Observing non-Gaussian sources in heavy-ion reactions

D. A. Brown[1] and P. Danielewicz[2]

[1]*University of Washington, Seattle, Washington 98195*
and Lawrence Livermore National Laboratory, Livermore, California 94551
[2]*Michigan State University, East Lansing, Michigan 48824*

We examine the possibility of extracting non-Gaussian sources from two-particle correlations in heavy-ion reactions. Non-Gaussian sources have been predicted in a variety of model calculations and may have been seen in various like-meson pair correlations. As a tool for this investigation, we have developed an improved imaging method that relies on a basis spline expansion of the source functions with an improved implementation of constraints. We examine under what conditions this improved method can distinguish between Gaussian and non-Gaussian sources. Finally, we investigate pion, kaon, and proton sources from the p-Pb reaction at 450 GeV/nucleon and from the S-Pb reaction at 200 GeV/nucleon studied by the NA44 experiment. Both the pion and kaon sources from the S-Pb correlations seem to exhibit a Gaussian core with an extended, non-Gaussian halo. We also find evidence for a scaling of the source widths with particle mass in the sources from the p-Pb reaction.

I. INTRODUCTION

Two-particle correlations have proven to be an important tool for experimentally accessing the space-time extent of heavy-ion collisions. For like-meson pair correlations (e.g., π's and K's), the correlation is dominated by the so-called Hanbury Brown–Twiss (HBT) effect (in other words, Bose-Einstein symmetrization of the meson-pair wave function) and the Coulomb corrected correlations are usually adequately parametrized by Gaussians [1–6]. Since meson final-state interactions (FSI) can usually be neglected, the Coulomb corrected correlation function becomes very nearly the Fourier transform of a source function. Thus, a Gaussian correlation corresponds to a Gaussian source function. *In general, there is no reason to expect the source to be Gaussian.* In fact, non-Gaussian sources may already have been observed in data [7–12].

There are several reasons to expect non-Gaussian sources: contributions from resonance decays should lead to an exponential halo [13–15], effects of space-momentum correlations (caused by either flow [2,16] or string fragmentation [17]) should lead to a focusing of the source [16], and even simple geometry should lead to non-Gaussian sources. Experimentally distinguishing between Gaussian and non-Gaussian sources is difficult and may be complicated by FSI within the pair. Recently it was realized that, by applying imaging techniques to the correlation data, we may extract the two-particle source function directly [11,12]. The imaging has two main advantages over the traditional HBT approach: it is model independent, meaning that it may reveal non-Gaussian features in the source, and it can clearly separate the effects of the FSI and symmetrization from effects due to the source itself. This last point requires elaboration: imaging extracts source functions that may be directly compared using correlations that cannot easily be compared when they arise from completely different particles. It is this feature that allows us to compare proton, kaon, and pion sources from p-Pb and S-Pb reactions from NA44. Indeed, a direct comparison of the proton and kaon sources from the p-Pb reaction suggests a simple scaling of the source widths that one should expect based on Lund-type string phenomenology in a fragmenting string.

Extracting the source function, $S_{\mathbf{P}}(\mathbf{r}')$, begins by noting that $S_{\mathbf{P}}(\mathbf{r}')$ is related to the experimentally measured two-particle correlation, $C_{\mathbf{P}}(\mathbf{q}')$, through a simple linear integral equation [2,18]:

$$\mathcal{R}_{\mathbf{P}}(\mathbf{q}') \equiv C_{\mathbf{P}}(\mathbf{q}') - 1 = \int d\mathbf{r}' K(\mathbf{q}', \mathbf{r}') S_{\mathbf{P}}(\mathbf{r}'). \tag{1}$$

Thus, ''imaging the source'' means somehow inverting this equation. Here primes denote quantities in the pair center-of-mass (c.m.) frame. Although for imaging purposes it is simplest to write Eq. (1) in the pair c.m., Eq. (1) may be written in any frame as $\mathcal{R}_{\mathbf{P}}(\mathbf{q}')$ is a Lorentz invariant observable. In Eq. (1), $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$ is the total momentum of the pair in the lab frame. The $\mathbf{P}$ subscript indicates the boost from the lab to the pair c.m. frame ($\mathbf{P}/P_0$ is the boost velocity between the frames). The kernel of Eq. (1) is

$$K(\mathbf{q}', \mathbf{r}') = |\Phi_{\mathbf{q}'}^{(-)}(\mathbf{r}')|^2 - 1. \tag{2}$$

The wave function $\Phi^{(-)}$ describes the propagation of the pair from a relative separation of $\mathbf{r}'$ in the pair c.m. to the detector with relative momentum $\mathbf{q}' = \frac{1}{2}(\mathbf{p}_1' - \mathbf{p}_2')$. The source function itself is the quasiprobability of emitting the pair a distance of $\mathbf{r}'$ apart, in the c.m. frame. We write the source as a convolution of Wigner functions, $D(\mathbf{r}, t, \mathbf{P}/2)$,

$$S_{\mathbf{P}}(\mathbf{r}') \equiv \int dt' \int d^3 R \, dT \, D(\mathbf{R} + \mathbf{r}/2, T + t/2, \mathbf{P}/2)$$
$$\times D(\mathbf{R} - \mathbf{r}/2, T - t/2, \mathbf{P}/2), \tag{3}$$

where the variables in the lab frame are understood as functions of the variables in the pair c.m. frame. Here the Wigner functions are normalized particle emission rates

$$D(\mathbf{r},t,\mathbf{p}) = \frac{Ed^7N}{d^3rdtd^3p} \bigg/ \frac{Ed^3N}{d^3p}, \qquad (4)$$

and may be computed directly from a transport model as discussed in Refs. [11,12,16]. Due to the time integral in Eq. (3), we cannot distinguish whether a given $\mathbf{r}'$ is associated with a time separation or a spatial separation.

Inverting Eq. (1) is generally an ill-posed problem. This means that small fluctuations in the data, even if well within statistical or systematic errors, can lead to large changes in the imaged source function. Ill-posedness stems from experimental factors (e.g., limited statistics, finite sized momentum bins, etc.) and the intrinsic resolution of the kernel in Eq. (2). In other fields, this stability problem is attacked using a variety of tactics including forcing the source function to obey known constraints or choosing a representation of the problem in which the kernel's resolution may be optimized. Both of these techniques were exploited in Ref. [12]. While the imaging in Ref. [12] was successful, the restored sources were represented in a basis that does not exhibit the continuity that we expect to see in the source. In this paper, we report a dramatic improvement of the imaging by using a representation of the source in which we have direct control over the continuity of the source. Our choice of representation still allows us to utilize constraints and to optimize the resolution of the kernel.

This paper is organized as follows. First, we will set up the problem of inverting angle-averaged correlations (i.e., expressed in terms of $q_{inv} = \sqrt{\mathbf{q}^2 - q_0^2}$) and outline the improved imaging method. The details of the imaging method and our representation of the source are contained in the appendices. Next, we apply the imaging method to correlations corresponding to Gaussian and non-Gaussian sources. This will orient us to some of the issues we will face when examining real data. Finally, we will confront like-pion, like-kaon, and two-proton correlation data from S-Pb collisions at 200 GeV/nucleon from NA44 and $p+$Pb collisions at 450 GeV/nucleon.

II. STATEMENT OF THE PROBLEM

In this section, we will set up the imaging problem. To simplify our discussion, we will consider only angle-averaged correlations and sources. First, we will outline the one-dimensional imaging problem and mention some of our expectations based on experience with Fourier transforms. Second, we will outline how we utilize the basis spline representation. Finally, we will describe our solution using a Bayesian approach to imaging. The details of the basis spline basis and the imaging itself are included in the Appendices.

A. The one-dimensional imaging problem

The source function and correlation in Eq. (1) both may be expanded in spherical harmonics and the relations between the angular coefficients are listed in Ref. [11]. With this expansion, we may image the individual components of the correlation function and compile a full three-dimensional imaged source. When doing this, one must take care in in-

terpreting the results: imaging is formulated in the pair c.m. frame as opposed to a frame in which we have more intuition, e.g., the lab frame. In this paper, we work with only the first term in the spherical expansions, i.e., the angle-averaged source and correlation. The drawback of performing a one-dimensional analysis is that the angular information is lost and the resulting source function is even more difficult to interpret.

The angle-averaged version of Eq. (1) is

$$\mathcal{R}(q) = C(q) - 1 = 4\pi \int dr r^2 K(q,r) S(r). \qquad (5)$$

Here $q = |\mathbf{q}'|$. For like pairs in the pair c.m. frame, $q_0' = 0$. This implies that $|\mathbf{q}'| = q_{inv} = \sqrt{\mathbf{q}^2 - q_0^2}$.

In Eq. (5), the kernel is simply the kernel in Eq. (2), but averaged over the angle between $\mathbf{q}$ and $\mathbf{r}$:

$$K(q,r) = \frac{1}{2} \int_{-1}^{1} d(\cos\theta_{\mathbf{qr}}) K(\mathbf{q},\mathbf{r}). \qquad (6)$$

For identical spin-zero bosons with no FSI, this kernel is

$$K(q,r) = \sin(2qr)/2qr, \qquad (7)$$

while with FSI it is

$$K(q,r) = \sum_{l \text{ even}} \frac{|g_l(r)|^2}{(2l+1)} - 1. \qquad (8)$$

Here l is the orbital angular momentum quantum number. Finally, for protons, the spin-averaged kernel is

$$K(q,r) = \frac{1}{2} \sum_{jsll'} (2j+1)[g_{js}^{ll'}(r)]^2 - 1. \qquad (9)$$

Here l and l' are both orbital angular momentum quantum numbers and j and s are the total angular momentum and spin quantum numbers. In the last two cases, g is the relative final-state radial wave function. For uncorrected meson data, g is the solution of the Klein-Gordon equation including the Coulomb potential. For protons, g is the solution of the Schrödinger equation using the Coulomb potential and REID93 [19] nucleon-nucleon potential.

Given that the identical particle kernels in Eqs. (1) or (5) are Fourier transform kernels at large distances, we expect our transforms to behave like Fourier transforms. If Eq. (5) were a Fourier transform, then by discretizing Eq. (5), we would be converting the imaging problem into a finite Fourier transform. In this case, the sampling theorem tells us how the sizes of the bins and the numbers of bins in the appropriate Fourier spaces are related,

$$\Delta r = \frac{\hbar c \pi}{q_{max}} \quad \text{and} \quad \Delta q = \frac{\hbar c \pi}{r_{max}}, \qquad (10)$$

where $q_{max} = N\Delta q$, $r_{max} = N\Delta r$, and N is the number of bins in both r and q space.

Using these relations, we may get a feeling for how structure in the data affect the imaged source. For example, the

low-q structure in the data sets the large length scale behavior of the source. Conversely, the high-q portion of the data sets the short length scale behavior of the source and therefore sets the size of the smallest features we could hope to resolve in the source. For example, if the correlation dies off around a $q \approx 80$ MeV/c, then we should not expect to resolve structure smaller than $\Delta r \approx 8$ fm. Owing to the fact that our kernel is not a trigonometric function in general, these estimates are qualitative at best. Nevertheless, we will often appeal to Fourier theory for explanations of some of the effects that we see while imaging.

