Partial correlation analysis method in ultrarelativistic heavy-ion collisions

Adam Olszewski[1,*] and Wojciech Broniowski[2,1,†]
[1]*Institute of Physics, Jan Kochanowski University, 25-406 Kielce, Poland*
[2]*The H. Niewodniczański Institute of Nuclear Physics, Polish Academy of Sciences, 31-342 Cracow, Poland*
(Received 15 June 2017; revised manuscript received 29 August 2017; published 13 November 2017)

We argue that statistical data analysis of two-particle longitudinal correlations in ultrarelativistic heavy-ion collisions may be efficiently carried out with the technique of partial covariance. In this method, the spurious event-by-event fluctuations due to imprecise centrality determination are eliminated via projecting out the component of the covariance influenced by the centrality fluctuations. We bring up the relationship of the partial covariance to the conditional covariance. Importantly, in the superposition approach, where hadrons are produced independently from a collection of sources, the framework allows us to impose centrality constraints on the number of sources rather than hadrons, that way unfolding of the trivial fluctuations from statistical hadronization and focusing better on the initial-state physics. We show, using simulated data from hydrodynamics followed with statistical hadronization, that the technique is practical and very simple to use, giving insight into the correlations generated in the initial stage. We also discuss the issues related to separation of the short- and long-range components of the correlation functions and show that in our example the short-range component from the resonance decays is largely reduced by considering pions of the same sign. We demonstrate the method explicitly on the cases where centrality is determined with a single central control bin or with two peripheral control bins.

DOI: 10.1103/PhysRevC.96.054903

I. INTRODUCTION

From the outset of the studies of correlations in ultrarelativistic heavy-ion collisions, it has been known that event-by-event fluctuations due to the choice of the centrality class of the sample lead to spurious effects that should be separated from the physical correlations. A simple example involves the forward-backward (FB) fluctuations of the multiplicity of produced hadrons, with the centrality defined from multiplicity in a suitable reference bin, or via some other quantity obtained from detector response, correlated to multiplicity. As the measurements in the physical F and B bins are correlated to the reference bin, the FB correlations determined via a naive definition depend strongly on the fluctuations in the reference bin, thus the character of the "true" FB correlations is obscured.

Several remedies have been proposed to cure the problem of *centrality fluctuations*. First, if the size of the data sample allows, then one may use sufficiently narrow centrality bins, such that centrality fluctuations are negligible. Then, to improve statistics, one may average the obtained covariance matrices over several narrow centrality bins within a broader class. This method, essentially based on the concept of *conditional correlation* (see, e.g., Ref. [1]), was successfully used in the analysis by the STAR Collaboration [2], with further proposals presented in Ref. [3].

In this paper we explore the so-called *partial correlation* analysis (see, e.g., Refs. [1,4]) in application to ultrarelativistic heavy-ion collisions. The method is closely related to the conditional correlations (see Appendix B); however, it offers an appealing simplicity as well as immediate insight into conditional independence of the studied variables. It has been widely used in other domains of statistical applications, ranging from physics (see, e.g., an interesting example from the x-ray spectroscopy, where partial correlations are used to remove the spurious effects of beam intensity fluctuation [5]) to medicine and psychology. The basic goal of the approach is to assess the correlation (or independence) of some "physical" variables, where the sample is determined by certain fluctuating control (or *external* or *nuisance*) variables, whose effects needs to be removed to accomplish the understanding of the relations between the physical variables.

We bring up the relationship of the partial covariance to the conditional covariance [6,7], which holds under conditions which are well satisfied in ultrarelativistic heavy-ion collisions. Then, the partial covariance with a control bin may be understood as conditional covariance with the fixed hadron multiplicity in the control bin. An important point, however, is that one should impose centrality constraints at the level of initial state rather than finally produced hadrons; that way, we limit the external fluctuations concerning the initial production, which are of our principal interest. One can accomplish this goal in the framework of the superposition model [8] of particle production, where hadrons are emitted from independent sources. We extend the partial correlation analysis to this physically interesting case. The result is a simple modification of the partial-correlation formulas, where variances have the autocorrelation terms removed. That way one can impose the constraints at the level of the initial sources, which is a nontrivial outcome of the partial covariance method in the superposition approach.

In our study, we use simulated events for ultrarelativistic heavy-ion collisions generated with event-by-event 3+one-dimensional (1D) viscous hydrodynamics [9] and `THERMINATOR` [10,11] to argue that the partial covariance technique is a practical tool to analyze two-particle correlations. We apply it to examples where centrality is determined with

*Adam.Olszewski.fiz@gmail.com
†Wojciech.Broniowski@ifj.edu.pl

the multiplicity in a single midrapidity control bin, as well as in two peripheral-rapidity control bins.

We verify in our examples that with the removed effects of resonance decays, the partial FB multiplicity correlations obtained from the simulated data reflect very closely the initial correlations in spatial rapidity, here implemented in the very simple Bzdak-Teaney (BT) [12] form for the wounded quark model [13]. This feature shows that the partial correlation technique may be used in practice to access information on the initial-state correlations. However, as in other approaches, this possibility relies on the separation of the short-range correlations (such as from the resonance decays, jets, or femtoscopic correlations) from the long-range correlations, generated in the initial state.

We recall that another approach separating centrality fluctuations, somewhat different in spirit and derived from the superposition approach, defines the so-called *strongly intensive measures* [14,15], which by construction do not depend on fluctuations of the number of sources (or the volume) which emit the observed hadrons. Yet another technique which can be used to accomplish the goal invokes the principal components analysis (PCA) [16]. We discuss the relation of the partial covariance technique to these methods in Appendix C.

The outline of our paper is as follows: In Sec. II we provide the basic definitions of the partial covariance and partial correlation and discuss their meaning linked to imposing conditions with control variables. Then in Sec. III we pass to the discussion of the FB multiplicity correlations, followed in Sec. IV by the description of the superposition approach to the multistage particle production process. Our key formulas, allowing us to impose centrality fixing conditions at the level of sources rather than hadrons, are derived there. In Sec. V we obtain partial correlations in the initial state for the BT model, which are confronted with the corresponding results obtained with simulated data in Sec. VI. In Sec. VII we recapitulate our results, summarizing in practical terms the method based on partial correlations, which can be used in experimental analyses. The appendices contain the discussion of the relation of the partial covariance analysis to other methods, as well as some technical details.

II. PARTIAL CORRELATION

In this section we establish the notation and provide the standard definitions, with more details presented in Appendix A.

The simplest case of partial correlations [1,4] involves two physical random variables, X and Y, and a single control random variable Z. One defines the element of the covariance matrix in the standard way as

$$\mathrm{c}(A,B) = \langle AB\rangle - \langle A\rangle\langle B\rangle, \quad A,B = X,Y,Z,$$

where $\langle . \rangle$ denotes the averaging over the sample of n events. The variance is, of course, the diagonal term,

$$\mathrm{v}(A) = \langle A^2\rangle - \langle A\rangle^2 = \mathrm{c}(A,A). \tag{1}$$

The partial covariance between X and Y with the control variable Z is defined by the formula (see Appendix A for interpretation)

$$\mathrm{c}(X,Y\bullet Z) = \mathrm{c}(X,Y) - \frac{\mathrm{c}(X,Z)\mathrm{c}(Z,Y)}{\mathrm{v}(Z)}. \tag{2}$$

The second term in Eq. (2) removes the piece of the covariance between X and Y which is due to their correlation to Z by means of projecting out the components of X and Y which are parallel to Z (in the n-dimensional space, with n denoting the number of events). The partial variance is, correspondingly,

$$\mathrm{v}(A\bullet Z) = \mathrm{v}(A) - \frac{\mathrm{c}(A,Z)^2}{\mathrm{v}(Z)} = \mathrm{c}(A,A\bullet Z), \quad A = X,Y. \tag{3}$$

In experimental studies one often uses the correlation function defined as the covariance scaled with the multiplicities, i.e.,

$$C(X,Y) = \frac{\mathrm{c}(X,Y)}{\langle X\rangle\langle Y\rangle}, \tag{4}$$

and, correspondingly, $V(A) = \mathrm{v}(A)/\langle A\rangle^2$. Then the partial C correlation following from Eq. (2) is

$$C(X,Y\bullet Z) = C(X,Y) - \frac{C(X,Z)C(Z,Y)}{V(Z)}. \tag{5}$$

Finally, one defines the partial analog of Pearson's ρ correlation [1,4],

$$\begin{aligned}\rho(X,Y\bullet Z) &= \frac{\mathrm{c}(X,Y\bullet Z)}{\sqrt{\mathrm{v}(X\bullet Z)\mathrm{v}(Y\bullet Z)}}\\ &= \frac{\rho(X,Y)-\rho(X,Z)\rho(Z,Y)}{\sqrt{1-\rho(X,Z)^2}\sqrt{1-\rho(Z,Y)^2}},\end{aligned} \tag{6}$$

where

$$\rho(A,B) = \mathrm{c}(A,B)/\sqrt{\mathrm{v}(A)\mathrm{v}(B)}. \tag{7}$$

As discussed in a greater detail in Appendix B, the partial covariance, under quite general assumptions [6,7] which are typically fulfilled in ultrarelativistic heavy-ion collisions, is related to the conditional covariance, namely

$$\mathrm{c}(X,Y\bullet Z) \simeq \mathrm{c}(X,Y|Z), \tag{8}$$

where the meaning of the condition in $\mathrm{c}(X,Y|Z)$ is that first Z is fixed to a very narrow class (for instance, if it describes the discrete multiplicity of hadrons in the reference bin, it can be fixed to a natural number equal to this multiplicity), then the covariance between X and Y is evaluated within this subsample, and, finally, averaging of thus-obtained covariances over various values of Z within the sample is performed.

