On the performance of orthogonal source separation algorithms

Jean-François Cardoso

École Nationale Supérieure des Télécommunications
Télécom Paris, Département Signal, 46 rue Barrault, 75634 Paris Cedex 13, France.
email: cardoso@sig.enst.fr, *fax: (33) 1 45 88 79 35.*

Abstract. *Source separation* consists in recovering a set of n independent signals from $m \geq n$ observed instantaneous mixtures of these signals, possibly corrupted by additive noise. Many source separation algorithms use second order information in a whitening operation which reduces the non trivial part of the separation to determining a unitary matrix. Most of them further show a kind of invariance property which can be exploited to predict some general results about their performance. Our first contribution is to exhibit a lower bound to the performance in terms of accuracy of the separation. This bound is independent of the algorithm and, in the i.i.d. case, of the distribution of the source signals. Second, we show that the performance of invariant algorithms depends on the mixing matrix and on the noise level in a specific way. A consequence is that at low noise levels, the performance does not depend on the mixture but only on the distribution of the sources, via a function which is characteristic of the given source separation algorithm.

1. Introduction.

This paper is concerned with the *source separation problem* which consists in recovering a set of n independent signals from $m \geq n$ observed instantaneous mixtures of these signals. Denoting $\mathbf{x}(t)$ the $m \times 1$ vector of observations (sensor outputs) at time t, possibly corrupted by additive noise $\mathbf{n}(t)$, the model is

$$\mathbf{x}(t) = A\mathrm{s}(t) + \mathbf{n}(t) = \mathbf{y}(t) + \mathbf{n}(t) \qquad (1)$$

where the $m \times n$ matrix A is called the 'mixing matrix' and where the n independent signals are collected in a $n \times 1$ vector denoted $\mathrm{s}(t)$. All the processes are assumed to be zero-mean stationary. The aim of source separation is to find a separating matrix, i.e. a $n \times m$ matrix such that $\widehat{\mathrm{s}}(t) = B\mathbf{x}(t)$ is an estimate of the source signals.

In the complex case, model (1) is the familiar linear model used in narrow band array processing. In this context, it is usually assumed that the columns of A depend on very few location parameters (such as DOAs) and this dependence is assumed to be known via the 'array manifold'. In contrast, we address here a problem of *blind array processing* in the sense that matrix A is completely unstructured: it is supposed to be a fixed full column rank matrix but no other assumptions are made.

The 'blind' approach is strongly motivated when **i)** one is interested in recovering the source signals (like in communication applications) but not in locating the emitting sources and **ii)** whenever the array manifold is unavailable or is expected to significantly depart from its model. Source separation is calibration-free and, by essence, insensitive to modelling errors. Blind source separation relies essentially on the assumption that the signals to be separated are mutually independent; this is a strong but often plausible assumption, which may be exploited using either the non-normality (if any) or the spectral differences (if any) of the source signals.

Several block-oriented source separation algorithms have been proposed in [1-7] which are based on a two step approach: whitening and rotating the observations. This paper is concerned about the performance of such algorithms. The idea here is to determine the general performance of this approach without specifying a particular algorithm. Although this paper addresses a statistical issue, emphasis is not statistical rigor but rather on exhibiting the algebraic mechanism by which the prewhitening mechaism affects the performance.

2. Performance and blind identication

2.1. Blind identification

A source separation algorithm also is a blind identification technique and may be represented as

$$\widehat{A} = \mathcal{A}(X_T) \qquad (2)$$

where an estimate $\widehat{A}$ is computed from a block of T samples by some algorithm which is represented here as a function $\mathcal{A}$ of the $m \times T$ matrix[1] X_T:

$$X_T = [\mathbf{x}(1), \dots, \mathbf{x}(T)]. \qquad (3)$$

Before going further, it is important to notice that the function $\mathcal{A}$ may not be perfectly defined because of some indeterminations of the blind identification problem. This is because the exchange of a complex factor between each source signal and the corresponding column of A leaves the observations unchanged. It follows that without any loss of generality, the source signals can be assumed to have unit variance. With this normalization *convention* (we insist that this is not an assumption), the covariance matrix of $\mathrm{s}(t)$ is the identity matrix:

$$R_s = \mathrm{E}\mathrm{s}(t)\mathrm{s}(t)^* = I_n. \qquad (4)$$

[1] We use the same notational convention for S_T, N_T, etc...

