Considerations on the combination procedure of correlated errors

Jens Erler

MWDays23

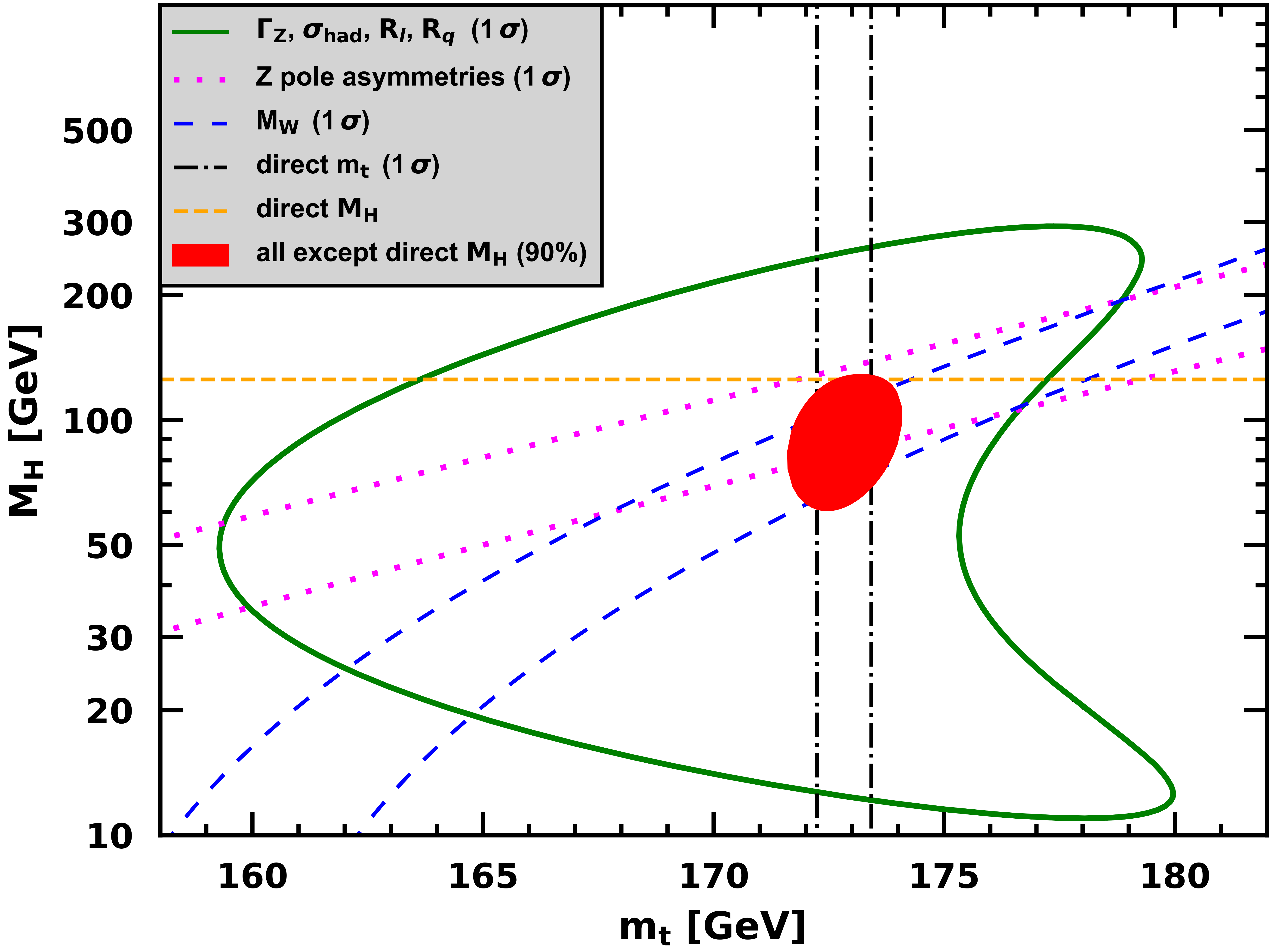
CERN, April 19, 2023

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Content

- Introduction
- Classes of uncertainties
- Correlations
- Simplified approach
- PDG scale factors
- Alternative to PDG scale factors
- Conclusions

