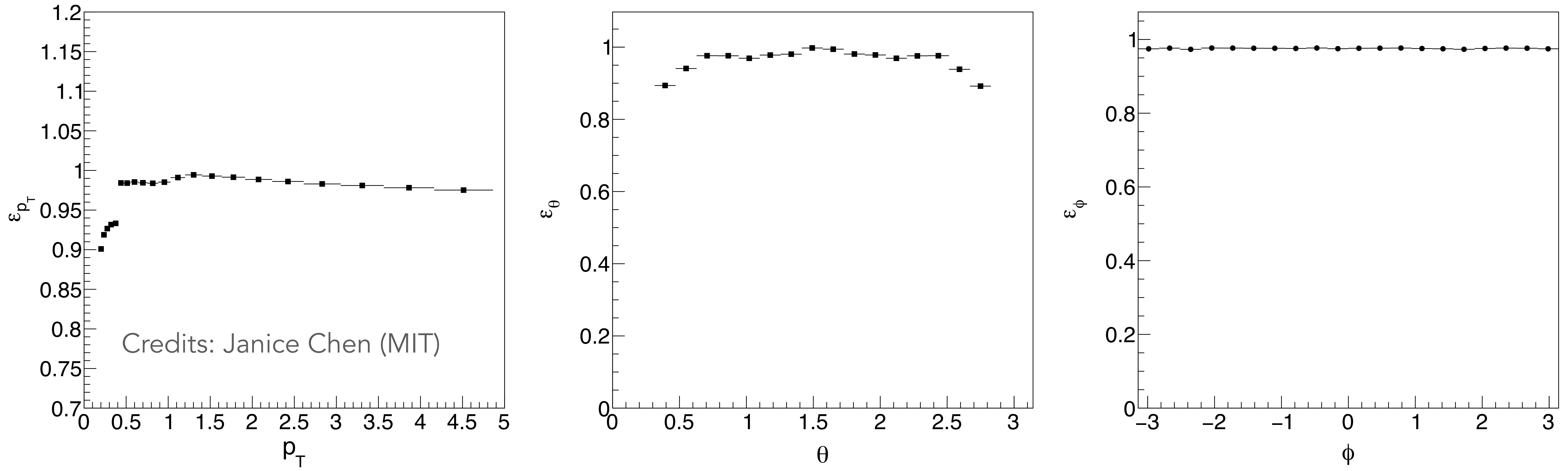
Hadronic selection removes dilepton final states

While highly efficient for hadronic final states, some fraction of hadronic events are removed

This removed fraction is corrected for during the unfolding procedure



Tracking Efficiency



Tracking efficiency typically applied in the past:

$$\varepsilon(p_T, \theta, \phi, N_{\mathrm{Trk}}^{\mathrm{Offline}}) = \left[\frac{d^3 N^{\mathrm{reco}}}{dp_T d\theta d\phi} \Big/ \frac{d^3 N^{\mathrm{gen}}}{dp_T d\theta d\phi} \right]_{N_{\mathrm{Trk}}^{\mathrm{Offline}}}$$

Here, the efficiency correction is applied implicitly within unbinned unfolding procedure

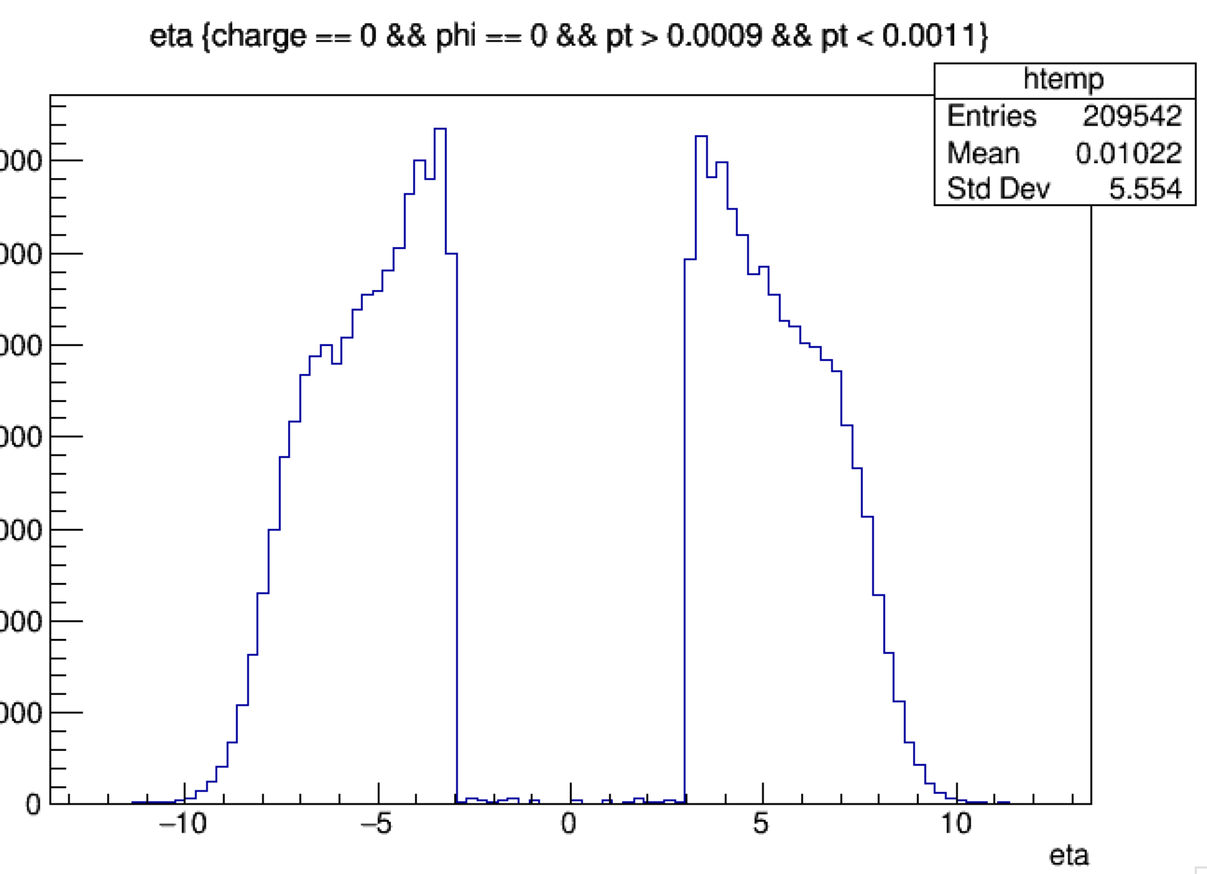
N_{trk} range	Fraction of data (%)	$\langle N_{\mathrm{trk}} \rangle$	$\langle N_{\mathrm{trk}}^{\mathrm{corr}} \rangle$
[5, 10)	3.1	8.2	8.9
[10, 20)	59.2	15.2	15.8
[20, 30)	34.6	23.1	23.4
[30, ∞)	3.1	32.4	32.6
[35, ∞)	0.5	36.9	37.2

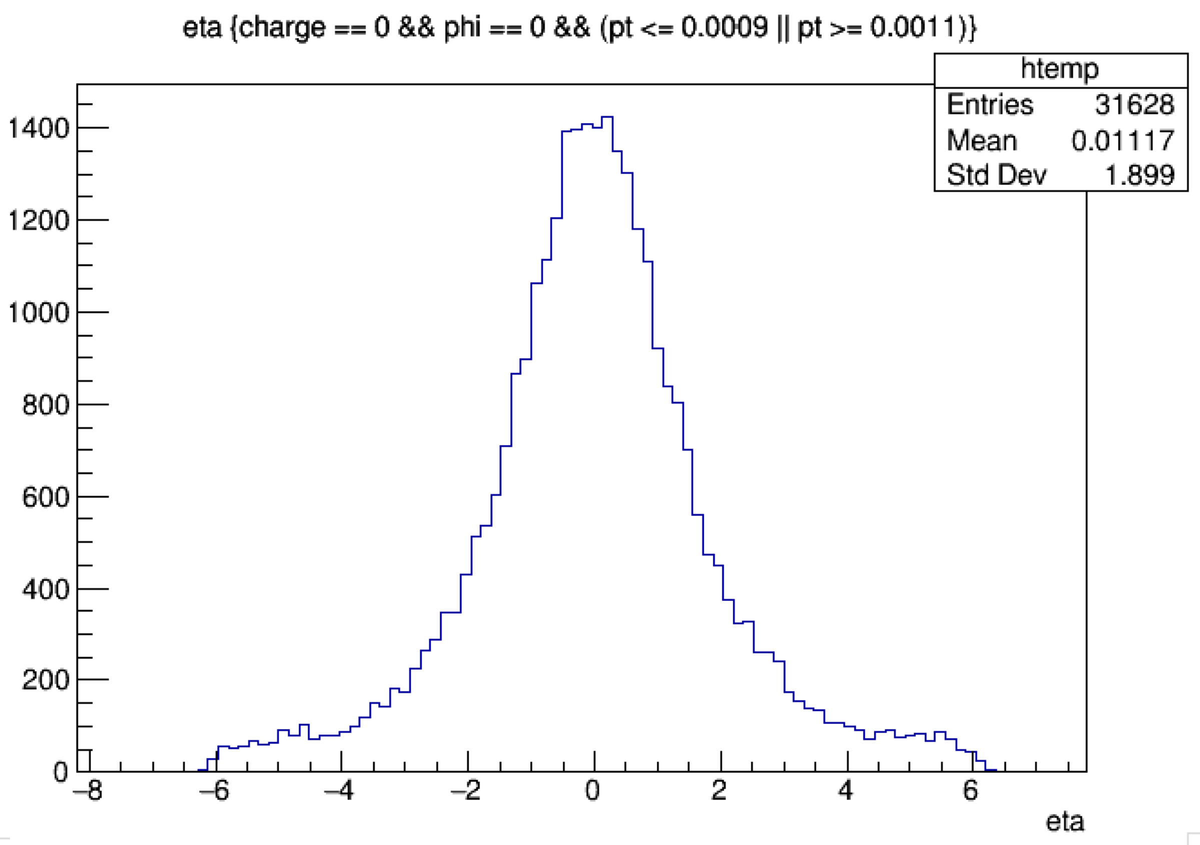
Phys. Rev. Lett. 123, 212002 (2019)

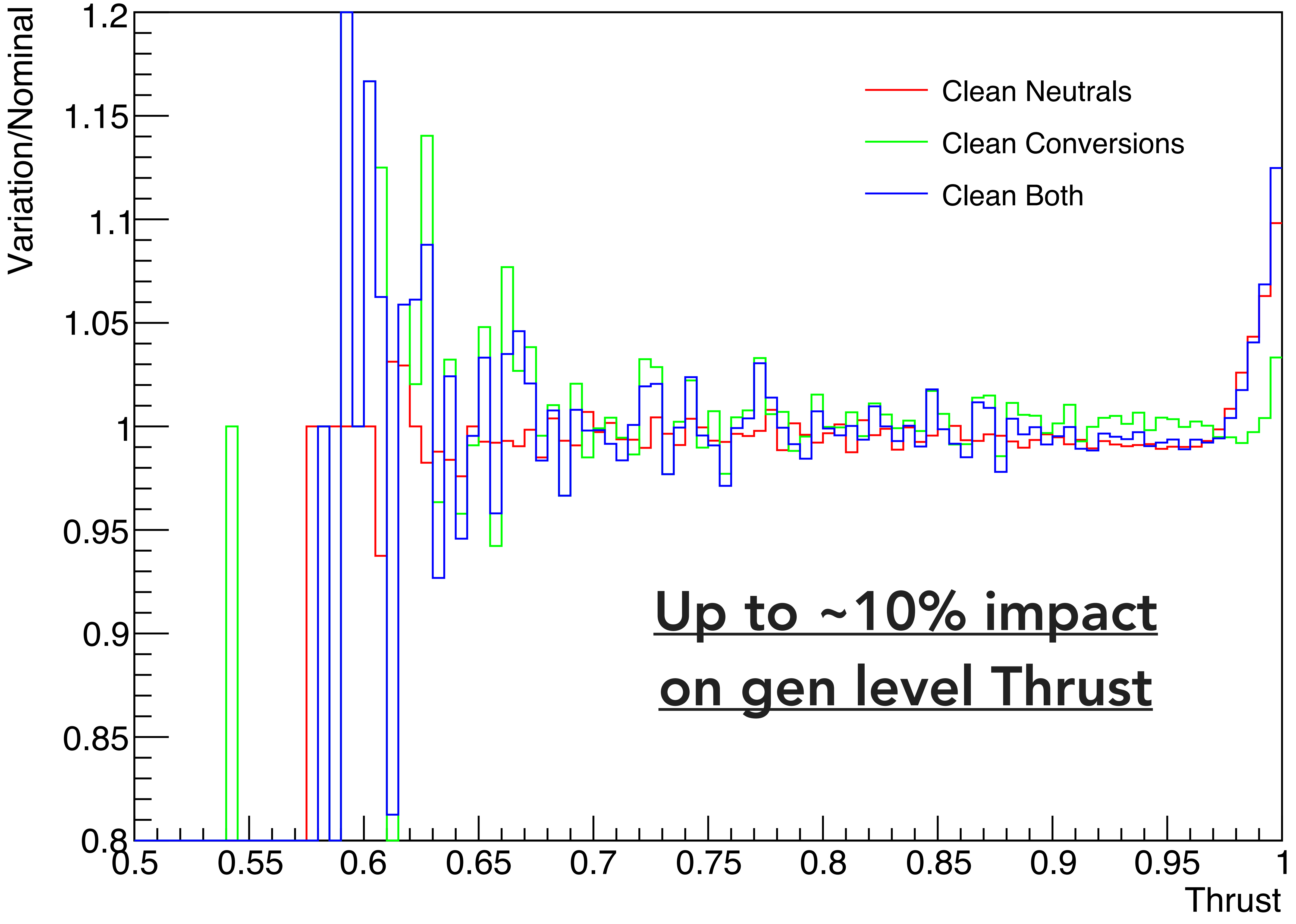
Electromagnetic (EM) Radiation Effects

Final state e⁺e⁻ pair from γ radiation (EM ISR/FSR) → Trace gen/reco history for subsequent e⁺e⁻ close in (θ,ϕ), Remove the e⁺e⁻ pair

Odd set of neutrals along beam pipe in MC → Likely radiated photons, Removed those particles



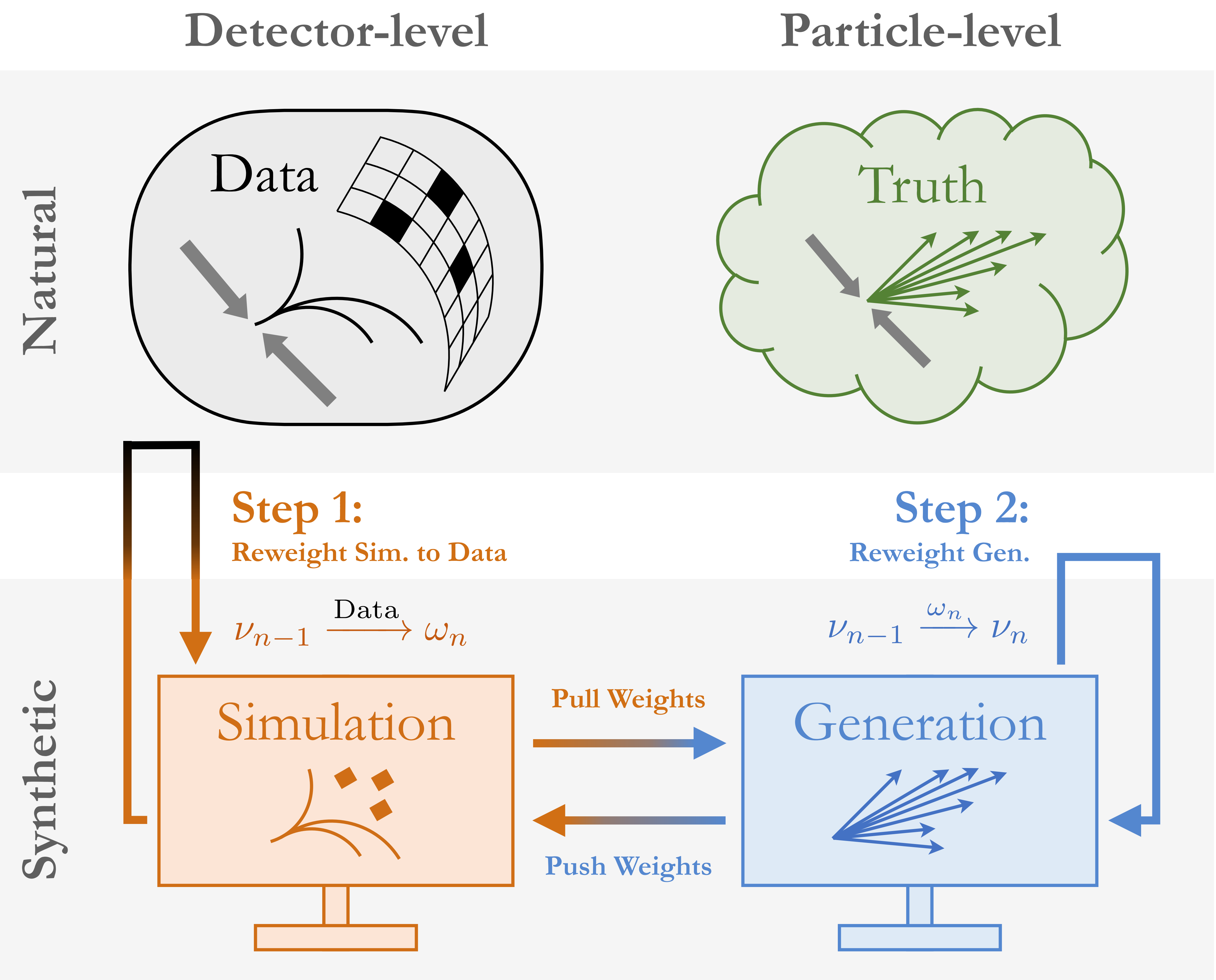




Correcting the Spectrum to Particle Level

- Implicit corrections and detector effects accounted for with unbinned unfolding algorithm OmniFold
- Applied to single observable $\log \tau$

As a high-level reminder, conceptually ~ unbinned iterative bayesian unfolding but very different procedure

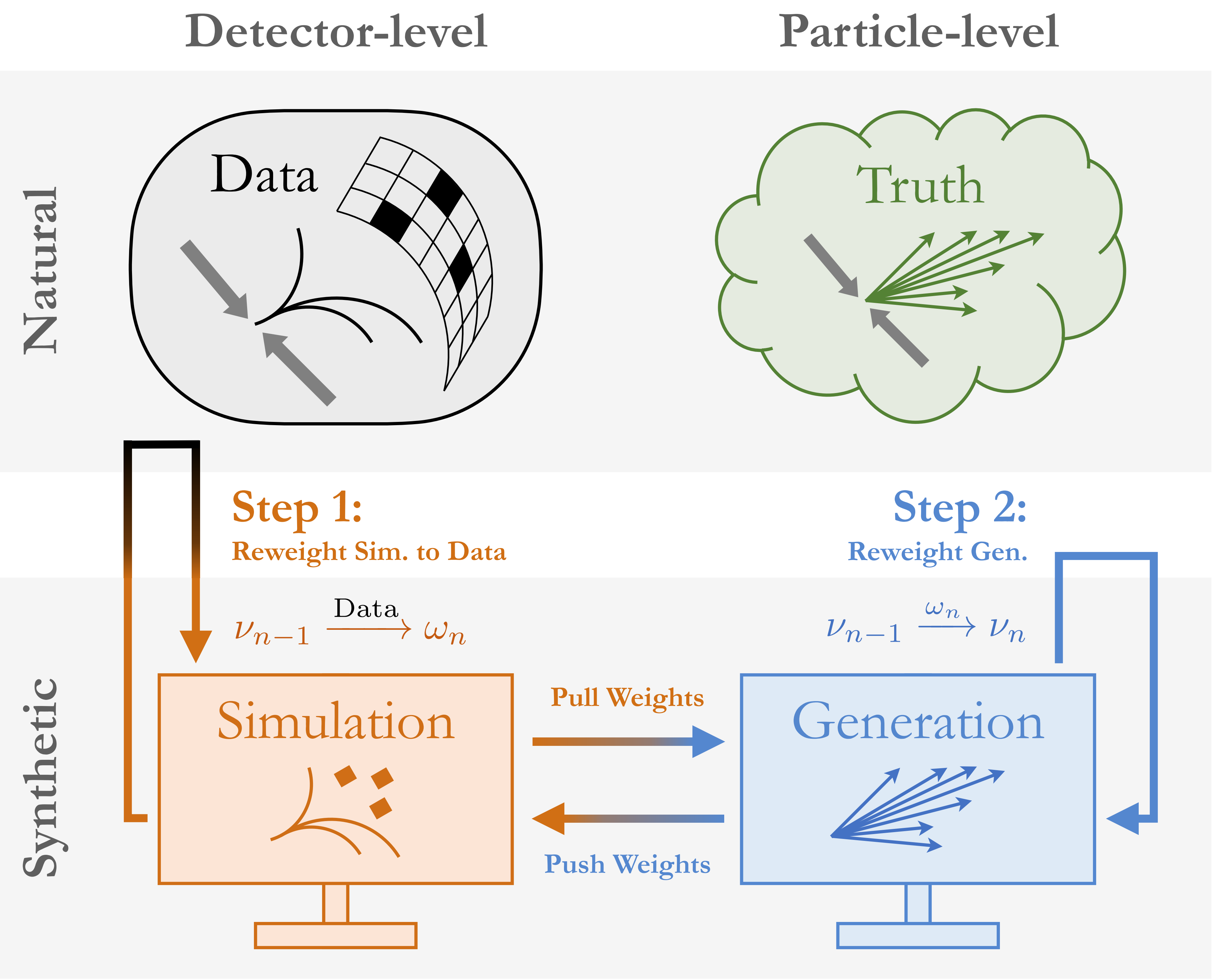


Andreassen et al. Phys. Rev. Lett. 124, 182001

How to Unfold with AI, CERN Courier, B. Nachman

OmniFold Algorithm

Sequential reweighting using neural networks (NN) to approximate likelihood ratios



Andreassen et al. Phys. Rev. Lett. 124, 182001

Step 1 (NN) $\omega_n(m) = \nu_{n-1}^{\text{push}}(m)\, L\left[(1,\text{Data}), \left(\nu_{n-1}^{\text{push}}, \text{Sim.}\right)\right](m)$

Step 2 (NN) $\nu_n(t) = \nu_{n-1}(t)\, L\left[\left(\omega_n^{\text{pull}}, \text{Gen.}\right), \left(\nu_{n-1}, \text{Gen.}\right)\right](t)$

where L is the likelihood ratio $L[(w,X),(w',X')](x) = \dfrac{p_{(w,X)}(x)}{p_{(w',X')}(x)}$

sim/gen matching used for $\nu_n^{\text{push}}(m) = \nu_n(t)$, $\omega_n^{\text{pull}}(t) = \omega_n(m)$

and the gen prior is $\nu_0(t)$

$$p^{(n)}_{\text{unfolded}}(t) = \nu_n(t) p_{\text{Gen.}}(t)$$

$$= \int dm\, p_{\text{data}}(m) \frac{p_{\text{Sim.|Gen.}}(m \mid t)\nu_{n-1}(t) p_{\text{Gen.}}(t)}{\int dt' p_{\text{Sim.|Gen.}}(m \mid t')\nu_{n-1}(t') p(t')}$$

$$\sim \sum_i m_i \frac{R_{ij} t_j}{\sum_k R_{ik} t_k}$$

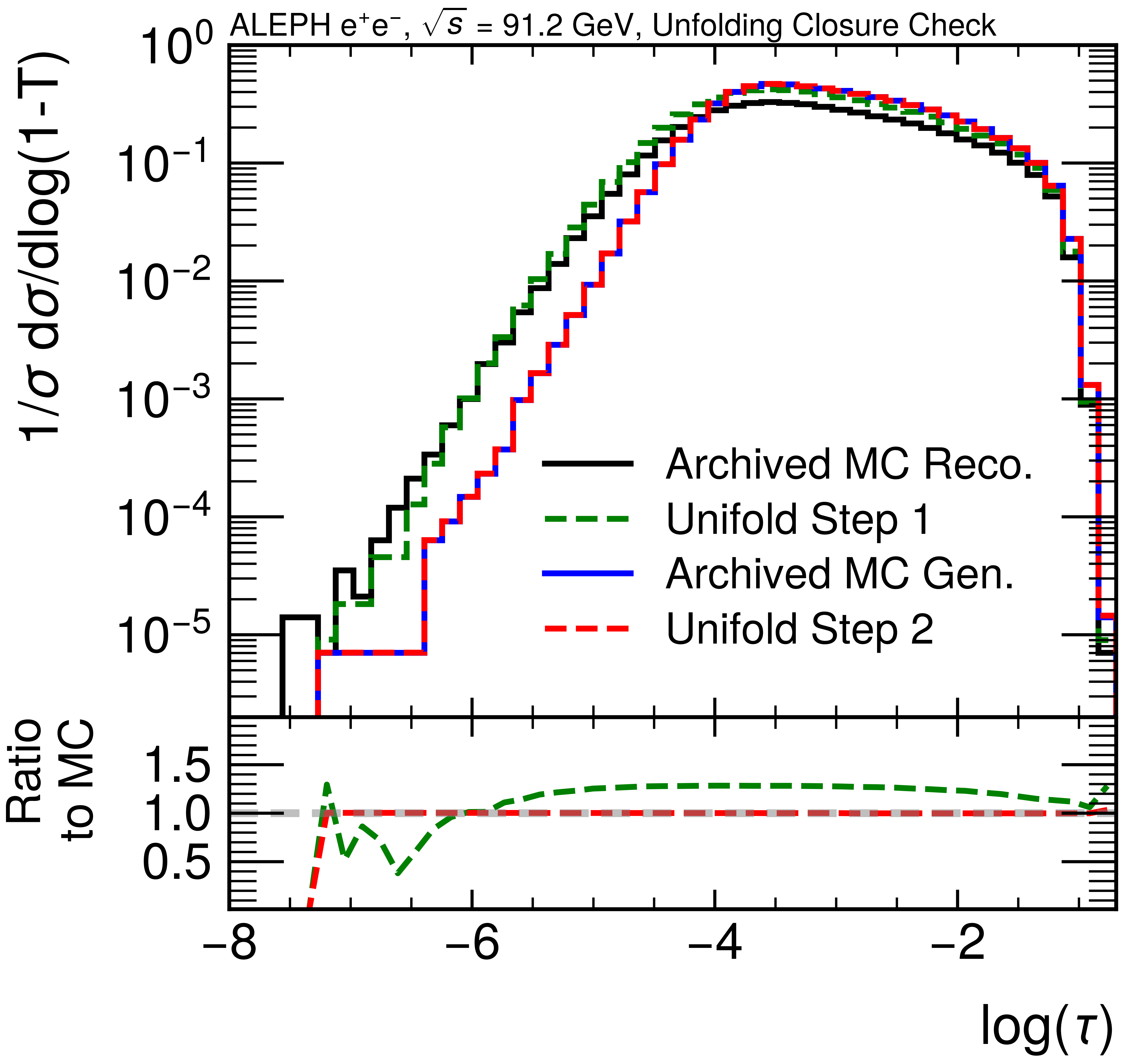
**** Unbinned IBU ****

Correcting the Spectrum to Particle Level

Successful closure check on MC. Replace data with MC reco, and run the full chain

Verify that the final result (red) converges to the gen level MC (blue) ✅

(Note: Unifold = Omnifold)

