Considerations concerning the fluctuations of the ratios of two observables

M. Hauer[1] and S. Vogel[2]

[1]*Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany and Institut für Theoretische Physik, Goethe-Universität, Frankfurt am Main, Germany*

[2]*SUBATECH, Laboratoire de Physique Subatomique et des Technologies Associées, University of Nantes–IN2P3/CNRS–Ecole des Mines de Nantes, 4 rue Alfred Kastler, F-44072 Nantes Cedex 03, France*

We discuss several possible caveats which arise with the interpretation of measurements of fluctuations in heavy-ion collisions. We especially focus on the ratios of particle yields, which have been advocated as a possible signature of a critical point in the QCD phase diagram. We conclude that current experimental observables are not well defined and are without a proper quantitative meaning.

I. INTRODUCTION

The investigation of fluctuations and correlations, i.e., an event-by-event analysis, of high-energy heavy-ion collision data has been ongoing for several years [1–14]. Recently, this discussion was refreshed [15–21] in the view that (event-by-event) fluctuations of multiplicity, or of ratios of multiplicities, might serve as an observable to pin down the critical behavior of the collective, and potentially at some point of its evolution-thermalized, system created when two heavy ions collide. Thus, the measurement of ratio fluctuations might make it possible to fix the values of temperature and baryon chemical potential of a critical point (see, e.g., [22,23]) in the phase diagram of quantum chromodynamics (QCD). In this paper we discuss the problems and possible caveats of such an analysis.

An event-by-event analysis ultimately deals with (joint) distributions of a certain set of observables. Parameters such as mean value, variance, skewness, and kurtosis can be employed to describe the shape of an obtained distribution. Their values, as the obtained distribution itself, will depend on many experimental factors. More complicated measures are often lacking a clear interpretation. In particular, we will investigate the commonly used measure of dynamical ratio fluctuations.

It is not the aim of this paper to suggest an improved observable or method capable of locating a critical point of strongly interacting matter, or one that could potentially be indicative of new physics. The goal of this paper is to discuss caveats of the existing methods which are usually not mentioned in the literature. The paper is organized as follows. In Sec. II we formulate some rather general statements about the particularities of an event-by-event analysis. In Sec. III we then focus on a particular method and discuss some difficulties one might face with its interpretation. A discussion of specific experimental and theoretical studies can be found in Sec. IV. Some concluding remarks are summarized in Sec. V.

II. GENERAL CONSIDERATIONS

Let us start by making a rather general observation. The statistical properties of a sample of events depend on the rules chosen to select this sample from an even larger sample of events and on the degree of completeness of the information available about each member of this sample. In the context of heavy-ion collision physics these two aspects translate into centrality class construction from minimum bias data and an experiment's capabilities of particle identification and momentum measurement. We will elaborate on these aspects in the following.

A. Centrality selection

We first discuss centrality selection. Two heavy ions collide with relativistic momenta. Being extended objects, they can do so in a variety of different ways. Roughly the following rule should apply (in the average sense): The larger the interaction region, the more particles of each species are produced. However, the initial state of the collision cannot be observed directly. All that can be observed is the final state. From this one can then infer the likelihood of a certain initial state. Yet, each single possible final state observable will generally suggest a slightly different initial state. Hence, the need arises to average over centrality classes. The problem is then, within any such a centrality class will be events with rather different initial states, altering the true correlation between two observables. Results will depend on, in particular, which trigger was chosen to construct these centrality classes, or subsamples, of events.

To illustrate this point we refer to Fig. 1. Shown is a result of a Hadron String Dynamics (HSD) transport simulation, taken from [24]. The two observables shown in Fig. 1 are the number of participating (or wounded) nucleons N_P, and the multiplicity of charged final-state hadrons N_{ch}^{ref} in a midrapidity acceptance window. This analysis is somewhat hypothetical,[1] but serves to illustrate a crucial point: Depending on which one of the two observables is used to define the sample of 20–22% most central events, two rather different subsamples, indicated by the vertical and horizontal lines, are created. These two samples could possibly have rather different statistical properties.

