Towards the combination of the LHC quarkonium results

A starting point for discussion

W. Adam, R. Frühwirth
Institute of High Energy Physics, Austrian Academy of Sciences

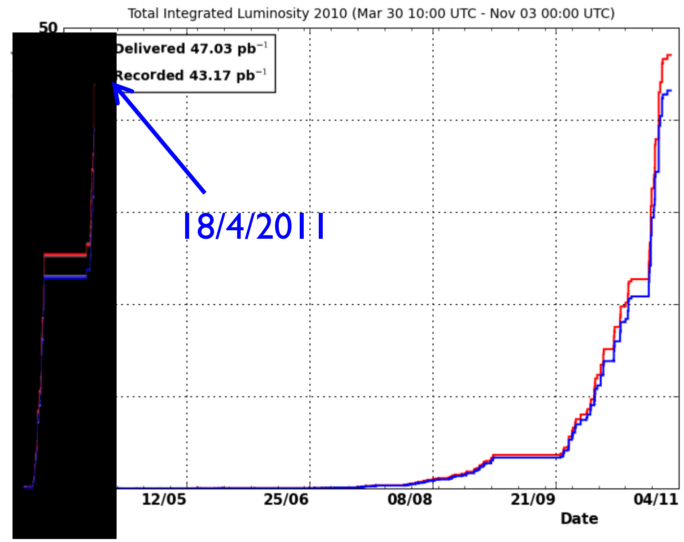
Combinations : why? what? how?

Combinations : why? what? how?

- Even with the results from the Tevatron we are left with many open questions concerning the production of heavy quarkonia states – and in particular of the J/ψ
- The LHC experiments have the possibility to add information and to help to resolve existing contradictions
  – at least from the experimental side!
  - the detectors work perfectly and they profit from larger / complementary ranges of acceptance w.r.t. CDF and D0
- Together we have the obligation of publishing the best possible results
  - in many cases this means – after termination of the work in the individual experiments – a combination

Combinations : why? what? how?

- Priorities are moving to discovery physics
  - the total bandwidth is limited
  - trigger thresholds are rising in order to cope with the luminosity



Integrated luminosity - CMS

- The best candidates for combinations are clearly measurements which are statistics limited or where we could profit from uncorrelated systematics
  - quarkonium + Q / γ / V / quarkonium
  - polarization

Combinations : why? what? how?

- In the following I will use the J/ψ polarization as example
  - the measurement of the polarization is challenging
  - it requires simultaneous determination of multiple parameters in a superposition of several components and in the presence of non-negligible systematic effects
  - Therefore it will be important to perform a detailed comparison of the results of the LHC experiments and a combination in order to reach the ultimate precision

Combinations : why? what? how?

- Obviously one needs agreement on which variables to measure, and in which kinematic regions
  - y and p_T ranges are typically adjusted in order to keep similar statistics and to coincide naturally with the detector segmentation
    - for a combination the results should also be made available in common bins
  - in the case of the polarization experiments need to agree on
    - the reference frames to be used
    - the parameters to be fitted: the classical set λ_θ, λ_ϕ, $\lambda_{\theta\phi}$ or directly using frame-independent variables?
    - concentrating only on the polarization for prompt J/ψs or including also the parameters for the non-prompt component