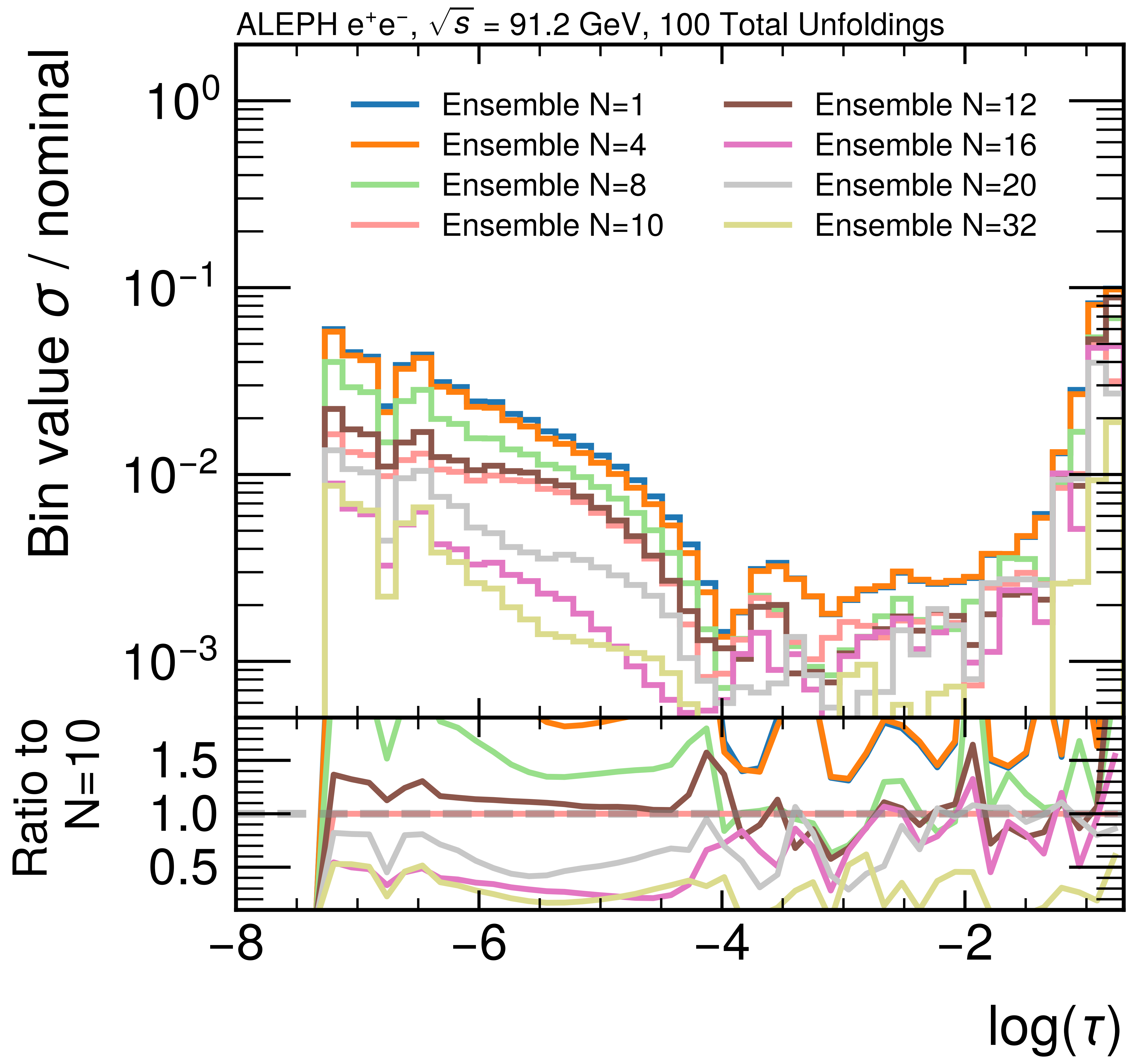


OmniFold Ensembling

NN output varies upon retraining by change in random seed→ Ensemble NN's to constrain effect to subdominant level

Example measuring the bin-by-bin spread over ensembles of different sizes

Converged on 1 "unfolding" = 100 training. Repeat for every variation



Experimental Uncertainties

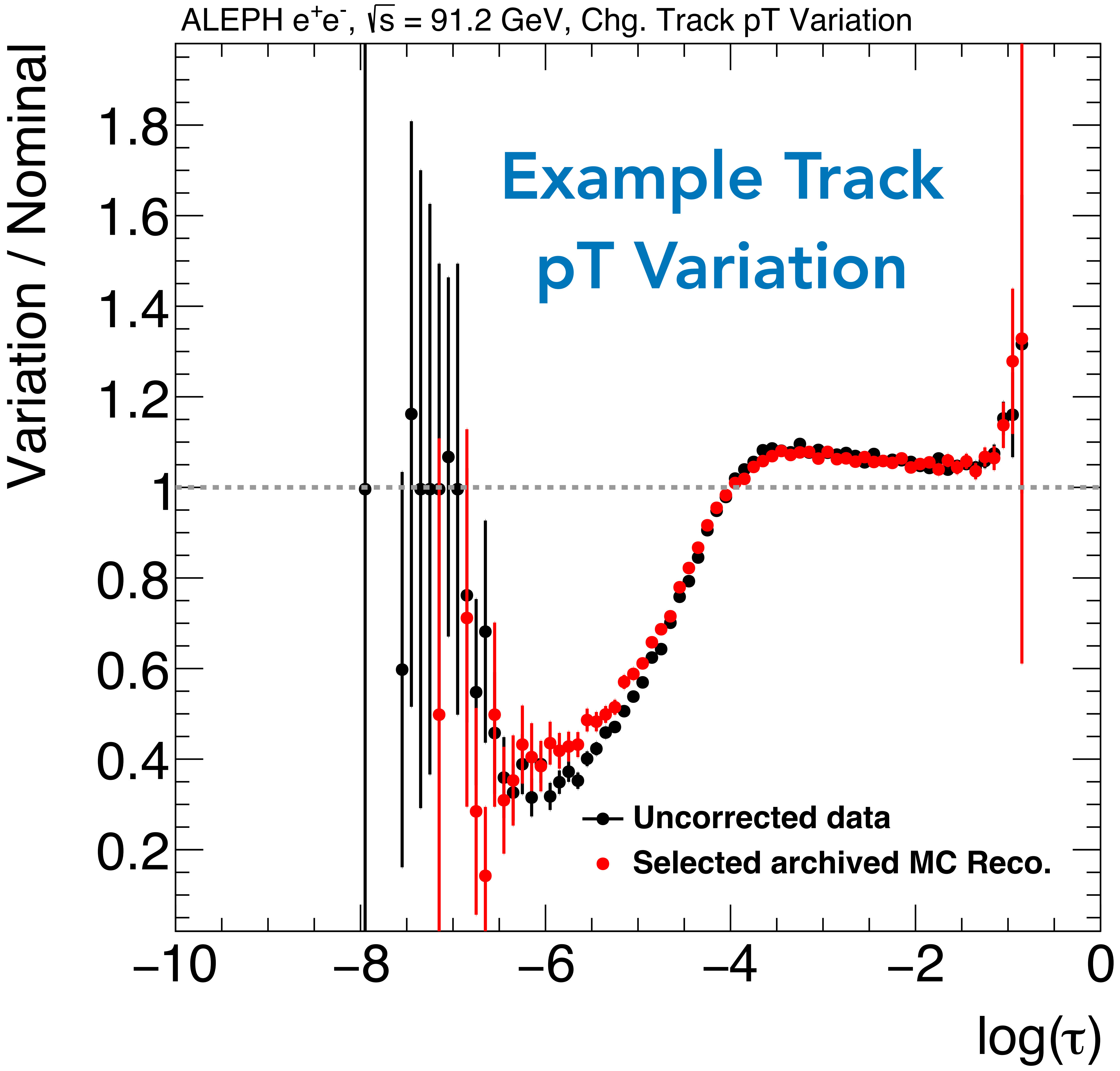
Vary selections, repeat unfolding, and measure change in thrust.

Standard ALEPH variation:

- Track NTPC ≥ 4 to 7
- **Track pT ≥ 0.2 to 0.4 GeV**
- $E_{charged}$ ≥ 15 GeV to 10 GeV

Custom variation:

- Neutral particle energy scale (NES)
- Neutral particle efficiency (NEE)
- Thrust w/ MET vector



Neutral Particle Variations

Vary selections, repeat unfolding, and measure change in thrust.

Standard ALEPH variation:

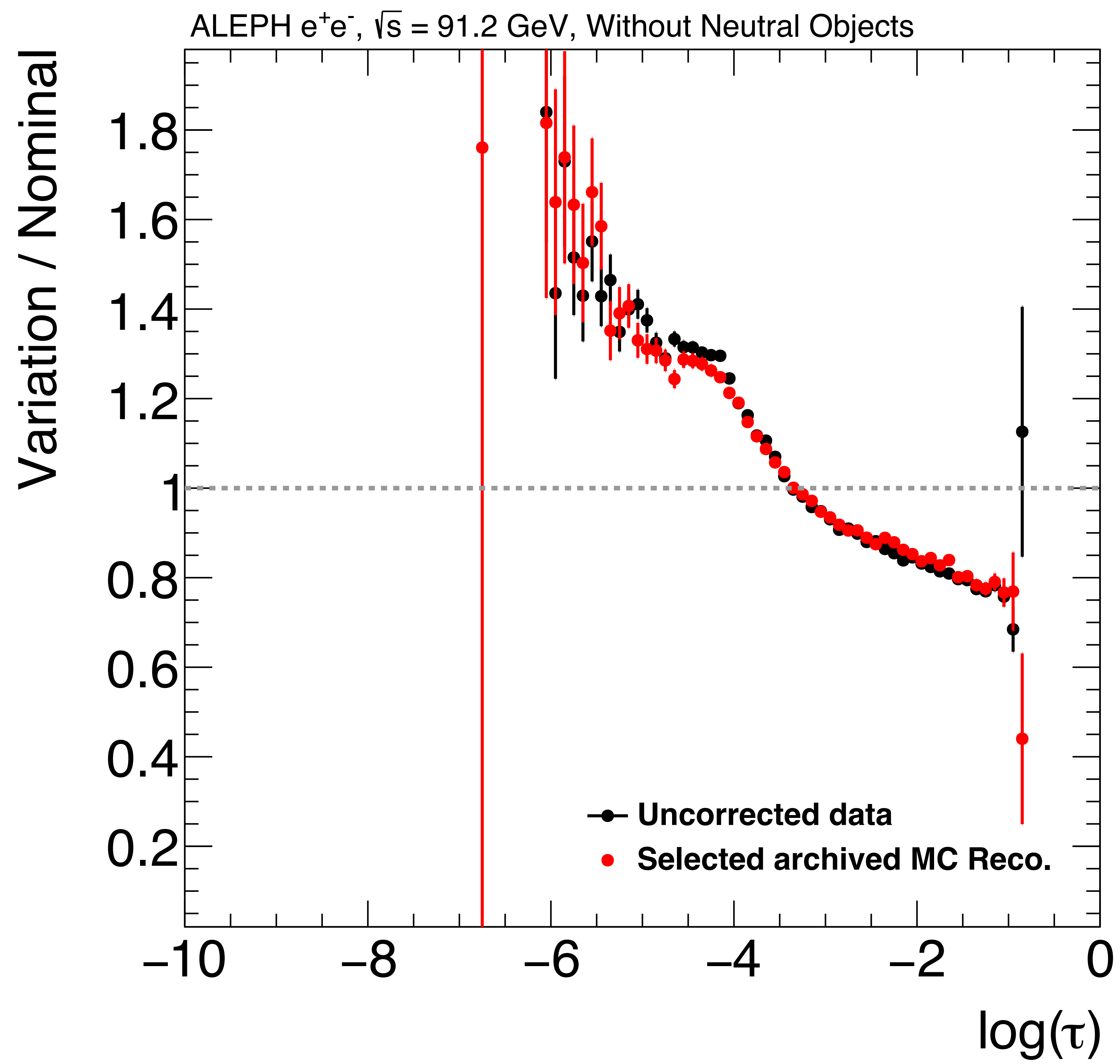
- Track NTPC ≥ 4 to 7
- Track pT ≥ 0.2 to 0.4 GeV
- $E_{charged}$ ≥ 15 GeV to 10 GeV

Custom variation:

- Neutral particle energy scale (NES)
- Neutral particle efficiency (NEE)
- Thrust w/ MET vector

Previous ALEPH result removed all 1-2 GeV neutrals to account for generator and detector mismodeling. Since then charged particle thrust created → **that systematic redefines observable**

→ Theory variation handles the generator mismodeling and we account for neutral particle energy scale and efficiency



Neutral Particle Energy Scale

Vary selections, repeat unfolding, and measure change in thrust.

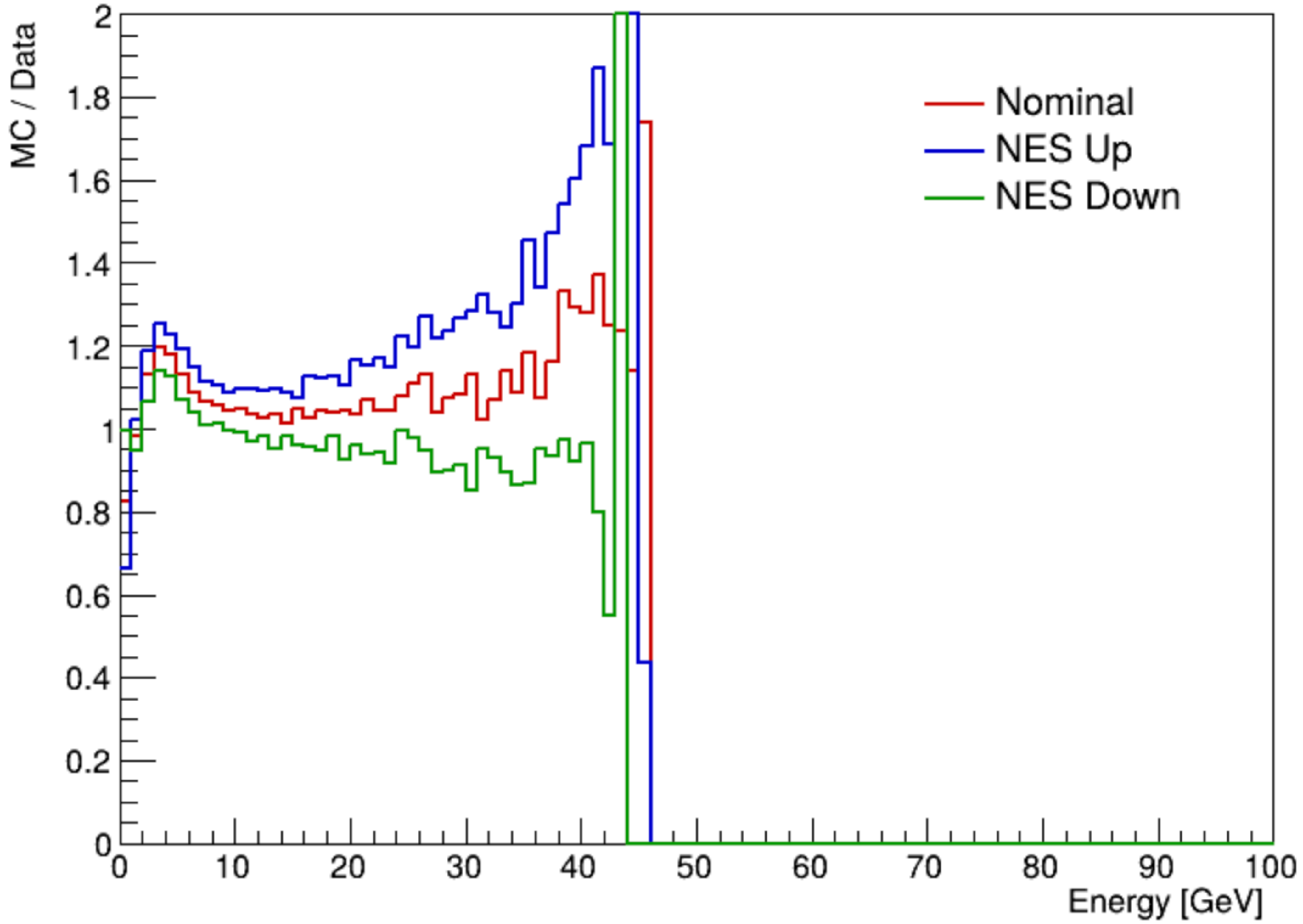
Standard ALEPH variation:

- Track NTPC ≥ 4 to 7
- Track pT ≥ 0.2 to 0.4 GeV
- $E_{charged}$ ≥ 15 GeV to 10 GeV

Custom variation:

- Neutral particle energy scale (NES)
- Neutral particle efficiency (NEE)
- Thrust w/ MET vector

Vary up/down the neutral particle energy spectrum before selections by 5%



2108.04877

Neutral Particle Efficiency

Vary selections, repeat unfolding, and measure change in thrust.

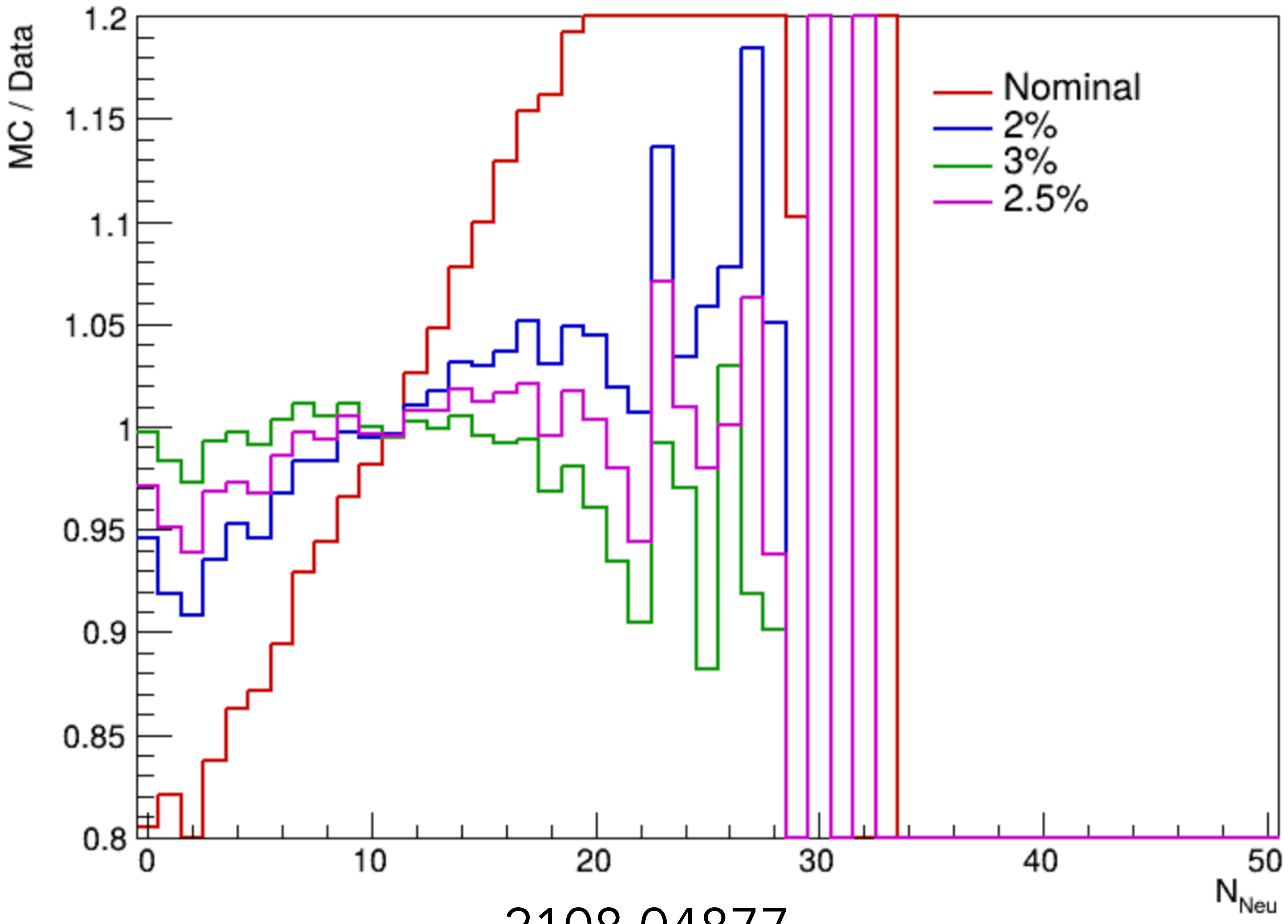
Standard ALEPH variation:

- Track NTPC ≥ 4 to 7
- Track pT ≥ 0.2 to 0.4 GeV
- $E_{charged}$ ≥ 15 GeV to 10 GeV

Custom variation:

- Neutral particle energy scale (NES)
- Neutral particle efficiency (NEE)
- Thrust w/ MET vector

Throw out neutral particles that pass selections with probability X% → **choose 2.5%**



2108.04877

Missing Momentum Vector Variation

Vary selections, repeat unfolding, and measure change in thrust.

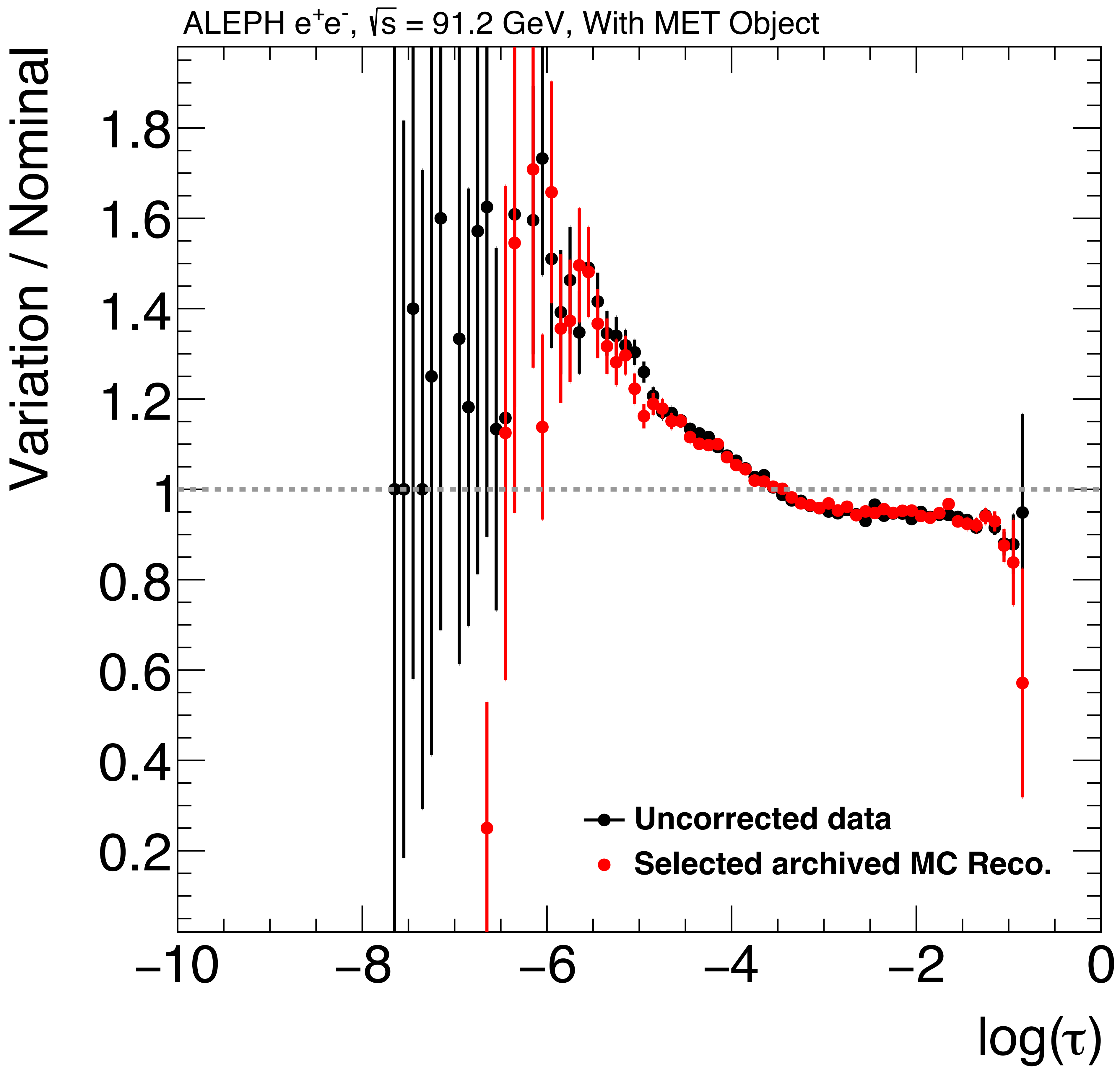
Standard ALEPH variation:

- Track NTPC ≥ 4 to 7
- Track pT ≥ 0.2 to 0.4 GeV
- $E_{charged}$ ≥ 15 GeV to 10 GeV

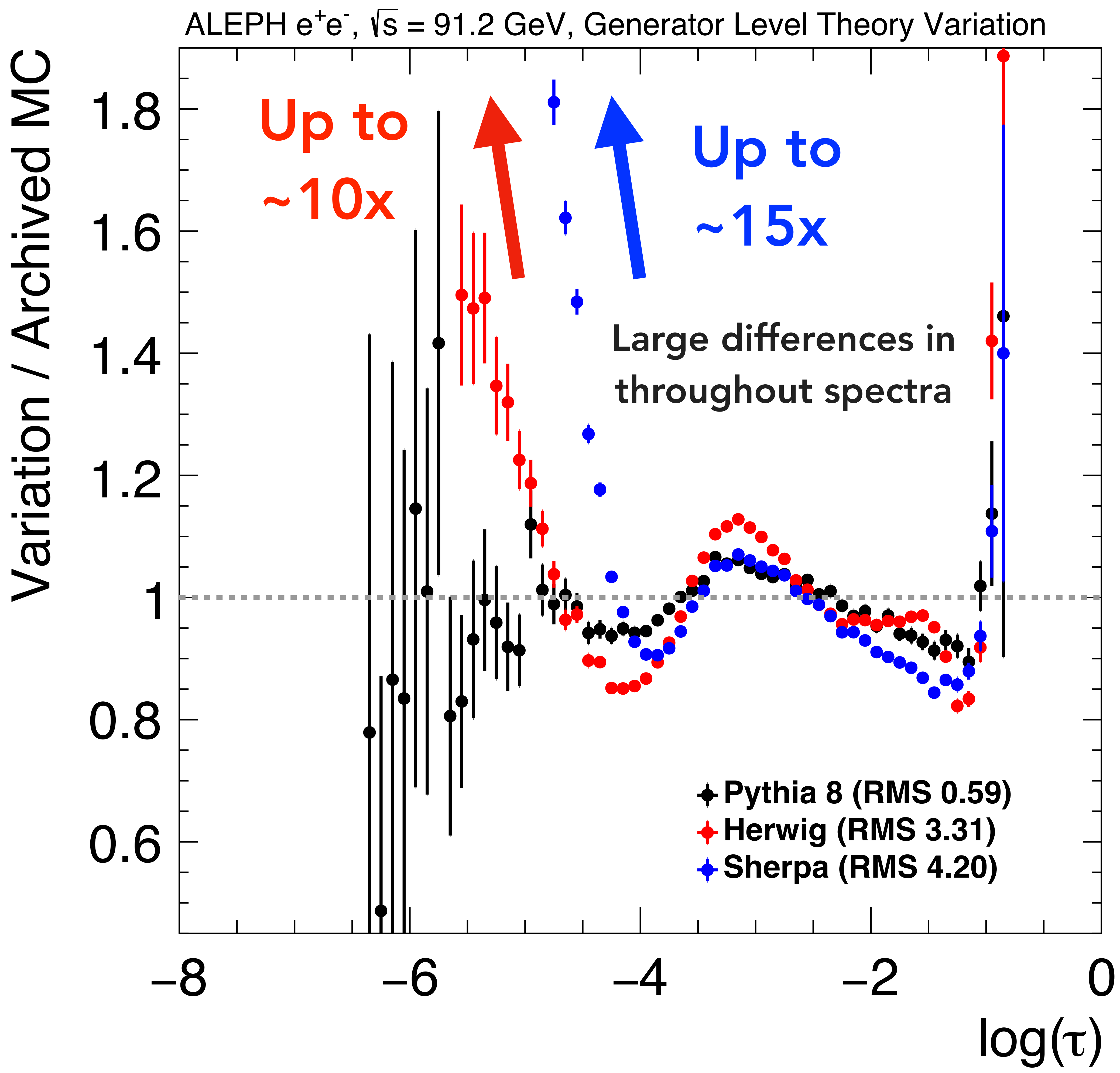
Custom variation:

