

Performing Nonlinear Blind Source Separation with Signal Invariants

David N. Levin, *Member, IEEE*

Abstract—Given a time series of multicomponent measurements $x(t)$, the usual objective of nonlinear blind source separation (BSS) is to find a “source” time series $s(t)$, comprised of statistically independent combinations of the measured components. In this paper, the source time series is required to have a density function in (s, \dot{s}) -space that is equal to the product of density functions of individual components. This formulation of the BSS problem has a solution that is unique, up to permutations and component-wise transformations. Separability is shown to impose constraints on certain locally invariant (scalar) functions of x , which are derived from local higher-order correlations of the data’s velocity \dot{x} . The data are separable if and only if they satisfy these constraints, and, if the constraints are satisfied, the sources can be explicitly constructed from the data. The method is illustrated by using it to separate two speech-like sounds recorded with a single microphone.

Index Terms—Blind, nonlinear, separation, signal, source

EDICS Category: MLR-SSEP

I. INTRODUCTION

SENSORY devices often receive signals from multiple physical stimuli that evolve simultaneously but are unrelated to one another. In many of these situations, it is necessary to create separate representations of one or more of these stimuli by blindly processing the observed signals (i.e., by processing them without prior knowledge of the nature of the stimuli). In recent years, there has been considerable progress in the solution of this “blind source separation” (BSS) problem for the special case in which the signals and source variables are linearly related. However, although nonlinear BSS is often performed effortlessly by humans, computational methods for doing this are quite limited [1].

Consider a time series of data $x(t)$, where x is a multiplet of N measurements (x_k for $k = 1, 2, \dots, N$). The usual objectives of nonlinear BSS are: 1) determine if these data are instantaneous mixtures of N statistically independent source components $s(t)$

$$x(t) = f[s(t)], \quad (1)$$

where f is a possibly nonlinear, invertible N -component mixing function; 2) if this is the case, compute the mixing function. In other words, the problem is to find a coordinate transformation f^{-1} that transforms the observed data $x(t)$ from the measurement-defined coordinate system (x) on state space to a special source coordinate system (s) in which the

components of the transformed data are statistically independent. Let $\rho_S(s)$ be the state space probability density function (PDF) in the source coordinate system, defined so that $\rho_S(s)ds$ is the fraction of total time that the source trajectory $s(t)$ is located within the volume element ds at location s . In the usual formulation of the BSS problem, the source components are required to be statistically independent in the sense that their state space PDF is the product of the density functions of the individual components

$$\rho_S(s) = \prod_{k=1}^N \rho_k(s_k). \quad (2)$$

In every formulation of BSS, multiple solutions can be created by permutations and component-wise transformations of any one solution. However, it is well known that the criterion in (2) is so weak that it suffers from a much worse non-uniqueness problem: namely, in this form of the BSS problem, multiple solutions can be created by transformations that mix the source variables (see [2] and references therein).

The issue of non-uniqueness can be circumvented by considering the data’s trajectory in (s, \dot{s}) -space ($\dot{s} = ds/dt$) instead of s -space (i.e., state space). First, let $\rho_S(s, \dot{s})$ be the PDF in this space, defined so that $\rho_S(s, \dot{s})dsd\dot{s}$ is the fraction of total time that the location and velocity of the source trajectory are within the volume element $dsd\dot{s}$ at location (s, \dot{s}) . An earlier paper [3] described a formulation of the BSS problem in which this PDF was required to be the product of the density functions of the individual components

$$\rho_S(s, \dot{s}) = \prod_{k=1}^N \rho_k(s_k, \dot{s}_k). \quad (3)$$

Separability in (s, \dot{s}) -space is a stronger requirement than separability in state space. To see this, note that (2) can be recovered by integrating both sides of (3) over all velocities, but the latter equation cannot be deduced from the former one. In fact, it can be shown that (3) is strong enough to guarantee that the BSS problem in (s, \dot{s}) -space has a unique solution, up to permutations and component-wise transformations [3]. Furthermore, this type of statistical independence has the virtue of being satisfied by almost all classical physical systems that are composed of non-interacting subsystems, which are the generators of most signals of interest.

The author previously demonstrated [3] that the (s, \dot{s}) -space PDF of a time series induces a Riemannian geometry on the state space, with the metric equal to the local second-order correlation matrix of the data’s velocity. Nonlinear BSS can be performed by computing this metric in the x coordinate system (i.e., by computing the second-order correlation of \dot{x} at each

D. Levin is with the Department of Radiology and the Committee on Medical Physics, University of Chicago. Mailing address: 1720 N. Lasalle Dr., Unit 25, Chicago, IL 60614. Telephone: 312-482-8624. Fax: 312-482-9758. Email: d-levin@uchicago.edu

Manuscript received April 1, 2009

point x), as well as its first and second derivatives with respect to x . However, although this is a mathematically correct and complete method of solving the nonlinear BSS problem, it suffers from a practical difficulty: namely, if the dimensionality of state space is high, a great deal of data is required to cover it densely enough in order to calculate these derivatives accurately. The current paper [4] shows how to perform nonlinear BSS by computing higher-order local correlations of the data's velocity, instead of computing derivatives of its second-order correlation. This approach is advantageous because it requires much less data for an accurate computation. For example, in the synthetic speech separation experiment in Section III, the new method can separate two synthetic utterances recorded with a single microphone after minutes of observation, rather than the hours of observation required by the differential geometric method.

