SENSITIVITY ANALYSIS FOR THE PROBLEM OF MATRIX JOINT DIAGONALIZATION*

BIJAN AFSARI†

Abstract. We investigate the sensitivity of the problem of Non-Orthogonal (matrix) Joint Diagonalization (NOJD). First, we consider the uniqueness conditions for the problem of Exact Joint Diagonalization (EJD), which is closely related to the issue of uniqueness in tensor decompositions. As a by-product, we derive the well-known identifiability conditions for Independent Component Analysis (ICA), based on an EJD formulation of ICA. We introduce some known cost functions for NOJD and derive flows based on these cost functions for NOJD. Then we define and investigate the noise sensitivity of the stationary points of these flows. We show that the condition number of the joint diagonalizer and uniqueness of the joint diagonalizer as measured by modulus of uniqueness (as defined in the paper) affect the sensitivity. We also investigate the effect of the number of matrices on the sensitivity. Our numerical experiments confirm the theoretical results.[1]

Key words. Joint Diagonalization, Independent Component Analysis(ICA), Simultaneous Diagonalization, Sensitivity Analysis, Perturbation Analysis, CANDECOMP/PARAFAC, Tensor Decompositions

AMS subject classifications. 15A23, 15A69, 15A52, 74P20, 49Q12

1. Introduction and a case study.

Many interesting recent problems and paradigms in blind signal processing can be formulated as the problem of matrix Joint Diagonalization (JD). This problem in its simplest form can be phrased as: given a set of N symmetric matrices $\{C_i\}_{i=1}^N$ of dimension $n \times n$, find a non-singular matrix B such that all BC_iB^T's are "as diagonal as possible," where B^T denotes the transpose of matrix B. Note that here diagonalization is meant in the sense of congruence. Matrix joint diagonalization problem is also referred to as simultaneous matrix diagonalization. In practice, i.e., when C_i's are constructed from empirical data we do not expect a B to exist such that all BC_iB^T's are diagonal. Therefore, maybe a more exact name for this problem can be Approximate Joint Diagonalization. Nevertheless, we choose to call this problem as joint diagonalization, where approximation is implicitly assumed and we refer to the problem when exact joint diagonalization is possible as Exact Joint Diagonalization (EJD).

Historically, the problem of matrix joint diagonalization in the signal processing community was first considered in the restricted form of Orthogonal Joint Diagonalization (OJD) in [9], where an efficient algorithm for it was proposed. In the OJD problem the joint diagonalizer is assumed to be orthogonal. This situation can happen for example when one tries to blindly separate non-Gaussian sources that are spatially whitened [9]. The orthogonality assumption on B is not justified in many occasions and one expects that by allowing more freedom in the search space, "more diagonalization" would be possible. We refer to the problem of joint diagonalization, when B is only assumed to be non-singular, as the problem of Non-Orthogonal Joint Diagonalization or NOJD. The focus of this paper is the NOJD problem.

*This research was supported in part by Army Research Office under ODDR&E MURI01 Program Grant No. DAAD19-01-1-0465 to the Center for Communicating Networked Control Systems (through Boston University).

†Department of Applied Mathematics, University of Maryland, College Park, 20740 MD,USA(bijan@glue.umd.edu).

[1]While this paper was under review few results in this paper were presented in ICASSP07 conference in Honolulu, HI [4].

Non-orthogonal joint diagonalization arises in a variety of problems. As a case study, we will see that how the problem of Independent Component Analysis (ICA) can be considered as an NOJD problem. In the problem of blind separation of non-stationary mixtures [21], one can perform NOJD on a set of correlation matrices to find the un-mixing matrix. Blind separation of instantaneous mixtures using only second order statistics, also results in NOJD of a set of covariance matrices [7]. Moreover, the NOJD problem is closely related to the problem of tensor decomposition and CNADECOMP/PARAFAC modeling [18, 11]. Since applications or algorithms are not the focus of this work, we will not cite numerous applications where the NOJD problem is useful. Instead, we consider the ICA problem, as a case study, to give the reader a feeling of the recurring situation, where the NOJD problem shows itself in numerous applications.

1.1. A case study: Independent Component Analysis (ICA). Independent Component Analysis (ICA) [10] is one of the major paradigms in which joint diagonalization and tensorial methods have proven useful. We refer the reader to [17] for further discussion on this issue. The basic model in ICA is:

$$\vec{\mathbf{x}}_{n\times 1} = A_{n\times n}\vec{\mathbf{s}}_{n\times 1} \tag{1.1}$$

where $\vec{\mathbf{s}}_{n\times 1}$ is a random vector of dimension n with independent components of zero mean, and A is an $n \times n$ non-singular matrix. We can think of $\vec{\mathbf{s}}_{n\times 1}$ to represent a source with independent components whose signals are mixed by the mixing matrix A and $\vec{\mathbf{x}}_{n\times 1}$ to represent the observed mixture. The problem is to find the matrix A or its inverse, assuming that only realizations or the moments of the random mixture $\vec{\mathbf{x}}_{n\times 1}$ are available . Obviously, we can only hope to find A up to column permutation and column scaling. The key assumption of independence of the elements of $\vec{\mathbf{s}}$ imposes some specific structure on certain matrices that can be formed from the cumulants of the observation $\vec{\mathbf{x}}$. The main theme here is that independence implies diagonality. We investigate this further. First, note that $R_{\mathbf{xx}}$, the covariance matrix of $\vec{\mathbf{x}}$, satisfies:

$$R_{\mathbf{xx}} = A\Lambda_{\mathbf{ss}}A^T, \tag{1.2}$$

where $\Lambda_{\mathbf{ss}}$ is the (diagonal) covariance matrix of $\vec{\mathbf{s}}$. We can trace this structure in higher cumulants of $\vec{\mathbf{x}}$ as well. The k^{th} order cumulant of a random vector $\vec{\mathbf{z}}_{n\times 1}$ is a tensor $\mathcal{C}^k_{\mathbf{z}}$ of order k and dimension $n \times ... \times n$. The cumulants are closely related to the moments, and they give information about the shape of the probability density function of $\vec{\mathbf{z}}_{n\times 1}$. In fact, the second order cumulant tensor is the covariance matrix. Each element of $\mathcal{C}^k_{\mathbf{z}}$ can be indexed by k indices $i_1, ...i_k$, where $1 \leq i_1, ..., i_k \leq n$. If we fix all but two indices and vary the remaining two indices we obtain a matrix slice of the tensor. The notation $\mathcal{C}^k_{\mathbf{z}}(i_1, i_2, ..., i_{k-2}, :, :)$ represents such a matrix that is found by fixing all but the last two indices. An important fact is that, if $\vec{\mathbf{z}}_{n\times 1}$ is of independent components, then its cumulant tensors of any order are diagonal. Since $\vec{\mathbf{s}}_{n\times 1}$ is of independent components, its cumulant tensors are diagonal, i.e. only the elements $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$ can be non-zero. Based on the multi-linear property of cumulants we can show that for $k \geq 3$:

$$\mathcal{C}^k_{\mathbf{x}}(i_1, i_2, ..., i_{k-2}, :, :) = A\Lambda_{i_1 i_2 ... i_{k-2}}A^T, \tag{1.3}$$

where $\Lambda_{i_1 i_2 ... i_{k-2}}$ is a diagonal matrix that depends on elements of A and $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s, the auto-cumulants of $\vec{\mathbf{s}}$, as:

$$[\Lambda_{i_1 i_2 ... i_{k-2}}]_{ii} = a_{i_1 i}a_{i_2 i}...a_{i_{k-2} i}\mathcal{C}^k_{\mathbf{s}}(i, ..., i), \quad 1 \leq i \leq n. \tag{1.4}$$

Note that (1.2) is also of this form except that the diagonal matrix $\Lambda_{\mathbf{ss}}$ does not depend on A. There is a profound difference between cumulant matrix slices of order higher than two and the covariance matrix of $\vec{\mathbf{x}}_{n\times 1}$, in that the latter is always positive definite whereas the former need not be of any definite sign, and their signs depend both on the signs of the $\mathcal{C}^k_{\mathbf{s}}(i,...,i)$'s as well as the elements of A. From (1.2) and (1.3)one can see how NOJD and ICA are related: in order to find A^{-1} search for a non-singular matrix B that jointly diagonalizes all the cumulant matrix slices, including the covariance matrix. In Section 2.2 we show that under certain conditions, which are basically the uniqueness conditions for the EJD problem, A can be found (up to the inherent indeterminacies) from the NOJD of the cumulant slices. The interesting point here is that restoration of diagonality can be equivalent to restoration of independence, and in this process we do not need to know much about the source $\vec{\mathbf{s}}_{n\times 1}$ or its statistical distribution.

1.2. Scope and organization of the paper. In [29, 3, 6, 28, 25, 1, 20] and many other works, different algorithms have been proposed to find the non-orthogonal joint diagonalizer of a given set of matrices. Although one might think of other ideas, the NOJD problem has been considered as a minimization problem whose solution gives the joint diagonalizer. There are not so many cost functions known that can be used for this purpose. Given a set of matrices:

$$C_i \approx A\Lambda_i A^T, \quad 1 \le i \le N, \tag{1.5}$$

where Λ_i's are diagonal; the hope of NOJD is that if a B if found such that all BC_iB^T's are "as diagonal as possible", then B is close to A^{-1} up to permutation and diagonal scaling. Therefore, the accuracy or usefulness of an NOJD algorithm depends on the actual algorithm and on the cost function used, in the sense that how its minimizers differ from A^{-1} when we have (1.5) instead of an equality. The focus of this work is on what factors affect the sensitivity of the NOJD cost functions. Using a perturbation analysis for the stationary points of certain minimization flows, we will show that this sensitivity is closely related to the uniqueness properties of the corresponding exact joint diagonalization problem. Also, non-unexpectedly, we show that if norm of A^{-1} is large, then again the NOJD will be sensitive. Note that this can happen if norm of A is small or if A is ill-conditioned. One of our main motivations in considering the sensitivity issue is to investigate the effect of the number of matrices included in the NOJD process. Inclusion of more matrices can not only help to reduce the harm of noise by an averaging effect but also by reducing the sensitivity through improvement of measures of uniqueness defined in Section 2.

The organization of this paper is as follows: in Section 2 we investigate the uniqueness conditions for the problem of exact joint diagonalization. We also use this result to derive the well known identifiability conditions for the ICA problem [10]. In Section 3 we introduce some of the known cost functions for NOJD and derive the corresponding flows whose stationary points characterize the joint diagonalizers. In Section 4 we perform a perturbation analysis on the stationary points of the introduced flows in order to find the sensitivity properties. We also elaborate on the effect of the number of matrices in the NOJD process. Numerical simulations in Section 5 confirm the derived results.

1.3. Notations. Throughout the paper all variables are real valued unless otherwise stated. Boldface small letters denote random variables. A and B both are $n\times n$ non-singular matrices unless otherwise stated. If X is a matrix, x_{ij} or X_{ij} or $[X]_{ij}$

denotes its entry at position (i,j). $\|X\|_F$ and $\|X\|_2$ denote the Frobenius norm and the 2-norm of the matrix X, respectively. X^T denotes the transpose of X, and X^{-T} denotes the transpose of the inverse of X. $\mathrm{tr}(X)$ is the trace of the square matrix X. $\mathrm{cond}(A)$ is the 2-norm based condition number of the matrix A. For a square matrix, $\mathrm{diag}(X)$ is the diagonal part of X, i.e., a diagonal matrix whose diagonal is equal to the diagonal of X. I or $I_{n\times n}$ denotes the $n \times n$ identity matrix. Unless otherwise stated, letters D and Π denote a non-singular diagonal matrix and a permutation matrix, respectively. For a vector x, $\mathrm{diag}(x)$ is a diagonal matrix with diagonal x. Λ_i is a diagonal matrix and we denote the k^{th} diagonal element of Λ_i by λ_{ik}. $\|x\|$ is the 2-norm of the vector x. We also define $X^\circ = X - \mathrm{diag}(X)$. $\mathrm{GL}(n)$ and $\mathrm{SO}(n)$ denote the Lie groups of non-singular $n \times n$ matrices and orthogonal $n \times n$ matrices with $+1$ determinant, respectively. T_pM denotes the tangent space of the manifold M at point p on the manifold. Notation $X \leftarrow Y$ means that: "the new value of X is Y."

