HAL
open science

Relating centrality to impact parameter in nucleus-nucleus collisions

Sruthy Jyothi Das, Giuliano Giacalone, Pierre-Amaury Monard, Jean-Yves Ollitrault

To cite this version:

Sruthy Jyothi Das, Giuliano Giacalone, Pierre-Amaury Monard, Jean-Yves Ollitrault. Relating centrality to impact parameter in nucleus-nucleus collisions. 2017. cea-01583430

HAL Id: cea-01583430

https://cea.hal.science/cea-01583430v1

Preprint submitted on 7 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Relating centrality to impact parameter in nucleus-nucleus collisions

Sruthy Jyothi Das,[1,2] Giuliano Giacalone,[3] Pierre-Amaury Monard,[1] and Jean-Yves Ollitrault[1]

[1] *Institut de physique théorique, Université Paris Saclay, CNRS, CEA, 91191 Gif-sur-Yvette, France*

[2] *Indian Institute of Science Education and Research (IISER) Dr. Homi Bhabha Road, Pashan, Pune 411008, India*

[3] *Centre de physique théorique, École polytechnique, Université Paris-Saclay, 91128 Palaiseau, France*

(Dated: August 2, 2017)

In ultrarelativistic heavy-ion experiments, one estimates the centrality of a collision using a single observable, say n, typically given by the transverse energy or the number of tracks observed in a dedicated detector. The correlation between n and the impact parameter, b, of the collision is then inferred by fitting a specific model of the collision dynamics, such as the Glauber model, to experimental data. The goal of this paper is to assess precisely which information about b can be extracted from data without any specific model of the collision. Under the sole assumption that the probability distribution of n is Gaussian for a fixed b, we show that the probability distribution of impact parameter in a narrow centrality bin can be accurately reconstructed up to 5% centrality. We apply our methodology to Relativistic Heavy Ion Collider (RHIC) and Large Hadron Collider (LHC) data. We propose a simple measure of the precision of the centrality determination, which can be used to compare different experiments.

arXiv:1708.00081v1 [nucl-th] 31 Jul 2017

I. INTRODUCTION

The impact parameter, b, of an ultrarelativistic nucleus-nucleus collision is a crucial quantity. It determines the size and transverse shape of the quark-gluon matter formed in the collision. Central collisions, at small b, yield large and round interaction regions, while peripheral collisions, characterized by large values of impact parameter, yield smaller interaction regions with a pronounced elliptical anisotropy. The centrality dependence of various observables provides, then, insight into their dependence on the global geometry. The energy loss of high-momentum particles [1, 2] or jets [3] is larger in central collisions, as it increases with the length of the path traversed by the particles inside the quark-gluon plasma. By contrast, elliptic flow [4, 5] originates from the elliptical shape of the nuclear overlap region, and is larger in peripheral events [6].

The impact parameter of a single collision, even though it is a perfectly well-defined quantity at ultrarelativistic energies (in the sense that the quantum uncertainty is negligible), is not directly measurable. In experiments, impact parameter is estimated using a single observable, which we denote generically by n [7]. Depending on the experiment, n is either the number of particles (multiplicity) in a given detector [8–11] or the transverse energy deposited in a calorimeter [12, 13]. The idea is that collisions with a small impact parameter produce on average larger values of n. However, the relation between n and b is not one-to-one, and the variation of n with b is not known a priori. This relation is usually inferred from a microscopic model of the collision, such as HIJING [14], or a two-component Glauber model [15] coupled with a simple model of particle production. The parameters of these models are tuned to reproduce the observed probability distribution of n. While these models offer a convenient parameterization, they may not describe the actual dynamics of a collision. This is suggested by the fact that different sets of parameters must be used for different colliding systems, and by the observation that the two-component Glauber model is disfavored by analyses of U+U collisions [16, 17].

The goal of this article is to assess which information about the actual values of impact parameter can be extracted from the measured distributions of n, with as little theoretical bias as possible. In particular, as we shall see in the following sections, we do not need to introduce the concept of "participant nucleon", which is a key ingredient of many microscopic models, but not a measurable quantity.

The term *centrality* originally refers to a classification according to impact parameter. In nowadays experiments, however, it refers to the classification of the collisions in terms of the parameter n. To avoid confusion, we call *b-centrality* the centrality determined with respect to impact parameter. The corresponding definitions are recalled in Sec. II. In Sec. III, we show that a correspondence between n and b can be drawn under the sole assumption that fluctuations of n for a given impact parameter are Gaussian. This Gaussian is characterized by a mean $\bar{n}$ and a width σ, which both depend on impact parameter. We test the validity of this assumption in Sec. III using model calculations. We argue that data allow to reconstruct unambiguously the full impact parameter dependence of the mean $\bar{n}$, and the value of the width σ for central collisions, and we explain how this can be done in practice. In Sec. IV, we validate the proposed procedure using model calculations, where the impact parameter is known. We show that the fluctuations of impact parameter at a fixed centrality can be unambiguously reconstructed, and we apply this method to experimental data in Sec. V.