The method described in this paper differs significantly from the methods proposed by other investigators because it uses a criterion of statistical independence in (s, \dot{s}) -space, instead of state space. In addition, there are technical differences between the proposed method and conventional ones. First of all, the technique in this paper exploits statistical constraints on the data that are *locally* defined in state space, in contrast to the usual criteria for statistical independence that are *global* conditions on the data time series or its time derivatives [5]. Furthermore, unlike many other methods [6], [7], the mixing function is derived in a constructive, deterministic, and non-parametric manner, without employing iterative algorithms, without using probabilistic learning methods, and without parameterizing it with a neural network architecture or other means. In addition, the proposed method can handle any differentiable mixing function, unlike some other techniques that only apply to a restricted class of mixing functions [8].

The next section describes how to separate two-dimensional data into two one-dimensional source variables. Section III illustrates the method by using it to separate two simultaneous speech-like sounds that are recorded with a single microphone. The implications of this work are discussed in the last section. The appendix describes how the method can be generalized to separate data of arbitrary dimensionality into possibly multidimensional source variables.

II. METHOD

The BSS procedure, which is described in this section, is initiated by constructing scalar functions on the data space from combinations of local velocity correlations. The values of these scalars are invariant under any nonlinear transformations of coordinates on the data space. It is relatively easy to show that separability imposes necessary conditions on these scalar functions in the source coordinate system. Because of their scalarity, these conditions can readily be transferred to the measurement-defined coordinate system (x) , where they can be tested with the data. If the data do not satisfy these necessary conditions, the data are simply not separable. If the data do satisfy these conditions, we show that there is only one possible source coordinate system, and it can be explicitly constructed. The data can then be transformed into

this putative source coordinate system to see if their PDF and/or correlations factorize there. The data are separable if and only if this factorization occurs.

The first step is to construct local correlations of the data's velocity, such as

$$C_{kl\dots}(x) = \langle (\dot{x}_k - \bar{\dot{x}}_k)(\dot{x}_l - \bar{\dot{x}}_l) \dots \rangle_x, \quad (4)$$

where $\bar{\dot{x}} = \langle \dot{x} \rangle_x$, where the bracket denotes the time average over the trajectory's segments in a small neighborhood of x , where $1 \leq k, l \leq N$, and where "... " denotes possible additional indices on the left side and corresponding factors of $\dot{x} - \bar{\dot{x}}$ on the right side. The definition of the PDF implies that this velocity correlation is one of its moments

$$C_{kl\dots}(x) = \frac{\int \rho(x, \dot{x})(\dot{x}_k - \bar{\dot{x}}_k)(\dot{x}_l - \bar{\dot{x}}_l) \dots d\dot{x}}{\int \rho(x, \dot{x})d\dot{x}}, \quad (5)$$

where $\rho(x, \dot{x})$ is the PDF in the x coordinate system. Incidentally, although (5) is useful in a formal sense, in practical applications, all required correlation functions can be computed directly from local time averages of the data ((4)), without explicitly computing the data's PDF. Also, note that velocity "correlations" with a single subscript vanish identically.

Next, let $M(x)$ be a local $N \times N$ matrix, and use it to define M -transformed velocity correlations

$$I_{kl\dots}(x) = \sum_{1 \leq k', l', \dots \leq N} M_{kk'}(x)M_{ll'}(x) \dots C_{k'l'\dots}(x), \quad (6)$$

where "... " denotes possible additional indices of I and C , as well as corresponding factors of $M(x)$. Because $C_{kl}(x)$ is positive definite at any point x , it is always possible to find an $M(x)$ such that

$$I_{kl}(x) = \delta_{kl} \quad (7)$$

$$\sum_{1 \leq m \leq N} I_{klmm}(x) = D_{kl}(x), \quad (8)$$

where $D(x)$ is a diagonal $N \times N$ matrix. Such an $M(x)$ can always be constructed from the product of three matrices: 1) a rotation that diagonalizes $C_{kl}(x)$, 2) a diagonal rescaling matrix that transforms this diagonalized correlation into the identity matrix, 3) another rotation that diagonalizes

$$\sum_{1 \leq m \leq N} C_{klmm}(x),$$

after the fourth-order correlation has been transformed by the first rotation and the rescaling matrix. As long as the last-diagonalized matrix is not degenerate, $M(x)$ is unique, up to arbitrary *local* permutations and reflections. In almost all realistic applications, the velocity correlations will be continuous functions of the state space coordinate x . Therefore, in any neighborhood of state space, there will always be a continuous solution for $M(x)$, and this solution is unique, up to arbitrary *global* reflections and permutations.