2. Uniqueness conditions for Exact Joint Diagonalization (EJD). Consider matrices:

$$C_i = A\Lambda_i A^T, \quad 1 \leq i \leq N, \tag{2.1}$$

where Λ_i's are diagonal matrices, i.e., $\Lambda_i = \mathrm{diag}([\lambda_{i1}, ...\lambda_{in}])$. One interesting problem is: given only $\{C_i\}_{i=1}^N$ find A. We call this problem the Exact Joint Diagonalization or the EJD problem. Note that with the only information that Λ_i's are diagonal A can be determined only up to permutation and diagonal scaling, i.e., if A is a solution then $AD\Pi$ is also a solution, for any D and Π. We say that the EJD has a unique[2] solution if the permutation and diagonal scaling are the only ambiguities in finding A. If the EJD has a unique solution then finding A is equivalent to finding a $B \in \mathrm{GL}(n)$ such that all BC_iB^T's are diagonal, hence the name "joint diagonalization."

The issue of uniqueness in the EJD problem can be considered as a special case of uniqueness in the CANDECMOP/PARAFAC model, which has been addressed in [14]. In order to quantify the uniqueness property, which as will be seen in Section 4 is closely related to the sensitivity issue of the NOJD problem, we re-phrase the necessary and sufficient conditions for uniqueness differently from the related literature.

Definition 2.1. *For the set of diagonal matrices $\{\Lambda_i\}_{i=1}^N$ let:*

$$\rho_{kl} = \frac{\sum_{i=1}^N \lambda_{ik}\lambda_{il}}{(\sum_{i=1}^N \lambda_{il}^2)^{\frac{1}{2}}(\sum_{i=1}^N \lambda_{ik}^2)^{\frac{1}{2}}}, \quad 1 \leq k \neq l \leq N, \tag{2.2}$$

with the convention that $\rho_{kl} = 1$ if $\lambda_{ik} = 0$ for some k and all i. Let ρ be equal to one of the ρ_{kl}'s that have the maximum absolute value among all. The Modulus of Uniqueness for this set is defined as $|\rho|$.

Note that $|\rho| \leq 1$ and $|\rho| = 1$ if and only if at least two columns of the matrix $[\Lambda]_{ij} = \lambda_{ij}$ are collinear, i.e., if there is a real number K and integers p and q such that $\lambda_{ip} = K\lambda_{iq}$ for $1 \leq i \leq N$. $|\rho|$ measures the maximum degree of collinearity between any two columns of the matrix $[\Lambda]_{ij} = \lambda_{ij}$. This measure to quantify collinearity may seem to be chosen arbitrarily, but as will be seen later it shows itself naturally in the analysis of certain cost functions for NOJD. Another measure, which also naturally appears in the analysis of the log-likelihood based cost (see Section 3.2.3) function is given as:

[2] In some works this is referred to as essential uniqueness.

DEFINITION 2.2. *For the set of positive definite diagonal matrices* $\{\Lambda_i\}_{i=1}^N$ *let:*

$$\mu_{kl} = \frac{1}{N^2}(\sum_{i=1}^{N} \frac{\lambda_{ik}}{\lambda_{il}})(\sum_{i=1}^{N} \frac{\lambda_{il}}{\lambda_{ik}}), \quad 1 \leq k \neq l \leq N. \tag{2.3}$$

Let μ *be the minimum value of* μ_{kl}*'s. The Modulus of Uniqueness of second type for this set is defined as* μ*.*

Note that $\mu \geq 1$ with equality if and only if $|\rho| = 1$. μ also measures the maximum collinearity between the columns of Λ, with the assumption that Λ_i's are positive definite.

If $N = 1$, then $|\rho| = 1$ and the diagonalizer is not unique. For $N > 1$, also the modulus of uniqueness captures the uniqueness property in an exact sense:

THEOREM 2.3. *Let* C_i*'s satisfy (2.1). The necessary and sufficient condition to have unique non-orthogonal joint diagonalizer is that* $|\rho| < 1$*.*

Proof. First we consider the case $n = 2$. If $|\rho| = 1$, then either: (a) there is a real number K such that: $\lambda_{i2} = K\lambda_{i1}$ for all $1 \leq i \leq N$, or (b) $\lambda_{i1} = 0$ for all $1 \leq i \leq N$ and $\lambda_{i2} \neq 0$ for some i, or (c) $\lambda_{i1} = \lambda_{i2} = 0$ for all i, which is a trivial situation. In case of (a) we have:

$$C_i = \lambda_{i1} A \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{|K|} \end{bmatrix}}_{D_K} \begin{bmatrix} 1 & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{|K|} \end{bmatrix} A^T. \tag{2.4}$$

We have denoted the diagonal matrix that includes $\sqrt{|K|}$ as D_K. Let us first assume that $K \neq 0$. Now, if $\rho = +1$, then let $B = Q_{+1} D_K^{-1} A^{-1}$, where:

$$Q_{+1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{2.5}$$

This B diagonalizes every C_i for all θ. If $\rho = -1$, then let $B = Q_{-1} D_K^{-1} A^{-1}$, where:

$$Q_{-1} = \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix}. \tag{2.6}$$

This B diagonalizes every C_i for all θ. If $K = 0$ (and hence $\rho = 1$), then let $B = Q_1 A^{-1}$ where:

$$Q_1 = \begin{bmatrix} 1 & \theta \\ 0 & 1 \end{bmatrix}. \tag{2.7}$$

This B diagonalizes every C_i for all θ. Also, in case of (b), $B = Q_1^T A^{-1}$ diagonalizes every C_i for all θ. Therefore for $|\rho| = 1$ the non-orthogonal joint diagonalizer is not unique. For $n > 2$, $|\rho| = 1$ means that the situation described for $n = 2$ happens between two diagonal elements, in the same positions within Λ_i's; and we can apply the previous argument to those elements. Hence, for $n > 2$ also $|\rho| = 1$ implies non-uniqueness of the joint diagonalizer. To see the necessary part, first note that existence of more than one exact joint diagonalizer means that there exists a C which differs from a permuted diagonal matrix such that the matrices $D_i = C\Lambda_i C^T$ are diagonal. For the moment, assume that one of the Λ_i's, say Λ_1, is non-singular. Then $D_i D_1^{-1} = C\Lambda_i\Lambda_1^{-1}C^{-1}$ for $1 < i \leq N$. These are the eigen decompositions of diagonal

matrices $\Lambda_i\Lambda_1^{-1}$'s, for $1 < i \leq N$. Non-uniqueness of C happens only when there are two integers $1 \leq k \neq l \leq n$, such that $\frac{\lambda_{ik}}{\lambda_{1k}} = \frac{\lambda_{il}}{\lambda_{1l}}$, $1 < i \leq N$. This means that $|\rho| = 1$. If C is not unique and all Λ_i's are singular, then two cases can happen. In the first case, all Λ_i's have one zero diagonal element at a common position, i.e., there exists an integer $1 \leq k \leq n$ such that for all $1 \leq i \leq N$ we have $\lambda_{ik} = 0$, which implies that $\rho = 1$. If the first case is not true, then there exists a linear combination of Λ_i's like Λ_0, which is non-singular, and $D_0 = C\Lambda_0 C^T$ is diagonal; then we are back to the non-singular case. This completes the proof. □

This result and more general ones have been referred to in [23] using the concept of Kruskal's rank.

2.1. On minimum number of matrices needed for EJD. Let A be an orthogonal matrix. Then equations in (2.1) are the eigen decompositions of C_i's. If C_1 or equivalently Λ_1 has distinct eigenvalues, then A can be found from eigen decomposition of C_1, uniquely up to permutations. If Λ_1 has only two equal diagonal elements at positions k and l, and if we can find another Λ_i with distinct values at those positions, then again A can be found uniquely, from eigen decompositions of C_1 and C_i. Therefore, if for each pair of k and l we can find an i for which $\lambda_{il} \neq \lambda_{ik}$, then A can be determined uniquely. As a result, in the generic case orthogonal joint diagonalization is in fact a one-matrix problem, and inclusion of more matrices can be justified by presence of noise. The uniqueness properties of OJD as well as its sensitivity analysis has been addressed in [8].

There is a huge difference between the uniqueness properties of the orthogonal and non-orthogonal joint diagonalization problems. From the proof of Theorem (2.3) it should be evident that $N = 1$ matrix is not enough to find a unique non-orthogonal (joint) diagonalizer. NOJD allows more degrees of freedom in finding the diagonalizer. Let us count the degrees of freedom in both sides of the equations in (2.1). Recall that a symmetric $n \times n$ matrix has $\frac{n(n+1)}{2}$ degrees of freedom, and A has $n^2 - n$ degrees of freedom (as far as the NOJD problem is concerned). Hence, the left hand side of (2.1) has total $N\frac{n(n+1)}{2}$ degrees of freedom, and its right hand side has $n^2 - n + Nn$ degrees of freedom. Equating the degrees of freedom from both sides and solving for N gives $N = 2$. Therefore, the minimum number of matrices to give enough equations to find a unique non-orthogonal joint diagonalizer is $N = 2$; and hence, the NOJD problem, in the generic case, is a two-matrix problem. For two arbitrary and generic matrices $\{C_1, C_2\}$ whether the equations in (2.1) yield a real valued solution for $\{A, \Lambda_1, \Lambda_2\}$ depends on the matrices[3]. It is well known that if one of the two matrices is positive definite, then they admit a (real) exact jointly diagonalizer [13, pp. 461-462]. As we will show in Section 4.4.1, for only two matrices, if their dimension is moderately large ($n > 20$, for example), the modulus of uniqueness is close to unity. This in turn, as will be shown in Section 4, means that the NOJD of the two matrices is an ill-conditioned problem; and hence, it is better to include more matrices in the NOJD process.

2.2. Identifiability of the ICA problem. Now we would like to apply the previous theorem to the case of the ICA problem. It is obvious that if we can find two cumulant matrix slices of $\vec{\mathbf{x}}_{n\times 1}$ for which $|\rho|$ is not unity, then the matrix A in (1.1) can be found uniquely. From (1.3) and (1.4), one can show that for the set

[3] Assuming C_1 is invertible (which is true for a generic matrix), in order for (2.1) to hold, we should have that $C_2 C_1^{-1} = A\Lambda_2\Lambda_1^{-1}A^{-1}$, which is an eigen decomposition. Again in a generic case, this would give a unique and (in general) complex valued $\{A, \Lambda_1, \Lambda_2\}$.