II. TWO DEFINITIONS OF CENTRALITY

Collisions can be classified according to their impact parameter, b. We define the centrality, c_b, as the cumulative probability distribution of b:

$$c_b \equiv \frac{1}{\sigma_{\rm inel}} \int_0^b P_{\rm inel}(b') 2\pi b' db', \qquad (1)$$

where $\sigma_{\rm inel}$ is the inelastic nucleus-nucleus cross section and $P_{\rm inel}(b)$ is the probability that an inelastic collision occurs at impact parameter b. We name c_b the *b-centrality* of the collision, in order to distinguish it from the usual centrality defined in heavy-ion experiments, to be discussed below. The probability distribution of c_b is flat by construction: $P(c_b) = 1$ for $0 < c_b < 1$. Neither b nor c_b can be measured experimentally. They are known only in model calculations[1].

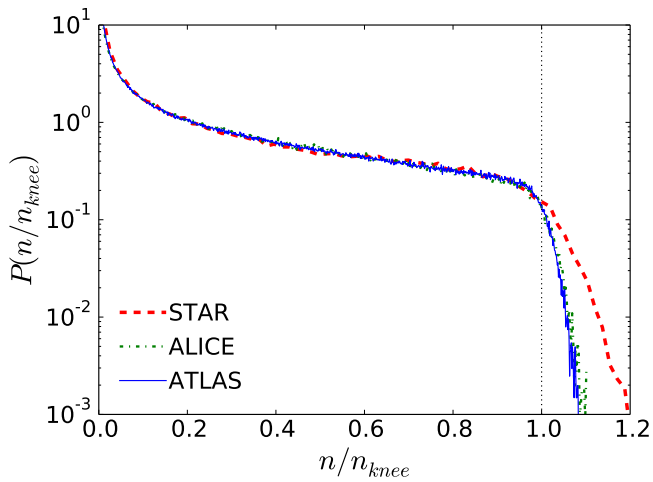

FIG. 1. (Color online) Histograms of the probability distribution of n measured by different experiments. Dashed line: STAR data on Au+Au collisions at $\sqrt{s} = 130$ GeV. Dot-dashed line: ALICE data on Pb+Pb collisions at $\sqrt{s} = 2.76$ TeV. Solid line: ATLAS data on Pb+Pb collisions at $\sqrt{s} = 2.76$ TeV. The horizontal axis of each histogram has been rescaled by the value of n at the *knee* (see text).

In experiments, collisions are instead classified according to a single observable, n. The STAR Collaboration [16] defines n as the number of tracks of charged particles detected in the pseudorapidity window $-1 < \eta < 1$. The ALICE Collaboration [11] uses the number of hits in two scintillators covering the windows $-3.7 < \eta < -1.7$ and $2.8 < \eta < 5.1$. The ATLAS Collaboration [13] uses the energy deposited in two calorimeters in the symmetric windows $-4.9 < \eta < -3.2$ and $3.2 < \eta < 4.9$. Figure 1 displays the probability distribution of n, $P(n)$, measured by these three experiments[2]. Since different detectors have different acceptance and efficiency, and n can refer to a multiplicity or an energy, we rescale the value of n by its value at the *knee*, to be defined precisely in Sec. III. Once rescaled, ATLAS and ALICE data are almost identical. STAR data differ in the tail, which is twice as broad[3].

The cumulative distribution of n defines the experimental measure of centrality, which we denote by c. It is defined by

$$c \equiv \int_n^\infty P(n') dn'. \qquad (2)$$

Note that the centrality classification is in ascending order for b and in descending order for n, which explains the different integration limits in Eqs. (1) and (2). The probability distribution of c is also flat by construction: $P(c) = 1$ for $0 < c < 1$.

We have thus defined two measures of the centrality: c_b and c, depending on whether one sorts events according to b or to n. If the relation between n and b is one-to-one, both measures coincide, $c = c_b$. In practice, one observes a range of value of n at a given value of b. The joint distribution of n and b is usually inferred from a specific model of the collision [14, 15].

III. RELATING CENTRALITY TO *b*-CENTRALITY

Here we simply assume that the probability of n for fixed b is Gaussian [7]:

$$P(n|c_b) = \frac{1}{\sigma(c_b)\sqrt{2\pi}} \exp\left(-\frac{(n - \bar{n}(c_b))^2}{2\sigma(c_b)^2}\right), \qquad (3)$$

where both the mean $\bar{n}$ and the width σ depend on the impact parameter or, equivalently, on c_b. We expect this to be a good approximation because of the central limit theorem: n is a multiplicity, or transverse energy, which gets contributions from many collision processes which are located at different points in the transverse plane, and, therefore, causally disconnected and independent.

It is useful to check this Gaussian approximation on a model[4]. We take n as the total entropy deposited in

[1] The results in this paper use the variable c_b, but one can easily express them in terms of b using the change of variables $c_b = \pi b^2/\sigma_{\rm inel}$. The value of $\sigma_{\rm inel}$ needs to be taken from either data or some collision model.

[2] Data on $P(n)$ collected by the CMS and PHENIX Collaborations are shown in [12] and [10], respectively.

[3] We use uncalibrated 130 GeV STAR data [9] rather than calibrated 200 GeV data [16]. We have checked that, once rescaled, the two distributions are very similar. The advantage of 130 GeV data is that errors are provided, so that we are able to assess the quality of our fits, as we shall see in Sec. V.

[4] We stress, though, that our subsequent analysis of the experimental results does not rely on whether or not the chosen model provides a good description of data.