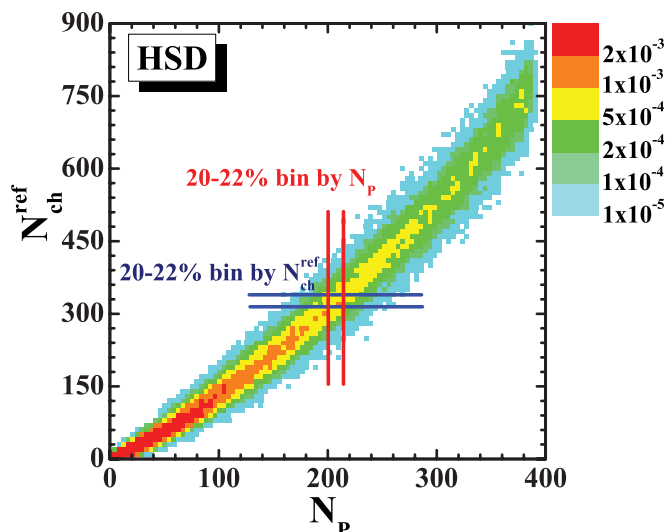[1]The number of participating nucleons is not accessible to experimental measurement.

FIG. 1. (Color online) Illustration of centrality selection being done in two different ways: once via the number of charged particles in the final states (N_{ch}^{ref}) and once via the number of wounded nucleons (N_P). This leads to two different subsamples, despite the fact that the same centrality class (20–22%) was constructed. The different colors depict the probability. The figure is taken from [24].

The two methods are hence not equivalent, and both introduce their very own bias. It deserves to be stressed that the smaller the chosen centrality class intervals (here 2%) are, the more distinct are the subsamples obtained from two different trigger conditions.

B. Acceptance in momentum space

The second aspect we want to discuss is particle acceptance. By this we mean particle identification, momentum measurement, and ultimately the geometric coverage of the interaction region by the detector. We ignore for now centrality selection and think of a perfectly prepared initial state and explore two limiting cases, the first one being the ideal detector. All final-state particles are observed, and any correlation can be measured to any degree. The opposite limit would be a very bad detector, capable of only detecting a particle every once in a while. Such a detector could surely measure the ratios of the occurrence of particles of different species (provided sufficient particle identification capabilities) and hence allows for a comparison of model results for ratios of averaged particle multiplicities to actual data [25–27]. However, it would be completely unable to inform us how the multiplicities of different particle species within one event are correlated or how this correlation would change in different segments of momentum space. Any realistic detector is in-between these limits.

Also, here Fig. 1 can serve as an illustration. Changing the acceptance window for charged final-state hadrons N_{ch}^{ref}, which was for the purpose of this analysis located around midrapidity (while assuming that N_P can still be measured with the same accuracy), will change the marginal distribution

$P(N_{ch}^{ref}) = \sum_{N_P} P(N_P, N_{ch}^{ref})$. In particular, mean value, variance, skewness, and kurtosis of the marginal distribution $P(N_{ch}^{ref})$ are affected. The distribution $P(N_{ch}^{ref})$ is often used as a reference distribution for construction of centrality classes by collider experiments. That means, depending on the value N_{ch}^{ref} of a particular event, one assigns this event to a certain centrality class.

III. FLUCTUATIONS OF THE RATIO OF TWO OBSERVABLES

The above considerations served to emphasize the importance of the way by which a certain data set was constructed and the degree of completeness of the information available about each element of ones data set. These two aspects also determine whether two data sets can really be compared to each other. We now focus on one method, the measurement of the fluctuations of the ratio of particle multiplicities, while leaving it to the reader to decide whether or not some aspects of this discussion could be shared by other methods.

A. Data distribution

First, a plain observation: Independent of a particular analysis technique, what is considered here is the ratio of two observables, each of which will have a certain value in each event. Directly (as done, e.g., in the STAR experiment [28]) or indirectly (as done, e.g., in the NA49 experiment [29]) one can then obtain the ratio of these two observables. In the following we will focus on the number of particles of a certain species in each event.

