STABILITY ANALYSIS OF COMPLEX-VALUED NONLINEARITIES FOR MAXIMIZATION OF NONGAUSSIANITY

Mike Novey and Tülay Adalı

University of Maryland Baltimore County, Baltimore, MD 21250
{mnovey1, adali@umbc.edu}

ABSTRACT

Complex maximization of nongaussianity (CMN) has been shown to provide reliable separation of both circular and noncircular sources. It is also shown that the algorithm converges to the principal component of the source distribution when studied in the estimation direction. In this paper, we study the local stability of the CMN algorithm and determine the conditions under which local stability is achieved by extending our previous work to all dimensions of the weight vector. We use these conditions of stability to quantify convergence performance for a number of complex nonlinear functions, and present simulation results to demonstrate the effectiveness of these functions.

1. INTRODUCTION

Complex nonlinear functions, *i.e.,*, functions that are $\mathbb{C} \mapsto \mathbb{C}$, are shown to be highly effective in extracting higher order statistical information for achieving independent component analysis (ICA) [1]. In [4], they are used to implicitly generate the higher order statistics for maximization of nongaussianity and to derive the complex maximization of nongaussianity (CMN) algorithm. By examining the stability of the algorithm in the direction of the estimated source, the algorithm is shown to converge to the principal component of the source distribution [4]. The stability of the algorithm, however, depends on the behavior of the nonlinear function across all dimensions. In this paper, we perform a complete second-order analysis of the CMN algorithm across the entire weight vector and derive expressions that explain the interaction between the nonlinearity and the source statistics. We use the results of this analysis to define a stability measure which indicates how stable a nonlinearity is when the sources are noncircular. We apply this measure to several complex nonlinearities to quantify their performance and present simulation results to support the theory.

2. COMPLEX ICA BY MAXIMIZATION OF NONGAUSSIANITY

2.1. Complex preliminaries

A complex variable z is defined in terms of two real variables z^R and z^I as $z = z^R + jz^I$. Statistics of a complex random vector $\mathbf{x} = \mathbf{x}^R + j\mathbf{x}^I$ is defined by the joint probability density function (pdf) $p_x(\mathbf{x}^R, \mathbf{x}^I)$ provided that it exists. The expectation of a complex random vector $\mathbf{x}$ is then given with respect to this pdf and is written as

$$E\{\mathbf{x}\} = E\{\mathbf{x}^R\} + jE\{\mathbf{x}^I\}.$$

The covariance matrix is written as

$$\mathrm{cov}(\mathbf{x}) = E\{\mathbf{x} - E\{\mathbf{x}\}\}E\{\mathbf{x} - E\{\mathbf{x}\}\}^H$$

where H denotes Hermitian transform and the pseudocovariance matrix is defined as $\mathrm{pcov}(\mathbf{x}) = E\{\mathbf{x} - E\{\mathbf{x}\}\}E\{\mathbf{x} - E\{\mathbf{x}\}\}^T$, where T denotes the transpose. These two quantities together define a complex random vector, and the random vector is second-order circular if $\mathrm{pcov}(\mathbf{x}) = \mathbf{0}$. A stronger definition of circularity is based on the pdf of the complex random variable such that for any α, the pdf of z and $e^{j\alpha}z$ are the same [5].

2.2. Complex ICA

In ICA, the observed data $\mathbf{z}$ are typically expressed as a linear combination of latent variables such that $\mathbf{z} = \mathbf{A}\mathbf{s}$ where $\mathbf{s} = [s_1, \ldots, s_N]^T$ is the column vector of latent sources, $\mathbf{z} = [z_1, \ldots, z_N]^T$ is the column vector of observed mixtures, and matrix $\mathbf{A}$ is the $N \times N$ mixing matrix. ICA then identifies the statistically independent sources given the observed mixtures typically by estimating a matrix $\mathbf{W}$ so that the source estimates become $\hat{\mathbf{s}} = \mathbf{W}\mathbf{z}$. We will assume without loss of generality that the sources have zero mean and unit variance, i.e., $E\{s_k s_k^*\} = 1$.