As a matter of fact, such a conditional procedure was used in the STAR experiment [2] to analyze the FB multiplicity correlations in Au+Au and $p+p$ collisions at $\sqrt{s_{NN}} =$ 200 GeV. In this context, relation (8) was derived by Lappi-McLerran [17] under the assumption of normal distributions, and by Bzdak [18] with the condition (B1), similarly as in Refs. [6,7].

The practical significance of Eq. (2) or (5) is that the imposition of external constraints (such as fixing centrality)

may be, under general assumptions, accomplished via the partial covariance technique.

III. FORWARD-BACKWARD MULTIPLICITY CORRELATIONS

In our study, X, Y, and Z random variables are the multiplicities of produced hadrons in, correspondingly, a forward pseudorapidity bin F, a backward pseudorapidity bin B, and a reference bin. We will explore two cases: reference bin C located at a central pseudorapidity bin, and the sum of bins L and R, located symmetrically at peripheral rapidities η_L and $\eta_R = -\eta_L$. Multiplicity in the reference bin determines the centrality of the event. The location of the reference bin is fixed, whereas F, centered around η_1, and B, centered around η_2, vary in the covered pseudorapidity range (in this more general arrangement we do not request that F is forward and B backward, but they assume any location in pseudorapidity within the acceptance range). We will also consider the case where the multiplicities in L and R bins are fixed independently, according to the case with two constraints described in Appendix A.

Experimental studies of the longitudinal multiplicity correlations have a long history. Early investigations of pp and $p\bar{p}$ collisions [19–25] and nuclear collisions [26,27] were followed by the studies of ultrarelativistic heavy-ion and pp reactions at the Relativistic Heavy Ion Collider [2,28] and at the Large Hadron Collider (LHC) [29–35]. Numerous physical models and theoretical methods have been invented in attempts to understand the mechanisms behind the generation of long-range correlations [3,12,17,18,36–63]. The importance of these investigations lies in the well-known fact that the correlations over a large rapidity separation can originate only from the earliest stages of the collision, and thus they may reveal fingerprints of the early dynamics of the system.

In our study we demonstrate the methodology based on partial correlations with simulated events for Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, described in detail in Sec. VI.

IV. SUPERPOSITION APPROACH

One should bear in mind that the particle production in ultrarelativistic heavy-ion collisions is an effect of a multistage process (see, e.g., Refs. [8,64]). First, we have an early production of entropy, resulting from partonic physics. It is modeled in various approaches, such as string formation [40,45], the color glass condensate theory [65,66], or the Glauber approach [67–72], used in this work. The initial partonic entropy is distributed in space in a correlated way.

This distribution, treated event by event, serves as an initial condition for the intermediate hydrodynamic evolution (for reviews see, e.g., Refs. [73,74] and references therein) or for transport modeling [75]. In our work the applied hydrodynamics is deterministic, and hence it does not introduce extra fluctuations, which should arise in a viscous system [76] but which according to estimates are not very significant [77,78].

The intermediate evolution continues until freeze-out, where the Cooper-Frye formalism [79] is applied at the freeze-out hypersurface defined with a constant temperature (we take $T_f = 150$ MeV). It generates *primordial* hadrons (stable and resonances) according to a thermal distribution. Subsequently, resonances undergo decays, possibly in cascades, into stable particles. Due to a statistical nature of the production process, the distribution of a hadron of a given species is Poissonian, and hence (trivial) fluctuations are generated due to the sampling with a finite number of particles.

The goal of the data analyses is to unfold of the trivial fluctuations [14,15,80–83], such as those from statistical hadronization, and acquire information on correlations generated in the earlier evolution phases. Below we describe how this is accomplished in the superposition approach.

Let us first bring up an important approximation underlying this approach, which may be termed as *no bin mixing*. The initial distribution of entropy may be divided into cells labeled with their space-time rapidity

$$\eta_\parallel = \frac{1}{2}\log\left(\frac{t+z}{t-z}\right) \tag{9}$$

(and also by the transverse coordinates x and y). The bins are "carried over" with hydrodynamics or transport to the freeze-out hypersurface, with the assumption of no mixing between the bins, such that the final pseudorapidity of the fluid element, η, is a function of $\eta_\parallel$. The hydrodynamic push in the longitudinal direction is not very strong, and we estimate

$$\eta \simeq k\eta_\parallel, \tag{10}$$

with $k = 1.2$ for the model we apply.

At freeze-out, there is some thermal dispersion of the momenta of hadrons, as particles originating from the same hydrodynamic fluid cell acquire rapidities spread with $\Delta\eta \sim 1$. This causes some bin mixing of the two-particle correlation function in pseudorapidity. In addition, resonance decays generate extra short-range correlations with the width of $\Delta\eta \sim 1$ (we will return to the issue of separating the short-range components when describing our results in Sec. VI).

We remark that inclusion of the detector acceptance into the framework does not lead to a new element. If the acceptance is expressed with a Bernoulli trial of success rate p, then its folding with a Poisson distribution from the thermal motion leads to a Poisson distribution with a mean enhanced by the factor p. Sums of various particle species with Poisson distributions also lead to a Poisson distribution with the mean expressed as a sum means of the added distributions.

Our basic methodology is as follows: We will use simulated data (with removed short-range component coming from resonance decays) to obtain the two-particle partial correlation function with the trivial fluctuations unfolded with the help our formalism. The result will be directly compared to the partial correlation function of the initial state used in the simulations (cf. Sec. V), with arguments shifted according to Eq. (10).

We now come to the derivation of relevant formulas. One may consider the fluid cells at the freeze-out hypersurface as *sources*, which emit hadrons. Each source, by definition, emits particles independently of other sources, but with the same (for the sake of simplicity) distribution. The number of fluid cells (sources) at the forward F and backward B pseudorapidity is denoted as S_F and S_B, respectively (recall that the fluid cells at a given rapidity are located at various transverse positions

and their number fluctuates). Then, the number of produced hadrons in each bin is

$$N_A = \sum_{i=1}^{S_A} m_i, \quad A = F, B, \tag{11}$$

where m_i denotes the number of particles produced from source i. We assume that each source produces particles with the same distribution, hence $\langle m_i \rangle = \langle m \rangle$ and $\mathrm{v}(m_i) = \mathrm{v}(m)$. The assumption of the independent production leads to simple formulas which connect moments of the number of sources with the moments of the number of the produced particles [8,64]:

$$\begin{aligned} \langle S_A \rangle \langle m \rangle &= \langle N_A \rangle, \\ \mathrm{c}(S_A, S_{A'}) \langle m \rangle^2 &= \mathrm{c}(N_A, N_{A'}) - \delta^{AA'} \omega(m) \langle N_A \rangle \\ &\equiv \overline{\mathrm{c}}(N_A, N_{A'}), \end{aligned} \tag{12}$$

where $\omega(X) = \mathrm{v}(X)/\langle X \rangle^2$ is the scaled variance. We notice a subtraction of $\omega(m)\langle N_A \rangle$ appearing for the diagonal term $A = A'$, i.e., for the variance:

$$\begin{aligned} \mathrm{v}(S_A) \langle m \rangle^2 &\equiv \mathrm{c}(S_A, S_A) \langle m \rangle^2 = \mathrm{v}(N_A) - \omega(m) \langle N_A \rangle \\ &\equiv \overline{\mathrm{v}}(N_A). \end{aligned} \tag{13}$$

The origin of this term is the presence of variance (or autocorrelation) of the particles produced from the same sources, $\mathrm{v}(m_i) = \mathrm{c}(m_i, m_i) > 0$, whereas the covariance of the overlaid distributions from different sources vanishes by the assumption of independent production, $\mathrm{c}(m_i, m_j) = 0$ for $i \neq j$.

For the special case of the Poisson distribution of the overlaid variable m (which is the case of our numerical study presented in Sec. VI), we have $\omega(m) = 1$ and $\overline{\mathrm{v}}(N_A) = \langle N_A(N_A - 1) \rangle - \langle N_A \rangle^2$, which corresponds to the subtraction of autocorrelations; the average number of pairs appears in the formula. However, Eq. (13) is more general, holding, under the assumptions of the superposition model, for any distribution of the overlaid variable m. In this paper, we refer to the subtraction of $\omega(m)\langle N_A \rangle$ from the variance as to "the removal of autocorrelations" and indicate it with a bar.

In an experimental study, where in principle one does not know the distribution of m, one may find $\omega(m)$ in a numerical way with the following procedure: One examines the covariance matrix $\mathrm{c}(N_A, N_{A'})$ as a two-dimensional matrix in A and A' indices. The diagonal term at $A = A'$ forms a sharp discontinuous ridge, sticking up from a smooth "background" function. One then adjusts the value of $\omega(m)$ to remove the sharp ridge from the function $\mathrm{c}(N_A, N_{A'}) - \delta^{AA'} \omega(m) \langle N_A \rangle$. This prescription conforms to Eq. (12), as the covariance of the sources $\mathrm{c}(S_A, S_{A'})$ is a smooth function of A and A'.

