SOLVING AN INVERSE EIGENVALUE PROBLEM WITH TRIPLE CONSTRAINTS ON EIGENVALUES, SINGULAR VALUES, AND DIAGONAL ELEMENTS

SHENG-JHIH WU* AND MOODY T. CHU†

Abstract. An inverse eigenvalue problem usually entails two constraints, one conditioned upon the spectrum and the other on the structure. This paper investigates the problem where triple constraints of eigenvalues, singular values, and diagonal entries are imposed simultaneously. An approach combining an eclectic mix of skills from differential geometry, optimization theory, and analytic gradient flow is employed to prove the solvability of such a problem. The result generalizes the classical Mirsky, Sing-Thompson, and Weyl-Horn theorems concerning the respective majorization relationships between any two of the arrays of main diagonal entries, eigenvalues, and singular values. The existence theory fills a gap in the classical matrix theory. The problem might find applications in wireless communication and quantum information science. The technique employed can be implemented as a first-step numerical method for constructing the matrix. With slight modification, the approach might be used to explore similar types of inverse problems where the prescribed entries are at general locations.

Key words. inverse eigenvalue problem, majorization relationships, projected gradient, projected Hessian, analytic gradient dynamics,

AMS subject classifications. 65F18, 90C52, 15A29, 15A45,

1. Introduction. The focus of this paper is on the existence of a solution to a new type of inverse eigenvalue problem (IEP) where triple constraints of eigenvalues, singular values, and diagonal entries must be satisfied simultaneously. Before we present our results and explore some possible applications of this particular type of IEP, it might be fitting to give a brief recount on the general scope of IEPs and why they are interesting, important, and challenging.

The basic goal of a general IEP is to reconstruct the parameters of a certain physical system from the knowledge or desire of its dynamical behavior. Since a dynamics often is characterized by the underlying natural frequencies and/or normal modes, a fundamental ingredient in forming an IEP is the spectral constraints. Since the model should further be subject to a certain inherent physical feasibility such as the nonnegativity of parameters, the specific triangulation of a finite element grid, or the preconceived inter-connection in a mass-spring system, it often becomes necessary to impose additional structural constraints on the construction. Depending on the application, the structural constraints appear in different forms and, thus, lead to different challenges in IEPs.

Inverse Sturm-Liouville Problem. The general concepts mentioned above for both continuous and discrete IEPs might be illustrated by considering the classical regular Sturm-Liouville problem:

$$-\frac{d}{dx}\left(p(x)\frac{du(x)}{dx}\right)+q(x)u(x)=\lambda u(x),\ a<x<b, \tag{1.1}$$

where $p(x)$ and $q(x)$ are piecewise continuous on $[a,b]$ and appropriate boundary conditions are imposed. As a direct problem, it is known that eigenvalues of the system (1.1) are real, simple, countable, and tend to infinity. As an inverse problem, the question is to determine the potential function $q(x)$ from eigenvalues. This inverse problem, closely tied to the one-dimensional inverse scatting problem and served as a building block for scores of other important applications, has generated many interests in the field, notably the celebrated work by Gel′fand and Levitan [23]. Without repeating the details, we mention that the main idea is to employ a transformation operator to build a linear integral equation, now known as the Gel′fand-Levitan equation, which the kernel associated with the transformation operator must satisfy. Thus, the inverse problem is reduced to the solution to this linear integral equation. Once the kernel is solved from the equation, the potential can be obtained. In this way, the necessary and sufficient conditions for the solvability of the inverse problem are completely resolved. Simply put, the fundamental result that "two" data sequences of eigenvalues corresponding to two different boundary conditions are required to uniquely determine a potential. A quick introduction to this subject can be found in [6, Chapter 3]. A more thoroughgoing discussion was done in the translated book [37]. The more recent

*Center for Advanced Statistics and Econometrics Research, School of Mathematical Sciences, Soochow University, Suzhou, China. (szwu@suda.edu.cn)

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu) This research was supported in part by the National Science Foundation under grant DMS-1316779.

monograph [21] uses the Sturm-Liouville operator as a model to describe the main ideas and methods for the general theory of inverse spectral problems and contains a chapter of specific applications.

On the other hand, discretization is probably the only apparatus we have in hand when tackling the Sturm-Liouville problems numerically [3, 49, 51, 63]. To demonstrate such an approach, consider the simple Sturm-Liouville operator when $p(x) \equiv 1$, and $[a, b] = [0, 1]$. Suppose that the central difference scheme with mesh size $h = \frac{1}{n+1}$ is applied. The differential equation (1.1) is reduced to the matrix eigenvalue problem

$$\left(-\frac{1}{h^2} J_0 + X\right) \mathbf{u} = \lambda \mathbf{u}, \tag{1.2}$$

where J_0 is the fixed tridiagonal matrix whose diagonal entries are all 2's and the super- and the sub-diagonals are all -1's, and X is the diagonal matrix representing the evaluation of the potential function $q(x)$ at the grid points. The analogue to the inverse Sturm-Liouville problem is to determine a diagonal matrix X so that the matrix $-\frac{1}{h^2} J_0 + X$ possesses a (finitely many) prescribed spectrum. See the thesis [27] on discrete Sturm-Liouville problems and the article [48] on the comparison between the continuous and discrete problems. Note how nature the structural constraint of the problem comes to place in (1.2) when discretizing (1.1) by the central difference scheme. A finite element discretization for partial differential equations such as the Helmholtz equation will result in other types of structured IEPs. In a different context, see also [17, 43] for correlation matrix structure, [36] for upper triangular structure, and [31] for the structure of prescribed entries at arbitrary locations.

Matrix inverse eigenvalue problems. Just like the IEPs for differential equations, the IEPs for matrices have been studied extensively with applications to system and control theory, geophysics, molecular spectroscopy, particle physics, structure analysis, numerical analysis, and many other disciplines. One common assumption in the application of inverse problems is that the underlying physical system is somehow representable in terms of matrices [35]. Studying IEPs for matrices is therefore equally important as for differential systems. Our book [14] identifies 21 major distinct characteristics in the IEP formulation, each with several variations, and describes many applications together with a list of over 430 references. The survey article [19] contains a more impressive list of 774 references on direct, semi-inverse and inverse eigenvalue problems for structures described by differential equations. The newly expanded seminal book [25] contains works and references in the engineering literature.

The singular values decomposition (SVD) and its inherent properties contain innate critical information of the data that a matrix represents. A wide range of applications such as image compression, dimension reduction, noise removal, and principal component analysis exploits features of the SVD [55]. A natural outgrowth of the inverse eigenvalue problems is the generalization to inverse singular value problems (ISVP) for model reconstruction [8, 42, 58, 60, 61]. The ISVP can be categorized as specially structured IEP [13].

IEP with three constraints. This paper considers a new type of IEP, demanding all three constraints, i.e., eigenvalues, singular values, and diagonals, be satisfied concurrently. To our knowledge, such an inverse problem has never been considered before and imposes immediate challenges to conventional methods. We propose an approach utilizing an eclectic mix of skills from analytic gradient dynamics and optimization theory to successfully tackle this new and challenging problem. Our primary goal is to establish the existence theory, but the proof itself can be employed as a numerical method as well. The technique, innovative in itself, might be useful for exploring other types of existence questions. For example, the prescribed entries in this paper are limited to the diagonal only. With little modification of the flow to be described below, we have experimented numerically the same technique with problems where the prescribed entries are given at locations other than the diagonal, which thus far has no known theory of existence yet. So as to stay focus on the technique, we shall not pursue this direction in this paper, but can furnish the empirical report upon request. Some related discussions on IEPs with prescribed eigenvalues and arbitrarily prescribed entries, can be found in [31, 41], but these problems do not involve prescribed singular values.

Although the results remain valid over the complex field, we limit our discussion to the real field for the ease of conveying the idea. With appropriate modifications, e.g., using unitary similarity transformations instead of orthogonal similarity transformations, our technique can be carried over to the general complex case. We shall leave the generalization to interested readers, but refer them to [10, Section 5] for a worked-out case on how such an extension can be accomplished.

2. Preliminaries. We point out immediately that not all prescribed sets of values are feasible as sets of singular values, eigenvalues, and diagonal entries simultaneously. There are limitations upon these constraints.

Let $\mathbf{d} \in \mathbb{R}^n$ denote the vector whose entries are the desirable diagonal elements and are arranged in the order $|d_1| \geq \ldots \geq |d_n|$, $\boldsymbol{\sigma} \in \mathbb{R}^n$ the nonnegative vector of desirable singular values in the order $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$, and $\boldsymbol{\lambda} \in \mathbb{C}^n$ the complex vector of desirable eigenvalues that are closed under complex conjugation and are ordered as $|\lambda_1| \geq \ldots \geq |\lambda_n|$.