2.3. CMN Algorithm

The CMN contrast function is written as

$$J(\mathbf{w}) = E\left\{\left|G(\mathbf{w}^H\mathbf{x})\right|^2\right\} \tag{1}$$

where $G : \mathbb{C} \mapsto \mathbb{C}$. This is similar to the complex variant of FastICA algorithm introduced in [3], which we refer to as the c-FastICA algorithm. In [3], nongaussianity is measured as

$$J(\mathbf{w}) = E\left\{G\left(\left|\mathbf{w}^H\mathbf{x}\right|^2\right)\right\} \tag{2}$$

and hence $G : \mathbb{R} \mapsto \mathbb{R}$. Comparing (1) and (2), it is clear that the nonlinearity, G, in c-FastICA is invariant to the phase information and hence assumes that the sources are circular. Incorporating phase information using (1) is shown to significantly improve the performance of the algorithm when the sources are noncircular [4].

The CMN algorithm requires a preliminary sphering or whitening transform $\mathbf{V}$, resulting in

$$\mathbf{x} = \mathbf{V}\mathbf{z} = \mathbf{V}\mathbf{A}\mathbf{s} = \hat{\mathbf{A}}\mathbf{s}$$

where $E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}$. We estimate each source, k, separately by finding a vector $\mathbf{w}$ such that

$$\hat{s}_k = \mathbf{w}_k^H\mathbf{x} = \mathbf{w}_k^H\hat{\mathbf{A}}\mathbf{s} = \mathbf{q}_k^H\mathbf{s} \tag{3}$$

where $\mathbf{q}_k = [0, \ldots, q_k, 0, \ldots]^T$. Constraining the source estimates such that $E\{\hat{s_k}\hat{s}_k^*\} = 1$, also constrains the weights to $|\mathbf{w}|^2 = 1$ due to the whitening transform. To calculate the optimal weight, i.e.,

$$\mathbf{w}_{opt} = \arg \max_{||\mathbf{w}||^2=1} E\{|G(\mathbf{w}^H \mathbf{x})|^2\}$$

we use gradient optimization followed by a normalization step, such that

$$\mathbf{w} = \mathbf{w} + \mu\nu \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^*}; \quad \mathbf{w} = \frac{\mathbf{w}}{||\mathbf{w}||^2}$$

where μ is the learning rate and ν is a parameter based on the sign of kurtosis as defined in [3].

3. STABILITY ANALYSIS

For the stability analysis, we use a second-order approximation of the cost function around the stable point. The approximation is a good one as long as the higher order terms of a Taylor series are negligible and the Hessian exists or is not singular at the stable point. For the cases where the second-order analysis fails to describe the true behavior, we demonstrate the properties graphically in section 4.

We examine the stability of $J(\mathbf{w})$ under the constraint $||\mathbf{w}||^2 = 1$. We use the orthogonal change of coordinates $\mathbf{q} = \mathbf{A}^H \mathbf{w}$, such that $J(\mathbf{q}) = E\{|G(\mathbf{q}^H \mathbf{s})|^2\}$ and $||\mathbf{q}||^2 = 1$. The optimal solution is given by $\mathbf{q}_k = [0, \ldots, q_k, \ldots, 0]^T$, where $q_k = e^{j\theta}$, the condition when $\mathbf{w}$ equals one of the rows of $\mathbf{A}^{-1}$. Without loss of generality, we assume a solution at $\mathbf{q}_1$.

Previous analysis in [4], using a Maclaurin series expansion along the direction of the estimated source, has shown that in the complex plane of $q_1 = e^{j\theta}$, the extrema are at

$$\theta = \frac{1}{2}\text{atan}\left(\frac{-2E\left((s_1^R)^3 s_1^I + (s_1^I)^3 s_1^R\right)}{E\left((s_1^R)^4 - (s_1^I)^4\right)}\right) + \frac{n\pi}{2} \qquad (4)$$

for integer n. This result shows us that as we traverse along the unit circle in the complex plane of q_1, there exist at least two minima and two maxima at $n = \{0, 1, 2, 3\}$. If the sources are circular however, all values of θ are stable, *i.e.*, the gradient is zero with respect to q_1. In this paper, we extend the stability conditions of the CMN algorithm to the other dimensions, namely $q_2, \ldots, q_N$ using a second-order (Hessian) analysis. We show that the overall stability is particularly important to show the interdependency of stability on multiple noncircular sources, which is ignored when concentrating only on the estimated source dimension, q_1.