The meaning of Eq. (13) is that the covariance of the number sources is proportional to the covariance of the observed hadron multiplicities, but with the autocorrelations removed. Passing to the scaled covariance (4), we have

$$C(S_A, S_{A'}) = C(N_A, N_{A'}) - \delta^{AA'} \frac{\omega(m)}{\langle N_A \rangle} \equiv \overline{C}(N_A, N_{A'}). \tag{14}$$

We are now ready to build the partial covariance for the superposition approach. Some introductory discussion is in place. As stated in Sec. II and further discussed in Appendix B, the meaning of the partial covariance is, essentially, an imposition of a condition on the value of the control variables. In the case of multiplicity correlations in ultrarelativistic heavy-ion collisions, a first instinct is to constrain the centrality fluctuations, i.e., the fluctuations of number of hadrons in a reference bin. However, a more desired constraint concerns the number of sources corresponding to the reference bin. Such a constraint is more directly related to a physical situation in the initial state. Suppose a reference bin, defined as C, has in a given event S_C sources which "determine the physics." On the other hand, the number of detected hadrons N_C is sensitive to the fluctuation in the production from sources. Therefore a fixed value of N_C corresponds to event-by event fluctuating values of S_C; constraining N_C does not completely constrain S_C and the physics of the initial condition remains washed out.

According to our formalism, the constraint imposed at the level of initial sources is realized with the equation

$$\begin{aligned} C(S_F, S_B \bullet S_C) &= \overline{C}(N_F, N_B) - \frac{\overline{C}(N_F, N_C)\overline{C}(N_B, N_C)}{\overline{\mathrm{v}}(N_C)} \\ &\simeq C(S_F, S_B | S_C). \end{aligned} \tag{15}$$

This is the key formula used in our analysis of the simulated data in the following sections.

If one wishes, however, to impose the constraint at the level of the produced hadrons, then using the formula

$$\mathrm{c}(S_A, N_{A'}) = \mathrm{c}\left(S_A, \sum_i^{S_{A'}} m_i\right) = \langle m \rangle \mathrm{c}(S_A, S_{A'}),$$

one arrives at the expression

$$\begin{aligned} C(S_F, S_B \bullet N_C) &= \overline{C}(N_F, N_B) - \frac{\overline{C}(N_F, N_C)\overline{C}(N_B, N_C)}{\mathrm{v}(N_C)} \\ &\simeq C(S_F, S_B | N_C). \end{aligned} \tag{16}$$

Note that the subtle but important difference between Eq. (15) and (16) is in the denominator of the subtracted term, where we find the variance of the multiplicity in the reference bin with autocorrelations subtracted or present. In our sample, the autocorrelations increase the variance by ∼100%, hence the effect is very important.

We stress that despite its simplicity, the meaning of Eq. (15) is nontrivial, as it allows us to impose a strict centrality constraint at the level of *sources* and infer partial correlation of sources, whereas the evaluation is based solely on measured multiplicities of the *produced hadrons*.

V. MODELING INITIAL CORRELATIONS

For our illustrative purposes, we use the wounded quark model [84,85] for the initial state. In this model, the initial sources are the wounded quarks, moving forward or backward, according to the motion of their parent nucleons from nuclei A and B. An advantage of using the wounded quarks compared to wounded nucleons [69] is that one obtains proper scaling [13,86–90] of the multiplicities on the number of participants with no need for the binary-collision component

[70]. The event-by-event distribution of wounded quarks in the transverse plane is obtained from the Glauber simulations with GLISSANDO [91,92], corresponding to the transverse location of the wounded quarks. The longitudinal profile in spatial rapidity $\eta_\parallel$ is taken according to the model of "triangles" [93–95], where each source has the entropy distributed preferentially in the direction of its motion, according to a simple formula

$$f_{A,B}(\eta_\parallel) = \frac{y_b \pm \eta_\parallel}{y_b} h(\eta_\parallel), \quad \text{for } |\eta_\parallel| < y_b, \tag{17}$$

where A and B denote the sources belonging to, respectively, the left- and right-moving nuclei, y_b is the rapidity of the beam ($y_b \simeq 8$ for Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV), and $h(\eta_\parallel)$ is an additional profile, typically of a flattened Gaussian form [95]. As $h(\eta_\parallel)$ cancels from the formulas for symmetric $A + B$ collisions, we do not need to specify it explicitly.

Let us introduce the notation Q_A and Q_B for the number of wounded quarks belonging to the A and B nuclei, and $Q_\pm = Q_A \pm Q_B$. Then, according to Eq. (17), the number of sources (combined from nucleus A and B) at location $\eta_\parallel$ is

$$S(\eta_\parallel) = \left(Q_+ + Q_- \frac{\eta_\parallel}{y_b}\right) h(\eta_\parallel). \tag{18}$$

For symmetric collisions the average over events yields $\langle Q_- \rangle = 0$, and hence $\langle S(\eta_\parallel)\rangle = \langle Q_+ \rangle h(\eta_\parallel)$.

Bzdak and Teaney [12] computed the correlation function in the model of triangles (in the variant with wounded nucleons). It yields a very simple result,

$$C(S_F, S_B) = \frac{\mathrm{v}(Q_+)}{\langle Q_+ \rangle^2} + \frac{\mathrm{v}(Q_-)}{\langle Q_+ \rangle^2} u_1 u_2, \tag{19}$$

where

$$u_{1,2} = \frac{\eta_{\parallel 1,2}}{y_b} = \frac{\eta}{k y_b} \tag{20}$$

[cf. Eq. (10)], and indices 1 and 2 correspond to labels F and B, respectively. Note that, as announced, the overall rapidity profile $h(\eta_\parallel)$ cancels out. The moments of $Q_\pm$ are read out from GLISSANDO simulations via averaging over some chosen class of events.

Next, we use Eq. (5) to derive the partial covariance function for the BT model with the control bin C placed at $\eta_\parallel = 0$. A short calculation yields

$$C(S_F, S_B \bullet S_C) = \frac{\mathrm{v}(Q_-)}{\langle Q_+ \rangle^2} u_1 u_2. \tag{21}$$

We note that Eq. (21) differs from Eq. (19) by not carrying the term with $\mathrm{v}(Q_+)$. This is clear from the point of view of the conditional correlation, as in the present case $C(S_F, S_B|S_C)$ corresponds to fixing the multiplicity at $\eta_\parallel = 0$. From Eq. (18) we get $S(0) = Q_+ h(0)$, hence this is equivalent to fixing Q_+, and thus $\mathrm{v}(Q_+) = 0$ in the calculation of the conditional covariance. The obtained consistency verifies in an obvious way the relation $C(S_F, S_B \bullet S_C) = C(S_F, S_B|S_C)$.

For the ρ correlation of Eq. (6) we find the very simple formula

$$\rho(S_F, S_B \bullet S_C) = \mathrm{sgn}(u_1 u_2), \tag{22}$$

where sgn denotes the sign function. This means that the partial ρ correlation of multiplicities in bins located at rapidities of the same (opposite) sign is $+1$ (-1), indicating maximum correlation (anticorrelation).

In experiments, it frequently happens that multiplicities in peripheral forward (R) and distant backward (L) bins are available and can be used for centrality determination. Below we consider two cases: (1) where the sum of the multiplicities in L and R is taken and Eq. (5) is used and (2) the multiplicities in L and R are taken as separate constraints according to Eq. (A11). We consider the case where the peripheral bins are symmetrically arranged, with R and L around pseudorapidities η_R and $\eta_L = -\eta_R$.

In case (1) the BT model yields exactly the same result as Eq. (21),

$$C(S_F, S_B \bullet S_L + S_R) = \frac{\mathrm{v}(Q_-)}{\langle Q_+ \rangle^2} u_1 u_2, \tag{23}$$

which follows from the fact that according to Eq. (18) $S(\eta_L) + S(\eta_R) \sim Q_+$, and the condition with symmetrically arranged peripheral bins fixes Q_+, as in the case of the central bin.

In case (2) Eq. (A11) gives for the BT model a vanishing result,

$$C(S_F, S_B \bullet S_L, S_R) = 0, \tag{24}$$

which is compatible with the simultaneous constraints $S(\eta_L) = 0$ and $S(\eta_R) = 0$, which fixes both Q_+ and Q_-, hence $\mathrm{v}(Q+) = \mathrm{v}(Q_-) = 0$, and $C(S_F, S_B|S_L, S_R) = 0$. We note that this is a specific feature of the BT model, whereas models which include extra fluctuations in the initial state, e.g., the fluctuating-string model of Ref. [60], would yield a nonzero result.

VI. RESULTS FOR THE SIMULATED EVENTS

Our final goal is to test to what extent the formulas obtained in the previous section are reproduced with simulated data and the master formula (15) and its equivalent for the case of left and right peripheral bins. Because of departures from assumptions of the superposition approach (bin mixing, resonance decays) this is not an academic exercise but a verification if the method may be practical in actual data analyses.