This convention still leaves undetermined the phase of each column of A as well as their ordering since the 'labelling' of the source signals is immaterial. In the following, two estimates $\widehat{A}$ and $\widehat{A}'$ of A are considered as equivalent if

$$\widehat{A}' = \widehat{A}J \tag{5}$$

when J is any $n \times n$ matrix with only one non-zero unit-norm entry in each row and each column. We call such a matrix a 'quasi identity'; it is a unitary matrix. A source separation algorithm (hence a function $\mathcal{A}$) is considered well defined if it is defined up to right multiplication by some quasi-identity.

2.2. Rejection rates

We shall characterize the quality of an estimate $\widehat{A}$ of A by the closeness of matrix $\widehat{A}^{\#}A$ to some quasi-identity matrix where supscript $\#$ denotes pseudoinversion. Indeed if $\widehat{A}^{\#}$ is used to estimate the source signals from the observations, then

$$\widehat{\mathbf{s}}(t) = \widehat{A}^{\#}X(t) = (\widehat{A}^{\#}A)\mathbf{s}(t) + \widehat{A}^{\#}\mathbf{n}(t) \tag{6}$$

so that the variance of the q-th signal at the p-th output of the separator is given by

$$P_{pq} \stackrel{\text{def}}{=} |(\widehat{A}^{\#}A)_{pq}|^2 \tag{7}$$

since our convention is that each source has unit variance. Hence, examination of the distribution P_{pq} provides a very intuitive measure of the performance of the identification.

Note that changing $\widehat{A}$ into $\widehat{A}J$ for J some quasi-identity matrix amounts to changing P_{pq} into $P_{p\sigma(q)}$ where σ is some permutation of $\{1, \ldots, n\}$. We assume in the following that the permutation has been removed (for instance on the basis of some additional *a priori* information) so that P_{pq} actually is the variance of the q-th signal in the estimate of the p-th signal. The phase indetermination is invisible in P_{pq}. Also note that when $\widehat{A}$ is close enough to A, them P_{pp} is close to 1: the quantities P_{pq} are readily normalized and can be directly read as 'rejection rates', 'interfer-to-signal ratio', etc...

2.3. Invariance

Assume for a while that $n = m$ and that no noise is present. Some source separation algorithms yields the estimate of A as the solution of

$$\frac{1}{T}\sum_{t=1}^{T} G(\widehat{A}^{-1}\mathbf{x}(t)) = 0 \tag{8}$$

where G is some vector-to-matrix mapping. A slight modification of [8] falls in this class and [9] is an adaptive solver of such an equation. The estimator $\mathcal{A}$ associated to (8) is such that,

$$\mathcal{A}(CX_T) = C\mathcal{A}(X_T)\ J \tag{9}$$

for any invertible matrix C and some quasi-identity matrix J. This property is called here *full invariance* (we consider later a restricted 'orthogonal invariance'). For a fully invariant estimator:

$$\widehat{A}^{\#}A = \mathcal{A}(X_T)^{-1}A = \mathcal{A}(AS_T)^{-1}A = J\ \mathcal{A}(S_T)^{-1} \tag{10}$$

for some J so that, up to a permutation, the distribution of P_{pq} depends only on S_T in the noiseless case. We conclude that in absence ofnoise, *the performance of a fully invariant estimator does not depend on mixing matrix A.*

This very brief discussion on invariant algorithms is intended to introduce the next two sections: these ideas can be, to some extent, generalized to source separation algorithms based on prewhitening even though they are not fully invariant and are supposed to operate in noisy situations.

3. Invariant orthogonal estimators

A few preliminary definitions are needed. The *signal subspace* is the range of A and the *noise subspace* is the orthogonal subspace. The orthogonal projectors onto these subspaces, respectively denoted as Π and $\Pi^{\perp}$, are given by

$$\Pi = AA^{\#} = A(A^HA)^{-1}A^H, \qquad \Pi^{\perp} = I_m - \Pi \tag{11}$$

We also define the following covariance matrices

$$R_x = \mathrm{E}\mathbf{x}(t)\mathbf{x}(t)^*, \quad R_y = \mathrm{E}\mathbf{y}(t)\mathbf{y}(t)^*, \quad R_n = \mathrm{E}\mathbf{n}(t)\mathbf{n}(t)^*,$$

