A comparative analysis of statistical hadron production

F. Becattini
Dipartimento di Fisica, Università di Firenze, and INFN Sezione di Firenze, Italy

P. Castorina
Dipartimento di Fisica, Università di Catania, and INFN Sezione di Catania, Italy

A. Milov
Department of Particle Physics, Weizmann Institute of Science Rehovot 76100, Israel

H. Satz
Fakultät für Physik, Universität Bielefeld, Germany

We perform a systematic comparison of the statistical model parametrization of hadron abundances observed in high energy $pp$, $AA$ and $e^+e^-$ collisions. The basic aim of the study is to test if the quality of the description depends on the nature of the collision process. In particular, we want to see if nuclear collisions, with multiple initial interactions, lead to “more thermal” average multiplicities than elementary $pp$ collisions or $e^+e^-$ annihilation. Such a comparison is meaningful only if it is based on data for the same or similar hadronic species and if the analyzed data has quantitatively similar errors. When these requirements are maintained, the quality of the statistical model description is found to be the same for the different initial collision configurations.

I. INTRODUCTION

One of the most striking observations in high energy multihadron production is that both species abundances and transverse momentum spectra (provided effects of collective flow and gluon radiation are removed) follow the thermal pattern of an ideal hadron-resonance gas, with a universal temperature $T \approx 160 – 170$ MeV \cite{1}. This behavior was initially attributed to a temperature limit arising from an exponentially growing resonance spectrum \cite{2, 3} and is today generally taken as a reflection of the quark-hadron transition temperature in QCD \cite{4}.

Thermal multihadron production has been investigated experimentally in a variety of collision processes, from $e^+e^-$ annihilation and $pp$-$p\bar{p}$ collisions to high energy nucleus-nucleus interactions. For sufficiently high energies, these studies all led to the same universal hadronic resonance gas temperature, even though there were other distinguishing features. In particular, it was observed that in elementary collisions, strangeness production suffered an overall suppression, quite likely due to the heavier mass of the strange quark. In nuclear interaction, this suppression is less or perhaps completely absent\cite{1}.

The origin of the success of a thermal picture for such a variety of different collision configurations has been an enigma for a long time, extensively discussed in the literature \cite{6}. In particular, given a large number of possible multi-hadron channels, why does nature always choose to maximize entropy at a universal temperature within a finite region? In heavy ion collisions, it is conceivable that this could have something to do with the relatively large volume, which makes the system confined for a long enough time to allow sufficient inelastic scattering to reach equilibration. However, this cannot account for the observation of thermal behaviour in elementary collisions, where such a mechanism cannot play any role; also, a hadronic-reshattering thermalization is hardly reconciled with the observation of a centrality-independent chemical freeze-out temperature in relativistic heavy-ion collisions \cite{7}. These facts are a strong indication that the thermal behaviour is a feature of hadronization itself. One explanation proposed for a universal spontaneous thermal hadron emission is that it arises through quantum tunnelling (Unruh radiation) at the color event horizon \cite{8}; the Schwinger mechanism \cite{9} is a special case of such Unruh radiation.

\footnote{It has recently been noted that whatever strangeness suppression remains in heavy ion collisions can be accounted for by residual ("corona") single nucleon-nucleon interactions \cite{5}.}
Among other ideas put forward to explain the mechanism of the apparent thermalization, it is worth mentioning quantum chaos and Berry’s conjecture \[1, 11\], still at a speculative stage.

In some recent studies \[12–14\], the extent of agreement of data with a thermal description was discussed for different processes. In a “thermalization” framework where inelastic collisions play a major role, it is suggestive that elementary processes should not lead to as good an agreement to the same, though approximated, statistical model formula, as it is found for nuclear collisions, with a sort of hierarchy from $e^+e^-$ to $AA$. Such a view would be supported if it was found that the agreement, as specified by e.g. $\chi^2$/dof, is simply much better for $AA$ data than it is for that from $e^+e^-$ annihilation. This has in fact been claimed \[13\]; however, if such comparisons are to be conclusive, several conditions have to be met.

- First of all, any comparison should be based on essentially the same set of species. Ideally, these should be narrow and hence long-lived resonance states, in order to avoid the difficulties encountered in determining the rates of broad short-lived states, concerning background separation, feed-down and branching ratios. This requirement leaves in general some 10 to 15 states and so still allows a significant comparative analysis.

- Next, if the comparison is to be based on the $\chi^2$/dof for the fits to be considered, the corresponding data sets should have similar experimental errors. If the estimated theoretical values are approximations, a more accurate data set may imply larger deviations between model and data in units of errors, so that a comparison of the $\chi^2$/dof of two different sets makes sense only if their errors are comparable.

- Finally, the basic idea of statistical hadronization can be implemented for specific process in different ways. Without decisive further information, a comparison of the quality of the fits remains the only tool to judge which scheme is closest to reality.

We have here emphasized the use of $\chi^2$/dof as a tool to compare different experimental configurations as well as different model implementations. The reason for this is, as we shall discuss in more detail in section III, that the absolute value of $\chi^2$/dof has to be interpreted with much care. Since the theoretical formulae employed in any statistical model fit are only approximations of the full dynamics, deviations must appear, as has been mentioned, once the measurements become sufficiently precise, and hence $\chi^2$/dof must then become large. This point was already made about 10 years ago \[15\], when discussing the effect of local fluctuations of thermodynamical parameters.

With these caveats in mind, we have selected three extensive data sets for our comparative analysis. In Section II, following a short summary defining the details of the underlying statistical model framework, we compare recent hadroproduction data from $pp$ collisions, taken by the STAR experiment at RHIC for $\sqrt{s} = 200$ GeV \[16\], to the corresponding results from $Au-Au$ collisions \[17\] at the same energy and measured by the experimental group. One very remarkable result of this analysis is, as we shall see, that the fit to the $pp$ data is as good as it can possibly be, with a $\chi^2$/dof $\approx 1$: the corresponding $Au-Au$ analysis leads to a less optimal fit. To elucidate the meaning of this result, we discuss in Section III more generally the relevant features of comparing the statistical hadronization model to data and testing the fits obtained. In Section IV, we then consider $e^+e^-$ data from LEP at 91.25 GeV \[18\] and compare the fits for this to the $Au-Au$ results. In Section V, we consider more generally tests of the different implementations of the statistical model; in this context, we also consider possible origins of recent apparently contradictory conclusions \[12, 13\] on the thermal description of $e^+e^-$ data.