M_H – m_t today



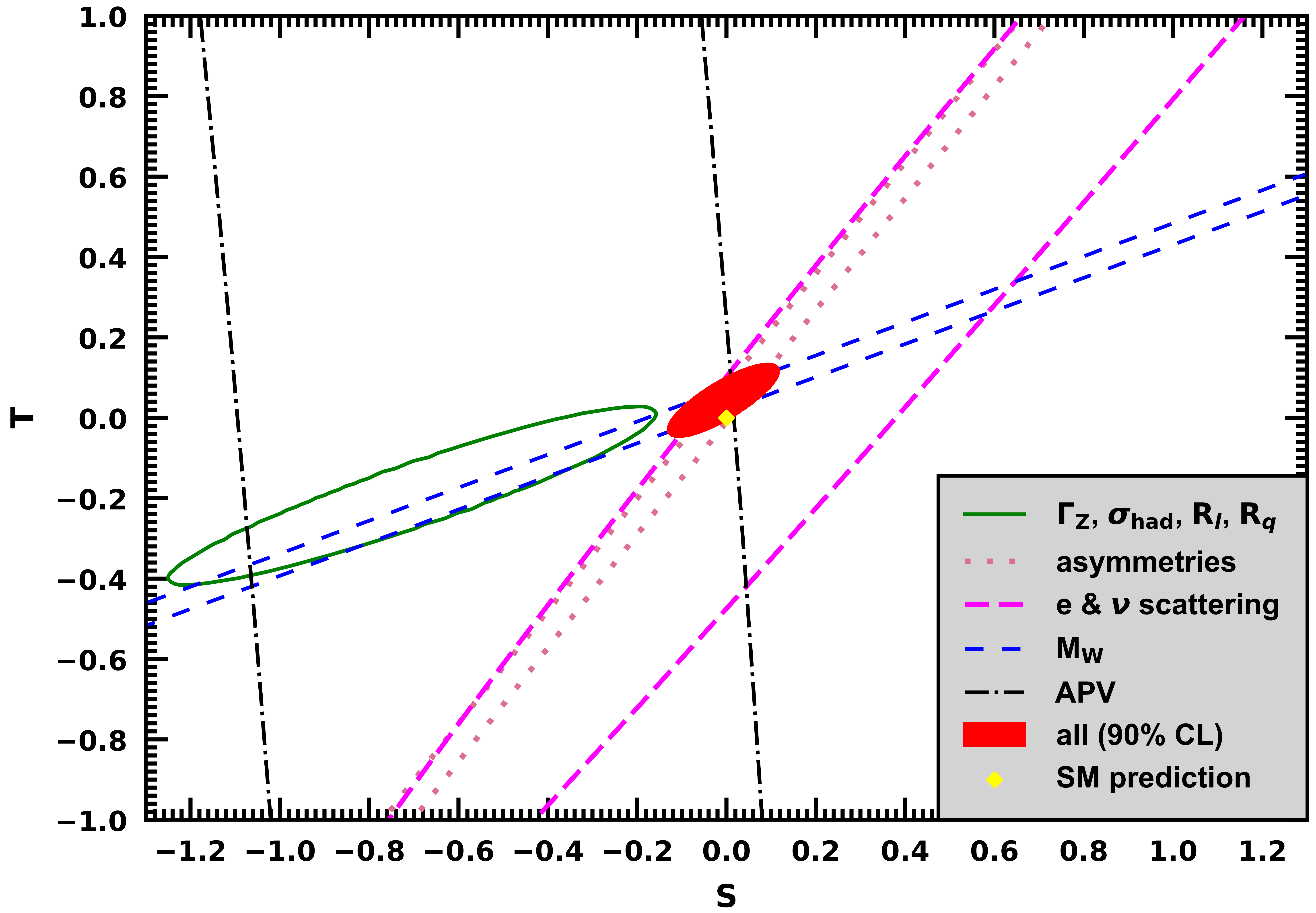
M_W and $\sin^2\theta_W$ at the heart of electroweak fits

