



OPTIMIZATION SCHEMES ON MANIFOLDS FOR STRUCTURED MATRICES WITH FIXED EIGENVALUES

Jean-Paul Chehab, Harry Oviedo, Marcos Raydan

► To cite this version:

Jean-Paul Chehab, Harry Oviedo, Marcos Raydan. OPTIMIZATION SCHEMES ON MANIFOLDS FOR STRUCTURED MATRICES WITH FIXED EIGENVALUES. 2023. hal-04239863v1

HAL Id: hal-04239863

<https://hal.science/hal-04239863v1>

Preprint submitted on 12 Oct 2023 (v1), last revised 3 Apr 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMIZATION SCHEMES ON MANIFOLDS FOR STRUCTURED MATRICES WITH FIXED EIGENVALUES

JEAN-PAUL CHEHAB*, HARRY OVIEDO†, AND MARCOS RAYDAN‡

Abstract. Several manifold optimization schemes are presented and analyzed for solving inverse structured symmetric matrix problems with prescribed spectrum. Some entries in the desired matrix are assigned in advance and cannot be altered, and some others should be nonzero. The rest of the entries are free. The reconstructed matrix must satisfy these requirements and its eigenvalues must be the given ones. This inverse eigenvalue problem is related to the problem of determining the graph, with weights on the undirected edges, of the matrix associated with its sparse pattern. Our optimization schemes are based on considering the eigenvector matrix as the only unknown and moving iteratively on the manifold of orthogonal matrices, forcing the additional structural requirements through a change of variables and a convenient differentiable objective function in the space of square matrices. We propose Riemannian gradient-type methods combined with either a penalization or an augmented Lagrangian strategy, and two different well-known retractions. We also present a block alternating technique that takes advantage of a proper separation of variables. Convergence properties of the penalty alternating approach are established. Finally, we present initial numerical results to demonstrate the effectiveness of our proposals.

Keywords: Riemannian optimization, Stiefel manifold, inverse eigenvalue problems, spectral graph theory, augmented Lagrangian, alternating direction method of multipliers

AMS Subject Classification: 15A29, 65F18, 15A83, 15B10, 05C50, 90C30

1. Introduction. Given a set of real eigenvalues, say $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, that defines a fixed diagonal matrix D whose diagonal entries are the given λ_i s, the problem of interest is to find a real symmetric $n \times n$ matrix A with additional structural requirements, such that its eigenvalues are the given ones. In other words, we need to find A such that $A = XDX^\top$ (Schur factorization) for an orthogonal matrix X whose columns are the eigenvectors of A .

There are, of course, some additional and fundamental requirements that add difficulty to the problem, otherwise $A = D$ is a valid solution. For some specified pairs $(i, j) \in I_0$ the a_{ij} entries must be zero at the solution. In addition, for some other specified pairs $(i, j) \in I_{fix}$ the a_{ij} entries must have some given (nonzero) values at the solution. For some other specified pairs $(i, j) \in I_{nz}$ the a_{ij} entries must be nonzero (any nonzero value is fine) at the solution. Note that the entries a_{ij} for $(i, j) \in I_{nz}$ are free, as long as they are not zero. Finally, for some other specified pairs $(i, j) \in I_{free}$ the a_{ij} entries are free variables that can take any real value. Here, the four given sets of indices: I_0, I_{fix}, I_{nz} , and I_{free} must form a partition of all possible pairs (i, j) , with i and j in $\{1, 2, 3, \dots, n\}$.

We note that all the given eigenvalues can be moved to the positive side of the real line by a constant shift $\alpha > 0$ for a sufficiently large α . In that case, the new positive eigenvalues correspond to the eigenvalues of $A + \alpha I$, that possesses the same eigenvectors of A . Once the problem is solved we can always return to the original matrix A by subtracting αI . Hence, without any loss of generality we can assume that the given eigenvalues are positive and that the unknown A is Symmetric and Positive Definite (SPD).

*Laboratoire LAMFA (UMR CNRS 7352), Université de Picardie Jules Verne, 33 rue Saint Leu, 80039 Amiens Cédex, France (Jean-Paul.Chehab@u-picardie.fr).

†Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diag. Las Torres 2640, Santiago de Chile, Chile (harry.oviedo@uai.cl).

‡Center for Mathematics and Applications (NovaMath), FCT NOVA, 2829-516 Caparica, Portugal (m.raydan@fct.unl.pt).

The desired pattern, described above for the unknown symmetric matrix A , can be associated with a weighted undirected graph with n vertices, in which two vertices i and j are connected (adjacent) if and only if a_{ij} is not zero. The described inverse problem and the connection between the undirected graph and the matrix A with preset eigenvalues appear in several real-life applications related to quantum chemistry, information theory, social networks, and combinatorial optimization; see, e.g., [7, 8, 32, 33] and the references cited therein. It is worth noting that not every arbitrary assignment of real eigenvalues can guarantee the existence of a certain structure of the associated symmetric matrix. This is a topic of recent interest in matrix analysis. For a unified theoretical development of whether a predetermined distribution of eigenvalues can be achieved with a given graph, and also how the graph of a symmetric matrix influences the possible multiplicities of its eigenvalues, we refer to the book by Johnson and Saito [16].

