A THEORETIC MODEL FOR LINEAR GEOMETRIC ICA

Fabian J. Theis[1], *Andreas Jung*[2], *Elmar W. Lang*[1], *Carlos G. Puntonet*[3]

[1]Institute of Biophysics, University of Regensburg, 93040-Regensburg, Germany
[2]Institute of Theoretical Physics, University of Regensburg, 93040-Regensburg, Germany
[3]Dept. of Architecture and Computer Technology, University of Granda, Spain
Email: fabian.theis@mathematik.uni-regensburg.de

ABSTRACT

Geometric algorithms for linear ICA have recently received some attention due to their pictorial description and their relative ease of implementation. The geometric approach to ICA has been proposed first by Puntonet and Prieto [1] [2] in order to separate linear mixtures. We will reconsider geometric ICA in a solid theoretic framework showing that fixpoints of geometric ICA fulfill a so called geometric convergence condition, which the mixed images of the unit vectors satisfy, too. This leads to a conjecture claiming that in the supergaussian unimodal symmetric case there is only one stable fixpoint, thus demonstrating uniqueness of geometric ICA after convergence.

1. INTRODUCTION

In independent component analysis (ICA), given a random vector, the goal is to find its statistically independent components. This can be used to solve the blind source separation (BSS) problem which is, given only the mixtures of some underlying independent sources, to separate the mixed signals thus recovering the original sources. In contrast to correlation-based transformations such as principal component analysis, ICA renders the output signals as statistically independent as possible by evaluating higher-order statistics. The idea of ICA was first expressed by Jutten and Herault [3] [4] while the term ICA was later coined by Comon in [5]. However the field became popular only with the seminal paper by Bell and Sejnowski [6] who elaborated upon the Infomax-principle first advocated by Linsker [7] [8].

Recently, geometric ICA algorithms have received further attention due to their relative ease of implementation [1] [9]. They have been applied successfully to the analysis of real world biomedical data [10] [11] and have been extended to non-linear ICA problems [12] also.

2. BASICS

For $m, n \in \mathbb{N}$ let $\mathrm{Mat}(m \times n)$ be the $\mathbb{R}-$vectorspace of real $m \times n$ matrices, and $\mathrm{Gl}(n) := \{W \in \mathrm{Mat}(n \times n) \mid \det(W) \neq 0\}$ be the general linear group in $\mathbb{R}^n$.

In the quadratic case of linear blind source separation (BSS), a random vector $X : \Omega \to \mathbb{R}^n$ called **mixed vector** is given; it originates from an independent random vector $S : \Omega \to \mathbb{R}^n$, which will be denoted as **source vector**, by mixing with a **mixing matrix** $A \in \mathrm{Gl}(n)$, i.e. $X = A \circ S$. Here Ω is a fixed probability space. Only the mixed vector is known, and the task is to recover A and therewith $S = A^{-1} \circ X$.

In the nonlinear case, where A is any function $\mathbb{R}^n \to \mathbb{R}^n$, little is known because, without further restrictions, the problem is generally ill-posed. But in the linear case described above, many different algorithms have been proposed with the Bell-Sejnowski maximum entropy algorithm [6] being the most popular and also most widely studied among them.

In this paper we consider a geometric approach to the source separation problem. As we need a certain uniqueness of the solution, we want at most one of the source variables $S_i := \pi_i \circ S$, where $\pi_i : \mathbb{R}^n \to \mathbb{R}$ denotes the projection on the i-th coordinate, to be Gaussian. Then any **solution to the BSS problem**, i.e. any $B \in \mathrm{Gl}(n)$ such that $B \circ X$ is independent, is **equivalent** to A^{-1}, where equivalent means that B can be written as $B = LPA^{-1}$ with an invertible diagonal matrix (scaling matrix) $L \in \mathrm{Gl}(n)$ and an invertible matrix with unit vectors in each row (permutation matrix) $P \in \mathrm{Gl}(n)$. This has been proved by Comon [5]. Vice versa, any matrix B which is equivalent to A^{-1} solves the BSS problem, since we calculate for the transformed mutual information

$$I(B \circ X) = I(LPA^{-1} \circ X) = I(A^{-1} \circ X) = I(S) = 0,$$

taking into account that the information is invariant under scaling and permutation of coordinates.