The simulated data are obtained as follows: For the initial condition, we apply the wounded quark model for Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, as described in Sec. V. We take a rather broad sample corresponding to centrality (as determined with Q_+) of 30–40%. Then we use the results of event-by-event 3+1D viscous hydrodynamics [96] obtained with the wounded-quark initial conditions. The statistical hadronization at freeze-out is carried out with THERMINATOR [10,11], which incorporates all hadrons from the Particle Data Tables and implements resonance decays. We label the results obtained with the products of resonance decays as "all charged," which consists of $\pi^\pm$, $K^\pm$, p, and $\bar{p}$. To reduce the

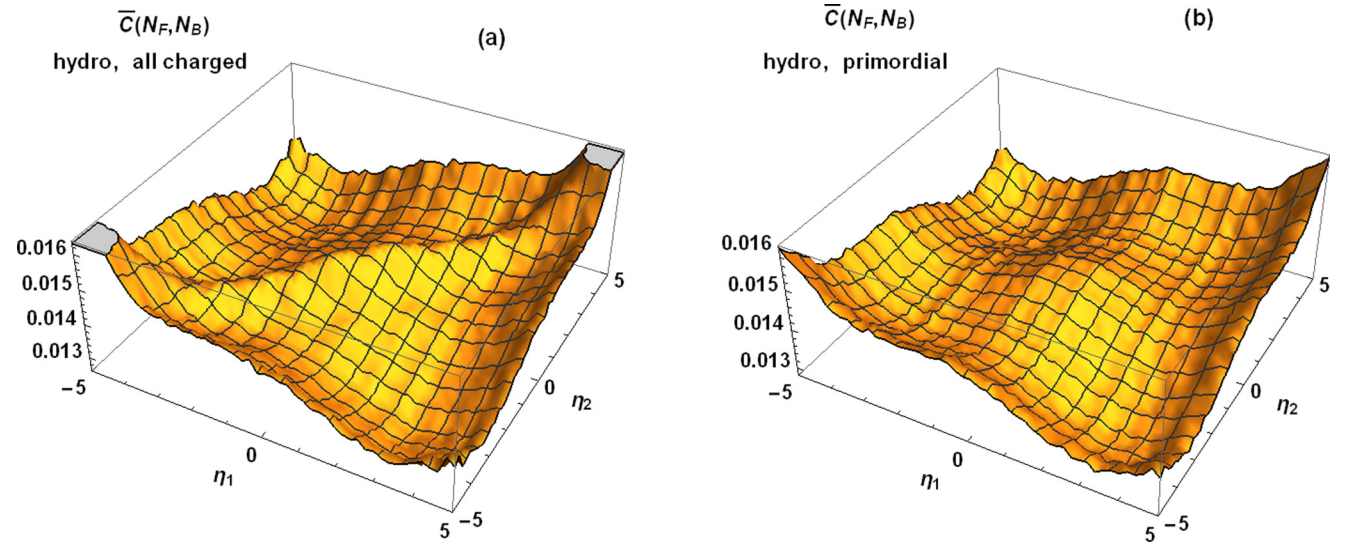

FIG. 1. Two-particle $\overline{C}$ correlation (with removed autocorrelations) in pseudorapidity of hadron multiplicities for the simulated data (a) for all charged particles after resonance decays and (b) for the primordial hadrons. Wounded quark model, 3+1D viscous event-by-event hydrodynamics, statistical hadronization. Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV in centrality class 30–40%.

correlations induced by the resonance decays, we also present the results obtained with the "primordial" particles generated at freeze-out, i.e., before the resonance decays.

As we wish to have a view as broad as possible, we take a large continuous acceptance window in our numerical study, with $|\eta_{1,2}| \leqslant 5.1$ divided into 51 *physical* bins of width of 0.2. Our coverage is thus (on purpose) much larger than accessible in the LHC experiments to better illustrate the approach. The central *reference* bin, where the number of hadrons in N_C, is taken as $\eta_C \leqslant 0.5$, and thus it has a width of five physical bins (other values of width of the reference bin could be taken, with similar results up to statistical uncertainties).

Figure 1 presents the basic output from the simulated data, namely the $\overline{C}$ correlation (with autocorrelations removed), obtained with (a) all charged particles (i.e., after resonance decays) and (b) for the primordial particles. In case (a) we note a hallmark ridge, wide by about one unit of pseudorapidity, extending along the diagonal. It is due to the resonance decays, which is the basic difference between cases (a) and (b). We also note other nontrivial structures in the $\overline{C}$ correlation arising from the applied hydrodynamic model, such as its rise at the boundaries; however, understanding these intricate details is not the goal of this work. Rather, we take the simulated data (which, as we see, are not trivial) as they are and then carry out the partial correlation analysis outlined in the previous sections.

In Fig. 2 we compare the partial C-correlation function for the source multiplicities with the central reference bin,

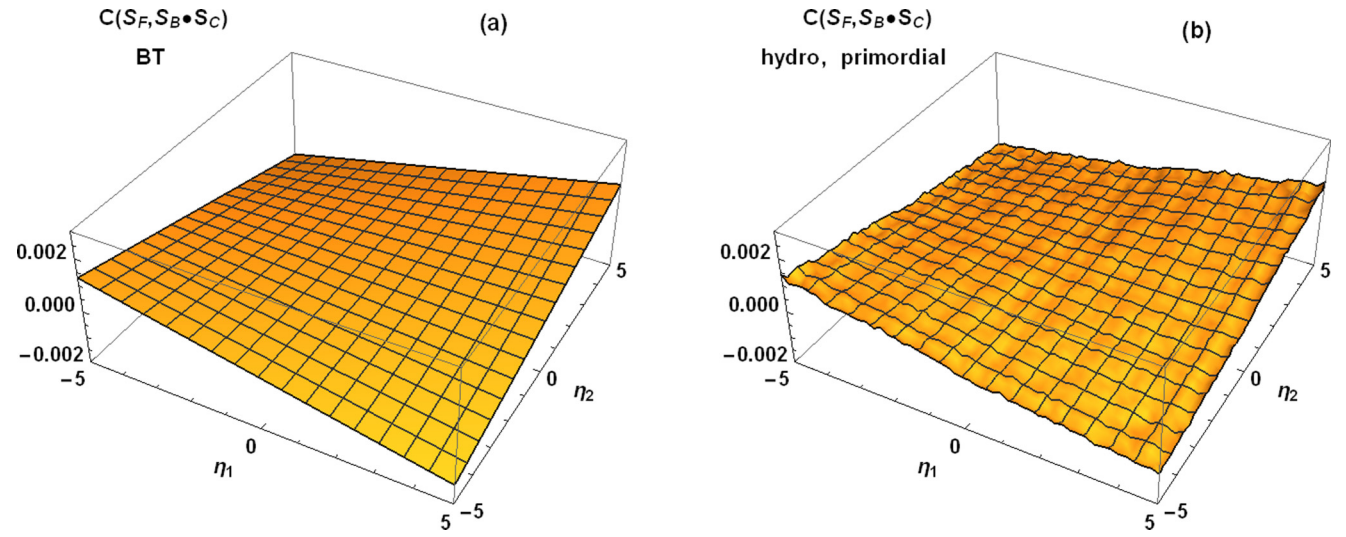

FIG. 2. Partial C-correlation function for the source multiplicities with the central reference bin (a) for the BT model and (b) for the simulated data with primordial hadrons.

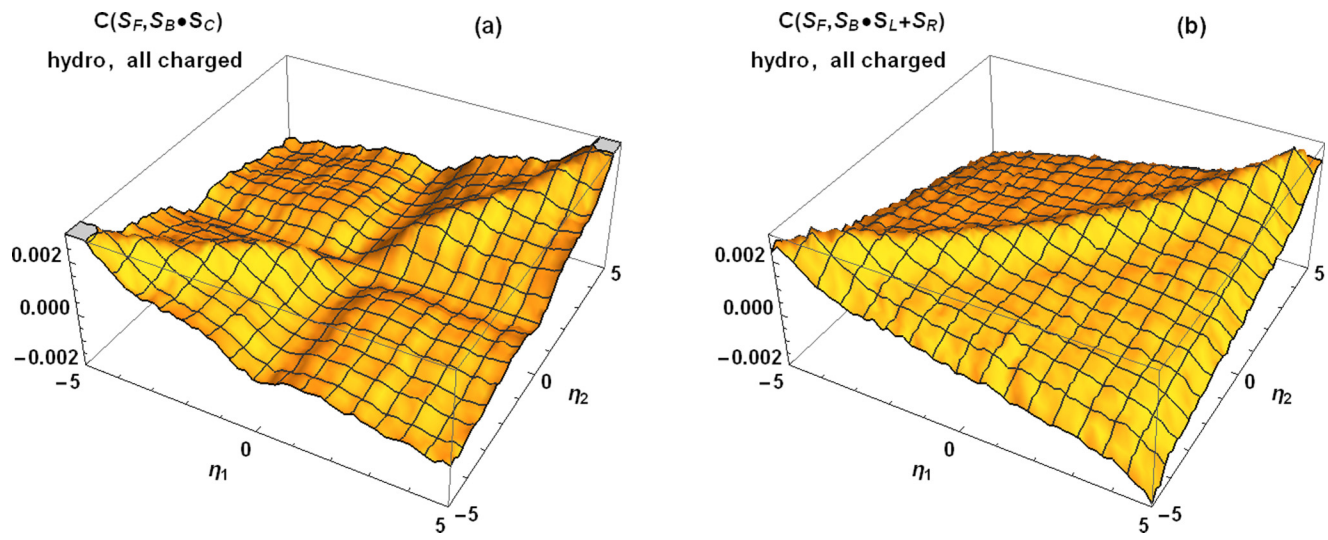

FIG. 3. Partial C-correlation function for the source multiplicities for the simulated data for all charged hadrons, obtained with (a) the central reference bin and (b) with the sum of left and right peripheral bins.

obtained for the BT model with Eq. (21), shown in panel (a), to the same quantity obtained from the simulated data with *primordial* hadrons with Eq. (15). We note that the overall agreement of (a) and (b) is quite remarkable. Hence, if we had the possibility of separating the resonance decays (and in real data, also other sources of short-range correlations from later stages of the evolution), then we could infer the initial state correlations with the presented methodology. Note that the application of Eq. (15) leads to large cancellations when passing from Fig. 1(b) to Fig. 2(b). Also, the nonuniformities are "miraculously" smoothed out, and hence they had to originate from centrality fluctuations.

Separating other sources of correlations is of course far from simple. One method (as done by the ATLAS collaboration in Ref. [97]) is to do a numerical fit to a function describing the short-range corrections and then simply subtract it. In our illustration, it would correspond to "skimming" the ridge in Fig. 1(a), and then carrying out the calculation, which would lead to a result similar to Fig. 2(b).

We note that our model does not account for the correlations reflecting the conservation laws, which is not essential for testing the partial covariance approach.