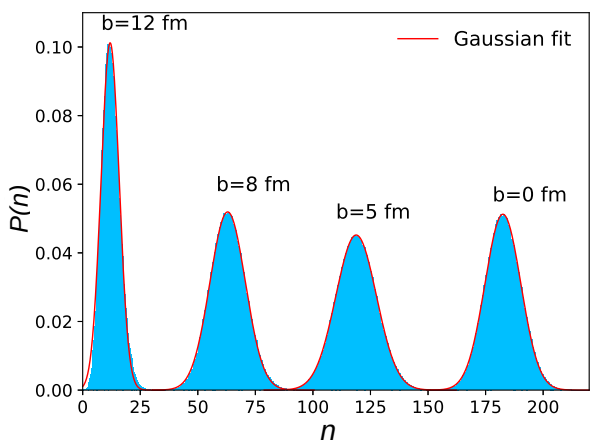

FIG. 2. (Color online) Shaded areas: Histograms of the probability distribution of n for fixed impact parameter in Pb+Pb collisions at $\sqrt{s} = 2.76$ TeV in the TRENTo model [17]. The values $b = 0$, 5, 8, 12 fm were used, which correspond to $c_b = 0$, 10%, 26%, 58%, respectively. We have generated 5×10^5 events for each value of b. Solid lines are Gaussian fits. The quantity n is in arbitrary units.

the transverse plane by the TRENTo model of initial conditions [17] in Pb+Pb collisions at $\sqrt{s} = 2.76$ TeV. We employ TRENTo with reduced thickness parameter $p = 0$, a setup which is known to reproduce accurately a wide range of observables [18, 19], in particular, distributions of multiplicity at the LHC [17]. Figure 2 displays the distribution of total entropy in this setup of TRENTo for different fixed values of impact parameter. This plot shows that the Gaussian approximation is valid in the model up to $\sim 60\%$ b-centrality. Although we do not know whether the breakdown of the Gaussian approximation is a mere default of the model or not, in the following analysis we will exclude the most peripheral collisions. This is also motivated by the fact that fluctuations of n for large b are expected to be large, and large fluctuations are in general non-Gaussian. Moreover, if fluctuations are large, the condition $\bar{n}(c_b) \gg \sigma(c_b)$ is not guaranteed, and integrating Eq. (3) over n from 0 to $+\infty$ may not yield 1.

A crucial quantity which we will use throughout this work is the position of the *knee* of the distribution of n. We define it as the mean value of n at zero impact parameter:

$$n_{\rm knee} \equiv \bar{n}(0). \tag{4}$$

The observed distribution of n, $P(n)$, is eventually obtained by integrating Eq. (3) over c_b, i.e.,

$$P(n) = \int_0^1 P(n|c_b) dc_b. \tag{5}$$

In this paper, we determine smooth functions $\bar{n}(c_b)$ and $\sigma(c_b)$ such that $P(n)$ matches experimental data. This problem is underconstrained, in the sense that one cannot determined two unknown functions $\bar{n}(c_b)$ and $\sigma(c_b)$ from a single function $P(n)$. We shall argue that one can only constrain the mean $\bar{n}(c_b)$, and the value of the width for central collisions, $\sigma(0)$. Since the variation of σ with c_b cannot be determined from data alone, we test two different scenarios:

- (A) $\sigma(c_b) = \sigma(0)\sqrt{\bar{n}(c_b)/\bar{n}(0)}$
- (B) $\sigma(c_b) = \sigma(0)$.

The first scenario, (A), assumes that the variance is proportional to the mean, which would be true if n were the sum of contributions from independent nucleon-nucleon collisions. Scenario (B) is motivated by the observation that the width of the histograms observed in Fig. 2 varies little between $b = 0$ and $b = 8$ fm. We state that this is an artifact of the Monte Carlo model, where particle production is essentially determined by the participant nucleons. Since the number of participant nucleons is bounded by the total number of nucleons, fluctuations of n are consequently reduced by the presence of this upper cutoff. There is, however, no deep theoretical reason to believe that this particular feature of the Monte Carlo models is realistic.

For each scenario, (A) or (B), we need a smooth function $\bar{n}(c_b)$, and a constant, $\sigma(0)$, such that $P(n)$ defined by Eqs. (3) and (5) fit experimental data. We use the following functional form of $\bar{n}(c_b)$, which guarantees positivity:

$$\bar{n}(c_b) = n_{\rm knee} \exp\left(-a_1 c_b - a_2 c_b^2 - a_3 c_b^3\right). \tag{6}$$

One could as well use other functional forms, requiring the final fitting function to be a smooth, positive, monotonically decreasing function of c_b, with no singularities in the interval $0 \le c_b < 1$. We carry out a 5-parameter fit to $P(n)$ using Eqs. (3) and (5), with parameters given by $n_{\rm knee}$, a_1, a_2, a_3 and $\sigma(0)$. In order to eliminate peripheral collisions from the fit, we only use values of n above a cutoff $n_{\rm min}$, which we specify in each case.

IV. VALIDATION OF THE METHOD

We now validate this procedure of relating n to c_b using Monte Carlo simulations, where both b and n are known in each event. We simulate Pb+Pb collisions at $\sqrt{s} = 2.76$ TeV using the same setup of TRENTo as in Fig. 2, and again we use the entropy of each event to construct the probability distribution $P(n)$. We generate 10^7 events. We determine c_b of each event by sorting events according to b, and c by sorting them according to n. Symbols in Fig. 3 correspond to the distribution $P(n)$ from this Monte Carlo calculation.