3. GEOMETRIC CONSIDERATIONS

The basic idea of the geometric separation method lies in the fact that in the source space $\{s_1, \dots, s_\lambda\} \subset \mathbb{R}^n$, where s_i

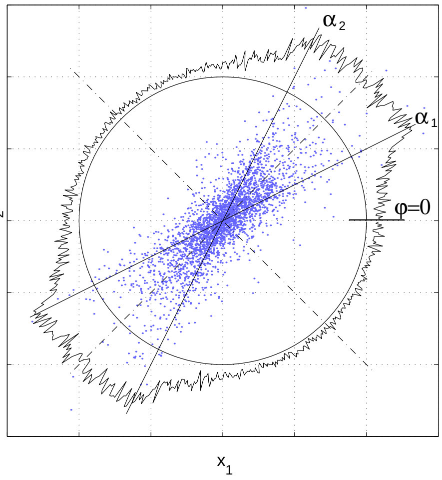

Fig. 1. Example of a two-dimensional scatter plot of a mixture of two Laplacian signals with identical variance. The signals have been mixed by a maxtrix A mapping the unit vectors onto vectors inclined with angle α_i to the x_1-axis. Striked lines show borders of the receptive fields.

represent a fixed number of samples of the source vector S with zero mean, the clusters along the axes of the coordinate system are transformed by A into clusters along different lines through the origin. The detection of those n new axes allows to determin a demixing matrix B which is equivalent to A^{-1}, see figure 1.

We now consider the learning process to be terminated already and describe precisely how to recover the matrix A then, i.e. after the axes, which span the observation space, have been extracted from the data successfully. Let

$$L := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \exists i\, x_i > 0, x_j = 0 \text{ for all } j \neq i\}$$

be the set of positive coordinate axes. Denote $L' := AL$ the image of this set under A.

We claim that L' intersects the unit $(n-1)$-sphere

$$S^{n-1} := \{x \in \mathbb{R}^n \mid |x| = 1\}$$

in exactly n distinct points $\{p_1, \dots, p_n\}$. For this note that $L' \cap S^{n-1}$ is the image of the unit vectors $\{e_1, \dots, e_n\}$ under the map

$$\begin{array}{ccccccc} f: & \mathbb{R}^n \setminus \{0\} & \longrightarrow & \mathbb{R}^n \setminus \{0\} & \longrightarrow & S^{n-1} \\ & x & \longmapsto & Ax & \longmapsto & \frac{Ax}{|Ax|} \end{array}$$

so we have (after a possible reordering of the p_i's) $f(e_i) = p_i$. Since A is bijective, $p_i = p_j$ induces $e_i = e_j$ and hence $i = j$. Furthermore, we note that the p_i's span the whole $\mathbb{R}^n$, so they form a basis of $\mathbb{R}^n$.

Define the matrix $M_{p_1,\dots,p_n} \in \mathrm{Gl}(n)$ to be the linear mapping of e_i onto p_i for $i = 1, \dots, n$, i.e.

$$M_{p_1,\dots,p_n} = (p_1 | \dots | p_n)\,.$$

This matrix thus effects the linear coordinate change from the standard coordinates $(e_i)_i$ to the new basis $(p_i)_i$. We then have the following lemma:

Lemma 3.1. *For a permutation $\sigma \in \mathcal{S}_n$, the two matrices $M_{p_1,\dots,p_n}$ and $M_{p_{\sigma(1)},\dots,p_{\sigma(n)}}$ are equivalent.*

Proof.

$$M_{p_1,\dots,p_n} = PM_{p_{\sigma(1)},\dots,p_{\sigma(n)}}$$

for a permutation matrix P. □

Theorem 3.2 (Uniqueness of the geometric method). *The matrix $M_{p_1,\dots,p_n}$ is equivalent to A.*