B. The representation of the problem

In our calculations, we expand the imaged source in a function basis

$$S(r) = \sum_{i=1}^{N_M} S_i B_i(r) \tag{11}$$

and, in this basis, the error on the source is

$$\Delta S(r) = \sqrt{\sum_{i,j=1}^{N_M} \Delta^2 S_{ij} B_i(r) B_j(r)}. \tag{12}$$

Here $\Delta^2 S$ is the covariance matrix of the source coefficients. Once we average the kernel over momentum bins to account for the experimental binning, our inversion problem reduces to the following matrix equation:

$$\mathcal{R}_i \equiv \mathcal{R}(q_i) = \sum_{j=1}^{N_M} K_{ij} S_j, \tag{13}$$

where the kernel matrix is

$$K_{ij} = \frac{4\pi}{\Delta q} \int_{q_i - \Delta q/2}^{q_i + \Delta q/2} dq \int_0^\infty dr\, r^2 K(q,r) B_j(r). \tag{14}$$

Here Δq is the momentum bin size. Our source vector is made of the coefficients S_j of the basis function representation of the source and our data vector is made of the correlation values $\mathcal{R}_i$.

The function basis that we use to represent our source function must have several properties: (1) it must be efficient, i.e., requiring few coefficients to represent a realistic source, (2) it should be continuous, or at least have continuity as an option, and (3) it should have an adjustable parameter that we might use to optimize the resolution in a manner analogous to Ref. [12]. One obvious possibility is to either use a Laguerre expansion [20] (so that the first term is an exponential fitted to the source) or an Edgeworth expansion [20,21] (so that the leading term is a Gaussian fitted to the source). The down side of either of these choices is that it is difficult to adjust the terms in one's expansion to maximize the resolution of the inversion. Furthermore, one could argue that if one picks one of these bases and keeps only a few terms in the expansion, then one biases the inversion to give, for example, only Gaussian sources.
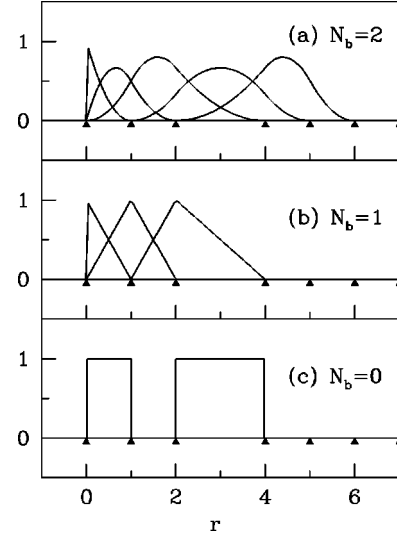

FIG. 1. Sample plots of N_b^{th} degree b-splines. In all panels, the knots are marked by carets and the knots at $r=0$ are actually $N_b + 1$ regular knots piled together.

We choose to represent the source function in a basis spline (also known as b-spline) basis [22] as this basis has all of the features we require for a good representation. Plots of some sample b-splines are shown in Fig. 1 and this basis is detailed in Appendix A. B-splines are piecewise polynomials and are continuous up to the degree of these polynomials. The 0th degree b-spline is the box-spline, making our b-spline expansion a natural generalization of the approach in Refs. [11,12]. Furthermore, in the b-spline basis the concept of the ''edge of a bin'' in the box-spline basis is replaced with the concept of a knot [22]. A knot is simply the place where the polynomials that make up the b-spline are patched together. In the ''optimized discretization'' scheme of Ref. [12], the edges of the box-splines are varied to minimize the relative error of the source. We may generalize this idea to the b-splines easily by varying the locations of the knots. We will give examples of choosing the knots in Sec. III C and we will explain in detail how to choose the ''optimal knots'' in Appendix A.

C. The reconstruction

Once we have converted the inversion problem into a matrix inversion by choosing a representation of the source, we proceed as in Refs. [11,12,23,24] and extract the source. The details of the Bayesian approach to imaging are discussed in Appendix B and we summarize the main results here. To obtain the coefficients of the source, we seek the source that minimizes the χ^2:

$$\chi^2 = (K \cdot \mathbf{S} - \boldsymbol{\mathcal{R}})^T \cdot (\Delta^2 \mathcal{R})^{-1} \cdot (K \cdot \mathbf{S} - \boldsymbol{\mathcal{R}}), \tag{15}$$

where $\Delta^2 \mathcal{R}$ is the covariance matrix of the correlation data. The source that does this is

$$\mathbf{S} = \Delta^2 S \cdot K^T \cdot (\Delta^2 \mathcal{R})^{-1} \cdot \boldsymbol{\mathcal{R}}. \tag{16}$$

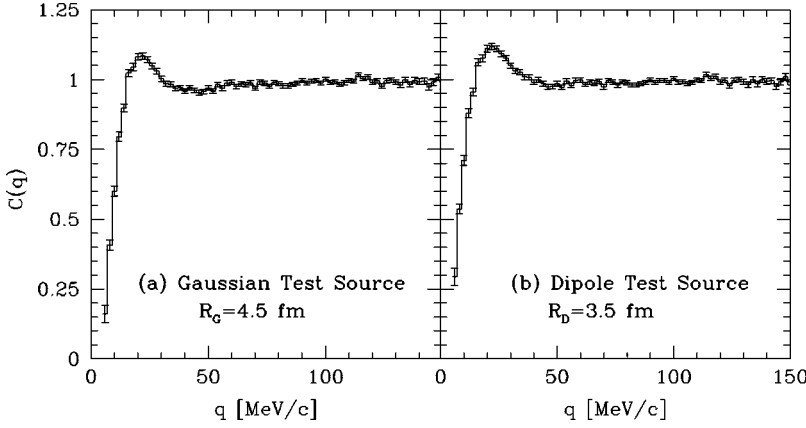The covariance matrix of this source is

FIG. 2. Model proton correlation corresponding to (a) Gaussian proton source function and to (b) dipole proton source.

$$\Delta^2 S = [K^T \cdot (\Delta^2 \mathcal{R})^{-1} \cdot K]^{-1}. \tag{17}$$

One should note the dependence on the experimental uncertainty $\Delta^2 \mathcal{R}$ in Eqs. (16) and (17). In Eq. (16), the points with the largest error contribute to the source determination the least. Also, in Eq. (17) the points that are most affected by the points with the large error also have the largest uncertainty.

In order to stabilize the inversion, we can take advantage of prior information in the form of equality constraints [25]. An equality constraint is a condition on the vector of source coefficients that has the generic form $\mathcal{C} \cdot \mathbf{S} = \mathbf{c}$. One example of such a constraint is that the source has slope zero at the origin. Such a situation arises if the normalized particle emission rates, D [cf. Eq. (4)], have a maximum. In this case we write

$$S'(r \to 0) = \sum_{i=1}^{N_M} S_i B_i'(r \to 0) = 0. \tag{18}$$

Thus, this case corresponds to $\mathcal{C}_i = B_i'(r \to 0)$ and $c = 0$. We can implement this type of constraint by adding a penalty term to the χ^2:

$$\chi^2 + \lambda (\mathcal{C} \cdot \mathbf{S} - \mathbf{c})^2. \tag{19}$$

Here λ is a trade-off parameter and we may vary it in order to emphasize stability in the inversion (by making λ huge) or to emphasize goodness of fit (by setting λ to zero). Such an ability to trade-off stability for goodness of fit is discussed in Ref. [26] in detail. With this modification of the χ^2, the imaged source is

$$\mathbf{S} = \Delta^2 S \cdot [K^T \cdot (\Delta^2 \mathcal{R})^{-1} \cdot \mathcal{R} + \lambda \mathcal{C}^T \cdot \mathbf{c}], \tag{20}$$

and the covariance matrix of source now is

$$\Delta^2 S = [K^T \cdot (\Delta^2 \mathcal{R})^{-1} \cdot K + \lambda \mathcal{C}^T \cdot \mathcal{C}]^{-1}. \tag{21}$$

There is another way in which prior information enters into the inversion—in the representation we choose for the source. By using, say, $N_b = 2$ in a b-spline expansion of the source, we are really assuming that our source and its first and second derivatives are all continuous.

The reader should note that, when we image, we are really finding a probability density for the source given the correlation data rather than the source itself. The set of source coefficients and the covariance matrix of the source characterize the height and width of this probability distribution. In the end, we use the source coefficients as an estimator of the true source.

III. TESTS OF THE IMAGING

We now explore the imaging in the b-spline basis by inverting some simple model correlations. We consider two model sources, a Gaussian source:

$$S(r) = \frac{\lambda}{(2\sqrt{\pi}R_G)^3} \exp\left(-\frac{r^2}{4R_G^2}\right) \tag{22}$$

and a source with a dipole form-factor-like shape:

$$S(r) = \lambda \frac{2}{\pi^2} \frac{R_D}{(r^2 + 4R_D^2)^2}. \tag{23}$$

This second source has a roughly Gaussian peak and an extended tail that one could imagine corresponds to long-time emission of particles. We chose this source to facilitate comparison to the experimental results in the next section. We pick $R_G = 4.5$ fm, $R_D = 3.5$ fm, and $\lambda = 1$. To generate the correlations, we convolute the source with the proton kernel in Eq. (9) and bin the correlation in 6-MeV/c sized q bins. To simulate realistic data, we take the error bars from the real data of Ref. [27] and add statistical scatter consistent with these error bars. The data in [27] are plotted in the same size momentum bins as in our test, this data is fully corrected for various experimental effects, and these effects are properly reflected in their estimates for the experimental uncertainty. The resulting correlations are shown in Fig. 2. In all of our tests, we confine ourselves to proton correlations because the proton FSI are more important than meson FSI and place a more demanding test on the imaging.

Our first test is to examine how the quality of the source reconstruction depends on the number of coefficients in the source expansion. In this test we will use only box-splines. In the second test, we will use higher degree b-splines in the reconstruction. In this test, we will use a fixed separation
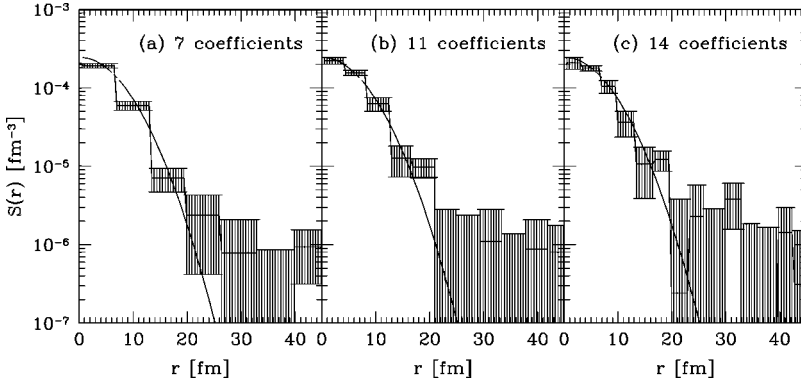
FIG. 3. Reconstructions of the Gaussian source with different numbers of coefficients in the source image. In both panels, the model sources are the solid curves and the reconstructed source is the curve with the error band.

between the knots (equivalent to using equal width box-splines). In the third test, we will demonstrate the use of the "optimal knots" in analogy to the "optimized discretization" method of Ref. [12]. In the final test, we will demonstrate the practical use of equality constraints.

A. Number of coefficients in the source

Before we begin our imaging tests, we must decide on the size of our imaging region and set the number of coefficients that we wish to reconstruct. Naively, to set r_{max} we might use the Fourier estimates from Eqs. (10) giving $r_{max} = 103$ fm. Experience has shown us that the source is usually lost in statistical noise and is consistent with zero to within the errors long before this r_{max}. So, we set $r_{max} = 45$ fm, roughly half of what the naive Fourier estimate suggests. If we find this is too conservative, we may increase it later.