## II. A COMPARISON BETWEEN $pp$ AND $AA$ COLLISIONS

In this Section we perform a statistical analysis of hadroproduction data at $\sqrt{s} = 200$ GeV at RHIC with $pp$ and $Au-Au$ as initial collision configurations. The data in both cases are centre-of-mass midrapidity densities from the same experiment \[16, 17\]. For $pp$ collisions, there are 18 species of measured secondaries \[16\], including several short-lived strange meson and hyperon resonant states; for $Au-Au$ interactions, we use a set of 12 rapidity densities of long-lived states at midrapidities \[17\], already studied in ref. \[19\], with updated experimental errors. The observed abundances are listed in table \[11\].
| Particle | Measured $dN/dy$ (E) | Relative error | Model $dN/dy$ (M) | Residual $(M - E)/E$ (%) |
|----------|----------------------|----------------|------------------|------------------------|
| $\pi^+$  | 1.44 ± 0.11          | 0.076          | 1.403            | -0.34                  | -2.62                  |
| $\pi^-$  | 1.42 ± 0.11          | 0.077          | 1.384            | -0.33                  | -2.59                  |
| $K^+$    | 0.150 ± 0.013        | 0.087          | 0.1522           | 0.17                   | 1.48                   |
| $K^-$    | 0.145 ± 0.013        | 0.090          | 0.1460           | 0.076                  | 0.68                   |
| $p$      | 0.138 ± 0.012        | 0.087          | 0.1491           | 0.92                   | 7.42                   |
| $\bar{p}$| 0.113 ± 0.010        | 0.088          | 0.1120           | 0.66                   | 5.56                   |
| $\phi$   | 0.0180 ± 0.0029      | 0.16           | 0.01130          | -2.31                  | -59.3                  |
| $\Lambda$| 0.0436 ± 0.0041      | 0.094          | 0.04348          | -0.030                 | -0.28                  |
| $\Lambda$| 0.0398 ± 0.0038      | 0.095          | 0.03686          | -0.77                  | -7.96                  |
| $\Xi^-$  | 0.0026 ± 0.00092     | 0.35           | 0.003070         | 0.51                   | 15.3                   |
| $\Xi^+$  | 0.0029 ± 0.00104     | 0.36           | 0.002728         | -0.17                  | -6.29                  |
| $\Omega + \bar{\Omega}$| 0.00034 ± 0.00019 | 0.56           | 0.0005712        | 1.22                   | 40.5                   |

Au-Au collisions at $\sqrt{s_{NN}} = 200$ GeV:

| Particle | Measured $dN/dy$ (E) | Relative error | Model $dN/dy$ (M) | Residual $(M - E)/E$ (%) |
|----------|----------------------|----------------|------------------|------------------------|
| $\pi^+$  | 322 ± 25             | 0.078          | 330.0            | 0.32                   | 2.41                   |
| $\pi^-$  | 327 ± 25             | 0.077          | 331.9            | 0.19                   | 1.46                   |
| $K^+$    | 51.3 ± 6.5            | 0.13           | 57.65            | 0.98                   | 11.0                   |
| $K^-$    | 49.5 ± 6.2            | 0.13           | 54.44            | 0.80                   | 9.07                   |
| $p$      | 34.7 ± 4.4            | 0.13           | 42.23            | 1.71                   | 17.8                   |
| $\bar{p}$| 26.7 ± 3.4            | 0.13           | 31.24            | 1.34                   | 14.5                   |
| $\Lambda$| 16.7 ± 1.12           | 0.067          | 14.44            | -2.02                  | -15.7                  |
| $\bar{\Lambda}$| 12.7 ± 0.92         | 0.072          | 11.10            | -1.74                  | -14.4                  |
| $\phi$   | 7.95 ± 0.74           | 0.093          | 6.697            | -1.09                  | -18.7                  |
| $\Xi^-$  | 1.83 ± 0.206          | 0.092          | 2.024            | -0.73                  | -7.20                  |
| $\Xi^+$  | 2.17 ± 0.20           | 0.11           | 1.676            | -0.75                  | -9.16                  |
| $\Omega + \bar{\Omega}$| 0.53 ± 0.057       | 0.11           | 0.6529           | 2.16                   | 18.8                   |

TABLE I: Measured and fitted mid-rapidity densities in pp and Au-Au collisions at 200 GeV; data from STAR experiment. For pp collisions, none of the quoted experimental numbers are corrected for weak decay feed-down [16], while for Au-Au collisions all multiplicities are feed-down corrected, except protons and antiprotons [17]. Our model calculations were carried out accordingly.

For elementary collisions, the abundance $\langle n_j \rangle$ of hadron species $j$ is in the statistical hadronization model given by (for a detailed description, see ref. [12])

$$\langle n_j \rangle_{\text{primary}} = \frac{VT(2J_j + 1)}{2\pi^2} \sum_{n=1}^{\infty} \gamma_n \Lambda_n (\mp 1)^{n+1} \frac{m_j^2}{n} K_2 \left( \frac{m_j}{T} \right) \frac{Z(Q - nq_j)}{Z(Q)} ,$$

where the temperature $T$, the strangeness suppression $\gamma_n$ and the normalization volume $V$ are taken as free parameters; $Q = (Q, B, S, \ldots)$ is the array of conserved charges and $q_j$ the corresponding array for the $j$th hadron species. The “chemical” factors $Z(Q - nq_j)/Z(Q)$ are ratios of partition functions and replace the more familiar fugacities when the exact conservation of the initial charges is to be taken into account, a typical feature of small systems also known as canonical suppression. To the primary production for each species $j$ one then adds the decay products of heavier states, using the experimentally
Formulae (1) and (2) apply in principle to full phase space multiplicities, since possible charge-momentum correlations are integrated out. In order to apply it nevertheless in a comparison of midrapidity data from pp and Au-Au collisions at the same energy, we assume that particle ratios at midrapidity are essentially the same as particle ratios in full phase space. While such an assumption is certainly not tenable at low collision energy, it might improve its validity in high energy collisions with large rapidity coverage. We thus assume that the primary rapidity density of each species in collision energy, it might improve its validity in high energy collisions with large rapidity coverage. We thus assume that the primary rapidity density of each species in pp collisions is given by

\[
\langle n_j \rangle = \langle n_j \rangle_{\text{primary}} + \sum \langle n_k \rangle B R(k \to j).
\]

(2)

where \( A \) is a common normalization factor taking into account the ratio of production in the mid-rapidity interval to the overall rate. Note that \( A \) cannot be absorbed into the overall volume \( V \), since \( V \) still appears as a key parameter in the chemical factors \( Z(Q - nq_j)/Z(Q) \). Also, it is assumed that subsequent decays do not alter the midrapidity densities, i.e. that also formula (2) holds.

