New measures of the quality and of the reliability of fits applied to forward hadronic data at $t = 0$ [1]

J.R. Cudell
Inst. de Physique, Bât. B5, Univ. de Liège, Sart Tilman, B4000 Liège, Belgium

K. Kang
Physics Department, Brown University, Providence, RI, U.S.A.

V.V. Ezhela, Yu.V. Kuyanov, S.B. Lugovsky, N.P. Tkachenko
COMPAS group, IHEP, Protvino, Russia

P. Gauron, B. Nicolescu
LPTPE, Université Pierre et Marie Curie, Tour 12 E3, 4 Place Jussieu, 75252 Paris Cedex 05, France

(COMPETE collaboration)

Abstract

We develop five new statistical measures of the quality of fits, which we combine with the usual confidence level to determine the models which fit best all available data for total cross sections and for the real part of the forward hadronic amplitude.

[1] presented by J.R. Cudell at the 6^{th} workshop on non-perturbative QCD, American University of Paris, 5-9 June 2001.

Phenomenological studies compare models with data, in order to determine whether these agree together. The typical criterion is that there is good agreement if the χ^2 per degree of freedom (χ^2/dof) is of the order of 1, or if the confidence level is bigger than a typical value. To take an explicit example, one can imagine fitting all available elastic hadron-hadron amplitudes (their imaginary part being provided by the total cross sections, and their real part by the ρ parameters) to analytic parametrisations[1]. Although we shall only treat this case explicitly, the remarks and tools given here do apply to most situations.

There are several problems with this conventional approach: first of all, models usually do not apply everywhere, hence the comparison should hold only for part of the data. In the case of hadronic amplitudes, the models are smooth analytic functions, which are expected to work at high enough energy. Some of these models are also expected to have large (unitarising) corrections at large energies. Hence in this simple case, there is a range of energy over which the comparison should be meaningful. However, as is often the case, the exact range over which the models hold is not predicted by the theory, and should emerge from the fits themselves.

Secondly, once one fits many points, a large discrepancy between the theory and a few of the data points can be overshadowed by a good overall agreement. This has the drawback that it is precisely these points which may point out to new physics, but at the same time this may reveal problems with the data. Hence some uniformity in the description of the data is needed, and studies such as the present one, applied to soft hadronic amplitudes, lead to a reassessment of the data used in the fit.

Thirdly, the fits to models lead to some values for the physical parameters of the model. Here, one must take into account the interplay between these values and the data sub-sample for which the model applies. This means that if one changes that sample slightly, the parameters extracted for the models should be stable. A typical counter-example[2] is found in the fits to a simple-pole pomeron, which give (wrongly) an increasing pomeron intercept once the minimum energy of the data is below 9 GeV.

Because we want to be able to consider a large data sample and many possible models, we are aiming at the development of an automatic decision-making procedure, and hence we want to *measure* the above criteria. Although we are not entirely finished with this program, we can present what seems a reasonable set of measures which reflect the above aspects of the fits. All these measures, or *indicators*, are constructed so that the higher their value the better is the quality of the data description.

The first indicator concerns the sample of data that can be fitted. In the case of hadronic amplitudes, we shall consider the range of energies where a given model applies, defined as the region in energy where the fit has a confidence level (CL) bigger than 50%. Its size will be one of the measures of the quality of the fit: we define the **applicability** A of model M as:

$$A_j^M = w_j \log\left(\frac{E_j^{M,high}}{E_j^{M,low}}\right), \quad A^M = \frac{\sum_j A_j^M}{N_{sets}} \tag{1}$$

where $E_j^{M,high}$ (resp. $E_j^{M,low}$) are respectively the highest and lowest values of the energy in the area of applicability of model M in the data subset j and w_j is the weight determined from the best fit in the same interval. After we have defined where the model may work, we can check how well it fits, although by definition all models will provide a satisfactory fit. We may consider the usual confidence level, $C_1^M = CL(\%)$, where the CL refers to the whole area of applicability of the model M, or a reduced one C_2^M limited to the intersection of the areas of applicability of all models qualified for the comparison.

The next measure of quality has to do with the number of parameters of the model, given the number of data points in the range of applicability. Hence we define the **rigidity** R_1 as:

$$R_1^M = \frac{N_{dp}^M(A)}{1 + N_{par}^M} \tag{2}$$

All the information on a given fit is contained in the error matrix, and we shall use it to define the new measures. Hence the first condition to check is whether the error matrix itself is reliable, *i.e.* whether the correlations between parameters are minimal. Hence we define the **reliability** R_2 as:

$$R_2^M = \frac{2}{N_{par}(N_{par}-1)} \cdot \sum_{i>j=1}^{N} \Theta(90.0 - C_{ij}^M) \tag{3}$$

where C_{ij}^M is the correlation matrix element in % calculated in the fit at the low edge of the applicability area.