which, by independence of signal and noise, are related by

$$R_x = R_y + R_n = AR_sA^H + R_n = AA^H + R_n.$$

3.1. A two step procedure

Source separation algorithms based on prewhitening compute estimates of A in the form

$$\widehat{A}^{\#} = \widehat{U}^H\widehat{W} \tag{12}$$

where $\widehat{W}$ is a $n \times m$ matrix called the 'whitening matrix' and is obtained from an estimate of R_y:

$$\widehat{W} = \mathcal{W}(\widehat{R}_y) \tag{13}$$

and $\widehat{U}$ is a $n \times n$ unitary matrix computed from the whitened data:

$$\widehat{U} = \mathcal{U}(\widehat{W}X_T). \tag{14}$$

More specifically, $\widehat{W}$ is any $n \times m$ matrix verifying

$$\widehat{W}\widehat{R}_y\widehat{W}^H = I_n, \qquad \widehat{W}\widehat{\Pi}^{\perp} = 0 \tag{15}$$

where $\widehat{\Pi}^{\perp} = I_m - \widehat{\Pi}$ and $\widehat{\Pi}$ is the orthogonal projector on the range of $\widehat{R}_y$.

The idea here is recover the source signals by whitening (second-order decorrelation) and then rotating thenm to further satisfy a stronger independence criterion.

Our approach here is *not* to specify the particular algorithm $\mathcal{U}$ used to find the 'missing rotation' because we are interested in properties which are *shared* by all algorithms based on prewhitening. However, we assume that $\mathcal{U}$ satisfies an *orthogonal invariance* property:

$$\forall V \text{ unitary}, \ \ \mathcal{U}(VZ_T) = V\mathcal{U}(Z_T)\ J \tag{16}$$

for almost any realization of Z_T and for some (irrelevant) quasi-identity matrix J. Such a property arises naturally in our context. For instance the maximum contrast estimation [4] of U according to

$$\widehat{U} = \underset{U \text{ unitary}}{\mathrm{Argmax}} \sum_{i=1,n} |\widehat{\mathrm{Cum}}^{(4)}(U^H\mathbf{z}(t))_i|^2 \tag{17}$$

is easily seen to be an invariant estimator. In fact, any reasonable estimator of U obtained by optimizing under unitary constraint a functional of the (empirical) distribution of $U^H\mathbf{x}$ enjoys the orthogonal invariance property.

3.2. The noiseless case

As a first step in investigating performance, we show that in the limit where the noise can be neglected, invariant orthogonal estimators become fully invariant in the sense of equation (9).

Note that if no noise is present, the signal subspace can be determined exactly (as the range of $\widehat{R}_x$ for instance); source separation algorithms can then operate entirely in this n-dimensional subspace. As a consequence, we can deal with the case $n = m$ without loss of generality. Hence in this subsection (only !), we take $n = m$ so that all the matrices of interest are invertible. Denoting $\widehat{R}_s = T^{-1}\sum_{t=1,T} \mathrm{s}(t)\mathrm{s}(t)^*$, eq. (15) becomes

$$I_n = \widehat{W}\widehat{R}_x\widehat{W}^H = \widehat{W}A\widehat{R}_sA^H\widehat{W}^H. \quad (18)$$

It follows that $\widehat{W}$ satisfies $\widehat{W}A = V\widehat{R}_s^{-1/2}$ where $\widehat{R}_s^{-1/2}$ is the positive hermitian square root of $\widehat{R}_s^{-1}$ and V is some undetermined unitary matrix. Then

$$\widehat{A}^\# A = \mathcal{U}(\widehat{W}X_t)^H\widehat{W}A = \mathcal{U}(V\widehat{R}_s^{-1/2}S_t)^HV\widehat{R}_s^{-1/2}, \quad (19)$$

and by the orthogonal invariance property (16):

$$\widehat{A}^\# A = J\ \mathcal{U}(\widehat{R}_s^{-1/2}S_t)^H\widehat{R}_s^{-1/2} \quad (20)$$

for some quasi-identity matrix J. This shows that the value taken by $\widehat{A}^\# A$ depends only on S_T, the particular realization of the source signals and not on the mixing matrix. In fact, property (20) implies that orthogonal invariant algorithms become fully invariant in the noiseless case (again, we disregard the irrelevant quasi-identity matrix).