Proof. By construction of $M_{p_1,\dots,p_n}$, we have

$$M_{p_1,\dots,p_n}(e_i) = p_i = f(e_i) = \frac{Ae_i}{|Ae_i|},$$

so there exists a $\lambda_i \in \mathbb{R} \setminus \{0\}$ such that

$$M_{p_1,\dots,p_n}(e_i) = \lambda_i A e_i.$$

Setting

$$L := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

yields an invertible diagonal matrix $L \in \mathrm{Gl}(n)$, such that

$$M_{p_1,\dots,p_n} = LAe_i.$$

This shows the claim. □

Corollary 3.3. *The matrix $M^{-1}_{p_1,\dots,p_n}$ solves the BSS problem.*

4. THE GEOMETRIC ALGORITHM

We will restrict ourselves to the two-dimensional case; in [13], we show that for higher dimensions the number of samples needed to guarantee a fixed error rate grows exponentially, at least for uniform distributions. Therefore, in practice, higher-dimensional ICA solutions are sometimes constructed from their two-dimensional counterparts by projecting onto $\mathbb{R}^2$ along different coordinate axes and then reconstructing the multi-dimensional matrix from the different two-dimensional solutions [14] [10] [11]. However, this only works if the mixing matrix A is close to the identity up to scaling and permutation.

So, let $S : \Omega \longrightarrow \mathbb{R}^2$ be an independent two-dimensional Lebesgue-continuous random vector describing the source pattern distribution; its density function is denoted by ρ :

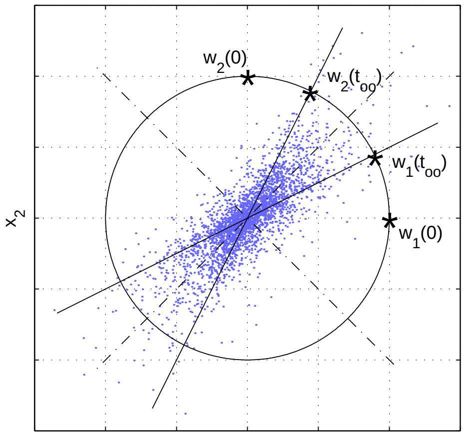

Fig. 2. Visualization of the geometric algorithm with starting points $w_1(0)$ and $w_2(0)$ and end points $w_1(\infty)$ and $w_2(\infty)$.

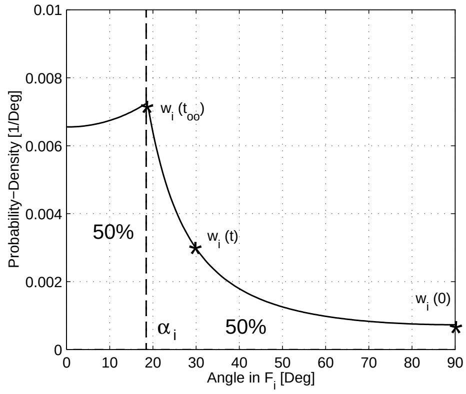

Fig. 3. Plot of the density ρ_Y of a mixture of two Laplacian signals with identical variance. The weight adaptation by the geometric algorithm is also visualized, also see figure 2.

$\mathbb{R}^2 \longrightarrow \mathbb{R}$. As S is independent, ρ factorizes in the following way

$$\rho(x, y) = \rho_1(x)\rho_2(y),$$

with the marginal source density functions $\rho_i : \mathbb{R} \longrightarrow \mathbb{R}$. We assume the source variables S_i tobe distributed symmetrically, i.e. $\rho_i(x) = \rho_i(-x)$ for $x \in \mathbb{R}$ and $i = 1, \ldots, n$. In particular, $E(S) = 0$. For stability of the geometric algorithm, later we will have to assume that the signals are supergaussian and unimodal – in practice these restrictions are often met at least approximately.