$\{\mathcal{C}^k_{\mathbf{x}}(i_1, i_2, ..., i_{k-2}, :, :)\}_{1\leq i_1,...,i_{k-2}\leq n}$ with $k > 2$ we have $|\rho| \neq 1$, if and only if none of $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s are zero. To see this, first note that if $\mathcal{C}^k_{\mathbf{s}}(i, ..., i) = 0$ for some i then $|\rho| = 1$. Now assume that none of $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s are zero, and $|\rho| = 1$. Since $|\rho| = 1$, there are two columns of A like j and l and a real number K such that:

$$a_{i_1 j} a_{i_2 j} ... a_{i_{k-2} j} \mathcal{C}^k_{\mathbf{s}}(j, ..., j) = K a_{i_1 l} a_{i_2 l} ... a_{i_{k-2} l} \mathcal{C}^k_{\mathbf{s}}(l, ..., l), \tag{2.8}$$

for all $1 \leq i_1, ..., i_{k-2} \leq n$. Because none of the $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s are zero, and since there is at least one non-zero element like a_{pj} in the j^{th} column of A, by setting $i_2 = ... = i_{k-2} = p$ we have that there is another real number K' such that:

$$a_{i_1 j} = K' a_{i_1 l}, \tag{2.9}$$

for all $1 \leq i_1 \leq n$. This contradicts the invertibility of A. Hence, with an invertible A, $|\rho|$ can not be unity unless at least one of $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s is zero.

Now assume that the covariance matrix of $\vec{\mathbf{s}}_{n\times 1}$ is non-singular, i.e., there is no source component with zero variance. Then by inclusion of the covariance matrix of $\vec{\mathbf{x}}_{n\times 1}$ in the above set we can weaken the uniqueness condition, i.e., for $\{R_{\mathbf{xx}}, \mathcal{C}^k_{\mathbf{x}}(i_1, i_2, ..., i_{k-2}, :, :)\}_{1\leq i_1,...,i_{k-2}\leq n}$ with $k > 2$, we have $|\rho| \neq 1$ if and only if at most one of $\mathcal{C}^k_{\mathbf{s}}(i, ..., i)$'s is zero. Therefore, if we start with the covariance matrix of $\vec{\mathbf{x}}_{n\times 1}$ and then include its third order cumulant slices, and if at most one of the skewness' $\mathcal{C}^3_{\mathbf{s}}(i, ..., i)$ is zero, then A can be determined uniquely. If at least two $\mathcal{C}^3_{\mathbf{s}}(i, ..., i)$'s are zero then we can go to the cumulants of higher orders and check the same condition. Note that this process fails if and only if there are at least two source elements s_p and s_q for which $\mathcal{C}^k_{\mathbf{s}}(p, ..., p) = \mathcal{C}^k_{\mathbf{s}}(q, ..., q) = 0$ for all $k \geq 3$. It is well known that such random variables have Gaussian distribution. As a result, exact non-orthogonal joint diagonalization of the set of all cumulant matrix slices of $\vec{\mathbf{x}}_{n\times 1}$ gives A uniquely, unless $\vec{\mathbf{s}}_{n\times 1}$ has at least two Gaussian components. To summarize we state this theorem (cf. Corollary 13 in [10]):

THEOREM 2.4. *(Identifiability of ICA- EJD formulation) Consider the model (1.1). About $\vec{\mathbf{s}}_{n\times 1}$ assume that its covariance matrix is non-singular, its k^{th} order cumulants (for some $k > 2$) exist and at most one of them is zero. Then exact joint diagonalization of the set $\{R_{\mathbf{xx}}, \mathcal{C}^k_{\mathbf{x}}(i_1, i_2, ..., i_{k-2}, :, :)\}_{1\leq i_1,...,i_{k-2}\leq n}$ results in finding A up to column permutation and scaling. For a source vector with finite cumulants of all orders, this process fails to identify A if only if the source vector has more than one Gaussian component.*

This result suggests that exact joint diagonalization can be used as a basis to define a contrast function [10] for ICA. Note that, this identifiability condition is derived solely based on the algebraic structure of the ICA model, and we have not used the Skitovich-Darmois Theorem [10, 16]. OJD or NOJD of cumulant matrix slices of order three, four or even higher have been suggested in many works, e.g. [9, 28, 29, 19, 17, 2]. The OJD scenario arises when one assumes that the mixture is already uncorrelated or whitened.

3. Cost functions for joint diagonalization. The joint diagonalization problem has been posed, in the literature, mostly as an optimization problem [9, 28, 25, 20]. We mention that in [28, 25] the joint diagonalization problem has been addressed with a different formulation than ours. As mentioned before, generically, in the OJD problem one matrix ($N = 1$), and in the NOJD problem two matrices ($N = 2$) are enough to find a unique joint diagonalizer. However, it is believed that inclusion of more

matrices is useful in making the solution less vulnerable to noise. Therefore, the proposed cost functions for joint diagonalization are designed to mitigate the effect of noise via averaging.

3.1. A cost function for orthogonal joint diagonalization. The OJD problem was introduced earlier than the NOJD problem. In [9] a natural cost function together with an efficient algorithm for OJD was introduced. The cost function $J_1 : \mathrm{SO}(n) \to \mathbb{R}$ for OJD, introduced in [9], is:

$$J_1(\Theta) = \sum_{i=1}^{n} \left\| \Theta C_i \Theta^T - \mathrm{diag}(\Theta C_i \Theta^T) \right\|_F^2, \tag{3.1}$$

where $\{C_i\}_{i=1}^N$ is the set of symmetric matrices to be diagonalized. If Θ minimizes J_1, then we call Θ an orthogonal joint diagonalizer of $\{C_i\}_{i=1}^N$. Note that, since $\mathrm{SO}(n)$ is a compact manifold, a priori we know that a minimizer exists for J_1. Whether, generically, this cost function has only global minimum on $\mathrm{SO}(n)$, and whether the minimizers are unique up to permutation are not known.

3.2. Cost functions for non-orthogonal joint diagonalization. Introducing a cost function for NOJD has been a challenge. First, note that a simple extension of J_1 from $\mathrm{SO}(n)$ to $\mathrm{GL}(n)$ is not effective. We remind that the NOJD problem in the exact case is a scale-invariant problem, i.e., if $B \in \mathrm{GL}(n)$ is an EJD for a set of matrices, then DB also should be a joint diagonalizer for any non-singular diagonal D. However, $J_1(DB) \neq J_1(B)$. In fact, we can reduce $J_1(B)$ just by reducing the norm of B; and $J_1(B)$ has a global infimum at $B = 0$.

3.2.1. A non-holonomic flow for NOJD based on J_1. For the derivations in this subsection we refer the reader to [2]. We also refer the reader to [15] for more comprehensive treatment of gradient flows for optimization on manifolds. On the Lie group of non-singular matrices, we can define a right-invariant Riemannian metric[4]that matches the group structure as:

$$\begin{aligned} \langle ., . \rangle_B : T_B GL(n) \times T_B GL(n) \to \mathbb{R} \\ \langle \xi_1, \xi_2 \rangle_B = \mathrm{tr}((\xi_1 B^{-1})^T \xi_2 B^{-1}), \end{aligned} \tag{3.2}$$

where $T_B\mathrm{GL}(n)$ is the tangent space to $\mathrm{GL}(n)$ at B. In general, a tangent vector ξ at a point B on $\mathrm{GL}(n)$ (and any Lie group) can be written as $\xi = \zeta B$ where ζ belongs to the tangent space at the identity. Also, the tangent spaces at B and the identity are isomorphic. Let $s \mapsto \tilde{B}(s)$ be any smooth curve with $\tilde{B}(0) = B$. With respect to the Riemannian metric in (3.2), the gradient of $J_1 : \mathrm{GL}(n) \to \mathbb{R}$ is defined as a vector field ∇J_1 that satisfies:

$$\dot{J}_1 = \langle \nabla J_1, \dot{B} \rangle_B, \tag{3.3}$$

where $\dot{J}_1 = \frac{dJ_1(\tilde{B}(s))}{ds}\big|_{s=0}$ and $\dot{B} = \frac{d\tilde{B}(s)}{ds}\big|_{s=0}$. From this, it is easy to verify that up to a scaler factor:

$$\nabla J_1(B) = \Omega_1 B, \tag{3.4}$$

[4]The significance of the right-invariant metric is that it matches the invariance property of the NOJD problem, which as mentioned is that the joint diagonalizer does not change by left multiplication by non-singular diagonal matrices. A discretization of a right-invariant flow such as $\frac{dB}{ds} = \Omega B$ has the form $B_{k+1} = (I + \Omega_k) B_k$.

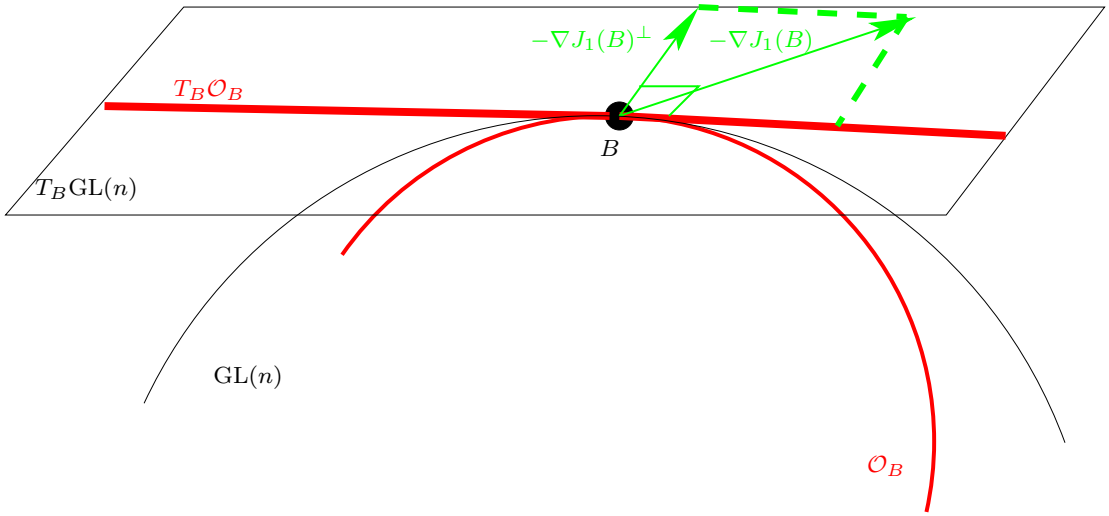

FIG. 3.1. *The group of non-singular diagonal matrices acts on the manifold $GL(n)$ at B via left multiplication. $\mathcal{O}_B$ is the orbit of this action. The liberalization of this orbit (i.e., the tangent space to it) at B is $T_B\mathcal{O}_B \subset T_B GL(n)$. This figure shows how $-\nabla J_1(B)$ should be projected onto the orthogonal complement of $T_B\mathcal{O}_B$ in order to have a flow for NOJD based on J_1 which is not a scale-invariant cost function for NOJD.*

where:

$$\Omega_1 = \sum_{i=1}^{N} (BC_iB^T)^{\circ} BC_iB^T. \tag{3.5}$$

We can show that the stationary points of J_1, i.e., values of B for which $\nabla J_1(B) = 0$ and hence $\Omega_1 = 0$ satisfy, $BC_iB^T = \mathrm{diag}(BC_iB^T)$. Therefore, if C_i's do not have an exact joint diagonalizer, then J_1 will have no stationary points on $\mathrm{GL}(n)$. A gradient flow for minimization of J_1 has the form $\frac{dB}{ds} = -\nabla J_1(B) = -\Omega_1 B$. As we mentioned before, the problem with minimizing J_1 as a cost function for NOJD is that it can be reduced by diagonal matrices. At each point $B \in \mathrm{GL}(n)$, we can project the gradient of J_1 (or more accurately the negative of the gradient) to directions that do not correspond to diagonal scaling. The group of non-singular diagonal matrices of dimension n can act on the group $\mathrm{GL}(n)$ via left multiplication. At B the orbit of this action is simply: $\mathcal{O}_B = \{DB | D = \text{non-signgular and diagonal}\}$ and it is in fact a sub-manifold which we would like our NOJD flow to avoid. The linearization or the tangent space to the orbit at B is $T_B\mathcal{O}_B = \{DB | D = \text{diagonal}\}$, which is a linear subspace of $T_B\mathrm{GL}(n)$. The orthogonal complement of $T_B\mathcal{O}_B$ in the tangent space $T_B\mathrm{GL}(n)$, with respect to the defined Riemannian metric, is $(T_B\mathcal{O}_B)^{\perp} = \{\Xi B | \Xi \in \mathbb{R}^{n\times n},\ \mathrm{diag}(\Xi) = 0\}$. Therefore, the projection of ∇J_1 onto $(T_B\mathcal{O}_B)^{\perp}$ is $\nabla J_1^{\perp} = \Omega_1^{\circ} B$. Figure (3.1) shows the process of constraining the negative of the gradient to directions along $(T_B\mathcal{O}_B)^{\perp}$, at each point B. The corresponding non-holonomic[5] flow for NOJD is:

$$\frac{dB}{ds} = -\nabla J_1^{\perp} = -\Omega_1^{\circ} B. \tag{3.6}$$

[5]A Non-holonomic flow is a flow whose velocity vector field is constrained by non-integrable constraints.