4. Orthogonal estimators in noise

To deal with the noisy situation and $m > n$, we consider the *polar decomposition* of matrix $\widehat{W}A$ as

$$\widehat{W}A = VH \quad (21)$$

where V is a $n \times n$ unitary matrix and H is a $n \times n$ hermitian, (almost surely) positive matrix. Combining these symmetry properties with (15), matrix H is found to be the unique $n \times n$ positive hermitian matrix verifying

$$H^2 = A^H\widehat{R}_y^\# A. \quad (22)$$

If R_y is known exactly, i.e. if $\widehat{R}_y = R_y = AA^H$, then $H = I_n$ so that the distance of H to the identity is a measure of the whitening error.

Some algebra upon (12), (15), (16) and (21) produces

$$\widehat{A}^\# A = J\ \mathcal{U}(H(S_T + A^\#\tilde{N}_T))^H H \quad (23)$$

$$\tilde{N}_T \stackrel{\text{def}}{=} (I_m + (\widehat{\Pi}\Pi)^\#\Pi^\perp)N_T. \quad (24)$$

for some quasi-identity matrix J.

This is the noisy equivalent of (20) and forms the basis to understanding how the invariance properties are affected by the additive noise. The key point here is that matrix A has 'almost' disappeared form (23): it only enters in a noise term $A^\#\tilde{N}_T$ and also affects the distribution of H.

A pairwise lower bound. Orthogonal algorithms rely on second-order prewhitening and are safe in this respect. However this very procedure introduces a lower bound the separating performance. The idea here is that errors in the whitening step cannot be compensated by any unitary matrix $\widehat{U}$. The required lemma is that H being an hermitian matrix, then for $p \neq q$ and any unitary matrix $\tilde{U}$:

$$|(\tilde{U}H)_{pq}|^2 + |(\tilde{U}H)_{qp}|^2 \geq \frac{|(H^2)_{pq}|^2}{(H^2)_{pp} + (H^2)_{qq}}. \quad (25)$$

Hence even the most clever choice of $\mathcal{U}$ in (23) cannot bring $P_{pq} + P_{qp}$ closer to zero than indicated by the right hand side of (25). This later term depends only on H and is then independent of a specific orthogonal algorithm. See (29) for an explicit evaluation in the i.i.d. case.

5. Asymptotics

To get further insights, we specify how $\widehat{R}_y^\#$ is computed from $\widehat{R}_x$ and go to the asymptotic domain.

5.1. Distribution of the whitening errors.

We consider the case of white Gaussian noise with covariance $R_n = \sigma I$ which allows for a simple estimation of R_y in the following standard fashion, based on an eigendecomposition of $\widehat{R}_x$. The estimated noise variance $\widehat{\sigma}$ is first obtained as the average of the $m - n$ smallest eigenvalues of $\widehat{R}_x$. Then, denoting $\lambda_1, \ldots, \lambda_n$ and $\mathbf{h}_1, \ldots, \mathbf{h}_n$ the n largest eigenvalues and associated eigenvectors of $\widehat{R}_x$, $\widehat{R}_y^\#$ is computed as $\widehat{R}_y^\# = \sum_{i=1,n}(\lambda_i - \widehat{\sigma})^{-1}\mathbf{h}_i\mathbf{h}_i^*$.

The first order expression of $\widehat{R}_y$ as a function of $\widehat{R}_x = R_x + \delta R_x$ may be computed by the standard perturbation technique. It is:

$$A^H\widehat{R}_y^\# A = I - A^\#\delta R_x A^{\#H} + \frac{Tr\ \Pi^\perp\delta R_x}{m-n}(A^HA)^{-1} + o(\delta R_x). \quad (26)$$

Inserting $\delta R_x = T^{-1}X_TX_T^* - R_x$, in the above shows, after minor rewriting, that $H^2 = A^H\widehat{R}_y^\# A$ depends on matrices $\sigma(A^HA)^{-1}$, S_T, $A^\# N_T$ and $\sigma^{-1/2}\Pi^\perp N_T$ which are mutually independent under the current assumptions on the noise. The distribution of $A^\# N_T$ depends only on the covariance matrix of $A^\#\mathbf{n}$ which is $\sigma(A^HA)^{-1}$, while the distribution of $\sigma^{-1/2}\Pi^\perp N_T$ does not depend on any parameter. We conclude that *asymptotically, in spatially and temporally white Gaussian noise, the distribution of H depends only on the distribution of the sources and on matrix $\sigma(A^HA)^{-1}$.*

