Using state space differential geometry for nonlinear blind source separation

David N. Levin[a)]
Department of Radiology, University of Chicago, Chicago, Illinois 60637, USA

(Received 11 December 2006; accepted 24 October 2007; published online 25 February 2008)

Given a time series of multicomponent data, the usual objective of nonlinear blind source separation (BSS) is to find a "source" time series, comprised of statistically independent combinations of the measured components. In this paper, we seek a source time series that has a phase-space density function equal to the product of density functions of individual components. In an earlier paper, it was shown that the phase space density function induces a Riemannian geometry on the data's state space with the metric equal to the local velocity correlation matrix of the data. From this geometric perspective, the vanishing of the curvature tensor is a necessary condition for BSS. If the curvature tensor is zero, there is only one possible set of source variables (up to permutations and transformations of individual components), and it is possible to compute these explicitly and to determine if they do have a factorizable phase space density function. The method is illustrated by using it to separate two simultaneous synthetic "utterances" recorded with a single microphone. A more general method that performs nonlinear multidimensional BSS is described in Appendix A, where it is illustrated with analytic and numerical examples. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2826943]

I. INTRODUCTION

Consider a set of data consisting of $\tilde{x}(t)$, a time-dependent multiplet of n measurements ($\tilde{x}_k$ for $k=1,2,\ldots,n$). The usual objectives of nonlinear blind source separation (BSS) are: (1) to determine if these observations are instantaneous mixtures of n statistically independent source components $x(t)$,

$$\tilde{x}(t) = f[x(t)], \tag{1}$$

where f is an unknown, possibly nonlinear, n-component mixing function, and, if so, (2) to compute the mixing function. In most approaches to this problem,[1,2] the desired source components are required to be statistically independent in the sense that their density function $\rho(x)$ is the product of the density functions of the individual components

$$\rho(x) = \prod_{k=1}^{n} \rho_k(x_k). \tag{2}$$

However, it is well known that this problem always has many solutions (see Ref. 3 and references therein). Specifically, any observed density function can be integrated in order to construct an entire family of functions f^{-1} that transform it into a separable (i.e., factorizable) form.

The observed trajectories of many classical physical systems can be characterized by density functions in *phase space* [i.e., $(\tilde{x},\dot{\tilde{x}})$ space]. Furthermore, if such a system is composed of noninteracting subsystems, the state space variables $\tilde{x}$ can be transformed so that the system's phase space density function is separable (i.e., is the product of the phase space density functions of the subsystems)

$$\rho(x,\dot{x}) = \prod_{k=1}^{n} \rho_k(x_k,\dot{x}_k). \tag{3}$$

This fact motivates the approach to BSS described in this paper:[4,5] we search for a function of the observations $\tilde{x}(t)$ that transforms their phase space density function $\tilde{\rho}(\tilde{x},\dot{\tilde{x}})$ into a separable form. It was previously demonstrated[6] that the phase space density function of a time series induces a Riemannian geometry on the data's state space and that its metric can be directly computed from the local velocity correlation matrix of the data. In the following section, we show how this differential geometry can be used to solve the BSS problem. Unlike conventional BSS, this "phase space BSS problem" has a unique solution in the following sense: either the data are inseparable, or they can be separated by a mixing function that is unique, up to transformations that do not affect separability (permutations and possibly nonlinear transformations of individual source components). This form of the BSS problem has a unique solution because separability in phase space is a stronger requirement than separability in state space.

In addition to using a stronger criterion of statistical independence, the proposed method differs from earlier approach on the technical level. For example, the proposed method exploits statistical constraints on source time derivatives that are *locally* defined in the state space, in contrast to the usual criteria for statistical independence that are *global* conditions on the source time series or its time derivatives.[7] Furthermore, the nonlinearities of the mixing function are unraveled by imposition of local second-order statistical constraints, unlike many conventional approaches that rely on higher-order statistics.[1,2] In addition, the constraints of statistical independence are used to construct the mixing function in a "deterministic" manner, without the need for parameter-

[a)]Electronic mail: d-levin@uchicago.edu.

izing it (with a neural network architecture or other means), without employing iterative methods, and without using probabilistic learning techniques.[8,9] And, unlike some other techniques that only apply to a restricted class of mixing functions,[10] the proposed method can handle any differentiable invertible mixing function. Finally, the use of differential geometry in this paper should not be confused with existing applications of differential geometry to BSS. In our case, the observed data trajectory is used to derive a metric on the system's *state space*, and the vanishing of the curvature tensor is shown to be a necessary condition for separability of the data. In contrast, other authors[11] define a metric on a completely different space, *the search space of possible mixing functions*, so that "natural" (i.e., covariant) differentiation can be used to expedite the search for the function that optimizes the fit to the observed data.

This paper is organized in the following manner. The next section describes the theoretical framework of the differential geometric approach to the blind source separation problem. Section III illustrates the method by using it to separate two simultaneous synthetic utterances that are recorded with a single microphone. The implications and limitations of this work are discussed in the last section. Appendix A 1 describes a generalization of the procedure in Sec. II, which can be used to perform multidimensional blind source separation (sometimes called multidimensional independent component analysis[12] or independent subspace separation[13]). Here, the objective is to separate the data into statistically independent *groups* of source variables, each of which may contain statistically dependent components. This method is illustrated with analytic and numerical examples in Appendices A 2 and A 3.

II. THEORY

This section describes how to determine if data are an instantaneous, possibly nonlinear mixture of statistically independent source variables, and, if so, how to compute the mixing function. The resulting mixing function is unique, up to transformations that do not affect separability (permutations and component-wise transformations). Figure 1 is a diagram of the procedure for achieving these objectives.

Let $x=x(t)$ (x_k for $k=1,2,\ldots,n$) denote the trajectory of a time series. Suppose that the trajectory densely covers a patch of $(x,\dot{x})$ space (i.e., phase space) and suppose that there is a phase space density function $\rho(x,\dot{x})$, which measures the fraction of total time that the trajectory spends in each small neighborhood $dxd\dot{x}$. As discussed in Appendix A 2, if a classical physical system in thermal equilibrium with a "bath" is observed by measuring the temporal evolution of its configuration, the measurement trajectory will have such a phase space density function: namely, the Maxwell–Boltzmann distribution.[14] Next, define $g^{kl}(x)$ to be the local second-order velocity correlation matrix[6]

$$g^{kl}(x) = \langle (\dot{x}_k - \bar{\dot{x}}_k)(\dot{x}_l - \bar{\dot{x}}_l) \rangle_x, \tag{4}$$

where the bracket denotes the time average over the trajectory's segments in a small neighborhood of x and where $\bar{\dot{x}} = \langle \dot{x} \rangle_x$, the local time average of $\dot{x}$. This means that g^{kl} is a combination of first and second moments of the local velocity distribution described by ρ. Because this correlation matrix transforms as a symmetric contravariant tensor, it can be taken to be a contravariant metric on the state space. Furthermore, as long as the local velocity distribution is not confined to a hyperplane in velocity space, this tensor is positive definite and can be inverted to form the corresponding covariant metric g_{kl}. Thus, under these conditions, the time series induces a nonsingular metric on state space. This metric can then be differentiated to compute the affine connection $\Gamma^k_{lm}(x)$ and Riemann–Christoffel curvature tensor $R^k{}_{lmn}(x)$ of state space by means of the standard formulas of differential geometry [Eqs. (A7) and (A8)].

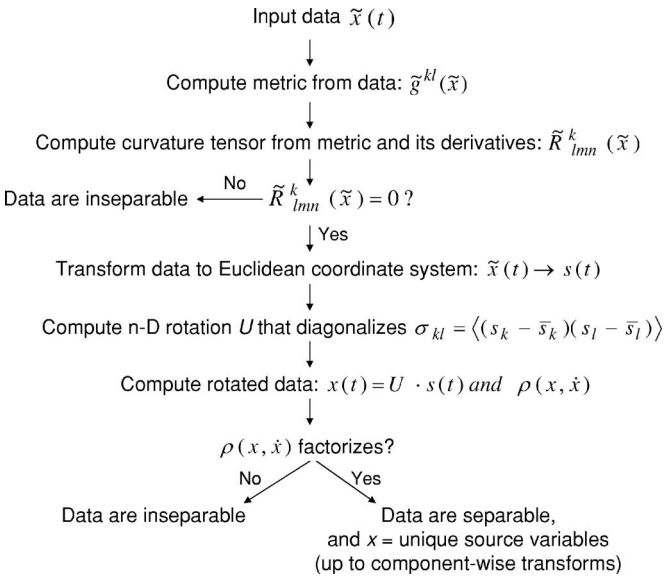

FIG. 1. Differential geometric method for one-dimensional BSS.