- Neutral particle energy scale (NES)
- Neutral particle efficiency (NEE)
- Thrust w/ MET vector

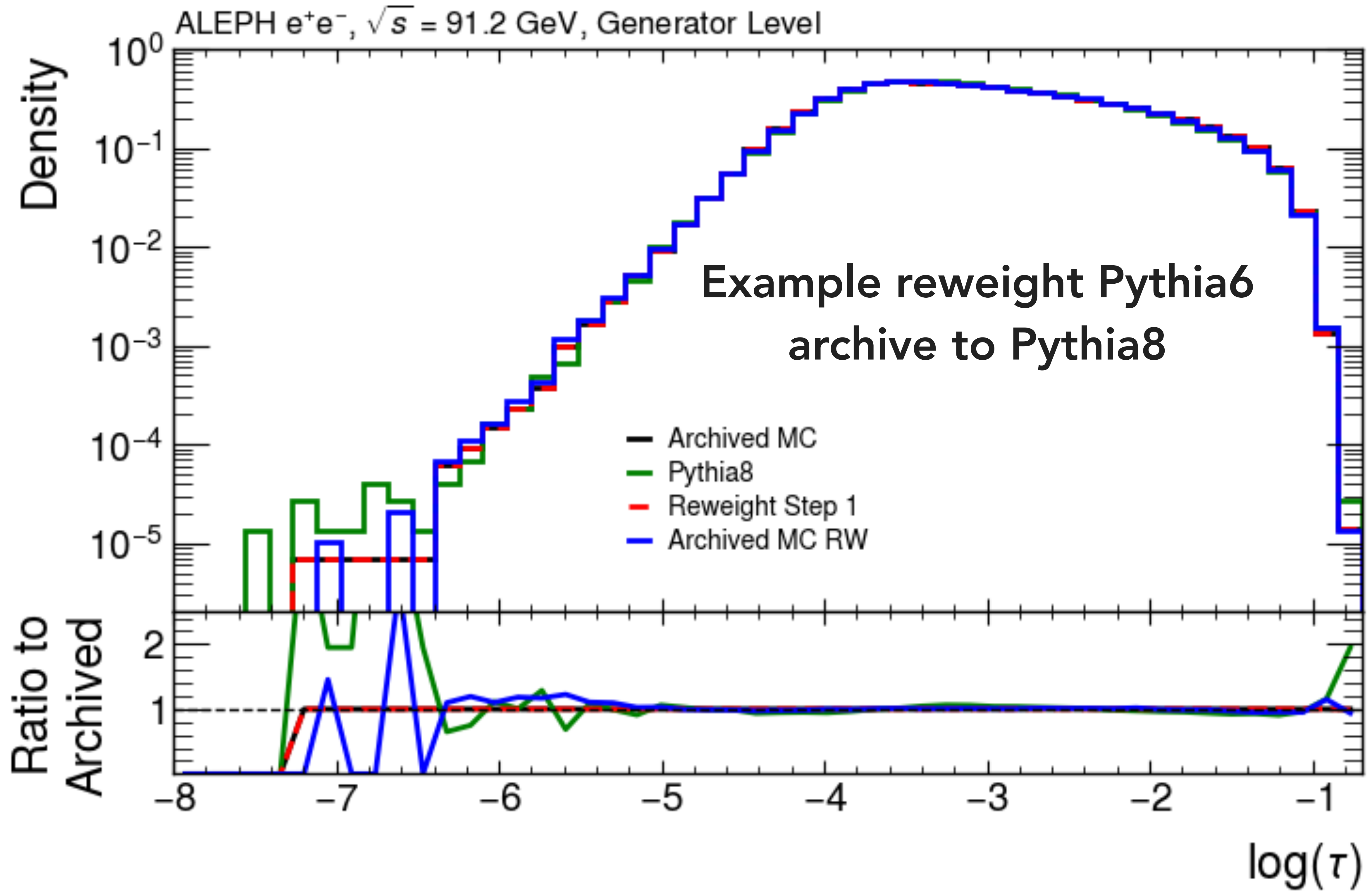
Include MET vector as particle in thrust calculation. Effect is well modeled in MC



Theory MC Prior Unfolding Uncertainty



Reweight gen. level Pythia 6 to Pythia 8, Herwig, Sherpa. Propagate weights to detector level. Repeat unfolding, measure difference w.r.t nominal

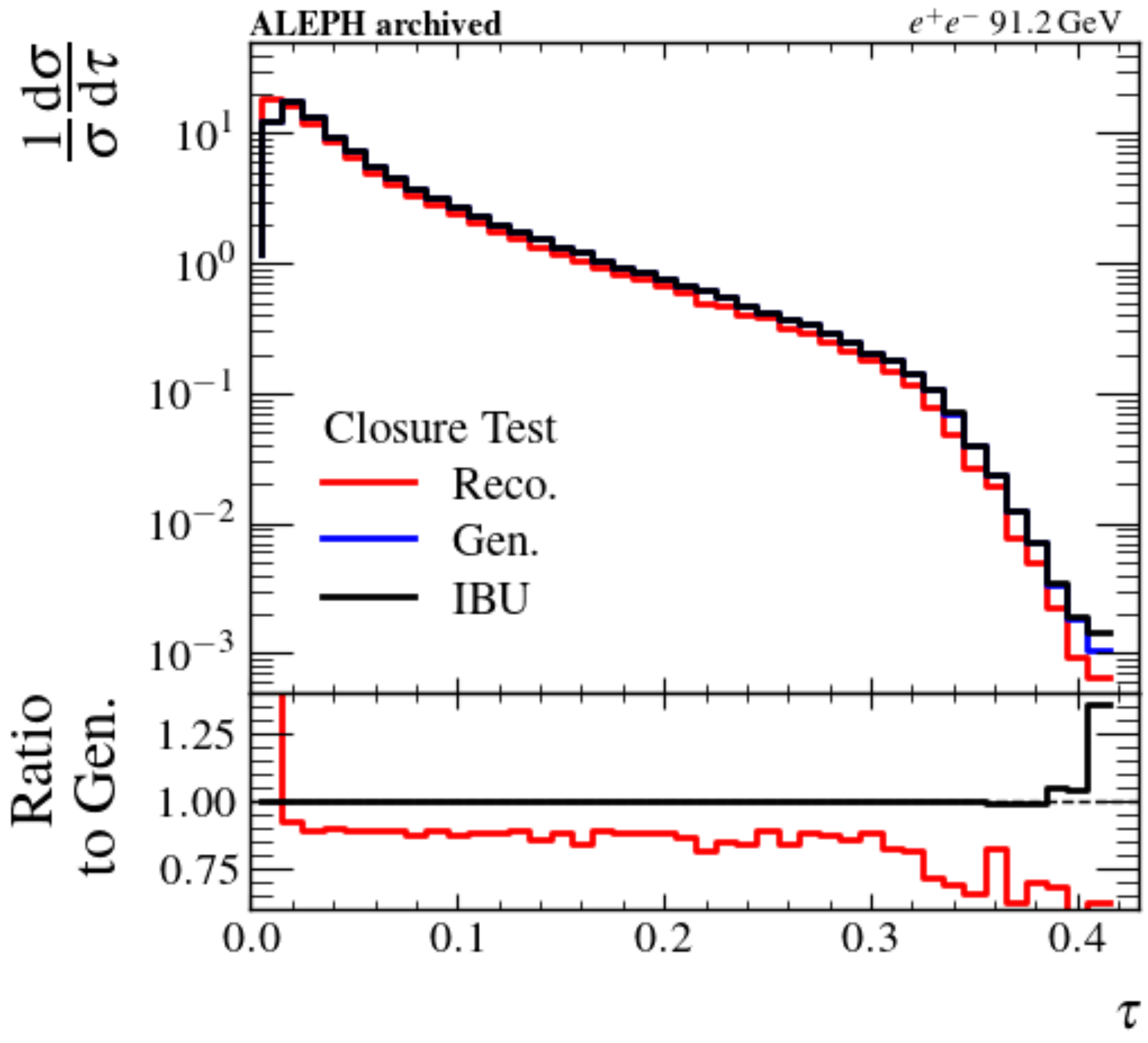


IBU Comparison

To assess the impact of OF, repeated with IBU. Notable other differences:

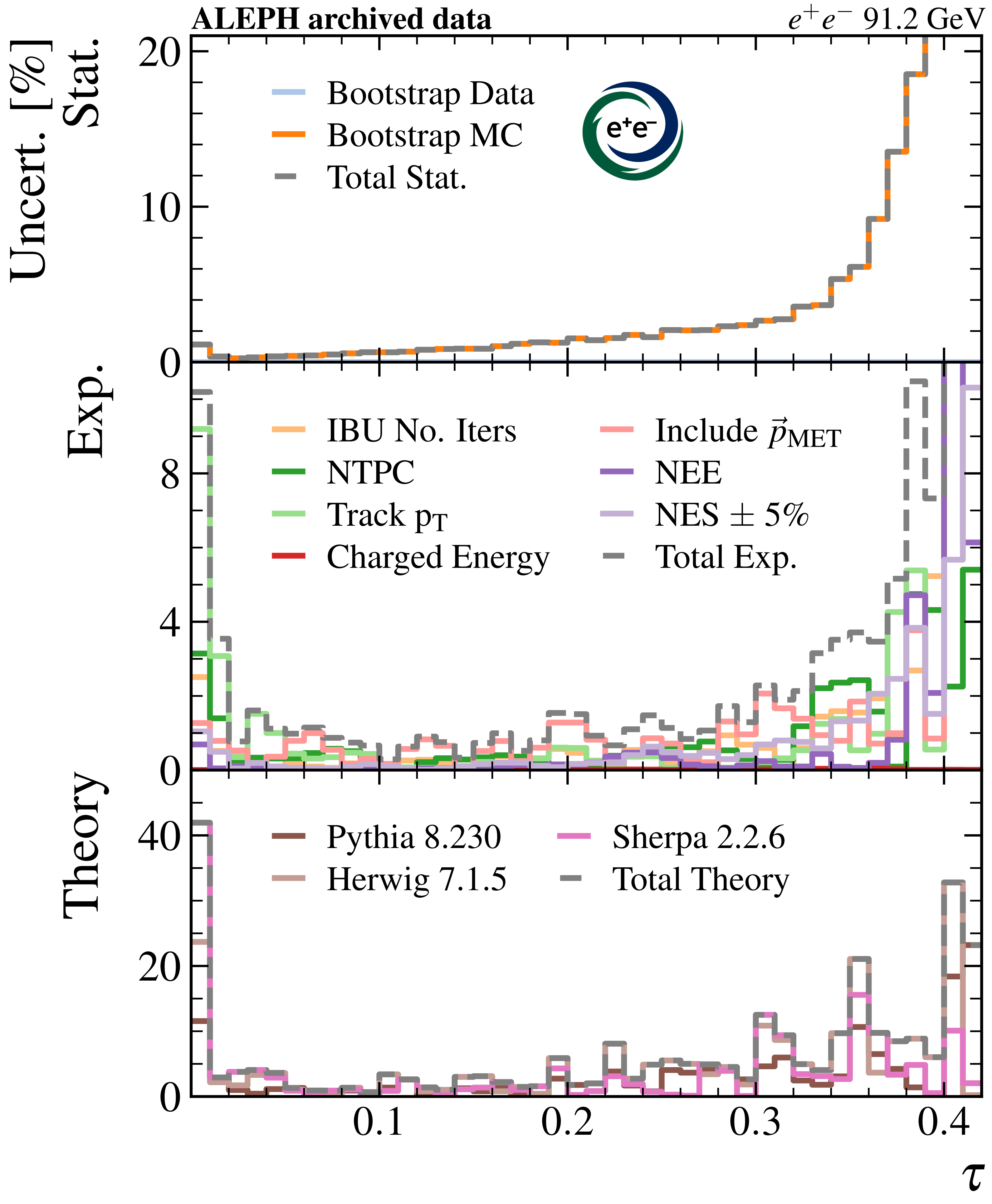
- Hadronic event selection corrected post unfolding with bin-by-bin correction
- No tracking efficiency correction applied after unfolding