5.2. The lower bound.

The asymptotic rejection rate $\mathcal{I}_{pq}$ is defined by

$$\mathcal{I}_{pq} \stackrel{\text{def}}{=} \lim_{T\to\infty} T\mathrm{E}P_{pq} = \lim_{T\to\infty} T\mathrm{E}|(\widehat{A}^\# A)_{pq}|^2. \quad (27)$$

Assume i.id. signals and noise so that the variance of $\widehat{R}_x$ decreases as T^{-1} and $(H^2)_{pp} \approx 1$. According to (25)

$$P_{pq} + P_{qp} \geq \frac{|(H^2)_{pq}|^2}{(H^2)_{pp} + (H^2)_{qq}} \approx \frac{1}{2}|(H^2)_{pq}|^2 \quad (28)$$

so that $\mathcal{I}_{pq} + \mathcal{I}_{qp} \geq \lim_{T\to\infty} T\mathrm{E}|(H^2)_{pq}|^2/2$ which is easily computed in the i.i.d. case using (26) and turns out not to depend on the distributions of the source signals:

$$\mathcal{I}_{pq} + \mathcal{I}_{qp} \geq \frac{1}{2}\left\{(1 + \sigma\rho_{pp})(1 + \sigma\rho_{qq}) + \frac{\sigma^2\rho_{pq}\rho_{qp}}{m-n}\right\} \quad (29)$$

where ρ_{pq} is the (p,q)-th entry of matrix $(A^H A)^{-1}$. In particular, in the low noise limit, the mean rejection rates are lower bounded by a numerical constant:

$$\lim_{\sigma\to\infty} \frac{\mathcal{I}_{pq}+\mathcal{I}_{qp}}{2} \geq \frac{1}{4}. \tag{30}$$

See a similar 1/4 factor due to whitening in an adaptive algorithm [9]. This bound is tight: it is reached for instance by the JADE algorithm [5, 10]. Note that $1/4T$ corresponds to a 26 dB rejection for 100 i.i.d. samples which is quite acceptable in many applications.

5.3. Rejection rates.

Herein, we assume circular i.i.d. signals and noise and large T. The distribution of $\widehat{A}^{\#}A$ can be characterized from (23). Note that matrix $(\widehat{\Pi}\Pi)^{\#}\Pi^{\perp}$ goes to zero as $\widehat{\Pi}$ converges to Π since $\Pi\Pi^{\perp}=0$ (it represents the amount of noise 'leaking' from the noise subspace into the signal subspace due to errors in $\widehat{\Pi}$) so that for large T, $\tilde{N}_T$ can be safely approximated by N_T in (23). By the same argument as above, we conclude that, for any orthogonal invariant algorithm in spatially and temporally white Gaussian noise, *the asymptotic distribution of $\widehat{A}^{\#}A$ depends only on the distributions of the sources and on matrix $\sigma(A^HA)^{-1}$.*

It follows that the asymptotic rejection rates take the form

$$\mathcal{I}_{pq} = f_{pq}(\sigma(A^HA)^{-1}, \mathcal{D}) \quad \forall p \neq q \tag{31}$$

where $\mathcal{D}$ represents the distributions of the sources. The functions f_{pq} depend on the specific orthogonal algorithm $\mathcal{U}$ used for the estimation of U, but our point here is that the dependence of the performance on the physical context is always via matrix $\sigma(A^HA)^{-1}$ which then quantifies the 'hardness' of the source separation problem. In particular, the significance of a given noise level may be determined by inspecting the entries of $\sigma(A^HA)^{-1}$.