To set the number of coefficients, we could use the naive Fourier estimate again. Doing so, we find that $\Delta r = 4.1$ fm and that we should use 11 coefficients. On the other hand, Eq. (13) suggests that the imaging problem is really a problem of simultaneously solving a set of linear equations. Given this, we look at the data and see that there are roughly 15–16 points that are different from one and hence contain useful information. This suggests that we try using something like 14 coefficients in the source.

This raises the question of the amount of information in a data set. If a correlation is Gaussian, than one could argue that it contains only two pieces of information: the height and width of the Gaussian. On the other hand if one bins the data, one could argue that there are really N pieces of information corresponding to the number of bins where there is

an apparent signal. We adopt the second viewpoint, but comment that the "amount of information" in a data set is an imprecise concept.

Since it is not clear as to which number of coefficients we should use, we will try both. Additionally, we will image using seven box-splines so that we may compare the results with the higher-degree results of the next section. In Figs. 3 and 4 we plot the inversions of the proton correlations in Fig. 2.

First we look at the Gaussian source images in Fig. 3. In all three panels, the inversions are reasonable representations of the true source. Only by looking at the χ^2 is it clear which image is the "best": for panel (a) $\chi^2 = 122$, for panel (b) $\chi^2 = 91$, and for panel (c) $\chi^2 = 76$. Since there are only 83 points in the proton correlations, the inversion with 14 coefficients is "too good" and 11 coefficients seems to be the best choice. Before moving on to discuss the dipole sources, we mention that the fluctuations in the imaged sources are not independent. If they were, then we would expect that roughly one-third of the bins would differ from the true source by at least a standard deviation.

Now we turn to the dipole sources. Looking at Fig. 4, none of the images are ideal—all three have large fluctuations that imply that there is a zero somewhere around 10 fm. In these three plots, the χ^2 is not much of a guide either. The χ^2's are 104, 90, and 77, respectively. Finally we comment that we cannot tell which images in Figs. 3 and 4 correspond to Gaussian sources or dipole sources.

We have seen that increasing the number of coefficients in the reconstruction helps to reproduce the source, however there is a practical limit as to how many coefficients we may add. As the number of source coefficients increases, they
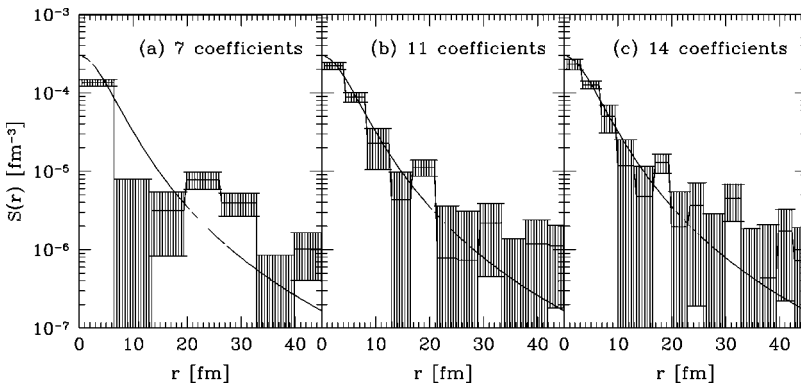

FIG. 4. Reconstructions of the dipole source with different numbers of coefficients in the source image. In both panels, the model sources are the solid curves and the reconstructed source is the curve with the error band.
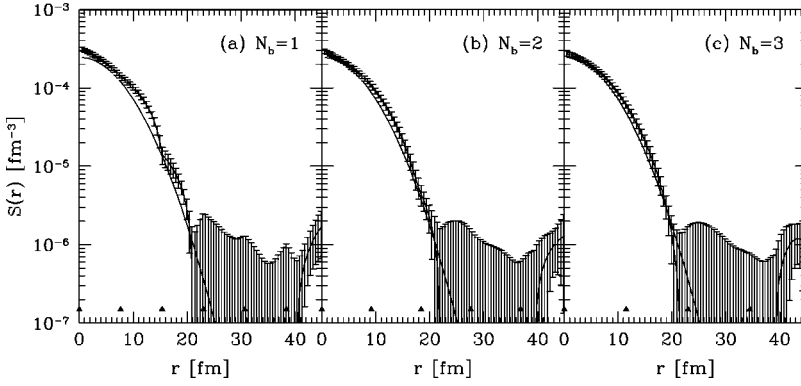
FIG. 5. In all three panels, the true Gaussian source is the solid curve and the reconstructed source is given by the points with errors. The knots in both panels are represented by carets.

become less constrained by the data. At some point there are more source coefficients than can be constrained by the data and then the extra coefficients only serve to reproduce the high frequency statistical fluctuations in the correlation data. At this point, the imaged source tends to oscillate about the true source as we have over-resolved the source. In general, we can never tell when we are over-resolving the source as we do not know the true source.

B. Using the basis spline representation

We expect the source function to be continuous as it is the convolution of two emission rates, yet we represent it by discontinuous box-splines. Thus, our first improvement over a simple box-spline representation of the source is to use higher degree b-splines.

We reimage the proton correlations in Fig. 2. In Figs. 5 and 6, we show the images obtained using the b-spline expansion with $N_M=7$ coefficients and either first, second, or third degree b-splines. Our choice of $N_M=7$ is somewhat arbitrary as we do not know how to extend our Fourier transform based estimates to our b-spline basis. However, the fact that the imaging works reasonably well points to the robustness of the method, despite the possibly suboptimal choice of N_M. In all plots, N_b+1 knots are placed at the end points of the imaging region (i.e., at $r=0$ fm and $r=45$ fm) and the rest of the knots are equally spaced between the end points.

In Fig. 5, the images are fairly accurate reconstructions of the source over two orders of magnitude, but the $N_b=3$ reconstruction is marginally better, due to the higher degree of continuity. In all cases, the inversions are better than any of the box-spline results. The corresponding χ^2's are 99 for

degree-1 splines, 94 for degree-2 splines and 93 for degree-3 splines. In these plots, the region past $r=17$ fm is lost in the noise from the correlation. We notice that all the plots exhibit the same kind of fluctuations seen in the $N_b=0$ images in the last section, however they are less noticeable because the b-splines are so delocalized. Finally, we mention that the unphysical rise out past 40 fm is most likely a result of aliasing. It is more obvious in these plots because the last b-spline has a cusp at the edge of the image.

We also image the dipole source in Fig. 2(b). Using the same settings as for the Gaussian source, we are able to reproduce the more complicated behavior of the dipole shaped source over two decades in source intensity. More importantly, upon comparing Figs. 5 and 6, we can clearly tell the difference between Gaussian and non-Gaussian source shapes on the logarithmic scale. This is something we could not do with the box-spline representation of the source. The χ^2 for these reconstructions are 95, 94, and 93, respectively. We comment that the cusp in the very first b-spline helps us represent the relatively sharp peak in the dipole source.

In all of the images shown so far, we see an unphysical rise in the source in the far right of the images. This rise is most likely a result of *aliasing*. Chapter 12 of *Numerical Recipes in C* [26] has a detailed discussion of this problem. Aliasing is a phenomenon that often occurs when approximating a Fourier transform over an infinite interval with a finite Fourier transform over a finite interval. Consider a function $f(r)$ and its Fourier cosine transform, $F(q)$:

$$F(q)=\int_0^\infty dr f(r)\cos(qr)\approx\int_0^{r_{max}} dr f(r)\cos(qr). \quad (24)$$
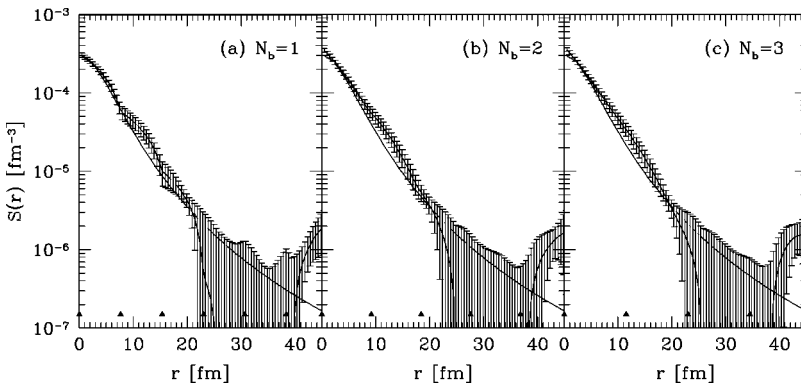

FIG. 6. In three panels, the true dipole source is the solid curve and the reconstructed source is given by the points with errors. The knots in both panels are represented by carets.
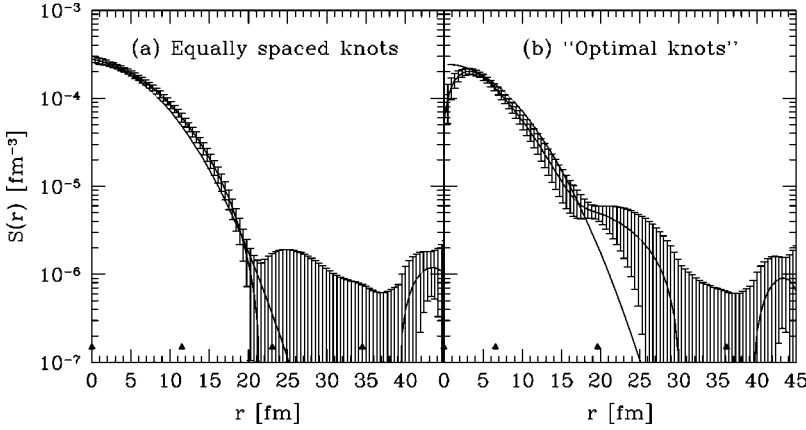
FIG. 7. In both panels, the true Gaussian source is the solid curve and the reconstructed sources are given by the points with errors. In panel (a) the knots are evenly spaced between the limits of the imaging region and in panel (b), the ''optimal knots'' are used. The knots in both panels are represented by carets. Note that the source from panel (c) in Fig. 5 is reproduced here as panel (a).

In the first line of this equation, low frequency structure in $F(q)$ corresponds to large distance structure in $f(r)$, which is neglected in the second line of this equation. Now, imagine beginning with $F(q)$ and attempting to infer $f(r)$ using a finite r_{max}. What often happens is that, whatever strength $f(r)$ should have out past r_{max} gets folded into the region $r<r_{max}$. In our inversion problem, the integral in Eq. (1) behaves like a Fourier transform. Since statistical fluctuations in the data are artificial high-frequency structure, we should not be surprised to see features reminiscent of aliasing when we image. Based on our experience, adjusting r_{max} or constraining the source at r_{max} can help cure this problem. However, the rise at the largest r is usually preceded by a region of the image that is consistent with zero so we can easily identify the usable part of the image and ignore any artifact due to aliasing.

C. Choosing the knots

For our next refinement, we examine how choosing the knots affects the inversion. Were the problem of imaging as simple as inverting a Fourier transform, the optimal bins in r would be evenly spaced and given by Eq. (10). However, the kernels we are interested in are often distorted by the Coulomb repulsion of the pairs as well as other FSI. Furthermore, some regions of the data have large errors and it would be useful if we could combine those bins somehow. Taken together, we must ask whether keeping equally spaced bins in the source is optimal. In Ref. [12], we found that we could

improve the imaging dramatically by choosing the size of the bins in r to minimize the error in the source relative to a test source. This technique can be generalized to the b-spline basis by simply varying the knots. To choose the ''optimal knots'' we proceed as mentioned earlier and detailed in Appendix A.