χ^2 / d.o.f. = 46.7 / 43
(before g_μ–2 update)

including
correlated theory errors

Freitas & JE (PDG 2022)

S and T



M_W & $\sin^2\theta_W$ provide oblique parameters

(for STU fits need extra constraint such as Γ_Z)

Freitas & JE (PDG 2022)

$\rho = 1.00087 \pm 0.00016$

→ 60 GeV splitting

of an extra color triplet

Classes of uncertainties

- experimental
  - statistical (rigorous, even if highly non-Gaussian or from small data samples)
  - systematic (sometimes determined by auxiliary measurements which may themselves be statistics dominated)
- theoretical
  - parametric (rigorous, even if highly non-Gaussian)
  - genuine theory errors (difficult to assess; usually non-Gaussian)
    - from systematically improvable approximations such as perturbative expansions
  - model errors (extremely difficult; often assumed to be small or simply ignored)
    - range from parametrizations to fundamental limits of the model applicability
- unknown unknowns (just because something is unknown, doesn't mean it's 0)

Correlations

- Of course, at the same level of approximation one also needs to assess correlations.
- In principle, the same issues apply to the off-diagonal entries of the covariance matrix as for the diagonal. However, in practice this is much harder. Last resort measures:
  - ignore correlations even when non-negligible (e.g. PDG m_t combination is a simple weighted average)
  - make ad hoc assumptions such as a common correlation coefficient (can be adjusted so that reduced $\chi^2 = 1$)
  - wait until the appropriate working group has produced the proper combination (so as to ignore the most recent and often most precise measurements)

Correlations: M_W as a prime example

- how to combine different channels?
- how to combine different experiments at the same collider?
- how to combine results from the same experiment from different collider runs
  - e.g. changes in energy, luminosity or polarization and their measurements
  - changes in detectors (deterioration, upgrades, …)
  - changes in analysis (triggers, cuts, new outside inputs, …)
    - can be dealt with within collaboration but is not always done
- how to combine results from different facilities (theory correlations)
- how to combine extractions from different observables (often no correlations)
- still, not all hope is lost…

Quick but not that dirty averages with correlations

- Many measurements are statistics dominated
- Leading systematic (or theory) errors often fully correlated of uncorrelated
- If leading systematic is only partially correlated, it is usually conservative to assume 100% correlation (but interesting anti-correlations may be missed)
- Refinements:
  - iteration: sub-combination of most correlated measurements first (e.g. analysis method → channels → data periods → collider → collider type → all)
  - sophistication: define finer sub-categories of uncertainty
- For PDG 2004 no recent m_t average by the Tevatron EWWG was available, but was needed for the EW fits. This method gave $m_t = 177.9 \pm 4.4$ GeV (December 2003) while hep-ex/0404010 (Tevatron EWWG) found 178.0 ± 4.3 GeV.
- Further simplification: only one (fully) correlated error source taking the smallest

M_W @ LEP [arXiv:1302.3415]

M_W [MeV]	central value	statistical	systematic	total
LEP (threshold scan)	80420	200	30	202
OPAL (leptonic)	80410	410	130	430
LEP (semi-leptonic)	80372	30	21	36
LEP (all hadronic)	80387	40	44	59
LEP	80376	25	22	33

M_W @ Tevatron [arXiv:1307.7627]

M_W [MeV]	central value	statistical	systematic	correlation	total
UA2	80360				370
CDF (4.4 pb^{-1})	79927.7				390
CDF (18.2 pb^{-1})	80377.3				181
CDF (84 pb^{-1})	80470.5			16 (e/μ)	89
DØ (95 pb^{-1})	80478.5				84
DØ (1.0 fb^{-1})	80401.8	21	38	(e only)	43
CDF (2.2 fb^{-1})	80387.3	12	15		19
DØ (4.3 fb^{-1})	80368.6	13	22	(e only)	26
Tevatron	80386.7 ∓ 0.1			10 ± 1	16.0 ± 0.4