Now, assume that the data are separable; i.e., assume that there is a set of source variables x for which the phase space density function ρ is equal to the product of density functions of each component of x, as in Eq. (3). It follows from Eq. (4) that the metric $g^{kl}(x)$ is diagonal and has positive diagonal elements, each of which is a function of the corresponding coordinate component. Therefore, the individual components of x can be transformed in order to create a new "Euclidean" state space coordinate system in which the metric is the identity matrix and the curvature tensor vanishes everywhere in state space. It follows that the curvature tensor must vanish in every coordinate system, including the coordinate system $\tilde{x}$ defined by the observed data

$$\tilde{R}^k{}_{lmn}(\tilde{x}) = 0. \tag{5}$$

In other words, the vanishing of the curvature tensor is a necessary consequence of separability. Therefore, if this data-derived quantity does not vanish, the data cannot be transformed so that their phase space density function factorizes.

On the other hand, if the data do satisfy Eq. (5), there is only one possible separable coordinate system (up to transformations that do not affect separability), and it can be explicitly constructed from the observed data $\tilde{x}(t)$. To see this, assume that the data satisfy Eq. (5), and note two properties

of such a flat manifold with a positive definite metric: (1) it is always possible to explicitly construct a Euclidean coordinate system for which the metric is the identity matrix and (2) if a coordinate system has a diagonal metric with positive diagonal elements that are functions of the corresponding coordinate components, it can be derived from this Euclidean one by means of an n-dimensional rotation, followed by transformations that do not affect separability (permutations and transformations of individual components). Therefore, because every separable coordinate system must have a diagonal metric with the aforementioned properties, all possible separable coordinate systems can be found by constructing a Euclidean coordinate system and then finding all rotations of it that are separable. The first step is to construct a Euclidean coordinate system in the following manner: at some arbitrarily chosen point $\tilde{x}_0$, select n vectors $\delta\tilde{x}_{(i)}$ $(i=1,2,\ldots,n)$ that are orthogonal with respect to the metric at that point (i.e., $\tilde{g}_{kl}(\tilde{x}_0)\delta\tilde{x}_{(i)k}\delta\tilde{x}_{(j)l}=\lambda^2\delta_{ij}$, where δ_{ij} is the Kronecker delta, λ is a small number, and repeated indices are summed). Then, (1) starting at $\tilde{x}_0$, use the affine connection to repeatedly parallel transfer all $\delta\tilde{x}$ along $\delta\tilde{x}_{(1)}$; (2) starting at each point along the resulting geodesic path, repeatedly parallel transfer these vectors along $\delta\tilde{x}_{(2)}$; … continue the parallel transfer process along directions $\delta\tilde{x}_{(3)}\ldots\delta\tilde{x}_{(n-1)}\ldots$; (n) starting at each point along the most recently produced geodesic path, parallel transfer these vectors along $\delta\tilde{x}_{(n)}$. Finally, each point is assigned the geodesic coordinate s $(s_k,k=1,2,\ldots,n)$, where s_k represents the number of parallel transfers of the vector $\delta\tilde{x}_{(k)}$ that was required to reach it. Given Eq. (5), differential geometry[15] guarantees that the metric will be a multiple of the identity matrix in the geodesic coordinate system constructed in this way. We can now transform the data into the corresponding Euclidean coordinate system and examine the separability of all possible rotations of it. The easiest way to do this is to compute the *global* second-order correlation matrix

$$\sigma_{kl}=\langle(s_k-\bar{s}_k)(s_l-\bar{s}_l)\rangle, \tag{6}$$

where the brackets denote the time average over the entire trajectory and $\bar{s}=\langle s\rangle$. If this data-derived matrix is not degenerate, there is a unique rotation U that diagonalizes it, and the corresponding rotation of the s coordinate system, $x=Us$, is the only candidate for a separable coordinate system (up to transformations that do not affect separability). In principle, the separability of the data in this rotated coordinate system can be determined by explicitly computing the data's phase space density function in order to see if it factorizes. Alternatively, if the amount of data is insufficient to accurately calculate the phase space density function of $x(t)$, the statistical independence of the x coordinates can be assessed by determining if higher-order correlations of x and $\dot{x}$ factorize into products of lower-order correlations, as required by Eq. (3).

III. EXAMPLE: SEPARATING SIMULTANEOUS SYNTHETIC "UTTERANCES" RECORDED WITH A SINGLE MICROPHONE

This section describes a numerical experiment in which two sounds were synthesized and then summed, as if they occurred simultaneously and were recorded with a single microphone. Each sound simulated an "utterance" of a vocal tract resembling a human vocal tract, except that it had fewer degrees of freedom (one degree of freedom instead of the 3–5 degrees of freedom of the human vocal tract[16,17]). The methodology described in Sec. II was blindly applied to the synthetic recording, in order to recover the time dependence of the state variable of each vocal tract (up to an unknown transformation on each voice's state space) and to synthesize a voice-converted version of each vocal tract's utterance.

The glottal wave forms of the two "voices" had different pitches (97 and 205 Hz), and the "vocal tract" response of each voice was characterized by a damped sinusoid, whose amplitude, frequency, and damping constant were linear functions of that voice's state variable.[18] For example, the resonant frequency of one voice's vocal tract varied linearly between 300 and 900 Hz as that voice's state variable varied on the interval $[-1,+1]$. For each voice, a 10 h utterance was produced by using glottal impulses to drive the vocal tract's response, which was determined by the time-dependent state variable of that vocal tract. The state variable time series of each voice was synthesized by smoothly interpolating among successive states randomly chosen at 100–120 ms intervals. The resulting utterances had energies differing by 0.7 dB, and they were summed and sampled at 16 kHz with 16-bit depth. Then, this "recorded" wave form was subjected to a short-term Fourier transform (using frames with 25 ms length and 5 ms spacing). The log energies of a bank of 20 mel-frequency filters between 0 and 8000 Hz were computed for each frame, and these were then averaged over each set of four consecutive frames. These log filter bank outputs were nonlinear functions of the two vocal tract state variables, which were statistically independent of each other.[18]

The remainder of this section describes how these data were analyzed in a completely blind fashion in order to discover the presence of two underlying independent source variables and to recover their time courses. In other words, the filter bank outputs were processed as a time series of numbers of unknown origin, without using any of the information in the preceding paragraph. For example, the analysis did not make use of the fact that the data were produced by driven resonant systems.

The first step was to determine if any data components were redundant in the sense that they were simply functions of other components. Figure 2(a) shows the first three principal components of the data during a typical recorded segment of the simultaneous utterances. Inspection showed that these data lay on a two-dimensional surface within the ambient 20-dimensional space, making it apparent that they were produced by an underlying system with two degrees of freedom. The redundant components were eliminated by using dimensional reduction[19] to establish a coordinate system $\tilde{x}$ ($\tilde{x}_k$ for $k=1,2$) on this surface and to find the trajectory of