2.1. Necessary conditions. Inherent to all matrices is a universal property that diagonal elements, eigenvalues, and singular values are necessarily related in peculiar way. Each relationship is characterized by a specific sequence of inequalities. Satisfying these inequalities is a prerequisite before we can proceed for construction. So that the paper is self-contained, we state these relationships in this section — pairs of $\mathbf{d}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\sigma}$ must comply with the following classical results in matrix theory [44].

The inequality relationship between the singular values $\boldsymbol{\sigma}$ and the eigenvalues $\boldsymbol{\lambda}$ is usually referred to as that $\boldsymbol{\lambda}$ is log majorized by $\boldsymbol{\sigma}$ from above.

THEOREM 2.1. *(Weyl-Horn Theorem [29, 59]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$ if and only if*

$$\prod_{i=1}^{k} |\lambda_i| \leq \prod_{i=1}^{k} \sigma_i, \quad k = 1, 2, \ldots, n-1, \tag{2.1}$$

and

$$\prod_{i=1}^{n} |\lambda_i| = \prod_{i=1}^{n} \sigma_i. \tag{2.2}$$

The relationship between the singular values $\boldsymbol{\sigma}$ and the main diagonal entries $\mathbf{d}$ is a combination of weak majorization and an additional inequality.

THEOREM 2.2. *(Sing-Thompson Theorem [54, 56]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with singular values $\boldsymbol{\sigma}$ and main diagonal entries $\mathbf{d}$, possibly in different order, if and only if*

$$\sum_{i=1}^{k} |d_i| \leq \sum_{i=1}^{k} \sigma_i, \quad k = 1, 2, \ldots, n, \tag{2.3}$$

and

$$\sum_{i=1}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{n-1} \sigma_i - \sigma_n. \tag{2.4}$$

With regard to the relationship between the diagonal entries $\mathbf{d}$ and the eigenvalues $\boldsymbol{\lambda}$, we have two separate results. The first result holds for general matrices and is in its most common form.

THEOREM 2.3. *(Mirsky Theorem [45]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with eigenvalues $\boldsymbol{\lambda}$ and main diagonal entries $\mathbf{d}$, possibly in different order, if and only if*

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} d_i. \tag{2.5}$$

Obviously the condition (2.5) is too general to be useful as it is applicable to all matrices. When a matrix is structured, a more restrictive condition than (2.5) should hold. For Hermitian matrices, the following set of inequalities sometimes is referred to as a majorization of $\mathbf{d}$ to $\boldsymbol{\lambda}$ from above.

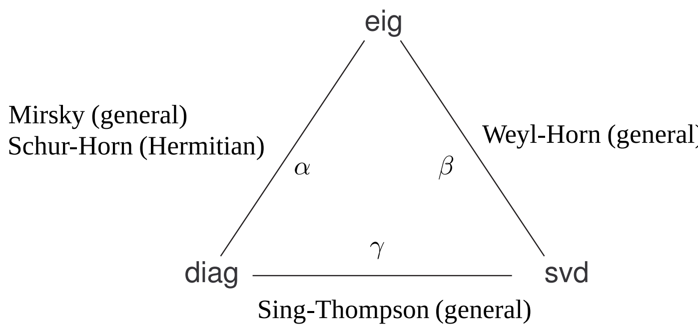

FIGURE 2.1. *Existing results concerning majorization.*

THEOREM 2.4. *(Schur-Horn Theorem [5]) Suppose that entries of $\boldsymbol{\lambda} \in \mathbb{R}^n$ are arranged in the order $\lambda_1 \geq \ldots \geq \lambda_n$ and entries of $\mathbf{d} \in \mathbb{R}^n$ in the order $d_1 \geq \ldots \geq d_n$. There exists a Hermitian matrix H with eigenvalues $\boldsymbol{\lambda}$ and diagonal entries $\mathbf{d}$, possibly in different order, if and only if*

$$\sum_{i=1}^{k} \lambda_{n-i+1} \leq \sum_{i=1}^{k} d_{n-i+1}, \quad k = 1, 2, \ldots, n-1, \tag{2.6}$$

and

$$\sum_{i=1}^{n} \lambda_{n-i+1} = \sum_{i=1}^{n} d_{n-i+1}. \tag{2.7}$$

It is intriguing that merely being a matrix, the structure alone induces these intrinsic inequalities, which we collectively refer to as majorization properties. For quick reference, we represent the mutual relationships by the three sides, denoted as α, β, and γ, respectively, of the triangle depicted in Figure 2.1.

2.2. Sufficient conditions. What makes the above results significant is that the conditions specified in each of the four theorems are both necessary and sufficient. Given a set of data satisfying any *one side* of the triangle, a matrix satisfying the prescribed characteristics does exist. A constructive proof of such a sufficient condition, an IEP, often can be converted into a numerical method, which has been extensively studied in the literature [7, 15, 11, 12, 18, 36, 62].

One common feature associated with these inverse problems is that the solution is not unique. An algorithm therefore may fail to single out a specific matrix. For instance, starting with a given matrix $A \in \mathbb{R}^{n \times n}$, we can calculate its eigenvalues $\boldsymbol{\lambda}$ and singular values $\boldsymbol{\sigma}$ which necessarily satisfy the inequalities (2.1) and (2.2). Applying the divide-and-conquer algorithm proposed in [12] to the set of data $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$, we can construct a matrix B which has the very same eigenvalues $\boldsymbol{\lambda}$ and singular values $\boldsymbol{\sigma}$. However, it is mostly the case that the newly constructed matrix B is entirely different from the original matrix A. Such discretion can easily be explained — There are more degrees of freedom in the matrix to be constructed than the prescribed data can characterize. Generally speaking, the inverse problem has multiple solutions and more conditions can be imposed.

Referring to Figure 2.1, we are curious to ask whether a matrix can satisfy any *two sides* of the triangle simultaneously. Clearly, satisfying any two of the three majorization conditions will automatically satisfy the third condition. Thus, this problem is equivalent to whether a matrix can be constructed to satisfy prescribed diagonal entries, eigenvalues, and singular values concurrently. Because such a matrix will satisfy the three sets of inequalities in Theorems 2.2, 2.1, and 2.3 all together, we shall refer to this structure as the Mirsky-Weyl-Horn-Sing-Thompson (MWHST) condition.

Note that we do not include the Schur-Horn condition which is for Hermitian matrices. In the event that symmetry is part of the desirable structure, singular values are the absolute values of eigenvalues and are automatically fixed. In this case, it suffices to consider the inverse problem of satisfying the Schur-Horn condition alone, which is already solved in [7, 15, 62]. The MWHST condition constitutes a new and harder problem because the matrix under construction has no symmetric structure.

3. Possible applications. The notion of majorization arises in a wide range of applications including statistics [2, 44], information theory [30, 33], system identification [4], wireless communication [34], quantum computing [47], and quantum mechanics [46, 53], to mention just a few. To review the complex theory in any of the applications is obviously beyond our capacity. At the risk of oversimplifying, we find that it often is the case that a certain majorization relationship is used as a criterium for determining the condition of a certain physical state in nature. We mention the criteria for checking the distillability of a bipartite quantum state [28] and the transformability between two pure entangled states [30] as two such instances. If the majorization relationship is also sufficient, then a solution to the corresponding inverse problem amounts to constructing the state under the desirable condition.

As to our specific inverse problem subject to the MWHST condition, we briefly outline two possible applications. There are far more details beyond the scope of this paper, so we only sketch the ideas.

Optimal signature matrix construction. One of the two major wireless communication standards in the industry is the code-division multiple access (CDMA) technology [52]. In the simplest setting, the base station receives a superposed signal

$$\mathfrak{s}(t) = \sum_{k=1}^{N} b_k(t)\sqrt{\omega_k(t)}\mathbf{s}_k(t) + \mathbf{v}(t)$$

at time t from N users, where the unit vector $\mathbf{s}_k$ stands for the kth user's unique signature vector, ω_k is the received power level, $\mathbf{v}$ is the additive noise, and the message is encoded in the complex number b_k. The task for the base station is to extract all b_k's from $\mathfrak{s}$. In practice, $\mathfrak{s}(t)$ is transmitted at fixed time intervals, so $\mathfrak{s}(t)$ is sampled as a time series, also known as a bitstream. So does the message $b_k(t)$ sent by the kth user. It can be shown that the chip-sampled, matched-filter outputs provide sufficient statistics for deciphering the signal. For the system to effectively extract the messages for each individual user, the signature vectors should be well separated from each other. This leads to the problem of constructing optimal weighted signature matrix

$$X(t) = [\omega_1(t)\mathbf{s}_1(t), \dots, \omega_1(t)\mathbf{s}_N(t)]$$

under white noise, where the optimization (of separation) is gauged against different structural constraints. In the context of CDMA, for example, the ideal case without constraint is that X is row-orthogonal with specified column norms which are the singular values of $X(t)$. Under constraints, constructing the minimally correlated vectors can be formulated as an inverse singular value problem [16, 57]. Note that $X(t)$ evolves dynamically in t. As time moves on, we may need to address the stability of $X(t)$. This can be done by embedding $X(t)$ in a square matrix, say, by padding with zeros, and control the growth of resulting eigenvalues. In this way, we are solving an inverse problem with prescribed singular values, eigenvalues, and prescribed entries at specified locations. We have complete theory when the specified entries are at the diagonal, but the gradient flow approach developed in this paper for prescribed diagonal entries is readily generalizable to prescribed values at arbitrary locations. Thus far, no known theory exists for the general problem, but we have conducted considerable numerical experiments. Our software package for the general problem is available upon request.