We again consider construction of centrality classes and the window of particle acceptance available. In general, one can conclude that the joint (two-dimensional) distribution $P_{data}(N_1, N_2)$ of the multiplicities of the two species "1" and "2" of particles will be determined by the choices of solutions to the two aforementioned aspects, i.e., acceptance and centrality selection. The distribution of ratios, $P_{data}(N_1/N_2)$, being constructed from the joint distribution, is then certainly no exception.

A measure for quantifying the width of the distribution of ratios $P_{data}(N_1/N_2)$, similar to the ones used by the STAR [28] and NA49 Collaborations [29], could read

$$\sigma^2 = \frac{\langle(\Delta N_1)^2\rangle}{\langle N_1\rangle^2} + \frac{\langle(\Delta N_2)^2\rangle}{\langle N_2\rangle^2} - 2\frac{\langle\Delta N_1\Delta N_2\rangle}{\langle N_1\rangle\langle N_2\rangle}. \quad (1)$$

Here $\langle(\Delta N_i)^2\rangle$ denotes the variance of observable N_i, $\langle N_i\rangle$ denotes the mean value of the observable N_i, and $\langle\Delta N_1\Delta N_2\rangle$ denotes the covariance of the observables N_1 and N_2.

The above discussion about centrality selection and particle acceptance certainly does not do justice to all the technical details which will ultimately determine the shape of the joint distribution $P_{data}(N_1, N_2)$ of the two observables N_1 and N_2. Yet, all five terms which make up the measure depend strongly on many of these aspects. Hence, there is no reason to believe that the measure σ^2, given in Eq. (1), would not depend on them.

B. Mixed-event background

The need arises then to remove the bias introduced. The mixed-event background subtraction should remove all "trivial" correlations due to acceptance and centrality selection, while leaving all "physical" correlations unchanged. Note, that different mixing procedures have been suggested [30], but since they do not change the argumentation of our analysis we do not comment on them separately. Before focusing on the mixed-event background, let us make another observation about the measure σ^2. This measure, being so general, could be applied to any sample as long as the observables N_1, N_2 are positive integers, independent of what N_1 and N_2 may represent. This implies that the event mixing procedure, once it is specified, could also be applied to a more general (or plainly different[2]) system with observables N_1 and N_2.

Let us consider a particular event mixing procedure similar to the one used by the NA49 Collaboration. First, a reference distribution $P_{\text{data}}(N)$ of the observable $N = N_1 + N_2$ is constructed:

$$P_{\text{data}}(N) = \sum_{N_1=0}^{N} P_{\text{data}}(N_1, N - N_1). \qquad (2)$$

A mixed distribution $P_{\text{mix}}(N_1, N_2)$ can be obtained by sampling the reference distribution $P_{\text{data}}(N)$ and assigning the values for N_1 and N_2 according to a binomial distribution. The mixed-event background distribution is then

$$P_{\text{mix}}(N_1, N_2) = q^{N_1}(1 - q)^{N_2}\binom{N_1 + N_2}{N_2} P_{\text{data}}(N_1 + N_2), \qquad (3)$$

where $q = \langle N_1 \rangle / (\langle N_1 \rangle + \langle N_2 \rangle)$ is the probability that a randomly drawn particle is of species 1. The distributions $P_{\text{data}}(N)$ and $P_{\text{mix}}(N)$ are then identical.

Stated differently, once the distribution $P_{\text{data}}(N_1, N_2)$ is chosen (or fixed), then the new mixed background $P_{\text{mix}}(N_1, N_2)$ is also similarly determined. This is to say, for any generic distribution $P(N_1, N_2)$ a mixed background and the measure of dynamical ratio fluctuations

$$\sigma_{\text{dyn}}^2 = \sigma_{\text{data}}^2 - \sigma_{\text{mix}}^2 \qquad (4)$$

are equally determined. This is independent of what N_1 and N_2 represent, and independent of how the original data set came about.