We are now in a position to define the stability of the model with respect to the data range considered. In the case of hadronic amplitudes, three possible changes can be considered: we can consider the variation that comes from modifying the energy threshold of the fit (**energy stability** S_1), or the fluctuation of the χ^2 from bin to bin for some data binning motivated by physics (in the case of hadronic amplitudes, we bin according to the process/observable) (**uniformity** U), or the reproducibility of the parameters values when fitting, with the same number of adjustable parameters, a reduced data sample and a reduced number of observables, in the case of hadronic amplitudes when excluding the real part (**r-stability** S_2). The latter is introduced in this case because the data for ρ parameter data may be less reliable than those for the cross section. Hence we obtain the three measures:

$$\begin{aligned}
\frac{1}{S_1^M} &= \frac{1}{N_{steps}N_{par}^M}\sum_{steps,ij}(P^t-P^{step})_i(W^t+W^{step})^{-1}_{ij}(P^t-P^{step})_j \\
\frac{1}{U^M} &= \frac{1}{N_{sets}}\sum_j \frac{1}{4}\left[\frac{\chi^2(t)}{N^t_{nop}}-\frac{\chi^2(j)}{N^j_{nop}}\right]^2 \qquad (4)\\
\frac{1}{S_2^M} &= \frac{1}{2N_{par}^M}\sum_{ij}(P^t-P^{t(no\,\rho)})_i(W^t+W^{t(no\,\rho)})^{-1}_{ij}(P^t-P^{t(no\,\rho)})_j,
\end{aligned}$$

where: P^t is the vector of parameters values obtained from the model fit to the whole area of applicability; P^{step} is the vector of parameters values obtained from the model fit to the reduced data set on the *step*, in our case *step* means a shift in the low edge of the fit interval to the right by 1 GeV; W^t and W^{step} are the error matrix estimates obtained from the fits to the total and to the reduced data samples from the domain of applicability, t denotes the total area of applicability, and $t(no\ \rho)$ the data sample with ρ data excluded.

Having defined these measures of the quality of fits, we want to use them to see whether we can decide which is the safest model to use to reproduce a given set of data. As already emphasized[2] in the case of hadronic amplitudes, standard methods do not allow one to decide which models are to be preferred, and several classes of parametrisations are possible. Using these new measures, we can try to decide which models are best. All models considered are the sum of several terms: the low-energy sector is described by an amplitude with the following imaginary part, with $s_1 = 1$ GeV^2:

$$Im(A^{ab}) = Y_1^{ab}\,(s/s_1)^{\alpha_1} \mp Y_2^{ab}\,(s/s_1)^{\alpha_2} \qquad (5)$$

The first term has charge-conjugation $C = +1$ whereas the second has $C = -1$ (with the $-$ sign for a positively charged beam). These two terms are symbolized by the notation RR in the following. The high-energy behaviour is dominated by a pomeron term, for which we consider the following terms, or their combination, in the imaginary part of the amplitude:

$$
\begin{aligned}
Im(A^{ab}) &= X^{ab}\,(s/s_1)^{\alpha_\wp} && (6)\\
Im(A^{ab}) &= Z^{ab}s && (7)\\
Im(A^{ab}) &= B^{ab}s\ln{(s/s_1)} && (8)\\
Im(A^{ab}) &= B^{ab}s\ln^2{(s/s_1)} && (9)
\end{aligned}
$$

which we denote respectively by E, P, L and $L2$. If we take for the pomeron a simple pole model we obtain in this notation RRE, whereas a double pole gives $RRPL$ and a triple pole $RRPLL2$, which we can write as $RRPL2$ with a scale s_0 instead of s_1 in the $\log^2$.

Furthermore, we have considered several possibilities to constrain the parameters. The following notations are attached as either superscript or subscript to the model variants in each case:
d means degenerate leading reggeon trajectories $\alpha_1 = \alpha_2$;
u means universal (independent of projectile hadron);
nf means that we have not imposed factorization for the residues of the pomeron term(s) in the case of the $\gamma\gamma$ and γp cross sections;
qc means that a quark counting rule is imposed on the residues of the amplitude for Σp scattering, constrained by the residues in pp and Kp;
c implies the use of the Johnson-Treiman-Freund relation for the cross section differences: $\Delta\sigma(N) = 5\Delta\sigma(\pi), \Delta\sigma(K) = 2\Delta\sigma(\pi)$.
Finally, the real parts of the amplitudes can be obtained through $s \to u$ crossing.