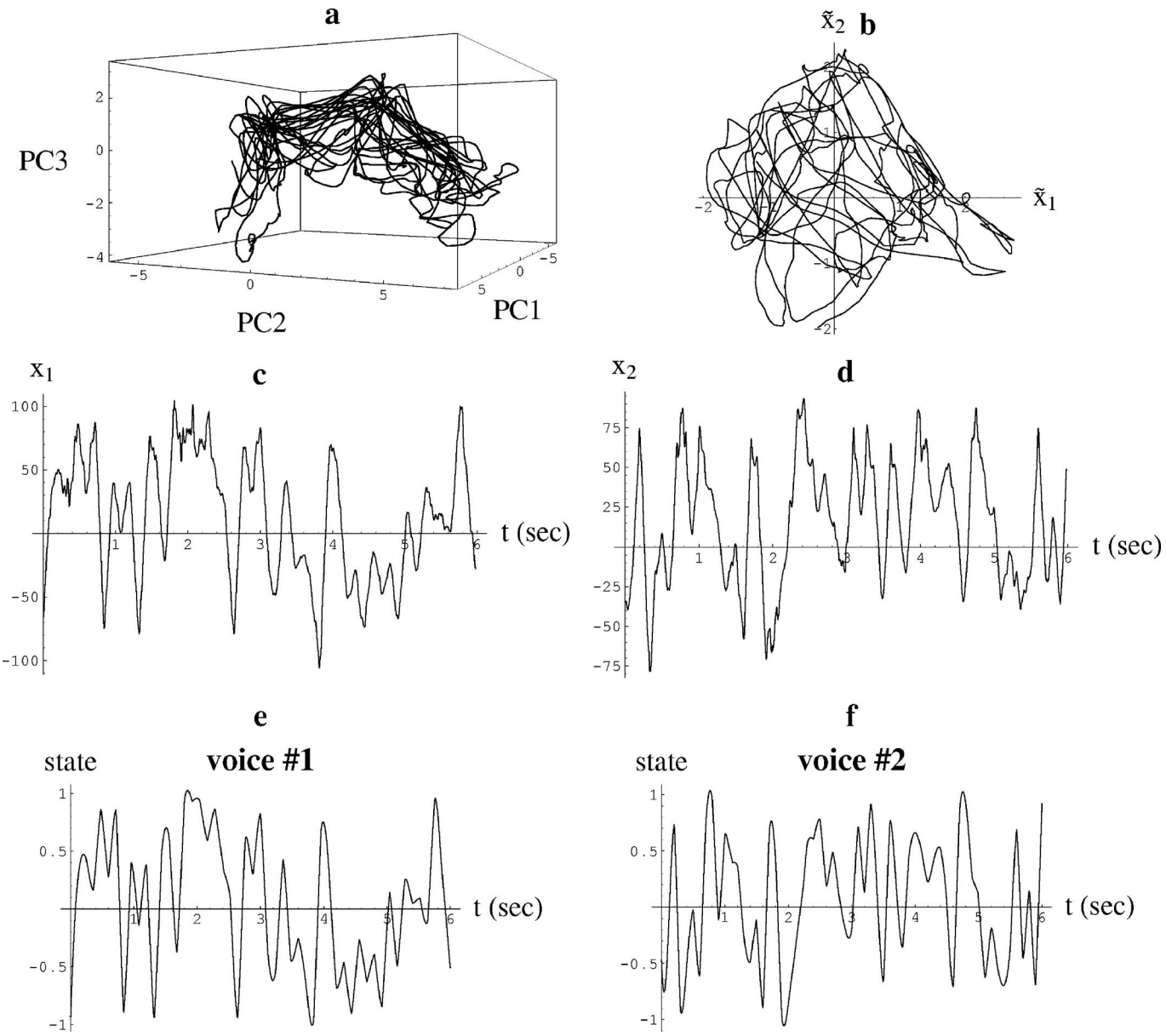

FIG. 2. A typical six-second segment of the synthetic recording. (a) The first three principal components of the recording's log filter bank outputs. (b) The trajectory in (a) after dimensional reduction. (c) and (d) The statistically independent source time series blindly derived from (b). (e) and (f) The state variable time series used to synthesize the utterances of the two voices.

the recorded sound $\tilde{x}(t)$ in that coordinate system [Fig. 2(b)]. In effect, $\tilde{x}(t)$ represents a relatively low bandwidth two-dimensional signal that was "hidden" within the higher bandwidth wave form recorded with the simulated microphone. The next step was to determine if the components of $\tilde{x}(t)$ were nonlinear mixtures of two source variables that were statistically independent of one another, in the sense that they had a factorizable phase space density function. Following the procedure in Sec. II, $\tilde{x}(t)$ of the entire recording was used to compute the metric in Eq. (4) on a 32×32 grid in $\tilde{x}$ space, and the result was differentiated to compute the affine connection and curvature tensor there. The values of the latter were distributed around zero, suggesting that the state space was flat [as in Eq. (5)], which is a necessary condition for the separability of the data. The procedure in Sec. II was then followed to transform the data into a Euclidean coordinate system s, and the resulting trajectory $s(t)$ was substituted into Eq. (6) to compute the state variable correlation matrix. Finally, the rotation that diagonalized this matrix was used to rotate the s coordinate system, thereby producing the x coordinate system, which was the only possible separable coordinate system (up to transformations of individual components). Figures 2(c)–2(f) show that the time courses of the putative source variables [$x_1(t)$ and $x_2(t)$] were nearly the same as the time courses of the statistically independent state variables, which were used to generate the voices' utterances (up to a transformation on each state variable space). Thus, it is apparent that the information encoded in the time series of each vocal tract's state variable was blindly extracted from the simulated recording of the superposed utterances.

It is not hard to show that the same results will be obtained if we similarly process input data consisting of *any* set of five or more spectral features that are functions of the two underlying state variables. For any choice of such features,

the system's trajectory in feature space will lie on a two-dimensional surface, on which a coordinate system $\tilde{x}$ ($\tilde{x}_k$ for $k=1,2$) can be induced by dimensional reduction. The Takens embedding theorem[20] almost guarantees that there will be an invertible mapping between the values of $\tilde{x}$ and the underlying state variables of the two vocal tracts. This means that $\tilde{x}$ will constitute a coordinate system on the state space of the system, with the nature of that coordinate system being determined by the choice of the measured spectral features. In other words, the only effect of measuring different spectral features is to influence the nature of the coordinate system in which the system's state space trajectory is observed. However, recall that the procedure for identifying source variables in Sec. II is coordinate-system-independent. Therefore, no matter what spectral features are measured, the same source variables will be identified (up to permutations and transformations of individual components).

It was not possible to use the single-microphone recording to recover the exact sound of each voice's utterance. However, the recovered state variable time series [e.g., Figs. 2(c) and 2(d)] were used to synthesize sounds in which the original separate "messages" could be heard. Specifically, the above-derived mapping from filter bank output space to x space was inverted and used to compute a trajectory in filter bank output space, corresponding to the recovered $x_1(t)$ [or $x_2(t)$] time series and a constant value of x_2 (or x_1). Then, this time series of filter bank outputs was used to compute a wave form that had similar filter bank outputs. In each case, the resulting wave form sounded like a crude voice-converted version of the original utterance of one voice, with a constant "hum" of the other voice in the background.

IV. DISCUSSION

This paper outlines an approach to nonlinear BSS, based on a notion of statistical independence that is characteristic of a wide variety of classical noninteracting physical systems. Specifically, the proposed method seeks to determine if the observed data are mixtures of source variables that are statistically independent in the sense that their phase-space density function equals the product of density functions of individual components (and their time derivatives). Given a time series of observations in a measurement-defined coordinate system ($\tilde{x}$), the basic problem is to determine if there is another coordinate system (a source coordinate system x) in which the density function is factorizable. The existence (or nonexistence) of such a source coordinate system is a coordinate-system-independent property of the time series of data (i.e., an intrinsic or "inner" property). This is because, in *all* coordinate systems, there either is or is not a transformation to such a source coordinate system. In general, differential geometry provides mathematical machinery for determining whether a manifold has a coordinate-system-independent property like this. In the case at hand, we induce a geometric structure on the state space by identifying its metric with the local second-order correlation matrix of the data's velocity.[6] Then, a necessary condition for BSS is that the curvature tensor vanishes in all coordinate systems (including the measurement coordinate system). Therefore, if this data-derived quantity is nonvanishing, the observations are not separable. However, if the curvature tensor is zero, the data are separable if, and only if, the density function is seen to factorize in a Euclidean coordinate system that can be explicitly constructed by using the data-derived affine connection. If it does factorize, these coordinates are the unique source variables (up to transformations that do not affect separability). In short, the BSS problem requires that one sift through all possible mixing functions in order to find one that separates the data, and this arduous task can be mapped onto the solved differential geometric problem of examining all possible coordinate transformations in order to find one that transforms a flat metric into the identity matrix.

Appendix A 1 describes the solution[4] of the more general multidimensional BSS problem, which is sometimes called multidimensional independent component analysis[12] (ICA) or independent subspace analysis.[13] Here, the source components are only required to be partitioned into statistically independent groups, each of which may contain statistically dependent components.[1,2] This more general methodology is illustrated with analytic examples, as well as with the detailed numerical simulation of an optical experiment, in Appendices A 2 and A 3, respectively. Note that many of the most interesting natural signals (e.g., speech, music, electroencephalographic data, and magnetoencephalographic data) are likely to be generated by multidimensional sources. Therefore, multidimensional blind source separation will be necessary in order to separate those sources from noise and from one another.

In this paper, we implicitly sought to determine whether or not the data were separable everywhere (i.e., globally) in state space. However, this was done by determining whether or not local criteria for statistical independence [e.g., Eqs. (5) and (A10)–(A12)] were true globally. However, these local criteria could also be applied separately in each small neighborhood of state space in order to determine the degree of separability of the data in that "patch." In this way, one might find that given data are inseparable in some neighborhoods, separable into multidimensional source variables in other neighborhoods, and separable into one-dimensional sources elsewhere. In other words, the methodology of this paper could be used to explore the local separability of data in the same way that Riemannian geometry can be used to assess the local intrinsic geometry of a manifold.