For the Gaussian source in Fig. 7, we show the inversions using third degree b-splines using seven coefficients for both fixed knots [as in Fig. 6(c)] and the ''optimal knots.'' Several things are apparent in this figure. First, the fixed width knot reconstruction is markedly better than the lower-degree b-spline reconstructions in the previous section, simply due to the higher degree of continuity. The χ^2 of this reconstruction is 93. Using the ''optimal knot'' reconstruction, the source is everywhere consistent with the true source except at the lowest r ($\lesssim 5$ fm) where the source drops nearly an order of magnitude. This drop is unphysical for a source that is the convolution of two single particle sources, each with one or more maxima. This drop is due to the close packing of the knots at the lowest r and can be remedied by lowering the number of coefficients in the reconstruction, increasing the size of the fit region, or by using equality constraints (as we will show in the next section). The χ^2 of this reconstruction is 90.

In Fig. 8, we show a similar set of inversions for the dipole shaped source. Both inversions seem to do a reasonable job of representing the source, except at the lowest r where the cusp of the first b-spline is a bit higher than the
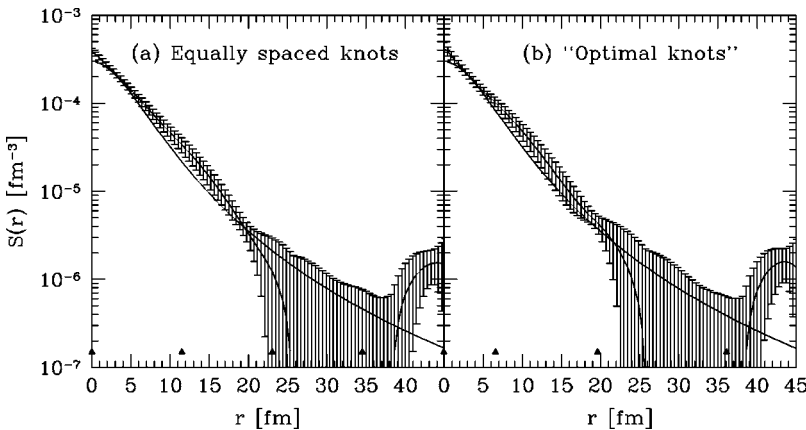

FIG. 8. In both panels, the true dipole source is the solid curve and the reconstructed sources are given by the points with errors. In panel (a) the knots are evenly spaced between the limits of the imaging region and in panel (b) the ''optimal knots'' are used. The knots in both panels are represented by carets. Note that the source from panel (c) in Fig. 6 is reproduced here as panel (a).

TABLE I. Equality constraints on the *b*-spline representation of spherically symmetric sources.

Constraint	Continuous representation	*b*-spline representation
Flat at $r=0$	$\dfrac{\partial S}{\partial r}(r\to 0)=0$	$\displaystyle\sum_{j=1}^{N_M} S_j B_j'(r\to 0)=0$
Normalized to λ	$4\pi\int_0^\infty dr\, r^2 S(r)=\lambda$	$\displaystyle\sum_{j=1}^{N_M} 4\pi S_j\int_0^\infty dr\, r^2 B_j(r)=\lambda$
Zero outside of imaged region	$S(r_{max})=0$	$\displaystyle\sum_{j=1}^{N_M} S_j B_j(r_{max})=0$
Flat at $r=r_{max}$	$\dfrac{\partial S}{\partial r}(r_{max})=0$	$\displaystyle\sum_{j=1}^{N_M} S_j B_j'(r_{max})=0$

true source. The ''optimal knot'' reconstruction is marginally better than the equally spaced knot reconstruction as it has a χ^2 of 92 compared to a χ^2 of 93 for the equally spaced knot reconstruction.

Given the inconsistent performance of the ''optimal knot'' reconstruction, we ask ourselves why this refinement does not always help. To find the optimal knots, we move the knots to minimize the error on the source relative to a test source (which is the same for all of the inversions). The error on the source depends *only* on the kernel and the error in the data, so the ''optimal knots'' do not know anything about the true source. If the true source has interesting structure some place where we are not sensitive to it, the ''optimal knots'' will be widely spaced there and we will not have the resolution to see the structure. Conversely, the ''optimal knots'' may end up giving very high resolution exactly where we do not need it, witness Fig. 7(b). On the other hand, the ''optimal knots'' can give resolution exactly where we need it, as in Fig. 8(b).

D. Equality constraints

Now we come to the final refinement, the use of equality constraints. As we have mentioned before, a constraint is a piece of prior information such as knowing that the first derivative at **r**=0 should vanish for a differentiable angle-averaged source, $S'(r\to 0)=0$. Using constraints amounts to adding information, so we imagine that we will be able to use more coefficients in the reconstructions. This we will see illustrated below. A list of constraints we could use are shown in Table I.

In Fig. 9, we show inversions using the $S'(r\to 0)=0$ constraint for the Gaussian source. We used the ''optimal knots,'' third degree *b*-splines and seven coefficients [in panel (a)] and nine coefficients [in panel (b)] in these inversions. We see that we have solved the pathological behavior of the imaged source at low r and the agreement with the true source appears good. Upon examining the χ^2 [107 for panel (a) and 93 for panel (b)] we see that the seven coefficient source is a lot worse than it appears as it is routinely higher than the true source in the region from 10–20 fm. In Fig. 10, we show the results of using the same constraint to image the dipole-shaped source. Here we see that, for seven coefficients [panel (a)], the quality of the image has gone down considerably: we no longer match the height of the peak and we cannot resolve any of the tail. The χ^2 for this inversion is a comparatively large 108. For nine coefficients [panel (b)], the situation is much better. We now get the peak and can resolve the tail. The χ^2 here is 89.

We see that this one constraint gave us the ability to add another two points in the reconstructions without over-resolving the source. At a practical level, the first few *b*-spline coefficients must be adjusted together in order to
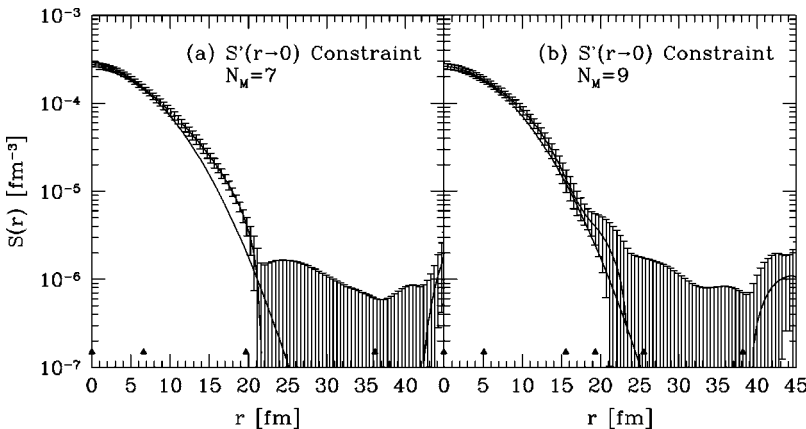

FIG. 9. In both panels, the true Gaussian source is the solid curve and the reconstructed sources are the points with errors. In both reconstructions, the source is constrained to have zero derivative at the origin. In panel (a), we use seven coefficients and, in panel (b), nine coefficients. The knots in both panels are represented by carets.
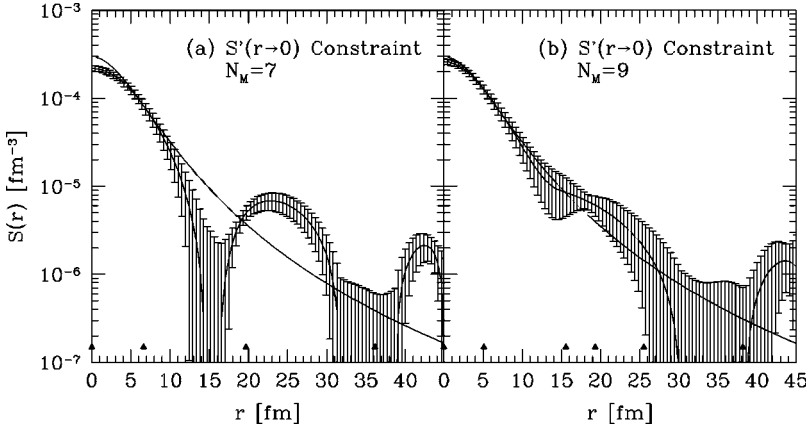
FIG. 10. In both panels, the true dipole source is the solid curve and the reconstructed sources are the points with errors. In both reconstructions, the source is constrained to have zero derivative at the origin. In panel (a), we use seven coefficients and, in panel (b), nine coefficients. The knots in both panels are represented by carets.

satisfy the constraint, in effect leaving fewer coefficients to fit the data. Thus, we must add more coefficients if we want to simultaneously fit the data and satisfy the constraint. This observation leads us to posit a rule of thumb: (amount of information in the data) + (number of constraints) ⩾ (number of coefficients in expansion). Additionally, one should pick the number of coefficients somewhere near what one would estimate based on the Fourier estimates discussed earlier.

Finally, by introducing all three refinements of the imaging (b-splines, optimal knots, and equality constraints) we are able to reproduce the height of the source at $r=0$ quite accurately. The value of the source at $r=0$ is essential for extracting the space-averaged phase-space density [11,28].

IV. ANALYSIS OF NA44 CORRELATIONS

Since we can reliably image a source from correlations using the Bayesian approach to imaging in a b-spline representation, we turn to the analysis of NA44 correlations. In a series of papers [7–9,29], NA44 detail their measurements of angle-averaged pion, kaon, and proton correlations from p+Pb collisions at $450A$ GeV/c and central S+Pb collisions at $200A$ GeV/c. In two of the earlier papers [7,8], they claim to have detected non-Gaussian kaon and pion correlations. To bolster their claim, they fit the Coulomb corrected correlations to Gaussians and to exponentials. In particular, they fit to the following functional forms:

$$\mathcal{R}(Q_{inv}) = \lambda \exp(-Q_{inv}^2 R_G^2) \qquad (25)$$

implying the Gaussian source of Eq. (22) and

$$\mathcal{R}(Q_{inv}) = \lambda \exp(-2Q_{inv}R_D) \qquad (26)$$

implying the source with a dipolelike shape of Eq. (23). Here $Q_{inv} = 2q_{inv} = \sqrt{-(p_1 - p_2)^2}$, the relative momentum variable traditionally used in the analysis of meson correlations. The NA44 correlations that we image are collected in Fig. 11.

In this section, we first image the NA44 correlations. Second, we compare the images to the results of some of NA44's fits. Next, we discuss NA44's (RQMD) simulations

of the S-Pb reaction and the implications for the source function images. Finally, we discuss the sources from the NA44 p-Pb data.

A. Imaging analysis

The results of the imaging analysis are presented in Fig. 12. As a crosscheck, in Fig. 11 we plot the correlations corresponding to the inverted sources along with the original data. In these inversions, we used either the noninteracting meson kernel in Eq. (7) (for the Coulomb corrected pion and kaon correlations) or the proton kernel in Eq. (9). Due to the differences in kernels, binning, and quality of the various data sets, each image had to be hand tuned separately. Since we do know the true sources that correspond to the data, we used a set of three criteria to decide when we have a good source.