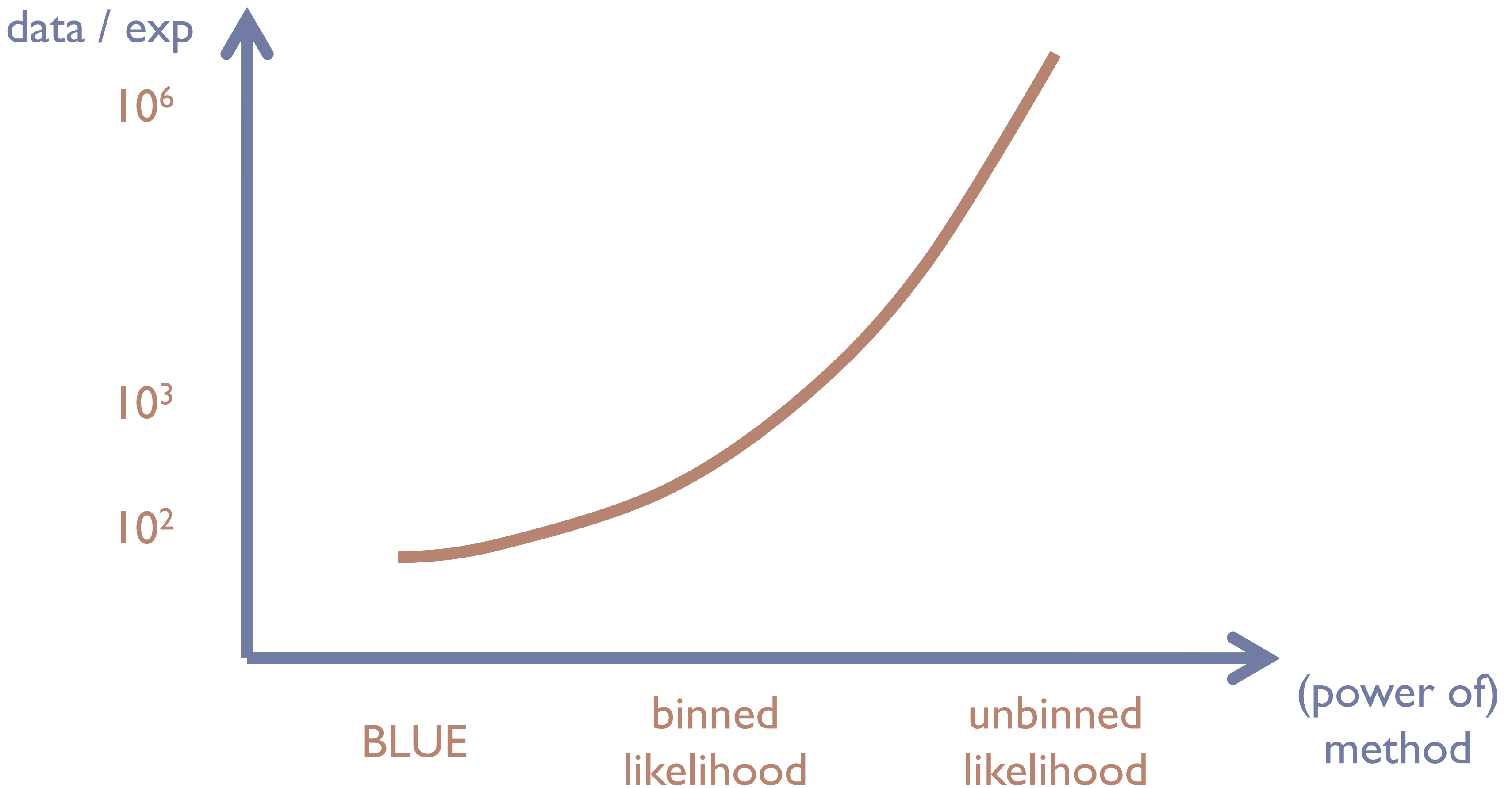
      (only the second choice will make use of the full information)

Combinations : why? what? how?

- The combination of results of (competing) experiments has a long tradition: LEP, Tevatron, B-factories
- We can distinguish two groups, even if the underlying statistical concepts can be in common
  - combination of exclusion limits (e.g., LEP Higgs working group)
  - combination of measurements (e.g., EWK working groups)
- In all cases this required a coordinated effort beyond what is needed for the publication of a single result
  - in particular true with the trend to extract the maximum information from data by the application of more advanced statistical methods
  - the experiments will need to plan in advance in order to approve the necessary information

Complexity of combination

- Ideal case: use the full information also in the combination
  - but this has a high price in terms of data being exchanged



The classical method of combination

- The traditional treatment for the combination of correlated measurements of one or several variables is based on a least-squared ansatz (BLUE)
  - following e.g. Lyons et al (1988) and Valassi (2003)
  - BLUE is optimal if
    - the problem is linear (true for combinations)
    - the distributions are multivariate normal
  - The combination can be done analytically
- Different implementations of BLUE has been applied to dozens of combinations, including complex ones like
  - Z-lineshape: 9 or 11 parameters
  - EW HF sector: 18 parameters
-

