Separation of a Mixture of Independent Signals Using Time Delayed Correlations

L. Molgedey, and H. G. Schuster

Institut für Theoretische Physik, Olshausenstr. 40, D-24118 Kiel 1, Germany

Abstract

The problem of separating n linearly superimposed uncorrelated signals and determing their mixing coefficents is reduced to an Eigenvalue problem which involves the simultaneous diagonalisation of two symmetric matrices whose elements are measureable time delayed correlation functions. The diagonalisation matrix can be determined from a cost function whose number of minima is equal the number of degenerate solutions. Our approach offers the possibility to separate also nonlinear mixtures of signals.

PACS numbers:02.50.+s, 05.40.+j, 06.50.-x, 87.71.-p

The problem of source separation appears in many contexts. The most simple situation occurs for two speakers. If the mixture of their voices reaches two microphones one wants to separate both sources such that each detector registers only one voice [1]. Typical examples involving many sources and many receivers are the separation of radio or radar signals by an array of antennas [2], the separation of odors in a mixture by an array of sensors, the parsing of the environment into different objects by our visual system [3] or the separation of biomagnetic sources by an array of SQUIDS in magnetoencephalography [4].

In 1986 Jutten and Herault [5] proposed an adaptive neural network to perform this task. It decorrelates the incoming signals via an inhibitory interaction between the output neurons. These authors [5–7] and recently Hopfield [8] have demonstrated by ways of numerical simulations that their method often works. However is range of applicability ist still open and there are situations where it fails.

In this Letter we include more information about the time structure of the sources into the adaptation process for the inhibitory interactions, i.e., we require that not only the equal time but also the time delayed correlations between the different output signals vanish. This leads to the following results:

- The problem of separating n linear superimposed uncorrelated sources and determing their mixing coefficents is reduced to an eigenvalue problem which requires the simultaneous diagonalization of two symmetric matrices.
- The learning rule for the lateral inhibitory interactions between the neurons is given by the gradient of a cost function whose number of minima is equal to the number of degenerate solutions.
- For Gaussian sources we find qualitatively the same equations of motion for the inhibitory interactions as Jutten and Herauld but augmented by contributions arising from the delay terms that are necessary for convergence.

The source separation problem can be stated mathematically as follows. Assuming that the number of sources and detectors are equal, the input I_i $(i = 1 \ldots n)$ to each receiver is a linear mixture $I_i(t) = \sum_{j=1}^{n} C_{ij} a_j(t)$ of statistical independent equilibrium signals i.e. $\langle a_i(t) a_j(t') \rangle = K_i(|t - t'|)\delta_{ij}$. Without restriction we assume that the mean value of the signals is zero $\langle a_i(t) \rangle = 0$. The problem is now to determine the coefficients C_{ij} and the source strengths $\lambda_i = K_i(0)$ from a measurement of $I_i(t)$.

Since the matrix $\mathbf{C}$ is generally not symmetric it is not suffficient to measure the symmetric correlation matrix $\langle I_i(t) I_j(t) \rangle = M_{ij}$. Jutten and Herault [6] proposed to measure nonlinear correlations like $\langle I_i(t) I_j(t)^3 \rangle$ which are nonsymmetric. Instead we suggest to measure in addition to M_{ij} the time delayed correlation matrix $< I_i(t) I_j(t+\tau) \rangle = \bar{M}_{ij}$. This yields $n(n+1)$ equations

$$M_{ij} = \sum_l C_{il} C_{jl} \lambda_l \qquad \text{and} \qquad \bar{M}_{ij} = \sum_l C_{il} C_{jl} \bar{\lambda}_l \tag{1}$$

for the $n(n+1)$ unknowns $C_{i \neq j}$, λ_i and $\bar{\lambda}_i = K_i(\tau)$.

If the mixing is linear independent i.e. $\det \mathbf{C} \neq 0$ and the time delay parameter τ has been chosen such that $K_i(0) K_j(\tau) = \lambda_i \bar{\lambda}_j \neq \bar{\lambda}_i \lambda_j = K_i(\tau) K_j(0)$ for all $i \neq j$, the problem is solvable up to $n!$ trivial permutations.