What are the limitations on the application of this method? As discussed in Sec. II, the metric certainly exists if the trajectory is described by a density function in phase space, and, in Appendix A 2, we showed that this condition is satisfied by trajectories describing a wide variety of physical systems. More generally, the metric is expected to be well defined if the data's trajectory densely covers a region of state space and if its local velocity distribution varies smoothly over that region. In practice, one must have observations that cover state space densely enough in order to compute the metric, as well as its first and second derivatives (required to compute the affine connection and curvature tensor). In the numerical simulation in Appendix A 3, approximately 8.3×10^6 short trajectory segments (containing a total of 56×10^6 points) were used to compute the metric and

curvature tensor on a $32\times32\times32$ grid on the three-dimensional state space. Of course, if the dimensionality of the state space is higher, even more data will be needed. So, a relatively long time series of data must be recorded in order to be able to separate them. However, this requirement is not surprising. After all, our task is to examine the huge search space of all possible mixing functions, and the data must be sufficiently abundant and sufficiently restrictive to eliminate all but one of them. There are few other limitations on the applicability of the technique. In particular, computational expense is not prohibitive. The computation of the metric is the most CPU-intensive part of the method. However, because the metric computation is local in state space, it can be distributed over multiple processors, each of which computes the metric in a small neighborhood. The observed data can also be divided into "chunks" corresponding to different time intervals, each of which is sent to a different processor where its contribution to the metric is computed. As additional data are accumulated, they can be processed separately and then added into the time average of the data that were used to compute the earlier estimate of the metric. Thus, the earlier data need not be processed again, and only the latest observations need to be kept in memory.

As mentioned earlier, separability is an intrinsic property of a time series of data in the sense that it does not depend on the coordinate system in which the data are represented. However, there are many other intrinsic properties of such a time series that can be learned by a "blinded" observer. For instance, in Ref. 6, the data-derived parallel transfer operation was employed to describe relative locations of the observed data points in a coordinate-system-independent manner. As a specific example of such a description, suppose that states $\tilde{x}_A$, $\tilde{x}_B$, and $\tilde{x}_C$ differ by small state transformations, and suppose that a more distant state $\tilde{x}_D$ can be described as being related to $\tilde{x}_A$, $\tilde{x}_B$, and $\tilde{x}_C$ by the following procedure: "$\tilde{x}_D$ is the state that is produced by the following sequence of operations: (1) start with state $\tilde{x}_A$ and parallel transfer the vectors $\tilde{x}_B-\tilde{x}_A$ and $\tilde{x}_C-\tilde{x}_A$ along $\tilde{x}_B-\tilde{x}_A$ 23 times and (2) start at the end of the resulting geodesic and parallel transfer $\tilde{x}_C-\tilde{x}_A$ along itself 72 times." Because the parallel transfer operation is coordinate-system-independent, this statement is a coordinate-system-independent description of the relative locations of $\tilde{x}_A$,$\tilde{x}_B$,$\tilde{x}_C$, and $\tilde{x}_D$. The collection of all such statements about relative state locations constitutes a rich coordinate-system-independent representation of the data. Such statements are also observer independent because the only essential difference between observers equipped with different sensors is that they record the data in different coordinate systems (e.g., see Sec. III and Ref. 6). Different observers can use this technology to represent the data in the same way, even though they do not communicate with one another, have no prior knowledge of the observed physical system, and are blinded to the nature of their own sensors. The current paper shows how such blinded observers can glean another intrinsic property of the data, namely its separability.

It is interesting to speculate about the relationship of the proposed methodology to biological phenomena. Many psychological experiments suggest that human perception is remarkably sensor independent. Specifically, suppose that an individual's visual sensors are changed by having the subject wear goggles that distort and/or invert the observed scene. Then, after a sufficiently long period of adaptation, most subjects perceive the world in approximately the same way as they did before the experiment.[21] An equally remarkable phenomenon is the approximate universality of human perception: i.e., the fact that perceptions seem to be shared by individuals with different sensors (e.g., different ocular anatomy and different microscopic brain anatomy), as long as they have been exposed to similar stimuli in the past. Thus, many human perceptions seem to represent properties that are "intrinsic" to the time series of experienced stimuli in the sense that they do not depend on the type of sensors used to observe the stimuli (or on the nature of the sensor-defined coordinate system on state space). In many situations, people are also able to perform source separation in a blinded or nearly blinded fashion. For example, they can often separate a speech signal from complex superposed noise (i.e., they can solve the "cocktail party problem"). This means that the human brain perceives another coordinate-system-independent property of the data, namely its separability. This paper and an earlier one[6] describe methods of finding such inner properties of a sufficiently dense data time series. Is it possible that the human brain somehow extracts these particular geometric invariants from sensory data? The only way to test this speculation is to perform biological experiments to determine if the brain actually utilizes the specific data-derived metric and geometric structure described in these two papers.

ACKNOWLEDGMENTS

The author is grateful to Stephen McLaughlin for suggesting the application of the methodology in Ref. 6 to the blind source separation problem. The insightful comments of Michael A. Levin are appreciated, as are those of Aapo Hyvärinen and Shun-ichi Amari, who reviewed the original manuscript and drew the author's attention to the uniqueness issue. This work was performed with the partial support of the National Institute of Deafness and Communicative Disorders.

APPENDIX A: MULTIDIMENSIONAL BLIND SOURCE SEPARATION

1. Theory

This subsection describes the solution of the more general BSS problem[12,13] in which the source components are only required to be partitioned into groups, each of which is statistically independent of the others but each of which may contain statistically dependent variables.[1,2] Figure 3 summarizes the method[4] described in this subsection. This procedure is illustrated with analytic and numerical examples in the next two subsections.

Let $\tilde{x}(t)$ ($\tilde{x}_k$ for $k=1,2,\ldots,n$) denote a time series of measurements. Suppose that we want to determine if these data are instantaneous mixtures of source variables $x(t)$, having a factorizable density function

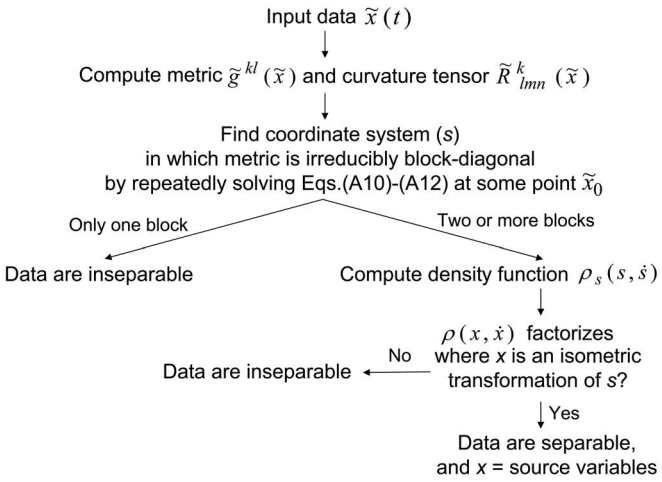

FIG. 3. Differential geometric method for multidimensional BSS.

$$\rho(x,\dot{x}) = \rho_A(x_A,\dot{x}_A)\rho_B(x_B,\dot{x}_B), \tag{A1}$$

where x_A and x_B contain components of x with indices $x_{Ak}=x_k$ for $k=1,2,\ldots,n_A<n$ and $x_{Bk}=x_k$ for $k=n_A+1,n_A+2,\ldots,n_A+n_B=n$. A necessary consequence of Eq. (A1) is that the metric given by Eq. (4) is block-diagonal in the x coordinate system; i.e.,

$$g_{kl}(x) = \begin{bmatrix} g_A(x_A) & \mathbf{0} \\ \mathbf{0} & g_B(x_B) \end{bmatrix}_{kl}, \tag{A2}$$

where g_A and g_B are $n_A\times n_A$ and $n_B\times n_B$ matrices, respectively, and each $\mathbf{0}$ symbol denotes a null matrix of appropriate dimensions. Notice that each block on the right side of Eq. (A2) is a function of the corresponding block of coordinates; in the following, we use the term "block-diagonal" to refer to metrics with this property. The necessary condition for separability in Eq. (A2) suggests the following strategy. The first step is to find the transformation from the measurement-defined coordinate system $\tilde{x}$ to a coordinate system in which the metric is irreducibly block-diagonal (in the sense that each block cannot be further block-diagonalized). If there is a source coordinate system, the x_A source variable must be comprised of the coordinate components in a group of these irreducible blocks (or mixtures of the components in a group of such blocks), and the x_B source variable must consist of mixtures of the coordinate components in the complimentary group of blocks. Therefore, once the irreducibly block-diagonal coordinate system has been found, the BSS problem is reduced to the following tasks: (1) transform the density function into that coordinate system and (2) determine if the density function is the product of factors, each of which is a function of the coordinate components in one of a set of mutually exclusive groups of blocks. If the density function does not factorize in the irreducibly block-diagonal coordinate system, the data are simply not separable.