In this form, the statistical model for \( pp \) collisions leads to a four-parameter fit: besides \( T \), \( \gamma_s \), and \( V \), there is the rapidity-cut related normalization factor \( A \). For heavy ion collisions, we have exactly the same number of parameters; because of the large volume, formula (3) here becomes

\[
\langle n_j \rangle_{\text{primary}}^{\gamma_s = 0} = \frac{AVT(2J + 1)}{2\pi^2} \sum_{n=1}^{\infty} \gamma_s^{N_n(\pm 1)^{n+1}} \frac{m_j^2}{n} \ \text{K}_2 \left( \frac{nm_j}{T} \right) \ \frac{Z(Q - nq_j)}{Z(Q)},
\]

(3)

with the chemical factors now replaced by the fugacities. In this case, the factor \( A \) can be absorbed in \( V \) and the free parameters are \( T \), \( \gamma_s \), \( \mu_B \) and \( AV \).

The results of the fits to the midrapidity densities are shown in table II for both \( pp \) and Au-Au and fig. III the fit parameter values and the corresponding \( \chi^2/\text{dof} \) are listed in table III. The first block of this table, labelled “full \( pp \) fit”, gives the parameters for the full analysis of all \( pp \) species. In the middle block of table III, we then compare the \( pp \) and Au-Au fits using the same data sample.

The most striking feature seems to us the high quality of the full \( pp \) fit. The resulting \( \chi^2/\text{dof} \sim 1 \) is the best value to be hoped for; it is as good as any thermal fit ever made for any high energy collision configuration.

Next, it is worth noting that the extracted temperature values are almost identical for \( pp \) and Au-Au collisions; the \( \gamma_s \) value for the \( pp \) data agree with previous \( pp \) analyses [21, 22]. The error on \( VT^3 \) here is large because this parameter is essentially determined by the chemical factors in eq. (4).

In spite of the extremely similar values obtained for the hadronization temperature, we thus find that the \( \chi^2/\text{dof} \) of the \( pp \) fits is a factor two better than that for the Au-Au analysis. A naive conclusion of this comparison could then be that the statistical model leads to a better agreement with \( pp \) collisions than for \( Au – Au \). Can we conclude that \( pp \) collisions provide a "more thermal" configuration than heavy ion collisions? Before answering this question, a general discussion about the statistical model formulae and the meaning of statistical tests is now appropriate.

III. THEORETICAL MODELS AND \( \chi^2 \) TESTS

Here we want to discuss more in detail two points that have been mentioned in the Introduction, namely the meaning of a \( \chi^2/\text{dof} \) fit given only an approximate theoretical description, and how one can compare two fits to such an incomplete input.

---

2 A recent analysis [20] of \( p–p \) data at a considerably lower energy, \( \sqrt{s} = 17 \text{ GeV} \), also finds very good agreement with a different implementation of the statistical model.
|                              | pp $\sqrt{s} = 200$ GeV | Au-Au $\sqrt{s_{NN}} = 200$ GeV |
|------------------------------|--------------------------|-----------------------------|
| Overall fit                 |                          |                             |
| $T$(MeV)                    | 170.1± 3.5               | 168.5 ± 4.0                 |
| Normalization               | 0.027± 0.011             | 13.6 ± 0.58                 |
| $VT^3$                      | 135 ± 60                 |                             |
| $\gamma_S$                  | 0.569± 0.031             | 0.932 ± 0.040               |
| $\mu_B/T$                   | 0.173 ± 0.052            |                             |
| $\chi^2$/dof               | 15.6/14                  | 22.2/8                      |
| Fit with standard sample    |                          |                             |
| $T$(MeV)                    | 169.8± 4.2               | 168.5 ± 4.0                 |
| Normalization               | 0.028± 0.012             | 13.6 ± 0.58                 |
| $VT^3$                      | 131 ± 60                 |                             |
| $\gamma_S$                  | 0.600± 0.033             | 0.932 ± 0.040               |
| $\mu_B/T$                   | 0.173 ± 0.052            |                             |
| $\chi^2$/dof               | 15.0/8                   | 22.2/8                      |
| Fit with standard sample and same relative errors |                          |                             |
| $T$(MeV)                    | 170.9± 5.2               | 169.6 ± 4.8                 |
| Normalization               | 0.031± 0.015             | 12.6 ± 0.73                 |
| $VT^3$                      | 118 ± 63                 |                             |
| $\gamma_S$                  | 0.595± 0.038             | 0.999 ± 0.069               |
| $\mu_B/T$                   | 0.163 ± 0.065            |                             |
| $\chi^2$/dof               | 13.4/8                   | 9.6/8                       |

TABLE II: Comparison between the statistical model fits in $pp$ and Au-Au collisions at $\sqrt{s} = 200$ GeV. The parameter referred to as “Normalization” is $A$ in eq. (3) for $pp$ collisions and $AVT^3 \exp(-0.7GeV/T)$ in eq. (4) for Au-Au collisions.

In general, the test of a physical model is usually based on a fit of data and a statistical test of this fit. Thus a $\chi^2$ test tells us how likely it would be to obtain a value larger than the minimum $\chi^2$ of our fit, provided that the hypothesis, i.e., the model, is correct. Most often, however, the theoretical formulae that we take as hypotheses are only an approximate representation of the underlying model or theory. In other words, they can be expected to reproduce the data only up to some reasonably small deviation. There are many instances of this situation; a simple example is the lowest-order perturbative expansion of a differential cross section in high-energy collisions. If measurements are more accurate than the estimated deviation, the $\chi^2$ statistical test will obviously fail, indicating that corrections to the lowest-order theoretical formula are necessary. Sometimes the theoretical description of the process is fully under control and corrections are relatively easy to calculate (as for electroweak processes in $e^+e^-$ collisions), sometimes they are not. This latter is the case in this work, where we test the statistical hadronization model.

The basic premise of the statistical model is that high energy collisions lead to the formation of multiple clusters, emitted sequentially in rapidity and decaying into hadrons according to their relative phase space weights. The formulae (1) and (4) are specific implementations of this idea, based on further additional assumptions besides the basic postulate of the model. In elementary collisions, it is assumed that the probability of distributing the conserved charges among the actually produced clusters has a special form: for instance, a statistical distribution of charges among the clusters, leading to the equivalence with one global cluster (this is assumed in our work here). But other charge distribution schemes are obviously conceivable, see ref. \cite{23}, and we shall return to this aspect in Section V. If it is therefore clear that if experimental multiplicities were known to a very good accuracy, discrepancies with the predictions or fitted values of formula (1) could show up, even if the basic idea of purely statistical decays of clusters/fireballs remains true. One would then have to correct (1) for effects of a non-statistical distribution of charges among the clusters, etc. Unfortunately, the definition of such corrections requires a more complete picture of the production process than we presently have. Of course, this does not mean that approximate models cannot be disproved. It only means that, as long as corrections to the leading-
order formula are not available, we have to be content with a statement like "the statistical hadronization model in its simplest approximation reproduces the data up to 10\%", and we should take deviations from formula (1) with great care.