Performed similar closure check to validate procedure

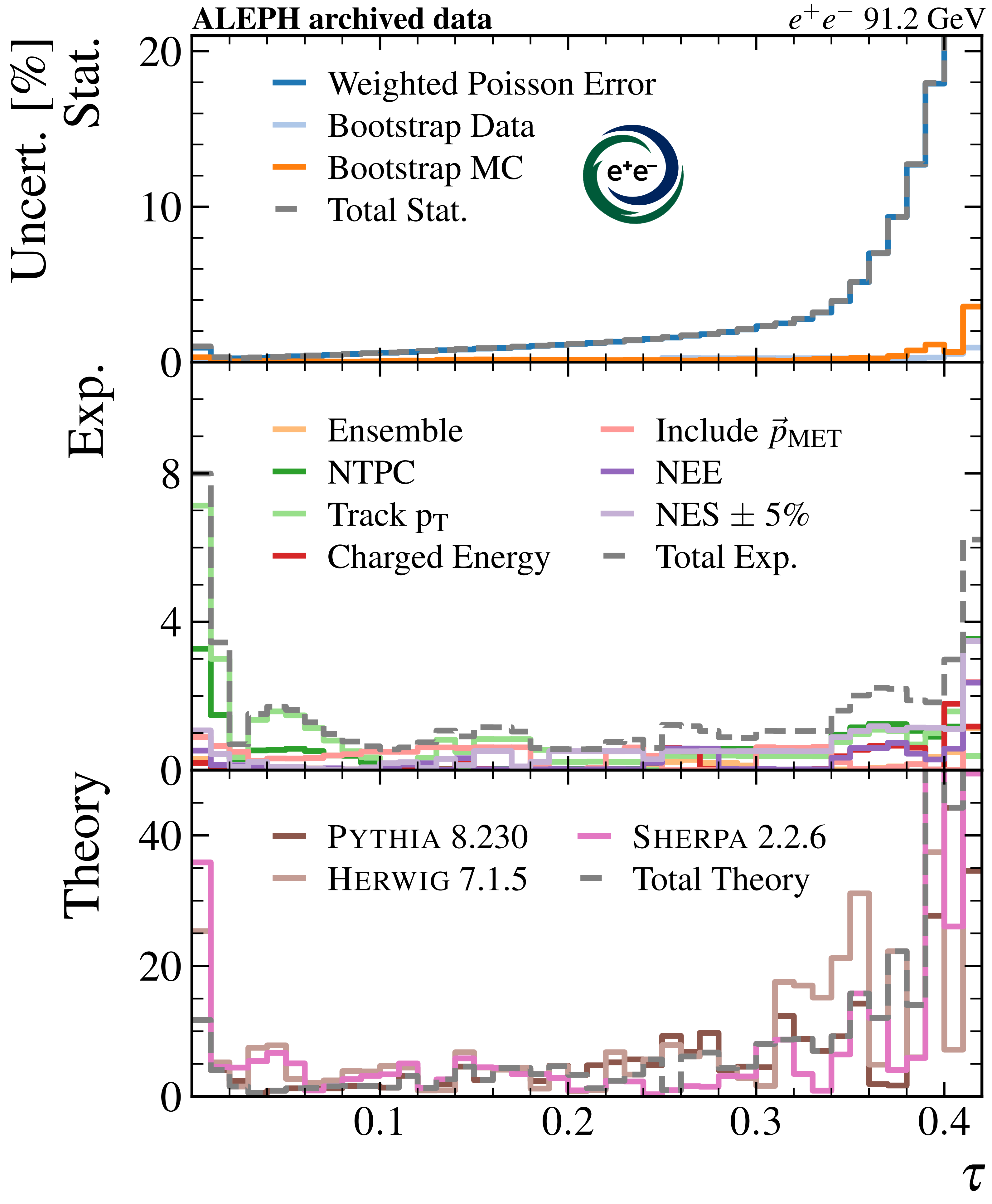


Uncertainty Breakdown

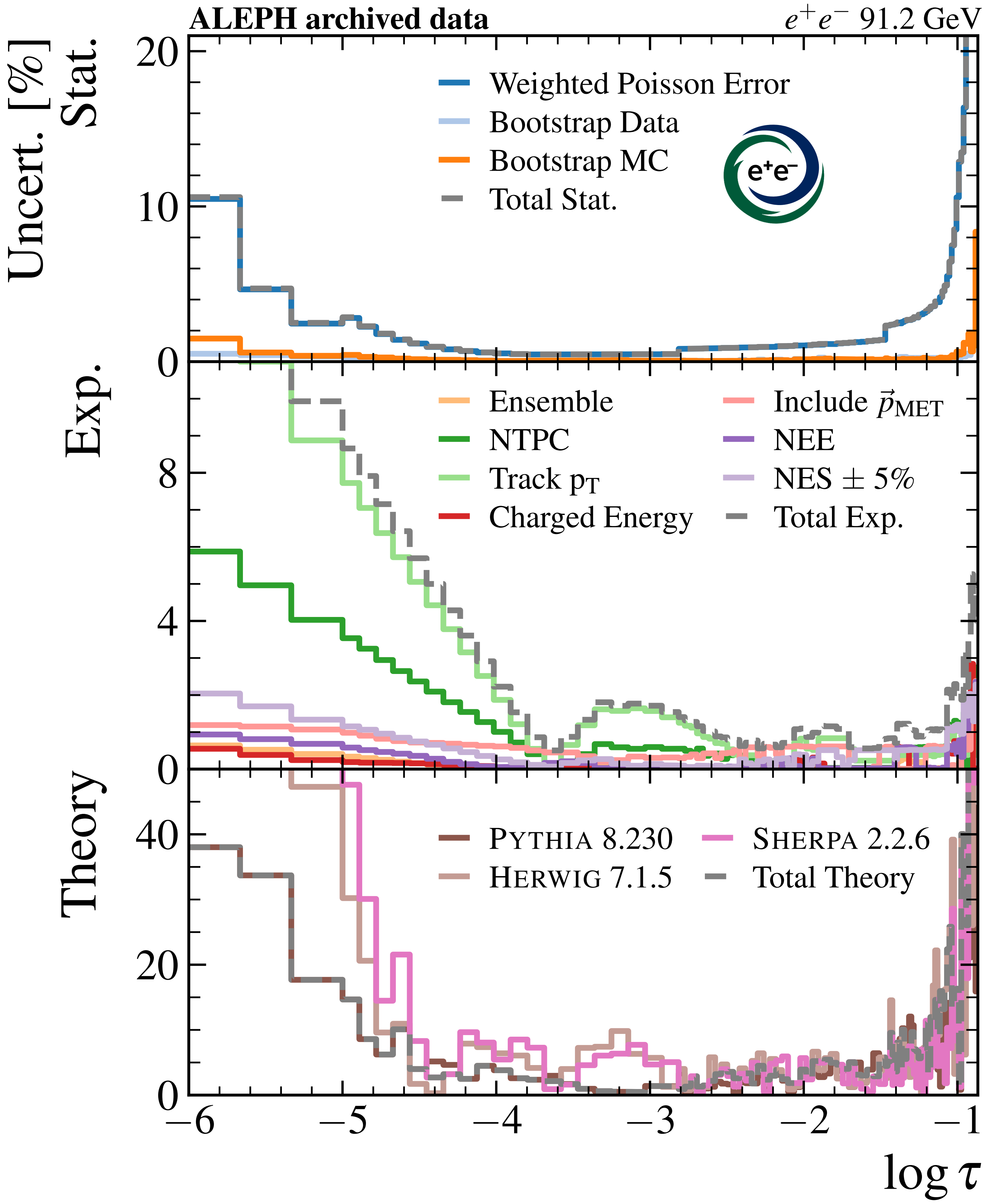
IBU τ, ALEPH Bins



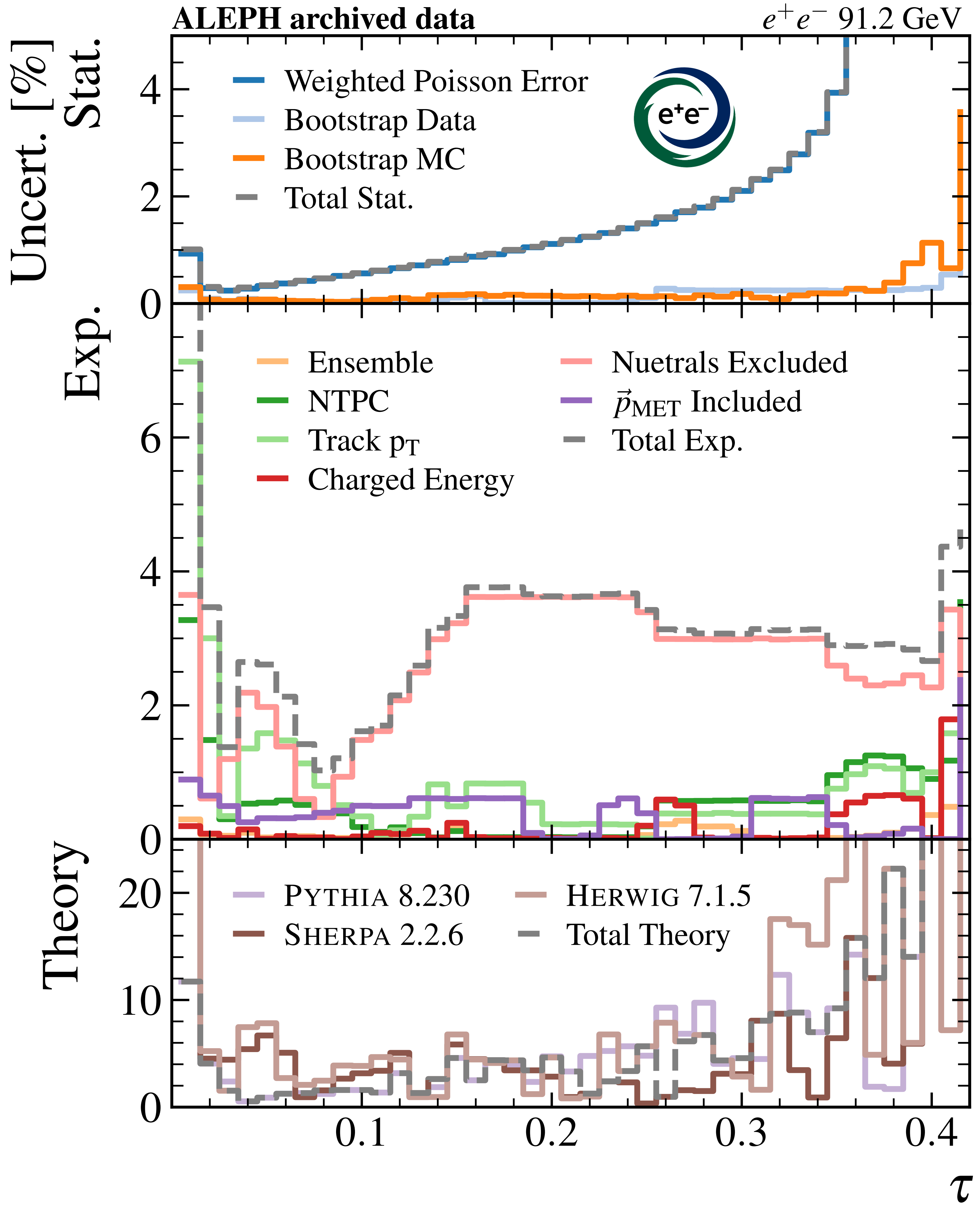
OF τ, ALEPH Bins

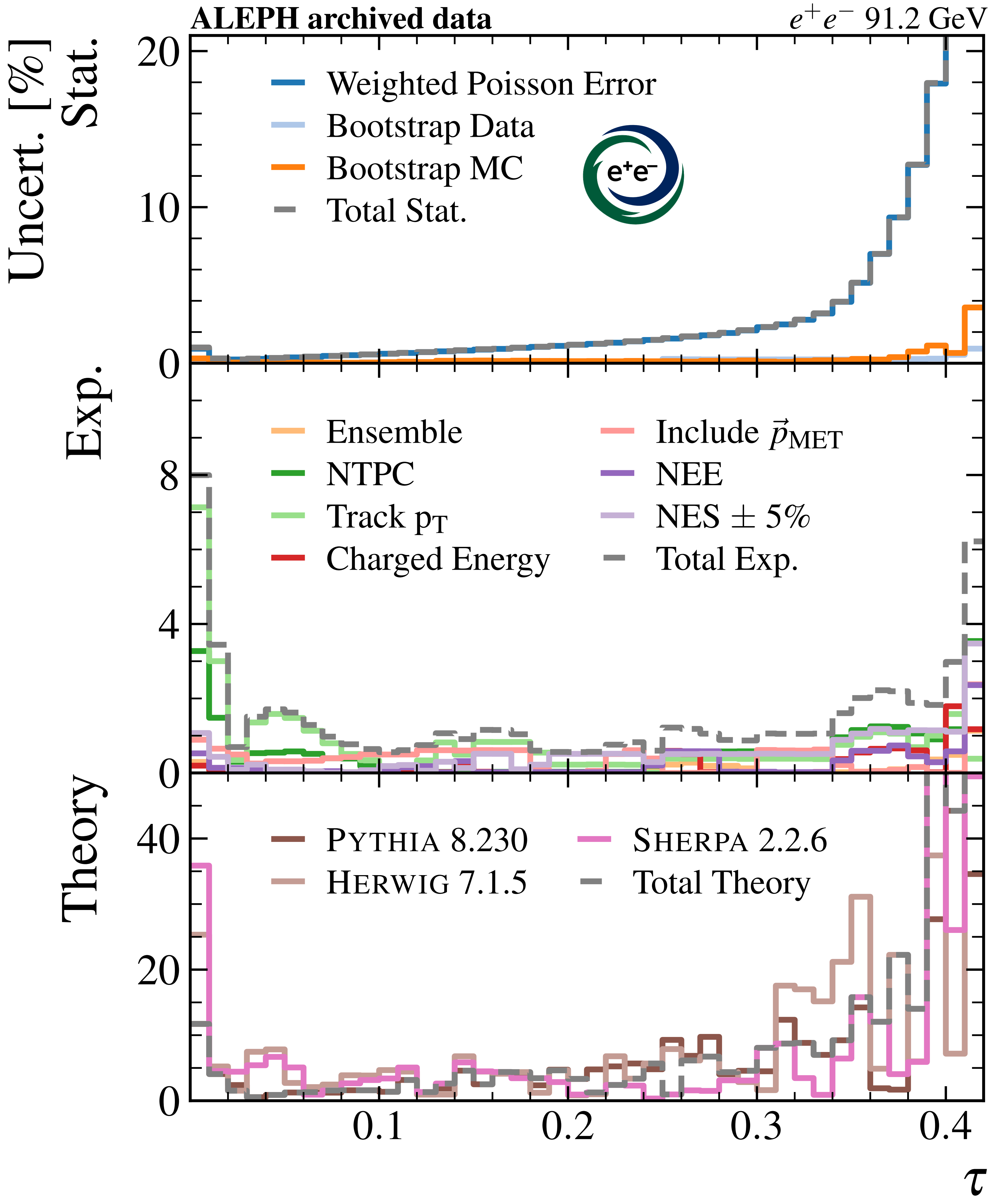


OF $\log \tau$, Custom Bins

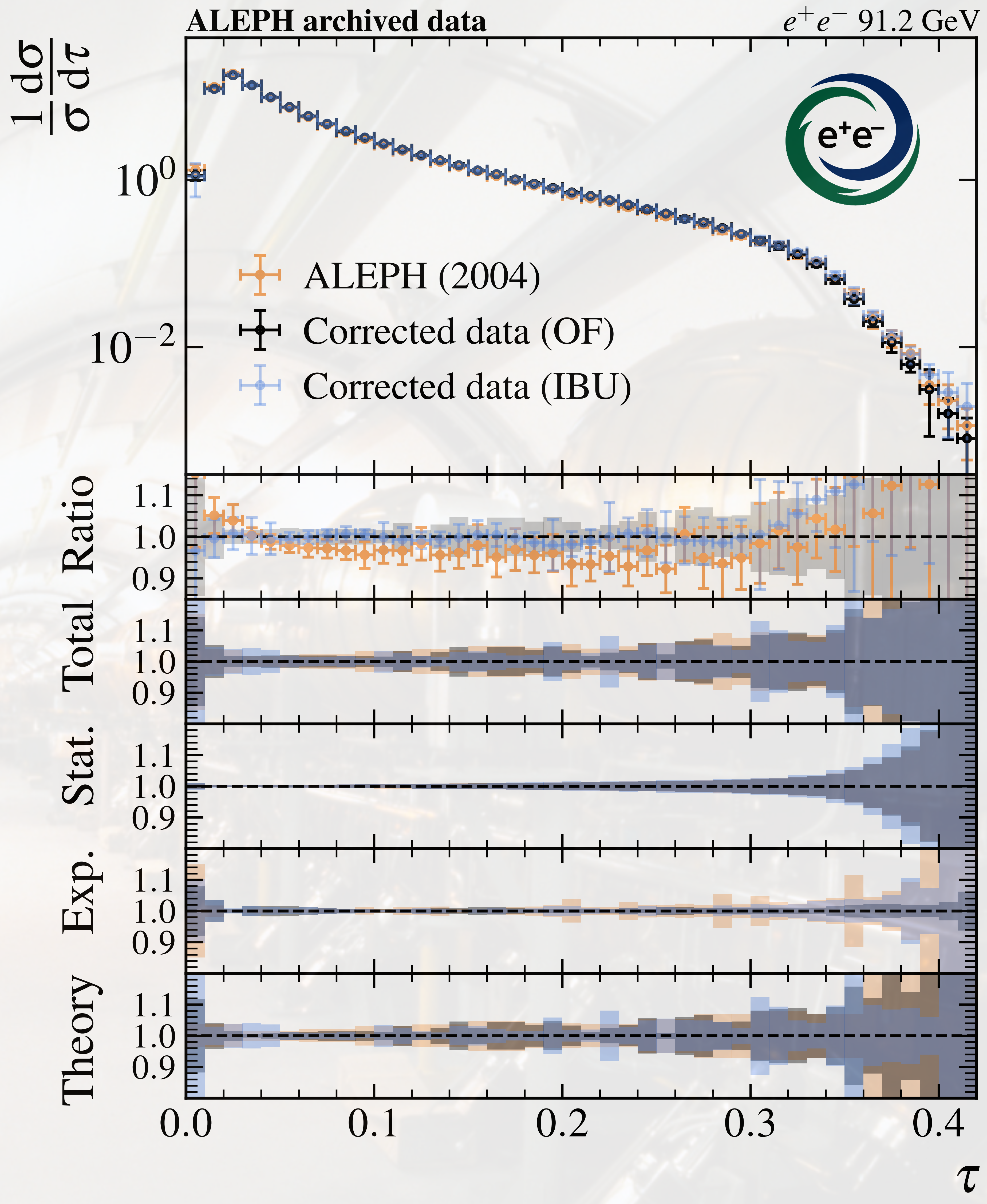


Neutrals Excluded → New Handling



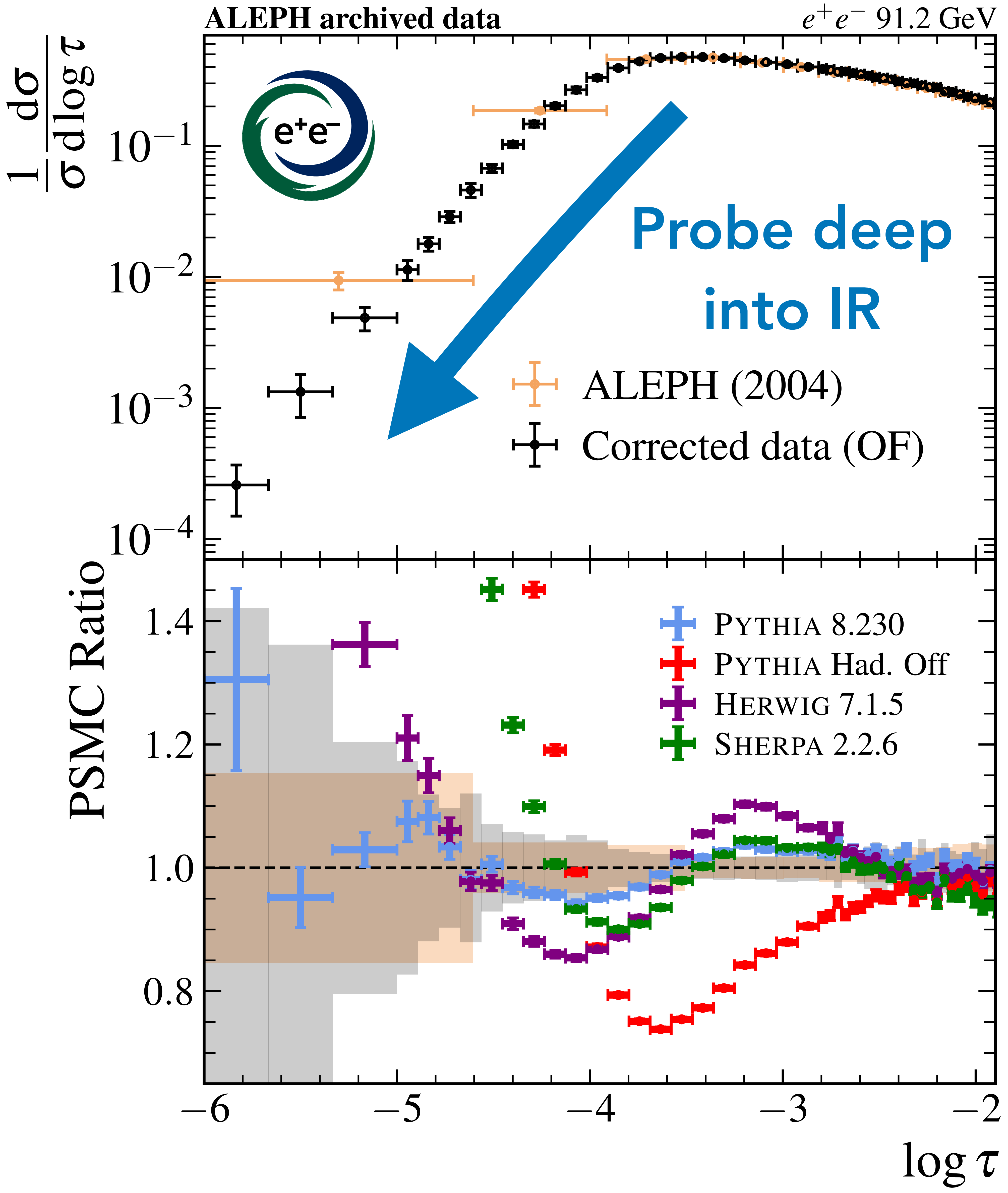


Fully Corrected Thrust Spectrum



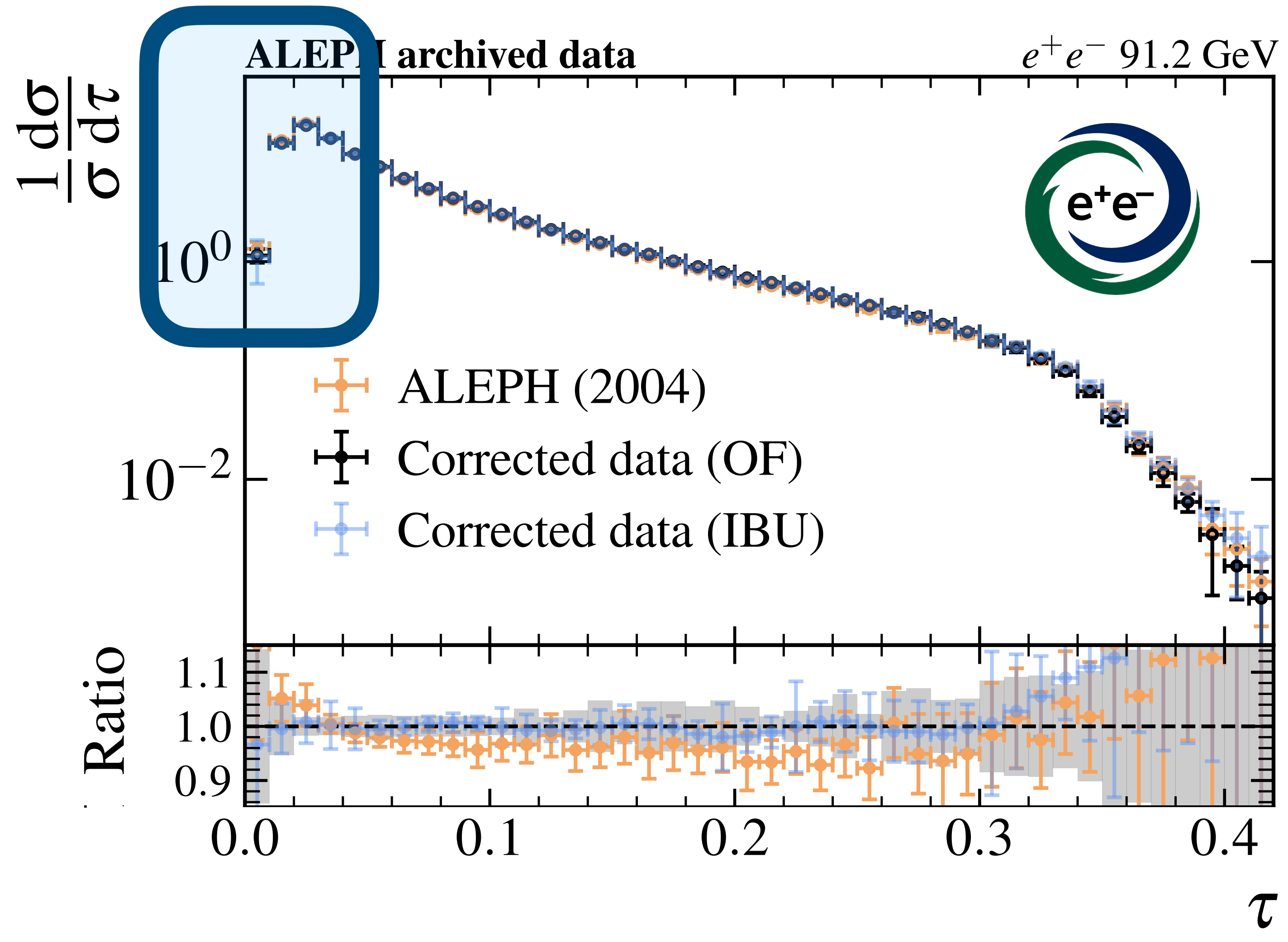
- Good agreement with previous ALEPH result and IBU result
- Possible hint of a systematic broadening of the τ spectrum, with potential α_S implications
- Possible sufficient precision to expand bins deep in dijet peak region for higher granularity study of NP regime

Fully Corrected Thrust Deep Into the IR



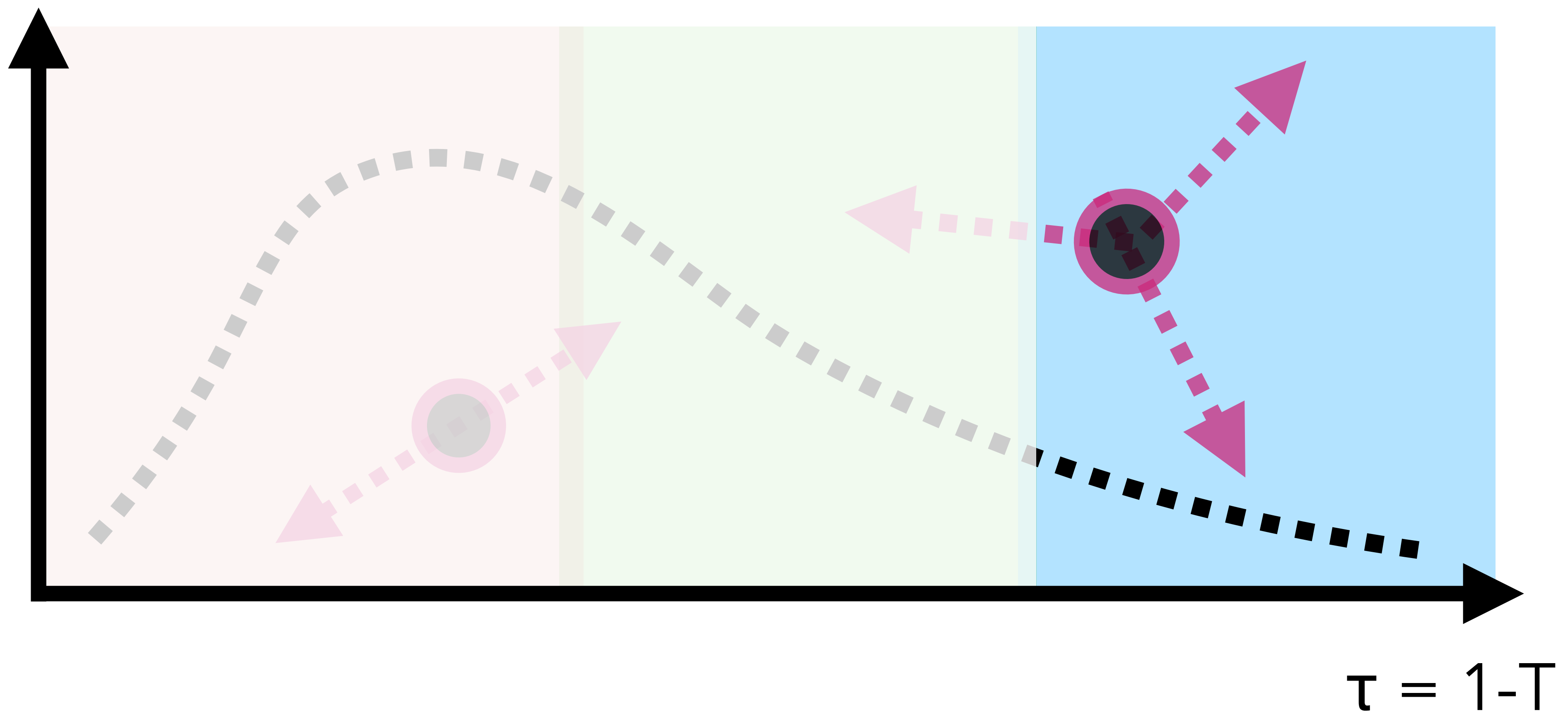
- Unbinned result allows us to test with the same measurement if the dataset supports dedicated fine binning in specific regimes
- New result probes deep into the IR region where the observable $\tau \sim \Lambda_{QCD}/\sqrt{s}$ and NP QCD effect become significant

Expansion of ~3 bins. ALEPH collaboration could have done this but at the time the main interest was on pQCD

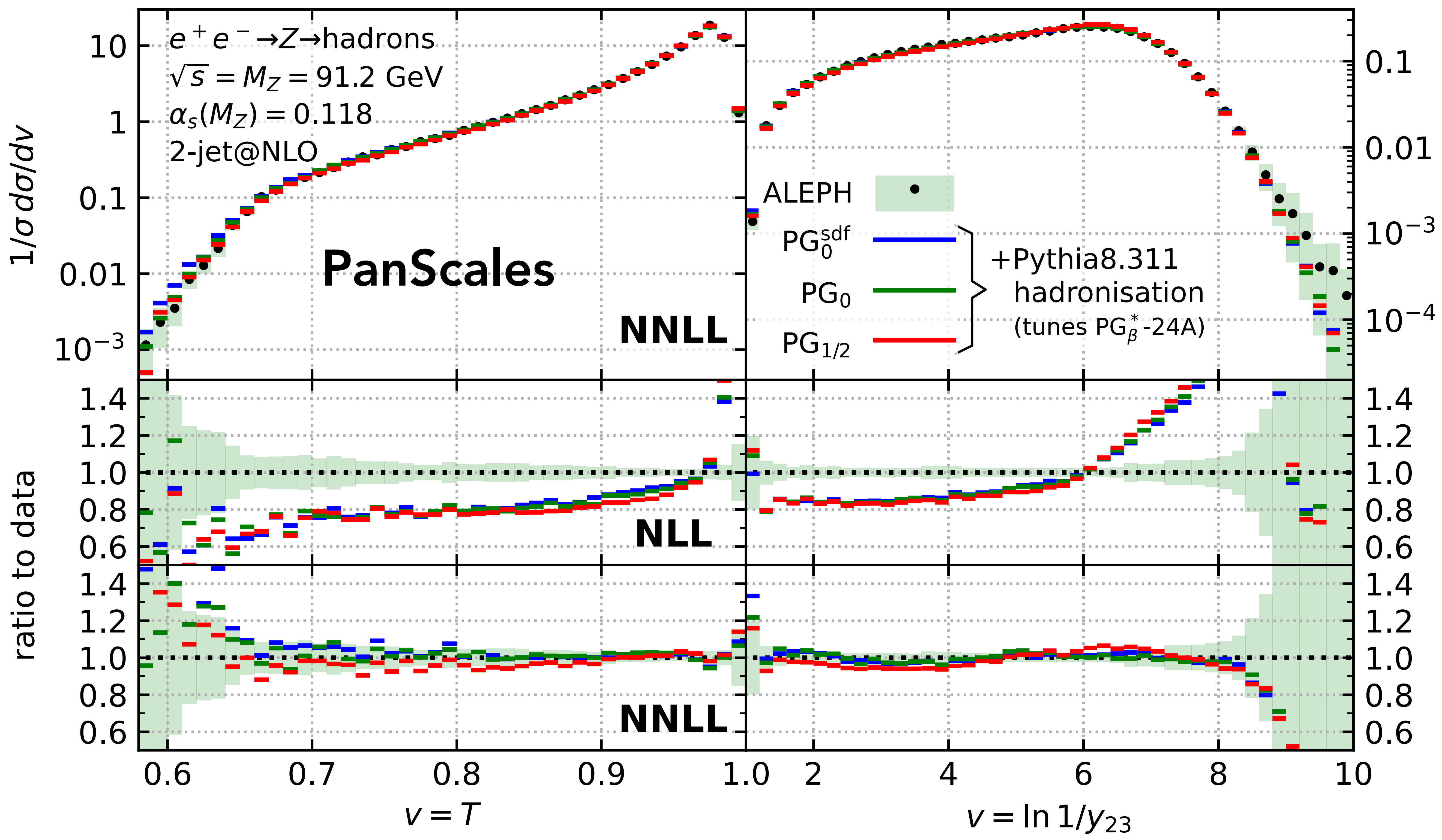


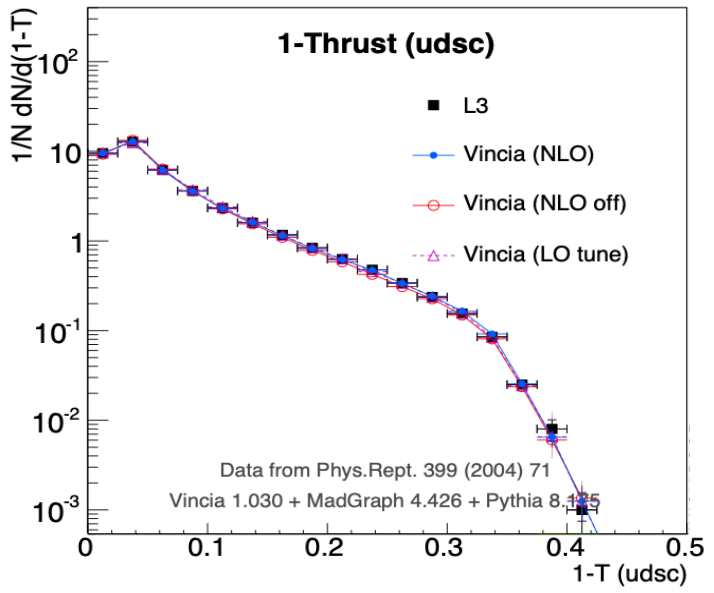
Modern Interest in Thrust Regime (III)

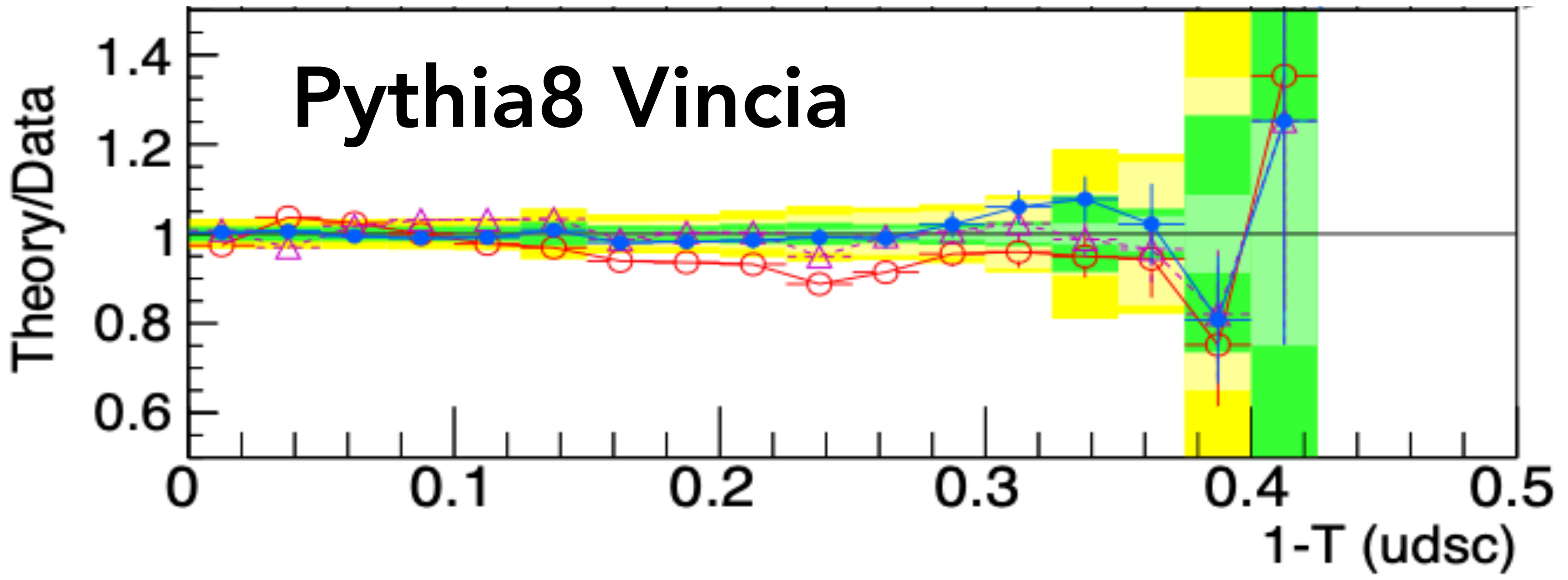
3(multi)-jet region



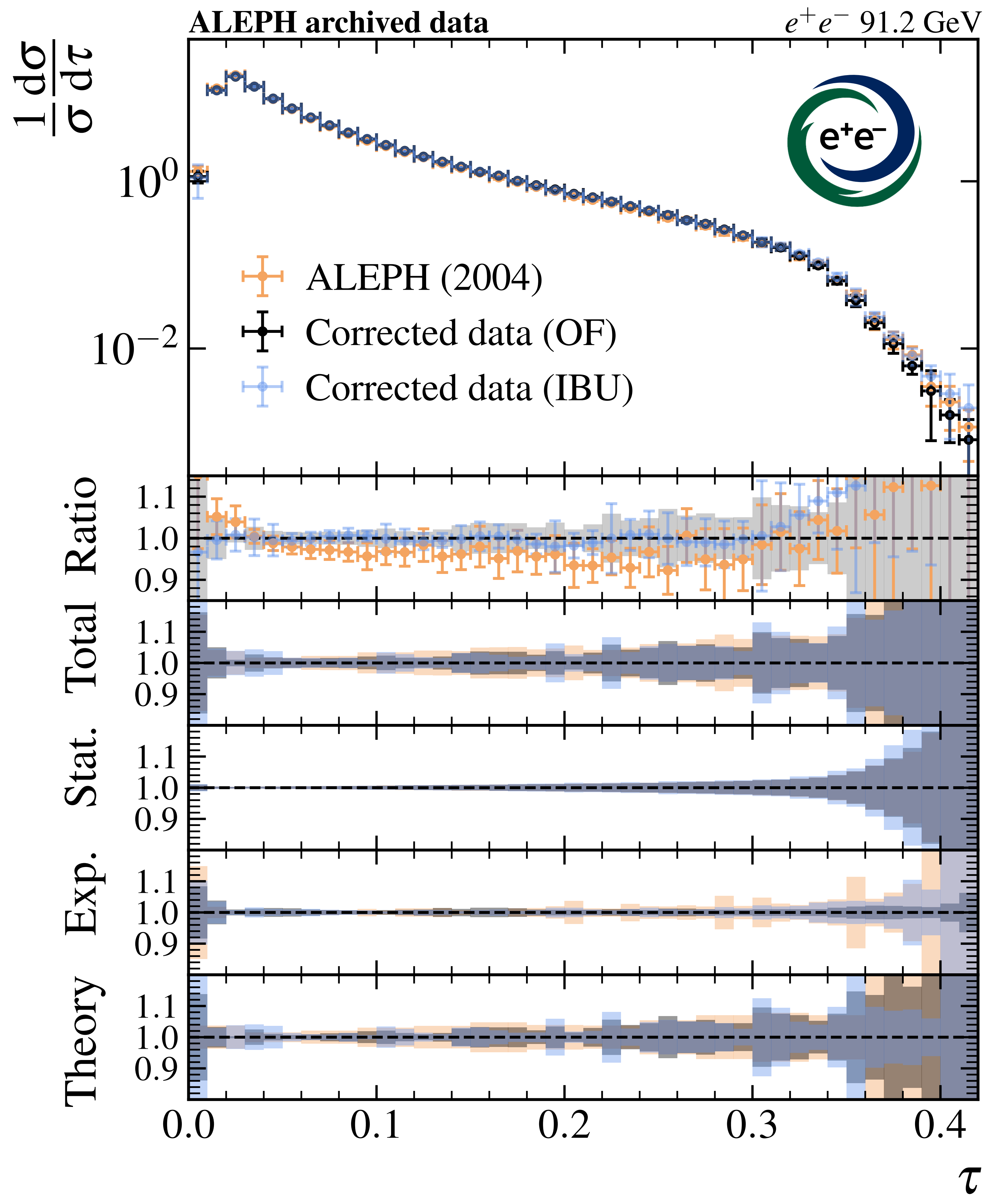
P QCD predictions in multi-jet regimes and parton shower MC validated and tuned with LEP event shapes