(1) Is the image stable—i.e., when we tweak a parameter (e.g., the number of bins, r_{max}, number of constraints, etc.) does it change much?

(2) Does the imaged source give a correlation consistent with the original?

(3) Is the χ^2 as small as we can make it?

In all cases, we used third degree b-splines. The parameters of the inversions are collected in Table II. Only the p-Pb pp source was imaged without the $r=0$ smoothness constraint. We did not use this constraint because the knot density is too low at low r. Turning on the constraint widens the peak artificially as the next few b-splines have to be tuned to get the zero slope at the origin, dramatically raising the χ^2.

When looking at the images, several things are apparent. First, each of the sources from the p-Pb reactions are roughly a factor of 2 narrower than the corresponding sources from the S-Pb reaction. This is most likely a result of the different system sizes. Second, a comparison of the sources from the same reaction reveals that the pion sources are wider than the kaon sources and the kaon sources are wider than the proton sources. Next, it is apparent that all six of the sources have main peaks that appear Gaussian. However, both the pion and kaon sources in the S-Pb reaction have significant non-Gaussian tails. These tails are most likely not due to aliasing as r_{max} in both plots is at 35 fm, but, in order to show all six
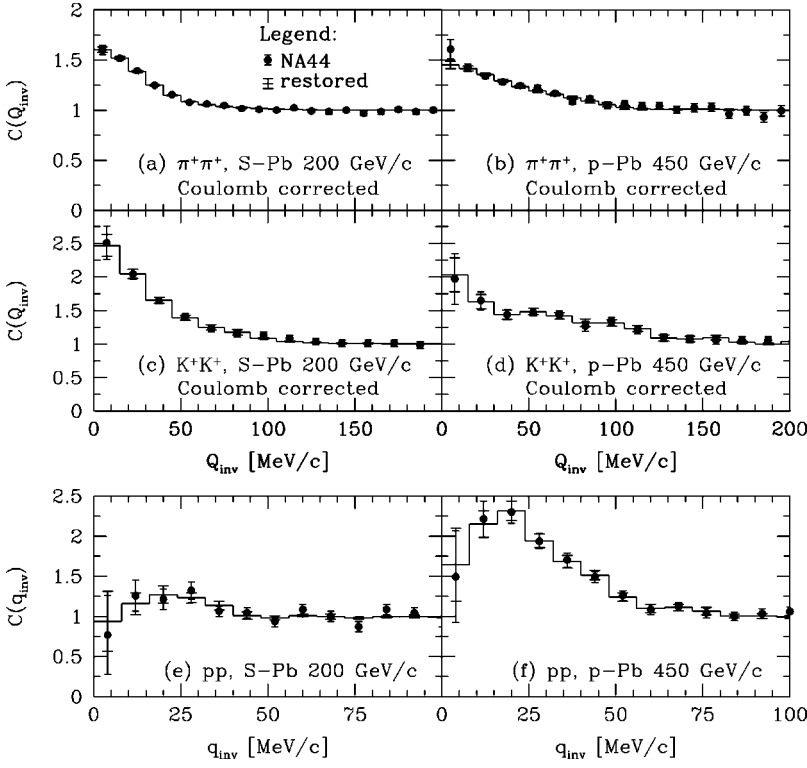
FIG. 11. Pion, kaon and proton correlations for the S-Pb and p-Pb reactions. The points and narrow error bars correspond to the experimentally measured correlations. The histogram and wide error bars correspond to the restored correlations from the imaging analysis. The pion, kaon, and proton correlations are from Refs. [9], [7], and [29], respectively.

sources on the same scale we truncated the plots at 20 fm. Unfortunately, this means that we cannot display that the kaon and pion sources are consistent with zero in the region from 25–30 fm nor can one see the rise that is obviously due to aliasing past the region 30–35 fm in the kaon source. No aliasing is apparent in the pion source or any of the proton sources. In the pion and kaon sources from the p-Pb reaction, aliasing is apparent on the far right side of the plots.
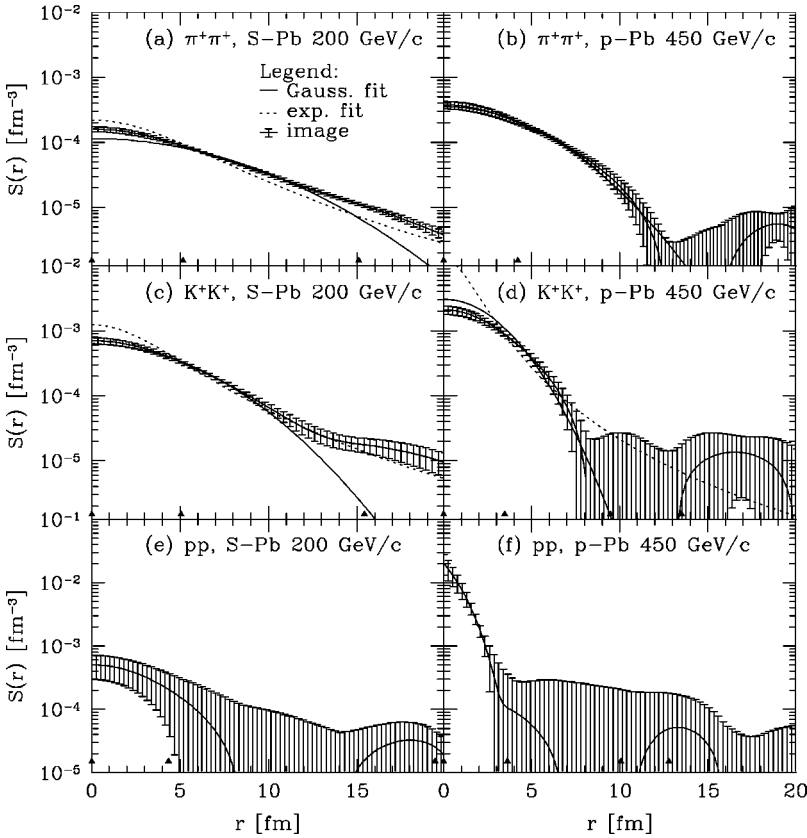

FIG. 12. Sources imaged from the S-Pb and p-Pb reactions. Where applicable, we have also plotted the Gaussian and dipole-shaped sources corresponding to NA44's fits.

TABLE II. Parameters used in the reconstruction of the NA44 sources. The numbers in parentheses in the number of data points column is our estimate of the number of points which contain usable information.

		r_{max} (fm)	No. of coeffs	$r=0$ constraint?	No. of data pts.	χ^2
S-Pb	$\pi^+\pi^+$	35	7	Yes	29(7)	19.8
	K^+K^+	35	7	Yes	16(8)	5.0
	pp	26	6	Yes	20(6)	14.6
p-Pb	$\pi^+\pi^+$	21	5	Yes	29(9)	24.8
	K^+K^+	26	8	Yes	29(9)	23.1
	pp	26	8	No	20(8)	7.6

B. NA44 fits

Following the imaging analysis of the NA44 correlations, it is useful to compare those results to the various fits performed by NA44. We have in mind two sets of fits: the one-dimensional fits to Gaussians and exponentials in Refs. [7,8] and the three-dimensional Gaussian fits in Ref. [30].

We first compare the imaged sources to the results of NA44's one-dimensional fits. In addition to the imaged sources, in Fig. 12 we show the fits as solid curves (for the Gaussian fits) or dashed curves (for the exponential fits). NA44's fit parameters are summarized in Tables III and IV. In the S-Pb sources in Fig. 12, both the kaon and pion sources seem to be consistent with both the Gaussian and the dipole-shaped curves in the range from 4 fm out to about 12 fm. At the lowest r the pion image seems to split the difference between the two fits and the kaon image is below both fits. Both sources exhibit long non-Gaussian tails. In the pion source, this tail is higher than the dipole fit and in the kaon source, the tail is consistent with the dipole fit. For the p-Pb sources in Fig. 12, both the kaon and pion sources appear very nearly Gaussian. At the lowest r, the kaon source is a bit below the Gaussian fit.

One might ask if the tails in the sources are simply due to the nonspherical geometry of the full three-dimensional source. To answer this question, we compare the imaged sources to sources corresponding to NA44's three-dimensional fits in Ref. [30]. These correlations were acquired as follows: NA44 first measured like-charged pion and kaon correlations in the longitudinal co-moving system (LCMS), Coulomb corrected their data, and then fit their correlations to a Gaussian form. The form they fit to is

$$C(\mathbf{Q}) = 1 + \lambda \exp(-Q_{Ts}^2 R_{Ts}^2 - Q_{To}^2 R_{To}^2 - Q_L^2 R_L^2). \quad (27)$$

The results of these fits are shown in Table V for π^+'s and K^+'s. Now, as in the one-dimensional case, a three-

dimensional Gaussian correlation corresponds to a three-dimensional Gaussian source. For the correlation in Eq. (27), the corresponding source is

$$S(\mathbf{r}) = \frac{\lambda}{(2\sqrt{\pi})^3 R_{Ts} R_{To} R_L} \exp\left[-\frac{1}{4}\left(\frac{r_{Ts}^2}{R_{Ts}^2} + \frac{r_{To}^2}{R_{To}^2} + \frac{r_L^2}{R_L^2} \right) \right]. \quad (28)$$

Looking at the various transverse fit radii obtained by NA44, it seems safe to set $R_{Ts} = R_{To} \equiv R_T$ in this source. To convert this three-dimensional source to a one-dimensional source, we only need to angle average it, $S(r) = 1/4\pi \int d\Omega_{\mathbf{r}} S(\mathbf{r})$, to obtain

$$S(r) = \frac{\lambda}{16\pi R_T^2 R_L} \exp\left(-\frac{r^2}{4R_T^2} \right)$$

$$\times \begin{cases} \dfrac{2R_{eff}}{r} \mathrm{erf}\left(\dfrac{r}{2R_{eff}} \right) & \text{for } R_T > R_L \\[2ex] \dfrac{2R_{eff}}{r} \mathrm{erfi}\left(\dfrac{r}{2R_{eff}} \right) & \text{for } R_T < R_L, \end{cases} \quad (29)$$

where $R_{eff} = 1/\sqrt{|R_L^{-2} - R_T^{-2}|}$ and $\mathrm{erfi}(x) = i\,\mathrm{erf}(-ix)$. With this result, we are now able to compare the imaged sources to the angle-averaged three-dimensional fits from NA44. This is shown in Fig. 13.

For these fits, we used λ and R_L as given in Table V and R_T was chosen to be the average of R_{Ts} and R_{To}. Examining Fig. 13, it is clear that despite the fact that NA44's three-dimensional sources correspond to non-Gaussian angle-averaged sources, they are not able to account for the extended tails we find in the images. There is a caveat with this conclusion: their fits were performed in the LCMS and our source is imaged in the pair c.m. frame. Since the two frames coincide only in the limit that $p_T \to 0$, we computed the pion source using the low-p_T data set only. Only one data set is

TABLE III. Gaussian fit parameters as obtained by NA44.

	Ref.	λ	R_G (fm)	χ^2/NDF
K^+K^+ (S-Pb)	[7]	0.92 ± 0.08	3.22 ± 0.20	53/31
$\pi^+\pi^+$ (S-Pb)	[8]	0.46 ± 0.04	4.50 ± 0.31	18.1/16
K^+K^+ (p-Pb)	[7]	0.68 ± 0.06	1.71 ± 0.17	65/54
$\pi^+\pi^+$ (p-Pb)	[8]	0.38 ± 0.03	2.89 ± 0.30	16/25

TABLE IV. Exponential fit parameters as obtained by NA44.