Above and beyond these model-dependent details, we should also stress that any analytical multiplicity formulae, such as (1) and (4) or variations thereof, provide by construction only an approximate description. The use of these formulae will necessarily lead to deviations for sufficiently precise measurements. To illustrate this situation, which is ubiquitous in physics, we recall a familiar example. The spectrum of the sunlight measured at the top of the atmosphere is usually fitted to a black body formula, yielding a temperature of about 5780\(^\circ\)K. The formula is accurate to about 10\%, but the fit quality is extremely bad, with a huge \(\chi^2\) (see e.g. ref. [24]). The reason for such discrepancy is that on a microscopic level, stars are not perfect radiators; different effects, such as the absorption of light by atoms and/or ions blocking part the outward radiation path, surface non-uniformities, etc., all cause deviations from the simple black-body spectrum, and these are revealed when the accuracy of the photometric measurement is better than about 10\%. The effects responsible for these deviations are difficult to embody in an analytical formula and can only be studied numerically. Still, the failure of the lowest-order Planck formula fit in passing a rigorous statistical \(\chi^2\) test has not led anyone to the conclusion that the surface of the sun is not a thermal system.

In addition to such theoretical caveats, one must also bear in mind the role of experimental complications. In fact, most of the measurements which are used to perform a fit include hidden correlations which are difficult to assess (e.g., multiplicities of particles measured with the same detector). Therefore the usual assumption of independent errors entering in the \(\chi^2\), which we have retained here throughout, is only an approximation, and the actual absolute value of the best \(\chi^2/dof\) could well be different.

Hence, for theoretical as well as for experimental reasons, the use of absolute \(\chi^2/dof\) values to judge the quality of statistical model descriptions, as suggested in ref. [14], appears to us as not really permissible.

But even with these caveats in mind, a comparative assessment of fit results in different collision configurations is still possible. To be specific, if the fit quality to a given formula was better in AA collisions than in \(e^+e^-\) or \(pp\) interactions under the same conditions, this could of course mean that the deviations observed in \(e^+e^-\) and \(pp\) stem from a genuine failure of the model for this case, rather than from the approximate nature of the employed formula. However, for such a comparison to make any sense, it must

---

**FIG. 1:** Above: fitted vs measured midrapidity densities in \(pp\) collisions at \(\sqrt{s} = 200\) GeV. Below: residual distributions.
fulfill the essential prerequisite noted in the Introduction. Thus, if the comparison is to be based on a \( \chi^2 \) test for the fits to be considered, the corresponding data sets must have comparable average experimental errors. It is quite clear that if data sets with largely different accuracy are used, a comparison of \( \chi^2/\text{dof} \) values could be misleading. To illustrate: if the data set \( A \) (example: hadronic multiplicities in \( pp \)) with an average accuracy of 10% yields a fit with \( \chi^2/\text{dof} = 1 \) and the set \( B \) (example: hadronic rapidity densities in heavy ion collisions) with an average accuracy of 1% yields a fit with \( \chi^2/\text{dof} = 3 \), a blind comparison of \( \chi^2 \)'s would lead us to the erroneous conclusion that the model can be used to describe \( A \), but not \( B \).

We now return to our comparison of Sect. 2 and the hypothetical conclusion that the statistical model works better in \( pp \) collisions than in Au-Au collisions. A comparison of \( \chi^2 \) values requires that measurements have the same experimental accuracy, but this is not the case here: as can be deduced from table I, the relative experimental errors for the same species sets differ, with an average value of about 18% for the \( pp \) data but only 10% for Au-Au.

A possible way of comparing the quality of the two fits would be to look at the average relative deviation between theoretical and experimental values. For each particle species this is shown in the last column of table I. Indeed, the average deviation of the 12 common species are very similar: 12.5% in \( pp \) and 11.7% in Au-Au. This result is a strong indication that the statistical model indeed yields the same quality of agreement in the examined cases. To illustrate this more precisely, we have made new fits with the experimental errors rescaled such that the relative errors of the measurements are exactly equal and determined by the largest error for each species. For instance, the \( K^+ \) multiplicity in \( pp \) collisions is now assigned an error equal to its relative error in heavy ion collisions times its multiplicity in \( pp \) collisions, that is \( 0.13 \times 0.150 = 0.0195 \), giving it the same relative error as the corresponding measurement in Au-Au collisions. Conversely, for the \( \Omega^+ + \bar{\Omega} \) in Au-Au collisions, the measurement is far more accurate than that in \( pp \) collisions; hence here a larger error is assigned to the heavy ion value in the same manner. This procedure artificially enhances the experimental errors in both samples and can thus be used only for illustrative purposes, neither to extract the best estimate of the model parameters, nor to make a proper statistical test.

The results of these fits are shown in table I (third block), and they show that the \( \chi^2/\text{dof} \) in Au-Au is now slightly better than in \( pp \) collisions at the same energy. This reinforces our first assessment that the simple statistical model formulae (3) and (4) agree with the data up to \( \approx 10\% \) both in \( pp \) and Au-Au collisions and that none of the examined systems can be claimed to be “more thermal” in this respect.

IV. A COMPARISON BETWEEN \( e^+e^- \) AND \( AA \) COLLISIONS

Here we focus our attention on the largest and the most accurate \( e^+e^- \) sample, data from LEP at \( \sqrt{s} = 91.25 \text{ GeV} \); this we compare again to the heavy ion sample from RHIC at \( \sqrt{s_{NN}} = 200 \text{ GeV} \), and we again choose as common as possible a set of long-lived particles. Since in \( e^+e^- \) collisions, the multiplicities of particle and antiparticle are obviously equal, we now find only 7 common species (\( \pi^\pm \), \( K^\pm \), \( \Lambda \), \( p \), \( \phi \), \( \Xi^\pm \), \( \Omega \)). Therefore we have retained all 12 RHIC measurements (\( \pi^\pm \), \( K^\pm \), \( \Lambda \), \( p \), \( \bar{p} \), \( \phi \), \( \Xi^\pm \), \( \Omega + \bar{\Omega} \)) and added to the \( e^+e^- \) data sample \( \pi^0 \), \( K^0_{S,L} \), and the three long-lived \( \Sigma \) states.