The stationary points or equilibria of this flow are defined by $\Omega_1^\circ = 0$, or

$$\sum_{i=1}^{N} \big((BC_iB^T)^\circ BC_iB^T\big)^\circ = 0. \tag{3.7}$$

Hence, if a non-orthogonal joint diagonalizer of $\{C_i\}_{i=1}^N$ based on the above non-holonomic flow exists, it should satisfy (3.7). In [3, 29] and many other works minimization schemes for J_1 are proposed, which try to find the stationary points in (3.7).

3.2.2. A Frobenius norm scale-invariant cost function. Note that $J_1 : \mathrm{SO}(n) \to \mathbb{R}$ in (3.1) can also be written as:

$$J_1(\Theta) = \sum_{i=1}^{n} \big\|C_i - \Theta^{-1}\mathrm{diag}(\Theta C_i \Theta^T)\Theta^{-T}\big\|_F^2. \tag{3.8}$$

Let $J_2 : \mathrm{GL(n)} \to \mathbb{R}$ be the extension of this form of J_1 to $\mathrm{GL}(n)$ defined by:

$$J_2(B) = \sum_{i=1}^{n} \big\|C_i - B^{-1}\mathrm{diag}(BC_iB^T)B^{-T}\big\|_F^2. \tag{3.9}$$

Then it is easy to check that $J_2(\Pi DB) = J_2(B)$ for any non-singular diagonal D and permutation Π. Therefore, J_2 is a scale and permutation invariant cost function for NOJD. Note that J_1 and J_2 are scaled versions of each other in the sense that:

$$\frac{J_1(B)}{n^2\|B\|_2^4} \le J_2(B) \le n^2\|B^{-1}\|_2^4 J_1(B). \tag{3.10}$$

Also, note that we can reduce J_2 without changing norm of B. This means that reducing J_2, if norm of B is not changed can, result in reduction of the upper bound of J_1. This cost function has been introduced in [5, 3].

3.2.3. Log-likelihood function for NOJD. In [20], another cost function for NOJD of a set of positive definite matrices $\{C_i\}_{i=1}^N$ has been introduced. This cost function has the form:

$$J_3(B) = \sum_{i=1}^{N} \log\big(\frac{\det \mathrm{diag}(BC_iB^T)}{\det BC_iB^T}\big). \tag{3.11}$$

A matrix $B \in \mathrm{GL}(n)$ that minimizes J_3 is the joint diagonalizer of $\{C_i\}_{i=1}^N$. It can be shown that $J_3(B) \ge 0$, and equality holds if and only if all BC_iB^T's are diagonal. It is easy to check that J_3 is scale and permutation invariant, i.e., $J_3(\Pi DB) = J_3(B)$, for any Π and D. The specific form of this cost function is imposed by the log-likelihood function of correlation matrices of Gaussian non-stationary sources [20]. Let us consider the same right-invariant Riemannian metric as in 3.2.1. Using the well known identity $\frac{\partial}{\partial B}\log\det B = (B^T)^{-1}$ we can show (see (3.3) also):

$$\dot{J}_3 = 2\sum_{i=1}^{N} \mathrm{tr}\Big((\dot{B}B^{-1})^T\big((\mathrm{diag}(BC_iB^T))^{-1}BC_iB^T - I\big)\Big). \tag{3.12}$$

As a result, with respect to the above Riemannian metric, the gradient vector field of J_3 up to a scalar factor is:

$$\nabla J_3(B) = \frac{1}{N}\sum_{i=1}^{N}\big(\mathrm{diag}((BC_iB^T))^{-1}BC_iB^T - I\big)B := \Omega_3 B. \tag{3.13}$$

It is interesting to note that $\mathrm{diag}(\Omega_3) = 0$ (cf. (3.5) and (3.6)). A gradient flow for NOJD based on minimization of J_3 is: $\frac{dB}{ds} = -\Omega_3 B$. The stationary points for this flow are characterized by $\Omega_3 = 0$; and if B is a joint diagonalizer it should satisfy

$$\frac{1}{N}\sum_{i=1}^{N} BC_iB^T\big(\mathrm{diag}(BC_iB^T)\big)^{-1} = I. \tag{3.14}$$

4. Sensitivity analysis. An interesting question to ask is: "which set of matrices are hard to be jointly diagonalized?" In other words, which factors affect the condition or sensitivity of the joint diagonalization problem? Consider the matrices $C_i = A\Lambda_iA^T,\ 1 \le i \le N$, where Λ_i's are diagonal. Obviously, $\{C_i\}_{i=1}^N$ have a joint diagonalizer $B = A^{-1}$. Note that, here equality is understood up to permutation and diagonal scaling. Now, we add noise to the matrices as:

$$C_i = A\Lambda_iA^T + tN_i, \quad t \in [-\delta, \delta], \delta > 0, \tag{4.1}$$

where $\{N_i\}_{i=1}^N$ are symmetric error or noise matrices, and t shows the noise gain or contribution. The joint diagonalizer of this noisy set will deviate from A^{-1} as t deviates from zero. If the sensitivity is high, then the deviation from A^{-1} will be large. In this case, we say that the NOJD problem is very sensitive or ill-conditioned. Note that the true goal of NOJD is to find A and not just diagonalizing the matrices $\{C_i\}_{i=1}^N$. It is in this context that the sensitivity of the problem is defined. If the modulus of uniqueness for $\{\Lambda_i\}_{i=1}^N$ is unity, then $\{C_i\}_{i=1}^N$ has already infinite sensitivity; since the joint diagonalizer can change even in absence of noise. Hence, one should expect the sensitivity for joint diagonalization to be closely related to the issue of uniqueness. To quantify this relation, we will perform a perturbation analysis of the stationary points of the NOJD cost functions or flows defined in Section 3. The results for different cost functions are very much similar.

4.1. Sensitivity for NOJD based on J_1. The non-orthogonal joint diagonalizer of $\{C_i\}_{i=1}^N$, B based on J_1 is defined by equation (3.7). As t deviates from zero in (4.1), $B(t)$, the joint diagonalizer, varies smoothly. For small enough δ, from the implicit function theorem and basic properties of Lie groups, we have:

$$B(t) = (I + t\Delta)A^{-1} + o(t), \quad t \in [-\delta, \delta], \tag{4.2}$$

where $\Delta \in \mathbb{R}^{n\times n}$ with $\mathrm{diag}(\Delta) = 0$ and $\frac{\|o(t)\|}{t} \to 0$ as $t \to 0$. The restriction $\mathrm{diag}(\Delta) = 0$ matches the structure the non-holonomic flow for NOJD derived in Section 3.2.1. Note that $\|\Delta\|$ measures the sensitivity of the NOJD problem to noise. Our goal is to calculate Δ. Using $B(0) = A^{-1}$ and $\frac{dB}{dt}(0) = \Delta A^{-1}$, and after plugging (4.1) into (3.7), and then differentiating with respect to t, we can easily verify that:

$$\sum_{i=1}^{N}\big(\Delta\Lambda_i + \Lambda_i\Delta^T\big)\Lambda_i = -\sum_{i=1}^{N}\big(A^{-1}N_i(A^{-1})^T\big)^{\circ}\Lambda_i. \tag{4.3}$$

The right hand side of above equation manifests the possible noise amplification that can happen due to large $\|A^{-1}\|$, i.e., when A is small in norm or more importantly when A is ill-conditioned. Equation (4.3) is a linear equation in terms of Δ. Let us define:

$$\mathcal{T} = \sum_{i=1}^{N} (A^{-1} N_i (A^{-1})^T)^{\circ} \Lambda_i. \tag{4.4}$$

Now, it is easy to check that the two entries Δ_{kl} and Δ_{lk} decouple from the rest of the entries of Δ and we have:

$$\begin{bmatrix} \sum_{i=1}^N \lambda_{il}^2 & \sum_{i=1}^N \lambda_{ik}\lambda_{il} \\ \sum_{i=1}^N \lambda_{ik}\lambda_{il} & \sum_{i=1}^N \lambda_{ik}^2 \end{bmatrix} \begin{bmatrix} \Delta_{kl} \\ \Delta_{lk} \end{bmatrix} = - \begin{bmatrix} \mathcal{T}_{kl} \\ \mathcal{T}_{lk} \end{bmatrix}, \ 1 \leq k < l \leq n. \tag{4.5}$$

Recall definition of ρ_{kl} (Definition 2.1). Also, let

$$\gamma_{kl} = (\sum_{i=1}^{N} \lambda_{ik}^2)^{\frac{1}{2}} (\sum_{i=1}^{N} \lambda_{il}^2)^{\frac{1}{2}}, \quad \eta_{kl} = \frac{(\sum_{i=1}^N \lambda_{ik}^2)^{\frac{1}{2}}}{(\sum_{i=1}^N \lambda_{il}^2)^{\frac{1}{2}}}. \tag{4.6}$$

We denote the coefficients matrix in (4.5) by M_{kl}:

$$M_{kl} = \gamma_{kl} \begin{bmatrix} \eta_{kl}^{-1} & \rho_{kl} \\ \rho_{kl} & \eta_{kl} \end{bmatrix}, \qquad 1 \leq k < l \leq n. \tag{4.7}$$

Then Equation (4.5) is equivalent to:

$$\begin{bmatrix} \Delta_{kl} \\ \Delta_{lk} \end{bmatrix} = -M_{kl}^{-1} \begin{bmatrix} \mathcal{T}_{kl} \\ \mathcal{T}_{lk} \end{bmatrix} = \frac{-1}{\gamma_{kl}(1-\rho_{kl}^2)} \begin{bmatrix} \eta_{kl} & -\rho_{kl} \\ -\rho_{kl} & \eta_{kl}^{-1} \end{bmatrix} \begin{bmatrix} \mathcal{T}_{kl} \\ \mathcal{T}_{lk} \end{bmatrix}, \ \ 1 \leq k < l \leq n. \tag{4.8}$$

Note that $\lambda_{max} \geq \lambda_{min}$, the eigenvalues of M_{kl}^{-1}, are

$$\lambda_{max}, \lambda_{max} = \frac{\eta_{kl} + \eta_{kl}^{-1} \pm \sqrt{(\eta_{kl} + \eta_{kl}^{-1})^2 - 4(1-\rho_{kl}^2)}}{2\gamma_{kl}(1-\rho_{kl}^2)}. \tag{4.9}$$

Also, it is easy to check that:

$$\frac{\eta_{kl} + \eta_{kl}^{-1} - 1}{\gamma_{kl}(1-\rho_{kl}^2)} \leq \lambda_{max} < \frac{\eta_{kl} + \eta_{kl}^{-1}}{\gamma_{kl}(1-\rho_{kl}^2)}, \tag{4.10}$$

and

$$\frac{1}{\gamma_{kl}(\eta_{kl} + \eta_{kl}^{-1})} \leq \lambda_{min} \leq \frac{1}{\gamma_{kl}}. \tag{4.11}$$