As a worked out example, the JADE algorithm [5, 10] has an asymptotic performance given by

$$\mathcal{I}_{pq} = \sum_{r=0}^{4} \sigma^r g_{pq}^{(r)} (k_p^2+k_q^2)^{-2} \tag{32}$$

$$g_{pq}^{(0)} = k_q^4 + l_p k_p^2 + l_q k_q^2 \tag{33}$$

$$\begin{aligned} g_{pq}^{(1)} &= \rho_p p (k_p^2+k_q^2)^2 + \rho_p p[k_p^2(5k_p+6)+l_q k_q^2] \\ &+ \rho_q q[k_q^2(5k_q+6)+l_p k_p^2] \end{aligned} \tag{34}$$

$$g_{pq}^{(2)} = \cdots$$

where $l_p \stackrel{\text{def}}{=} E|s_p(t)|^6 - E^2|s_p(t)|^4$, $k_p \stackrel{\text{def}}{=} E|s_p(t)|^4 - 2E^2|s_p(t)|^2$, and $\rho_p = (A^HA)^{-1}_{pp}$. The performance index shows terms with the noise variance raised to the 4th power because the $\mathcal{U}$ function of JADE is based on 4th-order cumulants. Only the first two terms are given here, but the following ones also involve the coefficient $|\rho_{pq}|$. The lowest degree term shows the necessity of a *pairwise* bound since, if the q-th source is Gaussian (so that $k_q=0$) and the p-th source has a constant modulus (so that $l_p=0$), then $\mathcal{I}_{pq}=0$! This means that a Gaussian source, in this case, experiences super-rejection (numerical experiments show that its residual variance in the estimate of the p-th source decreases as T^{-2}.)

Conclusion

The most common class of *orthogonal invariant* algorithms for the blind separation of independent sources has been studied in terms of performance, quantified by rejection rates $\mathcal{I}_{pq}$. Emphasis was on the dependence of the performance on the physical parameters: the mixing matrix and the noise level. The following results were obtained.

- It exists a pairwise lower bound on the performance, introduced by the orthogonal approach.
- In the noiseless case, orthogonal invariant algorithms are fully invariant; hence, their performance does not depend on the mixing matrix; the lower bound is purely numerical for i.i.d. source signals.
- In the i.i.d. noisy case, performance depends only on the distribution of sources and on matrix $\sigma(A^HA)^{-1}$ which then characterizes the 'hardness' of source separation. The lower bound has a very simple expression (29) and does not depend on the distribution of the sources.

Proofs have been only sketched but detailed calculations are available upon request. We believe that further general results could be obtained along the lines of this paper. In particular, the dependence on $\sigma(A^HA)^{-1}$ may probably be specified: for instance, expression (34) shows that the linear term in σ in $\mathcal{I}_{pq}$ depends on the (p,p)-th and (q,q)-th entry of $\sigma(A^HA)^{-1}$. Generalizing this would be more informative than the general dependence outlined in this paper.

References

[1] L. Féty, "New methods for signal separation," in *Proc. of 4th Int. Conf. on HF radio systems and techniques*, (London), pp. 226–230, IEE, Apr. 1988.

[2] M. Gaeta and J.-L. Lacoume, "Source separation without a priori knowledge: the maximum likelihood solution," in *Proc. EUSIPCO*, pp. 621–624, 1990.

[3] L. Tong and R. Liu, "Blind estimation of correlated source signals," in *Proc. Asilomar conference*, Nov. 1990.

[4] P. Comon, "Independent component analysis," in *Proc. Int. Workshop on Higher-Order Stat., Chamrousse, France*, pp. 111–120, 1991.

[5] J.-F. Cardoso and A. Souloumiac, "An efficient technique for blind separation of complex sources," in *Proc. IEEE SP Workshop on Higher-Order Stat., Lake Tahoe, USA*, pp. 275–279, 1993.

[6] L. Tong, Y. Inouye, and R. Liu, "Waveform preserving blind estimation of multiple independent sources," *IEEE Tr. on SP*, vol. 41, pp. 2461–2470, July 1993.

[7] K. Abed Meraim, A. Belouchrani, J.-F. Cardoso, and E. Moulines, "Asymptotic performance of second order blind source separation," in *Proc. ICASSP*, vol. 4, pp. 277–280, Apr. 1994.

[8] D. T. Pham, P. Garrat, and C. Jutten, "Separation of a mixture of independent sources through a maximum likelihood approach," in *Proc. EUSIPCO*, pp. 771–774, 1992.

[9] B. Laheld and J.-F. Cardoso, "Adaptive source separation without prewhitening," *these proceedings.*

[10] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *IEE Proceedings-F*, vol. 140, pp. 362–370, Dec. 1993.