	Ref.	λ	R_D (fm)	χ^2/NDF
K^+K^+ (S-Pb)	[7]	1.80 ± 0.18	2.64 ± 0.22	26/31
$\pi^+\pi^+$ (S-Pb)	[8]	0.77 ± 0.08	3.54 ± 0.33	12.0/16
K^+K^+ (p-Pb)	[7]	1.10 ± 0.10	1.04 ± 0.19	55/54

TABLE V. Three-dimensional Gaussian fit parameters from the S-Pb reaction as obtained by NA44.

	$\langle p_T \rangle$ (MeV/c)	λ	R_{Ts} (fm)	R_{To} (fm)	R_L (fm)
$\pi^+\pi^+$	150	0.56 ± 0.02	4.15 ± 0.27	4.02 ± 0.14	4.73 ± 0.26
$\pi^+\pi^+$	450	0.55 ± 0.02	2.95 ± 0.16	2.97 ± 0.16	3.09 ± 0.19
K^+K^+	240	0.82 ± 0.14	2.55 ± 0.20	2.77 ± 0.12	3.02 ± 0.20

available for the kaons. Nevertheless, it seems unlikely that the pion or kaon fit parameters would change dramatically if the fits were performed in the pair c.m.

C. Discussion of S-Pb data

In addition to the various fits to the correlations, members of the NA44 Collaboration performed RQMD simulations [7–9,13,29] of the S-Pb and p-Pb collisions. Rather than repeat this work, we will summarize it and explain its implications for the source shape. In all but the pp case, the simulated RQMD correlations compared favorably with the data. In the pp case, the simulations overestimate the height of the correlation peak by roughly 30%.

Sullivan *et al.* [13] explain that the width of the kaon correlation (corresponding to the width of the imaged kaon sources) is determined mainly by the size of the kaon production region. First, kaons are mainly produced directly in the reaction (either from fragmenting strings or hadronic reactions) or from the decay of K^* resonances. Now, the reaction zone is roughly the size of the sulfur nucleus (R_{rms} $=3.3$ fm) and this defines the size of the Gaussian core. The K^*'s are also produced in a region of roughly this size. However, since their lifetime is roughly $\tau\approx4$ fm/c, K^*'s do not travel far before decaying, giving rise to a non-Gaussian halo surrounding the Gaussian core. This halo is neither exponential nor Gaussian, but rather a convolution of a Gaussian source with an exponential [14,15]. Since the kaons have a much larger mean-free path in nuclear matter than either pions or nucleons (at least in RQMD), they do not rescatter as the system evolves. This means that their last interaction point is very nearly the size of this production zone.

The pion width in the S-Pb reaction is also described well by the RQMD calculations. The initially produced pions (produced mainly from hadronic reactions, although there is a string component) also have a source width set by a combination of the geometrical overlap of the colliding nuclei and subsequent dynamics of the system. From their production until they freeze-out, the pions interact strongly with the system. Because this evolution involves longitudinal flow, there are strong position-momentum correlations in the freeze-out positions of the pions. The longitudinal size of the region where one can find pions with a low relative momentum is comparable to the transverse size of the system at freeze-out giving a Gaussian core somewhat larger than the initial system size. In addition to this core, the pions also have a large contribution from the decay of various resonances, mainly the ρ, ω, η', and η^0. Now the η's are not capable of altering the pion source size or shape as the η's lifetimes are much too long. On the other hand, both the ρ (whose lifetime is ≈2 fm/c) and the ω (whose lifetime is ~23 fm/c) both contribute to a non-Gaussian halo in a manner analogous to the K^*'s above. Since the ρ's lifetime is smaller than the ω's, it contributes to the shorter distance part of the tail and the ω to the longer distance part.

In the case of the protons in the S-Pb reaction, the size of the source is mainly set by the size of the region where we find protons with low-relative momentum at freeze-out. Looking at the plots of RQMD correlations in [29], the RQMD simulations are somehow unable to reproduce the sizes of these regions. Since the size of this region is a direct function of the geometry of the source and the space-momentum correlations in the source (these correlations are caused primarily by flow), RQMD's failure here is somewhat of a mystery. Unfortunately, our image is not detailed enough to give much of a clue where the discrepancy arises. A higher resolution measurement of this correlation would help immensely.

D. Discussion of p-Pb data

Unlike the S-Pb reaction, RQMD is able to describe all of NA44's measured p-Pb correlations quite well [7,9,29]. Because the p-Pb system is so small, the reaction is most likely dominated by the formation of a few color strings and/or ropes. RQMD uses the Lund string model to model string formation and fragmentation [31], so to understand what is happening in the sources it is useful to understand some of
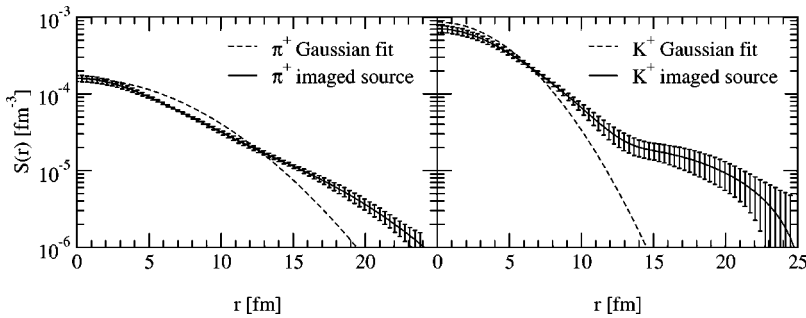
FIG. 13. Pion and kaon sources imaged from the S-Pb reactions compared to the NA44's angle-averaged three-dimensional sources.
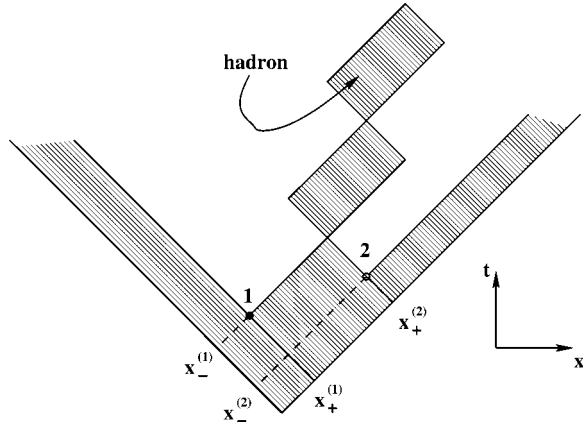
FIG. 14. String fragmentation in the Lund-string model. The hatched areas indicate regions of nonvanishing color field as well as the space-time region swept out by the oscillating strings. Points 1 and 2 indicate the space-time point where the hadron of interest breaks off of the main string.

the features of this type of model. In the Lund model, momentum and spatial rapidities are tightly correlated. We will show that this correlation leads to an approximate scaling of the K and p source radii with mass

$$m_K R_K \approx m_p R_p . \tag{30}$$

As we will point out, this scaling is borne out by the images. We mention that the pion source *should not* follow this scaling because a large fraction of the pions result from resonance decays that distort the shape of the pion source.

In a Lund-type string model [32], a meson is viewed as a $q\bar{q}$ pair attached by a string and this pair oscillates back and forth in the linear confining potential of the string. This type of model can also describe baryons if one replaces one of the quarks at the end of the string with a diquark. In this discussion, we ignore the transverse extent of the string and imagine the string lives in a (1+1)-dimensional space. As the $q\bar{q}$ pair oscillates, in one period it sweeps out an invariant area given by

$$\Delta x_+ \Delta x_- = \frac{m^2}{\sigma^2} \tag{31}$$

in light-cone coordinates. Here m is the mass of the hadron and σ is the string tension. A hadron breaking off from a string is pictured in Fig. 14.

To assess any correlations between the mass and the production location of a hadron, we examine the distribution of breakup points of the string that produces the hadron. In Fig. 14, points 1 and 2 are the locations where a $q\bar{q}$ pair separates from a string. Assuming a constant breakup probability per unit time per unit length P the distribution of breakup points is given by

$$dP \propto \exp(-P x_-^{(1)} x_+^{(2)}) \theta(x_-^{(1)} - x_-^{(2)}) \theta(x_+^{(2)} - x_+^{(1)}). \tag{32}$$

The exponential in this expression gives the probability that the string does not break up before the q and $\bar{q}$ are made at points 1 and 2 and the theta functions ensure the proper ordering of the coordinates. We define the creation point of the hadron of interest as the mean of the breakoff points of the $q\bar{q}$ pair

$$X_\pm = \frac{1}{2}(x_\pm^{(1)} + x_\pm^{(2)}). \tag{33}$$

We also write

$$\Delta x_\pm = \pm(x_\pm^{(2)} - x_\pm^{(1)}). \tag{34}$$

In terms of these, we may write Eq. (32) as

$$dP \propto \theta(\Delta x_-)\theta(\Delta x_+)\exp\left(-P\left[X_+ X_- + \frac{1}{4}\Delta x_+ \Delta x_- \right.\right.$$
$$\left.\left. + \frac{1}{2}(\Delta x_- X_+ + \Delta x_+ X_-)\right]\right). \tag{35}$$

For a hadron at midrapidity, from Eq. (31) we have $\Delta x_+ = \Delta x_- = m/\sqrt{2}\sigma$. Writing the hadron position back in Cartesian coordinates, we find

$$dP \propto \exp\left[-\frac{P}{2}\left(T^2 - X^2 + \frac{m}{\sigma}T\right)\right]\theta(T - |X|). \tag{36}$$

This last theta function makes causality explicit.

Since dP expresses the probability of creating a hadron at position X and at time T, it is proportional to the emission rate in Eq. (4). Thus, we can imagine doing the convolution in Eq. (3) to obtain the source size. Given that the emission rates in this case are not Gaussian, we should not expect the source function itself to be Gaussian. Nevertheless, we can still estimate the width of the source function from Eq. (36). We take the source width to be the distance at which the magnitude of the source function drops by $1/e$. From Eq. (36), we see that the probability of creating a hadron drops by $1/e$ by roughly

$$T \sim X \sim \frac{2\sigma}{mP}. \tag{37}$$

This implies that the source function itself will have a width that is correlated with the mass of the hadron and given by

$$R \sim \frac{2\sqrt{2}\sigma}{mP}. \tag{38}$$

Here, the factor of $\sqrt{2}$ arises from the convolution in Eq. (3). If we make the reasonable assumption that the string tension and breakup probability are both universal, than we have the scaling relation in Eq. (30). Looking at the kaon source in Fig. 12(d), the kaon source $1/e$ width is $R_K \approx 6$ fm. Following the scaling, the proton source should drop by $1/e$ by roughly $R_p \approx m_K / m_p R_K = 3$ fm. This is roughly borne out in the image in Fig. 12(f). With Eq. (38), we can go further than just checking this scaling and try to compute the source

width directly. Taking typical values for the string breakup probability, $P=1$ fm^2, and for the string tension, $\sigma=1$ GeV/fm, we find that the proton source radius should be about $R=2.9$ fm and the kaon source radius should be about $R=5.7$ fm, again in rough agreement with the images. It would be very interesting to see if this scaling persists in sources imaged from other like-pair correlation in p-A collisions such as $\bar{p}\bar{p}$, K_0K_0, or $\Lambda\Lambda$.