The first results are quite generic and based on a study of the χ^2 alone:

(1) All analytic descriptions of the data based on the above terms break down at $\sqrt{s} \leq 4$ GeV;

(2) Most models require a non-degeneracy of lower trajectories. Degeneracy can be accommodated only by $RRPL2$;

(3) Simple pole pomerons fail to reproduce the real part of the cross section, and all models have problems in reproducing some of the ρ data;

(4) Cosmic ray data are well reproduced by the best parametrisation, with no need of re-analysis of the published data;

Table 1: Best models for total cross sections

	ACCURRS	ACCURRS_{20}	AURS_{20}
1	$RRL2_{qc}$	$R_{qc}R_cL2_{qc}$	$R_{qc}R_cL2_{qc}$
2	$(RR)_dPL2_u$	$(RR_c)_dPL2_u$	$(RR_c)_dP_{qc}L2_u$
3	$(RR_c)_dPL2_u$	$(RR)_dPL2_u$	$R_{qc}R_cL_{qc}$
4	$R_{qc}R_cL2_{qc}$	$RRL2_{qc}$	$(RR_c)_dPL2_u$

(5) Although quark counting rules can be approximately implemented, it is also possible to have a universal rising term for the pomeron[3].

Furthermore, we can use all the information contained in our indicators to define the best models. Several schemes are possible:
i) the ACCURRSS scheme: we take all indicators, including the C_1 and C_2, and for each indicator we order the N models considered from rank 1 to N according to the value of the indicator. We then sum the 8 numbers obtained, and the best model is the one with highest rank overall;
ii) As the indicators are statistical measures, we can do the same as above but consider that a model is better than another (and give it one point) only if its indicator is bigger than that of the other model by *e.g.* 20%. This leads to the ACCURRSS_{20} scheme;
iii) Finally, one may argue that all the CL are acceptable, and that the number of parameters is not relevant as we have chosen functional forms, and hence in principle an infinite number of parameters. Using a statistical ranking similar to ii) leads then to the AURSS_{20} scheme.

Using these ranking schemes, we obtain the best models (out of the order of 30 variations on the terms used in the models) given in Table 1 for fits to total cross sections only, and in Table 2 for fits to all data for hadronic amplitudes. As can be seen from these table, simple-pole pomerons are never preferred, and models containing a $\log^2 s$ rise in the cross section always provide the best fits to the data. To reach a more restrictive conclusion, one needs to use the ρ data, in which case the preferred model is consistently $RRPL2_u$. The problem however is that the ρ data are poorly reproduced by all models considered in this study, hence one cannot be sure that this preference will survive future iterations of the cross assessments with new models and new data added, such as off-forward cross sections, or DIS structure functions, or data from future experiments. As is always the case in

Table 2: Best models for hadronic amplitudes

	ACCURRSS	ACCURRSS_{20}	AURSS_{20}
1	$RRPL2_u$	$RRPL2_u$	$RRPL2_u$
2	RRL_{nf}	$RR_cL2_{qc} = RRL2_{qc}$	$RRL2$
3	$(RR_c)_dPL2_u$		$R_{qc}R_cL2_{qc}$
4	$(RR)_dPL2_u$	$(RR)_dPL2_u = R_{qc}R_cL2_{qc} = RRL2$	$(RR)_dPL2_u$

these studies, it would be of utmost interest to have higher-energy data for other beams than protons and antiprotons. This would of course enable one to determine directly whether the pomeron counts quarks, or has a universal component.

Acknowledgments

COMPAS was supported in part by the Russian Foundation for Basic Research grants RFBR-98-07-90381 and RFBR-01-07-90392. K.K. is in part supported by the U.S. D.o.E. Contract DE-FG-02-91ER40688-Task A. We thank the president C.W. Kim of the Korea Institute for Advanced Study, Yonsei University and Professor J.-E. Augustin of LPNHE-University Paris 6 for their hospitality during various stages of this work, which made it possible.

References

[1] J. R. Cudell, V.V. Ezhela, P. Gauron, K. Kang, Yu.V. Kuyanov, S.B. Lugovsky, B. Nicolescu and N.P. Tkachenko, hep-ph/0107219.

[2] J. R. Cudell, V. Ezhela, K. Kang, S. Lugovsky and N. Tkachenko, Phys. Rev. **D 61** (2000) 034019 [hep-ph/9908218]. Erratum-ibid. **D63** 059901 (2001); J. R. Cudell, K. Kang and S. K. Kim, Phys. Lett. B **395** (1997) 311 [hep-ph/9601336].

[3] P. Gauron and B. Nicolescu, *Phys. Lett.* **B486** (2000) 71 [hep-ph/0004066]455.