Therefore, we also can establish the bounds:

$$\frac{1}{\gamma_{kl}(\eta_{kl} + \eta_{kl}^{-1})} \| \begin{bmatrix} \Delta_{kl} \\ \Delta_{lk} \end{bmatrix} \| \leq \| \begin{bmatrix} \Delta_{kl} \\ \Delta_{lk} \end{bmatrix} \| < \frac{\eta_{kl} + \eta_{kl}^{-1}}{\gamma_{kl}(1-\rho_{kl}^2)} \| \begin{bmatrix} \mathcal{T}_{kl} \\ \mathcal{T}_{lk} \end{bmatrix} \|. \tag{4.12}$$

Because γ_{kl} is not scale-invariant, $\gamma_{kl} \approx 0$ by itself does not imply a high sensitivity. The definitions of the parameters reveal that γ_{kl} plays more a scaling role, whereas ρ_{kl} plays a structural role. Hence, as far as sensitivity to noise is concerned, the

interesting situation (approximate singularity) happens when $|\rho_{kl}| \approx 1$. Note that as $|\rho_{kl}| \to 1$, λ_{max} and λ_{min} approach their upper and lower bounds, respectively. Moreover, in that case λ_{max} grows unboundedly, and λ_{min} remains bounded. Since $\mathcal{T}$ depends on random noise, there always will be a component of $\begin{bmatrix} \mathcal{T}_{kl} \\ \mathcal{T}_{lk} \end{bmatrix}$ along the direction of the eigenvector of M_{kl}^{-1} corresponding to λ_{max}. Therefore, when $|\rho_{kl}|$ approaches unity, $\| \begin{bmatrix} \Delta_{kl} \\ \Delta_{lk} \end{bmatrix} \|$ tends towards the upper bound in (4.12). Hence, the upper bound is the more interesting one, and it is not a loose bound in the sense that it can be achieved very closely when $|\rho_{kl}| \approx 1$[6]. One can easily check that:

$$\|\Delta\|_F < \frac{\alpha}{(1-\rho^2)}\|\mathcal{T}\|_F \leq \frac{n\alpha\|A^{-1}\|_2^2}{(1-\rho^2)} \sum_{i=1}^{N} \|N_i\|_2 \|\Lambda_i\|_2, \tag{4.13}$$

where $\alpha = \max_{k\neq l} \frac{\eta_{kl}+\frac{1}{\eta_{kl}}}{\gamma_{kl}}$, and $|\rho|$ is the modulus of uniqueness for the set $\{\Lambda_i\}_{i=1}^N$ as defined before. Since an approximate non-uniqueness of the joint diagonalizer can happen when only one of $|\rho_{kl}|$'s is close to unity, the above bound might seem exaggerative. Again one can imagine a worse case scenario in which all $|\rho_{kl}|$'s are close to unity, and the bound would not be very loose. In summary, we have:

THEOREM 4.1. *Let $C_i = A\Lambda_i A^T + tN_i, 1 \leq i \leq N$ ($t \in [-\delta, \delta]$). Let us define $B(t)$ the non-orthogonal joint diagonalizer for $\{C_i\}_{i=1}^N$ as the minimizer of J_1 under the non-holonomic flow with equilibria defined in (3.7). Then for small enough δ, the joint diagonalizer can be written as: $B(t) = (I + t\Delta)A^{-1} + o(t)$, where Δ (with diag$(\Delta) = 0$) satisfies (4.8) as well as (4.13).*

The bound in (4.13) confirms the intuition that if the joint diagonalizer is close to non-uniqueness, as measured by the modulus of uniqueness, or if it is ill-conditioned, then the sensitivity of the NOJD problem will be high. Note that our derivations suggest that there is another scenario that can result in high sensitivity, which is when A and Λ_i (i.e., γ_{kl}'s) are small in norm. Of course, this is not an interesting scenario. We could have avoided this by imposing a constraint on the norm of the noise in (4.1). For example, we could assume that $\|N_i\|_2 \leq \|A\|_2^2\|\Lambda_i\|_2$. This choice makes the bound (4.13) such that it is unchanged if A or Λ_i's all are scaled by a scaler. Hence, one might be tempted to define $\frac{\text{cond}(A)^2}{1-\rho^2}$ as the condition number for the NOJD problem based on J_1.

4.2. Sensitivity for NOJD based on J_2. We can follow the same path as in the previous subsection and perform a perturbation analysis for the stationary points of J_2 in presence of noise. In [5], it is shown that the stationary points of J_2 satisfy:

$$\begin{cases} \sum_{i=1}^N \big(\Psi_i \text{diag}\big(BC_iB^T\big) - \text{diag}\big(\Psi_i\big)BC_iB^T\big) = 0 \\ \Psi_i = (BB^T)^{-1}\big(BC_iB^T - \text{diag}(BC_iB^T)\big)(BB^T)^{-1} \end{cases} \tag{4.14}$$

$B(t)$, the minimizer of J_2 with C_i's defined in (4.1, can be written as (4.2). Our goal is to find Δ, when $B(t)$ satisfies (4.14). Similar to previous subsection, we can show that Δ satisfies:

$$\sum_{i=1}^{N} \Big(A^T A(\Delta\Lambda_i + \Lambda_i\Delta^T)A^T A\Lambda_i \Big)^\circ = -\sum_{i=1}^{N} \Big(A^T A\big(A^{-1}N_i(A^{-1})^T\big)^\circ A^T A\Lambda_i \Big)^\circ. \tag{4.15}$$

[6]It is common in matrix perturbation theory to have bounds that only handle the worse cases well. For a discussion on this issue see [22, p. 124].

Presence of the terms A^TA in (4.15) makes the decoupling that we saw in (4.3) not possible here. Note that in (4.3) the effect of A and Λ_i are separated very much, but in (4.15) this is not the case. This is because J_2 is not congruence preserving, i.e., it is not expressed in terms of only the BC_iB^T's. Note that if A is close to a diagonal multiple of an orthogonal matrix, i.e., if $A \approx QD$, where D is a non-singular diagonal matrix and Q is orthogonal, then $A^TA \approx D^TD$; and as a result, we have that (4.15) reduces to

$$\sum_{i=1}^{N} \big(\Delta\Lambda_i + \Lambda_i\Delta^T\big)\Lambda_i \approx -\sum_{i=1}^{N} \big(A^{-1}N_i(A^{-1})^T\big)^{\circ}\Lambda_i, \tag{4.16}$$

which is the approximated version of (4.3). This case is of practical interest. Many algorithms for NOJD try to iteratively reduce the data matrices C_i's, by congruence transforms, to diagonal matrices, i.e., $C_i \leftarrow B_kC_iB_k^T$, where B_k is the local joint diagonalizer found at step k. After a number of iterations, and when the matrices under transformation become close to diagonal we have $C_i = BA\Lambda_i(BA)^T + tBN_iB^T$, where B is the product of the local joint diagonalizers. In this case, the new A ($A \leftarrow BA$) is close to diagonal, so is A^TA. Also, in another case if one of the C_i's, say C_1, is positive definite, then we can apply the transformation $C_i \leftarrow C_1^{-\frac{1}{2}}C_i(C_1^{-\frac{1}{2}})^T = C_1^{-\frac{1}{2}}A\Lambda_i(C_1^{-\frac{1}{2}}A)^T + tC_1^{-\frac{1}{2}}N_i(C_1^{-\frac{1}{2}})^T$, where $C_1^{\frac{1}{2}}$ is a square root of C_1, i.e., $C_1^{\frac{1}{2}}(C_1^{\frac{1}{2}})^T = C_1$. Again, we can show that if the noise is not too strong, for the new A ($A \leftarrow C_1^{-1/2}A$), we have $A \approx QD$, for some orthogonal Q and non-singular diagonal D. This case can, for example, correspond to the so-called pre-whitening step in the ICA problem, where C_1 is a covariance matrix. We can maintain that the sensitivity properties of NOJD based on J_2 are very similar to those of NOJD based on J_1.

4.3. Sensitivity for NOJD based on J_3. A stationary point $B(t)$ of J_3, when C_i's are of the form (4.1) and Λ_i's are positive definite, satisfies equation (3.14). Similar to previous derivations, by differentiating (3.14) with respect to t and considering (4.1) and (4.2), we have that Δ, with $\mathrm{diag}(\Delta) = 0$, satisfies:

$$\sum_{i=1}^{N} \Delta + \Lambda_i\Delta^T\Lambda_i^{-1} = -\sum_{i=1}^{N} \big(A^{-1}N_i(A^{-1})^T\big)^{\circ}\Lambda_i^{-1}. \tag{4.17}$$

We have used the fact $\frac{d}{dt}X^{-1} = -X^{-1}(\frac{d}{dt}X)X^{-1}$, where X is a non-singular differentiable matrix function of t. Let us define:

$$\tau_{kl} = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_{ik}}{\lambda_{il}} \quad , \quad \mu_{kl} = \tau_{kl}\tau_{lk} = \frac{1}{N^2}(\sum_{i=1}^{N}\frac{\lambda_{ik}}{\lambda_{il}})(\sum_{i=1}^{N}\frac{\lambda_{il}}{\lambda_{ik}}). \tag{4.18}$$

Also let:

$$\mathcal{S} = \textstyle\sum_{i=1}^{N} \big(A^{-1}N_i(A^{-1})^T\big)^{\circ}\Lambda_i^{-1} \quad H_{kl} = N\begin{bmatrix} 1 & \tau_{kl} \\ \tau_{lk} & 1 \end{bmatrix}. \tag{4.19}$$

Here $\mathcal{S}$ is very similar to $\mathcal{T}$ and represents possible noise amplification due to small norm or ill-conditioning of A. The structure of $\mathcal{S}$ also shows that if Λ_i's are close to singularity, then noise amplification can happen. Note that the cost function J_3 requires Λ_i's to be positive definite, and as this condition is close to violation (by

one of Λ_i's being almost singular), then J_3 becomes very sensitive to noise. Equation (4.17) decouples as:

$$H_{kl}\begin{bmatrix}\Delta_{kl}\\ \Delta_{lk}\end{bmatrix} = -\begin{bmatrix}\mathcal{S}_{kl}\\ \mathcal{S}_{lk}\end{bmatrix},\ 1\le k<l\le n, \tag{4.20}$$

or equivalently:

$$\begin{bmatrix}\Delta_{kl}\\ \Delta_{lk}\end{bmatrix} = \frac{1}{N(\mu_{kl}-1)}\begin{bmatrix}1 & -\tau_{kl}\\ -\tau_{lk} & 1\end{bmatrix}\begin{bmatrix}\mathcal{S}_{kl}\\ \mathcal{S}_{lk}\end{bmatrix},\ 1\le k<l\le n. \tag{4.21}$$

It is easy to see that $\|H_{kl}^{-1}\|_F^2 = \frac{2+\tau_{kl}^2+\tau_{lk}^2}{\left(N(\mu_{kl}-1)\right)^2}$, and hence σ_{max}, the larger singular value of H_{kl}^{-1}, satisfies:

$$\frac{1}{\sqrt{2}}\frac{\sqrt{\tau_{kl}^2+\tau_{lk}^2+2}}{N(\mu_{kl}-1)} \le \sigma_{max} < \frac{\sqrt{\tau_{kl}^2+\tau_{lk}^2+2}}{N(\mu_{kl}-1)}. \tag{4.22}$$