As above, let X denote the mixed vector and A the invertible mixing matrix such that $X = AS$. Without loss of generality, assume that A is of the form

$$A = \begin{pmatrix} \cos\alpha_1 & \cos\alpha_2 \\ \sin\alpha_1 & \sin\alpha_2 \end{pmatrix},$$

where $\alpha_i \in [0, \pi)$ are two angles. The **geometric learning algorithm** for symmetric distributions in its simplest form then goes as follows:

Pick four starting neurons w_1, w_1', w_2 and w_2' on S^1 such that w_i and w_i' are opposite each other, i.e. $w_i = -w_i'$ for $i = 1, 2$, and w_1 and w_2 are linearly independent vectors in R^2. Usually, one takes the unit vectors $w_1 = e_1$ and $w_2 = e_2$. Furthermore fix a learning rate $\eta : \mathbb{N} \longrightarrow \mathbb{R}$ such that $\eta(t) > 0$, $\sum_{n\in\mathbb{N}} \eta(n) = \infty$ and $\sum_{n\in\mathbb{N}} \eta(n)^2 < \infty$. Then iterate the following step until an appropriate abort condition has been met:

Choose a sample $x(t) \in \mathbb{R}^2$ according to the distribution of X. If $x(t) = 0$ pick a new one – note that this case happens with probability zero since the probability density function (pdf) ρ_X of X is assumed to be continuous. Project $x(t)$ onto the unit sphere and get $y(t) := \frac{x(t)}{|x(t)|}$. Let i be in $\{1, 2\}$ such that w_i or w_i' is the neuron closest to y with respect to an Euclidean metric. Then set

$$w_i(t + 1) := \pi(w_i(t) + \eta(t)\,\mathrm{sgn}(y(t) - w_i(t))),$$

where $\pi : \mathbb{R}^2 \setminus \{0\} \longrightarrow S^1$ is the projection, and

$$w_i'(t + 1) := -w_i(t + 1).$$

The other two neurons are not moved in this iteration.

In figures 2 and 3 the learning algorithm has been visualized both on the sphere and after the projection onto $[0, \pi)$.

This algorithm may be called **absolute winner-takes-all learning**. It is Kohonen's learning algorithm for self-organizing maps with a trivial neighbourhood function (0-**neighbour algorithm**) but with the modification that the step size along the direction of a sample does not depend on distance, and that the learning process takes place on S^1 not in $\mathbb{R}$.

5. FORMAL MODEL OF THE GEOMETRIC ALGORITHM

Now, we present a formal theoretical framework for geometric ICA which will be used in the next section to formulate a proper convergence condition.

First, we show using the symmetry of S that it is in fact not necessary to have two neurons w_i and w_i' moving around on the same axis. Indeed, we should not speak of neurons but of lines in $\mathbb{R}^2$ – so our neurons would be living in the real projective space $\mathbb{RP}^1 = S^1/\sim$, where $\sim$ identifies antipodal points. This is the manifold of all 1-dimensional subvectorspaces of $\mathbb{R}^2$. A metric is defined by setting

$$d([x], [y]) := \min\{|x - y|, |x + y|\}$$

for $[x], [y] \in \mathbb{RP}^1$. Alternatively, one can picture the neurons in $S^1_+ := S^1 \cap \{(x_1, x_2) \in \mathbb{R}^2 | x_2 \geq 0\}/\sim$, where $\sim$ identifies the two points $(1, 0)$ and $(-1, 0)$. Let $\eta : S^1 \longrightarrow S^1_+$ be the canonical projection. Furthermore, it is useful to introduce polar coordinates $\varphi : S^1_+ \longrightarrow [0, \pi)$ on S^1_+ with the stratification $\varphi' : \mathbb{R} \longrightarrow S^1_+$ such that $\varphi' \circ \varphi = \mathrm{id}$. Let $\chi := \varphi \circ \varphi' : \mathbb{R} \longrightarrow [0, \pi)$ be the 'modulo π' map. We want to somehow approximate the transformed random variable

$$Y := \varphi \circ \eta \circ \pi \circ X : \Omega \longrightarrow [0, \pi).$$

Note, that the density function ρ_Y of Y can be calculated from the density ρ_X of X by

$$\begin{aligned}
\rho_Y(\varphi) &= \int_{-\infty}^{\infty} \rho_X(r \cos\varphi, r \sin\varphi) r dr \\
&= |\det A|^{-1} \int_{-\infty}^{\infty} \rho(A^{-1}(r \cos\varphi, r \sin\varphi)^\top) r dr \\
&= 2|\det A|^{-1} \int_{0}^{\infty} \rho(A^{-1}(r \cos\varphi, r \sin\varphi)^\top) r dr
\end{aligned}$$