Equation (1) shows that by construction the matrix $\mathbf{C}$ diagonalizes $\mathbf{M}$ and $\bar{\mathbf{M}}$ simultaneously i.e. $\mathbf{C}^{-1}\mathbf{M}(\mathbf{C}^{\mathrm{T}})^{-1} = \mathbf{\Lambda}$ and $\mathbf{C}^{-1}\bar{\mathbf{M}}(\mathbf{C}^{\mathrm{T}})^{-1} = \bar{\mathbf{\Lambda}}$. But the elements of $\Lambda_{ij} = \lambda_i \delta_{ij}$ and $\bar{\Lambda}_{ij} = \bar{\lambda}_i \delta_{ij}$ are not simply the eigenvalues of the matrices $\mathbf{M}$ and $\bar{\mathbf{M}}$ because generally $\mathbf{C}$ is not an orthogonal matrix. Instead equation (1) leads to the eigenvalue problem

$$(\mathbf{M}\bar{\mathbf{M}}^{-1})\mathbf{C} = \mathbf{C}(\mathbf{\Lambda}\bar{\mathbf{\Lambda}}^{-1}). \tag{2}$$

We note that usually $\mathbf{M}\bar{\mathbf{M}}^{-1}$ is not symmetric and the diagonal elements of $\mathbf{C}$ are normalized to unity. Equation (2) can be solved by standard techniques of numerical linear algebra.

In order to compare our method to that of Jutten and Herault [5] and Hopfield [8] we next we proceed to solve eqn. (2) by an neural network whose architecture for $n = 2$ is shown in Figure 1. We follow [5–8] and use linear neurons such that the output is determined by:

$$u_i(t+1) = -\sum_{j=1}^{n} T_{ij} u_j(t) + I_i(t), \tag{3}$$

where $\mathbf{T}$ is the matrix of inhibitory connections with zero diagonal elements. We also assume as in [5–8] that the time variation of the signals is slow, so that eqn. (3) can be solved as

$$\vec{u}(t) = (\mathbf{1} + \mathbf{T})^{-1} \vec{I}(t). \tag{4}$$

The matrix $\mathbf{T}$ is determined by the minima of the cost function

$$V\{T_{pq}\} = \sum_{i \neq j} \langle u_i(t) u_j(t) \rangle^2 + \langle u_i(t) u_j(t+\tau) \rangle^2, \tag{5}$$

which occurs if the output correlations between different neurons vanish i.e. $\langle u_i(t)u_j(t)\rangle$=$\langle u_i(t)u_j(t+\tau)\rangle = 0$ for $i \neq j$.

By using the explicit form of $\vec{u}$ according to eqn. (4) this means that the matrices $(\mathbf{1}+\mathbf{T})^{-1}\mathbf{M}(\mathbf{1}+\mathbf{T}^{\mathrm{T}})^{-1}$ and $(\mathbf{1}+\mathbf{T})^{-1}\bar{\mathbf{M}}(\mathbf{1}+\mathbf{T}^{\mathrm{T}})^{-1}$ are diagonalised by $(\mathbf{1}+\mathbf{T})$. Therefore at the minima of $V\{T_{pq}\}$ the interaction matrix yields up to a permutation $\mathbf{P}$ the elements of the mixing matrix: $(\mathbf{1}+\mathbf{T}) = \mathbf{P}\,\mathbf{C}$, where $\mathbf{P}$ is a permutation matrix. The elements T_{ij} can be determined from V by gradient descent:

$$\dot{T}_{pq} \propto -\frac{\partial V}{\partial T_{pq}}. \tag{6}$$

To compare our result with that of Jutten and Herault, we consider the case $n = 2$ for Gaussian signals. Then we obtain from eqn. (6)

$$\begin{aligned} \dot{T}_{12} &\propto \langle I_2(t)u_2(t)\rangle\langle u_1(t)u_2(t)\rangle + \langle I_2(t)u_2(t+\tau)\rangle\langle u_1(t)u_2(t+\tau)\rangle \\ \dot{T}_{21} &\propto \langle I_1(t)u_1(t)\rangle\langle u_1(t)u_2(t)\rangle + \langle I_1(t+\tau)u_1(t)\rangle\langle u_1(t)u_2(t+\tau)\rangle \end{aligned} \tag{7}$$