From (4.21) and the previous bound we have:

$$\Big\|\begin{bmatrix}\Delta_{kl}\\ \Delta_{lk}\end{bmatrix}\Big\| \le \sigma_{max}\Big\|\begin{bmatrix}\mathcal{S}_{kl}\\ \mathcal{S}_{lk}\end{bmatrix}\Big\| < \frac{\sqrt{\tau_{kl}^2+\tau_{lk}^2+2}}{N(\mu_{kl}-1)}\Big\|\begin{bmatrix}\mathcal{S}_{kl}\\ \mathcal{S}_{lk}\end{bmatrix}\Big\|. \tag{4.23}$$

It is also easy to establish this bound:

$$\|\Delta\|_F < \frac{\beta}{N(\mu-1)}\|\mathcal{S}\|_F \le \frac{n\beta\|A^{-1}\|_2^2}{N(\mu-1)}\sum_{i=1}^{N}\|N_i\|_2\|\Lambda_i^{-1}\|_2, \tag{4.24}$$

where $\beta = \max_{k\neq l}\sqrt{\tau_{kl}^2+\tau_{lk}^2+2}$ and $\mu = \min_{k\neq l}\mu_{kl}$. In summary we have:

THEOREM 4.2. *Let $C_i = A\Lambda_i A^T + tN_i, 1\le i\le N$ ($t\in[-\delta,\delta]$) with Λ_i's positive definite. Let us define $B(t)$, the non-orthogonal joint diagonalizer for $\{C_i\}_{i=1}^N$, as the minimizer of J_3. Then for small enough δ the joint diagonalizer can be written as: $B(t) = (I+t\Delta)A^{-1}+o(t)$, where Δ (with $diag(\Delta)=0$) satisfies (4.20) as well as (4.24).*

Note that here, similar to the case of NOJD based on J_1, the modulus of uniqueness (μ) and the condition number of A affect the sensitivity. If one of Λ_i's is close to singularity i.e., if $\|\Lambda_i^{-1}\|$ is large, then the NOJD problem can be ill-conditioned. Therefore, almost similar to Section 4.1, we might impose the constraint $\|N_i\|_2 \le \|A\|_2^2\|\Lambda_i^{-1}\|_2^{-1}$ and define the condition number for the NOJD problem based on J_3 as $\frac{\mathrm{cond}(A)^2}{\mu-1}$. The imposed condition simply means that if $\|\Lambda_i^{-1}\|_2$ is large, then $\|N_i\|_2$ must be small or $\|A\|_2$ should be large.

4.4. Effect of the number of matrices. One of our motivations in performing sensitivity analysis for the problem of NOJD has been to consider the effect of the number of matrices on the accuracy of the solution. $N=2$ matrices are enough to find a unique non-orthogonal joint diagonalizer if $|\rho|<1$. However, to combat noise, we may want to include more matrices. Inclusion of more matrices can have two effects: one on how $\mathcal{T}$ in (4.13) or $\mathcal{S}$ in (4.24) changes and the other one on how ρ,γ and α in (4.13) or on how μ and α in (4.24) may change. The first effect is related to noise cancelation through averaging; and the second one is related to improvement of uniqueness measures. Both, of course, depend on how N_i's and Λ_i's are statistically

distributed. Let us consider a J_1-based-NOJD problem. Assume that the elements of N_i's are i.i.d with zero mean, and that the elements of Λ_i's are i.i.d with mean m and variance σ^2. Also, assume that matrices are independent from each other. Then, by the strong law of large numbers, we have that $\|\frac{\mathcal{T}}{N}\| \to 0$, $\rho \to \frac{m^2}{\sigma^2+m^2}$ and $N\alpha \to \frac{2}{\sigma^2+m^2} < \infty$ as $N \to \infty$ with probability one. Hence, $\|\Delta\| \to 0$ as $N \to \infty$ with probability one. Note that this might not happen if N_i's and Λ_i's are of non-zero mean. For small values of N such as $N = 2, 3$ or 4, and especially when n is large, $|\rho|$ can be very close to unity (for $N = 1$, $|\rho| = 1$). Moreover the cancelation or averaging effect that we expect to happen for large values of N in $\mathcal{T}$ is not likely to happen for small N. Hence, for small N the NOJD problem can be very sensitive.

4.4.1. More on the number of matrices and modulus of uniqueness. Our claim that for small N the modulus of uniqueness $|\rho|$ can be close to unity deserves more elaboration. From Definition 2.1, we can interpret ρ_{kl} as the cosine of the angle between two N dimensional vectors $(\lambda_{1k}, ..., \lambda_{Nk})^T$ and $(\lambda_{1l}, ..., \lambda_{Nl})^T$. Without loss of generality, we can assume that the vectors are of unit length, i.e., they represent points on the unit sphere in $\mathbb{R}^N$. Now $|\rho|$ is the maximum of the absolute value of the cosine of the angles between n points on the unit sphere in $\mathbb{R}^N$. Since $|\rho|$ is independent of the direction of the vectors, we can assume that all the points lie on the same hemisphere on the unit sphere in $\mathbb{R}^N$. The fact that $|\rho|$ can be large when n is much larger than N, is related to the fact that among n points on the unit hemisphere in $\mathbb{R}^N$ at least two of them can not be very far apart from each other. In other words, there are at least two of the points which are closer to each other than a deterministic distance, which depends on n and N. As n increases and the points become denser this deterministic distance decreases and $|\rho|$ approaches unity. Note that $|\rho|$ can become large if there is a large obtuse angle (an angle with negative cosine) between two of the points as well. However, we can only argue that as the number of points increases, they should become denser at some region, and hence, the upper bound on the minimum angular distance between them should decrease. We can not account for a lower bound on the maximum obtuse angles between the points unless we know a specific distribution for the points.

Unfortunately, finding the mentioned deterministic bound is difficult for arbitrary N. However, for $N = 2$ it is surprisingly easy to find. Assume that we have $n \geq 3$ points on the unit semicircle. Then we can divide the circumference of the semicircle into $n - 1$ arcs each of length $\frac{\pi}{n-1}$. Then by putting n points on the unit semicircle at least two of them will lie on the same arc. Hence, for $N = 2$ we should have $|\rho| \geq \cos\frac{\pi}{n-1}$, for $n \geq 3$. This implies that for two 20×20 matrices $|\rho| > 0.98$. Of course, for typical matrices this can be worse, since the bound we found is a lower bound. Again, note that this bound is solely based on an upper bound on the minimum of angular distances between the points. In fact, the configuration that achieves the bound $\frac{\pi}{n-1}$ has two antipodal points for which $\rho = -1$. Therefore, the bound is conservative. With more information about the points, we can find better lower bounds. For example, if the matrices are positive definite (i.e., $\lambda_{ik} > 0$ and hence $\rho > 0$), then we can consider points on a quarter of a circle, and hence have a lower bound of $\cos\frac{\pi}{2(n-1)}$ on ρ. As a result, for two positive definite matrices of dimension only $n = 10$ we have that $\rho > 0.98$. Maybe the most interesting finding of this paper is that, the NOJD of two matrices, if their dimension is fairly large, is ill-conditioned. Although, as explained before we might be able to find an exact non-orthogonal joint diagonalizer for the two matrices. Therefore, in general it is better to use more matrices not only to combat noise but also to improve the sensitivity.

The problem of finding an upper bound for the minimum distance between n points on the unit sphere in $\mathbb{R}^N$ is an old problem in the set of problems known as sphere packing problems. Tight bounds for these problems are in general very difficult to find. This specific problem is known as Tammes' problem or "dictators on a planet problem "[12]. One well-known result about it, is a bound for $N = 3$. According to this result [24], for $n \geq 3$ points on the unit sphere in $\mathbb{R}^3$, there are at least two points whose spherical (angular) distance is smaller than $d_n = \cos^{-1} \frac{\cot^2 \omega_n - 1}{2}$, where $\omega_n = \frac{n}{n-2} \frac{\pi}{6}$. Unfortunately, the proof for this result does not allow an extension to a parallel result for points on the hemisphere. However, we might argue, via homogeneity, that an approximate bound for n points on the hemisphere can be obtained by setting $2n$ points on the sphere and using d_{2n}. Hence, we have $\frac{\cot^2 \omega_{2n} - 1}{2}$ as an approximate lower bound for $|\rho|$ with $N = 3$. If we ignore the effect of the edge of the hemisphere, this scaling argument sounds quite plausible. Note that the scaling argument becomes more plausible for dense points. We expect the lower bound on $|\rho|$ to be smaller for $N = 3$ than that for $N = 2$, with equal n; and this is exactly what we observe. In Figure (4.1) we have plotted four curves. The lower two curves show the deterministic lower bounds on $|\rho|$ for $N = 2$ and $N = 3$ in terms of n. As can be seen, the bound for $N = 2$ is higher than the one for $N = 3$. The upper two curves show the average $|\rho|$ in terms of n, this time, for the uniform distribution of points on the circle and sphere. By uniform distribution on the sphere we mean that if $0 \leq \phi < 2\pi$ and $0 \leq \theta \leq \pi$ are the spherical coordinates of a point on the sphere, then these two random variables are uniformly distributed on their domains. Therefore, we generate n points (on the circle or the sphere), find the $|\rho|$ for them, and repeat this experiment $10,000$ times, and find the average $|\rho|$. As can be seen these values of $|\rho|$ are much higher than the bounds.

How about for other values of N? Let us pretend that we could extend the simple argument for the circle to higher dimensions. This helps us unveil the main dynamics between n and N in affecting $|\rho|$. Denote the surface area of the unit sphere in $\mathbb{R}^N$ by S_N. Assume we could divide the surface of the hemisphere into $n - 1$ congruent hyper-spherical regular polygons. This, of course, is a very difficult assumption to make. Let the angular diameter of each polygon be θ. If n is large, then we can approximate the area of the polygon by $V_{N-1}(\frac{\theta}{2})^{N-1}$, where V_{N-1} is the volume of the unit hyper-sphere in $\mathbb{R}^{N-1}$. Hence, we have $(n-1) \times V_{N-1}(\frac{\theta}{2})^{N-1} \approx \frac{S_N}{2}$, or $\theta \approx 2(\frac{1}{n-1})^{\frac{1}{N-1}}(\frac{S_N}{2V_{N-1}})^{\frac{1}{N-1}}$. One can see (for example from the explicit formulae for the surface and volume of the hyper-sphere in [27] and the formulae related to the Gamma function in [26]) that $(\frac{S_N}{2V_{N-1}})^{\frac{1}{N-1}}$ is of order $O(1)$ for large N; and in fact, it converges to 1. Here $O(.)$ is the big O notation. Now, $\theta \approx 2(\frac{1}{n})^{\frac{1}{N}}$ for large n and N, which in turn implies:

$$|\rho| \geq \cos(\theta) \approx 1 - 2(\frac{1}{n})^{\frac{2}{N}}. \tag{4.25}$$

This is in agreement, at least in its form, with a much more rigorous bound given in [12, p. 28 Equation (66)]. To be accurate, the result in [12] states that: if n is the maximum number of spherical caps of angular diameter $0 < \theta < 63^\circ$ that can be placed on the surface of the unit sphere in $\mathbb{R}^N$ without overlapping, then for large N

$$\cos\theta \gtrsim 1 - (\frac{1}{4})^{0.099}(\frac{1}{n})^{\frac{2}{N}} \approx 1 - 0.87(\frac{1}{n})^{\frac{2}{N}}. \tag{4.26}$$