The major result of this section, derived in the Appendix based on a second-order analysis, is that a local minimum (*resp.* maximum) is achieved for a given source ($i = 2, 3, \ldots$) when

$$\gamma(s_1) \pm |E\{s_i^2\}\beta(s_1)| > 0 \qquad (\textit{resp.} < 0) \qquad (5)$$

where we have defined

$$\gamma(s_1) = E\left\{g(e^{-j\theta}s_1)g^*(e^{-j\theta}s_1) - G(e^{-j\theta}s_1)g^*(e^{-j\theta}s_1)s_1^* e^{j\theta}\right\}$$

and $\beta(s_1) = E\{G^*(e^{-j\theta}s_1)g'(e^{-j\theta}s_1)\}$, and used the notation $g(z) = dG(z)/dz$ and $g'(z) = dg(z)/dz$. For circular sources, the expression reduces to

$$E\{g(s_1)g^*(s_1) - G(s_1)g^*(s_1)s_1^*\}\} > 0 \quad (\textit{resp.} < 0) \qquad (6)$$

which coincides with the result given in [3] assuming circularity that implies $E\{s_i^2\} = 0$.

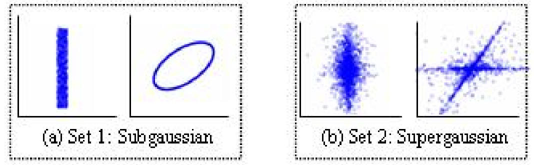


Fig. 1. Scatter plot (real vs. imaginary) of the two sets of sources used in the simulations.

We define a measure of performance, stability index (I_S), by solving for $E\{s_i^2\}$ in (5)

$$I_S \equiv \left|\frac{\gamma(s_1)}{\beta(s_1)}\right| > |E\{s_i^2\}|. \qquad (7)$$

Defining the stability index in this way provides: 1) a measure of the second-order noncircularity that can be tolerated and 2) an indication of the condition number (κ), ratio of the maximum eigenvalue to the minimum, providing an indication of the expected convergence behavior. We can see from (7) that a high I_S implies convergence for sources with large second-order noncircularity, $E\{s_i^2\}$. A large I_S also implies a condition number close to one, hence good convergence behavior, as seen by substituting $\gamma(s_1)$ and $\beta(s_1)$ into (23)

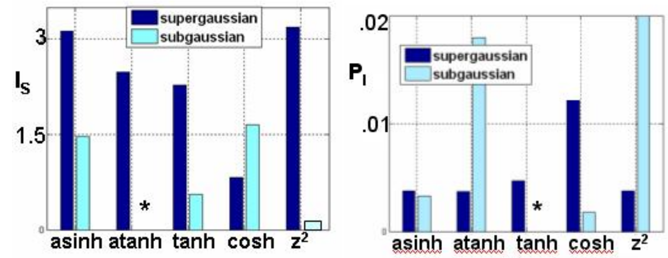
$$\kappa \equiv \frac{\gamma(s_1) + |E\{s_i^2\}\beta(s_1)|}{\gamma(s_1) - |E\{s_i^2\}\beta(s_1)|}. \qquad (8)$$

As shown in the simulations section, some nonlinearities yield robust results with both sub and supergaussian sources, whereas some might yield good results for only a certain type of input distributions.