The experimental values used for this comparison are a weighted average of the full phase space multiplicities measured by the four LEP experiments; these were also used in our previous analysis\(^3\). The resulting fit is to be compared to the Au-Au mid-rapidity densities measured by STAR\(^3\) and already used above. For comparison purposes, they are shown again in table I together with the corresponding \( e^+e^- \) fit values.

It should be noted that in the Au-Au collisions at \( \sqrt{s_{NN}} = 200 \text{ GeV} \) the experimental relative error is, in most cases, larger than for \( e^+e^- \) collisions at LEP energy; the average error in the \( e^+e^- \) data is 5.7%, to be compared to 10% for the heavy ion data. This suggests that the fit will result in a larger

---

\(^3\) We note that these weighted averages differ slightly from those compiled by the Particle Data Group\(^\text{[22]}\). In particular, in most cases our errors are slightly smaller than those quoted in ref.\(^\text{[22]}\) which makes the comparison even more conservative with regard to our final conclusions.
TABLE III: Measured and fitted $4\pi$ multiplicities in $e^+e^-$ collisions at $\sqrt{s} = 91.25$ GeV and mid-rapidity abundances in Au-Au collisions at 200 GeV. For $e^+e^-$ collisions, all quoted experimental numbers (for references see ref. [12]) include weak decays feed-down, while for Au-Au collisions all multiplicities are feed-down corrected but protons and antiprotons [17]. Our model calculations were carried out accordingly.

| Particle | Measured $dN/dy$ (E) | Relative error | Model $dN/dy$ (M) | Residual | $(M - E)/E$ (%) |
|----------|----------------------|----------------|-------------------|----------|-----------------|
| $\pi^0$  | 9.61 ± 0.29          | 0.030          | 9.865             | 0.89     | 2.6             |
| $\pi^+$  | 8.50 ± 0.10          | 0.012          | 8.460             | -0.37    | -2.44           |
| $K^+$    | 1.47 ± 0.026         | 0.023          | 1.040             | -1.79    | -4.3            |
| $K^{0}_S$| 1.2376 ± 0.0006      | 0.0093         | 1.040             | 0.25     | 0.24            |
| $p$      | 0.519 ± 0.018        | 0.035          | 0.5727            | 2.98     | 9.4             |
| $\phi$   | 0.0977 ± 0.0058      | 0.059          | 0.1179            | 3.47     | 17.1            |
| $\Lambda$| 0.1943 ± 0.0038      | 0.020          | 0.1867            | -1.98    | -4.0            |
| $\Sigma^+$| 0.0535 ± 0.0052     | 0.097          | 0.04331           | -1.96    | -23.6           |
| $\Sigma^0$| 0.0389 ± 0.0041     | 0.11           | 0.04392           | 1.22     | 11.4            |
| $\Sigma^-$| 0.0410 ± 0.0037     | 0.090          | 0.03949           | -0.40    | -3.7            |
| $\Xi^-$  | 0.01319 ± 0.00050    | 0.038          | 0.01285           | -0.65    | -2.7            |
| $\Omega$ | 0.00062 ± 0.00010    | 0.16           | 0.0009264         | 2.55     | 33.0            |

| Au-Au collisions at 200 GeV |
|-------------------------|
| $\pi^+$  | 322 ± 25          | 0.078          | 330.0             | 0.32     | 2.41            |
| $\pi^-$  | 327 ± 25          | 0.077          | 331.9             | 0.19     | 1.46            |
| $K^+$    | 51.3 ± 6.5        | 0.13           | 57.65             | 0.98     | 11.0            |
| $K^-$    | 49.5 ± 6.2        | 0.13           | 54.44             | 0.80     | 9.07            |
| $p$      | 34.7 ± 4.4        | 0.13           | 42.23             | 1.71     | 17.8            |
| $\bar{p}$| 26.7 ± 3.4        | 0.13           | 31.24             | 1.34     | 14.5            |
| $\Lambda$| 16.7 ± 1.12       | 0.067          | 14.44             | -2.02    | -15.7           |
| $\bar{\Lambda}$| 12.7 ± 0.92      | 0.072          | 11.10             | -1.74    | -14.4           |
| $\phi$   | 7.95 ± 0.74       | 0.093          | 6.697             | -1.69    | -18.7           |
| $\Xi^-$  | 1.83 ± 0.206      | 0.092          | 2.024             | -0.73    | -7.20           |
| $\Xi^+$  | 2.17 ± 0.20       | 0.11           | 1.676             | -0.75    | -9.16           |
| $\Omega + \bar{\Omega}$| 0.53 ± 0.057 | 0.11           | 0.6529            | 2.16     | 18.8            |

$\chi^2$/dof value for $e^+e^-$ than for Au-Au, and we see in table IV that this is indeed the case$^4$. Therefore, as emphasized above, a direct comparison of the $\chi^2$/dof values to assess the relative quality of the two fits would again be misleading.

As in the comparison between pp and Au-Au data, the average relative deviations between theoretical and experimental values are close: 9.4% for $e^+e^-$ annihilation at 91.25 GeV and 11.7% for Au-Au collisions at 200 GeV. This suggests again that the lowest order statistical model formulation yields the same quality of agreement in the two examined cases. To reinforce this, for illustrative purpose, we have also here made new fits with rescaled experimental errors, in order to make the relative errors of measurements equal in the two species for each species, in much the same way as for the comparison between pp and Au-Au collisions described in Sect. III. For the unmatched particles, we have taken an error correspondence $\pi^0 \rightarrow \pi^-$, $K^0_S \rightarrow K^-$, $\Sigma^{+} \rightarrow p$, $\Sigma^0 \rightarrow \Lambda$, $\Sigma^- \rightarrow \Xi^-$, where the first particle belongs to the $e^+e^-$ sample and the second to heavy ion sample. In fact, the only particles which have a larger error in $e^+e^-$ than in Au-Au collisions are the $\Omega$ and $\Sigma^0$, whose correspondent in Au-Au was taken as

---

$^4$ The $\chi^2$/dof value obtained here is slightly larger than that found in [12]; this is due to the fact, unlike in ref. [12], we have not included in the $\chi^2$ the contribution to errors owing to the uncertainties on masses, widths and branching ratios of resonances. This choice is motivated by the need of making the comparison between different collision systems as clean as possible.
|                | \(e^+e^-\) \(\sqrt{s} = 91.25\) GeV | Au-Au \(\sqrt{s}_{NN} = 200\) GeV |
|----------------|---------------------------------|----------------------------------|
| Fit with the standard samples | \(164.7\pm 0.9\) (1.9) | \(168.5 \pm 4.0\) |
| Normalization | \(23.2\pm 0.57\) (1.2) | \(13.6 \pm 0.58\) |
| \(\gamma_s\) | \(0.656\pm 0.0096\) (0.021) | \(0.932 \pm 0.040\) |
| \(\mu_B/T\) | \(0.173 \pm 0.052\) | \(0.173 \pm 0.052\) |
| \(\chi^2/\text{dof}\) | 41.5/9 | 22.2/8 |