and from eqn. (34) of [6]

$$\begin{aligned} \dot{T}_{12} &\propto \langle u_1(t)u_2(t)^3\rangle = 3\langle u_2(t)u_2(t)\rangle\langle u_1(t)u_2(t)\rangle \\ \dot{T}_{21} &\propto \langle u_2(t)u_1(t)^3\rangle = 3\langle u_1(t)u_1(t)\rangle\langle u_1(t)u_2(t)\rangle \end{aligned} \tag{8}$$

If we neglect in eqn. (7) the delay terms, then eqns. (7) and (8) yield via $\langle u_1(t)u_2(t)\rangle = g(T_{12}, T_{21}) = 0$ the same lines of fixed points shown in Figure 2a. Only the inclusion of the delay terms i.e. the full eqn. (7) drives the system to the correct pair of stable fixed points $T_{12} = C_{12}$, $T_{21} = C_{21}$ and $T'_{12} = 1/C_{21}$, $T'_{21} = 1/C_{12}$ depicted in Figure 2b.

In Figure 3 we compare the least squares method [9] with our approach for experimental speech signals (cries from different babies [10]) which have been mixed by a matrix with offdiagonal elements $C_{12} = 0.9$ and $C_{21} = 0.7$. It follows again that the use of time delayed correlation functions improves the source separation process.

Up to now we have only considered situations where the number of sources is equal to the number of detectors. If the number m of sources is smaller than the number of sensors n, i.e. $m \langle n$ the activity of $n-m$ neurons will vanish. The simplest case is the situation when one source is fed to two neurons. After the adaptation process the output of one neuron will be proportional to the source and the other neuron will be silent.

If the number of sources is larger than the number of neurons our potential yields always decorrelated outputs, but the mixing matrix $\mathbf{T}$ will not be correct. In order to decorrelate an unknown number of linearly mixed sources one must therfore apply our approch with an increasing number n of output neurons, until n ist so large, say $n = n^*$, that for the first time one neuron will remain silent, after the adaptation process. The number of sources is then $n^* - 1$ and one needs $n^* - 1$ neurons to decorrelate them [11].

Let us finally discuss the situation for nonlinear mixing. An example is shown in eqn. (9)

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} 1 & c_{12} \\ c_{21} & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_1 \end{pmatrix} a_1 a_2 \qquad (9)$$

where ε_1, ε_2 are nonlinearity parameters. In this case the neural network will completely decorrelate $\langle u_1(t)u_2(t)\rangle = \langle u_1(t)u_2(t+\tau)\rangle = 0$ but $\langle u_1(t)u_2(t+2\tau)\rangle$ is still a function of the nonlinearity parameters, as shown in Figure 4. Therefore our method enables us to detect nonlinearities in the mixing of the sources. On the other hand on could determine the linear and nonlinear mixing coefficients $c_{12}, c_{21}, \varepsilon_1, \varepsilon_2$ from the measureable time delayed correlation functions $\langle I_1(t)I_2(t+\tau)\rangle$ including more and more different delays [12]. In this sense our approch which involves time delayed correlation functions could be generalized to solve the source separation problem for nonlinearly mixed sources.

It is a pleasure to thank R. Hornreich for stimulating discussions during the initial steps of this work and the Deutsche Forschungsgemeinschaft for financial support.

REFERENCES

[1] H.W. Strube, Signal Processing **3**, 355, (1981).

[2] F. Cristopher and C. Morrisean, X. Coll. GRESI, 1017, (1985).

[3] C. Gray and W. Singer, Nature **338**, 334, (1989).

[4] S. Williamson, M. Hoke, G. Stroink and M. Kotani, eds. Advances in Biomagnetism (Plenum), New York (1990).

[5] C. Jutten, J. Herault, Neural Network for Computing, AIP Conference Proc. **151**, J. S. Decker, ed. 206, (1986).