The first step is to transform the metric into an irreducibly block-diagonal form. To do this, we assume that the metric can be transformed into a block-diagonal form with two, possibly reducible blocks [Eq. (A2)], and then we derive necessary conditions that follow from that assumption. It is helpful to define the A (B) subspace at each point x to be the hyperplane through that point with constant x_B (x_A). A vector at x is projected onto the A subspace by the $n\times n$ matrix $A^k{}_l$,

$$A^k{}_l(x) = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}_{kl}, \tag{A3}$$

where $\mathbf{1}$ is the $n_A\times n_A$ identity matrix. For example, if $\dot{x}$ is the velocity of the data's trajectory at x, then $A^k{}_l\dot{x}_l$ is the velocity's component in the A subspace, where we have used Einstein's convention of summing over repeated indices. The complementary projector onto the B subspace is $B^k{}_l=\delta^k{}_l-A^k{}_l$, where $\delta^k{}_l$ is the Kronecker delta. In any other coordinate system (e.g., the $\tilde{x}$ coordinate system), the corresponding projectors ($\tilde{A}^k{}_l$ and $\tilde{B}^k{}_l$) are mixed-index tensor transformations of the projectors in the x coordinate system; for example,

$$\tilde{A}^k{}_l(\tilde{x}) = \frac{\partial\tilde{x}_k}{\partial x_{k'}}\frac{\partial x_{l'}}{\partial\tilde{x}_l}A^{k'}{}_{l'}(x). \tag{A4}$$

Because the A and B projectors permit the local separation of the A and B subspaces, it will be useful to be able to construct them in the measurement ($\tilde{x}$) coordinate system. Our strategy for doing this is to find conditions that the projectors must satisfy in the x coordinate system and then transfer those conditions to the $\tilde{x}$ coordinate system by writing them in coordinate-system-independent form. First, note that Eq. (A3) implies that $A^k{}_l$ is idempotent

$$A^k{}_{k'}(x)A^{k'}{}_l(x) = A^k{}_l(x), \tag{A5}$$

its "trace" is an integer

$$A^k{}_k(x) = n_A, \tag{A6}$$

and it is unequal to the identity and null matrices. Next, consider the Riemann–Christoffel curvature tensor of the data space[15]

$$R^k{}_{lmn}(x) = -\frac{\partial\Gamma^k_{lm}}{\partial x_n} + \frac{\partial\Gamma^k_{ln}}{\partial x_m} + \Gamma^k_{im}\Gamma^i_{ln} - \Gamma^k_{in}\Gamma^i_{lm}, \tag{A7}$$

where the affine connection Γ^k_{lm} is defined in the usual way

$$\Gamma^k_{lm}(x) = \frac{1}{2}g^{kn}\left(\frac{\partial g_{nl}}{\partial x_m} + \frac{\partial g_{nm}}{\partial x_l} - \frac{\partial g_{lm}}{\partial x_n}\right). \tag{A8}$$

The block-diagonality of g_{kl} in the x coordinate system implies that Γ^k_{lm} and R^k_{lmn} are also block-diagonal in all of their indices. The block-diagonality of the curvature tensor, together with Eq. (A3), implies

$$R^j_{klm}(x)A^k{}_i(x) - A^j{}_k(x)R^k_{ilm}(x) = 0 \tag{A9}$$

at each point x. Covariant differentiation of Eq. (A9) will produce other local conditions that are necessarily satisfied by data with a block-diagonalizable metric. It can be shown that these conditions are also linear algebraic constraints on the subspace projector because the projector's covariant derivative vanishes.

Notice that both sides of Eqs. (A5), (A6), and (A9) transform as tensors when the coordinate system is changed.

Therefore, these equations must be true *in any coordinate system* on a space with a block-diagonalizable metric. In particular, in the $\tilde{x}$ coordinate system that is defined by the observer's choice of measurements, we have

$$\tilde{A}^k{}_{k'}(\tilde{x})\tilde{A}^{k'}{}_l(\tilde{x}) = \tilde{A}^k{}_l(\tilde{x}), \tag{A10}$$

$$\tilde{A}^k{}_k(\tilde{x}) = n_A, \tag{A11}$$

$$\tilde{R}^j{}_{klm}(\tilde{x})\tilde{A}^k{}_i(\tilde{x}) - \tilde{A}^j{}_k(\tilde{x})\tilde{R}^k{}_{ilm}(\tilde{x}) = 0, \tag{A12}$$

where $1 \leq n_A < n$. So far, we have shown that, if the metric can be transformed into the form in Eq. (A2), there must be solutions of Eqs. (A10)–(A12). Thus, block-diagonalizability imposes a significant constraint on the curvature tensor of the space and, therefore, on the observed data.

What is the intuitive meaning of Eq. (A12)? Because of the block-diagonality of the affine connection in the x coordinate system, it is easy to see that parallel transfer of a vector lying within the A (or B) subspace at any point produces a vector within the A (or B) subspace at the destination point. Consequently, the corresponding projectors ($\tilde{A}^k{}_l$ and $\tilde{B}^k{}_l$) at the first point parallel transfer into themselves at the destination point. In particular, parallel transferring one of these projectors along the ith direction and then along the jth direction will give the same result as parallel transferring it along the jth direction and then along the ith direction. It is not hard to show[15] that Eq. (A12) is a statement of this fact: namely, block-diagonalizable manifolds support local projectors that are parallel transferred in a path-independent manner. In contrast, if a Riemannian manifold is not block-diagonalizable, there may be no solutions of Eqs. (A10)–(A12). For example, on any intrinsically curved two-dimensional surface (e.g., a sphere), it is not possible to find a one-dimensional projector at each point (i.e., a direction at each point) that satisfies Eq. (A12). This is because the parallel transfer of directions on such a surface is path dependent.

Suppose we know that the metric can be transformed into the form in Eq. (A2), and suppose we can find the corresponding solutions of Eqs. (A10)–(A12). Then, we can use them to explicitly construct a transformation from the measurement-defined coordinate system ($\tilde{x}$) to a block-diagonal coordinate system (s). Let $\tilde{A}^k{}_l(\tilde{x}_0)$ and $\tilde{B}^k{}_l(\tilde{x}_0)$ be the solutions of Eqs. (A10)–(A12) at an arbitrarily chosen point $\tilde{x}_0$. Then, we can use these projectors to construct a geodesic coordinate system. To do this, first select n linearly independent small vectors $\delta\tilde{y}_{(i)}$ $(i=1,2,\ldots,n)$ at $\tilde{x}_0$, and use $\tilde{A}^k{}_l(\tilde{x}_0)$ and $\tilde{B}^k{}_l(\tilde{x}_0)$ to project them onto the local A and B subspaces. Then, use the results to create a set of n_A linearly independent vectors $\delta\tilde{x}_{(a)}$ $(a=1,2,\ldots,n_A)$ and a set of n_B linearly independent vectors $\delta\tilde{x}_{(b)}$ $(b=n_A+1,n_A+2,\ldots,n)$, which lie within the A and B subspaces, respectively. Finally, (1) starting at $\tilde{x}_0$, use the affine connection to repeatedly parallel transfer all $\delta\tilde{x}$ along $\delta\tilde{x}_{(1)}$; (2) starting at each point along the resulting geodesic path, repeatedly parallel transfer these vectors along $\delta\tilde{x}_{(2)}$; … continue the parallel transfer process along the directions $\delta\tilde{x}_{(3)},\ldots,\delta\tilde{x}_{(n-1)}\ldots$; (n) starting at each point along the most recently produced geodesic path, parallel transfer these vectors along $\delta\tilde{x}_{(n)}$. Each point in the neighborhood of $\tilde{x}_0$ is assigned the geodesic coordinate s $(s_k,k=1,2,\ldots,n)$, where each component s_k represents the number of parallel transfers of the vector $\delta\tilde{x}_{(k)}$ that was required to reach it. If these projection and parallel transfer procedures are visualized in the x coordinate system, it can be seen that the first n_A components of s (i.e., s_A) will be functions of x_A and the last n_B components of s (s_B) will be functions of x_B. In other words, s and x will just differ by a coordinate transformation that is block-diagonal with respect to the subspaces. Therefore, the metric will be block-diagonal in the s coordinate system, just like it is in the x coordinate system. But, because s is defined by a coordinate-system-independent procedure, the same s coordinate system will be constructed by performing that procedure in the measurement-defined coordinate system $\tilde{x}$. In summary: Eq. (A2) necessarily implies that there are subspace projectors satisfying Eqs. (A10)–(A12) at $\tilde{x}_0$ and that the metric will look like Eq. (A2) in the geodesic (s) coordinate system computed from those projectors.