**Fit with the standard samples and same relative errors**

|                | \(168.8\pm 5.2\) | \(167.8 \pm 4.1\) |
|----------------|-----------------|-----------------|
| Normalization | \(21.3\pm 3.4\) | \(13.15 \pm 0.61\) |
| \(\gamma_s\) | \(0.599\pm 0.029\) | \(0.968 \pm 0.044\) |
| \(\mu_B/T\) | \(0.200 \pm 0.057\) | \(0.200 \pm 0.057\) |
| \(\chi^2/\text{dof}\) | 11.0/9 | 16.8/8 |

**TABLE IV:** Comparison between fit results in \(e^+e^-\) collisions at LEP and Au-Au collisions at RHIC for a fit to a sample of 12 long-lived hadronic species. The parameter referred to as “Normalization” is \(V T^3\) for \(e^+e^-\) collisions and \(AV T^3 \exp(-0.7\text{GeV}/T)\) (see eq. (4)) for Au-Au collisions. The errors within brackets are the fit errors rescaled by \(\sqrt{\chi^2/\text{dof}}\) (see [12]).

The resulting fit parameters are included in table IV. It is seen that the final \(\chi^2/\text{dof}\) is now lower in \(e^+e^-\) collisions than in Au-Au collisions, although no strong statement can be made on this basis, as emphasized at the end of previous section. This exercise only demonstrates that, under equal error conditions, the agreement of the data with the statistical model predictions in the form of eq. (4) is approximately the same in \(e^+e^-\) and Au-Au collisions. Finally, we note that this result is found to be consistently independent of finer details, such as different particle species matching, number of particles, etc.

**V. COMPARING DIFFERENT STATISTICAL DESCRIPTIONS OF \(e^+e^-\) COLLISIONS**

An interesting question of quite general interest is to what extent a \(\chi^2\) test can be used to check if a specific model input is correct or not. In other words, we now want to compare the \(\chi^2/\text{dof}\) values obtained by fitting a specific data set to different theoretical schemes, rather than a given theoretical scheme to different data sets. This is again a well-posed question because it involves only a comparative assessment. More specifically, we can fit the data to different implementations of the statistical model, some of which are likely to be or are certainly incorrect. If the model indeed reflects the right physics, the \(\chi^2/\text{dof}\) of a fit to the data should be consistently larger for “incorrect” versions than for the most realistic one.

The \(e^+e^-\) data from LEP at 91.25 GeV contains enough species and is sufficiently precise to address this issue. Our test set will consists of 15 light-flavored, long-lived particles, having widths less than 10 MeV: \(\pi^0, \pi^\pm, \eta, \eta', K^+, K^0_N, \phi, p, \Lambda, \Sigma^+, \Sigma^0, \Sigma^-, \Xi^-\), and \(\Omega^-\). We now fit the measured abundances of these species to different implementations of the statistical model, using as basis eq. (1).

- We fit the abundances to the primary production form only, neglecting all resonance decay contributions.
- We fit the abundances to the primary production form only, neglecting all decay contributions from resonances of width greater than 10 MeV. In this case, strong decays are neglected, but the feed-down of weakly decaying heavy flavor states is included.
- We fit the abundances correctly taking into account all resonance decays, but we replace exact quantum number conservation (canonical suppression) by a grand canonical formulation. This means replacing the chemical factors \(Z(Q - q)/Z(Q)\) in formula (1) by the fugacities of (1).
The results are summarized in table V, where we have denoted the results from our global cluster formulation with “correct” inclusion of all effects as “full global implementation”. It is evident that the model indeed provides us with a $\chi^2$/dof hierarchy: the cruder the implementation, the larger the resulting $\chi^2$/dof.

| Fit condition                       | T (MeV) | $\gamma_S$ | $\chi^2$/dof |
|-------------------------------------|---------|------------|--------------|
| No resonance feed-down              | 158.7   | 0.494      | 1110/12      |
| No strong resonance feed-down       | 161.1   | 0.537      | 391/12       |
| No canonical suppression            | 144.9   | 0.690      | 131/12       |
| Full global implementation          | 164.8   | 0.654      | 44/12        |

TABLE V: Comparison of parameters for fits of 15 long-lived particles in $e^+e^-$ collisions at $\sqrt{s} = 91.2$ GeV, using different model inputs.

At this point, a specific feature of $e^+e^-$ annihilation is worth being emphasized. The annihilation process leads to the production of a pair of quarks, and about 40% of $e^+e^-$ annihilations produce a pair of primary $c$ or $b$ quarks, as predicted by the standard model. These then hadronize into heavy flavored hadrons, which in turn predominantly decay into strange hadrons. The secondary production of light-flavored particles from heavy flavored ones is a sizeable fraction of the overall particle production. It is thus crucial that the model correctly includes the (non-statistical) fractions of the different primary quark pairs, and fits neglecting these (e.g., in ref. [13] all except the data set at 91.2 GeV) cannot be considered as realistic.

Besides this aspect, however, there are others which distinguish different implementations of the statistical model concept. As already indicated in Section III, the distribution pattern of the conserved charges (baryon number, electric charge, strangeness, charm, bottom) is an issue to be decided when formulating a specific model. If a given overall quantum number (for illustration, consider the electric charge $Q$) in two-jet production is zero, the general production pattern will lead to a superposition of two produced jets each having zero charge, a pair with $Q = \pm 1$, one with $Q = \pm 2$, etc., as schematically illustrated in Fig. 2. To specify the model, we have to fix the weights $w_i$ of this superposition, and two particular schemes have been introduced for this [23].

![Fig. 2: Schematic distribution of conserved quantum numbers for two-jet hadron production](image)

- The simplest version is to enforce exact conservation of all discrete quantum numbers separately for each of the two jets formed in the annihilation, allowing no quantum number exchange between the jets; i.e., one puts $w_0 = 1$, $w_i = 0 \forall i \geq 1$. This is denoted as uncorrelated jet scheme.

- The model we have used here allows clusters with statistically distributed discrete quantum numbers, imposing exact overall conservation laws; i.e, the $w_i$ are distributed over $i$ according to the available multicluster phase space for an overall $Q = 0$. This is denoted as global cluster scheme.