---

[2]An example, unrelated to physics, could come to mind: The number of kindergartens and supermarkets in every city and town could be evaluated. Groups of cities sorted according to some criterion, such as their estimated number of citizens, would form centrality classes. Our ability to accurately determine the number of kindergartens and supermarkets would correspond to particle acceptance. This might not be a particular thoughtful example, yet it shows that such an analysis is not restricted to heavy-ion physics. And one still faces the same problem of having to give interpretation to the fluctuation measures, Eqs. (1) and (4).
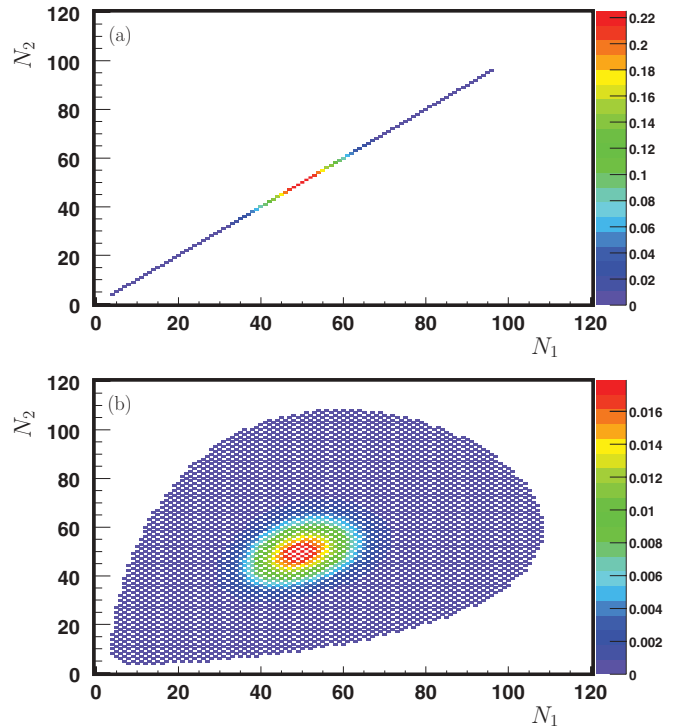

FIG. 2. (Color online) Illustration of a perfectly correlated data distribution (top). The mixed-event background (bottom) shows a wider distribution. The measured dynamical ratio fluctuations associated with the original distribution are then negative.

C. Illustrations

We are considering the example of a bivariate normal distribution (BND), with five parameters: marginal variances $\langle(\Delta N_1)^2\rangle_{\text{data}}$ and $\langle(\Delta N_2)^2\rangle_{\text{data}}$, their covariance $\langle\Delta N_1 \Delta N_2\rangle_{\text{data}}$, and two mean values, $\langle N_1\rangle_{\text{data}}$ and $\langle N_2\rangle_{\text{data}}$. Having specified these values, and having decided which event-mixing procedure should be applied, the measures σ_{data}^2, σ_{mix}^2, and σ_{dyn}^2 can be straightforwardly obtained. This implies the existence of a four-dimensional hypersurface in this parameter space representing BNDs with an equal amount σ_{dyn}^2 of dynamical ratio fluctuations.

Let us explore the limit of a BND, Fig. 2 (top), with completely correlated observables N_1 and N_2. The correlation coefficient

$$\rho = \frac{\langle\Delta N_1 \Delta N_2\rangle}{\sqrt{\langle(\Delta N_1)^2\rangle\langle(\Delta N_2)^2\rangle}} \qquad (5)$$

between them is then $\rho_{\text{data}}(N_1, N_2) = 1$. The ratio of N_1 and N_2 here in this example is then always equal to unity. The same will not be true for the mixed-event background, Fig. 2 (bottom), where the correlation will be weaker, $\rho_{\text{mix}}(N_1, N_2) < 1$. Following the above event-mixing procedure, one then quickly finds $\sigma_{\text{data}}^2 = 0$, $\sigma_{\text{mix}}^2 > 0$, and $\sigma_{\text{dyn}}^2 < 0$. In other words, despite the fact that the ratio is always one, the measure, discussed in Eqs. (1) and (4), suggests a degree of negative dynamical ratio fluctuations for this distribution.