M_W @ Tevatron [arXiv:1307.7627]

M_W [MeV]	central value	statistical	systematic	correlation	total
UA2	80360				370
CDF (4.4 pb^{-1})	79927.7				390
CDF (18.2 pb^{-1})	80377.3				181
CDF (84 pb^{-1})	804			16 (e/μ)	89
DØ (95 pb^{-1})	804				84
DØ (1.0 fb^{-1})	80401.8	21	38	(e only)	43
CDF (2.2 fb^{-1})	80387.3	12	15		19
DØ (4.3 fb^{-1})	80368.6	13	22	(e only)	26
Tevatron	80386.7 ∓ 0.1			10 ± 1	16.0 ± 0.4

PDF error of combination
and assumed equal correlation
to reproduce the average

Simplified combination of M_W (PDG 2022)

M_W [MeV]	central value	statistical	systematic	correlation	total
ATLAS	80369.5	6.8	17.2	9 ± 2	18.5
LHCb	80354	23	22	9* ± 2	32
LHC	$80366.1^{+0.3}_{-0.2}$			7 ± 2	16.9 ± 0.4
Tevatron	80387			7 ± 2	16
LEP	80376	25	22	–	33
World	80377.1 ± 0.1				$11.8^{+0.5}_{-0.4}$

Simplified combination of M_W (MWDays23 update)

M_W [MeV]	central value	statistical	systematic	correlation	total
ATLAS (p_T)	80360.1	4.9	15.5	8* ± 2	16.3
LHCb (μ)	80354	23	22	8 ± 2	32
LHC	$80359.0^{+0.2}_{-0.1}$			4 ± 1	15.2 ± 0.3
CDF Run II	80433.5	6.4	6.9	4* ± 1	9.4
DØ Run II	80375			4 ± 1	23
UA2 + Run I	80451			4 ± 1	57
LEP	80376	25	22	–	33
World	$80409.6^{+1.1}_{-0.8}$				7.8 ± 0.2

Simplified combination of M_W (MWDays23 update)

M_W [MeV]	central value	statistical	systematic	correlation	total
ATLAS (p_T)	80360.1	4.9	15.5	8* ± 2	16.3
LHCb (μ)	80354	23	22	8 ± 2	32
LHC	$80359.0^{+0.2}_{-0.1}$			4 ± 1	15.2 ± 0.3
CDF Run II	80433.5	6.4	6.9	4* ± 1	9.4
DØ Run II	80375			4 ± 1	23
UA2 + Run I	80451			4 ± 1	57
LEP	80376	25	22	–	33
World	$80409.6^{+1.1}_{-0.8}$				7.8 ± 0.2

new CDF / DØ combination corresponds to ≈ 5 MeV

Comments

- Use of PDF error as correlation estimate: the PDF error should be expected to be only partially correlated, but this should be partly compensated by further sources of correlations such as from radiative corrections
- The OPAL purely leptonic channel was merged with their semi-leptonic result.
- global fit except M_W and Γ_W: $M_W = 80356 \pm 6$ MeV