4. NUMERICAL STUDY

In this section, we first show the performance of different nonlinearities for $G(z)$, asinh z, atanh z, tanh z, cosh z, acosh z, and z^2 using the measureI_S. We use two sets of sources, a subgaussian and a supergaussian set both shown in Figure 1. The subgaussian set includes a uniform source and a correlated complex sinusoid and the supergaussian set includes Laplacian sources. The first source of each set has an asymmetry ratio of 0.1, defined as the ratio of real to imaginary standard deviation σ^R/σ^I and the second source includes correlated real and imaginary parts of same variance. The performance measure used is the mean-squared error from the permutation matrix $Q = \hat{\mathbf{A}}^H \mathbf{W}$ as in [3] and is denoted by P_I. Figure 2 shows the results of 200 runs for I_S and P_I for each nonlinearity. An asterisk (*) indicates that data were not available due to convergence problems. In Figure 2, we observe a direct correlation between the values for I_S and P_I for each nonlinearity, a higher tolerance to second-order noncircularity (measured by I_S) leading to a better performance index P_I. For example, z^2 (kurtosis), still converges with subgaussian sources but the value of I_S explicitly tells us that $E\{s_2^2\} < I_S$. In this example, if the second source made a small increase in noncircularity, convergence may be compromised.

Figure 3 illustrates how convergence rate is affected by noncircularity of subgaussian sources. Here we have number of iterations to convergence as a function of the asymmetry measured by σ^R/σ^I. As the sources become more noncircular ($E\{s_k^2\}$ increases), the condition number κ increases. This increase in κ increases the convergence rate as seen in the Figure, but affects asinh less then cosh. We also observe this property in the I_S value of asinh being greater then



(a) I_S vs. nonlinearity (b) P_I vs. nonlinearity

Fig. 2. Numerical values of stability index, I_S, and ICA performance measure P_I ('*' indicates data were not available)

.

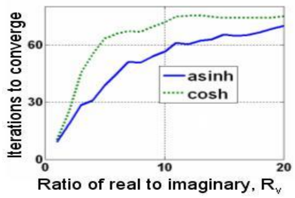


Fig. 3. Number of iterations to convergence for asinh and cosh nonlinearities as a function of σ^R/σ^I. Results are shown for the supergaussian source set.

that of cosh in Figure 2(a). As expected the I_S value is an indicator of κ and therfore the expected convergence behavior.

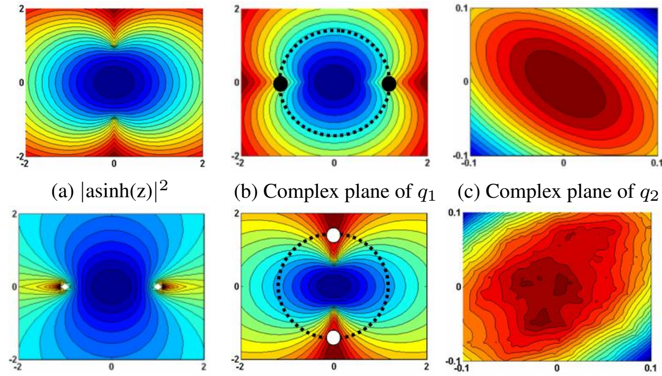
We next show a case when the second-order approximations fail to accurately describe the optimization landscape around the stable point $\mathbf{q}_1$, and demonstrate the behavior of the nonlinearities graphically. Two nonlinearities, asinh and atanh, are studied with the same set of noncircular subgaussian sources shown in Figure 1.

Figure 4 shows the optimization landscape in the complex plane of q_1 and q_2 along with the magnitude of the nonlinearities. As seen in Figures 4(b) and 4(e), both stable points are found by maximizing the contrast function shown as black/white dots on the constraint circle. In the q_2 direction, 4(c) and 4(f), asinh provides a smooth optimization landscape while the atanh nonlinearity shows many maxima/minima, albeit far from the optimal solution of $q_2 = 0$. Note the singularities for atanh at ± 1 in 4(d) and at $\pm j$ in 4(e), leading to the non-convex landscape in the q_2 direction rendering second-order analysis inaccurate. This result explains as to why some nonlinearities perform well with supergaussian sources and not subgaussian or vice versa.

If, on the other hand, we study the landscape with the supergaussian source set, the stable point is now at a minimum which corresponds to the point $e^{j\pi/2}$ in Figure 4(b) for asinh and e^{j0} in Figure 4(e) for atanh. The stable point now falls into a smooth region for both asinh and atanh indicating good performance for both. This is also seen in Figure 2(a) where atanh has a high value of I_S with supergaussian sources only.