In order to show that the M -transformed velocity correlations (i.e., the $I_{klm\dots}(x)$) transform like scalars, imagine constructing these quantities in some other coordinate system

x' . An M -matrix that satisfies (7) and (8) in the x' coordinate system is given by

$$M'_{kl}(x') = \sum_{1 \leq m \leq N} M_{km}(x) \frac{\partial x_m}{\partial x'_l}(x'), \quad (9)$$

where M is a matrix that satisfies (7) and (8) in the x coordinate system. To prove this, substitute this equation into the definition of $I'_{kl\dots}(x')$. Because velocity correlations transform as contravariant tensors, the partial derivative factors within M' transform correlations from the x' coordinate system to the x coordinate system, leading to

$$\begin{aligned} I'_{kl\dots}(x') &= \sum_{1 \leq k', l', \dots \leq N} M'_{kk'}(x') M'_{ll'}(x') \dots C'_{k'l' \dots}(x') \\ &= \sum_{1 \leq k', l', \dots \leq N} M_{kk'}(x) M_{ll'}(x) \dots C_{k'l' \dots}(x) \\ &= I_{kl\dots}(x). \end{aligned}$$

Therefore, because $I_{kl}(x)$ and $I_{klmn}(x)$ satisfy (7) and (8), so do $I'_{kl}(x')$ and $I'_{klmn}(x')$, thereby proving that (9) is one of the solutions for $M'(x')$ in the x' coordinate system. All other solutions for $M'(x')$ differ from this one by global reflections and permutations. Similar reasoning shows that, for any choice of M' and M , each of the functions $I'_{kl\dots}(x')$ equals the corresponding function $I_{kl\dots}(x)$, up to possible global permutations and reflections. In other words,

$$I'_{kl\dots}(x') = \sum_{1 \leq k', l', \dots \leq N} P_{kk'} P_{ll'} \dots I_{k'l' \dots}(x), \quad (10)$$

where $P_{kk'}$ denotes an element of a product of permutation, reflection, and identity matrices. In other words, the functions $I_{kl\dots}(x)$ transform as scalar functions on the state space, except for possible reflections and index permutations.

We now assume that the system is separable and derive some necessary conditions on these scalar functions in the source coordinate system (s). Because these separability conditions involve scalar functions, they can then be transferred to the measurement-defined coordinate system (x), where they can be tested with the data. In order to make the notation simple, it is assumed that $N = 2$ in the following. However, the appendix describes how the methodology can be generalized in order to separate higher-dimensional data into possibly multidimensional source variables.

Separability implies that there is a transformation f^{-1} from the x coordinate system to a source coordinate system (s) in which (3) is true. Because of (5), the velocity correlation functions in the s coordinate system are products of correlations of the independent sources

$$C_{S1\dots 2\dots}(s) = C_{S1\dots}(s_1) C_{S2\dots}(s_2), \quad (11)$$

where $1\dots$ and $2\dots$ denote arbitrary numbers of indices equal to 1 and 2, respectively. It follows from this equation and from the vanishing of all velocity ‘‘correlations’’ with one index that the source variable correlations $C_{Skil}(s)$ and

$$\sum_{1 \leq m \leq N} C_{Skilm}(x)$$

are diagonal. Therefore, in the s coordinate system, (7) and (8) are satisfied by a diagonal matrix $M_S(s)$ of the form

$$M_S(s) = \begin{pmatrix} M_{S1}(s_1) & 0 \\ 0 & M_{S2}(s_2) \end{pmatrix}. \quad (12)$$

It follows from (11) and (12) that the scalar functions $I_{Skil\dots}(s)$ with all subscripts $kl\dots$ equal to 1 (2) must equal the corresponding functions derived for subsystem 1 (2), and these latter functions depend on s_1 (s_2) alone. Although these constraints were derived in the s coordinate system, scalarity ((10)) implies that these separability conditions are true in all coordinate systems, except for possible permutations and reflections. Therefore, in the measurement-defined coordinate system (x), the functions $I_{kl\dots}(x)$ with all subscripts equal to 1 must be functions of either $s_1(x)$ or $s_2(x)$. Likewise, the functions $I_{kl\dots}(x)$ with all subscripts equal to 2 must be functions of the other source variable ($s_2(x)$ or $s_1(x)$, respectively).

This coordinate-system-independent consequence of separability can be used to perform nonlinear BSS in the following manner:

- 1) Use (4) to compute velocity correlations $C_{kl\dots}(x)$ from the data $x(t)$.
- 2) Use linear algebra to find a continuous matrix $M(x)$ that satisfies (7) and (8).
- 3) Use (6) to compute the functions $I_{kl\dots}(x)$.
- 4) Plot the values of the triplets

$$I_A(x) = \{I_{111}(x), I_{1111}(x), I_{11111}(x)\} \quad (13)$$

$$I_B(x) = \{I_{222}(x), I_{2222}(x), I_{22222}(x)\} \quad (14)$$

as x varies over the measurement-defined coordinate system.

- 5) If the plotted values of I_A and/or I_B do not lie in one-dimensional subspaces within the three-dimensional space of the plots, $I_A(x)$ and/or $I_B(x)$ cannot be functions of single source components ($s_1(x)$ or $s_2(x)$) as required by separability, and the data are not separable.
- 6) If the plotted values of both I_A and I_B do lie on one-dimensional manifolds, define one-dimensional coordinates (σ_A and σ_B , respectively) on those subspaces. Then, compute the function $\sigma(x) = (\sigma_A(x), \sigma_B(x))$ that maps each coordinate x onto the value of σ , which parameterizes the point $(I_A(x), I_B(x))$. Notice that, because of the Takens’ embedding theorem [9], x is invertibly related to the six components of $I_A(x)$ and $I_B(x)$, and, therefore, it is invertibly related to σ .
- 7) Transform the PDF (or correlations) of the measurements from the x coordinate system to the σ coordinate system. The data are separable if and only if the PDF factorizes (the correlations factorize) in the σ coordinate system.