Observable preserving nearest separable system approximation. One of the most fundamental challenges in quantum information science is the entanglement of subsystems. The simplest setting to see the entanglement is the bipartite system which corresponds to tensor product of two matrices. We use real-valued general matrices in this note to simply convey the idea. The basic question is whether a given matrix $A \in \mathbb{R}^{mn\times mn}$ can be written in the form

$$A = \sum_{i=1}^{k} X_i \otimes Y_i,$$

where $X_i \in \mathbb{R}^{m\times m}$, $Y_i \in \mathbb{R}^{n\times n}$, and $\otimes$ stands for the Kronecker product. If yes, we say that A is separable; otherwise, it is tangled [30]. This problem is closely related to so called tensor decomposition, if A is regarded as an order-4 tensor. In that regard, X_i and Y_i are required to be rank-1 matrices and the minimal value of k is called the tensor rank of A. In the context of quantum states, there are additional constraints — all matrices involved should be density matrices, i.e., Hermitian, positive-definite matrices with unit trace. As if it is a natural phenomenon, most quantum systems are tangled. It is therefore interesting to find its nearest separable system. Eigenvalues, singular values, and diagonal entries are important quantities known as the observables. For the approximation to be meaningful, we would like to see that the composition of the subsystems preserves these observables as much as possible. Consider the case $k = 1$. Then it is known that eigenvalues of A are $\lambda_i\mu_j$, if $\{\lambda_i\}$ and $\{\mu_j\}$ are the spectra of X_1 and Y_1, respectively. Likewise, singular values and diagonal entries of A can be expressed algebraically in terms of those of X_1 and Y_1, respectively. We thus want the subsystems to satisfy these algebraic relationships between eigenvalues, singular values, and diagonal entries. Untangling these scalars (observables) of A subject to the MWHST condition is equivalent to an inequality constrained optimization problem. The task is not trivial, but certainly is easier than untangling A itself while preserving the approximate observables. Once the approximated separation of the observables is achieved, we expect the subsystems X_1 and Y_1 to respect this essential information in a matrix. That is, as feasible candidates, X_1 and Y_1 should have prescribed eigenvalues, singular values, and diagonal entries. For $k > 1$, the separation of the observables for sum of matrices must satisfy additional inequality conditions [22, 38] which we will not elaborate here. After the approximate observables are obtained, we assign them to each X_i and Y_i and solve the corresponding inverse problems, respectively.

4. Our contributions. In regard to our particular inverse problem subject to the MWHST condition, our first motivation is by mathematical curiosity. It is of interest to ask whether the three sets of conditions (2.1) to (2.5) can ever be coordinated together as one sufficient condition for the existence of a common matrix with diagonal elements $\mathbf{d}$, eigenvalues $\boldsymbol{\lambda}$, and singular values $\boldsymbol{\sigma}$. How to construct such a matrix, if it exists, is another interesting question. This paper addresses these two questions.

Our contributions are twofold. First, our main theoretical result is summarized as follows. Second, the technique we employ along the way to prove this result is a gradient dynamical system which can be implemented as a numerical method. Although we have not studied its efficiency in this paper, the flow approach might be the first tool of its kind in the literature to tackle this three-constraint inverse problem.

THEOREM 4.1. *Given three sets of data, $\mathbf{d} = [d_i] \in \mathbb{R}^n$, $\boldsymbol{\sigma} = [\sigma_i] \in \mathbb{R}^n$, and $\boldsymbol{\lambda} = [\lambda_i] \in \mathbb{C}^n$, suppose that the entries can be arranged in the order $|d_1| \geq \ldots \geq |d_n|$, $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$, $|\lambda_1| \geq \ldots \geq |\lambda_n|$, and closed under complex conjugation. Suppose that $n \geq 3$. Then the MWHST condition is sufficient for the existence of a real-value matrix with $\mathbf{d}$ as its diagonal, possibly in a different order, $\boldsymbol{\sigma}$ as its singular values, and $\boldsymbol{\lambda}$ as its eigenvalues.*

The seemingly trivial problem for the case of $n = 2$ is very different from the general case. We analyze in Section 5 that the MWHST condition itself is not enough to guarantee the existence of a 2×2 matrix. We specify what other conditions must be satisfied for the existence. For the case $n \geq 3$, we argue in Section 6 by means of optimization that generically there does exist a matrix satisfying the MWHST condition. The proof starts with the assumption that the inverse problem associated with the Weyl-Horn theorem is already solved and follows a gradient flow to its equilibrium point which is a solution. Global convergence is guaranteed and, hence, the above theorem is proved.

5. The case of 2×2 is special. In searching for a matrix satisfying the MWHST condition, we first consider the 2×2 case. As will be seen below, this seemingly simple problem is actually quite complicated. There are mutually exclusive situations to be considered and additional conditions must be imposed to guarantee the existence. It might be even more astounding that such a difficulty does not occur when dealing with higher dimensional cases, which will be shown in the next section.

To fix the idea, let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

denote the 2×2 real matrix to be constructed. The Frobenius norm of A necessarily implies the equality

$$a^2 + b^2 + c^2 + d^2 = \sigma_1^2 + \sigma_2^2.$$

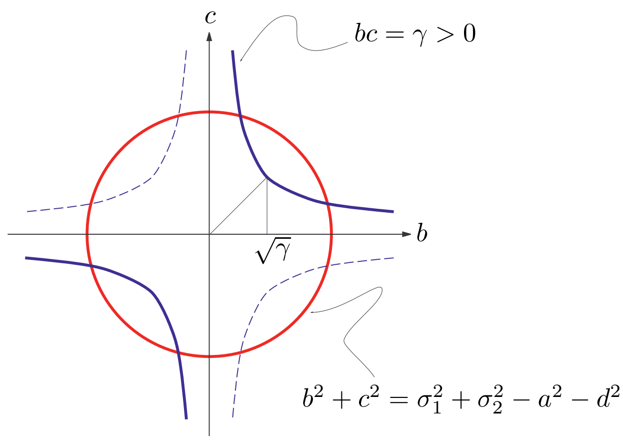

FIGURE 5.1. *Existence of $A \in \mathbb{R}^{2\times 2}$ satisfying the MWHST condition.*

Assuming that the main diagonal entries are already fixed, our goal is to determine the off-diagonal entries b and c to meet the prescribed eigenvalues and singular values, which translates to the system

$$\left\{\begin{array}{rcl} bc & = & ad - \lambda_1\lambda_2, \\ b^2 + c^2 & = & \sigma_1^2 + \sigma_2^2 - a^2 - d^2. \end{array}\right. \tag{5.1}$$

The existence of a 2×2 matrix A satisfying the MWHST condition therefore boils down to finding the intersection of a hyperbola and a circle as is indicated in Figure 5.1. Obviously, the system (5.1) is solvable for the off-diagonal entries b and c only if the vertex of the hyperbola lies within the disk, that is, when

$$2|ad - \lambda_1\lambda_2| \leq \sigma_1^2 + \sigma_2^2 - a^2 - d^2. \tag{5.2}$$

In this case, there are generically four solutions per given $\mathbf{d}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\sigma}$.

On the other hand, the MWHST condition requires that the following three sets of inequalities be held simultaneously:

$$\lambda_1 + \lambda_2 = a + d; \qquad (\mathrm{Mirsky}) \tag{5.3}$$

$$\left\{\begin{array}{rcl} |\lambda_1| & \geq & |\lambda_2|, \\ \sigma_1 & \geq & \sigma_2, \\ |\lambda_1| & \leq & \sigma_1, \\ |\lambda_1||\lambda_2| & = & \sigma_1\sigma_2; \end{array}\right. \qquad (\mathrm{Weyl-Horn}) \tag{5.4}$$

$$\left\{\begin{array}{rcl} |a| & \geq & |d|, \\ |a| + |d| & \leq & \sigma_1 + \sigma_2, \\ |a| - |d| & \leq & \sigma_1 - \sigma_2. \end{array}\right. \qquad (\mathrm{Sing-Thompson}) \tag{5.5}$$

We now examine how these inequalities play out to ensure (5.2), which guarantees a solution.

Given singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$, summarized in Figure 5.2 are various regions of $(a, d) \in \mathbb{R}^2$ over which the inequality (5.2) holds. Since a significant amount of information is contained in the drawing, we briefly explain its interpretation in the following theorem. The analysis is tedious but straightforward.