We are now in a position to systematically determine if the observed data can be decomposed into independent source variables. The first step is to use the observed measurements $\tilde{x}(t)$ to compute the metric [Eq. (4)], affine connection [Eq. (A8)], and curvature tensor [Eq. (A7)] at an arbitrary point $\tilde{x}_0$ in the data space. Next, we look for projectors $\tilde{A}^k{}_l(\tilde{x}_0)$ that are solutions of Eqs. (A10)–(A12) at that point. If a solution is not found, we conclude that the metric cannot be diagonalized into two or more blocks (i.e., there is only one irreducible block), and, therefore, the data are not separable. If one or more solutions are found, we search for one that leads to an s (geodesic) coordinate system in which the metric is block-diagonal. If there is no such solution (i.e., all the solutions are extraneous), we conclude that the metric has only one irreducible block, and, therefore, the data are not separable. If we do find such a solution, we use it to transform the metric into the form of Eq. (A2), and the foregoing procedure is then applied separately to each block in order to see if it can be further block-diagonalized into smaller blocks. In this way, we construct a geodesic coordinate system (s) in which the metric consists of a diagonal array of irreducible blocks. The only other irreducibly block-diagonal coordinate systems are those produced by permutations of blocks, intrablock coordinate transformations, and possible isometries of the metric[15] that mix coordinate components from different blocks.

As mentioned before, each multidimensional source variable must be comprised of the coordinate components in one group of a set of mutually exclusive groups of blocks in an irreducibly block-diagonal coordinate system (or mixtures of the coordinate components within such a group of blocks). In most practical applications, the data-derived metric will have no isometries that mix coordinate components from different blocks. In that case, the above-described geodesic (s) coordinate system is the only possible separable coordinate system (up to permutations and intrablock transformations). Then, the final step is to compute the density function of the

data in the s coordinate system and determine if it is the product of two or more factors, each of which is a function of the components (and their time derivatives) in one group of a set of mutually exclusive groups of irreducible coordinate blocks. If it does factorize, the corresponding groups of coordinate components comprise multidimensional source variables that are unique (up to permutations and transformations of each multidimensional source variable). If it does not factorize, the data are not completely separable.

If the metric does have isometries that mix coordinate components of different blocks, the factorizability of the density function must be tested in all coordinate systems derived by isometric transformations of the s coordinate system. Notice that this procedure will not produce a unique set of source variables if the density function factorizes in more than one of these isometrically related coordinate systems. In practical applications, the most important case in which there are metric isometries involves a metric that describes a multidimensional flat subspace. Specifically, suppose that the irreducible form of the metric includes n_E one-dimensional blocks, where $n_E \geq 2$. Because the metric is positive-definite and each diagonal element is a function of the corresponding coordinate component, each 1×1 metric block can be transformed to unity by a possibly nonlinear transformation of the corresponding variable. These components can then be mixed by any n_E-dimensional rotation, without affecting the metric (i.e., these rotations are isometries that mix the coordinate components of different one-dimensional blocks). For each value of this unknown rotation matrix, one must then determine if the density function factorizes. Thus, in this particular case, the proposed methodology reduces the nonlinear BSS problem to the linear BSS problem of finding which rotations of the data separate them.

In practice, we may not have enough data to accurately calculate the phase space density function and thereby directly assess the separability of the s coordinate system (and possible isometric transformations of it). However, in that case, we can still check a variety of weaker conditions that are necessary for separability. For example, we could compute σ_{kl} [Eq. (6)] in order to see if it has a block-diagonal form, in which the blocks correspond to mutually exclusive groups of the irreducible metric blocks. Likewise, we could check higher-order correlations of s and $\dot{s}$ in order to see if they factorize into products of lower-order correlations, which involve the variables within mutually exclusive groups of metric blocks. The only possible multidimensional source variables consist of the coordinate components in mutually exclusive groups of irreducible metric blocks for which these higher-order correlations are found to factorize.

2. Analytic examples

In this subsection, we demonstrate large classes of trajectories that satisfy the assumptions in Sec. II and Appendix A 1. In these cases: (1) the trajectory's statistics are described by a density function in phase space, (2) the trajectory-derived metric is well defined and can be computed analytically, and (3) there is a source coordinate system in which the density function is separable into the product of two density functions. Many of these trajectories are constructed from the behavior of physical systems that could be realized in actual or simulated laboratory experiments (see Appendix A 3).

First, consider the energy of a physical process with n degrees of freedom x (x_k for $k=1,2,\ldots,n$),

$$E(x,\dot{x}) = \frac{1}{2}\mu_{kl}(x)\dot{x}_k\dot{x}_l + V(x), \tag{A13}$$

where μ_{kl} and V are some functions of x and repeated indices are summed. Furthermore, suppose that

$$\mu_{kl}(x) = \begin{bmatrix} \mu_A(x_A) & \mathbf{0} \\ \mathbf{0} & \mu_B(x_B) \end{bmatrix}_{kl}, \tag{A14}$$

$$V(x) = V_A(x_A) + V_B(x_B), \tag{A15}$$

where μ_A and μ_B are $n_A \times n_A$ and $n_B \times n_B$ matrices, respectively, for $1 \leq n_A < n$ and $n_B = n - n_A$, where each $\mathbf{0}$ symbol denotes a null matrix of appropriate dimensions, and where $x_{Ak}=x_k$ for $k=1,2,\ldots,n_A$ and $x_{Bk}=x_k$ for $k=n_A+1,n_A+2,\ldots,n$. These equations describe the degrees of freedom (x_A and x_B) of almost any pair of classical physical systems, which do not exchange energy or interact with one another. A simple system of this kind consists of a particle with coordinates x_A moving in a potential V_A on a possibly warped two-dimensional frictionless surface with physical metric $\mu_{Akl}(x_A)$, together with a particle with coordinates x_B moving in a potential V_B on a two-dimensional frictionless surface with physical metric $\mu_{Bkl}(x_B)$. In the general case, suppose that the system intermittently exchanges energy with a thermal bath at temperature T. This means that the system evolves along one trajectory from the Maxwell–Boltzmann distribution at that temperature and periodically jumps to another trajectory randomly chosen from that distribution. After a sufficient number of jumps, the amount of time the system will have spent in a small neighborhood $dxd\dot{x}$ of $(x,\dot{x})$ is given by the product of $dxd\dot{x}$ and a density function that is proportional to the Maxwell–Boltzmann distribution[14]

$$\mu(x)\exp[-E(x,\dot{x})/kT], \tag{A16}$$

where k is the Boltzmann constant and μ is the determinant of μ_{kl}. The existence of this density function means that the local velocity covariance matrix is well defined and computation of the relevant Gaussian integrals shows that it is

$$\langle(\dot{x}_k - \bar{\dot{x}}_k)(\dot{x}_l - \bar{\dot{x}}_l)\rangle_x = kT\mu^{kl}(x), \tag{A17}$$

where μ^{kl} is the contravariant tensor equal to the inverse of μ_{kl}. It follows that the trajectory-induced metric on the state space is well defined and is given by $g_{kl}(x)=\mu_{kl}(x)/kT$. Furthermore, Eq. (A16) shows that the density function is the product of the density functions of the two noninteracting subsystems.