The last statement can be understood in the following manner. As shown above, separability implies that $I_A(x)$ must be a function of a single source variable ($s_1(x)$ or $s_2(x)$), and the Takens theorem implies that this function is invertible. Because I_A is also an invertible function of σ_A , it follows that σ_A must be invertibly related to one of the source variables,

and, in a similar manner, σ_B must be invertibly related to the other source variable. Thus, separability implies that σ_A and σ_B are themselves source variables. It follows that the data are separable if and only if the PDF factorizes in the σ coordinate system.

Although the above-described procedure will perform BSS for any mixing function, it is interesting to consider the special case in which source variables exist that are linearly related to the measurements; namely,

$$s_k = \sum_{1 \leq l \leq 2} W_{kl} x_l, \quad (15)$$

where W is a constant 2×2 matrix. In general, the above BSS procedure will construct source variables σ that are related to these “linear” source variables by

$$\sigma_1(x) = g_1(s_1(x)) = g_1(W_{11}x_1 + W_{12}x_2) \quad (16)$$

$$\sigma_2(x) = g_2(s_2(x)) = g_2(W_{21}x_1 + W_{22}x_2), \quad (17)$$

where g_1 and g_2 are some invertible nonlinear transformations determined by the choice of the σ_1 and σ_2 coordinates, respectively. Therefore, at each point x the partial derivatives

$$\begin{aligned} \partial\sigma_1/\partial x_k \\ \partial\sigma_2/\partial x_k \end{aligned}$$

will be proportional to constant (i.e., x -independent) vectors (denoted by U_{1k} and U_{2k}), which are themselves proportional to the first and second rows of W , respectively. Furthermore, these vectors can be used to construct other linearly-related source variables

$$\hat{s}_k = \sum_{1 \leq l \leq 2} U_{kl} x_l \quad (18)$$

that are just rescaled versions of the ones in (15). Consequently, given the source variables $\sigma(x)$ produced by the BSS procedure, the following process can be used to determine whether these can be transformed into source variables that are linearly related to the measurements: 1) compute the above-mentioned partial derivatives and determine if each is proportional to an x -independent vector; 2) if the partial derivatives do not satisfy this condition, there are no linearly-related source variables; 3) if the partial derivative do satisfy condition 1, transform the data into the \hat{s} coordinate system in order to see if the data’s PDF factorizes there. There are linearly-related source variables if and only if this factorization occurs.

III. NUMERICAL EXAMPLE: SEPARATING TWO SPEECH-LIKE SOUNDS RECORDED WITH A SINGLE MICROPHONE

This section describes a numerical experiment in which two speech-like sounds were synthesized and then summed, as if they were simultaneously recorded with a single microphone. Each sound simulated an “utterance” of a vocal tract resembling a human vocal tract, except that: 1) it had one degree of freedom, instead of the 3-5 degrees of freedom of the human vocal tract; 2) its impulse response was characterized by one pole pair, instead of the 4-6 pole pairs characteristic of

the human vocal tract. The methodology of Section II was blindly applied to a time series of two features extracted from 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). BSS was performed with only 16 minutes of data, instead of the hours of data required to separate similar sounds using a differential geometric method [3].

Each speaker was simulated by having a simulated glottis drive a simulated resonant cavity that represented the vocal tract. The glottal waveform of the each “voice” was a series of spikes separated by a pitch interval (100 Hz and 160 Hz). The impulse response of each “vocal tract” was taken to be a characteristic damped sinusoid, whose amplitude, resonant frequency, and damping were linear functions of a single state variable. For each voice, a 16 minute utterance was produced by convolving its glottal waveform with the impulse response of its vocal tract, which was a function of a slowly-varying state variable. The state variable time series of each voice was synthesized by smoothly interpolating among successive states. The latter were chosen at 100-120 msec intervals so that the state variable time series of the two voices were statistically independent of each other. The resulting utterances had energies differing by 2.4 dB, and they were summed and sampled at 16 kHz with 16-bit depth. Then, this “recorded” waveform was pre-emphasized and subjected to a short-term Fourier transform (using frames with 25 msec length and 5 msec spacing). The log energies of a bank of 12 mel-frequency filters between 0-8000 Hz were computed for each frame, and these were then averaged over pairs of consecutive frames. These log filterbank outputs were nonlinear functions of the two vocal tract state variables.