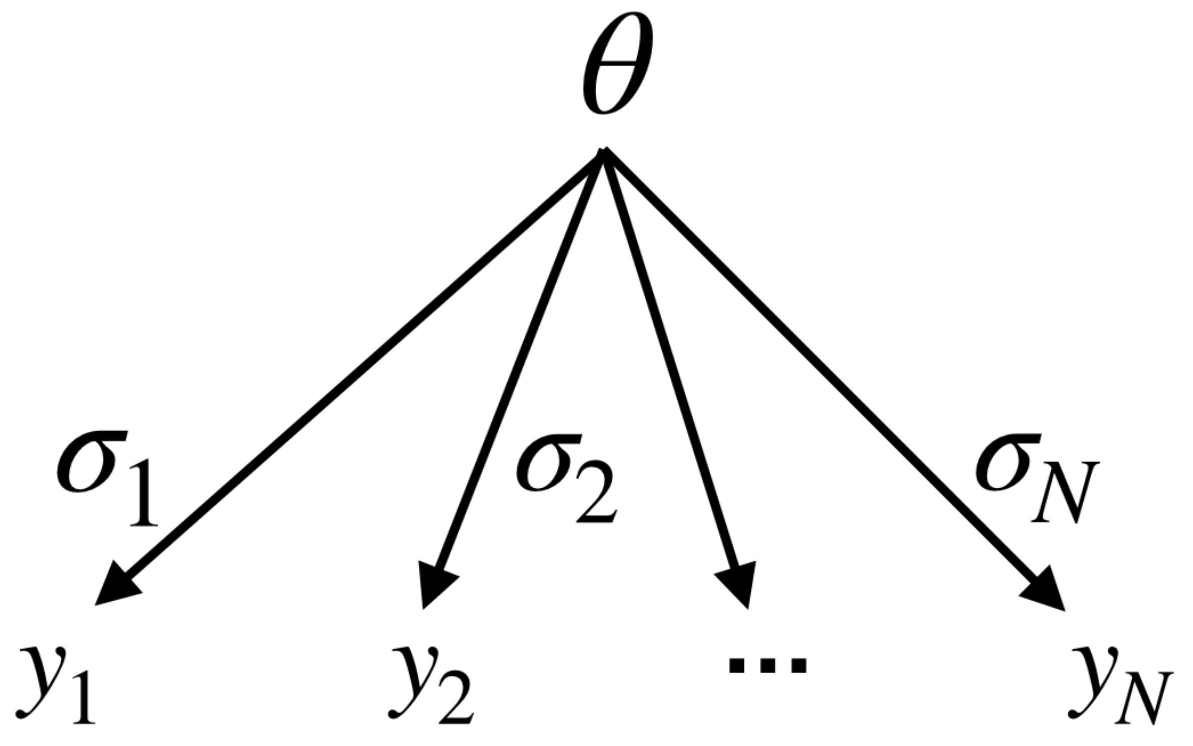
➡ 1.6 σ (5.4 σ) below 2022 (updated) average: $M_W = 80377 \pm 12$ MeV