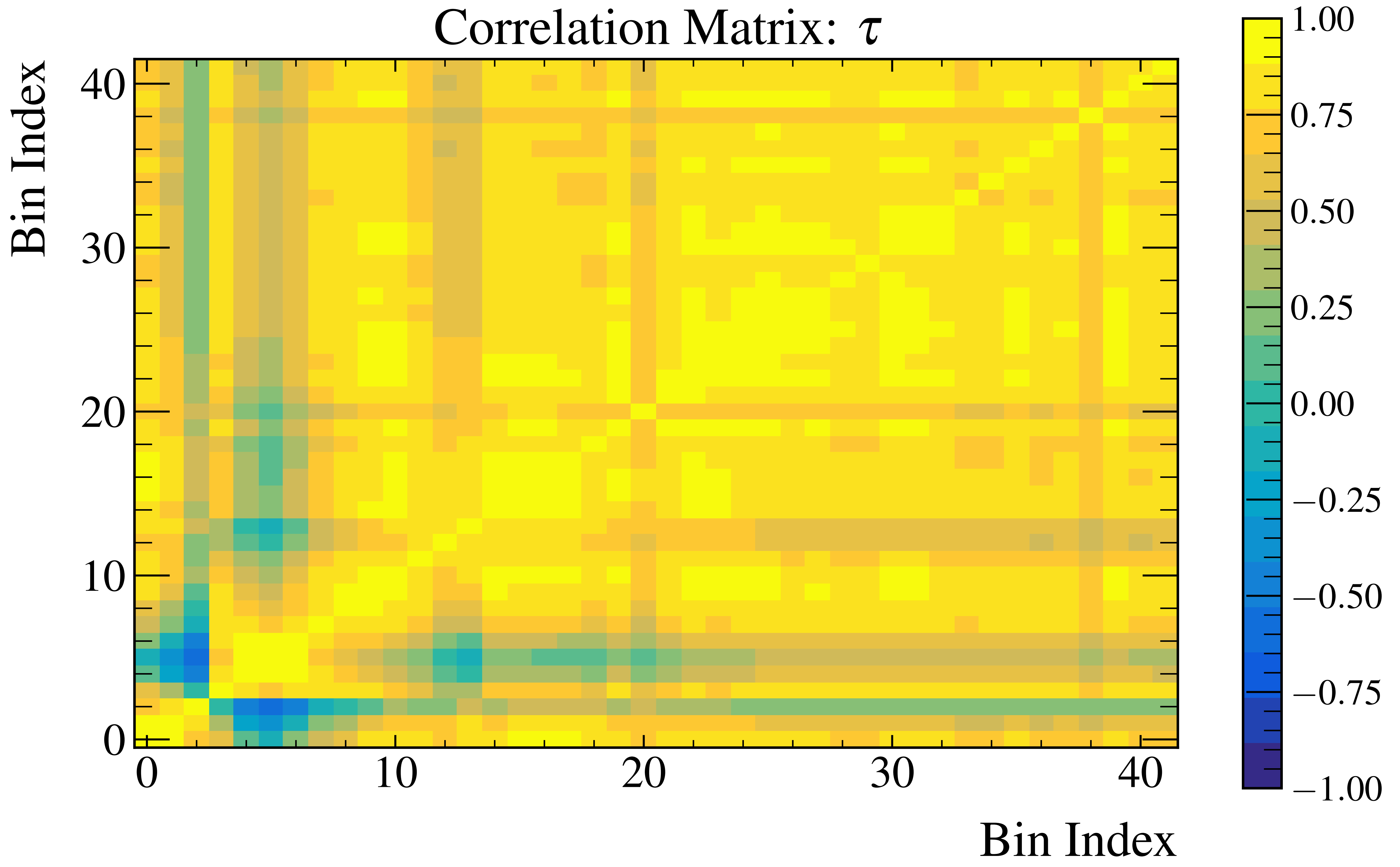






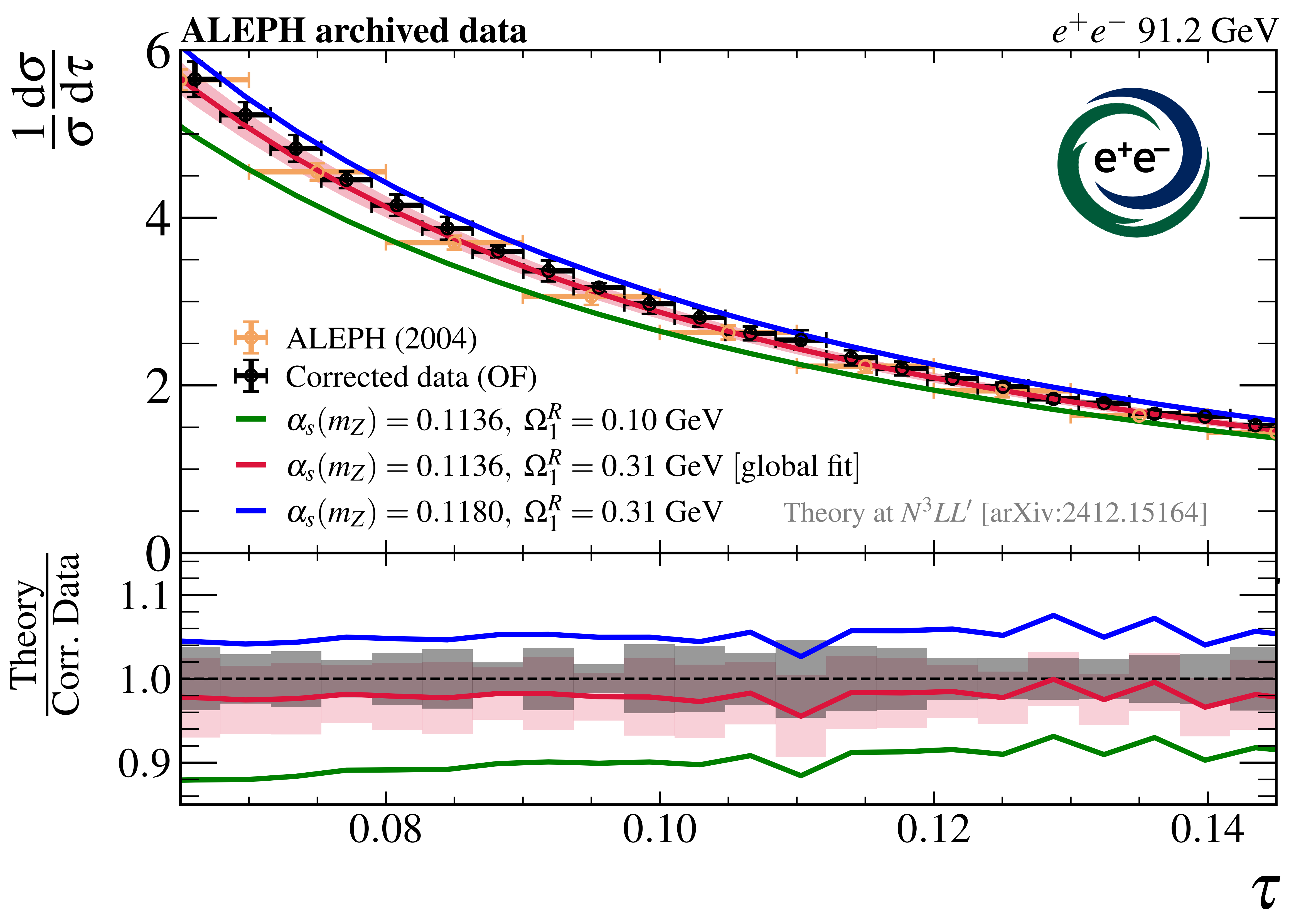
Covariance Matrices

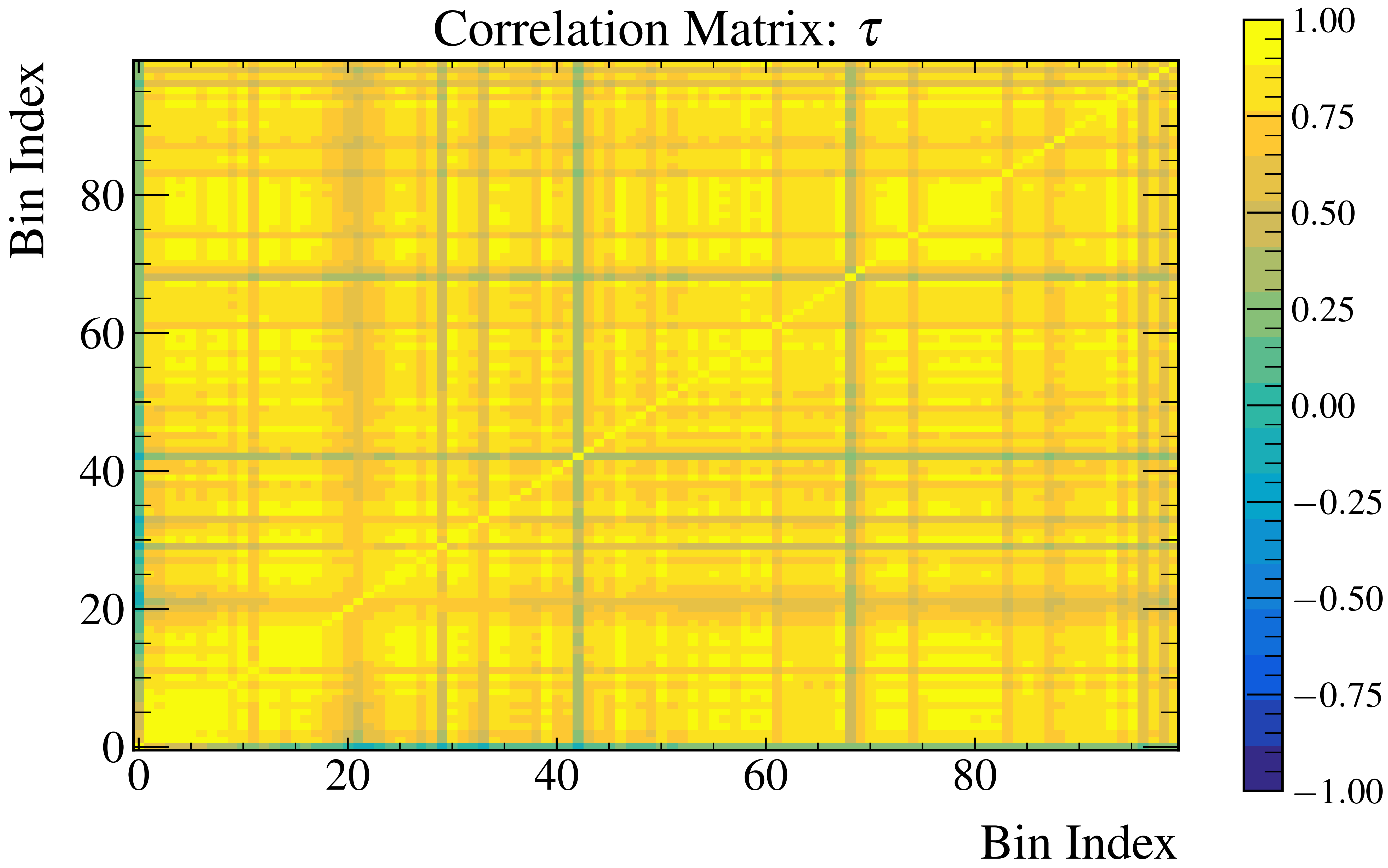




Sample covariance for bootstraps and ensembles.
Hessian calculation for theory and exp. systematics.

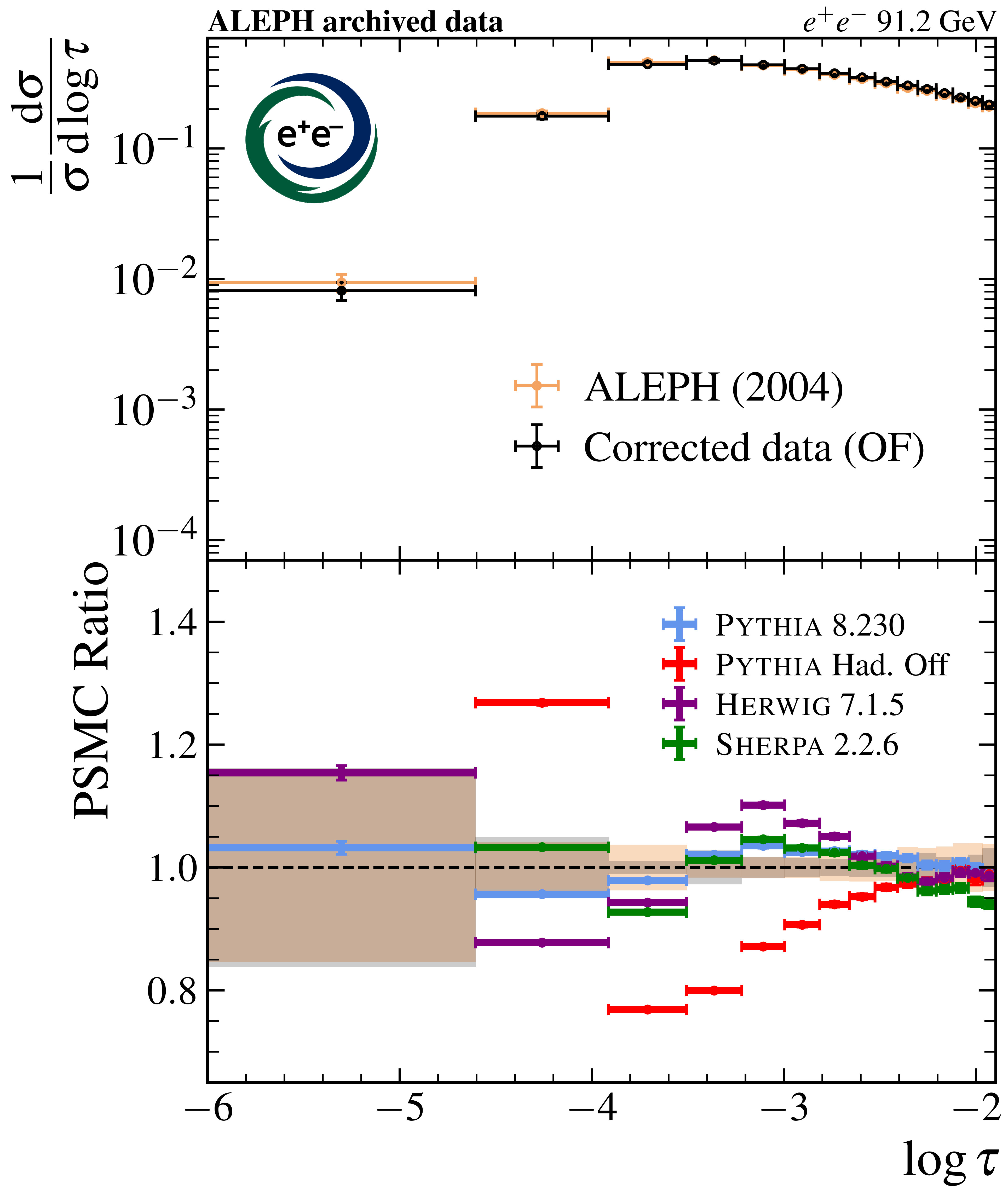
Covariance Matrices

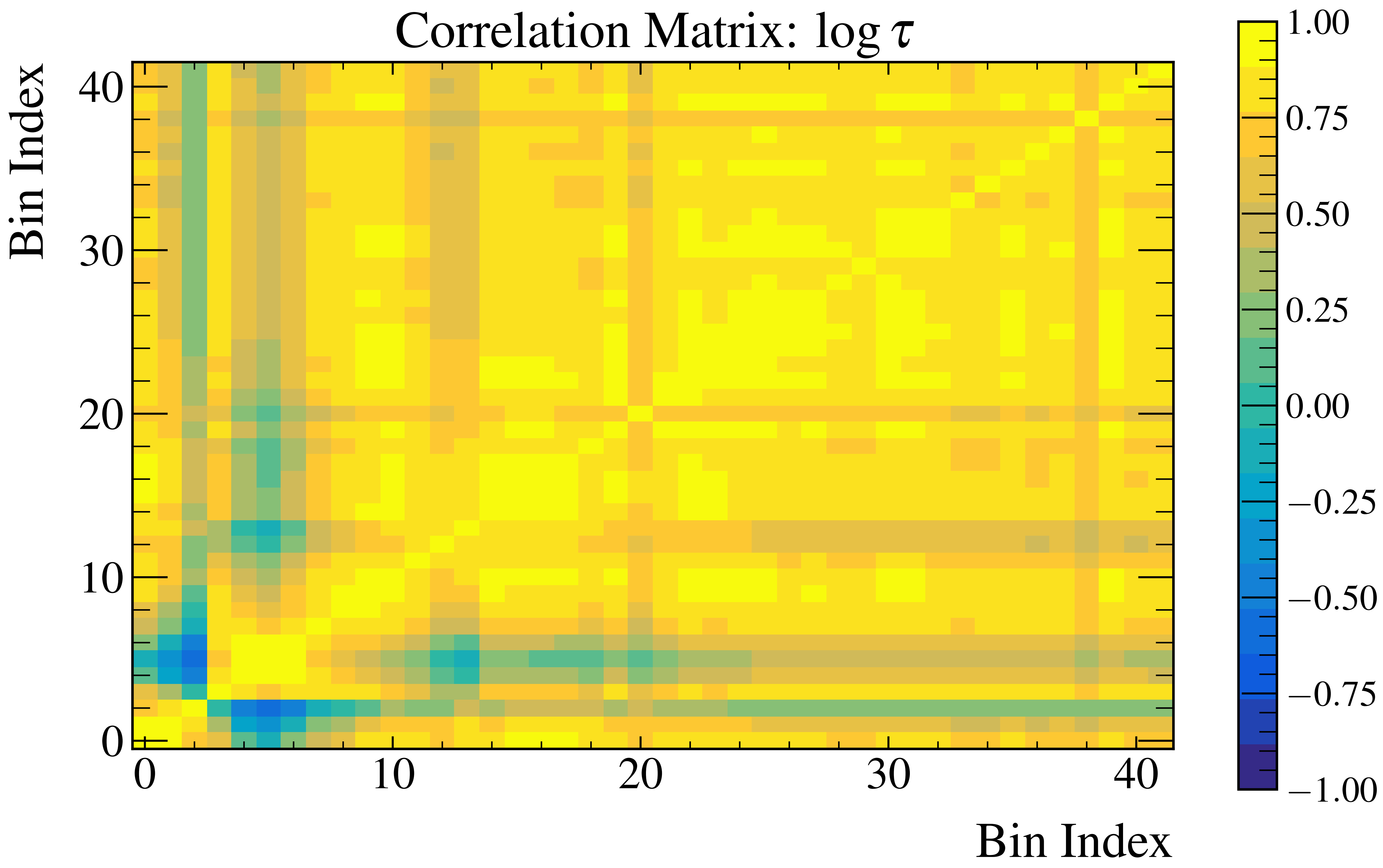




Sample covariance for bootstraps and ensembles.
Hessian calculation for theory and exp. systematics.

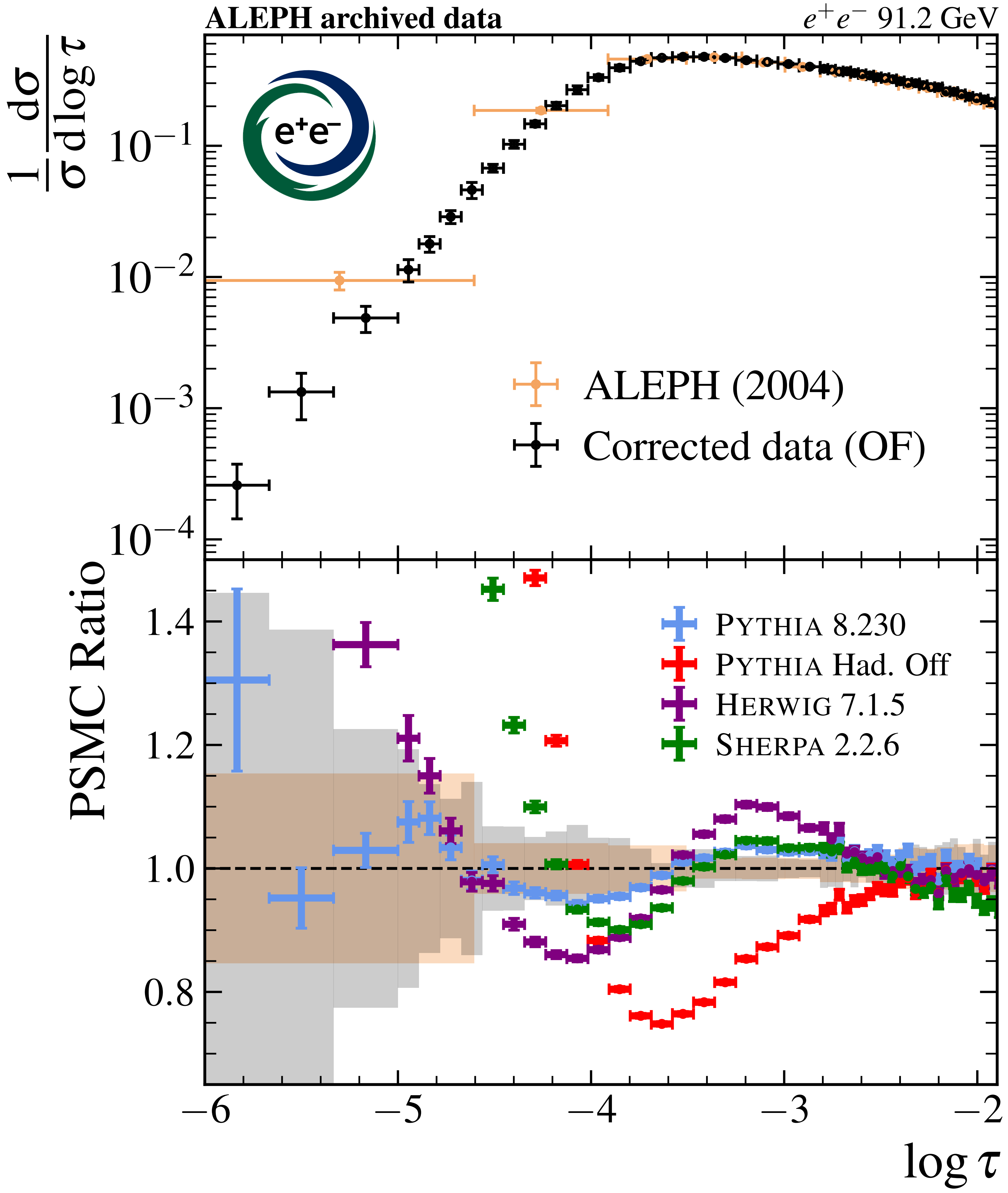
Covariance Matrices

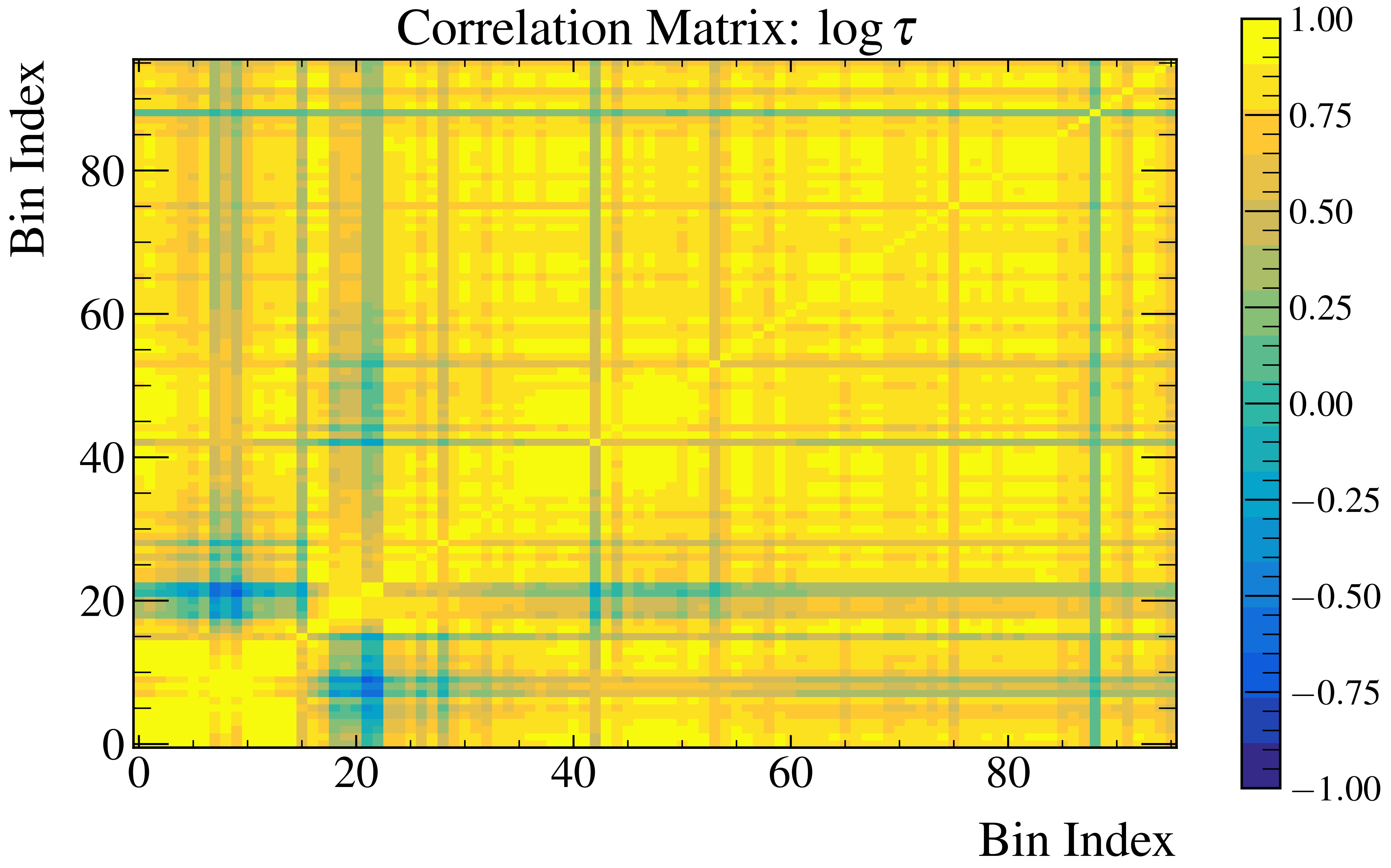




Sample covariance for bootstraps and ensembles.
Hessian calculation for theory and exp. systematics.