When the short-range effects from the resonance decays are kept, i.e., we are using the data from Fig. 1(a), then the resulting partial C correlation with the central reference bin has the shape shown in Fig. 3(a). We note the ridge along the $\eta_1 = \eta_2$ diagonal but also several other features. First, we can see a depletion along the lines $\eta_1 = 0$ and $\eta_2 = 0$. This is a simple artifact of the central reference bin placed at $\eta = 0$, since the definition (5) has the feature that when one of the measurement bins is equal to the reference bin, e.g., $Y = Z$, then $C(X, Z \bullet Z) = 0$ identically. Suppose we decompose the

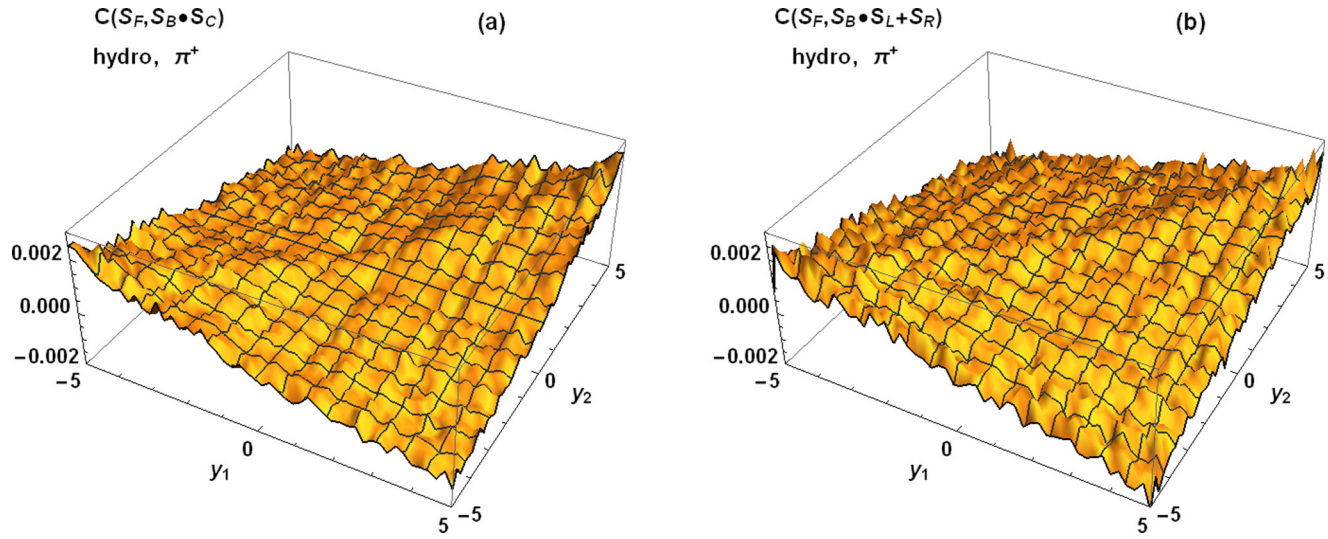

FIG. 4. Partial C-correlation function for the source multiplicities for the simulated data for π^+, with (a) the central reference bin and (b) with the sum of left and right peripheral bins.

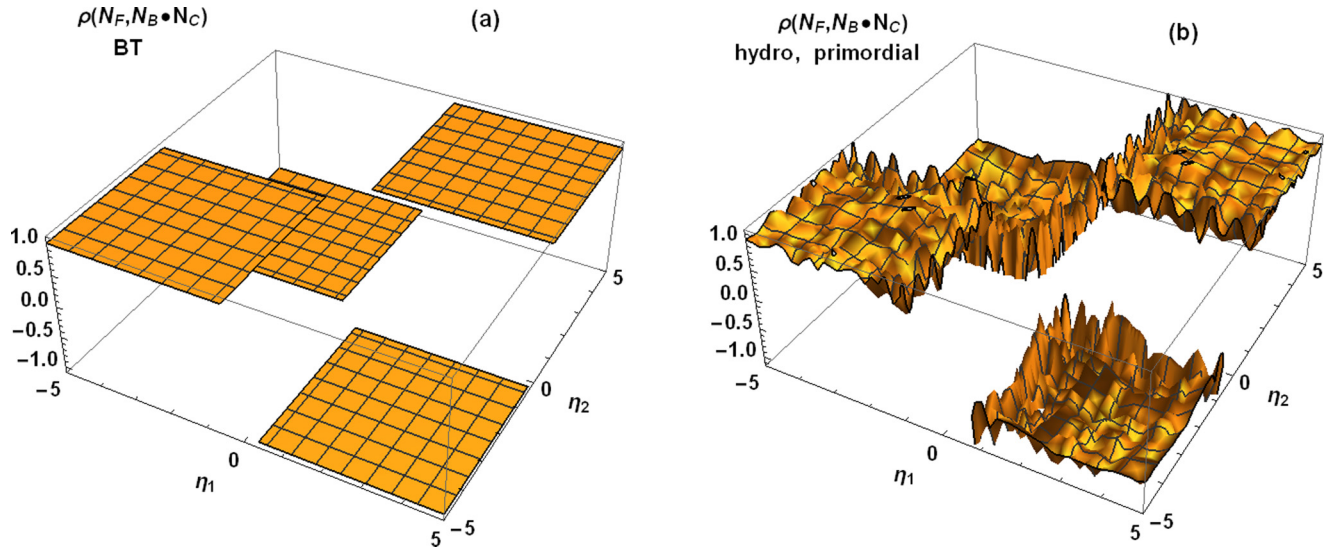

FIG. 5. Partial ρ correlation for the source multiplicities (a) for the BT model and (b) for the simulated data with primordial hadrons.

measured correlation function into the short- (s) and long-range (l) components, $\overline{C} = \overline{C}_s + \overline{C}_l$. Then Eq. (15) becomes

$$C(S_F, S_B \bullet S_C) = \overline{C}_s(N_F, N_B) + \overline{C}_l(N_F, N_B) - \frac{[\overline{C}_s(N_F, N_C) + \overline{C}_l(N_F, N_C)][\overline{C}_s(N_B, N_C) + \overline{C}_l(N_B, N_C)]}{\overline{\mathrm{v}}_s(N_C) + \overline{\mathrm{v}}_l(N_C)}. \tag{25}$$

We see from here all the features coming out in Fig. 3(a). When $F \simeq B$, we get the ridge from $\overline{C}_s(N_F, N_B)$, and when $F \simeq C$ or $B \simeq C$, the function drops to zero for the reason discussed above Eq. (25). Note, however, that even when F, B, and C are all sufficiently separated, we get still an artifact from the presence of v_s in the denominator of the second term in Eq. (25). It leads to a reduction of the subtraction and, consequently, larger values of $C(S_F, S_B \bullet S_C)$ in Fig. 3(a) (in the region where F, B, and C are separated) than in the BT model shown in Fig. 2(a).

Similar distortions are seen in the case of the $L + R$ reference bin, displayed in Fig. 3(b). We note that $C(S_F, S_B \bullet S_L + S_R)$ is "pulled down" at the peripheries and too large in the region where F, B, L, and R are well separated.

The conclusion of the above discussion is that the short-range components must be separated at the level of $\overline{C}(N_F, N_B)$ for the presented analysis to make practical sense.

A well-known method of reducing correlations from the resonance decays is to use particles of the same charge. We apply our procedures for π^+, as there are no resonances that decay into $\pi^+\pi^+$ pairs. Only some remnant correlation is expected from resonance decays proceeding in cascades. The result is shown in Fig. 4, with the central reference bin in panel (a) and the sum of the peripheral bins $L + R$ in panel (b). We note a very weak correlation from cascade decays, visible along the $y_1 = y_2$ diagonal. In the region away from the diagonal we see a very close agreement with Fig. 2. Therefore the use of hadrons of the same sign is an efficient way of getting rid of the short-range correlations from the resonance decays.

In Fig. 5 we show the partial ρ correlation of Eq. (6), with the central reference bin. We recall that in the BT model it is given by Eq. (22), visualized in panel (a). As expected, this shape is reproduced within statistical noise by the primordial particles of the simulated data. We remark that the ρ correlation, being the ratio of the covariance and variance, carries less information than the C correlation. From the C correlation we can read off both the scaled covariance and variance (along the diagonal).

Finally, we test the partial correlation approach for the case of the two peripheral reference bins discussed in Appendix A, with $-6.1 < \eta_L < -5.1$ and $5.1 < \eta_R < 6.1$. We recall that in the BT model the corresponding partial correlation vanishes, cf. Eq. (24). This is also the case, within numerical uncertainties, for the simulated data, as visualized in Fig. 6, for the case of primordial particles (a) and all positively charged pions (b).

VII. CONCLUSIONS

We have presented a simple method capable of providing information on the initial two-particle multiplicity correlations, which is insensitive to centrality fluctuations. The basic formalism relies on the concept of partial covariance, where a reference bin (or a few reference bins) are used to impose constraints on the sample. Application of the method to a superposition model, where particle production occurs in subsequent stages, allows us to unfold the trivial statistical fluctuations from the hadronization at freeze-out and gain insight into the correlations in the initial stage of the reaction.