THEOREM 5.1. *Given three sets of data $\mathbf{d}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\sigma}$ satisfying the MWHST condition, then a 2×2 matrix A exists with prescribed diagonal entries $\mathbf{d}$, eigenvalues $\boldsymbol{\lambda}$, and singular values $\boldsymbol{\sigma}$ if and only if the following (additional) conditions on $\mathbf{d}$ hold.*

1. *Given $\boldsymbol{\sigma}$, the shape of "kissing fish" in Figure 5.2 represents the feasible region of the diagonal (a, d) in order to satisfy the Sing-Thompson condition (5.5) alone. (See [11] for details.)*

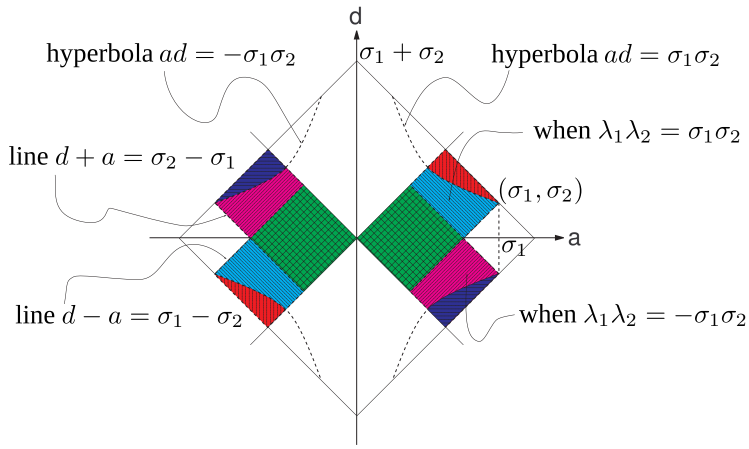

FIGURE 5.2. *Domain of feasible diagonal entries (a, d), given $\{\sigma_1, \sigma_2\}$ and $\{\lambda_1, \lambda_2\}$.*

2. *The Weyl-Horn condition (5.4) defines two mutually exclusive cases — either $\lambda_1\lambda_2 = \sigma_1\sigma_2$ or $\lambda_1\lambda_2 = -\sigma_1\sigma_2$.*
3. *When eigenvalues λ_1 and λ_2 are either complex conjugate or real valued with the same sign, then*
   - *(a) The only feasible diagonal entries must be further restricted to the union of the red, cyan, and green regions in Figure 5.2.*
   - *(b) When $ad \geq \sigma_1\sigma_2$, the (a, d) must come from the red region.*
4. *When eigenvalues λ_1 and λ_2 are of opposite sign,*
   - *(a) The only feasible diagonal entries must be further restricted to the union of the blue, purple, and green regions in Figure 5.2.*
   - *(b) When $ad \leq -\sigma_1\sigma_2$, the (a, d) must come from the blue region.*
5. *The green region is when both $|a + d| \leq \sigma_1 - \sigma_2$ and $|a - d| \leq \sigma_1 - \sigma_2$.*

The above elaboration on the 2×2 case is illuminating. It manifests that satisfying the MWHST condition alone by the prescribed diagonal entries, eigenvalues, and singular values is *not* sufficient to guarantee the existence of a 2×2 matrix. Indeed, depending on other factors such as the sign of $\lambda_1\lambda_2 = \pm\sigma_1\sigma_2$, the location of the diagonal entries (a, d) also comes to play in the solvability of the inverse problem for the 2×2 case. Such a simple fact is of interest in its own right. For instance, the following corollary is a special case of Theorem 5.1.

COROLLARY 5.2. *Suppose that the MWHST condition holds for three given sets of data $\mathbf{d}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\sigma}$. If, in addition, $|a + d| \leq \sigma_1 - \sigma_2$ and $|a - d| \leq \sigma_1 - \sigma_2$, then there exists a 2×2 real matrix with diagonal entries $\mathbf{d}$, eigenvalues $\boldsymbol{\lambda}$, and singular values $\boldsymbol{\sigma}$.*

The understanding of the 2×2 case seems to suggest that the pursuit for a matrix satisfying all inequalities simultaneously in the MWHST condition should have come to an end. It is not so. The 2×2 case discussed above is only an exception. In the next section, we argue that the inverse problem of constructing a real-valued matrix satisfying the MWHST condition is generically solvable when $n \geq 3$.

6. Existence in general case. A rough count of the dimensionality might give clue to the solvability, though in reality the MWHST condition consists of inequalities which make the dimensionality analysis not so straightforward. In the case $n = 2$, the task was to determine the two off-diagonal entries b and c so as to result in having two prescribed eigenvalues and two prescribed singular values. At first glance, this might seem to be an over-determined system. However, the Mirsky condition (5.3) and the last equality in the Weyl-Horn condition (5.4) imply that actually there are only one eigenvalue condition and one singular value condition to be satisfied. Thus, the reconstruction problem amounts to solving two equations in two unknowns, such as that of (5.1). In order to ensure that this nonlinear problem has a solution (indeed, four solutions generically, if it is ever solvable), we have concluded in the preceding section that some additional constraints are required. For the case $n \geq 3$, it is difficult to use a geometric argument directly. Instead, we prove the existence by an entirely different strategy.

6.1. Variational formulation. We begin with the assumption that an $n \times n$ real-valued matrix A satisfying the Weyl-Horn condition is already in hand. The existence of such a matrix with prescribed singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$ is guaranteed in theory and is obtainable numerically.

Referring to the diagram in Figure 2.1, we assume that the β side (Weyl-Horn condition) of the triangular relationship is already satisfied by A. If we can establish the α side (Mirsky condition) in the diagram, then the γ side follows automatically. In other words, given the desired main diagonal elements $\mathbf{d}$ satisfying both the Mirsky condition and the Sing-Thompson condition, our goal now is to somehow transform the matrix A so that the resulting diagonal elements agree with elements of the prescribed $\mathbf{d}$. A critical question is what kinds of transformations are allowed.

Foremost, we want to preserve the spectrum of the given A so as not to upset the Weyl-Horn condition. So we have to employ similarity transformations. Likewise, to preserve the singular values we must perform orthogonal equivalence transformations. To keep both the eigenvalues and the singular values invariant, therefore, the only option is to apply orthogonal similarity transformations to the matrix A.

Let $\mathcal{O}(n) \subset \mathbb{R}^{n\times n}$ denote the group of $n \times n$ real orthogonal matrices. Also, let $\operatorname{diag}(M)$ denote the diagonal matrix whose main diagonal is the same as that of the matrix M and $\operatorname{diag}(\mathbf{v})$ the diagonal matrix whose main diagonal entries are formed from the vector $\mathbf{v}$. Using the same notation diag for both M and $\mathbf{v}$ will prove convenient in the discussion. Any ambiguity can be clarified from the context. Our idea of driving the diagonal of $Q^\top AQ$ to that of the specified vector $\mathbf{d}$ is to formulate the minimization problem

$$\min_{Q\in\mathcal{O}(n)} F(Q) := \frac{1}{2}\|\operatorname{diag}(Q^\top AQ) - \operatorname{diag}(\mathbf{d})\|_F^2, \tag{6.1}$$

where $\|\cdot\|_F$ stands for the Frobenius matrix norm. Since $\mathbf{d}$ is already specifically ordered as we have premised in Section 2.1, included in the formulation (6.1) is an implicit sorting that, if convergence ever occurs, the orthogonal matrix Q should align $\operatorname{diag}(Q^\top AQ)$ to conform to that ordering.

Since the matrix A is real, $\operatorname{diag}(Q^\top AQ) = \operatorname{diag}(Q^\top A^\top Q)$. Therefore,

$$\operatorname{diag}(Q^\top AQ)) = \operatorname{diag}(Q^\top \frac{A + A^\top}{2} Q). \tag{6.2}$$

Define the matrix

$$S := \frac{A + A^\top}{2}. \tag{6.3}$$

It is more convenient to work on the (symmetrized) optimization problem[1]

$$\min_{Q\in\mathcal{O}(n)} F(Q) := \frac{1}{2}\|\operatorname{diag}(Q^\top SQ) - \operatorname{diag}(\mathbf{d})\|_F^2. \tag{6.4}$$

The optimizer Q of problem (6.1) is the same as that of problem (6.4), and vice versa.

Denote

$$\eta(Q) := \operatorname{diag}(Q^\top SQ) - \operatorname{diag}(\mathbf{d}). \tag{6.5}$$

If we can find an orthogonal matrix $Q \in \mathcal{O}(n)$ such that $\eta(Q) = 0$, then the very same Q will make the main diagonal entries, the eigenvalues, and the singular values of the matrix $Q^\top AQ$ satisfy the MWHST condition simultaneously. In the remainder of this paper, we focus on proving the following claim.