We now apply the fitting procedure described in the previous section to $P(n)$. For scenario (A), we vary the lower cutoff from $n_{\rm min} = 1.8$ to $n_{\rm min} = 9.1$, corresponding to centralities $c = 80\%$ and $c = 62\%$, and we find that

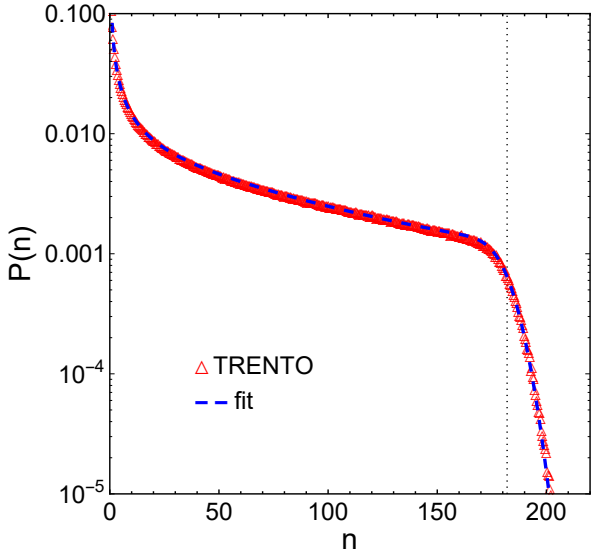

FIG. 3. (Color online) Symbols: $P(n)$ obtained from the TRENTo model of initial conditions. Dashed line: Fit of $P(n)$ using Eqs. (3) and (5) in the scenario (A). The vertical line indicates the position of the knee [see Table I]. The quantity n is in arbitrary units.

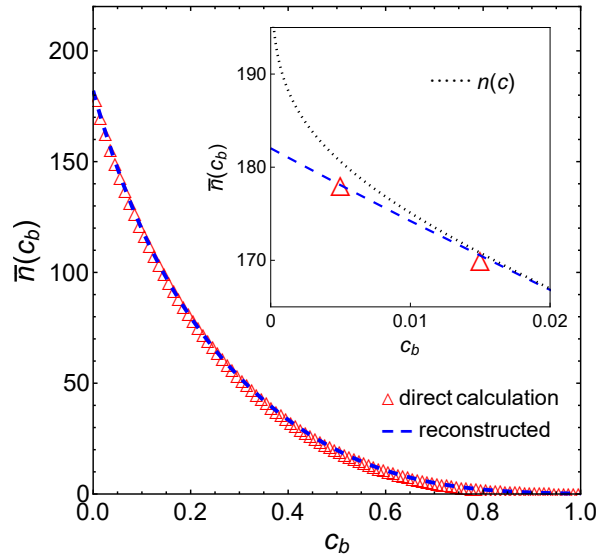

FIG. 4. (Color online) Mean value of n versus b-centrality. Symbols: calculated directly by binning TRENTo results in c_b. Dashed line: reconstructed from the fit of $P(n)$ alone. The inset shows a zoom of the most central collisions, where we compare $\bar{n}(c_b)$ to $n(c)$ (dotted line).

the results of the fit are stable. The solid line in Fig. 3 displays the fit corresponding to the larger value of $n_{\rm min}$. The fit provides an excellent description of $P(n)$ in the model. For case (B), fluctuations are larger and we use a larger cutoff, $n_{\rm min} = 36$. Above the cutoff, the two fits are indistinguishable in Fig. 3. The fit parameters are listed in Table I, for both scenarios. The error on the knee, $n_{\rm knee}$, returned by the fit is only 0.3%, while the error on the width $\sigma(0)$ is about 3%. Interestingly, the two fitting procedures (A) and (B) return essentially the same values of these two quantities.[5]

In Fig. 4, we show $\bar{n}(c_b)$ returned by the fit, i.e., we calculate Eq. (6) using the parameters of Table I for scenario (A). We display the comparison between this analytical estimate and the results directly obtained by binning TRENTo results in c_b, and computing the mean value of n in each bin. Agreement is within 0.5%, all the way up to $c_b = 70\%$, which corresponds to a value of n smaller than the lower cutoff applied to $P(n)$. We also calculate $\sigma(0)$ directly, by generating 10^5 collisions with $b = 0$ fm, and computing the standard deviation of n. The resulting value of $\sigma(0)$ agrees with the corresponding fit parameter.

It is useful to note that approximate values of $\sigma(0)$ and $\bar{n}(c_b)$ can be read off directly from the shape of $P(n)$. Specifically, $\sigma(0)$ can be inferred from the width of the tail of the distribution on the right of the knee. $\bar{n}(c_b)$ is instead related to the shape of $P(n)$ left of the knee. If $\sigma(c_b)$ is very small, the distribution of n for fixed b is very narrow, so that c and c_b tend to coincide. In this limit, $c(n)$ defined by Eq. (2) is equal to the inverse function of $\bar{n}(c_b)$ [7]. $P(n)$ can thus be obtained by differentiating the centrality with respect to n, according to Eq. (2):

$$P\left(n = \bar{n}(c_b)\right) \simeq -\left(\frac{d\bar{n}(c_b)}{dc_b}\right)^{-1}. \qquad (7)$$

In the inset of Fig. 4, we check the validity of this approximation in our TRENTo calculation by direct comparison of $\bar{n}(c_b)$ to $n(c)$. We find that $\bar{n}(c_b)$ deviates from $n(c)$ only above $n \simeq 170$, corresponding to 1.5% centrality. This means that $\bar{n}(c_b)$ carries an information which can not be inferred from data alone only in the vicinity of $n_{\rm knee}$.