5. CONCLUSION

We have derived the stability conditions for the CMN algorithm and defined a measure, I_S, for quantifying performance with noncircular sources. This measure provides information as to not only how well the algorithm performs but also how noncircular the sources can be and still convergence can be achieved. We also found that I_S provides a measure as to how sensitive the convergence rate is to noncircularity. Simulations were used to confirm these results. We also examined the effect of singularities in the neighborhood of the stable point and showed through graphical means that performance is degraded due to the existence of many extrema far from the optimal solution.



(a) $|\text{asinh}(z)|^2$ (b) Complex plane of q_1 (c) Complex plane of q_2

(d) $|\text{atanh}(z)|^2$ (e) Complex plane of q_1 (f) Complex plane of q_2

Fig. 4. Optimization landscapes in the complex planes of q_1 and q_2 with asinh (top row) and atanh (bottom row) nonlinearities and subgaussian sources. Figures (b) and (e) have the constraint circle and the stable points shown. Red indicates large values.

A. APPENDIX

A.1. Preliminaries

We use the mapping

$$\begin{pmatrix} z \\ z^* \end{pmatrix} = \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix} \begin{pmatrix} z^R \\ z^I \end{pmatrix}$$

to write vector $[\mathbf{z}^{R^T} \mathbf{z}^{I^T}]^T$ as $\mathbf{z} = [z_1, z_1^*, \ldots, z_N, z_N^*]^T$ through the transformation $\mathbb{R}^{2N} \mapsto \mathcal{M} \in \mathbb{C}^{2N}$.

The Taylor series expansions of $f : \mathbb{R}^{2N} \mapsto \mathbb{R}$ and $f_C : \mathcal{M} \mapsto \mathbb{R}$ are the same when the following gradient

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial z^R} - j\frac{\partial f}{\partial z^I}\right); \quad \frac{\partial f}{\partial z^*} = \frac{1}{2}\left(\frac{\partial f}{\partial z^R} + j\frac{\partial f}{\partial z^I}\right) \tag{9}$$

and Hessian forms

$$\nabla^2_{qq} f(\mathbf{z_0}) = \left.\frac{\partial^2 f(\mathbf{z})}{\partial \mathbf{z}^* \mathbf{z}^T}\right|_{\mathbf{z}=\mathbf{z}_0} \tag{10}$$

are used [6].

A.2. Derivation of two derivative forms

In this section, we show that for $H(z) = f_1(z^*x) f_2^*(z^*x)$ where f_1 and f_2 are analytic functions, we have

$$\frac{\partial H(z)}{\partial z} = x^* f_1(z^*x) f_2'^*(z^*x) \tag{11}$$

and

$$\frac{\partial H(z)}{\partial z^*} = x f_1'(z^*x) f_2^*(z^*x) \tag{12}$$

where $f'(z) = df(z)/dz$.

We first use (9) to write

$$\frac{\partial f^*(z^*x)}{\partial z} = \frac{1}{2}\left(\frac{\partial f^*(z^*x)}{\partial z^R} - j\frac{\partial f^*(z^*x)}{\partial z^I}\right). \quad (13)$$

We then write $f(z^*x)$ as the sum of two real-valued functions, $u(a,b)+jv(a,b)$. We use (13) and the chain rule to obtain

$$\begin{aligned}\frac{\partial f^*(z^*x)}{\partial z} &= \frac{1}{2}\left[\left(f_a^*\frac{\partial a}{\partial z^R} + f_b^*\frac{\partial b}{\partial z^R}\right) - j\left(f_a^*\frac{\partial a}{\partial z^I} + f_b^*\frac{\partial b}{\partial z^I}\right)\right] \\ &= \frac{1}{2}x^*(f_a^* + jf_b^*).\end{aligned}$$

Substituting $f_a^* = u_a - jv_a$, $f_b^* = u_b - jv_b$ and the Cauchy-Riemann conditions, $u_a = v_b$ and $u_b = -v_a$, we obtain

$$\frac{\partial f^*(z^*x)}{\partial z} = \frac{1}{2}x^*\left(u_a - jv_a + j(u_b - jv_b)\right) = x^* f'^*(z^*x).$$

Following similar algebraic steps and (9), one can find the following relationships:

$$\frac{\partial f(z^*x)}{\partial z} = 0, \quad \frac{\partial f^*(z^*x)}{\partial z^*} = 0, \quad \text{and} \quad \frac{\partial f(z^*x)}{\partial z^*} = xf'(z^*x).$$

Applying the chain rule and the above equations, we finally obtain (11) and similarly (12).