for $\varphi \in [0, \pi)$ using the symmetry of ρ. Then the geometric learning algorithm induces the following discrete Markov process $W(t) : \Omega \longrightarrow \mathbb{R}^2$ defined recursively by

$$W(0) = (w_1, w_2)$$

and

$$W(t + 1) = \chi(W(t) + \eta(t)\theta(Y(t) - W(t))),$$

where

$$\theta(x, y) := \begin{cases} (\mathrm{sgn}(x), 0) & |y| \geq |x| \\ (0, \mathrm{sgn}(y)) & |x| > |y| \end{cases}$$

and $Y(0), Y(1), \ldots$ is a sequence of independent identically distributed random variables $\Omega \longrightarrow \mathbb{R}^2$ with the same distribution as Y — we need them in order to represent the independence of the successive sampling experiments. Note that the right hand side of the algorithm has been plugged into χ in order to guarantee that $W(t + 1) \in [0, \pi)$. Indeed, this is just winner-takes-all learning with a signum function in $\mathbb{R}$, but taking into account the fact that we have to stay in $[0, \pi)$. Note that the metric used here now is the planar metric, which is obviously equivalent to the metric on S^1_+ induced by the Euclidean metric on $S^1 \subset \mathbb{R}^2$.

We furthermore can assume that after enough iterations there is one point $a \in S^1$ that will not be transversed any more, and without loss of generality, we assume a to be 0 (otherwise cut S^1 open at a and project along this resulting arc), so that the above algorithm simplifies to the planar case with the recursion rule

$$W(t + 1) = W(t) + \eta(t)\theta(Y(t) - W(t)).$$

This is exactly Kohonen's learning rule in the 0-neighbour case with an additional sign function. Without the sign function, and the additional fact that the probability distribution of Y is **log-concave**, it has been shown [15] [16] [17] that the process $W(t)$ converges to a unique constant fixpoint process $W \equiv w \in \mathbb{R}^2$ such that

$$\frac{1}{\lambda(F_i)} \int_{F_i} \varphi \rho_Y(\varphi) d\varphi = w_i \tag{1}$$

for $i = 1, 2$, where

$$\begin{aligned}
&F_i := F(w_i) := \\
&\{\varphi \in [0, \pi) \mid \chi(|\varphi - w_i|) \leq \chi(|\varphi - w_j|) \text{ for all } j \neq i\}
\end{aligned}$$

is the **receptive field** of the neuron w_i and $\lambda(F_i)$ is the volume of the field. However, it is not clear how to generalize the proof for our geometric case, especially because we do not have (and also do not want) log-concavity of Y because this would lead to a unique fixpoint. Therefore we will assume convergence in a sense stated in the following section. Refer to figure 1 for a picture of the distribution on the sphere together with the corresponding receptive fields (dotted lines).

6. LIMIT POINTS OF THE GEOMETRIC ALGORITHM

In this section, we want to study the end points of geometric ICA, so we will assume that the algorithm has already converged. The idea then is to formulate a condition which the end points will have to satisfy and to show that the solutions are among them.

Definition 6.1 (Geometric Convergence Condition). *Two angles $l_1, l_2 \in [0, \pi)$ satisfy the **Geometric Convergence Condition (GCC)** if they are the medians of Y restricted to their receptive fields i.e. if l_i is the median of $\rho_Y|F(l_i)$.*

Definition 6.2. *A constant random vector $\hat{W} \equiv (\hat{w}_1, \hat{w}_2) \in \mathbb{R}^2$ is called **fixpoint of geometric ICA in the expectation** if*

$$E(\theta(Y - \hat{W}(t))) = 0.$$

Hence, the expectation of a Markov process $W(t)$ starting at a fixpoint of geometric ICA will not be changed by the geometric update rule indeed, because

$$\begin{aligned}
E(W(t + 1)) &= E(W(t)) + \eta(t)E(\theta(Y(t) - W(t))) \\
&= E(W(t)).
\end{aligned}$$