In order to blindly analyze these data, we first determined if any data components were redundant in the sense that they were simply functions of other components. Fig. 1a shows the first three principal components of the log filterbank outputs during a typical short recording of the simultaneous utterances. Inspection showed that these data lay on a curved two-dimensional surface within the ambient 12-D space, making it apparent that they were produced by a “hidden” dynamical system with two degrees of freedom. The redundant components were eliminated by using dimensional reduction (principal components analysis in small overlapping neighborhoods of the data) to establish a coordinate system x on this surface and to find $x(t)$, the trajectory of the recorded sound in that coordinate system. Next, the BSS procedure in Section II was used to determine if $x(t)$ was a nonlinear mixture of two source variables that were statistically independent of one another. Following steps 1-4 of the BSS procedure, $x(t)$ of the entire recording was used to compute invariants $I_{kl\dots}(x)$ with up to five indices, and the related functions $I_A(x)$ and $I_B(x)$ were plotted, as illustrated in Figs. 1b-c. It was evident that the plotted values of both I_A and I_B lay in or close to one-dimensional subspaces. Following step 6 of the BSS procedure, a dimensional reduction procedure [12] was used to define coordinates (σ_A and σ_B) on these one-dimensional manifolds, and $\sigma(x) = (\sigma_A(x), \sigma_B(x))$ was computed. If the data were separable, σ must be a set of source variables, and

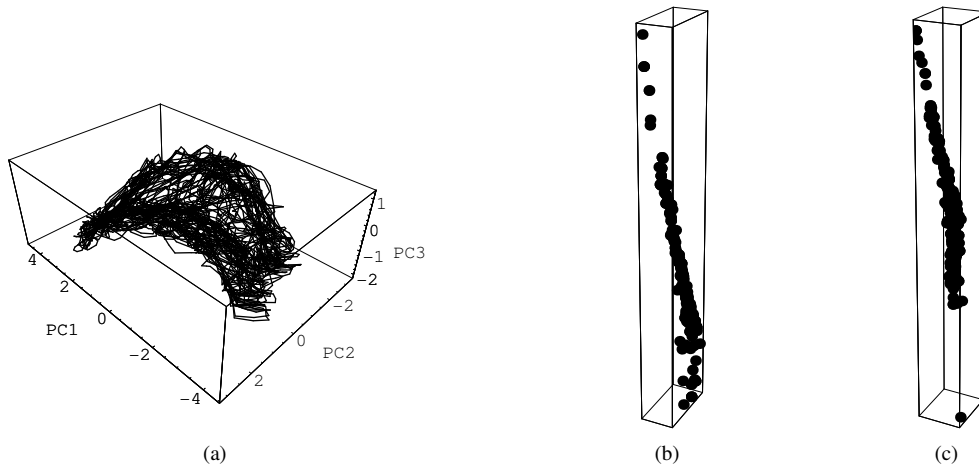


Fig. 1. (a) The first three principal components of log filterbank outputs of a typical short recording of two simultaneous speech-like sounds. (b) The distribution of the values of $I_A(x)$ ((13)), as x varied over the approximately two-dimensional manifold in (a). (c) The distribution of the values of $I_B(x)$ ((14)), as x varied over the approximately two-dimensional manifold in (a).

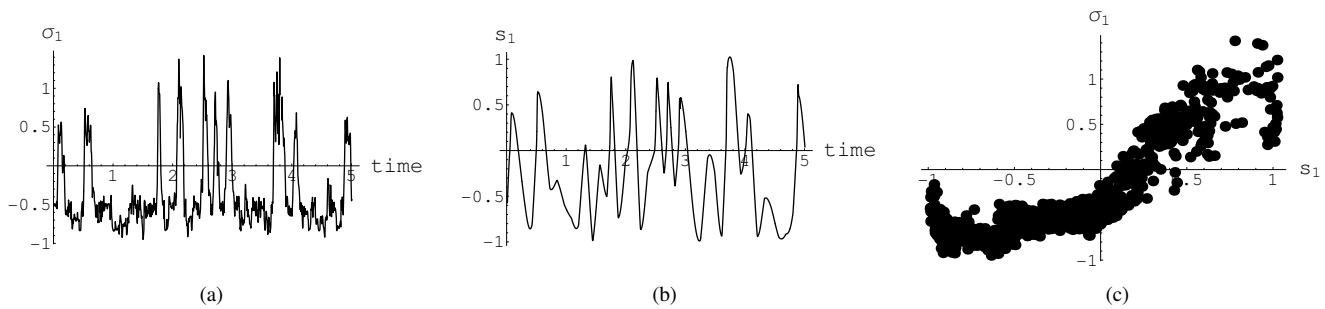


Fig. 2. (a) The time dependence of one of the source variables, blindly computed from a typical five-second segment of the data's trajectory $x(t)$, by finding the coordinates of $I_A[x(t)]$ on the one-dimensional manifold in Fig. 1b. (b) The state variable time series originally used to generate one of the speech-like sounds during the five-second recording analyzed in (a). (c) The scatter plot of the pairs of source and state variable values in (a) and (b).