[6] C. Jutten, J. Herault and Guerin A., Artifical Intelligence and Cognitive Sciences, 231-248, Manchester Press (1988).

[7] C. Jutten, J. Herault, Signal Processing **24**, 1, (1991).

[8] J. J. Hopfield, Proc. Natl. Acad. Sci. (USA) **88**, 6462, (1991).

[9] B. Widrow, J. M. McCool, M. G. Larimore and C. R. Johnson, Proc. IEEE **64**, 1151, (1976).

[10] J. Hirschberg and T. Szende, Pathologische Schreistimme - Stridor und Hustenton im Säuglinsalter, Fischer Stuttgart, (1985), p. 16 cry 3 and p. 19 cry 7.

[11] Even for a situation where the number of signals m is larger than the number of detectors n, ie. $m > n$ but $m \leq n(n+1)/2$, one could still determine the $m \times n$ mixing matrix C_{ij} and the Nm time delayed correlations of the sources $\langle a_i(t)a_i(t+l\tau)\rangle$ $(l = 0, ..., N-1)$ from the $Nn(n+1)/2$ measured correlation functions $\langle I_i(t)I_j(t+l\tau)\rangle$ $(l = 0, ..., N-1)$. One has only to chose the number N of delays large enough to ensure that the number of measurable variables $Nn(n+1)/2$ becomes larger then the number of unknowns $m(n+N-1)$. However the signals $a_i(t)$ $i = 1, ..., m$ cannot be extracted form $I_i(t)$ $i = 1, ..., n$ because the $m \times n$ mixing matrix C_{ij} cannot be inverted.

[12] To determine the mixing coefficients we have to solve the equations for $k = 0, 1, 2, 3$

$$\langle I_1(t) I_1(t + k\tau)\rangle = K_1(k\tau) + c_{12}^2 K_2(k\tau) + \varepsilon_1 \varepsilon_2 K_1(k\tau) K_2(k\tau)$$

$$\langle I_1(t) I_2(t + k\tau)\rangle = c_{21} K_1(k\tau) + c_{12} K_2(k\tau) + \varepsilon_1 \varepsilon_2 K_1(k\tau) K_2(k\tau)$$

$$\langle I_2(t) I_2(t + k\tau)\rangle = c_{21}^2 K_1(k\tau) + K_2(k\tau) + \varepsilon_1 \varepsilon_2 K_1(k\tau) K_2(k\tau).$$

This are 12 equations for the 12 unknown parameters $c_{12}, c_{21}, \varepsilon_1, \varepsilon_2, K_1(0), K_1(\tau), K_1(2\tau), K_1(3\tau), K_2(0), K_2(\tau), K_2(2\tau), K_2(3\tau)$. They can be solved by standard methods or by a nonlinear neural network using our potential approach.