Also other limiting cases could be explored. Here we mention the class of BNDs with covariance $\langle\Delta N_1 \Delta N_2\rangle = 0$, as illustrated in Fig. 3. The distribution $P_{\text{data}}(N_1, N_2)$ then
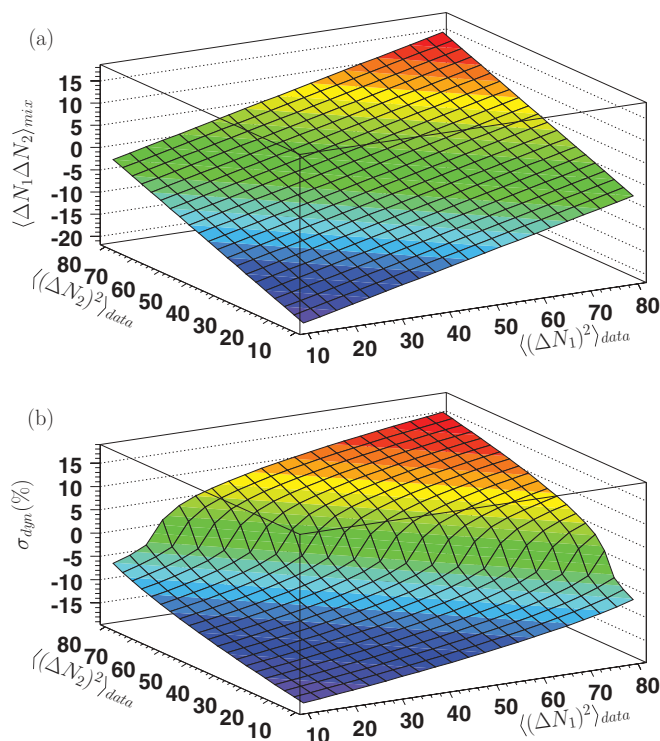
FIG. 3. (Color online) Illustration of uncorrelated ($\langle \Delta N_1 \Delta N_2 \rangle_{\text{data}} = 0$) data distribution leading to a correlated ($\langle \Delta N_1 \Delta N_2 \rangle_{\text{mix}} \neq 0$) mixed background (top) and resulting dynamic ratio fluctuations σ_{dyn} (bottom).

factorizes into the product of two normal distributions $P_{\text{data}}(N_1, N_2) = P_{\text{data}}(N_1) P_{\text{data}}(N_2)$. For the example shown in Fig. 3, a BND with mean values $\langle N_1 \rangle_{\text{data}} = \langle N_2 \rangle_{\text{data}} = 50$ and covariance $\langle \Delta N_1 \Delta N_2 \rangle_{\text{data}} = 0$ was chosen. In Fig. 3 the resulting covariance of the mixed-event background $\langle \Delta N_1 \Delta N_2 \rangle_{\text{mix}}$ (top) and the measure[3] of dynamical ratio fluctuations σ_{dyn} (bottom) are shown in their dependence on the variances $\langle (\Delta N_1)^2 \rangle_{\text{data}}$ and $\langle (\Delta N_2)^2 \rangle_{\text{data}}$. In general, one finds $\langle \Delta N_1 \Delta N_2 \rangle_{\text{mix}} \neq 0$ after event mixing, i.e., the mixed-event background now contains correlations, which the original sample did not.

IV. DISCUSSION

Having formulated rather general statements about the particularities of an event-by-event analysis of heavy-ion collision data, and having discussed a specific fluctuation measure, we feel the need to add a few more specific comments.

Measurements of fluctuations and correlations of particle multiplicities are often mentioned with regard to the detection of a critical point in the QCD phase diagram or other predictions for the properties of yet to be observed new states of matter. Certain aspects of distributions are supposed to

---

[3]Following the convention we present this number as $\sigma_{\text{dyn}} = \text{sgn}(\sigma_{\text{dyn}}^2)\sqrt{|\sigma_{\text{dyn}}^2|} \cdot 100\%$.

shed light on their (temporary) formation. Given the caveats discussed throughout this paper the authors are less optimistic.