Covariance Matrices



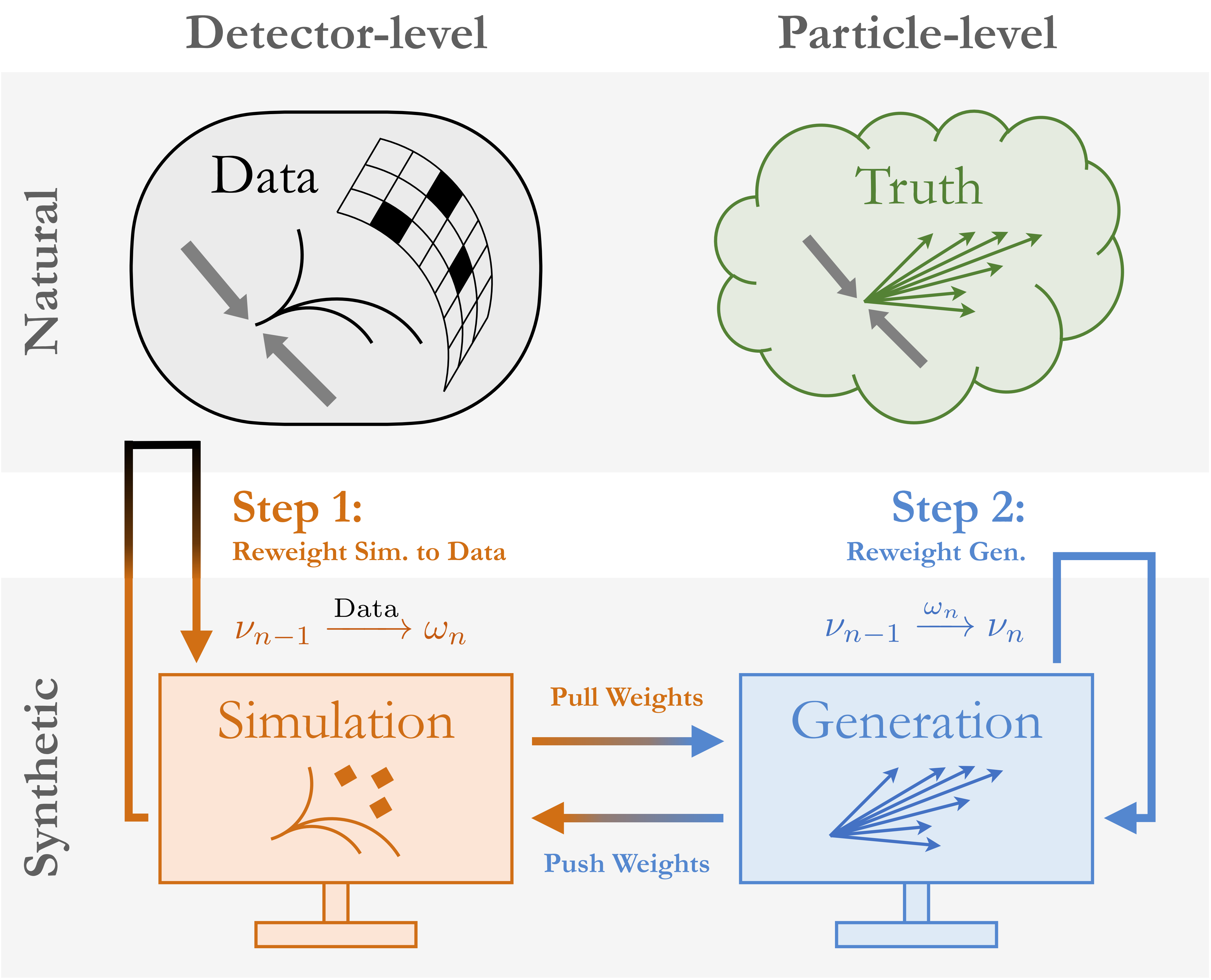


Sample covariance for bootstraps and ensembles.
Hessian calculation for theory and exp. systematics.

Correcting the Spectrum to Particle Level

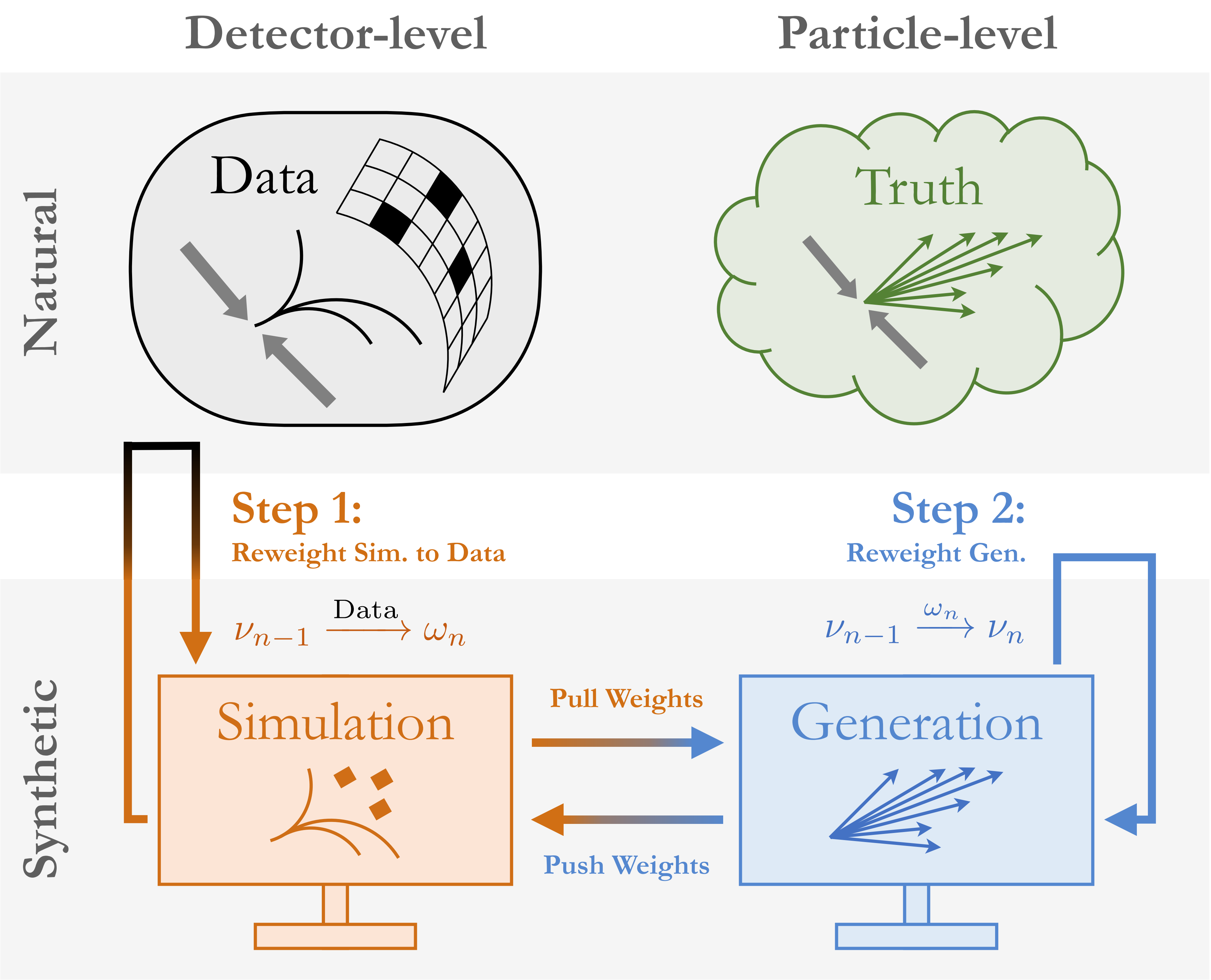
- Implicit corrections and detector effects accounted for with unbinned unfolding algorithm OmniFold
- Applied to single observable τ

As a high-level reminder, conceptually ~ unbinned iterative bayesian unfolding but very different procedure



Andreassen et al. Phys. Rev. Lett. 124, 182001

OmniFold Algorithm



Andreassen et al. Phys. Rev. Lett. 124, 182001

In theory the steps give us the following re-weightings

Step 1 $\omega_n(m) = \nu_{n-1}^{\text{push}}(m)\, L\left[(1,\text{Data}), \left(\nu_{n-1}^{\text{push}}, \text{Sim.}\right)\right](m)$

Step 2 $\nu_n(t) = \nu_{n-1}(t)\, L\left[\left(\omega_n^{\text{pull}}, \text{Gen.}\right), \left(\nu_{n-1}, \text{Gen.}\right)\right](t)$

where L is the likelihood ratio $L[(w,X),(w',X')](x) = \dfrac{p_{(w,X)}(x)}{p_{(w',X')}(x)}$

sim/gen matching used for $\nu_n^{\text{push}}(m) = \nu_n(t)$, $\omega_n^{\text{pull}}(t) = \omega_n(m)$

and the gen prior is $\nu_0(t)$

→ Assume we have these, follow the math

OmniFold Algorithm

Andreassen et al. Phys. Rev. Lett. 124, 182001

Step 2

$$\nu_n(t)p_{\text{Gen.}}(t) = (t)\nu_{n-1}(t)\frac{p_{(\omega_n^{\text{pull}},\text{Gen.})}(t)}{p_{(\nu_{n-1},\text{Gen.})}(t)}p_{\text{Gen.}}$$

Step 1

$$= \nu_{n-1}(t)\frac{\int dm\; p_{\text{Gen.|Sim.}}(t\mid m)p_{\text{Sim.}}(m)\omega_n(m)}{\nu_{n-1}(t)p_{\text{Gen.}}(t)}p_{\text{Gen.}}(t)$$

$$= \int dm\; \nu_{n-1}(t)\frac{p_{\text{Gen.|Sim.}}(t\mid m)p_{\text{Sim.}}(m)\nu_{n-1}^{\text{push}}(m)}{\nu_{n-1}p_{\text{Gen.}}(t)}\frac{p_{\text{data}}(m)}{p_{(\nu_{n-1}^{\text{push}},\text{Sim.})}(m)}p_{\text{Gen.}}(t)$$

→ Use that $\nu_{n-1}^{\text{push}}(m) = \nu_{n-1}(t)$

$$= \int dm\; p_{\text{data}}(m)\nu_{n-1}(t)\frac{p_{\text{Gen.|Sim.}}(t\mid m)p_{\text{Sim.}}(m)}{p_{\text{Gen.}}(t)}\frac{1}{p_{(\nu_{n-1}^{\text{push}},\text{Sim.})}(m)}p_{\text{Gen.}}(t)$$

→ Use Bayes' Rule: $P(m\,|\,t) = \dfrac{P(t\,|\,m)P(m)}{P(t)}$

$$= \int dm\; p_{\text{data}}(m)\nu_{n-1}(t)p_{\text{Sim.|Gen.}}(m\mid t)\frac{1}{p_{(\nu_{n-1}^{\text{push}},\text{Sim.})}(m)}p_{\text{Gen.}}(t)$$

→ Use that $p_{(\nu_{n-1}^{\text{push}},\text{Sim.})}(m) = \int dt' p_{\text{Sim.|Gen.}}(m\mid t')\nu_{n-1}(t')p(t')$

$$= \int dm\; p_{\text{data}}(m)\frac{p_{\text{Sim.|Gen.}}(m\mid t)\nu_{n-1}(t)p_{\text{Gen.}}(t)}{\int dt' p_{\text{Sim.|Gen.}}(m\mid t')\nu_{n-1}(t')p(t')}$$

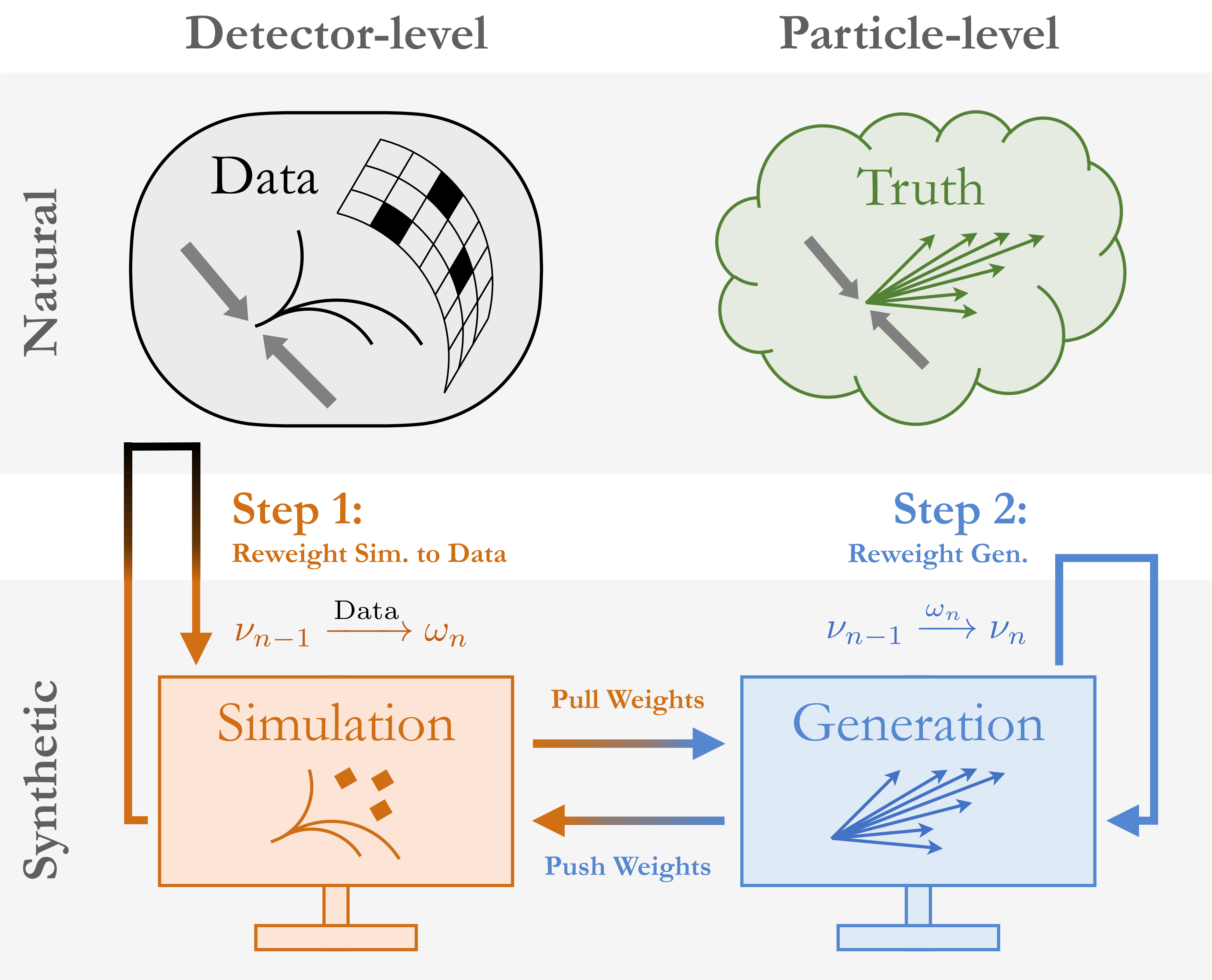
or as binned version

$$= \sum_i m_i \frac{R_{ij}t_j}{\sum_k R_{ik}t_k}$$

→ **Unbinned IBU**

→ Note: could derive from first principle using log-likelihood of unfolding problem

OmniFold Algorithm



Andreassen et al. Phys. Rev. Lett. 124, 182001

In theory the steps give us the following re-weightings

Step 1 $\omega_n(m) = \nu_{n-1}^{\text{push}}(m)\, L\left[(1,\text{Data}), \left(\nu_{n-1}^{\text{push}}, \text{Sim.}\right)\right](m)$

Step 2 $\nu_n(t) = \nu_{n-1}(t)\, L\left[\left(\omega_n^{\text{pull}}, \text{Gen.}\right), (\nu_{n-1}, \text{Gen.})\right](t)$

where L is the likelihood ratio $L[(w,X),(w',X')](x) = \dfrac{p_{(w,X)}(x)}{p_{(w',X')}(x)}$

sim/gen matching used for $\nu_n^{\text{push}}(m) = \nu_n(t)$, $\omega_n^{\text{pull}}(t) = \omega_n(m)$

and the gen prior is $\nu_0(t)$

$$p^{(n)}_{\text{unfolded}}(t) = \nu_n(t) p_{\text{Gen.}}(t)$$
$$= \int dm\, p_{\text{data}}(m) \frac{p_{\text{Sim.|Gen.}}(m \mid t)\nu_{n-1}(t) p_{\text{Gen.}}(t)}{\int dt' p_{\text{Sim.|Gen.}}(m \mid t')\nu_{n-1}(t') p(t')}$$

OmniFold Algorithm

Andreassen et al. Phys. Rev. Lett. 124, 182001

How to gain access to the likelihood ratios in Steps 1 and 2? Train a classifier to minimize the following

$$L_{\mathrm{BCE}}[f] = -\int \mathrm{d}x \left(p_A(x)\log(f(x)) + p_B(x)\log(1-f(x))\right)$$

$$\frac{\partial L}{\partial f} = -\frac{\partial}{\partial f}\left(p_A(x)\log(f(x)) + p_B(x)\log(1-f(x))\right)$$

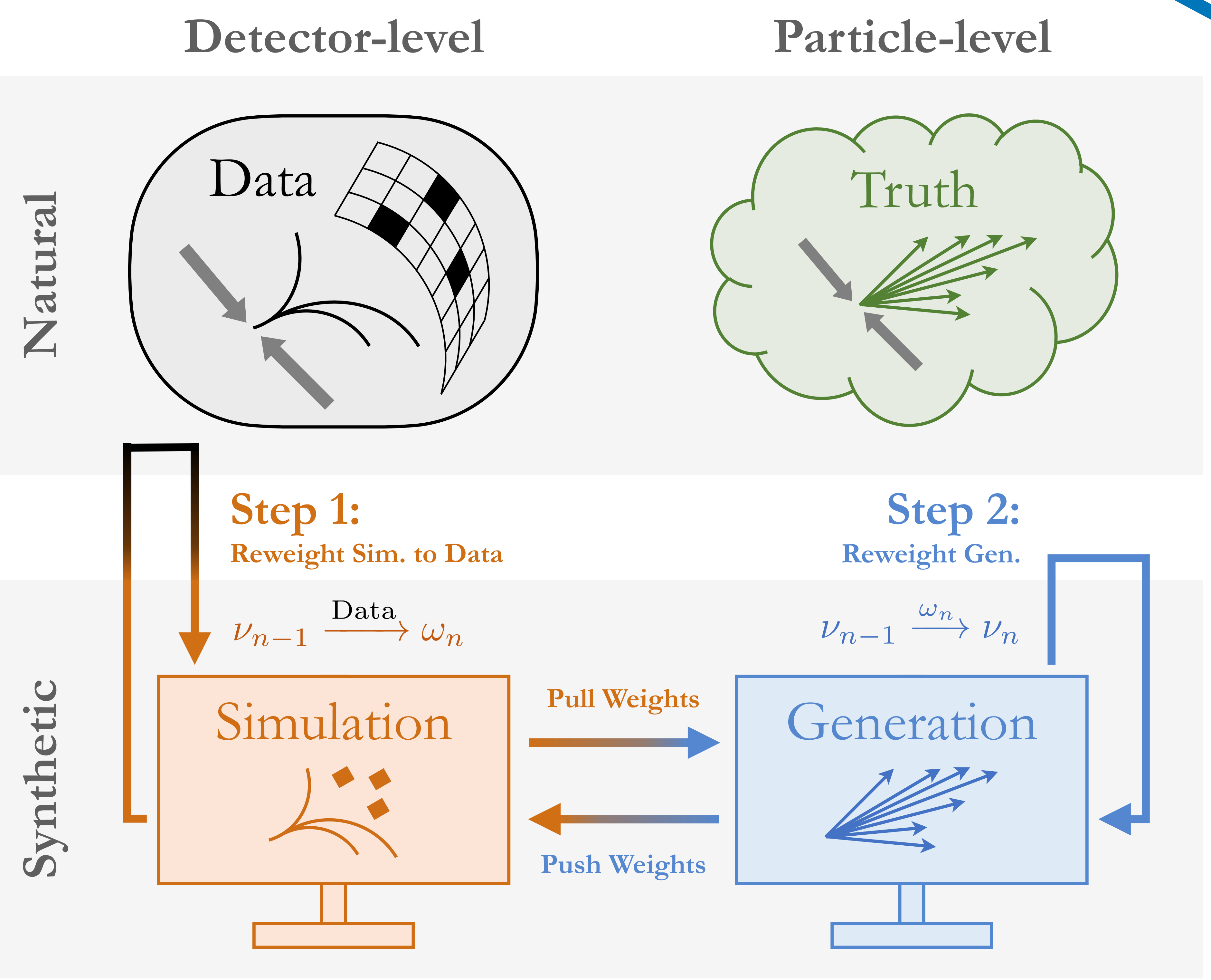
$$= -\left(\frac{p_A(x)}{f(x)} - \frac{p_B(x)}{1-f(x)}\right)$$

$$\frac{\partial L}{\partial f} = 0 \Rightarrow \frac{f(x)}{1-f(x)} = \frac{p_A(x)}{p_B(x)} \leftarrow$$

Likelihood ratio trick: output of classifier asymptotically converges to likelihood ratio

OmniFold Ensembling

Neural networks used to approximate likelihood ratios



Andreassen et al. Phys. Rev. Lett. 124, 182001

In theory the steps give us the following re-weightings

Step 1 (NN) $\omega_n(m) = \nu_{n-1}^{\text{push}}(m)\, L\left[(1,\text{Data}), \left(\nu_{n-1}^{\text{push}}, \text{Sim.}\right)\right](m)$

Step 2 (NN) $\nu_n(t) = \nu_{n-1}(t)\, L\left[\left(\omega_n^{\text{pull}}, \text{Gen.}\right), \left(\nu_{n-1}, \text{Gen.}\right)\right](t)$

where L is the likelihood ratio $L[(w,X),(w',X')](x) = \dfrac{p_{(w,X)}(x)}{p_{(w',X')}(x)}$

sim/gen matching used for $\nu_n^{\text{push}}(m) = \nu_n(t)$, $\omega_n^{\text{pull}}(t) = \omega_n(m)$

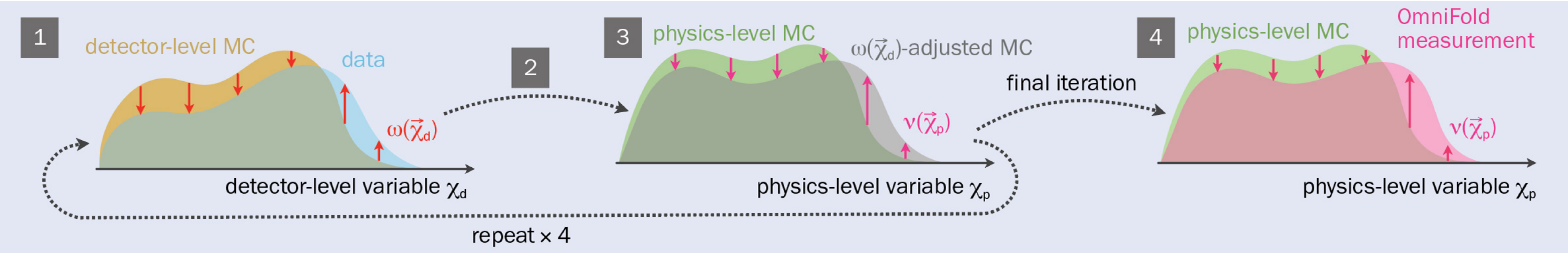
and the gen prior is $\nu_0(t)$

$$p^{(n)}_{\textbf{unfolded}}(t) = \nu_n(t) p_{\textbf{Gen.}}(t)$$

$$= \int dm\, p_{\textbf{data}}(m) \frac{p_{\textbf{Sim.|Gen.}}(m \mid t)\nu_{n-1}(t) p_{\textbf{Gen.}}(t)}{\int dt' p_{\textbf{Sim.|Gen.}}(m \mid t')\nu_{n-1}(t') p(t')}$$

OmniFold Visualization

How to Unfold with AI, CERN Courier, B. Nachman



Step 1

Step 2

$$p^{(n)}_{\text{unfolded}}(t) = \nu_n(t) p_{\text{Gen.}}(t)$$

DELPHI Thrust Backup

Corrections to the Spectrum

(1) Hadronic event selection:

- High efficiency, but removal of hadronic efficient
- Removed fraction corrected for via unfolding

(2) Tracking efficiency:

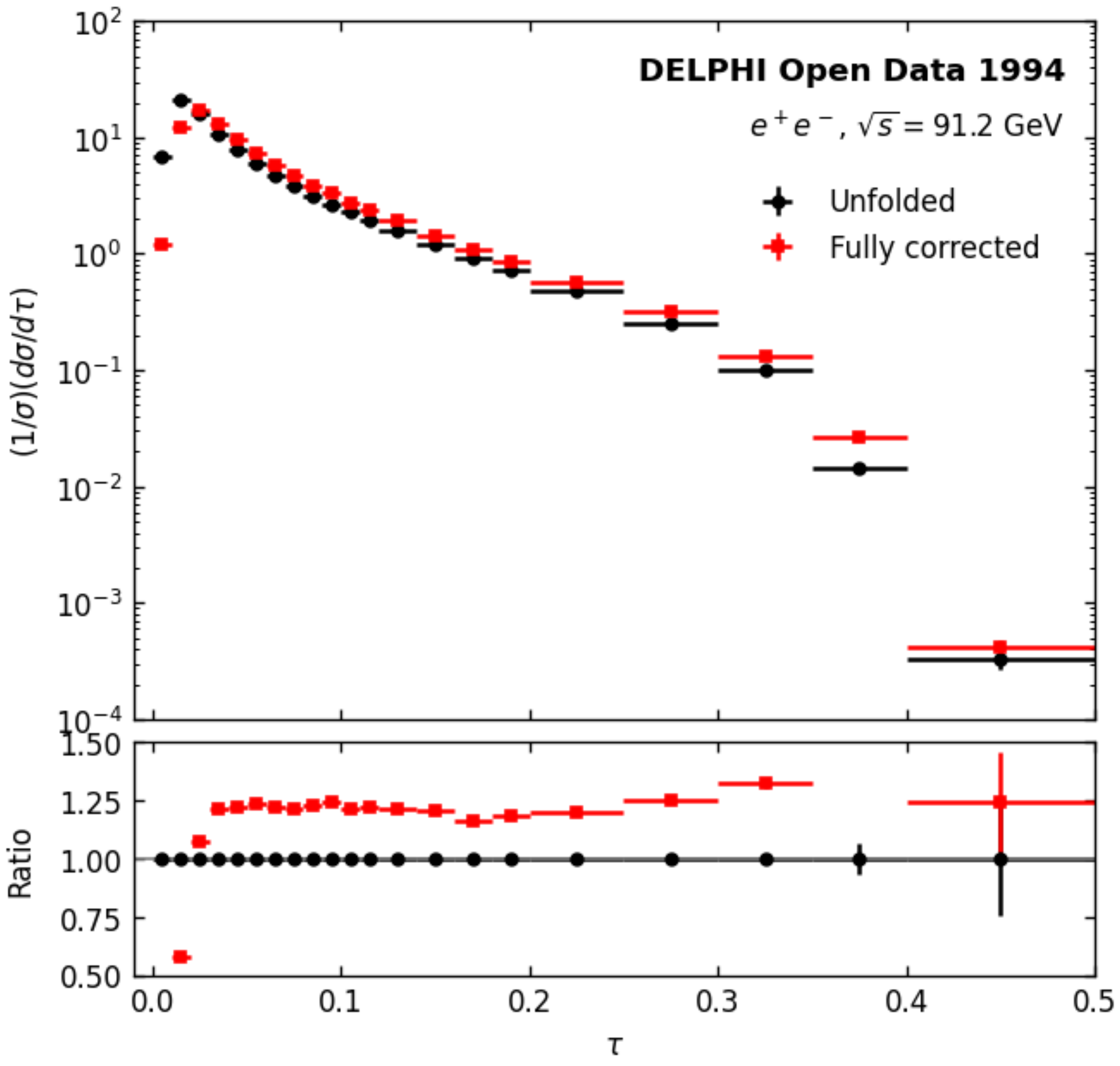
- Efficiency to reconstruct charged particles accounted for in unfolding

(3) 1D IBU D'Agostini unfolding with RooUnfold:

- Implicitly handle (1) - (2)
- Unfold to correct distribution for detector effects in fiducial phase space

(4) Full Phase Space & EM radiation effects:

- Bin-by-bin correction to full phase space without QED effects



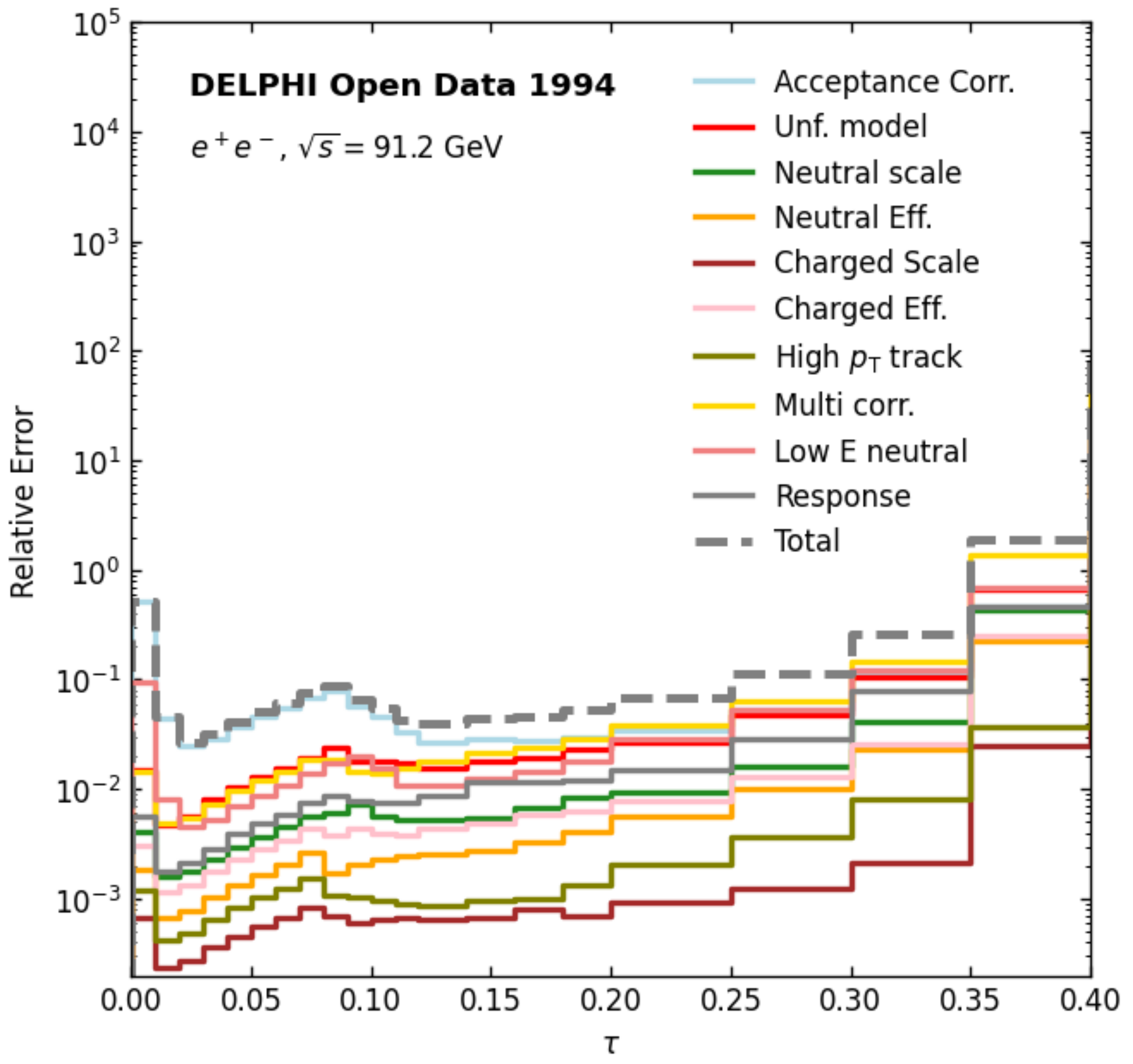
Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

Uncertainties

- **Experimental uncertainty:**
  - Variations in the selections
  - Variations in charged energy and efficiency
  - Variation of neutral energy and efficiency

- **Theory MC prior uncertainty:**
  - Max spread from variation of the MC prior in unfolding. 4 sets of MC samples used: Pythia 5.7 and ARIADNE (open data with DELPHI tune), Pythia 8.3 and Dire (new MC + DELSIM)

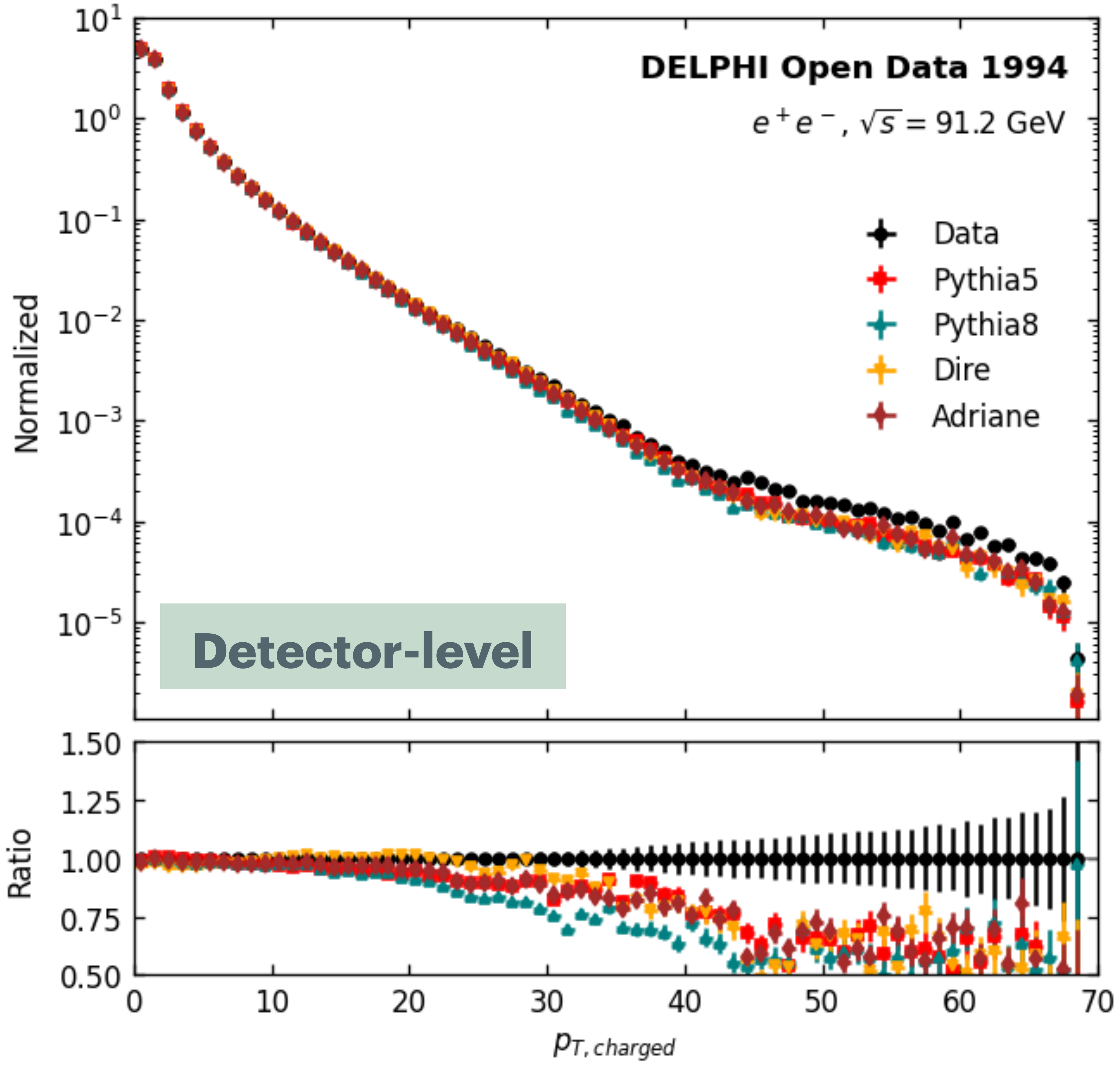
Not shown also major 1994 vs 1995 differences lead to asymmetric and anti-correlated uncertainties

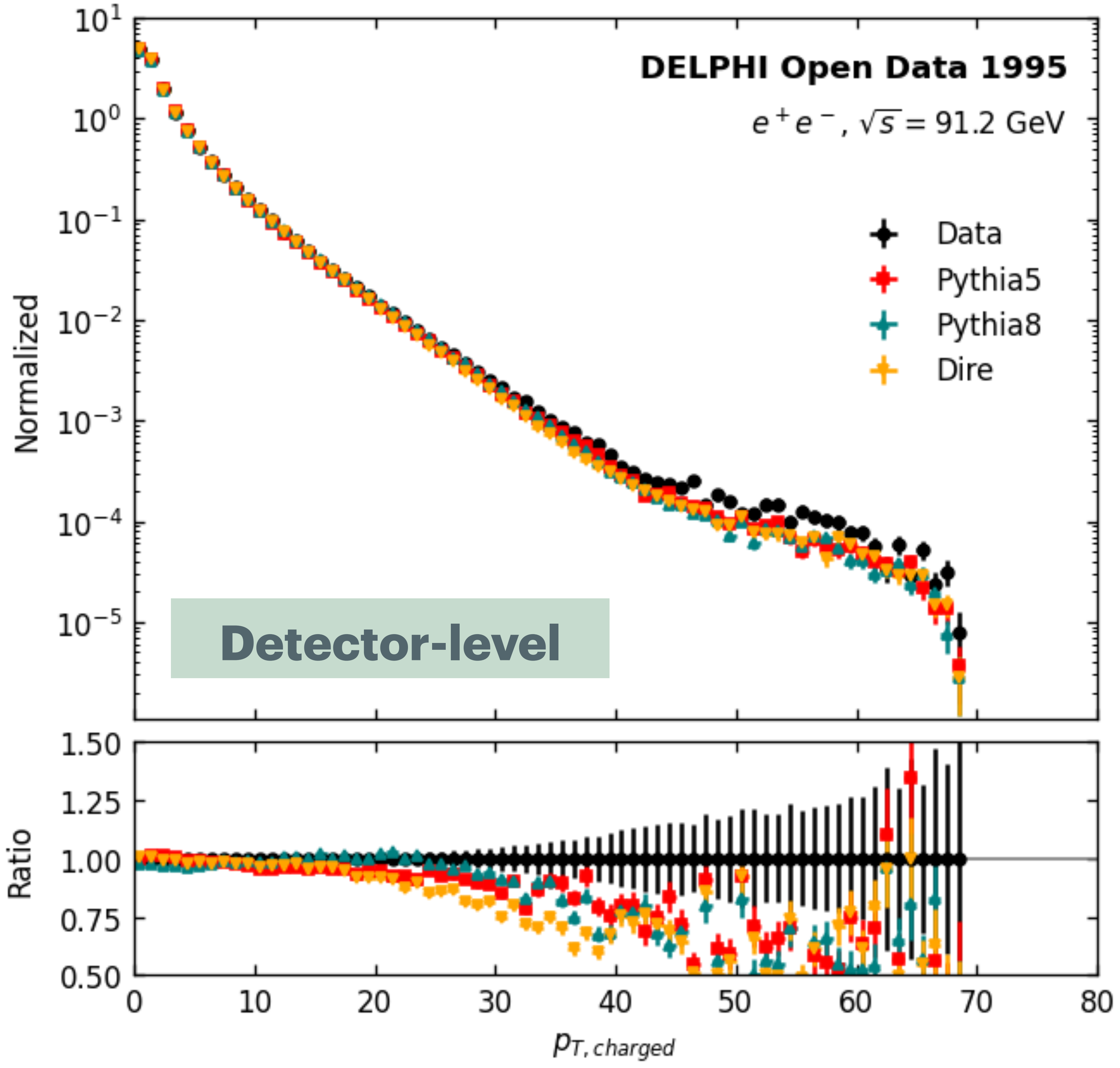


Credit: Jingyu Zhang and Luna (Yi) Chen (Vanderbilt)

Charged Particle Momentum

Credit: Jingyu Zhang, Luna Chen (Vanderbilt)

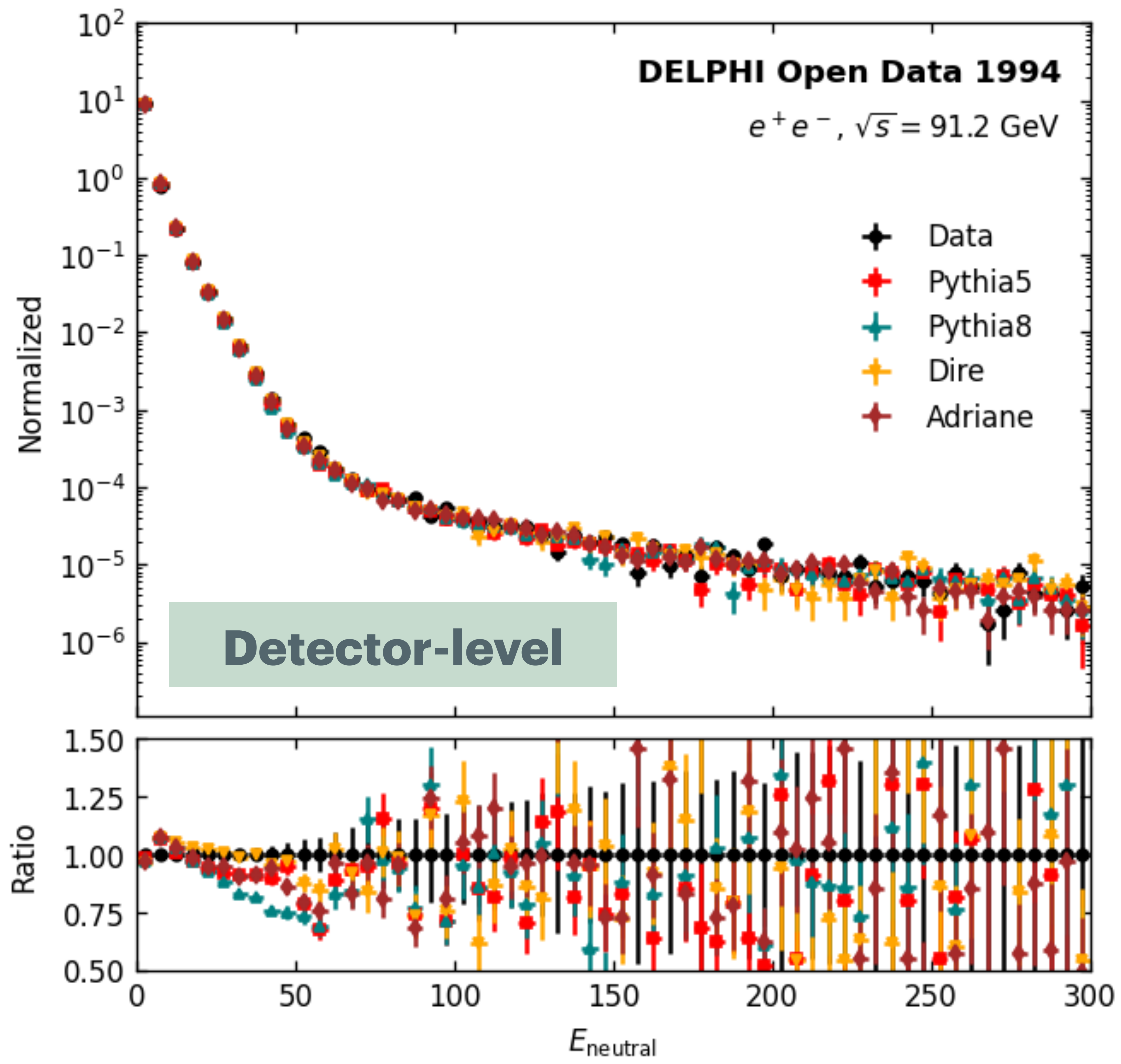


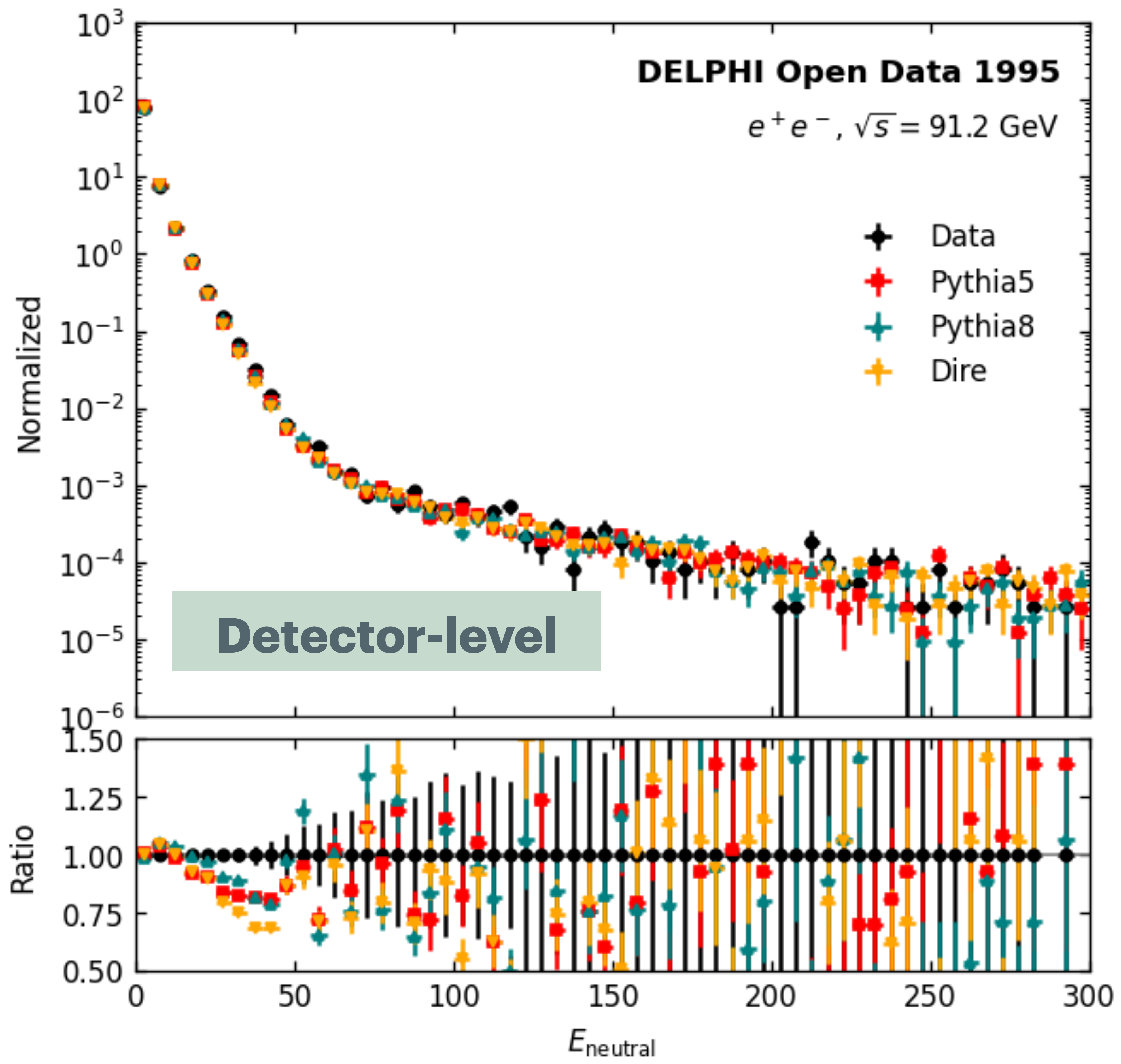


- MC describes charged particle pT (w.r.t beam line) well at low momentum
- Uncertainties included for the potential mismodeling of the detector effects

Neutral Particle Energy

Credit: Jingyu Zhang, Luna Chen (Vanderbilt)



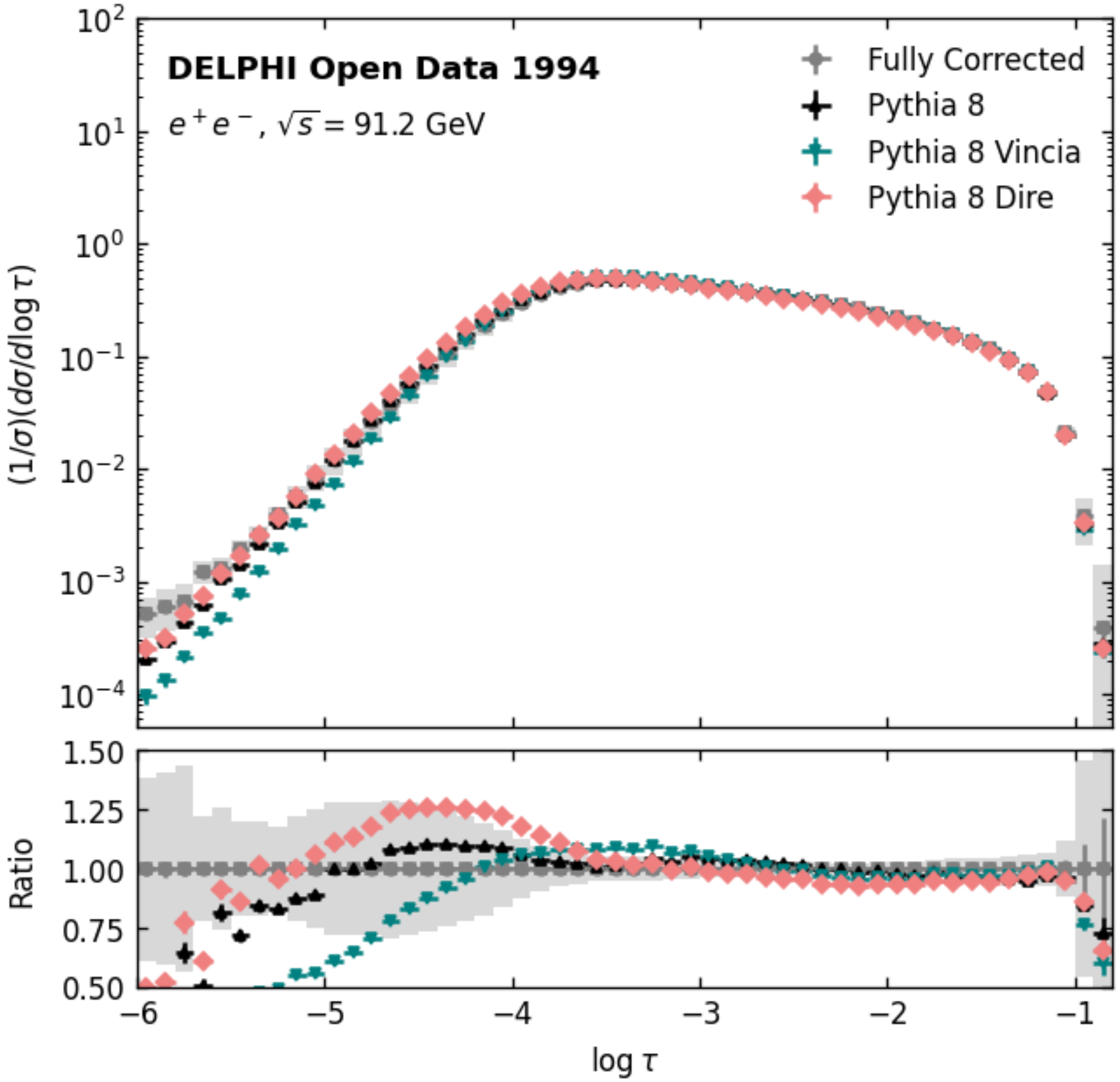


- MC describes neutral energy spectrum well at high energy (1994 vs 1995 diff.)
- Require neutral energy < 50 GeV to reduce noise

Fully Corrected Logarithmic Thrust

- Repeat full analysis (corrections, unfolding, systematics, etc.) with dedicated binning for $\log(\tau)$

- DELPHI successfully probes also deep in the IR regime where the observable $\tau \sim \Lambda_{QCD}/\sqrt{s}$



Analysis Strategy

Credit: Jingyu Zhang,
Luna Chen (Vanderbilt)

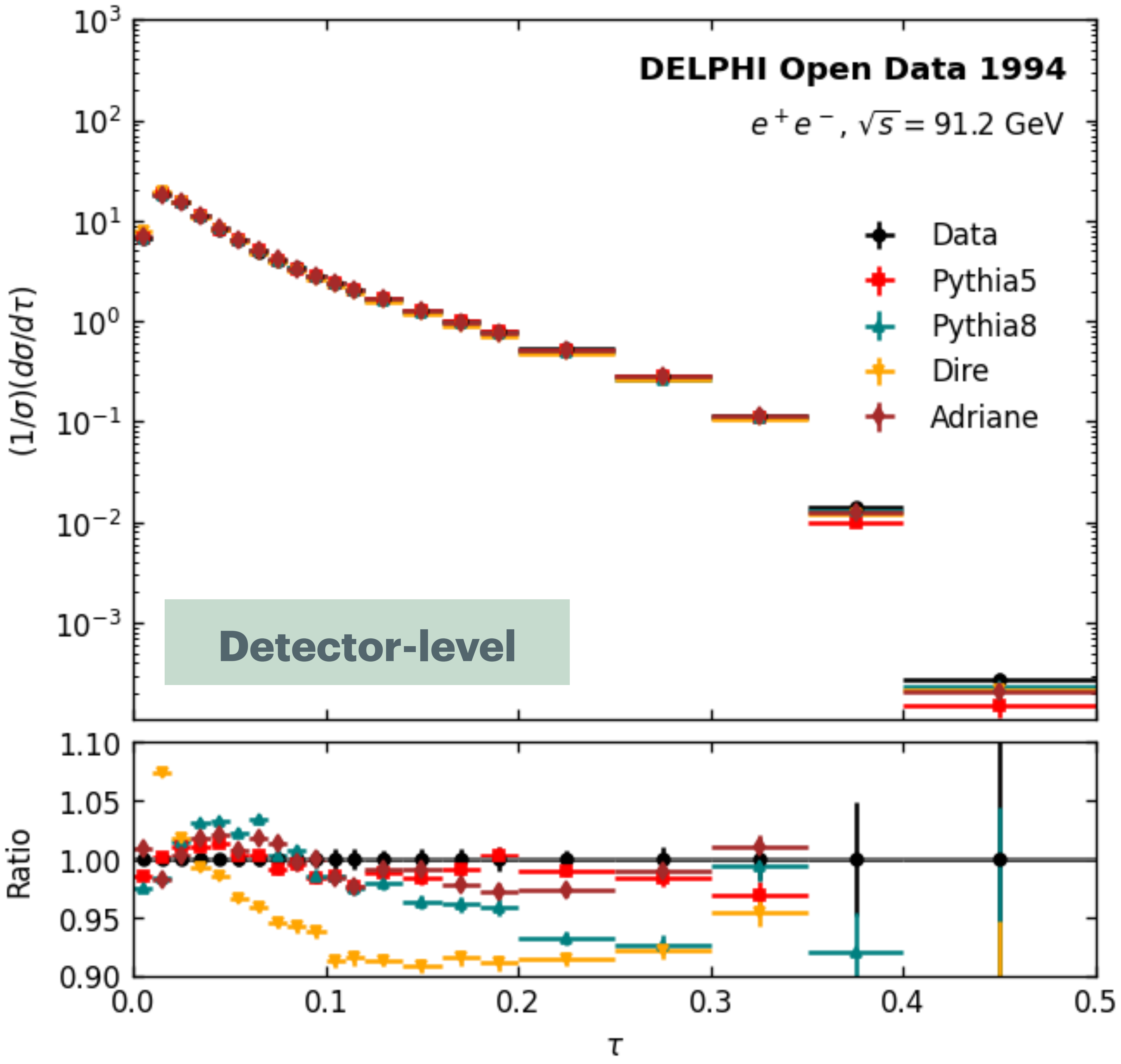
- Thrust calculated with exact algorithm using all particles and missing P
- 1D D'Agostini unfolding to correct distribution in fiducial phase space
- Bin-by-bin correction to full phase space without QED effects
- 4 sets of MC samples used: Pythia 5.7 and ARIADNE (open data with DELPHI tune), Pythia 8.3 and Dire (new MC + DELSIM)