	1) R_b	2) R_c	3) $A_\mathrm{FB}^{b\bar{b}}$ (−2)	4) $A_\mathrm{FB}^{c\bar{c}}$ (−2)	5) $A_\mathrm{FB}^{b\bar{b}}$ (pk)	6) $A_\mathrm{FB}^{c\bar{c}}$ (pk)	7) $A_\mathrm{FB}^{b\bar{b}}$ (+2)	8) $A_\mathrm{FB}^{c\bar{c}}$ (+2)	9) $\mathcal{A}_\mathrm{b}$	10) $\mathcal{A}_\mathrm{c}$	11) B (1)	12) B (2)	13) B (3)	14) $\overline{\chi}$	15) $f(\mathrm{D}^+)$	16) $f(\mathrm{D_s})$	17) $f(c_{\mathrm{bar.}})$	18) P
1)	1.00																	
2)	−0.18	1.00																
3)	−0.02	0.01	1.00															
4)	0.00	0.01	0.13	1.00														
5)	−0.10	0.03	0.03	0.01	1.00													
6)	0.07	−0.06	0.00	0.02	0.15	1.00												
7)	−0.04	0.01	0.01	0.01	0.08	0.02	1.00											
8)	0.03	−0.04	0.00	0.01	0.02	0.15	0.13	1.00										
9)	−0.08	0.04	0.01	0.00	0.06	−0.02	0.02	−0.01	1.00									
10)	0.04	−0.06	0.00	0.00	0.01	0.04	0.00	0.02	0.11	1.00								
11)	−0.08	0.05	0.00	0.01	0.00	0.18	0.00	0.07	−0.02	0.02	1.00							
12)	−0.03	−0.01	0.00	−0.02	−0.05	−0.23	−0.03	−0.08	0.02	−0.04	−0.24	1.00						
13)	−0.01	−0.29	0.00	0.02	0.00	−0.21	0.00	−0.14	0.03	−0.02	0.00	0.10	1.00					
14)	0.00	0.02	0.01	0.02	0.11	0.08	0.03	0.02	0.06	0.00	0.29	−0.23	0.16	1.00				
15)	−0.15	−0.10	0.00	0.00	0.01	−0.03	0.01	−0.02	0.00	0.00	0.04	0.02	0.00	0.02	1.00			
16)	−0.03	0.13	0.00	0.00	0.00	−0.02	0.00	−0.01	0.00	0.00	0.01	0.00	−0.01	−0.01	−0.40	1.00		
17)	0.11	0.17	0.00	0.00	−0.01	0.04	0.00	0.02	0.00	0.00	−0.02	−0.01	−0.02	0.00	−0.24	−0.49	1.00	
18)	0.13	−0.43	0.00	0.00	−0.02	0.04	−0.01	0.02	−0.02	0.02	−0.01	0.01	0.13	0.00	0.08	−0.06	−0.14	1.00

EWWG 2006

Classical methods of combination

- Each experiment *i* provides limited input
  - the vector of parameters m_i and the corresponding covariances, split into the uncorrelated and (possibly several) correlated parts

$$m = \begin{pmatrix} m_1 \\ m_2 \\ \dots \end{pmatrix} \qquad C = \sum_{s(ources)} \begin{pmatrix} C_1^s & (\gamma_1^{s^{-1}} \rho_{12} \gamma_2^{s^{-1}})^T & \dots \\ \gamma_1^{s^{-1}} \rho_{12} \gamma_2^{s^{-1}} & C_2^s & \dots \\ \dots & \dots & \dots \end{pmatrix} \qquad \gamma_i = \mathrm{diag}\,(C_{i,11}, \dots, C_{i,nn})^{1/2}$$