The effects discussed above are hard to quantify as this would have to be done specifically and systematically. The biases introduced by a particular centrality selection criterion and a given particle acceptance would have to be studied separately for each observable. Focus should be put on the different solutions realized in the two main groups of experimental setups, fixed target and collider geometries. This was not the aim of this paper. Yet some examples can be quoted. Here we want to refer to a recent review paper [17], summarizing the Ph.D. thesis of V. Konchakovski. In particular, we mention Secs. VI and VIII.

The first example we want to quote is the correlation coefficient [31,32] between the multiplicities of charged hadrons in two narrow, symmetrically around midrapidity, arranged acceptance intervals, usually termed "forward-backward correlations." These correlations have been studied in their dependence on choice of centrality trigger and size of centrality classes [17,24,33]. A wider definition of centrality classes leads to a stronger correlation between forward and backward hemispheres. A larger system will have many particles in both acceptance windows. A smaller system will obviously have less particles. Combining systems of different sizes into one sample will lead to (strong) correlations in this sample from geometric considerations alone. A different choice of the centrality trigger will result in a different centrality dependence of the ones observable. No unambiguous (final state) measure is available to select only events of the "same (initial) system size" into one sample.

For the second example we want to return to the fluctuations of hadron multiplicity ratios. An HSD simulation of heavy-ion collisions has been performed for several center-of-mass energies. At low center-of-mass energy the data was analyzed twice; once for the NA49 acceptance for 3.5% most central collisions selected via the projectile spectator nucleon signal and once for the STAR acceptance for 5% most central collisions selected via the reference charged hadron multiplicity distribution. From the above considerations, differences between these two setups are to be expected. The situation is summarized as follows [17]: For some ratios of the multiplicities of selected hadrons, stronger differences between the two samples emerge than for others. Given our difficulty in interpreting the measure of dynamical ratio fluctuations, we find it hard to make a positive conclusion about the apparent numerical similarity of the results of the NA49 [34] and STAR [35] Collaborations.

It is equally hard to disentangle different contributions to ones fluctuation and correlation signals. The fact remains that correlation must not be confused with causation. We cordially acknowledge the commonly quoted motivation for performing statistical analysis of heavy-ion collision data. However, the (conjectured) effects to be investigated, or hoped to be discovered and established, are not the only factors determining the statistical properties of ones data sample.

The view that a (sharp) nonmonotonic behavior of a suitably chosen observable evaluated on a suitably constructed and modified data set can straightforwardly be taken as evidence for new physics (e.g., the detection of a critical point) is not held by the authors.

V. SUMMARY

This discussion largely fell into two parts, each of which tried to raise two points. First, the results of any event-by-event analysis depend strongly on how one constructs subsamples from minimum bias data. Minimum bias data from different experimental setups can vary quite strongly, and no initial state can be observed. Secondly, and equally important, is an experiment's capability to record accurate and detailed information about each event. The patchier the data set, the less detailed and conclusive any analysis of it, or comparison to another data set, can be.

We then discussed a particular procedure. Here we also tried to raise two points. First, a lot of information is discarded in favor of a single number, σ_{dyn}^2. This number might not be very indicative of a system's properties. Secondly, plainly constructing a reference distribution from a given data distribution according to some set of rules does not necessarily help to subtract correlations. Unfortunately, no reference (data) sample (where a certain effect is simply not contained) is available in high-energy heavy-ion collision physics.

Given the above considerations we find it hard to compare and interpret results from different experiments. Especially the dependence on the center-of-mass energy of the colliding ions might then carry little information.