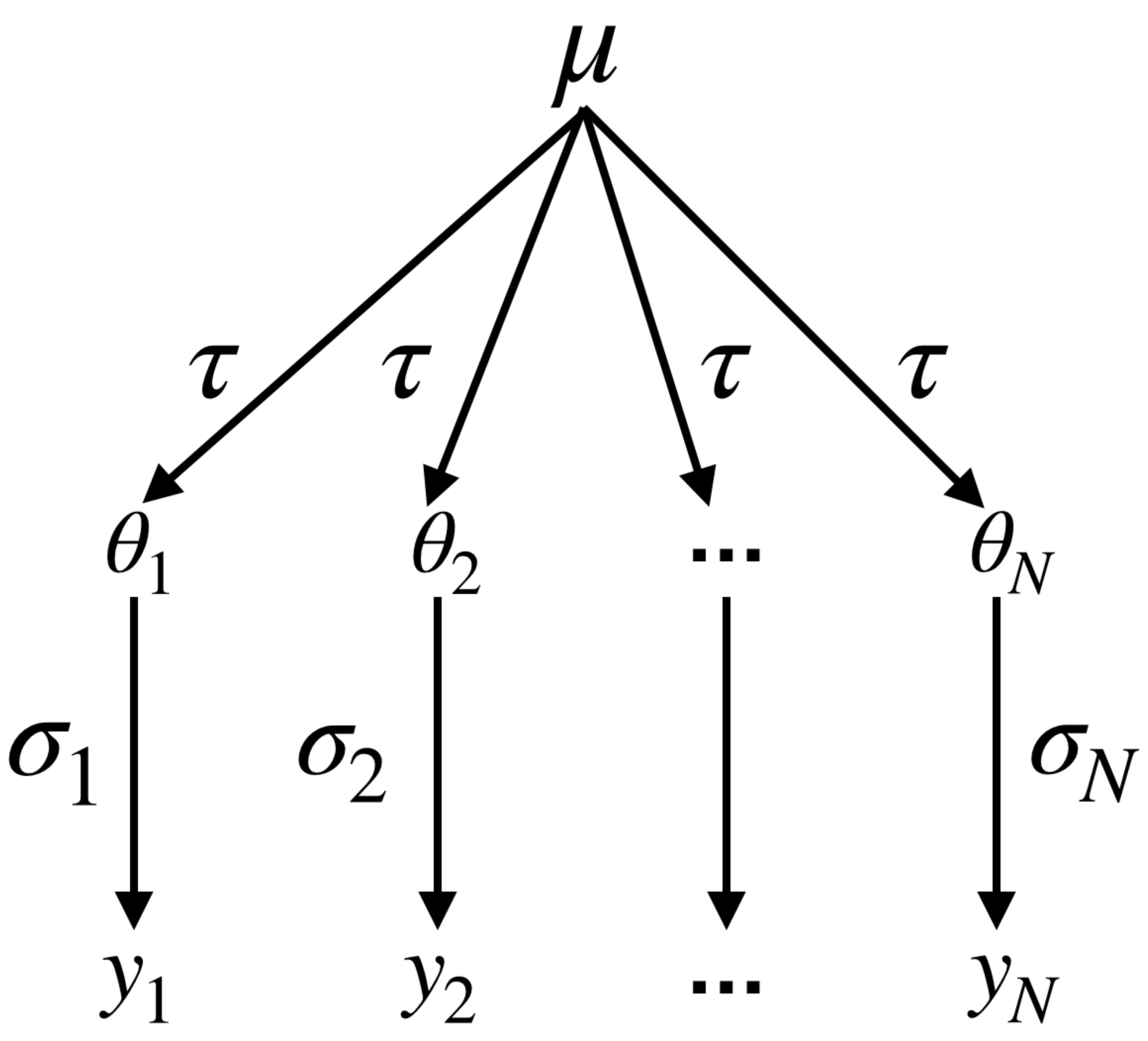
PDG scale factors

- <u>Rules</u>
  - If reduced $\chi^2 > 1$: re-scale all errors by $S \equiv [\chi^2 / N_{eff}]^{-½}$
  - otherwise do nothing
  - if some errors are much larger than some others, throw them out
  - do not change central values
- <u>Problems</u>
  - This is for conservatism … but in the most vulnerable case of only 1 measurement (no control measurement) there will never be $S \neq 1$
  - set of individual data points not well-defined (e.g. pre-averages)
  - if some data have already undergone PDG scaling, the iteration *does* change the central value
  - unscientific: result depends on the order in which information is added