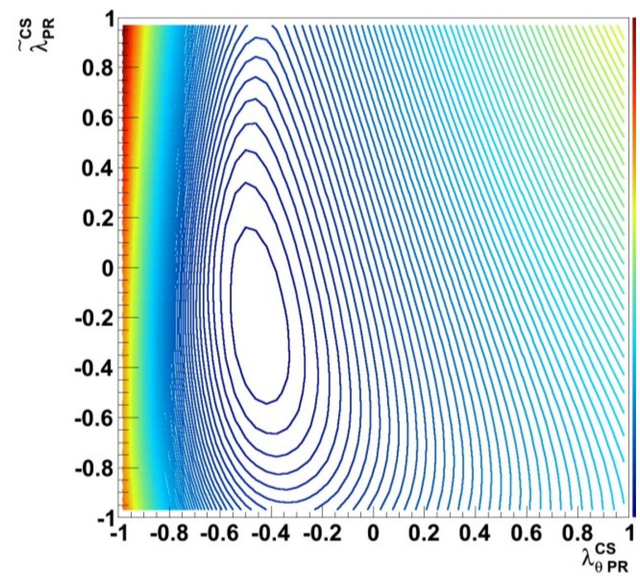
    - uncorrelated parts would include uncertainties on efficiency, scales, ..
    - correlated parts would include the dependence of acc*eff on the assumed p_T spectrum, uncertainties on the NP fraction (in case only P is fitted), …
- Estimation of the covariance matrices
  - uncorrelated statistical errors typically from parabolic approximation at the maximum of $\ln\mathcal{L}$
  - systematics typically from repeated fits or integrated as nuisance parameters
    - especially in the second case toy MC techniques might be necessary to estimate the correlations between experiments

Beyond BLUE?

- Need to combine likelihoods (or, in a Bayesian approach, posterior pdf's)
  - Exchange of likelihood functions for more than 2-3 variables is non-trivial
- How to treat systematics?
  - integrate nuisance parameters into the likelihood
    - the likelihood for experiment *i* can be expressed as a conditional probability w.r.t. the fitted parameters λ and some nuisance parameters n
      $$q_i(m,n \mid \lambda) = r_i(m \mid n,\lambda) \cdot s(n)$$
      the combined likelihood could be written as
      $$p(\lambda) \propto \prod_i r_i(m \mid n,\lambda) \cdot s(n) \cdot p_{prior}(\lambda)$$
  - If this can't be done the likelihoods after (coherent) variations in all experiments will be needed
- Positivity constraints should be applied to the common LH

Parameterized likelihoods

- The shape of ln $\mathcal{L}$ is not necessarily parabolic in the regions covered by the individual experiments
  - correlation terms are not sufficient
- The likelihood can be approximated by generalized distributions



distribution of lnL for a single toy MC dataset (thanks to V. Knünz and L. Gray!)

Parameterized likelihoods

- In recent years statisticians have worked on generalization of the multivariate normal distribution
  - e.g., Azzalini, A. Capitanio J.Roy.Stat.Soc. B61 (98) 3
    introduced a multivariate skew-normal distribution

    $$2 \cdot \Phi_n(\lambda - \lambda_0; C) \cdot \Phi(\alpha^T(\lambda - \lambda_0))$$

    where Φ and Φ_n are the 1- and the n-dim. normal densities and α is a vector of shape parameters
  - the statistical properties of the multivariate skew-normal distribution are known
- functions of this kind could be adapted to the individual ln$\mathcal{L}$ shapes and provide a better basis for combination at a limited cost in terms of data to be exchanged

Full combination of likelihoods

- By providing inputs for performing a binned LH fit
  - a reasonable compromise using (almost) the full information
  - $\cos\theta$-φ bins need to be adapted to the statistics
    - but can be different for different experiments

- By providing per-event information for unbinned fits
  - while technically feasible (e.g., using a RooDataSet) it would require exchanging large volumes of data
  - traditionally collaborations are not in favour of exchanging information at this level of detail

- Systematics
  - as mentioned before either conditional likelihoods with common priors for the nuisance parameters or several variations of the distribution will be needed

Conclusions

- The combination of LHC results can provide substantial gains in precision for the determination of the polarization and for more exclusive final states
- Combinations will need some additional effort and co-ordination
  - we will have to convince the management of the experiments
  - we need to agree on the basics: choice of parameters, kinematic ranges, ..
- Several statistical methods have been used in the past, but we also have new tools at our disposal
  - first tests could be done with a simple BLUE approach,
    in close contact with the statistic committees of the experiments
- It is essential to start the activity soon
  - while the collaborations are still focussed on these analyses
  - while the experts can contribute to the combination