THEOREM 6.1. *Given three sets of data, $\mathbf{d}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\lambda}$ satisfying the MWHST condition as in Theorem 4.1. Assume that A is a matrix with prescribed singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$ and that S defined in (6.3) is not identically zero. Then there exists an optimizer $Q \in \mathcal{O}(n)$ for the problem (6.4) such that $F(Q) = 0$.*

[1] Obviously, if A happens to be a skew-symmetric matrix, then $S = 0$ and $F(Q)$ is a constant. In this case, we cannot do anything with (6.4). However, the skew-symmetry is just another kind of symmetry and is easier to exploit than the general non-symmetry. Indeed, if A is skew-symmetric, then the discussion in this paper can be equally applied to (6.1) with appropriate changes of sign due to the fact that $A^\top = -A$. To save space, we shall not analyze this case in this article because it is merely a repetition of most of the arguments to be developed. We shall assume that generically the matrix A constructed to satisfy the Weyl-Horn condition is not skew-symmetric.

6.2. Projected gradient flow. We shall deal with the optimization problem (6.4) by applying the conventional optimization techniques to the matrices. Of significance is that we are able calculate the projected gradient and the projected Hessian analytically without resorting to the Lagrange multiplier theory [10].

Given two matrices $M = [m_{ij}]$ and $N = [n_{ij}]$ of the same size, denote their Frobenius inner product by

$$\langle M, N\rangle := \sum_{i,j} m_{ij} n_{ij}.$$

The Fréchet derivative $F'(Q)$ of F at Q is a linear operator mapping $\mathbb{R}^{n\times n}$ to $\mathbb{R}^{n\times n}$. Specifically, its action on an arbitrary matrix $H \in \mathbb{R}^{n\times n}$ is given by

$$F'(Q).H = \langle \eta(Q), \eta'(Q).H\rangle = 2\left\langle \eta(Q), \operatorname{diag}(Q^\top SH)\right\rangle = 2\left\langle \eta(Q), Q^\top SH\right\rangle = 2\left\langle SQ\eta(Q), H\right\rangle. \quad (6.6)$$

In (6.6), the second equality follows from the linearity of the operator diag; the third equality results from the fact $\eta(Q)$ is a diagonal matrix; and the fourth equality is due to the adjoint property. By the Riesz representation theorem, the gradient ∇F at Q can be represented as

$$\nabla F(Q) = 2SQ\eta(Q). \quad (6.7)$$

Because of the constraint that Q must be in $\mathcal{O}(n)$, we next calculate the projected gradient $g(Q)$ of $\nabla F(Q)$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(n)$ of $\mathcal{O}(n)$. Toward this end, we first recognize that the tangent space of $\mathcal{O}(n)$ at Q can be identified as the left translation of the subspace of skew-symmetric matrices, that is,

$$\mathcal{T}_Q\mathcal{O}(n) = \{QK \mid K \text{ is skew-symmetric}\}. \quad (6.8)$$

The projection operator onto the tangent space of $\mathcal{O}(n)$ can be obtained via the following formula [10].

LEMMA 6.2. *Let $Q \in \mathcal{O}(n)$ be a fixed orthogonal matrix. The projection of any given matrix $X \in \mathbb{R}^{n\times n}$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(n)$ is given by*

$$\mathbb{P}_{\mathcal{T}_Q\mathcal{O}(n)}(X) = \frac{1}{2}Q\left\{Q^\top X - X^\top Q\right\}. \quad (6.9)$$

In particular, the projected gradient $g(Q) := \mathbb{P}_{\mathcal{T}_Q\mathcal{O}(n)}(\nabla F(Q))$ of $\nabla F(Q)$ onto $\mathcal{O}(n)$ is given explicitly by

$$g(Q) = \frac{1}{2}Q\left(Q^\top \nabla F(Q) - \nabla F(Q)^\top Q\right) = Q\left[Q^\top SQ, \eta(Q)\right], \quad (6.10)$$

where, for convenience, we adopt the Lie bracket notation

$$[M, N] = MN - NM.$$

Define the dynamical system

$$\dot{Q} = -g(Q) = Q\left[\eta(Q), Q^\top SQ\right], \quad Q(0) = I. \quad (6.11)$$

By construction, the solution flow $Q(t)$ to the differential system (6.11) stays on the manifold $\mathcal{O}(n)$ and moves in the steepest descent direction for the objective function $F(Q)$.

In the next three subsections, we argue to make three points.

1. The asymptotically stable equilibria of this projected gradient flow are geometrically isolated.
2. The projected Hessian at an equilibrium point is explicitly computable.
3. Any equilibrium point of the projected gradient provides the orthogonal matrix Q at which $\eta(Q) = 0$.

In this way, we establish the existence of a matrix satisfying the MWHST condition.

6.3. Geometric isolation of equilibrium points. Consider the gradient flow[2]

$$\dot{\mathbf{x}} = -\nabla F(\mathbf{x}) \tag{6.12}$$

of the objective function $F(\mathbf{x})$. Let $\mathcal{C}$ denote the set of stationary points

$$\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^n \mid \nabla F(\mathbf{x}) = 0\}. \tag{6.13}$$

We are interested in the limiting behavior of the flow $\mathbf{x}(t)$. In particular, let $\omega(\mathbf{x}(0))$ denote the set of accumulation points of the flow $\mathbf{x}(t)$ starting from $\mathbf{x}(0)$

$$\omega(\mathbf{x}(0)) := \{\mathbf{x}^* \in \mathbb{R}^n \mid \mathbf{x}(t_\nu) \to \mathbf{x}^* \text{ for some infinite sequence } t_\nu \to \infty\}. \tag{6.14}$$

We are interested in knowing the topology of $\omega(\mathbf{x}(0))$. It is well known that if $\mathbf{x}(t)$ is a bounded semi-orbit of (6.12) and if F is differentiable, then $\omega(\mathbf{x}(0))$ is a non-empty, compact, and connected subset of $\mathcal{C}$. In general, it is possible that $\omega(\mathbf{x}(0))$ can form a periodic orbit [1]. The ω-limit set of a analytic gradient flow, nonetheless, enjoys a special convergence property.

THEOREM 6.3. *([1, 50]) Suppose that $F : U \to \mathbb{R}$ is real analytic in an open set $U \subset \mathbb{R}^n$. Then for any bounded semi-orbit $\mathbf{x}(t)$ of (6.12), there exists a point $\mathbf{x}^* \in \omega(\mathbf{x}(0))$ such that $\mathbf{x}(t) \to \mathbf{x}^*$ as $t \to \infty$.*

What has happened is that the semi-orbit of any analytic gradient flow is necessarily of finite arc length. Equivalently, the set $\omega(\mathbf{x}(0))$ of any analytic gradient flow $\mathbf{x}(t)$ is necessarily a singleton. The reason of finite arc length for an analytic gradient flow is a consequence of the Łojasiewicz inequality.

THEOREM 6.4. *(Łojasiewicz Inequality [9, 39, 40]) Suppose that $F : U \to \mathbb{R}$ is real analytic in an open set $U \subset \mathbb{R}^n$. Then for any point $\mathbf{p} \in U$, there exists a neighborhood W of $\mathbf{p}$, constants $\theta \in [\frac{1}{2}, 1)$ and $c > 0$ such that*

$$\|F(\mathbf{x}) - F(\mathbf{p})\|^\theta \le c\|\nabla F(\mathbf{x})\| \quad \textit{for all } \mathbf{x} \in W. \tag{6.15}$$

For a proof of Theorem 6.3 using this inequality, see [1, Theorem 2.2] and the notes [50]. In our case, note that vector field in (6.11) is a polynomial system which obviously is analytic in Q. We thus have obtained the first property of the orbit $Q(t)$.

COROLLARY 6.5. *The projected gradient flow $Q(t)$ defined by (6.11) converges to a single point.*

6.4. Projected Hessian. The optimization problem (6.4) may have many stationary points. A stationary point in the set $\mathcal{C}$ may be a local minimizer, a local maximizer, or a saddle point. To fully classify the local behavior of its stationary points, we need to rely on the second order optimality condition — the definiteness of the projected Hessian of $F(Q)$ over the tangent space at the stationary point [20, 24]. For a general constrained optimization problem, computing the projected Hessian is theoretically desirable but practically difficult. For problem (6.4), nonetheless, we can compute the projected Hessian explicitly by the technique developed in our earlier work [10].

The procedure goes as follows. First, we formally extend the projected gradient $g(Q)$ defined in (6.11) for $Q \in \mathcal{O}(n)$ to the function $G : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ for general Z via the definition

$$G(Z) := Z\left[Z^\top SZ, \eta(Z)\right]. \tag{6.16}$$

Note that $G(Z)$ is only a mechanical generalization of $g(Q)$. Second, we calculate action of the Fréchet derivative of G at $Z \in \mathbb{R}^{n\times n}$ on an arbitrary $H \in \mathbb{R}^{n\times n}$ as

$$G'(Z).H = H\left[Z^\top SZ, \eta(Z)\right] + Z\left[H^\top SZ + Z^\top SH, \eta(Z)\right] + Z\left[Z^\top SZ, \operatorname{diag}(H^\top SZ + Z^\top SH)\right]. \tag{6.17}$$

[2]Without causing ambiguity, we use the same notation F for a general objective function which eventually refers to the potential function whose gradient is our projected gradient $g(Q)$ in (6.11). An analytic expression of the potential function is obtainable, but not needed in our subsequent analysis.