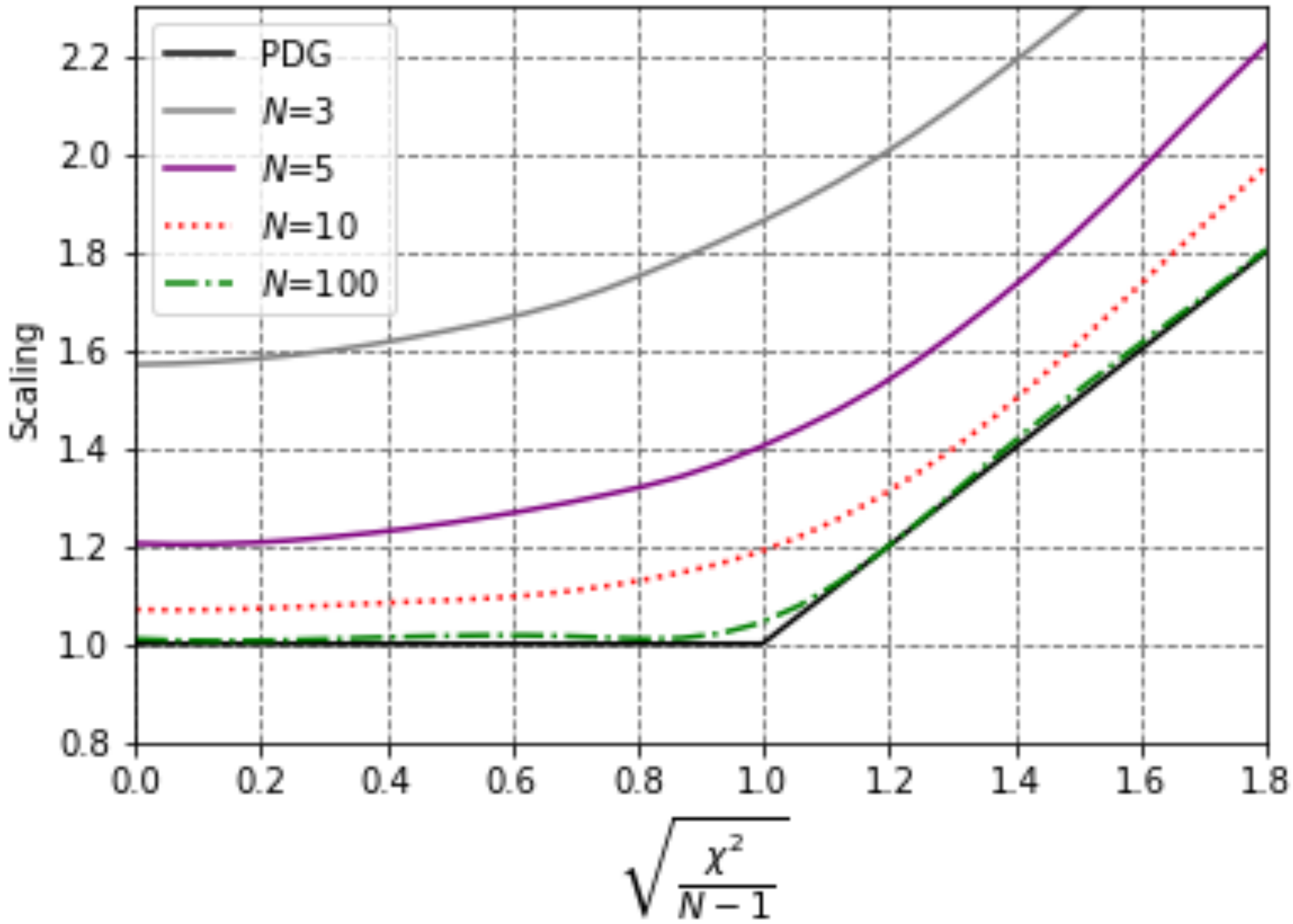
Alternative to PDG scale factors

- Hierarchical Bayesian model
- Idea: individual data points not independently and identically distributed (*iid*), but independently and similarly distributed
- *i.e.* parent distributions are permitted to vary somewhat to allow for unknown effects that could be different from one data point (measurement) to another
- We propose a hierarchical model where each measurement is assumed to determine a different parameter θ_i, each considered as having arisen as a random draw from a common parent distribution with hyper-parameters (μ, τ)
- The μ distribution is obtained by marginalizing over τ