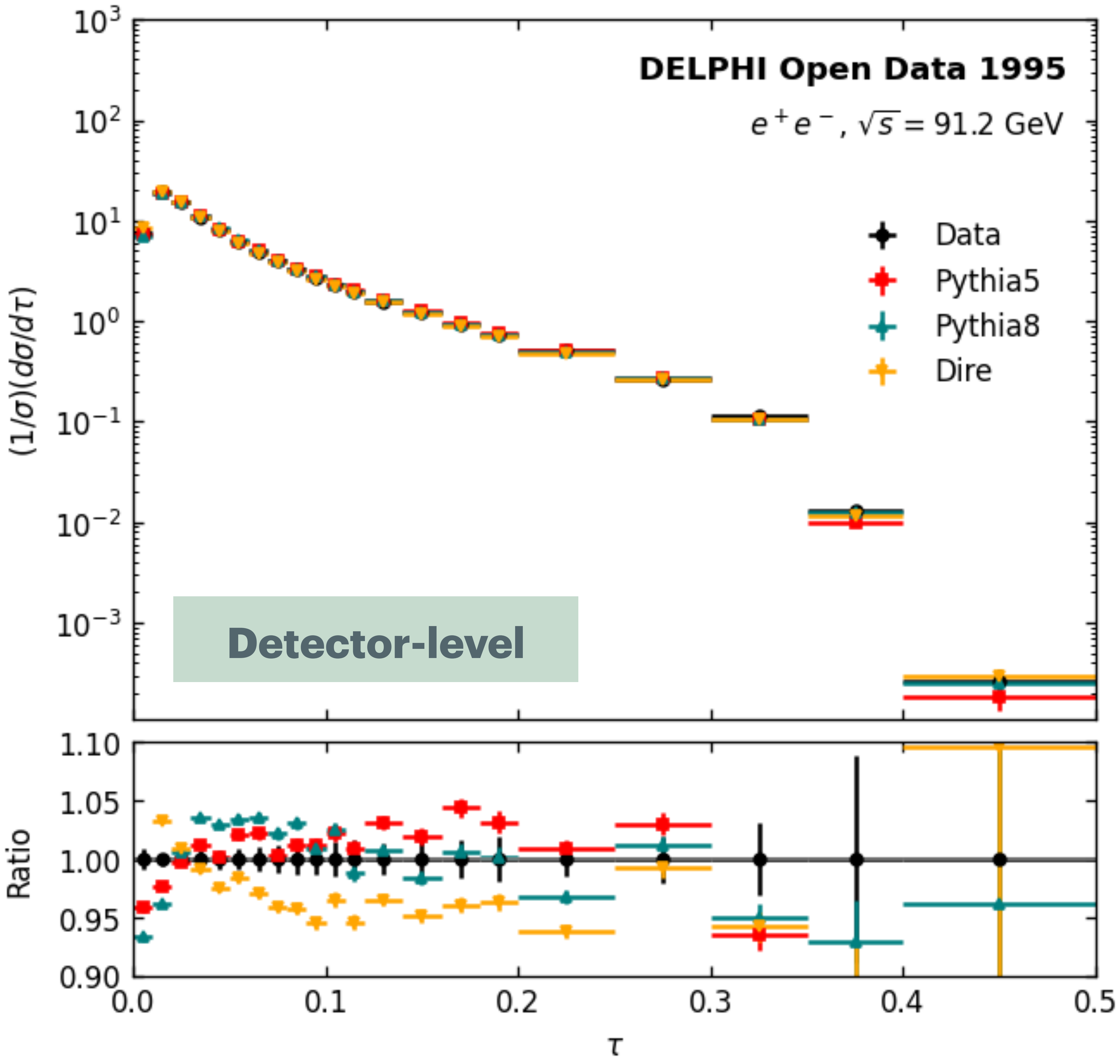
Hadronic Event Selection

neutral particle selection	$E \geq 0.5\,\text{GeV}$ $20° \leq \theta \leq 160°$
charged particle selection	$0.4\,\text{GeV} \leq p \leq 100\,\text{GeV}$ $\Delta p/p \leq 1.0$ measured track length ≥ 30 cm distance to I.P. in $r\phi$ plane ≤ 4 cm distance to I.P. in $z \leq 10$ cm $20° < \theta < 160°$
Standard event selection	$N_{\text{ch}} \geq 7$ $30° \leq \theta_{\text{Thrust}} \leq 150°$ $E_{\text{tot}} \geq 0.50 E_{\text{cm}}$

Uncorrected Thrust Spectrum

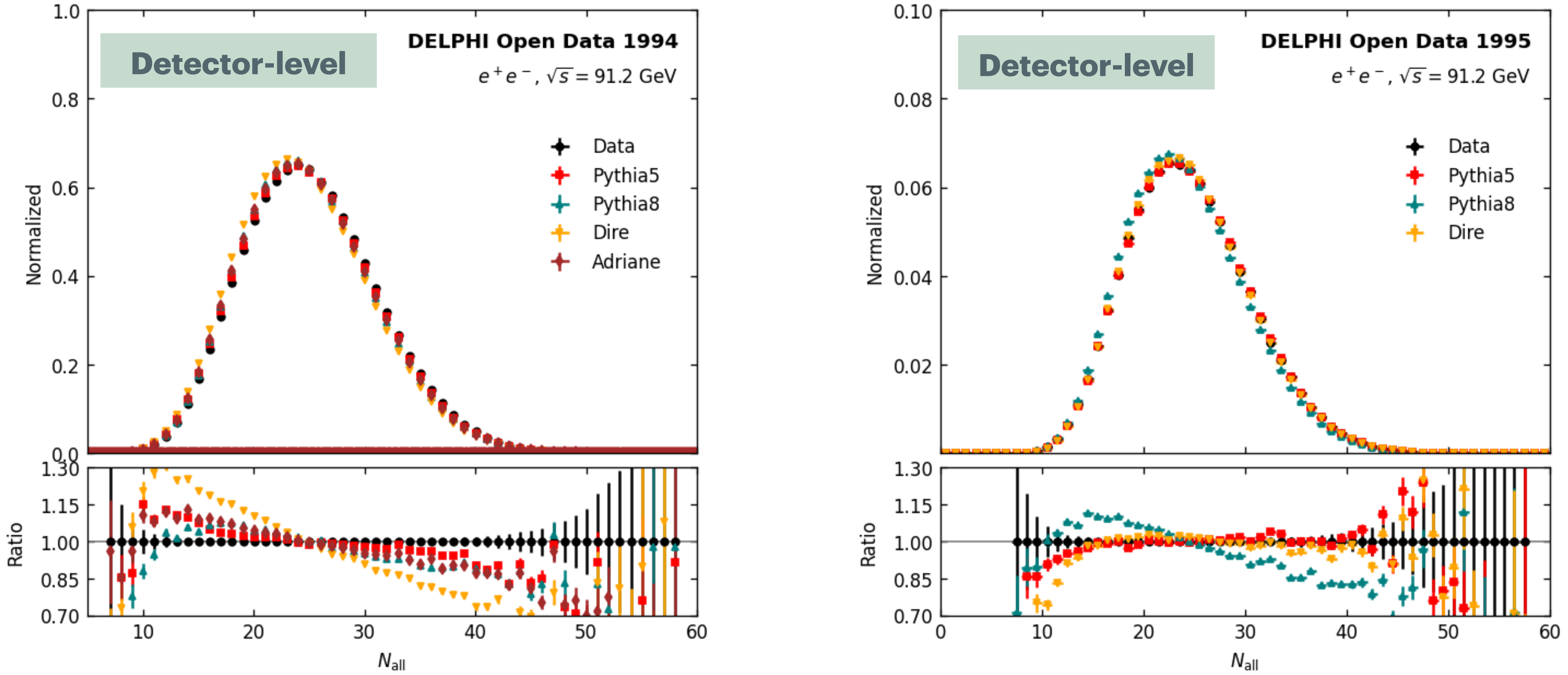
Credit: Jingyu Zhang, Luna Chen (Vanderbilt)





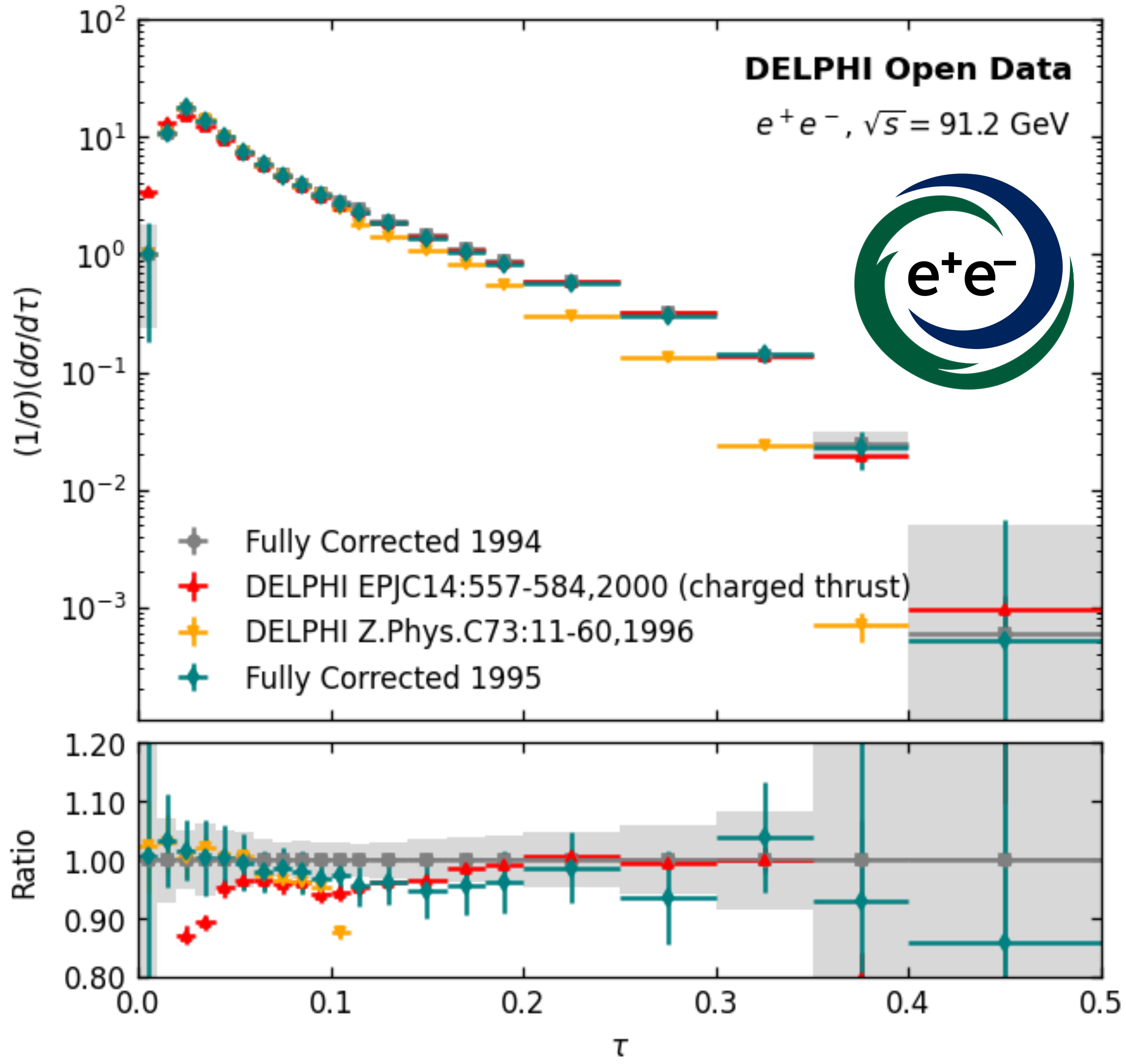
Particle Multiplicity

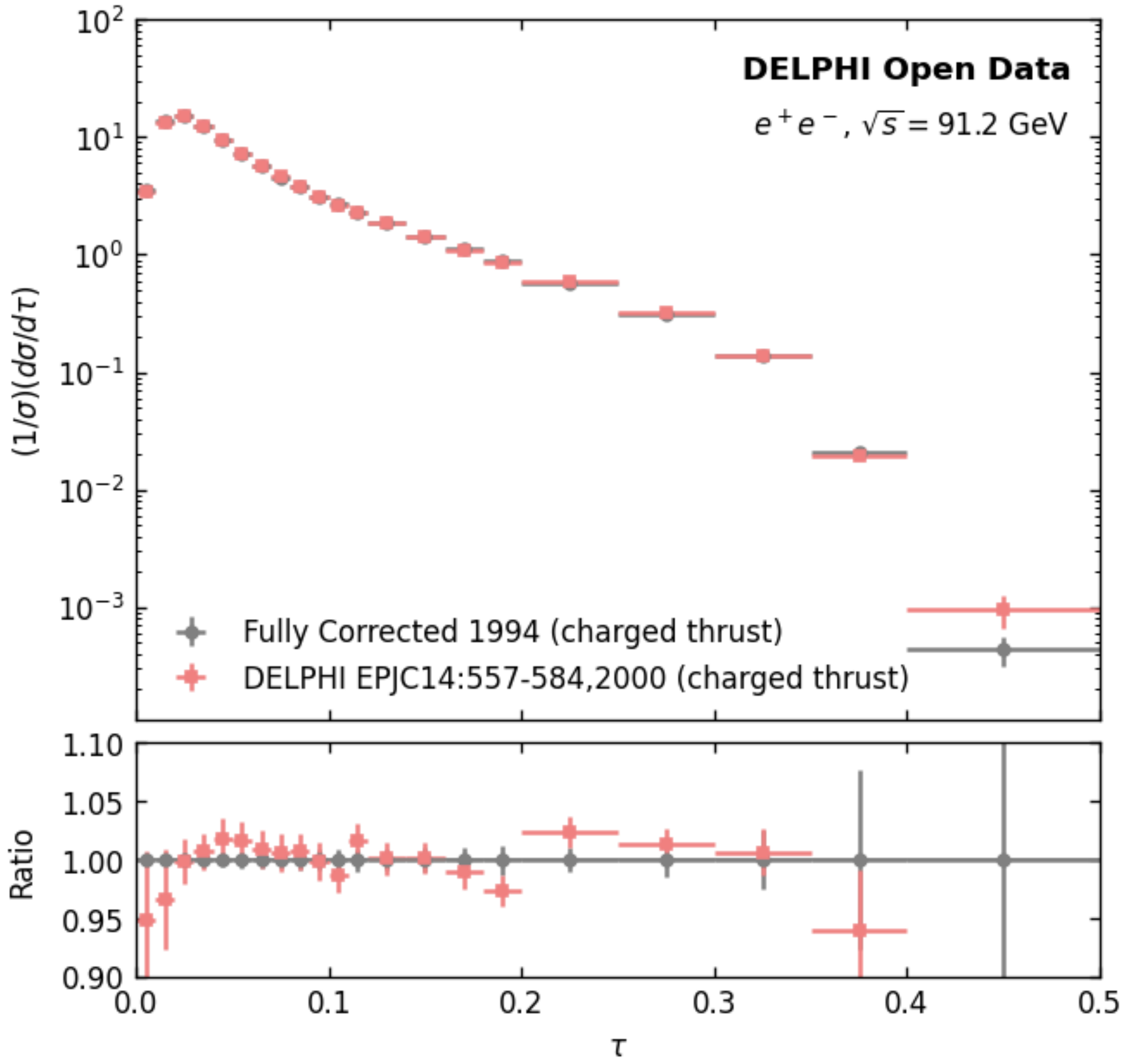
Credit: Jingyu Zhang, Luna Chen (Vanderbilt)



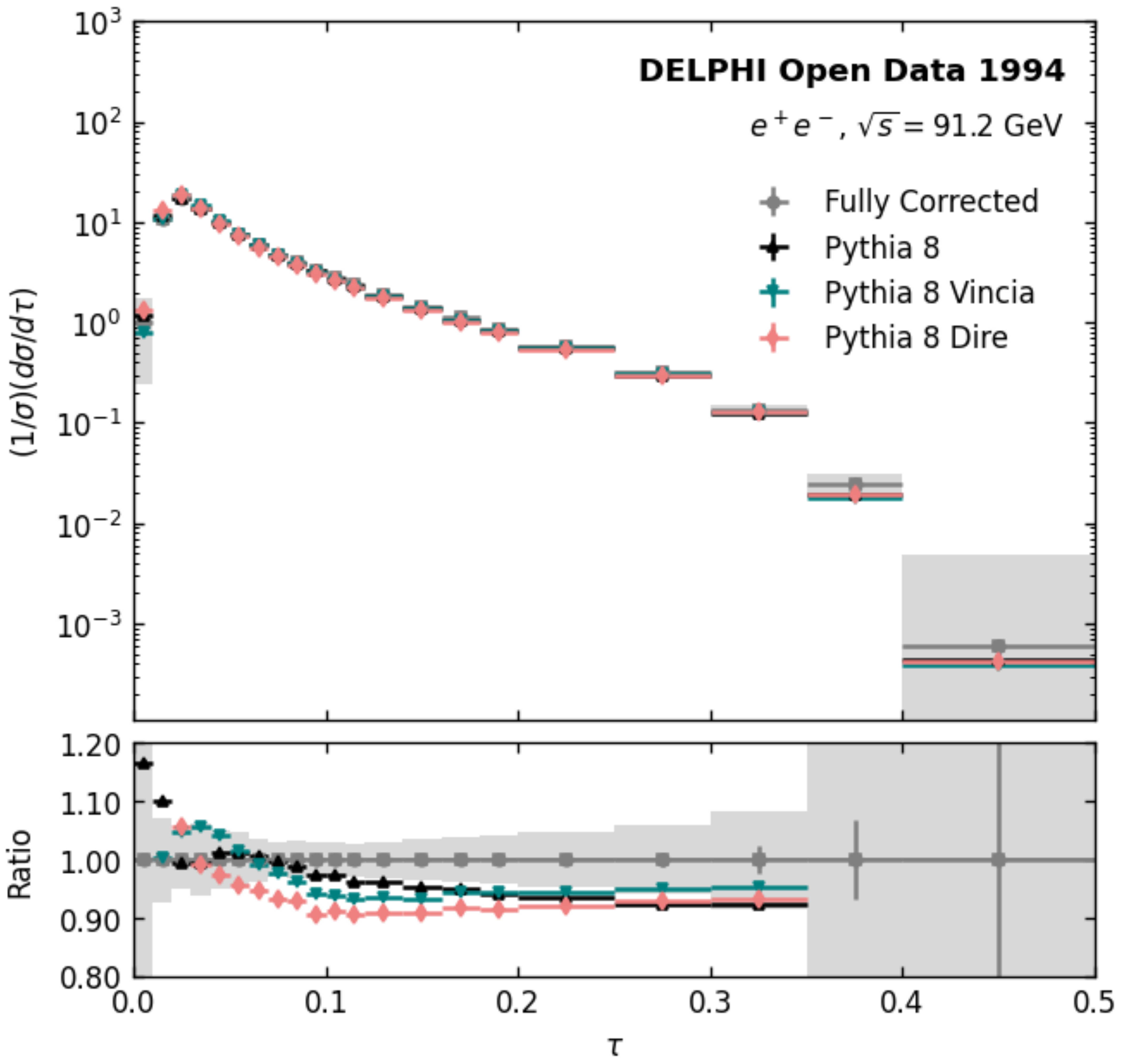
- Tuned Pythia 5 described data the best. Comparable to ALEPH
- Different modeling of 1994 vs 1995 conditions

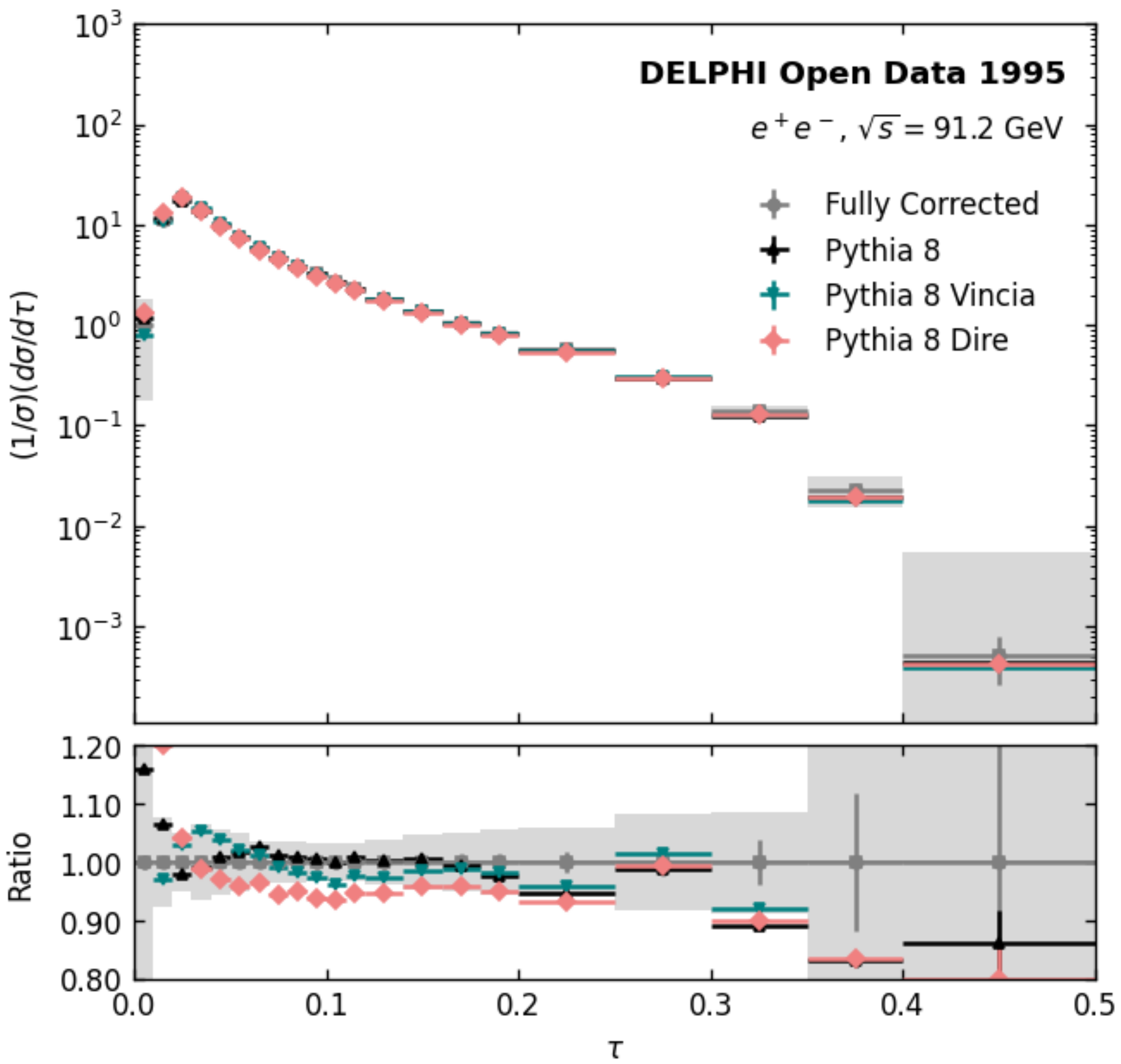
Fully Corrected Thrust Spectrum



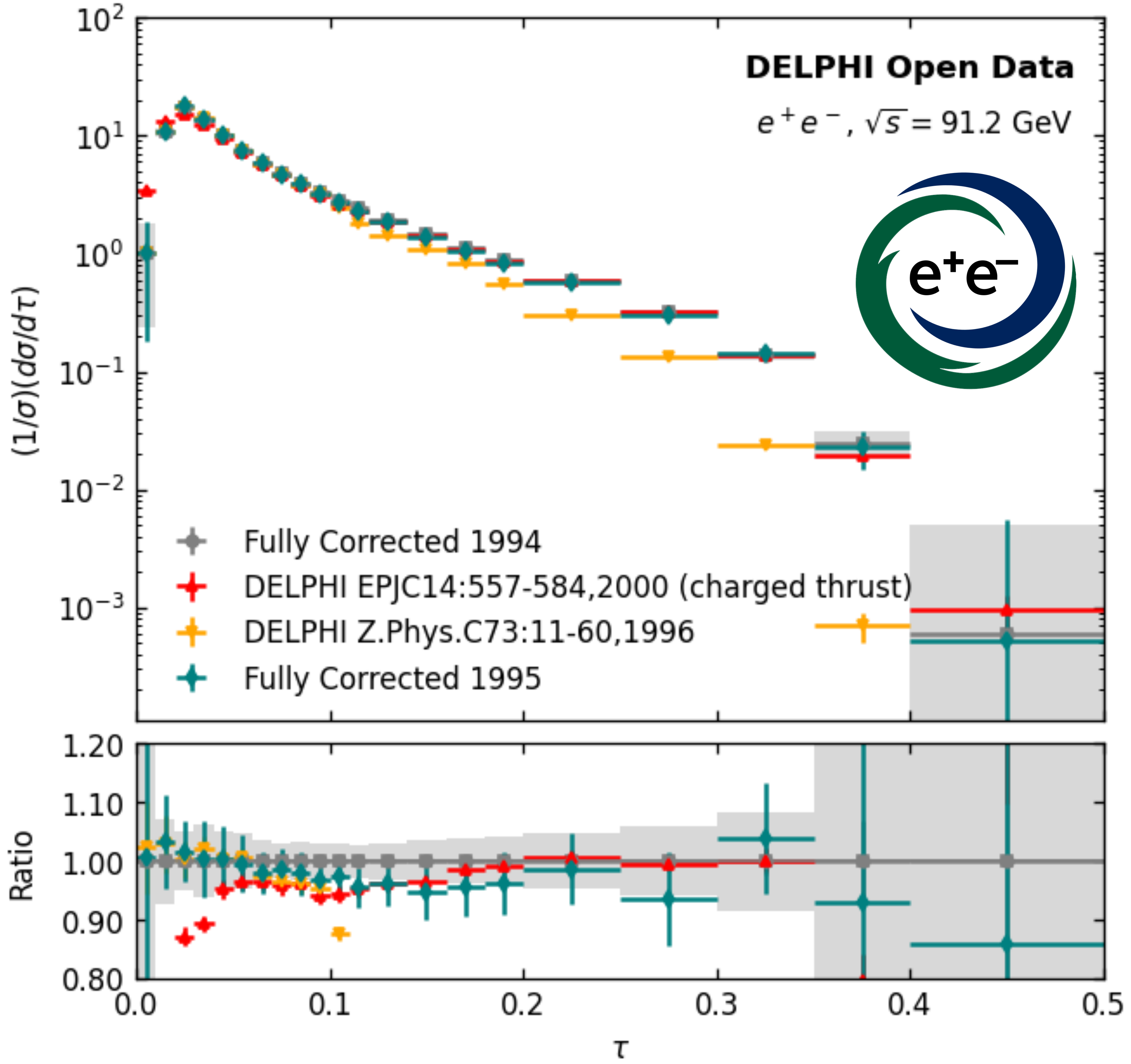


Fully Corrected Thrust Spectrum





Fully Corrected Thrust Spectrum



- For all particle measurement, a systematic shift towards higher thrust w.r.t previous DELPHI result
- For charged particle only measurement, cross-check (no systematics) agrees well with previous DELPHI result

Jingyu Zhang (Vanderbilt)

Luna Chen (Vanderbilt)