A.3. Derivation of the stability conditions

We study the stability of the CMN algorithm, by examining the Lagrangian function for the constrained optimization problem

$$L(\mathbf{q},\lambda) = J(\mathbf{q}) + \lambda(h(\mathbf{q})) \quad (14)$$

at $\mathbf{q}_1$, where λ is the real-valued Lagrange multiplier and $h(\mathbf{q})$ is the constraint function, $\|\mathbf{q}\|^2 - 1$. The second-order necessary and sufficient conditions for a local minimum (*resp.* maximum) are

$$\nabla_q L(\mathbf{q}_o) = 0, \quad \nabla_\lambda L(\lambda_o) = 0 \quad (15)$$

and

$$\mathbf{y}^H \nabla^2_{qq} L(\mathbf{q}_o)\mathbf{y} \geq 0 \quad (\textit{resp.} \leq 0) \quad (16)$$

with $\mathbf{y}$ defining the feasible directions $\{\mathbf{y}|\nabla h(\mathbf{q}_o)^T\mathbf{y} = 0\}$ [2]. For the unit norm constraint, the feasible directions thus become

$$[e^{-j\theta}, e^{j\theta}, 0, \ldots]\,[y_1, y_1^*, \ldots, y_N^*]^T = 0$$

which constrains $y_1^* = -y_1 e^{-j2\theta}$ and imposes no constraints on y_i, $i = 2, \ldots, N$.

We first evaluate condition (15) by using (11) and (12) in (14) and evaluating at $\mathbf{q}_1$. Noting that the sources are independent with zero mean, we find

$$\nabla_q L(\mathbf{q}_1) = E\begin{pmatrix} G(q_1^*s_1)g^*(q_1^*s_1)s_1^* \\ G^*(q_1^*s_1)g(q_1^*s_1)s_1 \\ 0 \\ \vdots \end{pmatrix} + \lambda_0 \begin{pmatrix} q_1^* \\ q_1 \\ 0 \\ \vdots \end{pmatrix}. \quad (17)$$

Setting the top row of (17) equal to zero and solving for λ_0, we obtain

$$\lambda_0 = -E\{G(e^{-j\theta}s_1)g^*(e^{-j\theta}s_1)s_1^*\}e^{j\theta} \quad (18)$$

where we made the substitution $q_1 = e^{j\theta}$.

To evaluate condition (16), we first find the second derivatives, and again using (11) and (12), we obtain

$$\begin{array}{ll} \frac{\partial^2 J}{\partial q_i^* q_j} = s_i s_j^* g g^* & \frac{\partial^2 J}{\partial q_i^* q_j^*} = s_i s_j G g'^* \\ \frac{\partial^2 J}{\partial q_i q_j} = s_i^* s_j^* G g'^* & \frac{\partial^2 J}{\partial q_i q_j^*} = s_i s_j^* g^* g \end{array} \quad (19)$$

where g' is the derivative of g. Substituting (19) into (10) and evaluating at $\mathbf{q}_1$ we obtain the block diagonal matrix for the Hessian

$$\nabla^2_{qq} L(\mathbf{q}_1) = E\begin{pmatrix} \mathbf{A} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{B}_N \end{pmatrix} + \lambda_0 \mathbf{I} \quad (20)$$

where $\mathbf{A}$ and $\mathbf{B}_i$ are each 2×2 matrices defined as

$$\mathbf{A} = \begin{pmatrix} s_1 s_1^* g^*(q_1^*s_1)g(q_1^*s_1) & s_1^2 G^*(q_1^*s_1)g'(q_1^*s_1) \\ (s_1^*)^2 G(q_1^*s_1)g'^*(q_1^*s_1) & s_1 s_1^* g^*(q_1^*s_1)g(q_1^*s_1) \end{pmatrix}$$

and

$$\mathbf{B}_i = \begin{pmatrix} g^*(q_1^*s_1)g(q_1^*s_1) & s_i^2 G^*(q_1^*s_1)g'(q_1^*s_1) \\ (s_i^*)^2 G(q_1^*s_1)g'^*(q_1^*s_1) & g^*(q_1^*s_1)g(q_1^*s_1) \end{pmatrix}. \quad (21)$$