There is a clear exception, however, to a transfer of quark quantum numbers: the heavy $c$ or $b$ quarks cannot be exchanged, since heavy quark production at hadronization is severely suppressed. In fact, it is an experimentally well established fact that heavy quark production originates entirely from the primary $e^+e^-$ annihilation or, as an almost negligible higher order perturbative correction, from a hard gluon.
is also a well established fact that primary heavy quarks show up in open-heavy flavored hadrons without reannihilating.

There seem to be no general physics grounds on which to exclude one or the other of these scenarios. However, color neutralization of the clusters created in $e^+e^-$ collisions requires, before hadron formation, the exchange of one or more quark-antiquark pairs between different clusters, and this in turn implies the appearance of non-vanishing integer additive charges for single clusters. Hence already the first study comparing the two schemes in fits to LEP $e^+e^-$ data at 91.2 GeV obtained for the uncorrelated jet fit $\chi^2$/dof values twice as large as for the global scheme [23]. For our present study, the comparison is shown in table VI and indicates that $\chi^2$/dof is more than a factor of two larger for the uncorrelated than for the global cluster scheme. Particularly, the $\phi$ meson is found to deviate about 8$\sigma$ from the data, which is not the case in the global cluster scheme (see table III); moreover, the temperature value becomes much larger in the uncorrelated jet scheme. The reason for such a behaviour is readily understood: in an uncorrelated jet scheme the phenomenon of canonical suppression is enhanced by the requirement of strict local conservation of quantum numbers and this drives the fit towards higher values of $T$ and/or $\gamma_S$ in order to reproduce the multiplicities; on the other hand, the $\phi$ meson is unaffected by canonical suppression and the increase of $T$ and/or $\gamma_S$ results in an overestimate of the production of this particle.

| Fit condition                  | $T$ (MeV)   | $\gamma_S$  | $\chi^2$/dof |
|-------------------------------|-------------|-------------|--------------|
| Uncorrelated jet scheme       | 196.9 ± 1.74| 0.622 ± 0.0096| 104/12       |
| Global cluster scheme         | 164.8 ± 0.93| 0.654 ± 0.0095| 44/12        |

TABLE VI: Comparison of parameters for fits of 15 long-lived particles in $e^+e^-$ collisions at $\sqrt{s} = 91.2$ GeV, using different conservation scheme for discrete quantum numbers. The errors within brackets are the fit errors rescaled by $\sqrt{\chi^2/dof}$ (see [12]).

The mentioned quantum number distribution among the produced jets is only one of the features to be specified for a concrete statistical model analysis code. In addition, different codes involve further technical input details, addressing e.g. the little-known decay of heavy resonances, the mass range of meson vs. baryon resonances to be included, etc. As a result, there exist several codes, differing in these rather technical details. Since one cannot give general physics grounds to judge one as “better” than another, the only tool we have is to compare the $\chi^2$/dof they provide for specific data sets. For the twelve species “standard” $e^+e^-$ data set at 91.2 GeV defined above in Table III, we had obtained a $\chi^2$/dof = 41.5/9 ≃ 4.6. Using the same code, we found [12] for a much larger set of 30 species, including short-lived resonances, a $\chi^2$/dof = 215/27 ≃ 8. This shows that also the choice of the data set affects the resulting $\chi^2$/dof; including broad resonances, with all the resulting experimental problems, significantly increases $\chi^2$/dof.

Using a different code, ref. [13] obtained for essentially the same large data set at 91.2 GeV a fit with $\chi^2$/dof = 499/28 ≃ 17.8, i.e., a value twice that which we had found. We can only conclude that the code used in ref. [13] must invoke physics features not in accord with the data. One such feature was already indicated: while we use global cluster scheme, ref. [13] retains only the first term of the expansion shown in Fig. 2 the uncorrelated jet scheme. To look at further details, we consider a comparison proposed in ref. [13]. Fixing the temperature $T = 158$ MeV, the strangeness suppression $\gamma_S = 0.8$ and the volume $V = 30$ fm$^3$, a set of species’ rates is calculated using the code of [13], using our code, and using the publicly available code THERMUS [26]. The results are shown in table VII At first sight, the yield looks fairly similar, and in particular those of our code and those of THERMUS indeed show satisfactory agreement, with rather small and fluctuating differences. A second look, however, shows that ref. [13] predicts for almost all species a larger multiplicity than the other two. This could be due to the inclusion of a larger number of heavy resonances. However, the relative differences are not uniformly distributed; the outstanding deviations are proton and $\Lambda$, which in ref. [13] lead to yields which are 1.5 - 1.7 times larger than those obtained in the other two codes. These two states are among the most accurately measured particles at LEP (3.5% error for the $p$, 2% for the $\Lambda$), and hence a sizeable difference in the predicted yields will have a large impact on the final fit. We believe that this illustrates once more our main point: it is not the absolute value of $\chi^2$/dof that matters, but rather the average relative deviation of fit to data. And here we expect that different codes shall (or should) lead to the same conclusion.
Our code & THERMUS & Code of ref. [13] \\
\pi^0 & 9.163 & 8.29 & 11.08 \\
\pi^+ & 7.763 & 7.14 & 9.27 \\
K^+ & 0.9953 & 0.945 & 1.014 \\
K_S^0 & 0.9706 & 0.920 & 0.976 \\
\eta & 1.047 & 0.890 & 1.090 \\
\rho^0 & 1.084 & 1.044 & 1.12 \\
K^{*0} & 0.3126 & 0.285 & 0.299 \\
\rho & 0.2840 & 0.334 & 0.487 \\
\phi & 0.1274 & 0.132 & 0.131 \\
\Lambda & 0.1170 & 0.120 & 0.182 \\
\Sigma^{*+} & 0.01512 & 0.0158 & 0.0197 \\
\Xi^- & 0.00862 & 0.0095 & 0.0101 \\
\Sigma^{**0} & 0.000342 & 0.00340 & 0.00384 \\
\Omega & 0.000541 & 0.000585 & 0.000625 \\

TABLE VII: Comparison between the output of our code, of THERMUS \[26\], and of the code used in ref. \[13\], for a chosen set of species and fixed parameter values.