We have demonstrated the feasibility of the method by carrying out an illustrative analysis on simulated data obtained with hydrodynamics, run event by event on initial conditions provided by the wounded quark model, and followed by statistical hadronization. We have shown that performing the calculations for the partial correlations of hadrons which do not

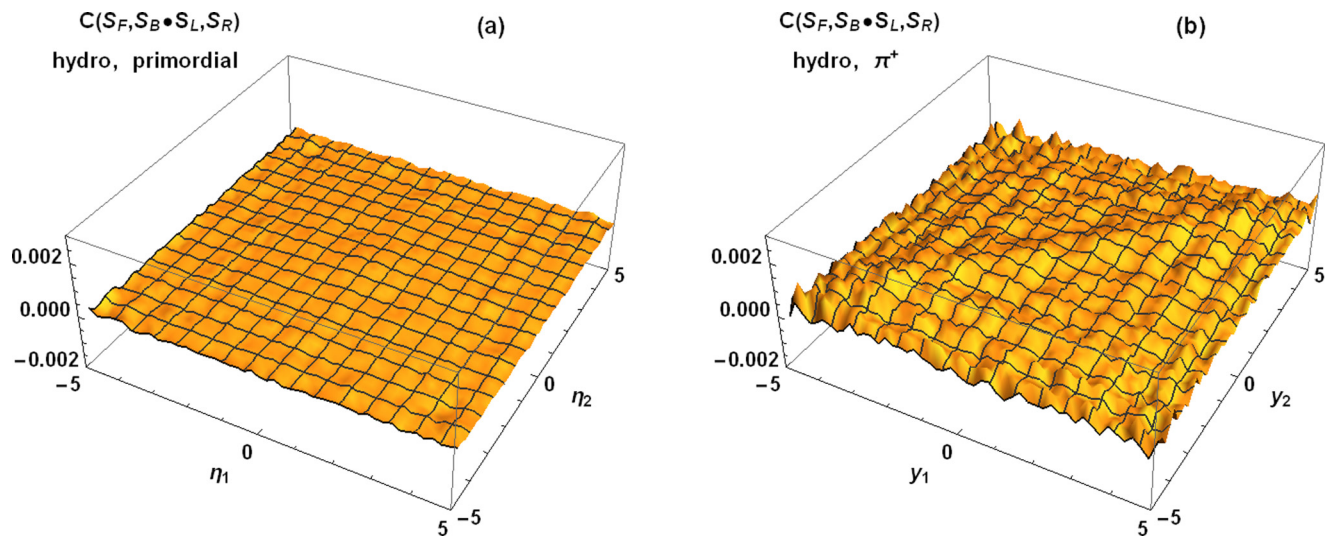

FIG. 6. Partial C-correlation function for the source multiplicities with two peripheral reference bins, obtained for the simulated data with (a) primordial hadrons and (b) all positively charged pions, π^+.

carry correlations from resonance decays (primordial hadrons, or pions of the same charge) reproduces efficiently the initial partial correlations. Thus the method can be used as a practical tool in experimental data analysis of two-particle correlations.

A nontrivial aspect of our approach is a simple way of unfolding trivial statistical fluctuations generated at statistical hadronization. In the superposition model, it amounts to removal of autocorrelations from the building blocks of the partial correlation function. Moreover, that way we are able to impose constraints on the number of sources in the reference bin, rather than the number of hadrons, which is desirable from the point of view of studying the initial state.

The method is directly extendable to imposition of more constraints, related to a possible involvement of more detectors. This allows for getting more information of the correlations generated in the initial state.

A crucial element of the correlation analyses is the separation of the short-range component, expected to be generated in later stages of the collision, and the long-range component, generated in the initial phase. We have demonstrated on the simulated data that the use of same-charge pions largely reduces the correlations due to resonances. Other sources of short-range correlations (jets, femtoscopy) should be removed with a suitable method.

To summarize, the procedure of obtaining the partial correlations in the initial state is as follows:

(i) Obtain the two-particle correlation function in (pseudo)rapidity from the data.
(ii) Remove autocorrelations and the short-distance component.
(iii) Apply Eq. (15) or its generalizations in the case of more control bins.
(iv) Up to corrections from bin mixing and spatial rapidity-pseudorapidity mapping, the result represents the correlations in the initial state of the collision with centrality fluctuations removed at the level of sources.

Finally, we note that the technique of partial correlations is applicable to other observables which correlate to the "centrality" determination, for instance, various charges or the transverse momentum.

ACKNOWLEDGMENTS

We are grateful to Piotr Bożek for providing us a sample of the hydrodynamic simulation results for the wounded quark model used for illustration of the data analysis methods discussed in this paper. This research was supported by the Polish National Science Centre Grant No. 2015/19/B/ST2/00937.

APPENDIX A: PARTIAL COVARIANCE WITH MORE CONTROL VARIABLES

In this Appendix we list some basic definitions and properties referring to the partial correlations. In a general case we have n physical random variables $X = (X_1, \ldots, X_n)$ and m control random variables $Z = (Z_1, \ldots, Z_m)$. The quantities X_i and Z_j are vectors in the $N_{\rm ev}$-dimensional space, where $N_{\rm ev}$ (the number of events) is the number of the data points. Averaging over events for a quantity A is defined as $\langle A\rangle = 1/N_{\rm ev}\sum_{k=1}^{N_{\rm ev}} A_i$. One defines the partial covariance matrix as

$$\Sigma_{XX\bullet Z} = \Sigma_{XX} - \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{ZX}, \tag{A1}$$

where Σ_{AB} is the standard covariance matrix defined as

$$\begin{aligned}(\Sigma_{AB})_{ij} &= \langle (A_i - \langle A_i\rangle)(B_j - \langle B_j\rangle)\rangle \\ &= \langle A_i B_j\rangle - \langle A_i\rangle\langle B_j\rangle,\ A,B = X,Z,\end{aligned} \tag{A2}$$

where i and j label the variable types. Mathematically, Eq. (A1) corresponds to projecting out from the vectors X the components belonging to the space spanned by the vectors Z (shifted to their central values). Indeed, introducing the

projected physical vectors

$$\overline{X}_i = X_i - (\Sigma_{XZ})_{ij}\left(\Sigma_{ZZ}^{-1}\right)_{jj'}(Z_{j'} - \langle Z_{jj'}\rangle), \tag{A3}$$

which by construction gives the orthogonality condition

$$\langle Z_m \overline{X}_i\rangle - \langle Z_m\rangle\langle \overline{X}_i\rangle = 0, \tag{A4}$$

we straightforwardly verify Eq. (A1).

The above formulas simplify when the control vectors $Z_j - \langle Z_j\rangle$ are orthonormal (which we can always achieve via the Gramm-Schmidt procedure or by diagonalization). Denoting these orthonormal vectors as $U_j \equiv Z_j - \langle Z_j\rangle$, $j = 1,\ldots,m$, we have

$$\overline{X}_i = X_i - (\Sigma_{XU})_{ij}U_j \tag{A5}$$

and

$$(\Sigma_{XX_\bullet Z})_{ii'} = (\Sigma_{XX})_{ii'} - \sum_{j=1}^{m}(\Sigma_{XU})_{ij}(\Sigma_{UX})_{ji'}. \tag{A6}$$

From construction, the diagonal terms are non-negative, $(\Sigma_{XX_\bullet Z})_{ii} \geqslant 0$, where the equality occurs when $X_i - \langle X_i\rangle$ is contained in the space spanned by the vectors U_k.

The interpretation of Eq. (A6) is straightforward: we subtract from the covariance of the physical variables the covariance proceeding via a correlation, one by one, to the control variables U_j. The meaning of the general case (A1) is the same, with the complication arising from the nonorthonormality of the control variables. Some further mathematical facts, in particular the connection to the *linear regression* analysis, may be found in Ref. [98].

Throughout the paper we use the shorthand notation

$$c(X_i, X_i \bullet Z) \equiv (\Sigma_{XX_\bullet Z})_{ii'},\ c(X_i, X_i) \equiv (\Sigma_{XX})_{ii'}. \tag{A7}$$

The partial covariance matrix scaled with the multiplicities is defined as

$$C(X_i, X_j \bullet Z) = \frac{c(X_i, X_j \bullet Z)}{\langle X_i\rangle\langle X_j\rangle}. \tag{A8}$$

The Pearson-like partial correlation coefficient between X_i and X_j is defined as

$$\rho(X_i, X_j \bullet Z) = \frac{c(X_i, X_j \bullet Z)}{\sqrt{c(X_i, X_i \bullet Z)c(X_j, X_j \bullet Z)}}, \tag{A9}$$

which makes sense as long as $c(X_i, X_j \bullet Z) > 0$. From the Schwartz inequality $-1 \leqslant \rho(X_i, X_j \bullet Z) \leqslant 1$, and for the diagonal terms $\rho(X_i, X_i \bullet Z) = 1$, as in the case of the standard Pearson's correlation $\rho(X_i, X_i)$.

For the simplest case of a single control variable, the above formulas reduce to Eqs. (2) and (6). For the special case of two physical and two control variables, explored in this paper, we have

$$\begin{aligned}&\mathrm{c}(X, Y \bullet Z_1, Z_2)\\ &= \mathrm{c}(X,Y) - \frac{\mathrm{c}(X,Z_1)\mathrm{v}(Z_1)\mathrm{c}(Z_1,Y) + \mathrm{c}(X,Z_2)\mathrm{v}(Z_2)\mathrm{c}(Z_2,Y) - \mathrm{c}(Z_1,Z_2)[\mathrm{c}(X,Z_1)\mathrm{c}(Z_2,Y) + \mathrm{c}(X,Z_2)\mathrm{c}(Z_1,Y)]}{\mathrm{v}(Z_1)\mathrm{v}(Z_2) - \mathrm{c}(Z_1,Z_2)^2}.\end{aligned} \tag{A10}$$

For the case of the scaled covariance Eq. (A10) becomes

$$\begin{aligned}&\mathrm{C}(X, Y \bullet Z_1, Z_2)\\ &= \mathrm{C}(X,Y) - \frac{\mathrm{C}(X,Z_1)V(Z_1)\mathrm{C}(Z_1,Y)\frac{\langle Z_1\rangle^2}{\langle Z_2\rangle^2} + \mathrm{C}(X,Z_2)V(Z_2)\mathrm{C}(Z_2,Y)\frac{\langle Z_2\rangle^2}{\langle Z_1\rangle^2} - \mathrm{C}(Z_1,Z_2)[\mathrm{C}(X,Z_1)\mathrm{C}(Z_2,Y) + \mathrm{C}(X,Z_2)\mathrm{C}(Z_1,Y)]}{V(Z_1)V(Z_2) + \mathrm{C}(Z_1,Z_2)^2},\end{aligned} \tag{A11}$$

where $V(Z_j) = \mathrm{v}(Z_j)/\langle Z_j\rangle^2 = C(Z_j,Z_j)/\langle Z_j\rangle^2$.