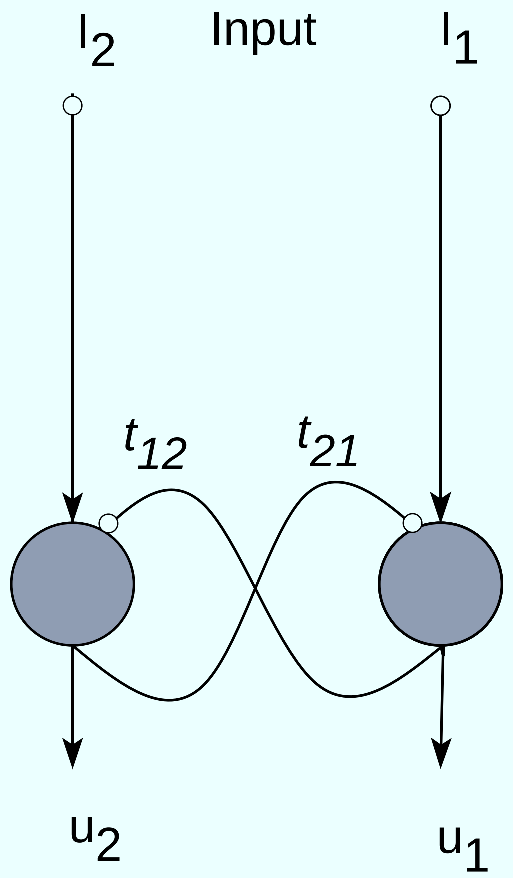
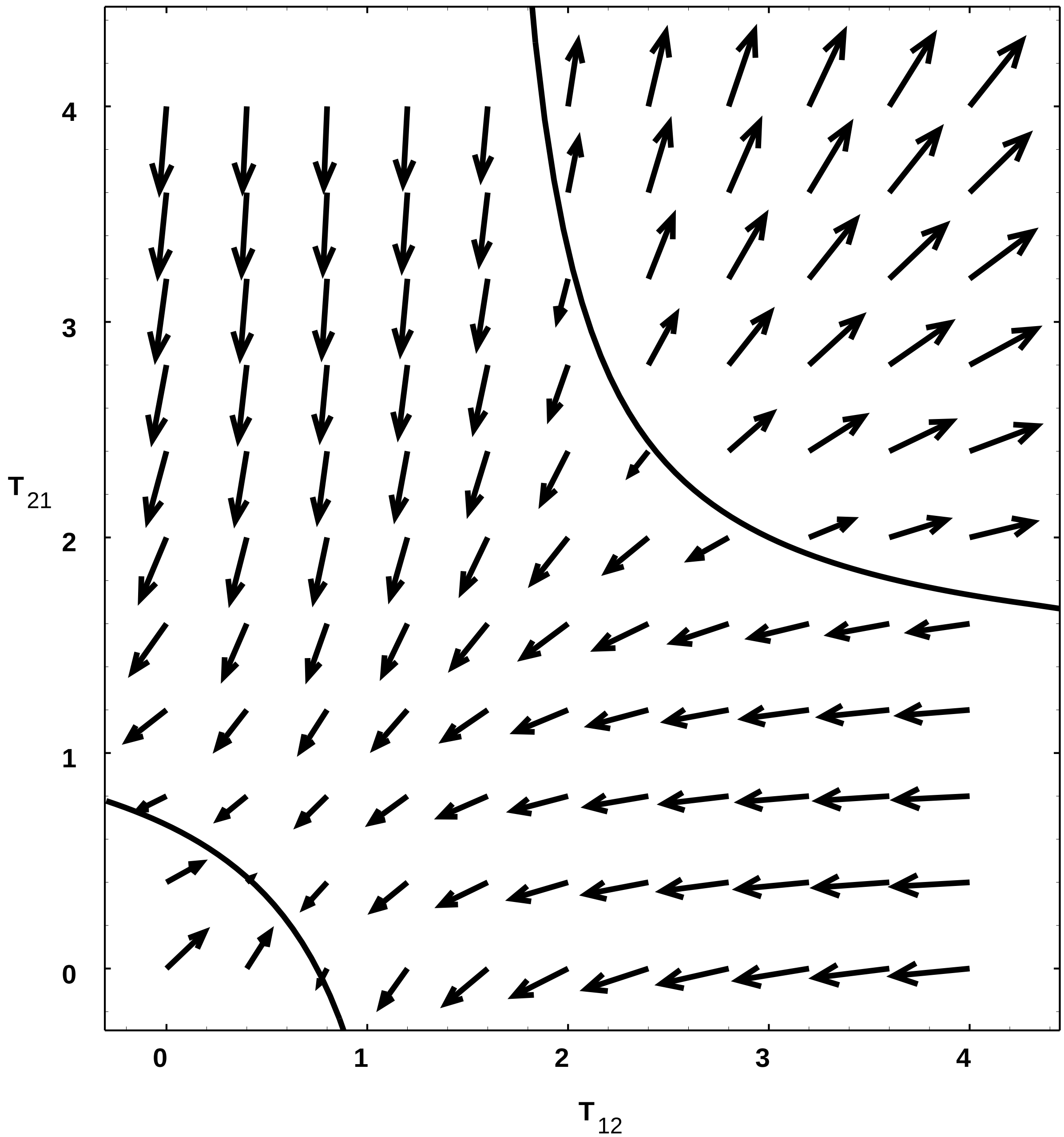