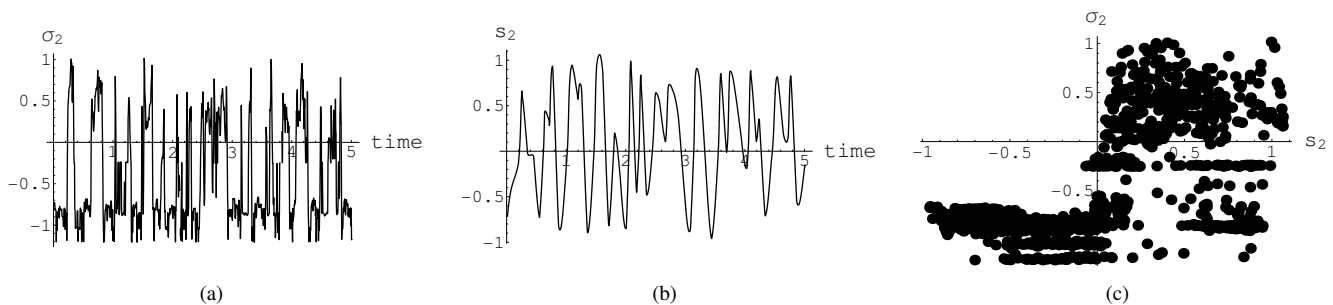


Fig. 3. (a) The time dependence of one of the source variables, blindly computed from a typical five-second segment of the data's trajectory $x(t)$, by finding the coordinates of $I_B[x(t)]$ on the one-dimensional manifold in Fig. 1c. (b) The state variable time series originally used to generate one of the speech-like sounds during the five-second recording analyzed in (a). (c) The scatter plot of the pairs of source and state variable values in (a) and (b).

$\sigma[x(t)]$ must describe the evolution of the underlying vocal tract states (up to invertible component-wise transformations). As illustrated in Figs. 2a-b and Figs. 3a-b, the time courses of the putative source variables ($\sigma_A[x(t)], \sigma_B[x(t)]$) did resemble distorted versions of the state variable time series that were originally used to generate the voices' utterances. The scatter plots in Fig. 2c and Fig. 3c show that, in each case, the recovered source variable and the corresponding state variable were related by a nonlinear transformation that was nearly monotonic, except for the effects of noise due

to the limited number of data samples. Thus, starting with a single-microphone recording, the BSS procedure was able to extract the information encoded in the time series of each speaker's state variable. The time course of the analogous multidimensional state variable of the human vocal tract contains the speech content of each utterance. This indicates that the BSS procedure is capable of recovering the speech content of superposed utterances, without recovering their original waveforms.

IV. DISCUSSION

In a previous paper [3], the nonlinear BSS problem was reformulated in (state, state velocity)-space, instead of state space as in conventional formulations. This approach is attractive because: 1) the reformulated 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 permutations and transformations of independent source variables); 2) statistical independence in (state, state velocity)-space is manifested by almost all classical physical systems that are composed of non-interacting subsystems. This paper [4] shows how a general solution of this problem can be constructed in a deterministic manner, which avoids the difficulties of the iterative, probabilistic, and parametric BSS techniques proposed by other investigators. Furthermore, an accurate computation can be performed with far less data than that required by the differential geometric solution, previously proposed by the author [3].

The BSS procedure in Section II shows how to compute $(\sigma_A[x(t)], \sigma_B[x(t)])$, the trajectory of each independent subsystem in a specific coordinate system on that subsystem's state space. In many practical applications, a pattern recognition "engine" has been trained to recognize the meaning of trajectories of one subsystem (e.g., "A") in another coordinate system (e.g., s_A) on that subsystem's state space. In order to use this information, it is necessary to know the transformation to this particular coordinate system ($\sigma_A \rightarrow s_A$). For example, subsystem A may be the vocal tract of speaker A , and subsystem B may be a noise generator of some sort. In this example, we may have trained an automatic speech recognition (ASR) engine on the quiet speech of speaker A (or, equivalently, on the quiet speech of another speaker who mimics A in the sense that their state space trajectories are related by an invertible transformation when they speak the same utterances). In order to recognize the speaker's utterances in the presence of B , we must know the transformation from the vocal tract coordinates recovered by BSS (σ_A) to the coordinates used to train the ASR engine (s_A). This mapping can be determined by using the training data to compute more than $2d_A$ invariants (like those in (6)) as functions of s_A . These must equal the invariants of one of the subsystems identified by the BSS procedure, up to a global permutation and/or reflection ((10)). This global transformation can be determined by permuting and reflecting the distribution of invariants produced by the training data, until it matches the distribution of invariants of one of the subsystems produced by the BSS procedure. Then, the mapping $\sigma_A \rightarrow s_A$ can be determined by finding paired values of σ_A and s_A that correspond to the same invariant values within these matching distributions. This type of analysis of human speech data is currently underway.

APPENDIX

SEPARATING DATA OF ANY DIMENSIONALITY INTO POSSIBLY MULTIDIMENSIONAL SOURCES

The procedure in Section II is capable of separating two-dimensional data into one-dimensional source variables. This

appendix describes the solution of the more general nonlinear BSS problem in which data of any dimensionality may be separated into possibly multidimensional source variables, each of which is statistically independent of the others but each of which may contain statistically dependent components. This is sometimes called multidimensional independent component analysis, subspace independent component analysis, or independent subspace analysis [10], [11].