VI. CONCLUSIONS

We have carried out statistical analyses of large multihadron production samples from \textit{pp}, Au-Au and \textit{e}^+\textit{e}^- interactions, containing the same number of species and, as far as possible, the same kind of species. We use RHIC data from STAR ($\sqrt{s} = 200$ GeV) for the first two cases, LEP data averaged over the four CERN experiments ($\sqrt{s} = 91.25$ GeV) for the last. The main results are summarized in table \[VIII\]. The hadronization temperatures are seen to agree extremely well, and the average deviation between the fit abundances and the data values (theory minus experiment/experiment) are around 10% for all three configurations.

| Collision | \textit{pp} | Au-Au | \textit{e}^+\textit{e}^- |
|-----------|--------------|-------|------------------|
| Temperature [MeV] | 169.8 ± 4.2 | 168.5 ± 4.0 | 164.7 ± 0.9 |
| Average relative deviation data vs. fit [%] | 12.5 | 11.7 | 9.4 |
| Average relative error of data [%] | 18 | 10 | 5.7 |
| $\chi^2$/dof | 15.0/8 ≃ 1.9 | 22.2/8 ≃ 2.8 | 41.5/9 ≃ 4.6 |

TABLE VIII: Summary of the fit results for a subset of 12 long-lived particles in high energy \textit{pp}, Au-Au and \textit{e}^+\textit{e}^- collisions.

On the other hand, the resulting $\chi^2$/dof values are approximately 2 for \textit{pp}, 3 for Au-Au and above 4 for \textit{e}^+\textit{e}^- interactions. We argue that this does not imply a corresponding hierarchy of agreement with a statistical description. Since the average deviations of the fitted abundances are essentially the same in the three cases, the observed differences in $\chi^2$/dof values are rather a reflection of the relative errors in the three experiments, also shown in table \[VIII\]. To test this, we have rescaled the errors in the comparison \textit{pp} vs. Au-Au and in \textit{e}^+\textit{e}^- vs. Au-Au, and in both cases, the resulting $\chi^2$/dof values then become comparable. We thus conclude that the hadroproduction abundances from high energy \textit{pp}, Au-Au and \textit{e}^+\textit{e}^- interactions agree equally well, i.e., to about 10%, with the best present statistical model parametrization.
References

[1] For an introduction and a recent summary of results see:
F. Becattini, An introduction to the Statistical Hadronization Model, arXiv:0901.3643 [hep-ph];
F. Becattini and R. Fries, arXiv:0907.1031 [nucl-th], to appear in Landolt-Börnstein vol. 1-23 A.

[2] R. Hagedorn, Nuovo Cim. Suppl. 3 (1965) 147;
Nuovo Cim. 56A (1968) 1027.

[3] K. Bardakci and S. Mandelstam, Phys. Rev. 184 (1969) 1640;
S. Puhani and G. Veneziano, Nuovo Cim. 64A (1969) 811.

[4] N. Cabibbo and G. Parisi, Phys. Lett. 59 B (1975) 67.

[5] F. Becattini and J. Manninen, J. Phys. G 35 (2008) 104013;
F. Becattini and J. Manninen, Phys. Lett. B 673 (2009) 19

[6] R. Hagedorn, Thermodynamics of Strong Interactions, CERN 71-12, 1971;
U. Heinz, Nucl. Phys. A 661 (1999) 140;
R. Stock, Phys. Lett. B 456 (1999) 277;
A. Bialas, Phys. Lett. B 466 (1999) 301;
H. Satz, Nucl. Phys. Proc. Suppl. 94 (2001) 204;
J. Hormuzdiar, S. D. H. Hsu, G. Mahlon Int. J. Mod. Phys. E 12 (2003) 649; V. Koch, Nucl. Phys. A 715(2003) 108;
L. McLerran, arXiv:hep-ph/0311028.

[7] U. Heinz and G. Kestin, PoS POD2006 (2006) 038 arXiv:nucl-th/0612105.

[8] P. Castorina, D. Kharzeev and H. Satz, Europ. Phys. J. C52 (2007) 187.

[9] J. Schwinger, Phys. Rev. 82 (1951) 664;
J. D. Bjorken, Lecture Notes in Physics (Springer) 56 (1976) 93;
A. Casher, H. Neuberger and S. Nussinov, Phys. Rev. D20 (1979) 179.

[10] See e.g.
R. Brout, R. Parentani and Ph. Spindel, Nucl. Phys. B 353 (1991) 209;
K. Srinivasan and T. Padmanabham, Phys. Rev. 60 (1999) 024007;
D. Kharzeev and K. Tuchin, Nucl. Phys. A 753 (2005) 316;
S. P. Kim, JHEP 0711:048 (2007).

[11] A. Krzywicki, arXiv:hep-ph/0204116.

[12] F. Becattini, P. Castorina, J. Manninen and H. Satz, Eur. Phys. J. C 56 (2008) 493.

[13] A. Andronic, F. Beutler, P. Braun-Munzinger, K. Redlich and J. Stachel, Phys. Lett. B 675, 312 (2009).

[14] T. Csorgo, arXiv:0901.0766 [hep-ph].

[15] U. W. Heinz, Nucl. Phys. A 661, 140 (1999).

[16] B. I. Abelev et al. [STAR Collaboration], Phys. Rev. C 79 (2009) 034909;
B. I. Abelev et al. [STAR Collaboration], Phys. Rev. C 75 (2007) 064901;
J. Adams et al. [STAR Collaboration], Phys. Rev. Lett. 92 (2004) 092301;
B. I. Abelev et al. [STAR Collaboration], Phys. Rev. C 79 (2009) 034909;
J. Adams et al. [STAR Collaboration], Phys. Lett. B 612 (2005) 181-189;
J. Adams et al. [STAR Collaboration], Phys. Rev. Lett. 98 (2007) 062301;
B. I. Abelev et al. [STAR Collaboration], Phys. Rev. Lett. 97 (2006) 132301.

[17] B. I. Abelev et al. [STAR Collaboration], Phys. Rev. C 79 (2009) 034909;
B. I. Abelev et al. [STAR Collaboration], Phys. Rev. C 79 (2009) 064903;
J. Adams et al. [STAR Collaboration], Phys. Rev. Lett. 98 (2007) 062301.

[18] For the list of references to experimental measurements and averages, see:
F. Becattini, P. Castorina, J. Manninen and H. Satz, Phys. Rev. C 56 (2008) 493
J. Manninen and F. Becattini, Phys. Rev. C 78 (2008) 054901 and references therein.

[19] I. Kraus, J. Cleymans, H. Oeschler and K. Redlich, arXiv:0910.3125

[20] F. Becattini and U. W. Heinz, Z. Phys. C 76 (1997) 269.

[21] F. Becattini and G. Passaleva, Eur. Phys. J. C 23 (2002) 551.

[22] F. Becattini, Z. Phys. C 69 (1996) 485.

[23] see e. g. W. Fussell, J. Schutt, NASA Technical note TN D-1845, 1964

[24] C. Amsler et al., Physics Letters B667, 1 (2008)

[25] S. Wheaton and J. Cleymans, J. Phys. G 31 (2005) S1069.