We can replace n with $2n$ to have a similar (approximate) result for the hyper-

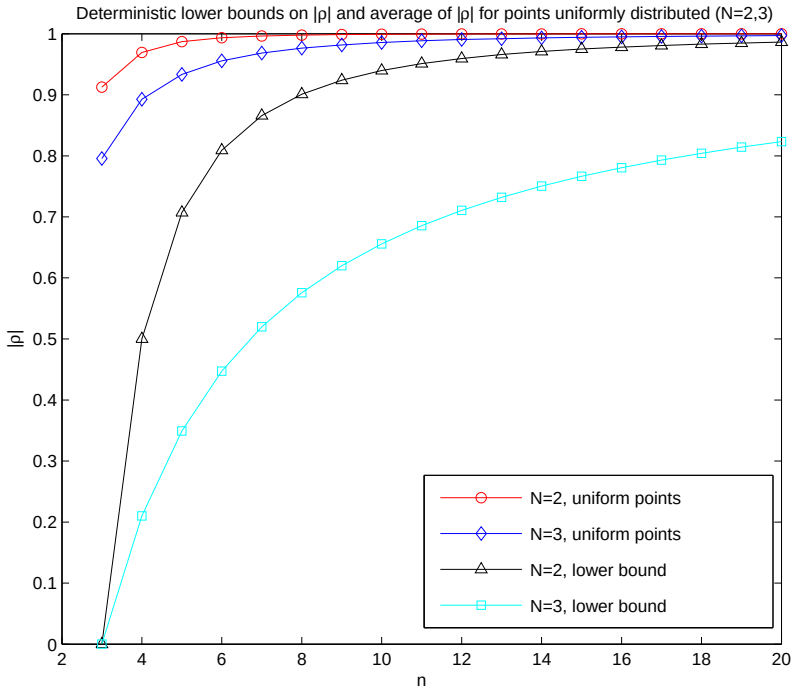

FIG. 4.1. *This graph shows two forms of variation of $|\rho|$ in terms of n for $N = 2$ and $N = 3$. The higher two curves show the average $|\rho|$ for points that are uniformly distributed on the circle and sphere. Here uniform means that the angular coordinates of the points in the spherical coordinate are uniformly distributed over the appropriate ranges. The lower two curves are the deterministic lower bounds described in the text. The deterministic bound for $N = 2$ is higher than the one for $N = 3$, as expected.*

hemisphere, which essentially does not change the asymptotic bound. Either bounds suggests that in order to control $|\rho|$, as n increases, it suffices to have $N = O(\log n)$, which is encouraging! As a result, we do not need to have too many matrices in order to avoid the ill-conditioning that happens due to a small number of matrices being used. Note that if there is a structural cause of ill-conditioning within Λ_i's then this recipe is irrelevant. Also, note that in (4.26) for fixed n as N increases $|\rho|$ does not decrease indefinitely; and there is an asymptotic non-zero lower bound of 0.13. Unfortunately, our approximate bound in (4.25) does not give an interesting answer in this case. The mentioned behavior is observed in our simulations. Figure (4.2) shows the experimental and fitted behavior of ρ in terms of N for $n = 20$. The experimental ρ comes from generating λ_{ik}'s independently from uniform distribution on $[0, 1]$. Each value of ρ is an average over 1000 runs. The graph also shows the curve $\tilde{\rho} = 1 - 0.20(\frac{1}{n})^{\frac{5.59}{N}}$ (with $n = 20$), which is fitted to the experimental data. These two curves obviously demonstrate the predicted dynamics between n and N in determining ρ. The mentioned asymptotic lower bound for ρ as $N \to \infty$ in this case is 0.8. The interesting point is that, for small N improvement of ρ is dramatic as N increases, whereas for larger N and better ρ increasing N does not improve the sensitivity significantly. Recall that the important quantity in the sensitivity is $\frac{1}{1-\rho^2}$

which drops rapidly at first few N's. In fact for the experimental data it drops from 10^4 at $N = 2$ to 8.6 at $N = 10$. In this case the NOJD of only ten 20×20 matrices is well-conditioned or safe. Of course, use of more matrices improves the answer via averaging out the noise.

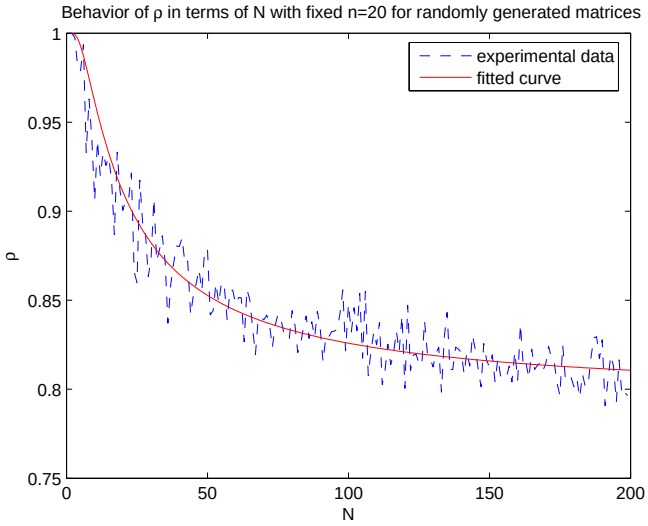

FIG. 4.2. *A typical behavior of $|\rho|$ in terms of N for fixed n. Here $n = 20$ and the λ_{ik}'s are generated from a uniform distribution on $[0, 1]$. The dashed curve shows the experimental ρ. Each point is an average over 1000 runs. The solid curve shows the curve: $\tilde{\rho} = 1 - 0.20(\frac{1}{n})^{\frac{5.59}{N}}$ (with $n = 20$), which is fitted to the data.*

The preceding discussion concerned the behavior of $|\rho|$ in terms of N and n. Unfortunately, a similar framework and analysis for μ do not seem obvious. Nevertheless, simulations show that, expectedly, whenever $\rho \approx 1$, μ is also close to unity. Therefore, our conclusion that "for small N and large n, the NOJD problem is ill-conditioned" stays valid when J_3 is used, as well.

5. Numerical experiments. In this section we perform some experiments to examine the derived results. The first example is just a toy example; the second one is a more realistic one in the context of Blind Source Separation(BSS).

5.1. Example 1. We investigate the effect of ρ, N and the condition number of A, cond(A), on the sensitivity of NOJD for matrices generated as in (4.1). We generate $\{\Lambda_i\}_{i=1}^N$ with elements that are i.i.d exponential random variables with mean 1. We choose $n = 10$. We also generate $A_{n\times n}$ randomly. Note that with probability one the joint diagonalizer for $\{C_i\}_{i=1}^N$ is unique. The noise matrices are with standard normal elements.

We consider the quality of joint diagonalization in terms of noise level t, ρ, and the condition number of A. We only consider J_1 and J_2 based methods. We use the QRJ2D algorithm[7] , introduced in [3], to find B and measure the error by:

$$\text{Index}(P) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{n}\frac{|p_{ij}|}{\max_k |p_{ik}|} - 1\Big) + \sum_{j=1}^{n}\Big(\sum_{i=1}^{n}\frac{|p_{ij}|}{\max_k |p_{kj}|} - 1\Big) \tag{5.1}$$

[7]Matlab code for this algorithm is available at http://www.isr.umd.edu/Labs/ISL/ICA2006/.

TABLE 5.1

(Left): Sensitivity of Index(BA) with respect to noise level t as N and hence ρ changes in Example 1. cond(A) = 25.11. (Right): Sensitivity of Index(BA) with respect to noise level t as cond(A) increases and $\|A\|_F = 1$. ($\rho = .68$)

Index(BA)	$t = 0$	$t = 0.01$
$N = 2, \rho = 0.9999$	3.9	17.0
$N = 4, \rho = 0.9959$	0.00	3.46
$N = 10, \rho = 0.9662$	0.00	1.46
$N = 100, \rho = 0.6903$	0.00	0.29
$N = 200, \rho = 0.60$	0.00	0.19

Index(BA) ($N = 100, \rho = .68$)	t=0	t=0.0001
cond(A) = 1	0.00	0.01
cond(A) = 2	0.00	0.01
cond(A) = 10	0.00	0.12
cond(A) = 50	0.00	3.02
cond(A) = 100	0.00	28.51

with $P = BA$. Index(BA) ≥ 0 and equality happens only when $BA = \Pi D$ (and hence $B = \Pi D A^{-1}$) for some Π and D. The smaller the Index is, the better joint diagonalization is, in the sense that B is closer to A^{-1}. We try different values of N (and hence ρ). We also investigate the effect of cond(A), by keeping the Λ_i's the same, and increasing cond(A) while $\|A\|_F$ is constant. Table (5.1) gives the results. The left sub-table shows the Index for different values of N (hence ρ) for two different noise levels $t = 0$ and $t = 0.01$. By increasing the number of matrices and hence improving the modulus of uniqueness, the sensitivity improves. Note that for $N = 2$, sensitivity is so high that the QRJ2D algorithm does not give a good answer even at zero noise. The right sub-table also shows the sensitivity degradation that happens because of increasing cond(A). In this experiment $\|A\|_F = 1$, $N = 100$, $t = 0$, and 0.0001. Sensitivity increases as conditioning of A degrades. Although the actual error values depend on the specific algorithm used, the trend of the error values as the parameters change gives an insight as to what factors affect the sensitivity.

5.2. Example 2: separation of non-stationary sources. Now we consider a more realistic situation, which is separation of non-stationary sources using NOJD of correlation matrices. This example also allows us to compare NOJD based on J_1 and J_3. The idea of using non-stationarity to separate sources has been described in [21]. Consider model (1.1) where the source vector is a Gaussian vector of independent components. Also, assume that the sources are non-stationary with varying variances. $R_{\mathbf{xx}}(t_i)$ the correlation matrix of the mixture at time t_i is:

$$R_{\mathbf{xx}}(t_i) = A\Lambda_s(t_i)A^T, \tag{5.2}$$

where $\Lambda_s(t_i)$ is the (diagonal) correlation matrix of the source at time t_i. Suppose that we gather the correlation matrices at times $t_1, ..., t_N$, and form the set $\{R(t_i)\}_{i=1}^N$. If $\Lambda_s(t_i)$ changes enough such that the modulus of uniqueness for this set is smaller than one, then NOJD of this set yields an estimation for A^{-1}; and hence it can result in separation of the mixture.

We have $n = 10$ sources. First we generate a random matrix $A_{n\times n}$. The condition number for this matrix is 75.11. Then we generate the sources as follows. We assume that the sources are stationary on short periods of $T = 100$ samples, and that they change their variances randomly at the end of each period. We consider $N = 20$ periods. During the i^{th} stationary period, the j^{th} source has Gaussian distribution with zero mean and a random variance λ_{ij}. We draw each random standard deviation $\sqrt{\lambda_{ij}}$ from a uniform distribution on $[0, 1]$. Also, during the stationary periods each source generates independent samples. The sources are mixed through A. In each

stationary period, we use the observed 100 samples of the mixture to estimate the correlation matrix for the mixture in that period. After the first stationary period, at the end of each stationary period, we perform an NOJD of the estimated correlation matrices gathered up to that time, in order to estimate A^{-1}. Also, as time passes, we compute ρ and μ for the set of true correlation matrices based on λ_{ij}'s. We use three different methods for NOJD of the estimated correlation matrices: (i) Pham's algorithm [20] which uses J_3 and requires positive definite matrices, (ii) QRJ2D algorithm, which is based on J_2, and (iii) FFDIAG [29] which is based on J_1. As we mentioned before, J_1 and J_2 based NOJD have similar sensitivity properties. We use QRJ2D and FFDIAG, since we want to have more evidence for comparing J_1-based-NOJD and J_3-based-NOJD. The output of each of these algorithms is an un-mixing matrix B. In order to measure the performance we use two measures. One is Index(BA), which we introduced before. The other one is the mean-squared Interference to Signal Ratio (ISR), which measures that at each restored source how much other sources are present. Note that from (4.2) for the recovered source vector $\vec{\mathbf{y}}$ we have:

$$\vec{\mathbf{y}} = B(t)\vec{\mathbf{x}} = B(t)A\vec{\mathbf{s}} \approx \vec{\mathbf{s}} + t\Delta\vec{\mathbf{s}}. \tag{5.3}$$

Here again, we have ignored the possible scaling and permutation ambiguity in the restored vector. In practice, of course, we compute Δ from $P = BA$, after re-ordering and normalizing the rows of P. As above equation suggests, $\|\Delta\|_F$ also measures the mean-squared ISR [8], i.e., how much interference from other sources is present in each recovered source. We use:

$$\mathrm{ISR} = 10 \log \frac{\|\Delta\|_F^2}{n} \tag{5.4}$$

as a measure of the interference. Note that in this example we have no noise and the source of error is the estimation error due to finite number of data samples. Another point that we want to examine is the sensitivity of NOJD based on J_3, in the case when one of the matrices becomes almost singular. For that purpose, in the last ($i = 20$) interval we set the standard deviation of six of the sources to 10^{-10}.