Separability implies that there is a transformation f^{-1} from the x coordinate system to a source coordinate system ($s = (s_A, s_B)$) in which

$$\rho_S(s, \dot{s}) = \rho_A(s_A, \dot{s}_A) \rho_B(s_B, \dot{s}_B), \quad (19)$$

where s_A is a possibly multidimensional source variable with d_A components and s_B is a possibly multidimensional source variable with $d_B = N - d_A$ components. Because of (5), the velocity correlation functions in the s coordinate system are products of correlations of independent sources

$$C_{S_{a\dots b\dots}}(s) = C_{S_{a\dots}}(s_A) C_{S_{b\dots}}(s_B), \quad (20)$$

where $a\dots$ and $b\dots$ denote arbitrary series of indices in the ranges $1 \leq a \leq d_A$ and $d_A + 1 \leq b \leq N$, respectively. It follows from this equation and from the vanishing of all velocity "correlations" with one index that the source variable correlations $C_{S_{kl}}(s)$ and

$$\sum_{1 \leq m \leq N} C_{S_{klmm}}(x)$$

have block-diagonal forms with $d_A \times d_A$ and $d_B \times d_B$ upper and lower blocks, respectively. Consequently, in the s coordinate system, (7) and (8) are satisfied by a block-diagonal matrix $M_S(s)$ of the form

$$M_S(s) = \begin{pmatrix} M_{SA}(s_A) & 0 \\ 0 & M_{SB}(s_B) \end{pmatrix}, \quad (21)$$

where M_{SA} and M_{SB} are matrices that satisfy (7) and (8) for the A and B subsystems, respectively. In order to prove that (21) satisfies (7), substitute it into the definition of $I_{S_{kl}}$, and note that each block of M_S is defined to transform the corresponding block of $C_{S_{kl}}$ into an identity matrix. In order to prove that (21) satisfies (8), substitute it into the definition of

$$\sum_{1 \leq m \leq N} I_{S_{klmm}}. \quad (22)$$

Then, note that: 1) when k and l belong to different blocks, each term in this sum vanishes because it factorizes into a product of a one-index correlation and a three-index correlation; 2) when k and l belong to the same block and are unequal, each term with m in the other block contains a factor equal to $I_{S_{kl}}$, which vanishes, as proved above; 3) when k and l belong to the same block and are equal, the sum over m in the same block vanishes, because each block of M_S is defined to satisfy (8) for the corresponding subsystem.

It follows from (6), (20), and (21) that the scalar functions $I_{S_{kl\dots}}(s)$ with all subscripts $kl\dots$ in the range $1 \leq k, l \leq d_A$ (or in the range $d_A + 1 \leq k, l \leq N$) must depend on s_A (or s_B) alone. Although these constraints were derived in

the s coordinate system, scalarity ((10)) implies that these separability conditions are true in all coordinate systems, except for possible permutations. Therefore, in the measurement-defined coordinate system (x), it must be possible to partition the indices of x ($k = 1, 2, \dots, N$) into A and B groups (containing d_A “ A ” indices and d_B “ B ” indices, respectively) so that the functions $I_{kl\dots}(x)$ with all subscripts in the A (or B) group are functions of $s_A(x)$ (or $s_B(x)$) alone.

This coordinate-system-independent consequence of separability can be used to perform nonlinear BSS in the following manner:

- 1) Use (4) to compute velocity correlations $C_{kl\dots}(x)$ from the data $x(t)$.
- 2) Use linear algebra to find a continuous matrix $M(x)$ that satisfies (7) and (8).
- 3) Use (6) to compute the functions $I_{kl\dots}(x)$.
- 4) Consider each choice of an integer d_A in the range $1 \leq d_A < N$, and consider each way of partitioning the data indices ($k = 1, 2, \dots, N$) into A and B groups (containing d_A “ A ” indices and $d_B = N - d_A$ “ B ” indices, respectively). For each of these choices, let $I_A(x)$ ($I_B(x)$) be any set of more than $2d_A$ ($2d_B$) of the functions $I_{kl\dots}(x)$ for which all subscripts belong to the A (B) group, and plot the values of $I_A(x)$ and $I_B(x)$ as x varies over the measurement-defined coordinate system.
- 5) Suppose that, for all of the choices in step 4, the plotted values of I_A and/or I_B *do not* lie in d_A -dimensional (d_B -dimensional) subspaces within the higher-dimensional space of the plots. Then, there is no way that $I_A(x)$ and $I_B(x)$ can be functions of single source variables ($s_A(x)$ and $s_B(x)$) as required by separability, and the data are not separable.
- 6) Suppose that, for one or more of the choices in step 4, the plotted values of both I_A and I_B *do* lie in d_A -dimensional and d_B -dimensional manifolds, respectively. In that case, define d_A -dimensional (d_B -dimensional) coordinates σ_A (σ_B) on those subspaces. Then, compute the function $\sigma(x) = (\sigma_A(x), \sigma_B(x))$ that maps each coordinate x onto the value of σ , which parameterizes the point $(I_A(x), I_B(x))$. Notice that, because of the Takens’ embedding theorem [9], x is invertibly related to the $2N + 2$ (or more) components of $I_A(x)$ and $I_B(x)$, and, therefore, it is invertibly related to σ .
- 7) Transform the PDF (or correlations) of the measurements from the x coordinate system to the σ coordinate system. The data are separable, and σ_A and σ_B are source variables if and only if the PDF factorizes (the correlations factorize) in a σ coordinate system created in this way.