To complete our procedure, we show how one can reconstruct the distribution of impact parameter for a given

	direct	fit (A) $\sigma(c_b) \propto \sqrt{\bar{n}(c_b)}$	fit (B) $\sigma(c_b)$ =const.
$n_{\rm knee}$	181.49 ± 0.13	182.03	181.91
$\sigma(0)$	7.79 ± 0.04	8.03	8.09
a_1	4.36 ± 0.02	4.41	4.48
a_2	-2.3 ± 0.1	-2.4	-3.0
a_3	4.8 ± 0.1	4.9	6.5

TABLE I. Fit parameters obtained from fitting $P(n)$ in the TRENTo calculation [Fig. 3].

[5] The larger difference in a_2 and a_3 is due to the different lower cutoff, $n_{\rm min}$, used for (B).

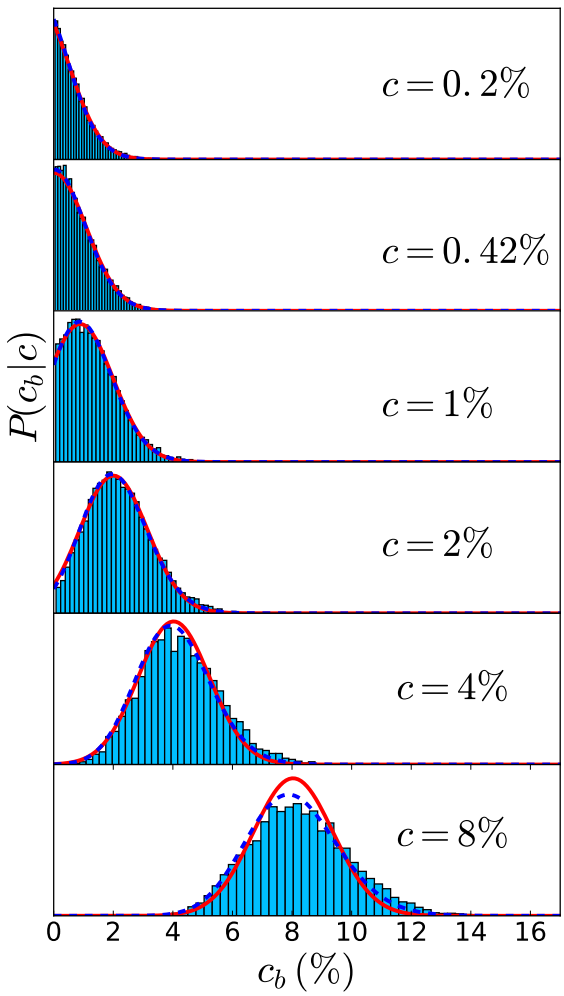

FIG. 5. (Color online) Probability distribution of b-centrality in a very narrow centrality bin. Histograms: direct calculation using TRENTo. Lines: Probability distributions of c_b reconstructed through the Bayes' theorem, Eq. (8). Solid and dashed lines correspond to scenarios (A) and (B), respectively. The second panel corresponds to $n = n_{\rm knee}$. Note that for $n \geq n_{\rm knee}$ (two uppermost panels), the most probable value of c_b is $c_b = 0\%$. Left of the knee, on the other hand, the most probable value of c_b is c [7], as clearly visible in the two lowermost panels. All curves displayed in this figure have area normalized to unity.

value of n, i.e., at a given value of centrality. In the TRENTo model, this can done directly by sorting events into very narrow centrality bins, and then looking at the distribution of c_b in each bin. We show such histograms of c_b for a few selected centrality bins in Fig. 5. Each panel corresponds to a bin of width 0.1%, centered around the displayed value of c (for instance, the panel with $c = 1\%$ shows the distribution of c_b for $0.95\% < c < 1.05\%$.)

	STAR [9]	ALICE [11]	ATLAS [13]
$n_{\rm knee}$	296.8	20406	3.5746
$\sigma(0)$	21.5	731	0.1131
a_1	3.55	4.11	4.05
a_2	0.8	-1.9	-1.5
a_3	1.6	4.4	4.1

TABLE II. Values of fit parameters for several experiments.

In an experimental situation where the impact parameter is not known, these distributions can be reconstructed from $P(n)$ using Bayes' theorem:

$$\begin{aligned} P(c_b|c) &= P(c|c_b) \\ &= \frac{P(n|c_b)}{P(n)}, \end{aligned} \tag{8}$$

where, in the first line, we have used the property that the distribution of c_b and c are uniform, i.e., $P(c_b) = P(c) = 1$. The distribution $P(n|c_b)$ is given by Eq. (3), where $\bar{n}(c_b)$ and $\sigma(c_b)$ can be obtained from the fitting procedure.

One first needs to determine the value of n corresponding to a given centrality, c. This can be done straightforwardly using the fitting function, which offers a smooth interpolation of $P(n)$.[6] Inserting Eq. (3) into Eq. (5), and exchanging the integration order, Eq. (2) yields

$$c = \int_0^1 \frac{1}{2}\,\mathrm{erfc}\left(\frac{n - \bar{n}(c_b)}{\sqrt{2}\sigma(c_b)}\right) dc_b, \tag{9}$$

where $\mathrm{erfc}(x)$ denotes the complementary error function. Eventually, once n is determined, $P(c_b|c)$ is given by application of Bayes' theorem, Eq. (8), to $P(n|c_b)$ in Eq. (3). The reconstructed distributions are shown as lines in Fig. 5. Scenarios (A) and (B) are represented by solid and dashed lines, respectively. Both are in perfect agreement with the direct calculation up to 4% centrality. Discrepancies between both scenarios and the direct calculation start to appear around $c = 8\%$, meaning that our approximated formulas for $\sigma(c_b)$ starts to break down around $c \sim 10\%$ in this model calculation. In a sense, this is a consequence of the fact that the variation of σ with c_b can not be inferred from the sole $P(n)$.