For the partial correlation coefficient the explicit formula reads

$$\begin{aligned}\rho(X, Y \bullet Z_1, Z_2) &= \frac{A(X,Y)}{\sqrt{A(X,X)A(Y,Y)}},\\ A(X,Y) &= [1 - \rho(Z_1,Z_2)^2]\rho(X,Y) + \rho(Z_1,Z_2)[\rho(Z_2,X)\rho(Z_1,Y) + \rho(Z_1,X)\rho(Z_2,Y)]\\ &\quad - \frac{\mathrm{v}(Z_1)}{\mathrm{v}(Z_2)}\rho(Z_1,X)\rho(Z_1,Y) - \frac{\mathrm{v}(Z_2)}{\mathrm{v}(Z_1)}\rho(Z_2,X)\rho(Z_2,Y).\end{aligned} \tag{A12}$$

APPENDIX B: RELATION OF THE PARTIAL COVARIANCE TO THE CONDITIONAL COVARIANCE

Lawrance [6] has shown that if a sample satisfies the affine condition

$$E(X|Y) = \alpha + BY, \tag{B1}$$

with α a constant and B a constant matrix, then the equality of the partial covariance and the conditional covariance follows,

$$\Sigma_{XX_\bullet Y} = \Sigma_{XX|Y}. \tag{B2}$$

The converse was shown by Baba, Shibata, and Sibuya [7], and hence conditions (B1) and (B2) are equivalent. The question whether (B1) holds can be tested on the actual data. In the context of ultrarelativistic heavy-ion collisions, it should hold for sufficiently narrow centrality classes [18]. Also, condition (B1) holds for normal distributions [7].

The important practical implication of Eq. (B2) is that the partial covariance method is a simple way of imposing constraints in the correlation analysis.

APPENDIX C: RELATION TO OTHER METHODS

One may relate the partial covariance method to the PCA [16]. Suppose all bins are the *measurement* bins X, obtained, for instance, by dividing the full pseudorapidity acceptance of the detector info narrower bins. Let U_j denote the eigenmodes of the covariance matrix Σ_{XX}. Then one may project out the eigenmodes (with the highest eigenvalues) from the covariance matrix, according to Eq. (A6). Thus, PCA may be viewed as a special case of the partial covariance method, where the constraints have the form of the eigenmodes of the covariance matrix.

Whereas the algebra (projection) in the two methods is the same, the accents are different. In PCA one does not separate the measurement and reference variables, treating all the bins "democratically," and there is a focus on a possible hierarchy in the eigenvalues (large gaps in the spectrum are linked to strong correlations, or collectivity, in the fluctuations). In the partial covariance method the measured variables are grouped from the outset into the *physical* bins X and the *reference* bins Y. This is more natural when the reference data come from different detectors (e.g., in ultrarelativistic heavy-ion collisions, from the peripheral detectors, transverse energy colorimeters, etc.) than the detector collecting the physical data (central TPC). In this case the constraints come entirely from the reference variables, and the constraint vectors U_j are built from the vectors Y (see Appendix A).

In PCA applied to pseudorapidity fluctuations, the multiplicity eigenmode with a highest eigenvalue, denoted in Ref. [16] as $v_0^{(1)}$, corresponds to an η-independent fluctuation, i.e., where multiplicities in all the bins vary by an equal amount. This may be interpreted as *centrality* fluctuation, and hence the removal of the highest eigenvalue mode in PCA is equivalent to getting rid of the centrality fluctuations, with centrality defined via the total multiplicity from all the bins.

The issue of separating centrality fluctuations is also focal the methods aiming at *strongly intensive* fluctuation measures [14,15,99,100]. The framework applied there is also based on the superposition model, but there is one common source (with its multiplicity traditionally termed as the "volume") and two types of particles, A and B, emitted from the source with multiplicities m_A and m_B, respectively. One may then use the superposition approach to relate the statistical moments of m_A and m_B to the corresponding moments of the observed multiplicities of A and B, N_A and N_B, in such a way the fluctuation of the sources cancels out. Thus the object of the study is the emission process from the source, and not the fluctuation of the number of sources (volume), considered trivial.

With the method described in this paper we study the case where there are multiple *types* of sources (each in a given rapidity bin) and (for simplicity) just one type of hadrons (an extension to more types of hadrons is possible). Our objective is the fluctuation of the number of sources (at various rapidities), whereas the fluctuations of the overlaid variable m are considered not interesting. Thus, the partial covariance method used with the superposition approach in this work is, in a sense, complementary to the studies based on the strongly intensive fluctuation measures.

APPENDIX D: THE A_{NM} COEFFICIENTS

In this Appendix we discuss the partial covariance for the case, where the $\overline{C}$-correlation function for hadrons (with autocorrelations removed) is expressed via the expansion [12]

$$\overline{C}(\eta_1,\eta_2) = \sum_{m,n=0}^{\infty} a_{nm} T_n\left(\frac{\eta_1}{Y}\right) T_m\left(\frac{\eta_2}{Y}\right), \tag{D1}$$

where $T_n(x)$ is a set of orthonormal functions and $[-Y,Y]$ is the pseudorapidity domain. The choice of Refs. [30,97,101] is

$$T_n(x) = \sqrt{2+1/2}\,P_n(x), \tag{D2}$$

where $P_n(x)$ denote the Legendre polynomials. The orthonormality condition $\int_{-1}^{1} dx T_n(x) T_m(x) = \delta_{nm}$ is satisfied. The a_{nm} coefficients are

$$a_{nm} = \int_{-Y}^{Y} \frac{d\eta_1}{Y} \int_{-Y}^{Y} \frac{d\eta_2}{Y} \overline{C}(\eta_1,\eta_2) T_n\left(\frac{\eta_1}{Y}\right) T_m\left(\frac{\eta_2}{Y}\right). \tag{D3}$$

The transformation to the partial C-correlation function of Eq. (15) leads, in general, to mixing of the a_{nm} coefficients, i.e., the coefficients a^C_{nm} for $C(S_F, S_B \bullet S_C)$ become complicated functions of the a_{nm} coefficients for $\overline{C}(N_F, N_B)$.

We can, however, derive a simple formula connecting these coefficients, introducing the expansion (for a fixed η_C of the reference bin)

$$\overline{C}(\eta_1,\eta_C) = \sum_{m}^{\infty} a_m(\eta_C) T_m\left(\frac{\eta_2}{Y}\right). \tag{D4}$$

Then

$$a^C_{nm} = a_{nm} - \frac{a_n(\eta_C) a_m(\eta_C)}{\overline{\nu}(\eta_C)} \tag{D5}$$

[up to the rescaling effects of Eq. (10), not included explicitly in the present discussion].

We remark that for the special case of the BT model with the central reference bin, $a^C_{11} = a_{11}$.

[1] H. Cramer, *Mathematical Methods of Statistics*, Princeton Mathematical Series, no. 9 (Princeton University Press, Princeton, NJ, 1946).

[2] B. Abelev *et al.* (STAR Collaboration), Phys. Rev. Lett. **103**, 172301 (2009).

[3] S. De, T. Tarnowsky, T. K. Nayak, R. P. Scharenberg, and B. K. Srivastava, Phys. Rev. C **88**, 044903 (2013).

[4] W. Krzanowski, *Principles of Multivariate Analysis*, Oxford Statistical Science Series (Oxford University Press, Oxford, 2000).