Figure (5.1) shows the results of the experiment. The top graph shows the Index(BA) in terms of i (which is the number of correlation matrices used) for different methods. The middle graph gives the ISR in terms of i. Note that the ISR measures for QRJ2D and FFDIAG are very close; however, the Index measure for these two methods differ. The bottom graph shows $\frac{1}{1-\rho^2}$ and $\frac{1}{\mu-1}$ in terms of i for the correlation matrices involved. As explained before, these two numbers, in fact, can be considered as condition numbers for NOJD based on J_1 and J_3, respectively (we have omitted the effect of A, i.e., $\mathrm{cond}(A)^2$, which is common in both condition numbers, see the last paragraphs of Sections 4.1 and 4.3). For $i = 2, 3$ the numbers are very high for both the cases. One can see that after the first few i, the condition numbers do not improve much. Note that the the condition number for the J_1-based-NOJD is higher than that of the J_3-based-NOJD; and at the same time the J_3-based-NOJD yields better separation except for $i = 2$ and $i = 20$. From our theoretical results this is certainly what we expect. However, we can not relate these two facts immediately, since the actual numbers depend on many factors. Note that NOJD for $i = 2, 3$ is not so effective, since the ISR measure is really poor (around or above -7dB). As a comparison, the best ISR that Pham's method achieves is -32dB, and the best one

[8] The author is thankful to one of the anonymous reviewers for reminding this observation.

that QRJ2D or FFDIAG achieve is about -23dB. The jumps at $i = 20$ in Index(BA) and in ISR for Pham's method are due to the fact that the last correlation matrix is almost non-singular. As we can see, the NOJD based on J_1 gives better separation in this case. Recall that J_1 does not require non-singular matrices, and the condition number we assigned for J_3-based-NOJD was based on the assumption that $\|\Lambda_i^{-1}\|_2$'s are not too large. At $i = 20$ this condition is violated and that is why despite the fact that $\frac{1}{\mu-1} < \frac{1}{1-\rho^2}$, the J_1-based-NOJD performs better. For the curious reader, we mention that despite some evidence in this example, for a given $\{\Lambda_i\}_{i=1}^N$ the conjecture that $\frac{1}{\mu_{kl}-1} \leq \frac{1}{1-\rho_{kl}^2}$ or $\frac{1}{\mu-1} \leq \frac{1}{1-\rho^2}$ is not true. However, note that for $\frac{1}{\mu-1} \leq \frac{1}{1-\rho^2}$ to hold it is sufficient to have $\mu \geq 2$, which can be achieved since the range for μ is the long half-line $[1, +\infty)$. This may explain why in most simulations and in this example $\frac{1}{\mu-1} < \frac{1}{1-\rho^2}$.

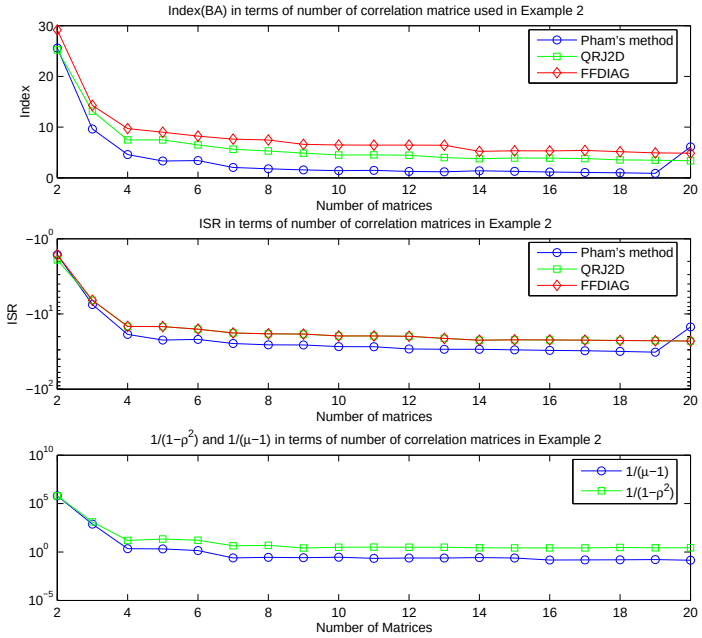

FIG. 5.1. *This figure shows the performance of source separation for non-stationary sources based on NOJD of correlation matrices at different times. As time passes more correlation matrices are used. Three different methods for NOJD are employed: Pham'a algorithm which is based on J_3, QRJ2D algorithm, which uses J_2 and FFDIAG which uses J_1. Top: Index(BA) in terms of number of correlation matrices used. Middle: ISR in terms of number of correlation matrices used. Bottom: $\frac{1}{1-\rho^2}$ and $\frac{1}{\mu-1}$ in terms of number of correlation matrices used. The jump seen at $i = 20$ in the graphs for J_3-based-NOJD is because in the last period some of the sources become extremely weak and the correlation matrix for that period becomes almost singular.*

6. Conclusions. We introduced the NOJD problem and the related EJD problem. We derived the uniqueness conditions for the EJD problem. We gave a joint diagonalization based formulation of ICA. Factors that affect the sensitivity of the NOJD problem were investigated. Modulus of uniqueness captures the uniqueness of the exact joint diagonalization problem and it affects the sensitivity of the NOJD

problem that arises from adding noise to clean matrices. Also we showed that if the sought joint diagonalizer is ill-conditioned, then sensitivity will be high. We tried to quantitatively show how dimension of the matrices and the number of matrices can affect the modulus of uniqueness. In particular, we showed that the NOJD problem can be very ill-conditioned if the number of matrices is small and they are fairly large. Sensitivity of the NOJD problem depends on the cost function used; and in one example we gave a comparison of the behaviors of two different cost functions for NOJD.

Acknowledgements. The framework outlined in Section 3.2.1 to derive the non-holonomic flow for NOJD has been conveyed to the author by Prof. P.S Krishnaprasad. The author also would like to thank him for his support of this work. The author feels indebted to the anonymous reviewers who have read this paper and whose insightful suggestions have helped improving the work. Also, the author cordially thanks Dr. Mehdi Kalantari, for his comments and suggestions to improve this paper.

REFERENCES

[1] P.-A. Absil and K.A. Gallivan, *Joint diagonalization on the oblique manifold for independent component analysis*, in Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), May 2006.

[2] B. Afsari, *Gradient flow based matrix joint diagonalization for independent component analysis*, master's thesis, Univeristy of Maryland, College Park, May 2004.

[3] ———, *Simple LU and QR based non-orthogonal matrix joint diagonalization.*, in ICA, J. P. Rosca, D. Erdogmus, J. C. P., and S. Haykin, eds., vol. 3889 of Lecture Notes in Computer Science, Springer, 2006, pp. 1–7.

[4] ———, *What can make joint diagonalization difficult?*, in Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP07), vol. 3, Honolulu, HI, April 2007, IEEE, pp. III–1377–III–1380.

[5] B. Afsari and P.S .Krishnaprasad, *A novel non-orthogonal joint diagonalization cost function for ICA*, tech. report, Inistute for Systems Research, University of Maryland 2005.

[6] B. Afsari and P. S. Krishnaprasad, *Some gradient based joint diagonalization methods for ICA*, in ICA, Carlos García Puntonet and Alberto Prieto, eds., vol. 3195 of Lecture Notes in Computer Science, Springer, 2004, pp. 437–444.

[7] A. Belouchrani, K.A. Merai m, J.-F. Cardoso, and E. Moulines, *A blind source separation technique using second-order statistics*, IEEE Transactions on Signal Processing, 45 (1997), pp. 434–444.

[8] J. F. Cardoso, *Perturbation of joint diagonalizers*, Tech. Report 94D023, Telecom Paris, 1994.

[9] J. F. Cardoso and A. Soloumiac, *Blind beamforming for non-gaussian signals*, IEE Proceedings-F, 140 (1993), pp. 362–370.

[10] P. Comon, *Independent component analysis: A new concept?*, Signal Processing, 36 (1994), pp. 287–314.

[11] ———, *Canonical tensor decompositions*, tech. report, Lab. I3S, June 2004.

[12] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer, 1999.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.

[14] R. Harshman, *Foundations of the PARAFAC procedure: model and conditions for an explanatory multi-mode factor analysis*, in UCLA Working Papers in Phonetics, vol. 16, 1970, pp. 1–84.

[15] U. Helmke and J.B. More, *Optimization and Dynamical Systems*, Springer-Verlag, 1994.

[16] A. M. Kagan, Y. V. Linnik, and C.R. Rao, *Characterization Problems in Mathematical Statistics*, Wiley, 1973.

[17] L. De Lathauwer, *Signal Processing Based on Multilinear Algebra*, PhD thesis, Katholiecke Univ., Leuven, Belgium, 1997.

[18] ———, *A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization*, SIAM Journal on Matrix Analysis and Applications, 28 (2006), pp. 642–666.

[19] E. Moreau, *A generalization of joint-diagonalization criteria for source separation*, IEEE Transactions on Signal Processing, 49 (2001), pp. 530–541.
[20] D. T. Pham, *Joint approximate diagonalization of positive definite hermitian matrices*, SIAM Journal of Matrix Analysis and Applications, 22 (2001), pp. 1136–1152.
[21] D-T Pham and J. F. Cardoso, *Blind separation of instantaneous mixtures of non stationary sources*, IEEE Transactions on Signal Processing, 49 (2001), pp. 1837–1848.
[22] G.W. Stewart and Ji guang Sun, *Matrix Perturbation Theory*, Academic Press, 1990.
[23] J. M.F. ten Berge and N.D. Sidiropoulos, *On uniqueness in CANDECOMP/PARAFAC*, Psychometrika, 67 (2002), pp. 399–409.
[24] L. Fejes Tóth, *On the densest packing of spherical caps*, The American Mathematical Monthly, 56 (1949), pp. 330–331.
[25] A. Van Der Veen, *Joint diagonalization via subspace fitting techniques*, in Proceedings of the 2001 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001), vol. V, Salt Lake City, UT, 2001, pp. 2773–2776.
[26] Eric W Weisstein, *Gamma function*. From MathWorld- A Wolfram Web Resource. http://mathworld.wolfram.com/GammaFunction.html.
[27] ———, *Hypersphere*. From MathWorld- A Wolfram Web Resource. http://mathworld.wolfram.com/Hypersphere.html.
[28] A. Yeredor, *Non-orthogonal joint diagonalization in the least-squares sence with application in blind source separation*, IEEE Transactions on Signal Processing, 50 (2002).
[29] A Ziehe, P Laskov, G Nolte, and K R Mueller, *A fast algorithm for joint diagonalization with non-orthogonal transformation and its application to blind source seperation*, Journal of Machine Learning Research, 5 (2004), pp. 777–800.