V. APPLICATION TO DATA

Our method of relating c to c_b being validated on model calculations, we apply now the fitting procedure to experimental data. We fit the experimental curves of $P(n)$

[6] Experimental data typically present $P(n)$ in a bin center, and the centrality, c, at the boundary between two bins.

shown in Fig. 1 using Eq. (5). The results shown in this section correspond to scenario (A), with a lower centrality cutoff. We have checked that results are stable if one varies the cutoff, or if one uses scenario (B). The values of the fit parameters extracted from experimental data are given in Table II, for all the analyzed experiments. The values of $n_{\rm knee}$ and $\sigma(0)$ are in the same units as n, which are usually arbitrary, and vary from one experiment to the other. The other fit parameters are dimensionless.

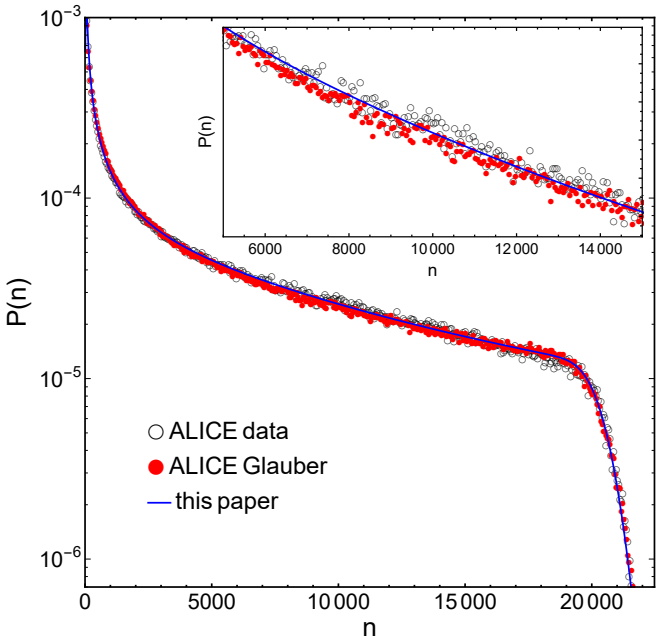

FIG. 6. (Color online) Empty symbols: distribution of the VZERO amplitude (denoted by n) used by the ALICE collaboration to determine centrality [11]. Full symbols: Glauber model used by the ALICE collaboration to fit the measured distribtuion. Line: Fit of data provided by Eq. (5). The inset is a zoom of the central part of the histogram. The quantity n is in arbitrary units.

In Fig. 6 we illustrate raw data on $P(n)$ (empty symbols) measured by the ALICE Collaboration, together with our fit (line). The χ^2 per degree of freedom of the fit is 1.2. Along with data, we plot also the distribution of n provided by the Monte Carlo Glauber model used by the ALICE collaboration (full symbols) to fit their $P(n)$, and, consequently, to perform the sorting of events into centrality bins. A remarkable outcome of our fitting method is that it provides a description of experimental data which is better than that provided by the Glauber model tuned to ALICE data, as evident from the inset of Fig. 6, where we zoom in the central body of the histogram. The fits of STAR and ATLAS data are of the same quality.

A convenient measure of the accuracy of the centrality determination is the fraction of events above the knee of the distribution, that is, the centrality of the knee, which we denote by $c_{\rm knee}$. It is obtained by replacing n with $n_{\rm knee}$ in Eq. (9). Its value for each of the three considered experiments is given in Table III. The recovered ordering of $c_{\rm knee}$ among the considered experiments is consistent with the curves shown in Fig. 1, since $c_{\rm knee}$ is proportional to the width of the tail of $P(n)$. We now derive a approximate expression of $c_{\rm knee}$ which provides a simple way to relate it to the fit parameters of Table II. If $n = n_{\rm knee}$, only small values of c_b contribute to the integral in Eq.(5). Therefore, we can expand $n(c_b)$ to first order in c_b in Eq. (3), obtaining $n(c_b) - n_{\rm knee} \simeq c_b (d\bar{n}/dc_b)|_{c_b=0}$. Neglecting the variation of the width with c_b, i.e., $\sigma(c_b) \approx \sigma(0)$, and replacing $\bar{n}(c_b)$ in Eq. (9), one obtains

$$c_{\rm knee} = -\frac{\sigma(0)}{\frac{d\bar{n}}{dc_b}\Big|_{c_b=0}\sqrt{2\pi}} = \frac{\sigma(0)}{n_{\rm knee} a_1 \sqrt{2\pi}}, \qquad (10)$$

where, in the last equality, we have used Eq. (6). We have checked that, using this simple estimate, the values of $c_{\rm knee}$ shown in Table II are reproduced to a good accuracy, within 1%.