[1] M. Bleicher *et al.*, Phys. Lett. B **435**, 9 (1998).
[2] H. Appelshauser *et al.* (NA49 Collaboration), Phys. Lett. B **459**, 679 (1999).
[3] S. V. Afanasev *et al.* (NA49 Collaboration), Phys. Rev. Lett. **86**, 1965 (2001).
[4] M. M. Aggarwal *et al.* (WA98 Collaboration), Phys. Rev. C **65**, 054912 (2002).
[5] J. Adams *et al.* (STAR Collaboration), J. Phys. G **34**, 799 (2007).
[6] Z. W. Chai *et al.* (PHOBOS Collaboration), J. Phys. Conf. Ser. **27**, 128 (2005).
[7] J. T. Mitchell, J. Phys. Conf. Ser. **27**, 88 (2005).
[8] T. J. Tarnowsky, J. Phys. Conf. Ser. **230**, 012025 (2010).
[9] M. M. Aggarwal *et al.* (STAR Collaboration), Phys. Rev. C **82**, 024912 (2010).
[10] T. J. Tarnowsky (STAR Collaboration), J. Phys. G **37**, 094037 (2010).
[11] M. A. Stephanov, K. Rajagopal, and E. V. Shuryak, Phys. Rev. Lett. **81**, 4816 (1998).
[12] M. A. Stephanov, K. Rajagopal, and E. V. Shuryak, Phys. Rev. D **60**, 114028 (1999).
[13] I. N. Mishustin, Phys. Rev. Lett. **82**, 4779 (1999).
[14] S. Jeon and V. Koch, in *Quark-Gluon Plasma 3*, edited by R. C. Hwa and X. N. Wang (World Scientific, Singapore), pp. 430–490.
[15] M. A. Stephanov, Phys. Rev. D **81**, 054012 (2010).
[16] V. Koch, PoS **CPOD2009**, 002 (2009).
[17] V. P. Konchakovski, M. I. Gorenstein, E. L. Bratkovskaya, and W. Greiner, J. Phys. G **37**, 073101 (2010).
[18] J. I. Kapusta, Phys. Rev. C **81**, 055201 (2010).
[19] C. Athanasiou, K. Rajagopal, and M. Stephanov, J. Phys. Conf. Ser. **270**, 012051 (2011).

[20] W.-j. Fu and Y.-l. Wu, Phys. Rev. D **82**, 074013 (2010).
[21] C. Athanasiou, K. Rajagopal, and M. Stephanov, Phys. Rev. D **82**, 074008 (2010).
[22] P. de Forcrand and O. Philipsen, in *Proceedings of Workshop on Strong and Electroweak Matter* (SEWM, Heidelberg, Germany, 2–5 Oct. 2002).
[23] J. Langelage and O. Philipsen, J. High Energy Phys. **01** (2010) 089.
[24] V. P. Konchakovski, M. Hauer, G. Torrieri, M. I. Gorenstein, and E. L. Bratkovskaya, Phys. Rev. C **79**, 034910 (2009).
[25] P. Braun-Munzinger, J. Stachel, and J. P. Wessels *et al.*, Phys. Lett. B **344**, 43 (1995).
[26] P. Braun-Munzinger, J. Stachel, and J. P. Wessels *et al.*, Phys. Lett. B **365**, 1 (1996).
[27] P. Braun-Munzinger, D. Magestro, and K. Redlich *et al.*, Phys. Lett. B **518**, 41 (2001).
[28] B. I. Abelev *et al.* (STAR Collaboration), Phys. Rev. Lett. **103**, 092301 (2009).
[29] C. Alt *et al.* (NA49 Collaboration), Phys. Rev. C **79**, 044910 (2009).
[30] G. Torrieri *et al.*, J. Phys. G **37**, 094016 (2010).
[31] T. Tarnowsky, R. Scharenberg, and B. Srivastava, Int. J. Mod. Phys. E **16**, 1859 (2007).
[32] T. J. Tarnowsky (STAR Collaboration), PoS **CPOD07**, 019 (2007).
[33] A. Bzdak, Phys. Rev. C **80**, 024906 (2009).
[34] C. Alt *et al.* (NA49 Collaboration), Phys. Rev. C **79**, 044910 (2009).
[35] S. Das (STAR Collaboration), J. Phys. G **32**, S541 (2006).