Appendix A 3 describes the numerical simulation of a physical system of this type, which was comprised of two noninteracting subsystems: one with two statistically dependent degrees of freedom and the other with one degree of freedom. The technique in Appendix A 1 was applied to the observed data to perform multidimensional BSS: i.e., to

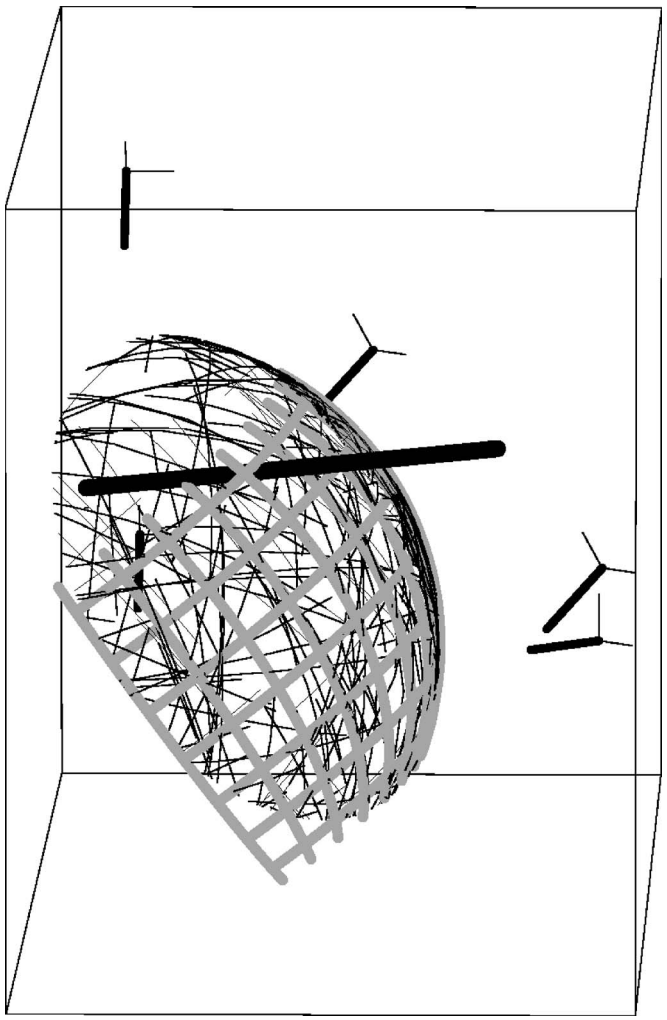

FIG. 4. The thin black curved lines comprise a small sample of the trajectory segments traversed by the particle that was confined to a spherical surface, and the long thick black line shows the corresponding trajectory segments of the second particle constrained to a straight line. The system was watched by five simulated pinhole cameras. Each small triplet of orthogonal straight lines shows the relative position and orientation of a camera, with the long thick line of each triplet being the perpendicular to a camera focal plane that was represented by the two short thin lines of each triplet. One camera is nearly obscured by the spherical surface. The thick gray curved lines show some latitudes and longitudes on the spherical surface.

blindly find the transformation from a measurement-defined coordinate system to a source coordinate system.

3. Numerical example: Optical imaging of two moving particles

In the following, the scenario described in Appendix A 2 is illustrated by the numerical simulation of a physical system with three degrees of freedom. The system was comprised of two moving particles of unit mass, one moving on a transparent frictionless curved surface and the other moving on a frictionless line. Figure 4 shows the curved surface, which consisted of all points on a spherical surface within one radian of a randomly chosen point. Figure 4 also shows that the curved surface and line were oriented at arbitrarily chosen angles with respect to the simulated laboratory coordinate system. Both particles moved freely, and they were in

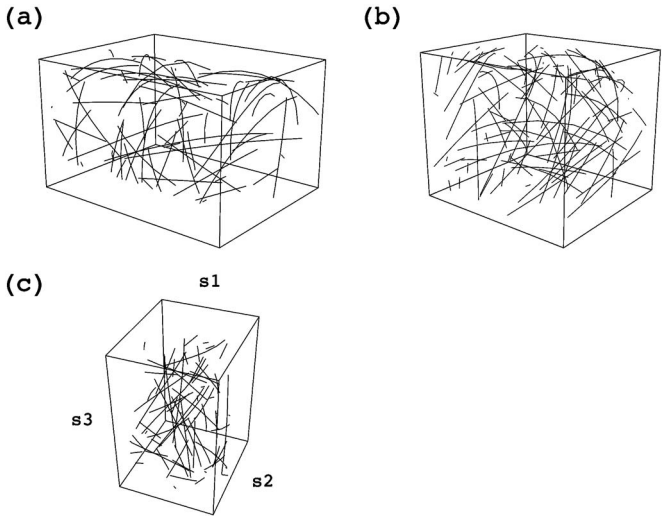

FIG. 5. (a) A small sample of segments of the system's trajectory in the 20-dimensional space of camera outputs. Only the first three principal components are shown. (b) A small sample of trajectory segments after dimensional reduction was used to map them from the 20-dimensional space onto the three-dimensional state ($\tilde{x}$) space. (c) A small sample of trajectory segments after they were transformed from the $\tilde{x}$ coordinate system to the geodesic (s) coordinate system (i.e., the "experimentally" determined source coordinate system).

thermal equilibrium with a bath for which $kT=0.01$ in the chosen units of mass, length, and time. As in Appendix A 2, the system's trajectory was created by temporally concatenating approximately 8.3×10^6 short trajectory segments randomly chosen from the corresponding Maxwell–Boltzmann distribution, given by Eqs. (A13)–(A16), where x_A and x_B denote coordinates on the spherical surface and on the line, respectively, where μ_A is the metric of the spherical surface, where μ_B is a constant, and where $V_A=V_B=0$. The factorizability of this density function makes it evident that $x=(x_A,x_B)$ comprised a source coordinate system. Figure 4 shows a small sample of the trajectory segments.

The particles were "watched" by a simulated observer $\tilde{O}b$ equipped with five pinhole cameras, which had arbitrarily chosen positions and faced the sphere/line with arbitrarily chosen orientations (Fig. 4). The image created by each camera was transformed by an arbitrarily chosen second-order polynomial, which varied from camera to camera. In other words, each pinhole camera image was warped by a translational shift, rotation, rescaling, skew, and quadratic deformation that simulated the effect of a distorted optical path between the particles and the camera's "focal" plane. The output of each camera was comprised of the four numbers representing the two particles' locations in the distorted image on its focal plane. As the particles moved, the cameras created a time series of sensor multiplets, each of which consisted of the 20 numbers produced by all five cameras at one time point. Figure 5(a) shows the first three principal components of the system's trajectory through the corresponding 20-dimensional space. A dimensional reduction technique[19] was applied to the full 20-dimensional time series in order to identify the underlying three-dimensional state space and to establish a coordinate system ($\tilde{x}$) on it, thereby eliminating redundant sensor data. Figure 5(b) shows

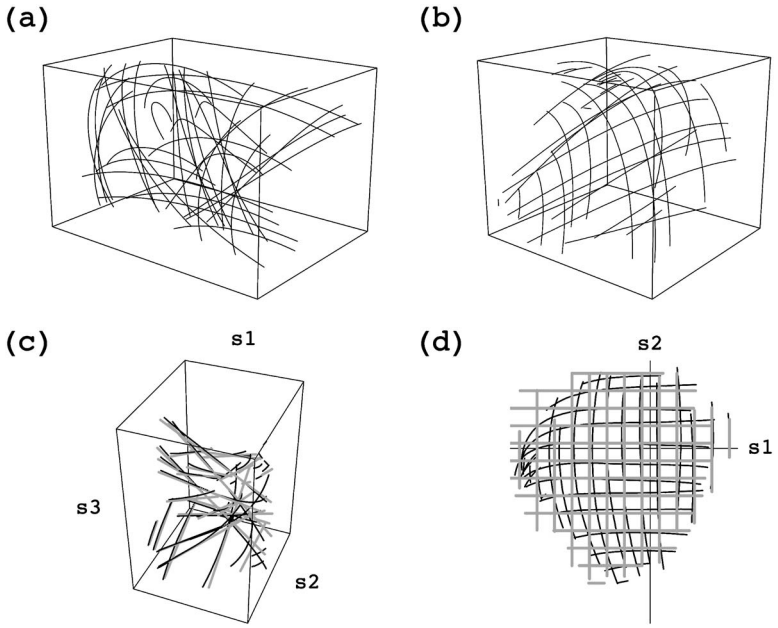

FIG. 6. (a) The test lines after they were mapped into the 20-dimensional space of camera outputs. The figure only depicts the resulting pattern's projection onto the space of the first three principal components of the 20-dimensional system trajectory. (b) The test lines after dimensional reduction was used to map them from the 20-dimensional space onto the three-dimensional state ($\tilde{x}$) space traversed by the trajectory segments. (c) The thin black lines show the test pattern after it was transformed from the $\tilde{x}$ coordinate system to the geodesic (s) coordinate system, which comprises the "experimentally" derived source coordinate system. The thick gray lines show the test lines in the comparable exact source coordinate system. (d) The thin and thick lines show the first two components of the test pattern in (c). These collections of lines represent the projection of the test pattern onto the experimentally derived two-dimensional independent subspace and onto the exactly known independent subspace, respectively.

typical trajectory segments in the $\tilde{x}$ coordinate system. Because the underlying state space had dimensionality $n=3$ and because the 20-dimensional sensor multiplets had more than $2n$ components, the Takens embedding theorem[20] virtually guaranteed that there was a one-to-one mapping between the system states and the corresponding values of $\tilde{x}$. In other words, it guaranteed that the $\tilde{x}$ coordinates were invertible instantaneous mixtures of the particle locations, as in Eq. (1). The exact nature of the mixing function depended on the positions, orientations, and optical distortions of the five cameras.