[5] L. J. Frasinski, J. Phys. B **49**, 152004 (2016).
[6] A. J. Lawrance, Am. Stat. **30**, 146 (1976).
[7] K. Baba, R. Shibata, and M. Sibuya, Austr. N. Z. J. Stat. **46**, 657 (2004).
[8] A. Olszewski and W. Broniowski, Phys. Rev. C **88**, 044913 (2013).
[9] P. Bożek, Phys. Rev. C **85**, 034901 (2012).
[10] A. Kisiel, T. Taluć, W. Broniowski, and W. Florkowski, Comput. Phys. Commun. **174**, 669 (2006).
[11] M. Chojnacki, A. Kisiel, W. Florkowski, and W. Broniowski, Comput. Phys. Commun. **183**, 746 (2012).
[12] A. Bzdak and D. Teaney, Phys. Rev. C **87**, 024906 (2013).
[13] P. Bożek, W. Broniowski, and M. Rybczyński, Phys. Rev. C **94**, 014902 (2016).
[14] M. Gaździcki and S. Mrówczyński, Z. Phys. C **54**, 127 (1992).
[15] M. I. Gorenstein and M. Gaździcki, Phys. Rev. C **84**, 014904 (2011).
[16] R. S. Bhalerao, J.-Y. Ollitrault, S. Pal, and D. Teaney, Phys. Rev. Lett. **114**, 152301 (2015).
[17] T. Lappi and L. McLerran, Nucl. Phys. **A832**, 330 (2010).
[18] A. Bzdak, Phys. Rev. C **85**, 051901 (2012).
[19] S. Uhlig, I. Derado, R. Meinke, and H. Preissner, Nucl. Phys. **B132**, 15 (1978).
[20] K. Alpgard *et al.* (UA5), Phys. Lett. **B123**, 361 (1983).
[21] G. J. Alner *et al.* (UA5), Phys. Rept. **154**, 247 (1987).
[22] V. V. Aivazian *et al.* (NA22), Z. Phys. C **42**, 533 (1989).
[23] R. E. Ansorge *et al.* (UA5), Z. Phys. C **37**, 191 (1988).
[24] I. Derado *et al.*, Z. Phys. C **40**, 25 (1988).
[25] T. Alexopoulos *et al.* (E735 Collaboration), Phys. Lett. **B353**, 155 (1995).
[26] J. Bächler *et al.* (NA35), Z. Phys.C **56**, 347 (1992).
[27] Y. Akiba *et al.* (E802), Phys. Rev. C **56**, 1544 (1997).
[28] B. B. Back *et al.* (PHOBOS), Phys. Rev. C **74**, 011901 (2006).
[29] G. Aad *et al.* (ATLAS), J. High Energy Phys. 07 (2012) 019.
[30] J. Jia, S. Radhakrishnan, and M. Zhou, Phys. Rev. C **93**, 044905 (2016).
[31] J. Adam *et al.* (ALICE), J. High Energy Phys. 05 (2015) 097.
[32] J. Jia (ATLAS), Nucl. Phys. **A956**, 405 (2016).
[33] M. Aaboud *et al.* (ATLAS), Phys. Rev. C **95**, 064914 (2017).
[34] I. Sputowska, Correlations in Particle Production in Nuclear Collisions at LHC Energies, Ph.D. thesis, Institute of Nuclear Physics PAN, Cracow, Poland, 2016.
[35] J. Adam *et al.* (ALICE), Eur. Phys. J.C **77**, 569 (2017).
[36] A. Capella and A. Krzywicki, Phys. Rev. **D18**, 4120 (1978).
[37] A. B. Kaidalov and K. A. Ter-Martirosian, Phys. Lett. **B117**, 247 (1982).
[38] T. T. Chou and C. N. Yang, Phys. Lett. **B135**, 175 (1984).
[39] A. Capella, U. Sukhatme, C.-I. Tan, and J. Tran Thanh Van, Phys. Rep. **236**, 225 (1994).
[40] N. S. Amelin, N. Armesto, M. A. Braun, E. G. Ferreiro, and C. Pajares, Phys. Rev. Lett. **73**, 2813 (1994).
[41] M. A. Braun, C. Pajares, and V. V. Vechernin, Phys. Lett. **B493**, 54 (2000).
[42] A. Giovannini and R. Ugoccioni, Phys. Rev. D **66**, 034001 (2002).
[43] M. A. Braun, R. S. Kolevatov, C. Pajares, and V. V. Vechernin, Eur. Phys. J. C **32**, 535 (2004).
[44] S. Haussler, M. Abdel-Aziz, and M. Bleicher, Nucl. Phys. **A785**, 253 (2007).
[45] P. Brogueira and J. Dias de Deus, Phys. Lett. **B653**, 202 (2007).
[46] N. Armesto, L. McLerran, and C. Pajares, Nucl. Phys. **A781**, 201 (2007).
[47] V. V. Vechernin and R. S. Kolevatov, Phys. Atom. Nucl. **70**, 1797 (2007) [Yad. Fiz. 70, 1846 (2007)].
[48] M. A. Braun, Nucl. Phys. **A806**, 230 (2008).
[49] V. P. Konchakovski, M. Hauer, G. Torrieri, M. I. Gorenstein, and E. L. Bratkovskaya, Phys. Rev. C **79**, 034910 (2009).
[50] A. Bzdak and K. Woźniak, Phys. Rev. C **81**, 034908 (2010).
[51] A. Bzdak, Phys. Rev. C **80**, 024906 (2009).
[52] P. Bożek, W. Broniowski, and J. Moreira, Phys. Rev. C **83**, 034911 (2011).
[53] Y.-L. Yan, D.-M. Zhou, B.-G. Dong, X.-M. Li, H.-L. Ma, and B.-H. Sa, Phys. Rev. C **81**, 044914 (2010).
[54] J. Dias de Deus and C. Pajares, Phys. Lett. **B695**, 211 (2011).
[55] A. Białas and K. Zalewski, Phys. Lett. **B698**, 416 (2011).
[56] V. V. Vechernin, Nucl. Phys. A **939**, 21 (2015).
[57] G.-L. Ma and A. Bzdak, Phys. Lett. **B739**, 209 (2014).
[58] A. Bzdak and P. Bożek, Phys. Rev. C **93**, 024903 (2016).
[59] A. Monnai and B. Schenke, Phys. Lett. **B752**, 317 (2016).
[60] W. Broniowski and P. Bożek, Phys. Rev. C **93**, 064910 (2016).
[61] R. He, J. Qian, and L. Huo, Phys. Rev. C **94**, 034902 (2016).
[62] R. He, J. Qian, and L. Huo, Phys. Rev. C **93**, 044918 (2016).
[63] A. Olszewski and W. Broniowski, Acta Phys. Polon. **B48**, 113 (2017).
[64] A. Olszewski and W. Broniowski, Phys. Rev. C **92**, 024913 (2015).
[65] A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. **D52**, 6231 (1995).
[66] E. Iancu, A. Leonidov, and L. D. McLerran, Nucl. Phys. **A692**, 583 (2001).
[67] R. Glauber, *High Energy Collision Theory*, Vol. 1 in Lectures in Theoretical Physics (Interscience, New York, 1959).
[68] W. Czyż and L. C. Maximon, Ann. Phys. **52**, 59 (1969).
[69] A. Białas, M. Błeszyński, and W. Czyż, Nucl. Phys. B **111**, 461 (1976).
[70] D. Kharzeev and M. Nardi, Phys. Lett. B **507**, 121 (2001).
[71] B. B. Back *et al.* (PHOBOS), Phys. Rev. C **65**, 031901 (2002).
[72] W. Broniowski, P. Bożek, and M. Rybczyński, Phys. Rev. C **76**, 054905 (2007).
[73] U. Heinz and R. Snellings, Ann. Rev. Nucl. Part. Sci. **63**, 123 (2013).
[74] C. Gale, S. Jeon, and B. Schenke, Int. J. Mod. Phys. A **28**, 1340011 (2013).
[75] Z.-W. Lin, C. M. Ko, B.-A. Li, B. Zhang, and S. Pal, Phys. Rev. C **72**, 064901 (2005).
[76] J. I. Kapusta, B. Muller, and M. Stephanov, Phys. Rev. C **85**, 054906 (2012).
[77] L. Yan and H. Gronqvist, J. High Energy Phys. 03 (2016) 121.
[78] S. Gavin, G. Moschelli, and C. Zin, Phys. Rev. C **95**, 064901 (2017).
[79] F. Cooper and G. Frye, Phys. Rev. D **10**, 186 (1974).
[80] Fu Jinghua and Liu Lianshou, Phys. Rev. C **68**, 064904 (2003).
[81] A. Białas, Acta Phys. Polon. **B35**, 683 (2004).
[82] B. Ling and M. A. Stephanov, Phys. Rev. C **93**, 034915 (2016).
[83] T. Nonaka, M. Kitazawa, and S. I. Esumi, Phys. Rev. C **95**, 064912 (2017).
[84] A. Białas, W. Czyż, and W. Furmański, Acta Phys. Polon. **B8**, 585 (1977).
[85] V. V. Anisovich, Yu. M. Shabelski, and V. M. Shekhter, Nucl. Phys. B **133**, 477 (1978).

[86] S. Eremin and S. Voloshin, Phys. Rev. C **67**, 064905 (2003).
[87] P. K. Netrakanti and B. Mohanty, Phys. Rev. C **70**, 027901 (2004).
[88] S. S. Adler *et al.* (PHENIX), Phys. Rev. C **89**, 044905 (2014).
[89] A. Adare *et al.* (PHENIX), Phys. Rev. C **93**, 024901 (2016).
[90] R. A. Lacey, P. Liu, N. Magdy, M. Csanád, B. Schweid, N. N. Ajitanand, J. Alexander, and R. Pak, arXiv:1601.06001 [nucl-ex].
[91] W. Broniowski, M. Rybczyński, and P. Bożek, Comput. Phys. Commun. **180**, 69 (2009).
[92] M. Rybczyński, G. Stefanek, W. Broniowski, and P. Bożek, Comput. Phys. Commun. **185**, 1759 (2014).
[93] A. Białas and W. Czyż, Acta Phys. Polon. **B36**, 905 (2005).
[94] A. Adil, M. Gyulassy, and T. Hirano, Phys. Rev. D **73**, 074006 (2006).
[95] P. Bożek and I. Wyskiel, Phys. Rev. C **81**, 054902 (2010).
[96] P. Bożek, Phys. Rev. C **81**, 034909 (2010).
[97] G. Aad *et al.* (ATLAS Collaboration) (2015), ATLAS-CONF-2015-051.
[98] Wikipedia, Partial correlation. Wikipedia, The Free Encyclopedia (2017) [online; accessed 23-August-2017], https://en.wikipedia.org/wiki/Partial_correlation.
[99] E. Sangaline (2015), arXiv:1505.00261 [nucl-th].
[100] W. Broniowski and A. Olszewski, Phys. Rev. C **95**, 064910 (2017).
[101] G. Aad *et al.* (ATLAS Collaboration) (2015), ATLAS-CONF-2015-020.