Third, we restrict the action of (6.17) to the case where $Q \in \mathcal{O}(n)$ is a stationary point and $H \in \mathcal{T}_Q\mathcal{O}(n)$, which gives out the information of the projected Hessian.

Specifically, we know that Q is a stationary point of (6.4) if and only if $\left[Q^\top SQ, \eta(Q)\right] = 0$ by (6.10). We know also that $H \in \mathcal{T}_Q\mathcal{O}(n)$ if and only if H is of the form $H = QK$ for some skew-symmetric matrix $K \in \mathbb{R}^{n\times n}$. So, upon substitution and simplification, the projected Hessian $G'(Q)$ acting on QK is given by

$$\begin{aligned}\langle QK, G'(Q).QK\rangle &= \left\langle QK, Q\left[\left[Q^\top SQ, K\right], \eta(Q)\right] + Q\left[Q^\top SQ, \operatorname{diag}\left[Q^\top SQ, K\right]\right]\right\rangle \\ &= \left\langle K, \left[\left[Q^\top SQ, K\right], \eta(Q)\right] + \left[Q^\top SQ, \operatorname{diag}\left[Q^\top SQ, K\right]\right]\right\rangle \\ &= \left\langle \left[Q^\top SQ, K\right], [K, \eta(Q)] + \operatorname{diag}\left[Q^\top SQ, K\right]\right\rangle. \end{aligned} \tag{6.18}$$

With the projected Hessian (6.18) in hand, the following theorem is simply the standard second-order optimality condition from classical optimization [20, 24] applied to our problem.

LEMMA 6.6. *Suppose that $Q \in \mathcal{O}(n)$ is a stationary point for (6.4). Then a necessary condition for Q to be a local minimizer is that*

$$\langle QK, G'(Q).QK\rangle \geq 0 \quad \textit{for all skew-symmetric matrices } K. \tag{6.19}$$

If the strict inequality in (6.19) holds at Q, then Q is guaranteed to be a local minimizer.

6.5. Asymptotically stable equilibrium. We now apply the above theory to establish the existence of a matrix satisfying the MWHST condition. As explained earlier, we need to find an orthogonal matrix Q such that $\eta(Q) = 0$. We argue by contradiction, namely, if $\eta(Q) \neq 0$ for a stationary point Q, then there is a direction along which the value of the objective function F can be further reduced. As such, our gradient flow $Q(t)$ will bypass[3] this stationary point and continue to descend until a local minimizer at which $\eta(Q) = 0$ is found.

We first make the following simple claim.

LEMMA 6.7. *Suppose that the three given sets of data $\mathbf{d}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\sigma}$ satisfy the MWHST condition. Suppose also that a matrix A satisfying the Weyl-Horn condition is already found and is not skew-symmetric. Let S be the symmetric matrix defined in (6.3). Then for any $Q \in \mathcal{O}(n)$, $\eta(Q)$ cannot be a constant diagonal unless it is identically zero.*

Proof. Suppose that $\eta(Q) = \operatorname{diag}(Q^\top SQ) - \operatorname{diag}(\mathbf{d}) = cI$ for some constant c. Summing over the diagonal entries, by the Mirsky condition, we have $\operatorname{trace}(\eta(Q)) = \operatorname{trace}(\operatorname{diag}(Q^\top AQ) - \sum_{i=1}^n d_i) = \sum_{i=1}^n \lambda_i - \sum_{i=1}^n d_i = 0 = nc$. It follows that $c = 0$. ◻

For simplicity, we assume the generic situation that all eigenvalues of S are distinct. The analysis for the case of equal eigenvalues is more involved[4], but the asymptotic behavior should be similar. By using the gradient flow, we now prove our major result on the solvability.

THEOREM 6.8. *Suppose that the symmetric matrix S has distinct eigenvalues. Let $Q \in \mathcal{O}(n)$ denote a stationary point for the problem (6.4). If $\eta(Q) \neq 0$, then there exists a skew-symmetric matrix K such that $\langle QK, G'(Q)QK\rangle < 0$.*

Proof. Suppose that $\eta(Q) \neq 0$. For the simplicity of describing the structure only [5], we may assume without loss of generality that $\eta(Q)$ is of the form

$$\eta(Q) = \operatorname{diag}\{\eta_1 I_{n_1}, \cdots, \eta_k I_{n_k}\}, \tag{6.20}$$

where I_{n_i} is the $n_i \times n_i$ identity matrix for $i = 1, \cdots, k$, and $\eta_1 > \cdots > \eta_k$. It is important to note that $k > 1$ because, by Lemma 6.7, $\eta(Q)$ must have more than one diagonal block.

[3]Unless $Q(t)$ happens to stay on a heteroclinic orbit, which is numerically unlikely due to the ubiquitous floating-point arithmetic errors.

[4]In this case, we still have $[\Pi, V] = 0$ as in the proof of Theorem 6.8. If Π has repeated entries, then V is block diagonal, but still orthogonally similar to the diagonal matrix $\eta(Q)$. From this point on, the same idea in the proof carries through with a little bit manipulation of block forms. See our numerical example in Section 7.

[5]It is easier to describe the structure in block form. The proof is still valid without the block form, except that we need to take extra efforts to describe the rows and columns corresponding to the same η_i for each $i = 1, \ldots, k$.

At such a stationary point Q, write $Y := Q^\top SQ$ for abbreviation. By (6.10), we have the commutativity $[Y, \eta(Q)] = 0$. It follows that Y must be of block diagonal form

$$Y = \operatorname{diag}\{Y_{11}, \cdots, Y_{kk}\}, \tag{6.21}$$

where Y_{ii}, $i = 1, \cdots, k$, is an $n_i \times n_i$ symmetric matrix.

Let the spectral decomposition of the symmetric matrix S be denoted by $S = U^\top \Pi U$. By rearranging the columns of orthonormal eigenvectors in $U^\top$ if necessary, we may write the diagonal matrix of eigenvalues as $\Pi = \operatorname{diag}\{\pi_1, \ldots, \pi_n\}$ with

$$\pi_1 > \ldots > \pi_n. \tag{6.22}$$

Define the matrix

$$V := (UQ)\eta(Q)(UQ)^\top. \tag{6.23}$$

Then $[\Pi, V] = 0$.

On one hand, since the diagonal matrix Π has distinct diagonal entries, V must also be a diagonal matrix, implying that (6.23) is a spectra decomposition of $\eta(Q)$ with columns of $(UQ)^\top$ as the eigenvectors of $\eta(Q)$. Write $V = \operatorname{diag}\{v_1, \cdots, v_n\}$. Because $\eta(Q)$ is itself a diagonal matrix, the set $\{v_1, \cdots, v_n\}$ is composed of exactly the diagonal entries of $\eta(Q)$.

On the other hand, because of the block structure specified in (6.20), the orthogonal matrix $(UQ)^\top$ must also be block structured accordingly with sizes $n_1, \ldots, n_k$, respectively. In each block the diagonal entries of $\eta(Q)$ is constant. The similarity transformation by UQ within that block therefore has no effect to (6.23). It follows that $V = \eta(Q)$. In particular, $\{v_1, \ldots, v_n\}$ must be in the ordering as

$$\underbrace{v_1 = \ldots = v_1}_{n_1 \text{ times}} > \underbrace{v_2 = \ldots = v_2}_{k_2 \text{ times}} > \ldots > \underbrace{v_k = \ldots v_k}_{n_k \text{ times}}. \tag{6.24}$$

Let $K \in \mathbb{R}^{n\times n}$ denote an arbitrary skew-symmetric matrix. Clearly, $\operatorname{diag}\left[Q^\top SQ, K\right] = 0$. With respect to this matrix K and at the stationary point Q, the projected Hessian (6.18) becomes

$$\begin{aligned} \langle QK, G'(Q).QK\rangle &= \left\langle \left[Q^\top SQ, K\right], [K, \eta(Q)]\right\rangle = -\left\langle V\tilde{K} - \tilde{K}V, \Pi\tilde{K} - \tilde{K}\Pi\right\rangle \\ &= -2\sum_{i<j}(\pi_i - \pi_j)(v_i - v_j)\tilde{k}_{ij}^2, \end{aligned} \tag{6.25}$$

where $\tilde{K} = \left[\tilde{k}_{ij}\right] := (UQ)K(UQ)^\top$ remains to be skew-symmetric since UQ is orthogonal. It is obvious from (6.22) and (6.24) that we may choose appropriate values of $\tilde{k}_{ij}$ such that $\langle QK, G'(Q).QK\rangle < 0$. □

COROLLARY 6.9. *Under the assumption of Theorem 6.8, if Q is a stationary point with $\eta(Q) \neq 0$, then Q is not a local minimizer for the objective function F in (6.4).*

Indeed, such a point is an unstable equilibrium for the gradient dynamics. There exists at least one tangent direction, i.e., a matrix QK with a certain skew-symmetric matrix K, along which $F(Q)$ can be further decreased. Therefore, the flow must continue until an isolated limit point at which $\eta(Q) = 0$ is found. Based on this understanding, we conclude that the convergence of $Q(t)$ to an asymptotically stable equilibrium point Q at which $\eta(Q) = 0$ is guaranteed.