Theorem 6.3. *If the geometric algorithm converges to a fixpoint $W(\infty) \equiv (w_1(\infty), w_2(\infty))$ of geometric ICA in the expectation, then the $w_i(\infty)$ satisfy GCC.*

Proof. Assume $w_1(\infty)$ does not satisfy GCC. Without loss of generality, let $[\beta_1, \beta_2]$ be the receptive field of $w_1(\infty)$ such that $\beta_i \in [0, \pi)$. Since $W(\infty)$ is a fixpoint of geometric ICA in the expectation, we have

$$E\left(\chi_{[\beta_1,\beta_2]}(Y(t))\,\mathrm{sgn}(Y(t) - w_1(\infty))\right) = 0,$$

where $\chi_{[\beta_1,\beta_2]}$ denotes the characteristic function of that interval. But this means

$$\int_{\beta_1}^{w_1(\infty)} (-1)\rho_Y(\varphi)d\varphi + \int_{w_1(\infty)}^{\beta_2} 1\rho_Y(\varphi)d\varphi = 0$$

and therefore

$$\int_{\beta_1}^{w_1(\infty)} \rho_Y(\varphi)d\varphi = \int_{w_1(\infty)}^{\beta_2} \rho_Y(\varphi)d\varphi,$$

so $w_1(\infty)$ satisfies GCC. The same calculation for $w_2(\infty)$ shows the theorem. □

As before, let $p_i := Ae_i$ be the transformed unit vectors, and let $q_i := \varphi \circ \eta \circ \pi(p_i) \in [0, \pi)$ be the corresponding angles for $i = 1, 2$.

Theorem 6.4. *The transformed angles q_i satisfy GCC.*

Proof. Because of the symmetry of the claim it is enough to show that q_1 satisfies GCC. Without loss of generality let $0 < \alpha_1 < \alpha_2 < \pi$ using the symmetry of ρ. Then, due to construction $q_i = \alpha_i$. Let $\beta_1 := \frac{\alpha_1+\alpha_2}{2} - \frac{\pi}{2}$ and $\beta_2 := \beta_1 + \frac{\pi}{2}$. Then the receptive field of q_1 can be written (modulo π) as $F(q_1) = [\beta_1, \beta_2]$. Therefore, we have to show that $q_1 = \alpha_1$ is the median of ρ_Y restricted to $F(q_1)$, which means $\int_{\beta_1}^{\alpha_1} \rho_Y(\varphi)d\varphi = \int_{\alpha_1}^{\beta_2} \rho_Y(\varphi)d\varphi$.

We will reduce this to the orthogonal standard case $A = \mathrm{id}$ by transforming the integral as follows:

$$\begin{aligned}
&\textstyle\int_{\beta_1}^{\alpha_1} \rho_Y(\varphi)d\varphi = \\
&= 2|\det A|^{-1} \textstyle\int_{\beta_1}^{\alpha_1} d\varphi \int_0^\infty r dr \rho(A^{-1}(r\cos\varphi, r\sin\varphi)^\top) \\
&= 2|\det A|^{-1} \textstyle\int_K dxdy \rho(A^{-1}(x,y)^\top)
\end{aligned}$$

where $K := \{(x,y) \in \mathbb{R}^2 | \beta_1 \le \arctan(y/x) \le \alpha_1\}$ denotes the cone of opening angle $\alpha_1 - \beta_1$ starting from angle β_1. Using the transformation formula, we continue $\int_{\beta_1}^{\alpha_1} \rho_Y(\varphi)d\varphi = 2\int_{A^{-1}(K)} dxdy\rho(x,y)$. Now note that the transformed cone $A^{-1}(K)$ is a cone ending at the x-axis of opening angle $\pi/4$, because A is linear; therefore we are left with the following integral:

$$\begin{aligned}
\int_{\beta_1}^{\alpha_1} \rho_Y(\varphi)d\varphi &= 2\int_0^\infty dx \int_{-x}^0 dy \rho(x,y) \\
&= 2\int_0^\infty dx \int_0^x dy \rho(x,-y) \\
&= 2\int_0^\infty dx \int_0^x dy \rho(x,y) \\
&= \int_{\alpha_1}^{\beta_2} \rho_Y(\varphi)d\varphi
\end{aligned}$$

where we have used the same calculation for $[\alpha_1, \beta_2]$ as for $[\beta_1, \alpha_1]$ at the last step. This completes the proof of the theorem. □