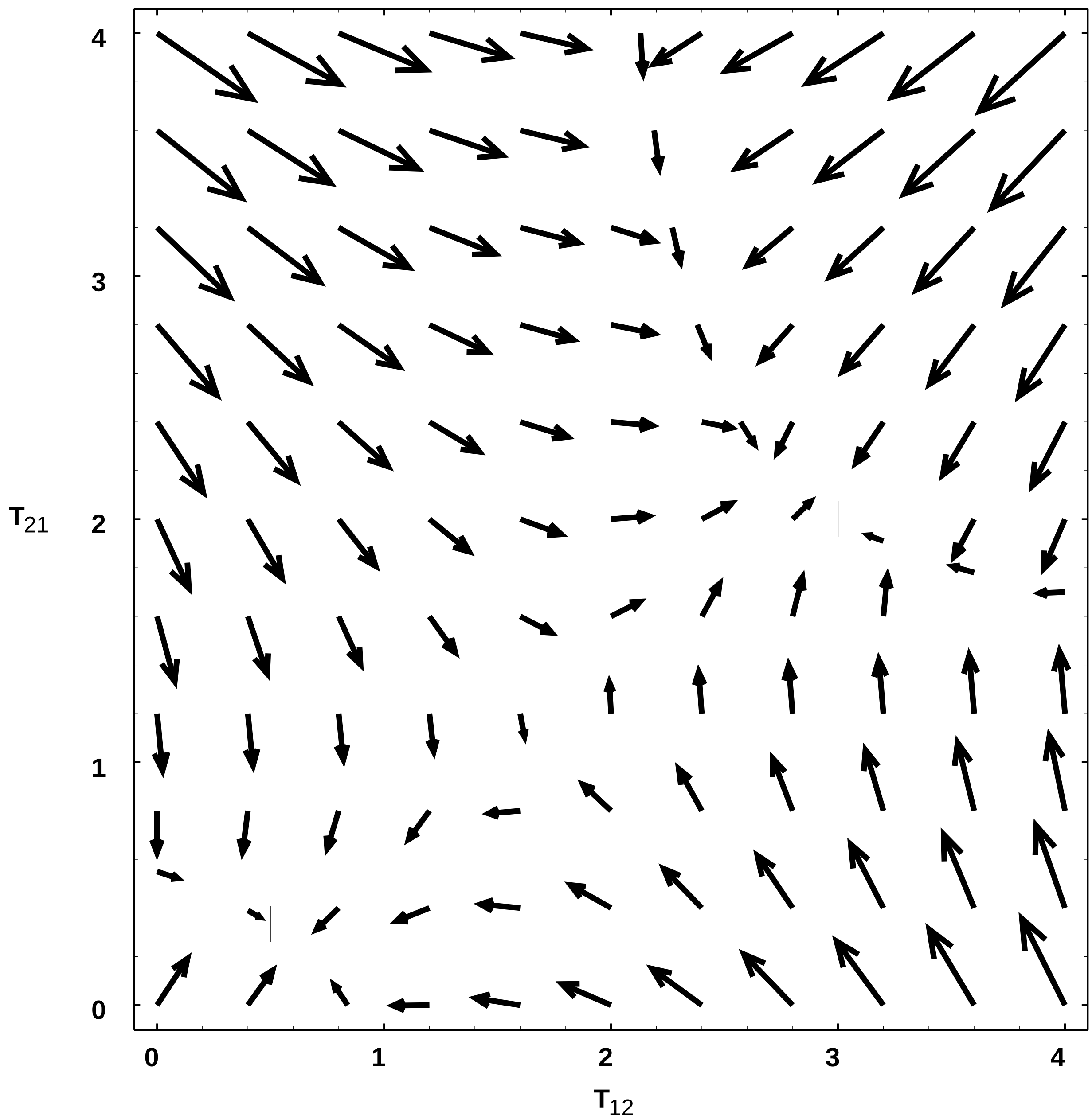