COROLLARY 6.10. *A local minimum for the objective function F in (6.4) is a global minimum.*

When this limit point Q is achieved, we use this Q to form the corresponding matrix $Q^\top AQ$ which now maintain the prescribed diagonal entries, eigenvalues, and singular values. The existence of a matrix satisfying the MWHST condition is hereby established.

7. Numerical example. The proof used to derive the theoretical existence can actually be implemented as a numerical scheme, albeit it might require additional effort to tune its efficiency. [26, 32]. For the time being, our numerical example is meant only to demonstrate workability of the differential system (6.11), which is the basis of our existence proof. To demonstrate the robustness of the approach, we even challenge ourselves with a case where multiple eigenvalues, near eigenvalues, and high ill-conditioning are all presented.

For convenience, we choose to use the standard routine `ode15s` from MATLAB as our integrator. The local error tolerance is set at $\mathsf{AbsTol} = \mathsf{RelTol} = 10^{-10}$. Consider the 8×8 Rosser matrix R with integer elements

$$R = \begin{bmatrix} 611 & 196 & -192 & 407 & -8 & -52 & -49 & 29 \\ 196 & 899 & 113 & -192 & -71 & -43 & -8 & -44 \\ -192 & 113 & 899 & 196 & 61 & 49 & 8 & 52 \\ 407 & -192 & 196 & 611 & 8 & 44 & 59 & -23 \\ -8 & -71 & 61 & 8 & 411 & -599 & 208 & 208 \\ -52 & -43 & 49 & 44 & -599 & 411 & 208 & 208 \\ -49 & -8 & 8 & 59 & 208 & 208 & 99 & -911 \\ 29 & -44 & 52 & -23 & 208 & 208 & -911 & 99 \end{bmatrix}.$$

The matrix R is known for its difficulty in that it has a double eigenvalue, three nearly equal eigenvalues, a zero eigenvalue, two dominant eigenvalues of opposite sign and a small nonzero eigenvalue. We use its diagonal entries $\mathbf{d}$ and the *computed* eigenvalues and singular values

$$\boldsymbol{\lambda} = \begin{bmatrix} -1.020049018429997e{+}03 \\ 1.020049018429997e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359278e{+}03 \\ 1.000000000000001e{+}03 \\ 9.999999999999998e{+}02 \\ 9.804864072152601e{-}02 \\ 4.851119506099622e{-}13 \end{bmatrix}, \boldsymbol{\sigma} = \begin{bmatrix} 1.020049018429997e{+}03 \\ 1.020049018429996e{+}03 \\ 1.020000000000000e{+}03 \\ 1.019901951359279e{+}03 \\ 1.000000000000000e{+}03 \\ 9.999999999999998e{+}02 \\ 9.804864072162672e{-}02 \\ 1.054603342667098e{-}14 \end{bmatrix}$$

as the test data, so the MWHST condition is automatically satisfied[6]. This example also serves to demonstrate the case that even though the matrix S in Theorem 6.8 does have multiple eigenvalues, the gradient flow still works.

Using the recursive algorithm proposed in [12] for the above $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$, we obtain first the following matrix A needed in Section 6.1:

$$A = \begin{bmatrix} 1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.0200e{+}03 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0199e{+}03 & 0 & 0 & 1.4668e{-}09 & 0 \\ 0 & 0 & 0 & 0 & 1.0000e{+}03 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0000e{+}03 & 0 & 0 \\ 0 & 0 & 0 & -1.5257e{-}05 & 0 & 0 & 9.8049e{-}02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.4045e{-}07 & 0 \end{bmatrix},$$

where for the convenience of running text we display all numbers in only 5 digits. Note that A is not symmetric. Defining S according to (6.3) and integrating our differential equation (6.11) numerically, we are able to find this matrix

$$B = \begin{bmatrix} 611.0000 & -184.6972 & 93.4026 & -177.9840 & 50.0103 & -66.4451 & -364.8367 & -230.1007 \\ -184.6972 & 899.0000 & 84.6701 & -18.4713 & 86.7552 & -70.4723 & -143.9197 & -161.0699 \\ 93.4026 & 84.6701 & 899.0000 & -136.2380 & -152.2550 & 106.6705 & -2.4645 & 191.7090 \\ -177.9840 & -18.4713 & -136.2380 & 611.0000 & -282.3676 & 196.8994 & -260.5891 & 85.8062 \\ 50.0103 & 86.7552 & -152.2550 & -282.3676 & 411.0000 & 592.8768 & 367.9171 & -228.2314 \\ -66.4451 & -70.4723 & 106.6705 & 196.8994 & 592.8768 & 411.0000 & -481.0590 & 348.2056 \\ -364.8367 & -143.9197 & -2.4645 & -260.5891 & 367.9171 & -481.0590 & 99.0000 & 459.5032 \\ -230.1007 & -161.0699 & 191.7090 & 85.8062 & -228.2314 & 348.2056 & 459.5032 & 99.0000 \end{bmatrix}.$$

It can been seen that the diagonal entries of B are almost identical to those of R. Indeed, the total difference is within a 2-norm of 2×10^{-9} in absolute error. Likewise, we can check that the eigenvalues and singular values of B agree with $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$ within the tolerance 10^{-9}. However, note that the off-diagonal entries of B are very different from those of the original R, indicating that the inverse problem has multiple solutions.

[6]Strictly speaking, the MWHST condition is satisfied only up to the machine precision. See, for example, the smallest computed eigenvalue is not exactly zero.

8. Conclusion. We studied the theoretical problem of whether a matrix with prescribed main diagonal entries, eigenvalues, and singular values exists when these data satisfy the equalities and inequalities entailed by the Mirsky, Weyl-Horn, and Sing-Thompson theorems simultaneously. We employ an argument involving an array of tools to establish the existence of such a matrix when $n \geq 3$. Extra conditions are needed for the case $n = 2$. The existence theory is new in the field. The dynamical system approach might be the first tool for constructing such a matrix.

REFERENCES

[1] P.-A. ABSIL, R. MAHONY, AND B. ANDREWS, *Convergence of the iterates of descent methods for analytic cost functions*, SIAM J. Optim., 16 (2005), pp. 531–547.

[2] T. ANDO, *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Linear Algebra Appl., 118 (1989), pp. 163–248.

[3] A. L. ANDREW, *Some recent developments in inverse eigenvalue problems*, Computational Techniques and Applications: CTAC93, ed. D. Stewart, H. Gardner and D. Singleton, (1994), pp. 94–102.

[4] R. W. BROCKETT, *Using feedback to improve system identification*, in Control of uncertain systems: modelling, approximation, and design, vol. 329 of Lect. Notes Control Inf. Sci., Springer, Berlin, 2006, pp. 45–65.

[5] E. A. CARLEN AND E. H. LIEB, *Short proofs of theorems of Mirsky and Horn on diagonals and eigenvalues of matrices*, Electron. J. Linear Algebra, 18 (2009), pp. 438–441.

[6] K. CHADAN, D. COLTON, L. PÄIVÄRINTA, AND W. RUNDELL, *An introduction to inverse scattering and inverse spectral problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[7] N. N. CHAN AND K. H. LI, *Diagonal elements and eigenvalues of a real symmetric matrix*, J. Math. Anal. Appl., 91 (1983), pp. 562–566. See also *Algorithm 115: A FORTRAN subroutine for finding a real symmetric matrix with prescribed diagonal elements and eigenvalues* in Algorithms Supplement, The Computer Journal (1983) 26(2): 184-186.

[8] R. H. CHAN, Z.-J. BAI, AND B. MORINI, *On the convergence rate of a Newton-like method for inverse eigenvalue and inverse singular value problems*, Int. J. Appl. Math., 13 (2003), pp. 59–69. OR and models (Hong Kong, 2003).

[9] R. CHILL, *On the Łojasiewicz-Simon gradient inequality*, J. Funct. Anal., 201 (2003), pp. 572–601.

[10] M. T. CHU, *Solving additive inverse eigenvalue problems for symmetric matrices by the homotopy method*, IMA J. Numer. Anal., 10 (1990), pp. 331–342.

[11] ———, *On constructing matrices with prescribed singular values and diagonal elements*, Linear Algebra Appl., 288 (1999), pp. 11–22.

[12] ———, *A fast recursive algorithm for constructing matrices with prescribed eigenvalues and singular values*, SIAM J. Numer. Anal., 37 (2000), pp. 1004–1020 (electronic).