Simplifications occur from applying the independence, unit variance and zero mean properties of the sources.

We can evaluate condition (16) by analyzing each submatrix in (20) separately. Submatrix $\mathbf{A}$ must satisfy $\mathbf{y}_1^H(\mathbf{A} + \lambda_0)\mathbf{y}_1 > 0$ (minimum), where $\mathbf{y}_1 = [y_1, -y_1 e^{-j2\theta}]^T$ from the feasible directions constraint. Substituting (18) into the above we obtain the condition along dimension q_1 for a minimum (*resp.* maximum)

$$2(E\{A_{11}\} + \lambda_0) > E\{A_{12}e^{-j2\theta} + A_{12}^* e^{j2\theta}\} \quad (\textit{resp.} <) \quad (22)$$

where A_{ij} is the i^{th} row and j^{th} column of $\mathbf{A}$. Expanding the nonlinearity G into its Maclaurin series expansion $G(z) = \alpha_0 + \alpha_1 z + \ldots + \alpha_k z^k$ yields the result in our previous work (4).

We examine the behavior along the remaining dimensions $[q_2, q_2^*, \ldots, q_N, q_N^*]^T$ through submatrix $\mathbf{B}_i$. Evaluating condition (16) for stability requires each $(\mathbf{B}_i + \lambda_0\mathbf{I})$ to be positive definite for a local minimum and negative definite for a local maximum. Noting that the form of $\mathbf{B}_i$ is $\mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^* & B_{11} \end{pmatrix}$ the eigenvalues are easily shown to be $\mathrm{eig}(\mathbf{B}+\lambda_0\mathbf{I}) = B_{11} + \lambda_0 \pm |B_{12}|$, which are real-valued by inspection of the structure of matrix $\mathbf{B}$. The conditions for a local minimum (*resp.* maximum) are found by using (21) as

$$\begin{aligned} &E\left\{g(e^{-j\theta}s_1)g^*(e^{-j\theta}s_1)\right\} + \lambda_0 \pm \\ &\left|E\{s_i s_i\}E\{G^*(e^{-j\theta}s_1)g'(e^{-j\theta}s_1)\}\right| > 0 \quad (\textit{resp.} < 0.) \end{aligned}$$

and is given in (5) where we substituted (18) for λ_0. Also of interest is the condition number (κ), ratio of largest to smallest eigenvalue, of each submatrix of $\nabla^2_{qq} L(\mathbf{q}_1)$ which is found to be

$$\kappa = \frac{e_{\max}}{e_{\min}} = \frac{(B_{11} + \lambda_0) + |B_{12}|}{(B_{11} + \lambda_0) - |B_{12}|}. \quad (23)$$

B. REFERENCES

[1] T. Adalı, T. Kim and V. Calhoun, "Independent component analysis by complex nonlinearities," in *Proc. ICASSP,* Montreal, Canada, May 2004, vol. 5, pp. 525–528.

[2] D. Bertsekas, Nonlinear Programming, Athena Scientific, 1995.

[3] E. Bingham and A. Hyvarinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *Int J. Neural Systems,* 10(1):1, pp. 1–8, 2000.

[4] M. Novey and T. Adalı, "ICA by maximization on nongaussianity using complex functions," in *Proc. MLSP,* Mystic, CT, 2005.

[5] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing,* vol 42, p. 3473, December 1994.

[6] A. van den Bos, "Complex gradient and Hessian," *IEE Proc. Vision, Image, and Signal Processing,* vol 141, no. 6, pp. 380–383, December 1994.