The last statement can be understood in the following manner. As shown above, separability implies that, for some choice of d_A and index partitioning, $I_A(x)$ must be a function of $s_A(x)$, and the Takens theorem implies that this function is invertible. Because I_A is also an invertible function of σ_A , it follows that σ_A must be invertibly related to s_A , and,

in a similar manner, σ_B must be invertibly related to s_B . Thus, separability implies that σ_A and σ_B are themselves source variables and, therefore, the PDF factorizes in the σ coordinate system. Finally, note that, if the data are separable, the same procedure can then be used to determine if each multicomponent source variable (σ_A or σ_B) can be further separated into lower-dimensional source variables.

The above-described procedure will perform BSS for any linear or nonlinear mixing. However, a few comments should be made about the special case in which source variables exist that are linearly related to the measurements; namely,

$$s_{Aa} = \sum_{1 \leq k \leq N} W_{Aak} x_k \quad (23)$$

$$s_{Bb} = \sum_{1 \leq k \leq N} W_{Bbk} x_k, \quad (24)$$

where $1 \leq a \leq d_A$, where $1 \leq b \leq d_B$, and where W_A and W_B are constant $d_A \times N$ and $d_B \times N$ matrices, respectively. In general, the above BSS procedure will construct source variables σ that are related to these “linear” source variables by

$$\sigma_A = g_A(s_A) \quad (25)$$

$$\sigma_B = g_B(s_B), \quad (26)$$

where g_A and g_B are some nonlinear transformations (with d_A and d_B components, respectively), determined by the choice of the σ_A and σ_B coordinates, respectively. Therefore, at each point x , the sets of partial derivatives

$$\begin{aligned} \partial \sigma_{Aa} / \partial x_k & \text{ for } 1 \leq a \leq d_A \\ \partial \sigma_{Bb} / \partial x_k & \text{ for } 1 \leq b \leq d_B \end{aligned}$$

will lie in the subspace spanned by the rows of W_A and W_B , respectively. Let U_{Aak} and U_{Bbk} (for $1 \leq a \leq d_A$ and $1 \leq b \leq d_B$) denote any sets of constant vectors that span these two subspaces. Then, another set of linearly-related source variables is given by

$$\hat{s}_{Aa} = \sum_{1 \leq k \leq N} U_{Aak} x_k \quad (27)$$

$$\hat{s}_{Bb} = \sum_{1 \leq k \leq N} U_{Bbk} x_k, \quad (28)$$

which are just linear combinations of the ones in (23) and (24), respectively. Consequently, given the source variables $\sigma(x)$ produced by the BSS procedure, the following process can be used to determine whether these can be transformed into source variables that are linearly related to the measurements: 1) compute the above-mentioned sets of partial derivatives and determine if each set is spanned by the appropriate number of x -independent vectors; 2) if one or both sets of partial derivatives do not satisfy this condition, there are no linearly-related source variables; 3) if both sets of partial derivative do satisfy condition 1, transform the data into the \hat{s} coordinate system in order to see if the data’s PDF factorizes there. There are linearly-related source variables if and only if this factorization occurs.

REFERENCES

- [1] C. Jutten and J. Karhunen, "Advances in nonlinear blind source separation," in *Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, April 2003.
- [2] A. Hyvärinen and P. Pajunen, "Nonlinear independent component analysis: existence and uniqueness results," *Neural Networks*, vol. 12, pp. 429-439, 1999.
- [3] D. N. Levin, "Using state space differential geometry for nonlinear blind source separation," *J. Applied Physics*, vol. 103, art. no. 044906, 2008.
- [4] D. N. Levin, "Using signal invariants to perform nonlinear blind source separation," in *Proc. of the 8th International Conference on Independent Component Analysis and Signal Separation*, LNCS, vol. 5441, T. Adali, C. Jutten, J. M. T. Romano, A. K. Barros (eds). Heidelberg: Springer, 2009, pp. 58-65.
- [5] S. Lagrange, L. Jaulin, V. Vigneron, C. Jutten, "Analytic solution of the blind source separation problem using derivatives," in *Independent Component Analysis and Blind Signal Separation*, LNCS, vol. 3195, C. G. Puntonet and A. G. Prieto (eds). Heidelberg: Springer, 2004, pp. 81-88.
- [6] H. H. Yang, S-I. Amari, A. Cichocki, "Information-theoretic approach to blind separation of sources in non-linear mixture," *Signal Processing*, vol. 64, pp. 291-300, 1998.
- [7] S. Haykin, *Neural Networks - A Comprehensive Foundation*. New York: Prentice Hall, 1998.
- [8] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures," *IEEE Trans. Signal Process.*, vol. 47, pp. 2807-2820, 1999.
- [9] T. Sauer, J. A. Yorke, M. Casdagli, "Embedology," *J. Statistical Physics*, vol. 65, pp. 579-616, 1991.
- [10] J.-F. Cardoso, "Multidimensional independent component analysis," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, Seattle, 12-15 May, 1998, pp. 1941-1944.
- [11] Y. Nishimori, S. Akaho, M. D. Plumbley, "Riemannian optimization method on the flag manifold for independent subspace analysis," in *Proc. 6th International Conference on Independent Component Analysis and Blind Source Separation*, LNCS, vol. 3889. Berlin: Springer, 2006, pp. 295-302
- [12] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323-2326, 2000.