In this consideration, we have neglected several things. First, although we have ignored the transverse degrees of freedom of the strings, the longitudinal length of the production region is much larger than the transverse extent so it is the longitudinal extent that determines the angle-averaged source radius. Second, we have ignored the finite mass of the quarks. Adding finite quark masses would change the trajectories of the $q\bar{q}$ pair and the shape of the area swept out by the pair, but it should not change our conclusions appreciably. Third, we have neglected any consideration of other strings in the p-Pb interaction zone. Next, if there is more than one string produced in the reaction, the possibility for string fusion exists and it is neglected here. Finally, we have neglected the Bjorken-like position-momentum correlations along the string length. These correlations will narrow the source [16] and possibly account for the minor difference between the predicted and measured proton source widths above.

V. CONCLUSION

In this paper, we have investigated the possibility of detecting non-Gaussian sources in heavy-ion collisions. In simple, but realistic, model calculations we have demonstrated that it is possible to distinguish between Gaussian and non-Gaussian source shapes using an improved imaging method and high resolution data. Imaging not only has achieved results comparable to Gaussian fits, but has now uncovered deviations from Gaussian behavior.

This improved imaging method has several features that make it superior to the previous methods in [11,12]. First, this method uses basis splines to represent the source, giving a continuous representation of the source. The resolution of the images is controlled by the placement of the knots (the points where the polynomials that make up the spline basis are matched). Since the knots are analogous to the edges of the bins of a box-spline representation, we may do as in [12] and use the ''optimal knots'' to further improve the image. In addition to these improvements, equality constraints are now implemented in a simple manner. As in the previous imaging method, imaging is still a least-square problem in which the coefficients of the basis spline representation are chosen to minimize the χ^2 of the data. In other words, the imaging gives the source that has the highest probability of representing the correlation data. Finally, we mention that the amount of information available in the data limits the amount of information we may extract in form of image and constraints ''increase'' the amount of information.

Using this improved imaging method, we analyzed the proton, kaon, and pion correlations from S-Pb and p-Pb reactions measured by NA44. We find evidence for non-

Gaussian halos in the source functions from the like-meson correlations in the S-Pb reaction. These halos are likely to be due to resonances decaying and producing the mesons. RQMD model simulations of the correlations reproduce the experimental data quite well except in the case of the proton correlations from the S-Pb reaction. Unfortunately the source image does not shed much light on the discrepancy. In the case of the p-Pb reaction, the imaged sources suggest a scaling of the source widths that we should expect on the basis of Lund-type string models. This scaling should be tested by examining other like-pair correlations (e.g., $\bar{p}\bar{p}$ and $\Lambda\Lambda$) where these other pairs do not suffer from large resonance contributions.

APPENDIX A: BASIS SPLINE REPRESENTATION OF SOURCES

In this paper and previously in [12], we write the source expanded in some basis:

$$S(r)=\sum_{i=1}^{N_M} S_i B_i(r). \tag{A1}$$

In [12] this basis was the box-spline basis. In the box-spline basis, the widths of the individual bins could be varied to increase or decrease the resolution of the kernel to minimize the relative error of the source. Unfortunately, in this representation the source is not continuous.

In this work, we expand our sources in a more general basis: that of basis splines, also known as b-splines [22]. b splines are piecewise continuous polynomials that can be made arbitrarily smooth by changing the degree of the splines; the 0th degree b-splines are actually box-splines used before. The b-splines are characterized by a set of knots that mark the points where the various polynomials that make up the b-spline are matched. In a sense, these knots generalize the ''edges of the bins'' of the box-splines. For this reason, we may vary the locations of the knots to minimize the relative error of the source, generalizing the method in [12].

1. The basis splines

Now we define the b-splines. A N_b^{th} degree b-spline is characterized by a set of knots, $\{t_k\}$ with $t_1 \leqslant t_2 \leqslant \cdots$

$\leqslant t_{N_{knots}}$. The number of knots must be chosen so that $N_{knots} \geqslant N_M + N_b + 1$ for N_M b-splines. For $N_b = 0$, the b-splines are box-splines, i.e.,

$$B_{0,i}(r) = X_i(r) \equiv \begin{cases} 1 & \text{if } t_i < r < t_{i+1} \\ 0 & \text{otherwise.} \end{cases} \quad \text{(A2)}$$

Note, if $t_i = t_{i+1}$ then $B_{0,i}(r) = 0$. The rest of the b-splines may be constructed from this first one through a set of recurrence relations

$$\begin{aligned} B_{N_b+1,i}(r) &= w_{N_b+1,i}(r) B_{N_b,i}(r) \\ &\quad + [1 - w_{N_b+1,i+1}(r)] B_{N_b,i+1}(r), \quad \text{(A3)} \end{aligned}$$

where the weight factor is

$$w_{N_b+1,i}(r) = \begin{cases} \dfrac{r - t_i}{t_{i+N_b} - t_i} & \text{if } t_i \neq t_{i+N_b} \\ 0 & \text{otherwise.} \end{cases} \quad \text{(A4)}$$

With these definitions, we may write a b-spline of any degree back in terms of the box-splines

$$B_{N_b,i}(r) = \sum_{j=i}^{i+N_b-1} b_{N_b,j} X_j(r), \quad \text{(A5)}$$

where $b_{N_b,j}$ is a polynomial in r of degree N_b that we will not write explicitly. We plot sample b-splines of different degrees in Fig. 1.

There are three other properties of the b-splines that are of note. First, b-splines are normalized so that

$$\int_{-\infty}^{\infty} dr B_i(r) = \frac{t_{i+N_b+1} - t_i}{N_b + 1}. \quad \text{(A6)}$$

Second, the ith b-spline is zero outside of the region $t_i \leqslant r \leqslant t_{i+N_b+1}$. Third, the b-splines are not orthogonal but expressions for their inner product exist [22].

From the definitions and from the figure, it is not clear how to pick the knots. The knots do not have to be equally spaced and, in many situations, it is best *not* to space them equally. In fact, one can even pile up to $N_b + 1$ knots on the same point. One can do this because $N_b + 1 =$ (number of knots at a point) + (number of continuity conditions at that point).

If a b-spline has $N_b + 1$ knots at a point then that b-spline is discontinuous there. Away from these multiple knots, the b-spline is still continuous up to derivatives of degree N_b. In the main text, we remove all assumptions about the continuity of the source at the boundaries by keeping $N_b + 1$ knots at the boundaries. We then optionally reinsert continuity conditions using equality constraints.

2. Optimizing resolution

Now we come to the point of choosing the best knots for the inversion, generalizing the ''optimized discretization''

scheme of Ref. [12]. First, the model covariance matrix [cf. Eq. (B7)] depends on the kernel of the inversion, the error on the data and whatever scheme we use to represent the source, *but not on the data or the model itself.* For a given kernel and set of data errors, we are free to change our representation of the source in order to minimize the error of the source. In particular, we may vary the location of the knots (at least not the knots fixed at the end points of the imaging region) to minimize the error of the source coefficients, $\Delta S_j = \sqrt{\Delta^2 S_{jj}}$, relative to some dummy source,

$$\sum_{j=N_b+2}^{N_{knots}-N_b-1} \left| \frac{\Delta S_j}{S_j^{dummy}} \right| = \min. \quad \text{(A7)}$$

The coefficients S_j^{dummy} are the coefficients of the expansion of a dummy source in b-splines. In this minimization, the first and last $N_b + 1$ knots are held fixed and the positions of all of the other knots are varied. The dummy source itself is chosen to be big roughly where one expects the source to be big and small where one expects the source to be small. Since the detailed shape of the dummy source should not be important, in this paper we chose an exponential dummy source with radius $R^{dummy} = 3.5$ fm given by $S^{dummy}(r) \propto \exp(-r/R^{dummy})$.

APPENDIX B: BAYESIAN APPROACH TO IMAGING

In this appendix, we will explain the technical details of the Bayesian approach to imaging and extend the approach of Ref. [12] by implementing constraints in a more consistent manner. In the previous works the constraints are implemented by Monte Carlo sampling the experimental errors, leading to statistical fluctuations in the extracted source. In that approach, no distinction is made between equality and inequality constraints. By equality constraints we mean constraints of the form

$$\int d\mathbf{r} f(\mathbf{r}) S_{\mathbf{P}}(\mathbf{r}) = \text{constant}, \quad \text{(B1)}$$

and by inequality constraints, we mean constraints of the form

$$\int d\mathbf{r} f(\mathbf{r}) S_{\mathbf{P}}(\mathbf{r}) \geqslant \text{constant}. \quad \text{(B2)}$$

Both of these types of constraints are forms of *prior information*, meaning information we have in hand before we began imaging. In this appendix, we will explain how to use prior information in an inversion. In particular, we will discuss two methods for implementing equality constraints *directly* into the inversion and we will discuss a better way to implement inequality constraints. We conclude this appendix with a brief discussion of inequality constraints.

1. General theory

Suppose we have an N_D dimensional vector of observed data $\mathbf{d}^{obs}$ with covariance matrix $\Delta^2 d$ that we wish to represent by some N_M-dimensional model vector $\mathbf{m}$. In the case of

source imaging, our model vector is the vector of coefficients of the function expansion of the source and the data are the raw correlation functions. We assume the data has a diagonal covariance matrix. In an ideal measurement of the data, there would be no experimental uncertainty and the data and the model would be related through some linear equation

$$\mathbf{d} = K \cdot \mathbf{m}. \tag{B3}$$

Here, K is the kernel of an integral equation.

In a real imaging problem, the observed data has errors and statistical scatter. To make progress, we adopt the so-called Bayesian approach to imaging where we seek the probability density, $\sigma(\mathbf{m})$, for a specific model $\mathbf{m}$ to represent the data [23,33]. With this density, we take the mean of the density as an estimator of the true model and the width of the density as an estimate of the uncertainty in our model. Neglecting the error in our determination of the kernel and assuming the uncertainty in our measurement of $\mathbf{d}^{obs}$ is Gaussian, we can write Bayes theorem as follows:

$$\sigma(\mathbf{m}) \propto \rho(\mathbf{m}) \exp\left(-\frac{1}{2}\chi^2_{data}\right). \tag{B4}$$

Here, $\rho(\mathbf{m})$ encodes all of the prior information we have about the model and χ^2_{data} is

$$\chi^2_{data} = (K \cdot \mathbf{m} - \mathbf{d}^{obs})^T \cdot (\Delta^2 d)^{-1} \cdot (K \cdot \mathbf{m} - \mathbf{d}^{obs}). \tag{B5}$$

Here the superscript T represents a matrix transpose. The dimension of the model vector N_M and the dimension of the data vector N_D need not be equal. Indeed, it is better to have many more data points than model parameters so that we may overconstrain the system. For the time being, assume we have no prior information so we may set $\rho(\mathbf{m}) = 1$.

We immediately see that the most probable model vector is the one that maximizes the probability and hence minimizes the χ^2_{data}. Following Refs. [11,12,23,24,26,33], we can find both the model vector that minimizes this χ^2_{data} as well as the covariance matrix of the model

$$\langle \mathbf{m} \rangle = \Delta^2 m \cdot K^T \cdot (\Delta^2 d)^{-1} \cdot \mathbf{d}^{obs} \tag{B6}$$

and

$$\Delta^2 m = [K^T \cdot (\Delta^2 d)^{-1} \cdot K]^{-1}. \tag{B7}$$

Equation (B6) usually goes by the name of a normal equation. The model covariance matrix is independent of $\mathbf{d}^{obs}$ and depends only on the error of the data and the kernel itself.