Combining both theorems, we have therefore shown:

Theorem 6.5. *Let Φ be the set of fixpoints of geometric ICA in the expectation. Then there exists $(\hat{w}_1, \hat{w}_2) \in \Phi$ such that $M^{-1}_{\hat{w}_1,\hat{w}_2}$ solves the BSS problem. The stable fixpoints in Φ can be found by the geometric ICA algorithm.*

Furthermore, we believe that in the special case of unimodal and supergaussian signals, the set Φ consists of only two elements: a stable and an unstable fixpoint, where the stable fixpoint will be found by the algorithm:

Conjecture 6.6. *Assume that the sources S_i are unimodal and symmetric. Then there are only two fixpoints of geometric ICA in the expectation.*

Note that then, the two fixpoints are related as follows: If $(\hat{w}_1, \hat{w}_2)$ is a fixpoint in Φ, then $\left(\frac{\hat{w}_1+\hat{w}_2}{2}, \frac{\hat{w}_1-\hat{w}_2}{2}\right)$ is the other fixpoint.

Conjecture 6.7. *Assume that the sources S_i are unimodal, symmetric and supergaussian. Then there is only one stable fixpoint $(\hat{w}_1, \hat{w}_2)$ of geometric ICA in the expectation, and $M^{-1}_{\hat{w}_1,\hat{w}_2}$ solves the BSS problem.*

If the two conjectures have been shown, using the above remark we can also perform geometric ICA for subgaussian signals with the standard algorithm. Then

$$M^{-1}_{\frac{\hat{w}_1+\hat{w}_2}{2}, \frac{\hat{w}_1-\hat{w}_2}{2}}$$

is the solution of the BSS problem for the subgaussian case. Simulations for mixed uniform distributions confirm this result.

7. UPDATE RULES WITHOUT SIGN FUNCTIONS

We have shown that the geometric update step requires the signum function as follows

$$w_i(t+1) = w_i(t) + \eta(t)\,\mathrm{sgn}(y(t) - w_i(t)).$$

Then (normally) the w_i converge to the medians in their receptive field. Note that the medians don't have to coincide with any maxima of the mixed density distribution on the sphere as shown in figure 4. Therefore, in general, any algorithm searching for the maxima of the distribution will not find the medians, which are the images of the unit vectors under the mixture. However under special restrictions to the sources (same super-gaussian distribution of each component, as for example speech signals), the medians correspond to the maxima [18].

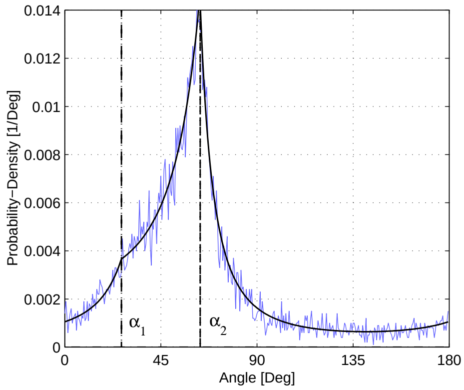

Fig. 4. Projected density distribution ρ_Y of a mixture of two Laplacian signals with different variances, with the mixture matrix mapping the unit vectors e_i to $(\cos\alpha_i, \sin\alpha_i)$ for $i = 1, 2$. (dark line = theoritical density function, gray line = histogram of a mixture of 10.000 samples)

8. CONCLUSION

The geometric ICA algorithm has been studied in a concise theoretical framework resembling the one of Kohonen's learning algorithm. The fixpoints of geometric ICA learning algorithm have been examined in detail. We have introduced a Geometric Convergence Condition, which has to be fullfilled by the fixpoints of the learning algoritm. We further showed it is also fulfilled by the mixed unit vectors. Hence geometric ICA can solve the BSS problem. Finally, we have given two conjectures for the unimodal case where the fixpoint property is expected to be very rigid.