[13] M. T. CHU AND M. A. ERBRECHT, *Symmetric Toeplitz matrices with two prescribed eigenpairs*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 623–635.

[14] M. T. CHU AND G. H. GOLUB, *Inverse eigenvalue problems: theory, algorithms, and applications*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.

[15] M. T. CHU AND J. W. WRIGHT, *The education testing problem revisited*, IMA J. Numer. Anal., 15 (1995), pp. 141–160.

[16] P. COTAE AND M. AGUIRRE, *On the construction of the unit tight frames in code division multiple access systems under total squared correlation criterion*, {AEU} - International Journal of Electronics and Communications, 60 (2006), pp. 724 – 734.

[17] P. I. DAVIES AND N. J. HIGHAM, *Numerically stable generation of correlation matrices and their factors*, BIT, 40 (2000), pp. 640–651.

[18] I. S. DHILLON, R. W. HEATH, JR., M. A. SUSTIK, AND J. A. TROPP, *Generalized finite algorithms for constructing Hermitian matrices with prescribed diagonal and spectrum*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 61–71 (electronic).

[19] I. ELISHAKOFF, *A selective review of direct, semi-inverse and inverse eigenvalue problems for structures described by differential equations with variable coefficients*, Arch. Comput. Methods Engrg., 7 (2000), pp. 451–526.

[20] R. FLETCHER, *Practical methods of optimization*, Wiley-Interscience [John Wiley & Sons], New York, second ed., 2001.

[21] G. FREILING AND V. YURKO, *Inverse Sturm-Liouville problems and their applications*, Nova Science Publishers, Inc., Huntington, NY, 2001.

[22] W. FULTON, *Eigenvalues, invariant factors, highest weights, and Schubert calculus*, Bull. Amer. Math. Soc. (N.S.), 37 (2000), pp. 209–249.

[23] I. M. GEL′FAND AND B. M. LEVITAN, *On the determination of a differential equation from its spectral function*, Amer. Math. Soc. Transl. (2), 1 (1955), pp. 253–304.

[24] P. E. GILL, W. MURRAY, AND M. H. WRIGHT, *Practical optimization*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981.

[25] G. M. L. GLADWELL, *Inverse problems in vibration*, vol. 119 of Solid Mechanics and its Applications, Kluwer Academic Publishers, Dordrecht, second ed., 2004.

[26] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric numerical integration*, vol. 31 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.

[27] O. H. HALD, *On discrete and numerical inverse Sturm-Liouville problems*, PhD thesis, New York University, 1972.

[28] T. HIROSHIMA, *Majorization criterion for distillability of a bipartite quantum state*, Phys. Rev. Lett., 91 (2003), p. 057902.

[29] A. HORN, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math., 76 (1954), pp. 620–630.

[30] R. HORODECKI, P. HORODECKI, M. HORODECKI, AND K. HORODECKI, *Quantum entanglement*, Rev. Mod. Phys., 81 (2009), pp. 865–942.

[31] K. D. IKRAMOV AND V. N. CHUGUNOV, *Inverse matrix eigenvalue problems*, J. Math. Sci. (New York), 98 (2000), pp. 51–136. Algebra, 9.

[32] A. ISERLES, H. Z. MUNTHE-KAAS, S. P. NØRSETT, AND A. ZANNA, *Lie-group methods*, in Acta numerica, 2000, vol. 9 of Acta Numer., Cambridge Univ. Press, Cambridge, 2000, pp. 215–365.

[33] M. JABBOUR, *Majorization relations in quantum information theory*, master's thesis, Ecole Polytechnique de Bruxelles, 2013.

[34] E. JORSWIECK AND H. BOCHE, *Majorization and Matrix-monotone Functions in Wireless Communications*, Found. Trends Commun. Inf. Theory, Now Publishers Inc., 2007.

[35] S. I. KABANIKHIN, *Inverse and ill-posed problems*, vol. 55 of Inverse and Ill-posed Problems Series, Walter de Gruyter GmbH & Co. KG, Berlin, 2012. Theory and applications.

[36] P. KOSOWSKI AND A. SMOKTUNOWICZ, *On constructing unit triangular matrices with prescribed singular values*, Computing, 64 (2000), pp. 279–285.

[37] B. M. LEVITAN, *Inverse Sturm-Liouville problems*, VSP, Zeist, 1987. Translated from the Russian by O. Efimov.

[38] C.-K. LI AND Y.-T. POON, *Sum of Hermitian matrices with given eigenvalues: inertia, rank, and multiple eigenvalues*, Canad. J. Math., 62 (2010), pp. 109–132.

[39] S. ŁOJASIEWICZ, *Une propriété topologique des sous-ensembles analytiques réels*, in Les Équations aux Dérivées Partielles (Paris, 1962), Éditions du Centre National de la Recherche Scientifique, Paris, 1963, pp. 87–89.

[40] S. ŁOJASIEWICZ AND M.-A. ZURRO, *On the gradient inequality*, Bull. Polish Acad. Sci. Math., 47 (1999), pp. 143–145.

[41] D. LONDON AND H. MINC, *Eigenvalues of matrices with prescribed entries*, Proc. Amer. Math. Soc., 34 (1972), pp. 8–14.

[42] W. MA AND Z.-J. BAI, *A regularized directional derivative-based Newton method for inverse singular value problems*, Inverse Problems, 28 (2012), pp. 125001, 24.

[43] G. MARSAGLIA AND I. OLKIN, *Generating correlation matrices*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 470–475.

[44] A. W. MARSHALL AND I. OLKIN, *Inequalities: theory of majorization and its applications*, vol. 143 of Mathematics in Science and Engineering, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.

[45] L. MIRSKY, *Matrices with prescribed characteristic roots and diagonal elements*, J. London Math. Soc., 33 (1958), pp. 14–21.

[46] M. A. NIELSEN, *An introduction to majorization and its applications to quantum mechanics*. Lecture notes, 2002. available at `http://michaelnielsen.org/blog/talks/2002/maj/book.ps`.

[47] M. A. NIELSEN AND I. L. CHUANG, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000.

[48] M. R. OSBORNE, *On the inverse eigenvalue problem for matrices and related problems for difference and differential equations*, in Conference on Applications of Numerical Analysis (Univ. Dundee, Dundee, 1971), Springer, Berlin, 1971, pp. 155–168. Lecture Notes in Math., Vol. 228.

[49] J. PAINE, *A numerical method for the inverse Sturm-Liouville problem*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 149–156.

[50] M. PIERRE, *Quelques applications de l'inégalit´e de Lojasiewicz à des discrétisations d'EDP*. SMAI, 2011. available at `http://smai.emath.fr/smai2011/slides/mpierre/Slides.pdf`.

[51] J. D. PRYCE, *Numerical solution of Sturm-Liouville problems*, Monographs on Numerical Analysis, The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[52] R. RAO AND S. DIANAT, *Basics of code division multiple access (CDMA)*, Bellingham, Wash.: SPIE Press, 2005.

[53] V. RUSSO, *Majorization criterion for undistillable quantum states*. Lecture notes, 2013. available at `https://cs.uwaterloo.ca/~vrusso/projects/Majorization_Notes.pdf`.

[54] F. Y. SING, *Some results on matrices with prescribed diagonal elements and singular values*, Canad. Math. Bull., 19 (1976), pp. 89–92.

[55] G. W. STEWART, *On the early history of the singular value decomposition*, SIAM Rev., 35 (1993), pp. 551–566.

[56] R. C. THOMPSON, *Singular values, diagonal elements, and convexity*, SIAM J. Appl. Math., 32 (1977), pp. 39–63.

[57] J. A. TROPP, I. S. DHILLON, AND R. W. HEATH, *Finite-step algorithms for constructing optimal cdma signature sequences*, IEEE Transactions on Information Theory, 50 (2004), pp. 2916–2921.

[58] S.-W. VONG, Z.-J. BAI, AND X.-Q. JIN, *An Ulm-like method for inverse singular value problems*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 412–429.

[59] H. WEYL, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. U. S. A., 35 (1949), pp. 408–411.

[60] S.-J. WU AND M. M. LIN, *Numerical methods for solving nonnegative inverse singular value problems with prescribed structure*, Inverse Problems, 30 (2014), pp. 055008, 14.

[61] S. YUAN, A. LIAO, AND G. YAO, *Parameterized inverse singular value problem for anti-bisymmetric matrices*, Numer. Algorithms, 60 (2012), pp. 501–522.

[62] H. ZHA AND Z. ZHANG, *A note on constructing a symmetric matrix with specified diagonal entries and eigenvalues*, BIT, 35 (1995), pp. 448–452.

[63] L. A. ZHORNITSKAYA AND V. S. SEROV, *Inverse eigenvalue problems for a singular Sturm-Liouville operator on* $[0, 1]$, Inverse Problems, 10 (1994), pp. 975–987.