We can write the χ^2_{data} directly in terms of the model covariance matrix and the χ^2_{data} minimizing model vector [23],

$$\chi^2_{data} = (\mathbf{m} - \langle \mathbf{m} \rangle)^T \cdot (\Delta^2 m)^{-1} \cdot (\mathbf{m} - \langle \mathbf{m} \rangle). \tag{B8}$$

In other words, if we have a N_M-dimensional model space, then the $\chi^2_{data} = 1$ hypersurface is an N_M-dimensional hyper-ellipsoid with principal axes given by the eigenvectors of the covariance matrix of the model, $\Delta^2 m$. These principal axes

do not need to correspond to the directions corresponding to the m_j components in the model space.

2. Equality constraints

Now let us discuss the role of prior information in the general inversion problem. We concentrate equality constraints such as in Eq. (B1). In a matrix form, equality constraints are written as

$$\mathcal{C} \cdot \mathbf{m} = \mathbf{c}. \tag{B9}$$

Here $\mathcal{C}$ is a matrix of constraint equations and $\mathbf{c}$ is a constant vector of constraint values. In the inversion problem in the main text, there are a variety of constraints we might use and they are listed in Table I. The equality constraint in Eq. (B9) corresponds to the prior probability density of

$$\rho(\mathbf{m}) \propto \delta(\mathcal{C} \cdot \mathbf{m} - \mathbf{c}). \tag{B10}$$

Equality constraints can be cast into a Gaussian prior probability density simply by writing this density as a Gaussian with vanishing width:

$$\rho(\mathbf{m}) \propto \lim_{\lambda \to \infty} \exp\left(-\frac{\lambda}{2}\chi^2_{equal}\right), \tag{B11}$$

where

$$\chi^2_{equal} = (\mathcal{C} \cdot \mathbf{m} - \mathbf{c})^2. \tag{B12}$$

Finding the most probable model then corresponds to minimizing a modified $\widetilde{\chi}^2$:

$$\widetilde{\chi}^2 = \chi^2_{data} + \lambda \chi^2_{equal}. \tag{B13}$$

The solution is straightforward and corresponds to the most probable model,

$$\langle \mathbf{m} \rangle = \Delta^2 m \cdot (K^T \cdot (\Delta^2 d)^{-1} \cdot \mathbf{d}^{obs} + \lambda \mathcal{C}^T \cdot \mathbf{c}) \tag{B14}$$

along with the model covariance matrix,

$$\Delta^2 m = [K^T \cdot (\Delta^2 d)^{-1} \cdot K + \lambda \mathcal{C}^T \cdot \mathcal{C}]^{-1}. \tag{B15}$$

It is clear that to correctly simulate a δ function prior information density, we must choose a large λ. Looking at Eq. (B13), we must choose a λ so that $\lambda \chi^2_{equal} \gg \chi^2_{data}$. We now estimate the sizes of χ^2_{equal} and χ^2_{data}. A good fit to the data should have the χ^2_{data} nearly at the number of degrees of freedom, i.e. $\chi^2_{data} \approx N_D - N_M$. To estimate χ^2_{equal}, we follow Eq. (B12). In the main text, a typical source is $\gtrsim 1 \times 10^{-6}$ fm^{-3} and the constraint matrix and vectors are typically ~ 1 fm^3, so $\chi^2_{equal} \sim N_M \times 10^{-12}$. Putting this together, we must choose

$$\lambda \gg \chi^2_{data}/\chi^2_{equal} \sim \left(\frac{N_D}{N_M} - 1\right) \times 10^{12}. \tag{B16}$$

For example, for $N_D = 83$ and $N_M = 8$ we need $\lambda \gg 10^{11}$. By adjusting strength of λ, we can adjust strength of various
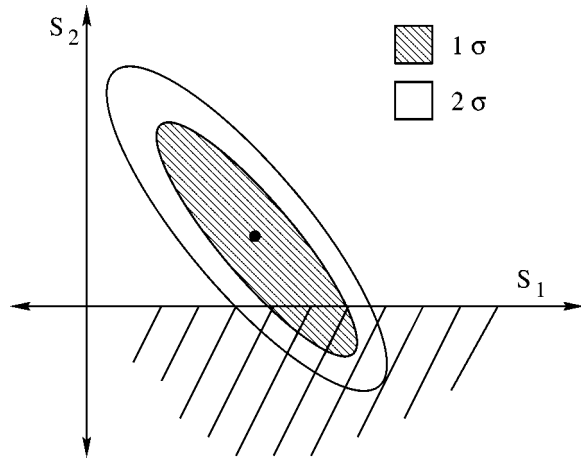
FIG. 15. An illustration of an inequality constraint cutting through 1σ band of a best-fit region.

terms, emphasizing stability of inversion (i.e., obeying constraints) over representing the data. Thus, λ functions as a trade-off parameter in the jargon of inverse theory. See Ref. [26] Sec. 18.4 for a more complete discussion.

A useful alternative to this scheme (and a way to do the $\lambda \rightarrow \infty$ limit exactly) is to use the Householder transformation to eliminate the constraints from the unmodified normal equations of Eqs. (B6) [26,34]. The tradeoff is that the Householder transformation may be somewhat unforgiving. Due to an unfortunate choice of basis functions, it may not be possible to satisfy two constraints simultaneously even if they can be satisfied simultaneously in the true answer. By keeping λ finite, we are never trying to satisfy the constraints exactly so we can do a reasonable job of obeying both constraints. Nevertheless, schemes based on Householder reductions of the constraints are complementary to ones using the Gaussian prior probability.

3. Inequality constraints

Now we ask how to use constraints of the form

$$\mathcal{C}\cdot\mathbf{m}\geq\mathbf{c}. \tag{B17}$$

Such constraints are called inequality constraints and there are many different ones we could use: the source is positive

definite, the derivative of the source is bounded (to ensure smoothness), or the source satisfies the Fourier transform test from Ref. [10]. Inequality constraints correspond to prior probability densities of the form

$$\rho(\mathbf{m})\propto\theta(\mathcal{C}\cdot\mathbf{m}-\mathbf{c}) \tag{B18}$$

which cannot be rendered into a Gaussian form.

In Ref. [11], we use a simple Monte Carlo sampling scheme to implement inequality constraints. In this scheme, one uses the experimental uncertainty to generate an ensemble of correlations, each consistent with the original. One then inverts each one to obtain a sample source and discards any sources that are not consistent with the inequality constraints. One then combines the samples that are consistent with the constraints to obtain an average source and an estimate of the errors on the source. The problem with this scheme is that it pushes the sources away from edges of the model space defined by the constraints.

We illustrate this problem with a simple example. Suppose we have an inversion problem where the goal is to determine two points S_1 and S_2 under the constraint that $S_2>0$. We sketch one possible outcome of the inversion in Fig. 15. In this picture, we see the best-fit value of S_1 and S_2 is consistent with our inequality constraint, but the constraint cuts through both the 1σ and 2σ contours. Using the Monte Carlo sampling scheme discussed above, we would actually be finding a false best-fit point which is slightly above and to the left of the true best-fit value because we throw out samples with $S_2<0$. The errors on these points would also be symmetrically placed around this point. In fact the correct way to solve the problem is just to quote the best-fit values of S_1 and S_2, with asymmetrical errors.

The way inequality constraints are implemented in most commercial inversion packages is through so-called ''active set methods'' [35]. In these methods, one finds the best-fit solution as one normally would have if there were no inequality constraints. If the best-fit solution lies in a region excluded by the inequality constraints, then the code finds the edge of the included region (the so-called active set) and searches along it until the code finds the solution that minimizes the χ^2. Such a scheme is powerful, but likely beyond what is needed for our problem.

[1] V. G. Grishin, G. I. Kopylov, and M. I. Podgoretski, Sov. J. Nucl. Phys. **13**, 638 (1971).

[2] S. Pratt, T. Csörgő, and T. Zimányi, Phys. Rev. C **42**, 2646 (1990).

[3] C. Gelbke and B. K. Jennings, Rev. Mod. Phys. **62**, 553 (1990).

[4] S. Pratt, Nucl. Phys. **A638**, 125c (1998).

[5] U. A. Wiedemann and U. Heinz, Phys. Rep. **319**, 145 (1999).

[6] U. Heinz and B. Jacak, Annu. Rev. Nucl. Part. Sci. **49**, Palo Alto, CA (1999).

[7] H. Beker *et al.*, Z. Phys. C **64**, 209 (1994).

[8] H. Bøggild *et al.*, Phys. Lett. B **302**, 510 (1993).

[9] H. Bøggild *et al.*, Phys. Lett. B **349**, 386 (1995).

[10] D. A. Brown, F. Wang, and P. Danielewicz, Phys. Lett. B **470**, 33 (1999).

[11] D. A. Brown and P. Danielewicz, Phys. Lett. B **398**, 252 (1997).

[12] D. A. Brown and P. Danielewicz, Phys. Rev. C **57**, 2474 (1998).

[13] J. P. Sullivan *et al.*, Phys. Rev. Lett. **70**, 3000 (1993).

[14] S. Nickerson, T. Csörgő, and D. Kiang, Phys. Rev. C **57**, 3251 (1998).

[15] T. Csörgő, B. Lorstad, J. Schmid-Sorensen, and A. Ster, Eur. Phys. J. C **9**, 275 (1999).

[16] S. Y. Panitkin and D. A. Brown, Phys. Rev. C **61**, 021901(R) (2000).

[17] B. Andersson and W. Hofmann, Phys. Lett. **169B**, 364 (1986).

[18] S. E. Koonin, Phys. Lett. **70B**, 43 (1977).

[19] W. G. J. Stoks, *et al.*, Phys. Rev. C **49**, 2950 (1994).

[20] T. Csörgő and A. T. Szerző, preprint, hep-ph/9912220.

[21] U. A. Wiedemann and U. Heinz, Phys. Rev. C **56**, 610 (1997).

[22] C. de Boor, *A Practical Guide to Splines* (Springer-Verlag, New York, 1978); MRC 2952 (1986) in *Fundamental Developments of Computer-Aided Geometric Modeling*, edited by L. Piegl (Academic Press, London, 1993).

[23] A. Tarantola, *Inverse Problem Theory* (Elsevier, New York, 1987).

[24] R. Parker, *Geophysical Inverse Theory* (Princeton University Press, Princeton, NJ, 1994).

[25] A. N. Tikhonov, Sov. Math. Dokl **4**, 1035 (1963).

[26] W. H. Press *et al.*, *Numerical Recipes in C* (Cambridge University Press, New York, 1992).

[27] W. G. Gong *et al.*, Phys. Rev. Lett. **65**, 2114 (1990); Phys. Rev. C **43**, 1804 (1991).

[28] D. A. Brown, S. Y. Panitkin, and G. Bertsch, Phys. Rev. C **62**, 014904 (2000).

[29] H. Böggild *et al.*, Phys. Lett. B **458**, 181 (1999).

[30] H. Beker *et al.*, Phys. Rev. Lett. **74**, 3340 (1995).

[31] H. Sorge, Phys. Rev. C **52**, 3291 (1995).

[32] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Phys. Rep. **97**, 31 (1983).

[33] W. P. Gouveia and J. A. Scales, Inverse Probl. **13**, 323 (1997).

[34] CERNLIB short writeups, http://wwwinfo.cern.ch/asd/cernlib/overview.html

[35] NEOSguide, http://www-fp.mcs.anl.gov/otc/Guide/