FIG. 1. A neural network which solves the source separation problem for two lineary mixed sources.



a)



b)

FIG. 2. Vector field plot for the two source separation problem of gaussian signals: a) for the Jutten-Herault model (eqn. 8) - the two lines correspond to stable and unstable fixed points. b) for using delay terms (eqn. 7) - only two fixed points exist which are stable.

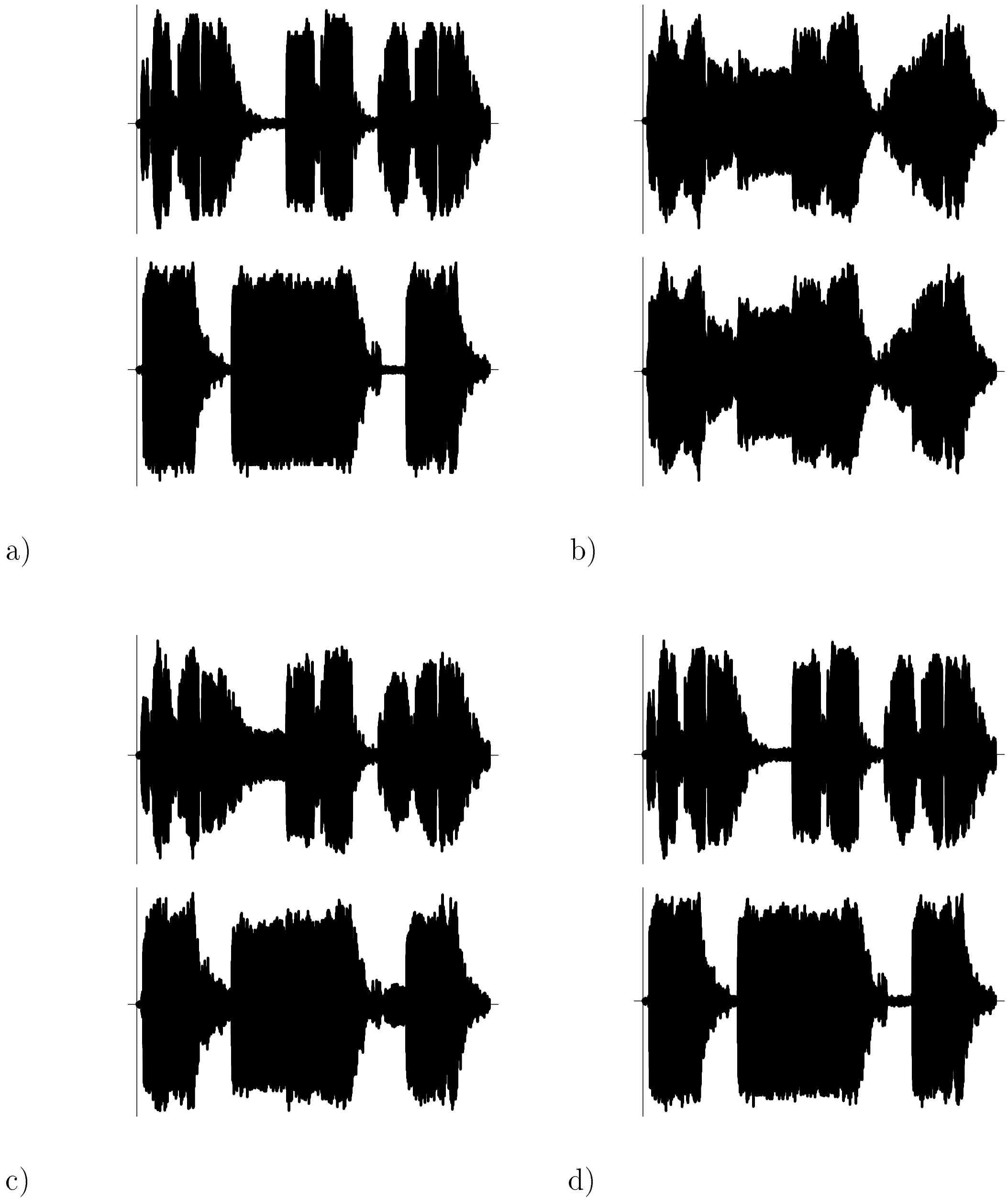


FIG. 3. Decorrelation of mixed signals: a) original speech signals produced by two independend crying babies, b) mixed signals with mixing matrix ((1,0.9),(0.7,1)), c) decorrelated signals using the least squares method [9], d) signals decorrelated using our method with delay parameter $\tau = 0.5$ms.