Ultimately, using Bayes' theorem, we can use the fits of experimental data to reconstruct the distribution of impact parameter for a fixed centrality. Figure 7 presents the distribution of c_b for a few selected values of centrality percentile. The distributions obtained from ALICE and ATLAS data are very similar, therefore, we show only a comparison between the curves derived from ALICE and STAR data. The distribution extracted from STAR data is much broader than the one extracted from ALICE data. This is a direct consequence of the wider tail of $P(n)$ measured by the STAR Collaboration [see Fig. 1]. Note that, as expected, the distributions extracted from STAR and ALICE data are almost identical if one rescales both c and c_b by $c_{\rm knee}$, as we show in Fig. 8.

VI. CONCLUSION

We have shown that even though the impact parameter of a nucleus-nucleus collision is not a measurable quantity, precise information about impact parameter is contained in available experimental data. We have delineated a procedure allowing to reconstruct accurately the probability distribution of impact parameter at a given centrality (as defined experimentally by a multiplicity or a transverse energy), up to $5-10\%$ centrality. This reconstruction does not involve the concept of participant

Experiment	$c_{\rm knee}$
STAR [9]	$0.81\% \pm 0.10\%$
ALICE [11]	$0.349\% \pm 0.023\%$
ATLAS [13]	$0.313\% \pm 0.011\%$

TABLE III. Fraction of events above the knee for various heavy-ion experiments.

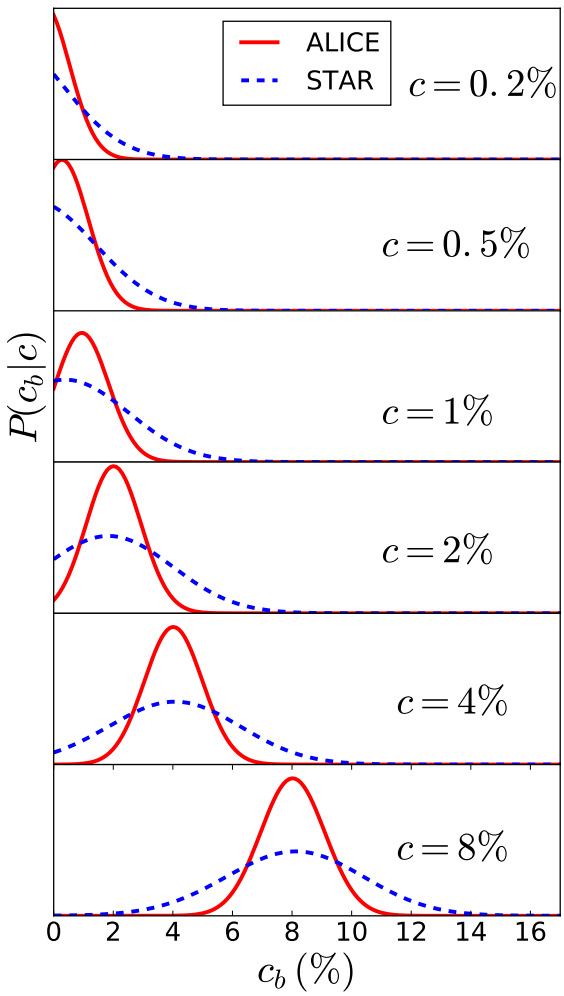

FIG. 7. (Color online) Probability distributions of b-centrality for selected values of the centrality reconstructed from STAR (dashed lines) and ALICE (solid lines) data using Bayes' theorem. All curves shown in this figure have area normalized to unity.

nucleons, or any microscopic model of the collision. Its sole inputs are the distribution $P(n)$, where n is the observable used to determine the collision centrality, along with the assumption that the distribution of n for a fixed b is Gaussian. We stress that this assumption is solidly rooted in the central limit theorem.

The fraction of events above the knee of $P(n)$, $c_{\rm knee}$, provides a simple measure of the precision of the centrality determination. It is below 0.4% at ATLAS and ALICE, and twice larger at STAR. We have shown that impact parameter fluctuations in the $0-10\%$ most central collisions are essentially determined by this quantity alone.

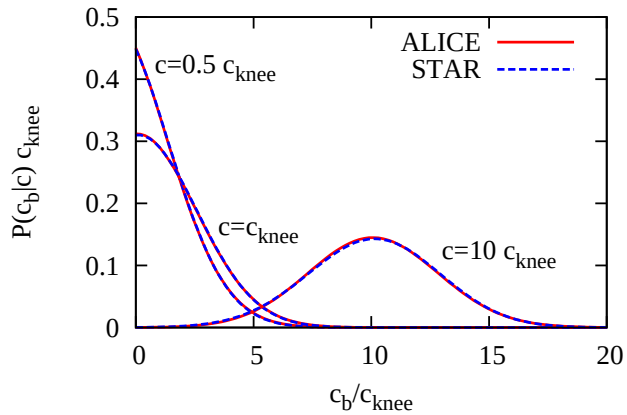

FIG. 8. (Color online) Probability distribution of $c_b/c_{\rm knee}$ extracted from STAR (solid lines) and ALICE (dashed lines) data for three selected values of $c/c_{\rm knee}$.

We have shown that the mean value of n at a fixed impact parameter can be reconstructed accurately up to 70% b-centrality. The standard deviation of n around the mean can instead be reconstructed only for $b = 0$: Its centrality dependence cannot be inferred from $P(n)$ alone. The mean and the standard deviation of n at fixed b are more natural quantities from a theory point of view than $P(n)$, as they can be directly obtained in a model by fixing the impact parameter. While the standard deviation may depend on detector details (in particular, purely statistical fluctuations are larger in relative value if the acceptance is smaller), the mean, $\bar{n}(c_b)$, provides a robust quantity for model comparisons.