In future work, besides treating non symmetric sources S, these theoretical concepts have to be extended to understand nonlinear geometric algorithms [11]. And of course, the two conjectures will have to be proven, as well as the Kohonen proof of convergence to be translated into the above model.

9. ACKNOWLEDGEMENTS

We want to thank Michaela Lautenschlager and Stefanie Ulsamer for their helpful comments when trying to prove convergence.

10. REFERENCES

[1] C. G. Puntonet and A. Prieto, "An adaptive geometrical procedure for blind separation of sources," *Neural Processing Letters*, vol. 2, 1995.

[2] C.G.Puntonet, A.Prieto, C.Jutten, M.R.Alvarez, and J.Ortega, "Separation of sources: a geometry-based procedure for reconstruction of n-valued signals," *Signal Processing*, vol. 46, pp. 267–284, 1995.

[3] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1–10, 1991.

[4] C.Jutten, J.Herault, P.Comon, and E.Sorouchiary, "Blind separation of sources, parts i, ii and iii," *Signal Processing*, vol. 24, pp. 1–29, 1991.

[5] P. Comon, "Independent component analysis - a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[6] A. J. Bell and T. J. Sejnowski, "An information-maximisation approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.

[7] R. Linsker, "An application of the principle of maximum information preservation to linear systems," *Advances in Neural Information Processing Systems*, vol. 1, 1989.

[8] R. Linsker, "Local synaptic learning rules suffice to maximize mutual information in a linear network," *Neural Computation*, vol. 4, pp. 691–702, 1992.

[9] C. G. Puntonet and A. Prieto, "Neural net approach for blind separation of sources based on geometric properties," *Neurocomputing*, vol. 18, pp. 141–164, 1998.

[10] Ch. Bauer, M. Habl, E.W. Lang, C.G. Puntonet, and M.R. Alvarez, "Probabilistic and geometric ICA applied to the separation of EEG signals," *M.H.Hamza, ed., Signal Processing and Communication (Proc.SPC'2000), IASTED/ACTA Press, Anaheim, USA*, pp. 339 – 346, 2000.

[11] C.G. Puntonet, Ch. Bauer, E.W. Lang, M.R. Alvarez, and B. Prieto, "Adaptive-geometric methods: application to the separation of EEG signals," *P.Pajunen, J.Karhunen, eds., Independent Component Analysis and Blind Signal Separation (Proc. ICA'2000)*, pp. 273–278, 2000.

[12] C.G.Puntonet, M.R.Alvarez, A.Prieto, and B.Prieto, "Separation of speech signals for nonlinear mixtures," *Lecture Notes in Computer Science*, vol. 1607, pp. 665–673, 1999.

[13] Andreas Jung, Fabian J. Theis, Carlos G. Puntonet, and Elmar W. Lang, "Fastgeo - a histogram based approach to linear geometric ica," *preprint*, 2001.

[14] Ch. Bauer, C.G. Puntonet, M. Rodriguez-Alvarez, and E.W.Lang, "Separation of EEG signals with geometric procedures," *C. Fyfe, ed., Engineering of Intelligent Systems (Proc. EIS'2000)*, pp. 104–108, 2000.

[15] Marie Cottrell and Jean-Claude Fort, "Étude d'un processus d'auto-organisation," *Annales de l'Institut Henri Poincaré*, vol. 23, no. 1, pp. 1–20, 1987, (in French).

[16] H. Ritter and K. Schulten, "Convergence porperties of Kohonen's topology conserving maps: Fluctuations, stability, and dimension selection," *Biological Cybernetics*, vol. 60, pp. 59–71, 1988.

[17] M. Benaim, J.-C. Fort, and G. Pagés, "Convergence of the one-dimensional Kohonen algorithm," *Adv. Appl. Prob.*, vol. 30, pp. 850–869, 1998.

[18] A. Prieto, B. Prieto, C.G. Puntonet, A. Canas, and P. Martin-Smith, "Geometric separation of linear mixtures of sources: Application to speech signals," *J.F.Cardoso, Ch.Jutten, Ph.Loubaton, eds., Independent Component Analysis and Signal Separation (Proc. ICA'99)*, pp. 295–300, 1999.