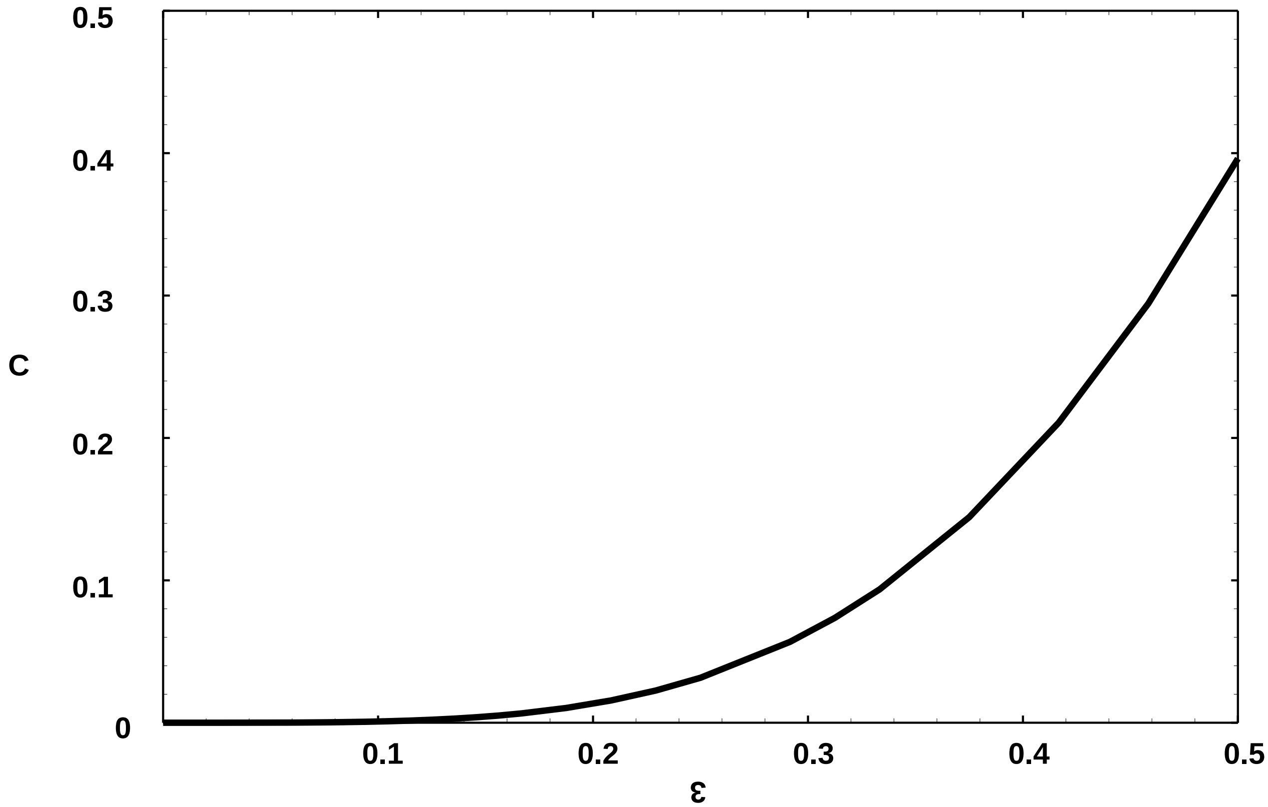


FIG. 4. The correlation $C = \langle u_1(t)u_2(t+2\tau)\rangle^2$ as a function of the nonlinearity parameter $\varepsilon = \varepsilon_1 = \varepsilon_2$ for eqn. 9. Only in the linear case ($\varepsilon = 0$) all time delayed correlations will vanish.