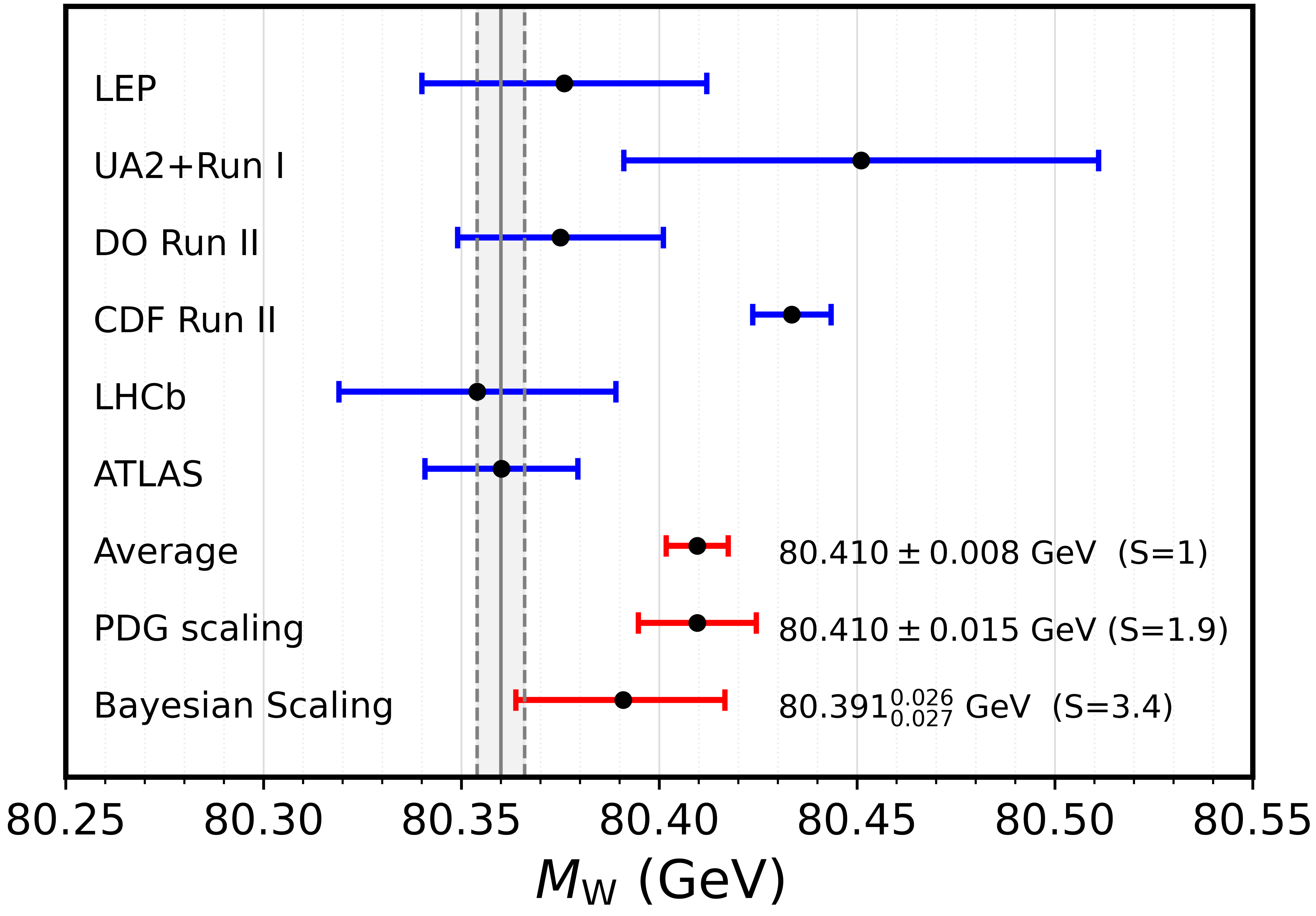


Alternative to PDG scale factors



$\lim_{N\to\infty} S(N) = S_{PDG}$

Multivariate Gaussians



Rodolfo Ferro & JE
Eur. Phys. J. **C 80** (2020) 541

figure by **Rodolfo Ferro**

Conclusions

- complicated averages should be done by the appropriate WG, but global fitters cannot wait for this ($\rightarrow$ Luca Silvestrini)
- in the meantime, averages with simplifying assumptions about correlations work well (better than one would probably expect)
- 13 months ago this gave: $M_W = 80377 \pm 12$ MeV
- MWDays23 update: $M_W = 80410 \pm 8$ MeV
- PDG scale factors is an *ad hoc* procedure and have some problems
- some (not all) of these problems mitigated by hierarchical Bayesian model which in the limit of an infinite number of data points approaches PDG scale factor