Given the measurements $\tilde{x}(t)$ and no other information, our task was to determine if they were instantaneous mixtures of statistically independent groups of source variables. This was accomplished by blindly processing the data with the technique described in Appendix A 1. First, Eqs. (4), (A7), and (A8) were used to compute the metric, affine connection, and curvature tensor in this coordinate system. Then, Eqs. (A10)–(A12) were solved at a point $\tilde{x}_0$. One pair of solutions was found, representing a local projector onto a two-dimensional subspace and the complementary projector onto a one-dimensional subspace. Following the procedure in Appendix A 1, we selected three small linearly independent vectors $\delta\tilde{y}_{(i)}$ $(i=1,2,3)$ at $\tilde{x}_0$, and we used the projectors at that point to project them onto the putative A and B subspaces. Then, the resulting projections were used to create a set of two linearly independent vectors $\delta\tilde{x}_{(a)}$ $(a=1,2)$ and a single vector $\delta\tilde{x}_{(3)}$ within the A and B subspaces, respectively. Finally, the geodesic (s) coordinate system was constructed by using the affine connection to parallel transfer these vectors throughout the neighborhood of $\tilde{x}_0$ [Fig. 5(c)]. After the metric was transformed into the s coordinate system, it was found to have a nearly block-diagonal form, consisting of a 2×2 block and a 1×1 block. Because the two-dimensional subspace had nonzero intrinsic curvature, the 2×2 metric block could not be decomposed into smaller (i.e., one-dimensional) blocks. Therefore, in this example, the only possible source coordinate system was the geodesic (s) coordinate system, which was unique up to coordinate transformations on each block and up to subspace permutations.

In order to demonstrate the accuracy of this separation process, we defined "test lines" that had known projections onto the independent subspaces used to define the system. Then, we compared those projections with the test pattern's projection onto the independent subspaces that were "experimentally" determined as described earlier. First, we defined an x coordinate system in which x_A was the position (longitude, latitude) of the particle on the spherical surface and in which x_B was the position of the other particle along the line (Fig. 4). In this coordinate system, the test lines consisted of straight lines that were oriented at various angles with respect to the $x_B=0$ plane and that projected onto the grid-like array of latitudes and longitudes in that plane. In other words, each line corresponded to a path generated by moving the first particle along a latitude or longitude of the sphere and simultaneously moving the second particle along its constraining line. The points along these test lines were "observed" by the five pinhole cameras to produce corresponding lines in the 20-dimensional space of the cameras' output [Fig. 6(a)]. These lines were then mapped onto lines in the $\tilde{x}$ coordinate system by means of the same procedure used to dimensionally reduce the trajectory data [Fig. 6(b)]. Finally, the test pattern was transformed from the $\tilde{x}$ coordinate system to the s coordinate system, the geodesic coordinate system that comprised the experimentally derived source coordinate system. As mentioned earlier, the s coordinate system was the only possible separable coordinate system, except for permutations and arbitrary coordinate transformations on each subspace. Therefore, it should be the same as the x coordinate system (an exactly known source coordinate system), except for such transformations. The nature of that coordinate transformation depended on the choice of vectors that were parallel transferred to define the geodesic (s) coordinate system on each subspace. In order to compare the test pattern in the experimentally derived source coordinate sys-

tem (s) with the appearance of the test pattern in the exactly known source coordinate system (x), we picked $\tilde{x}_0$ and the $\delta\tilde{y}$ vectors so that the s and x coordinate systems would be the same, as long as the independent subspaces were correctly identified by the BSS procedure. Specifically: (1) $\tilde{x}_0$ was chosen to be the mapping of the origin of the x coordinate system, which was located on the sphere's equator and at the line's center, (2) $\delta\tilde{y}_{(1)}$ and $\delta\tilde{y}_{(2)}$ were chosen to be mappings of vectors projecting along the equator and the longitude, respectively, at that point, and (3) all three $\delta\tilde{x}$ were normalized with respect to the metric in the same way as the corresponding unit vectors in the x coordinate system. Figure 6(c) shows that the test pattern in the experimentally derived source coordinate system consisted of nearly straight lines (narrow black lines), which almost coincided with the test pattern in the exactly known source coordinate system (thick gray lines). Figure 6(d) shows that the test pattern projected onto a grid-like pattern of lines (narrow black lines) on the experimentally determined A subspace, and these lines nearly coincided with the test pattern's projection onto the exactly known A subspace (thick gray lines). These results indicate that the proposed BSS method correctly determined the source coordinate system. In other words, the "blind" observer $\tilde{O}b$ was able to separate the state space into two independent subspaces, which were nearly the same as the independent subspaces used to define the system. Figures 6(c) and 6(d) also demonstrate an example of the observer-independence of statements about relative stimulus locations, as discussed in Sec. IV. Specifically, these figures show how many parallel transfers of the vectors $\delta\tilde{x}_{(i)}$ were required to reach each point of the test pattern, as computed by an observer $\tilde{O}b$ equipped with five pinhole camera sensors (narrow black lines) and as computed by a different observer Ob, who directly sensed the values of x (thick gray lines).

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis* (Wiley, New York, 2001).
[2] C. Jutten and J. Karhunen, Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation, Nara, Japan, April 2003.
[3] A. Hyvärinen and P. Pajunen, Neural Networks **12**, 429 (1999).
[4] D. N. Levin, "Using state space differential geometry for nonlinear blind source separation," 2006 (preprint posted at http://arxiv.org/abs/cs/0612096).
[5] D. N. Levin, *Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation*, Lect. Notes Comput. Sci. 4666 (Springer, Berlin, 2007), p. 65.
[6] D. N. Levin, J. Appl. Phys. **98**, 104701 (2005); also, see the papers posted at http://www.geocities.com/dlevin2001/.
[7] S. Lagrange, L. Jaulin, V. Vigneron, and C. Jutten, *Proceedings of the 5th International Symposium on Independent Component Analysis and Blind Source Separation*, Lect. Notes Comput. Sci. 3195 (Springer, Berlin, 2004), p. 81.
[8] S. Haykin, *Neural Networks-A Comprehensive Foundation* (Prentice Hall, New York, 1998).
[9] H. Yang, S. I. Amari, and A. Cichocki, Signal Process. **64**, 291 (1998).
[10] A. Taleb and C. Jutten, IEEE Trans. Signal Process. **47**, 2807 (1999).
[11] S. Amari, Neural Comput. **10**, 251 (1998).
[12] J.-F. Cardoso, Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, Seattle, 1998, p. 1941.
[13] Y. Nishimori, S. Akaho, and M. D. Plumbley, *Proceedings of the 6th International Conference on Independent Component Analysis and Blind Source Separation*, Lect. Notes Comput. Sci. 3889 (Springer, Berlin, 2006), p. 295.
[14] F. W. Sears, *Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics* (Addison-Wesley, Reading, MA, 1959).
[15] S. Weinberg, *Gravitation and Cosmology-Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972).
[16] N. Tishby, Proceedings of the 1990 IEEE International Conference on Acoustics, Speech and Signal Processing, 1990, p 365.
[17] B. Townshend, *Proceedings of the Workshop on Nonlinear Modeling and Forecasting*, Sante Fe, NM, 1990 (Addison-Wesley, New York, 1992).
[18] X. Huang, A. Acero, and H.-W. Hon, *Spoken Language Processing* (Prentice Hall PTR, Upper Saddle River, NJ, 2001).
[19] S. T. Roweis and L. K. Saul, Science **290**, 2323 (2000).
[20] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[21] R. Held and R. Whitman, *Perception: Mechanisms and Models* (Freeman, San Francisco, 1972).