It would be useful to extend this study to proton-nucleus collisions. However, our assumption that n has Gaussian fluctuations is not satisfied in model calculations, even for central collisions. We have checked that the fit procedure is less successful in describing p+Pb collisions at $\sqrt{s} = 5.02$ TeV [20]. In particular, the tail of the distribution of $P(n)$ is not as well reproduced, because it is not Gaussian. It will be worth investigating further the reasons of this failure of the Gaussian ansatz, to shed more light on the eventual mechanism of medium production in smaller collision systems.

ACKNOWLEDGEMENTS

We thank Alberica Toia and Jiangyong Jia for providing us with data of the ALICE and the ATLAS Collaborations. We thank Jean-Paul Blaizot, Matt Luzum, Sasha Milov, Stéphane Munier, and Lenka Zdeborova for useful discussions.

[1] C. Adler *et al.* [STAR Collaboration], Phys. Rev. Lett. **89**, 202301 (2002) doi:10.1103/PhysRevLett.89.202301 [nucl-ex/0206011].

[2] B. I. Abelev *et al.* [STAR Collaboration], Phys. Rev. Lett. **98**, 192301 (2007) Erratum: [Phys. Rev. Lett. **106**, 159902 (2011)] doi:10.1103/PhysRevLett.106.159902, 10.1103/PhysRevLett.98.192301 [nucl-ex/0607012].

[3] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. Lett. **105**, 252303 (2010) doi:10.1103/PhysRevLett.105.252303 [arXiv:1011.6182 [hep-ex]].

[4] K. H. Ackermann *et al.* [STAR Collaboration], Phys. Rev. Lett. **86**, 402 (2001) doi:10.1103/PhysRevLett.86.402 [nucl-ex/0009011].

[5] K. Aamodt *et al.* [ALICE Collaboration], Phys. Rev. Lett. **105**, 252302 (2010) doi:10.1103/PhysRevLett.105.252302 [arXiv:1011.3914 [nucl-ex]].

[6] J. Y. Ollitrault, Phys. Rev. D **46**, 229 (1992). doi:10.1103/PhysRevD.46.229

[7] W. Broniowski and W. Florkowski, Phys. Rev. C **65**, 024905 (2002) doi:10.1103/PhysRevC.65.024905 [nucl-th/0110020].

[8] B. B. Back *et al.* [PHOBOS Collaboration], Phys. Rev. Lett. **85**, 3100 (2000) doi:10.1103/PhysRevLett.85.3100 [hep-ex/0007036].

[9] C. Adler *et al.* [STAR Collaboration], Phys. Rev. Lett. **87**, 112303 (2001) doi:10.1103/PhysRevLett.87.112303 [nucl-ex/0106004].

[10] S. S. Adler *et al.* [PHENIX Collaboration], Phys. Rev. C **71**, 034908 (2005) Erratum: [Phys. Rev. C **71**, 049901 (2005)] doi:10.1103/PhysRevC.71.049901, 10.1103/PhysRevC.71.034908 [nucl-ex/0409015].

[11] B. Abelev *et al.* [ALICE Collaboration], Phys. Rev. C **88**, no. 4, 044909 (2013) doi:10.1103/PhysRevC.88.044909 [arXiv:1301.4361 [nucl-ex]].

[12] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1108**, 141 (2011) doi:10.1007/JHEP08(2011)141 [arXiv:1107.4800 [nucl-ex]].

[13] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **707**, 330 (2012) doi:10.1016/j.physletb.2011.12.056 [arXiv:1108.6018 [hep-ex]].

[14] X. N. Wang and M. Gyulassy, Phys. Rev. D **44**, 3501 (1991). doi:10.1103/PhysRevD.44.3501

[15] M. L. Miller, K. Reygers, S. J. Sanders and P. Steinberg, Ann. Rev. Nucl. Part. Sci. **57**, 205 (2007) doi:10.1146/annurev.nucl.57.090506.123020 [nucl-ex/0701025].

[16] L. Adamczyk *et al.* [STAR Collaboration], Phys. Rev. Lett. **115**, no. 22, 222301 (2015) doi:10.1103/PhysRevLett.115.222301 [arXiv:1505.07812 [nucl-ex]].

[17] J. S. Moreland, J. E. Bernhard and S. A. Bass, Phys. Rev. C **92**, no. 1, 011901 (2015) doi:10.1103/PhysRevC.92.011901 [arXiv:1412.4708 [nucl-th]].

[18] J. E. Bernhard, J. S. Moreland, S. A. Bass, J. Liu and U. Heinz, Phys. Rev. C **94**, no. 2, 024907 (2016) doi:10.1103/PhysRevC.94.024907 [arXiv:1605.03954 [nucl-th]].

[19] G. Giacalone, J. Noronha-Hostler and J. Y. Ollitrault, Phys. Rev. C **95**, no. 5, 054910 (2017) doi:10.1103/PhysRevC.95.054910 [arXiv:1702.01730 [nucl-th]].

[20] J. Adam *et al.* [ALICE Collaboration], Phys. Rev. C **91**, no. 6, 064905 (2015) doi:10.1103/PhysRevC.91.064905 [arXiv:1412.6828 [nucl-ex]].