In a recent work by Sutton [31], a few examples of the considered problem were fully described, including the desired matrix pattern and the associated graph. Moreover, an optimization scheme (based on off-the-shelf packages) to approximate the matrix A was also outlined. To the best of our knowledge, Sutton’s article is the first to describe in detail the general problem of finding a matrix with a certain default structure, including some fixed values at certain positions, and with predetermined eigenvalues. In addition, in [31] some numerical optimization ideas are indicated, such as the use of augmented Lagrangian strategies and manifold techniques. His pioneering work motivated us to study these problems in depth in order to propose a suitable objective function and to develop in detail robust alternating iterative optimization methods, for which we present a convergence analysis.

The rest of this document is organized as follows. In Section 2, we present the proposed optimization model that considers the eigenvector matrix X as the only unknown, such that the orthogonality of X is the only constraint that needs to be imposed on the optimization problem. The rest of the structural matrix requirements are included in the objective function using a penalization strategy. In Section 3, we describe in detail the proposed iterative method for solving the orthogonality constrained optimization problem. In particular, we apply a fast and robust Riemannian gradient scheme, using two well-understood retractions and a suitable non-monotone globalization strategy. To enhance the practical behavior of the overall procedure, in Section 4, we introduce a block alternating minimization algorithm that takes advantage of a convenient variable splitting. The proposed procedure iterates alternatively solving an optimization problem with orthogonality constraints, using the development of Section 3, and a linearly constrained quadratic problem for which a closed formula is obtained. In Section 5, we establish the convergence properties of the algorithm described in Section 4 under some mild assumptions. Furthermore, in Section 6 we develop an alternating direction scheme that, instead of using a penalization strategy, uses an augmented Lagrangian approach. In Section 7, we report the obtained numerical results, and give further insights into the proposed schemes. Finally, in Section 8 we present some concluding remarks.

2. Manifold approach and penalization strategy. The main idea is to minimize a conveniently defined function that only depends on the matrix of eigenvectors X for solving the problem described in Section 1. Since the matrix of eigenvectors must be orthogonal, the only constraint to be considered is $X^\top X = I$. This feasible set is the well-known Stiefel manifold that has been recently analyzed and for which a variety of effective iterative methods have been proposed; see, e.g., [15, 22, 24, 26].

Now, to consider only X as the variable matrix and to incorporate all the requirements described in Section 1, we need to introduce several suitable matrix operators. In here, the Hadamard product $A \circ B$ between matrices A and B of the same dimensions will play a key role. The main objective is to force the eigenvalues to be the given ones by building up a convenient differentiable objective function that guarantees that the eigenvalues of any of the iterates are always the diagonal entries of the given diagonal matrix D .

Let $P \in \mathbb{R}^{n \times n}$ be the matrix such that its entry (i, j) is given by

$$P_{ij} = \begin{cases} 0 & \text{if } (i, j) \in I_0, \\ 1 & \text{else} \end{cases}$$

This square matrix will be used to force the zero entries in the set of indices I_0 . Similarly, we need a matrix Q with entries 1 if $(i, j) \in I_{nz}$, and zero elsewhere. The matrix Q will be used to maximize those entries in the obtained iterates. We also need a matrix R with entries 1 if $(i, j) \in I_{fix}$, and zero elsewhere. The matrix R will be used to force the given specified nonzero entries, that will be represented by C , i.e., $C_{ij} = A_{ij}$ if the pair (i, j) is in I_{fix} , and zero elsewhere. All these matrices can be seen as pattern matrices, and clearly they satisfy: $P \circ P = P$, $R \circ C = C$, $Q \circ Q = Q$, and $R \circ R = R$. These properties will be used to establish the gradient expression of the objective function to be considered.

Let us consider:

- $\Phi : X \in \mathbb{R}^{n \times n} \mapsto X^\top DX$
- $\Psi : X \in \mathbb{R}^{n \times n} \mapsto \|P \circ (\Phi(X)) - \Phi(X)\|_F^2$
- $\Omega : X \in \mathbb{R}^{n \times n} \mapsto -\|Q \circ (\Phi(X))\|_F^2$
- $\Theta : X \in \mathbb{R}^{n \times n} \mapsto \|R \circ (\Phi(X)) - C\|_F^2$

Hence, our optimization problems is given by

$$\min_{X^\top X = I} F(X), \quad (2.1)$$

where

$$F(X) = -\|Q \circ (\Phi(X))\|_F^2 + \|P \circ (\Phi(X)) - \Phi(X)\|_F^2 + \|R \circ (\Phi(X)) - C\|_F^2. \quad (2.2)$$

Our next result is dedicated to obtain the gradient of F .

LEMMA 2.1. *It follows that*

- $\mathcal{D}\Phi(X)(Z) = X^\top DZ + Z^\top DX$.
- $\nabla(\Omega(X)) = -2(DX(Q \circ (X^\top DX)) + Q \circ (X^\top DX)X^\top D)$
- $\nabla(\Psi(X)) = DX(X^\top DX - P \circ (X^\top DX)) + (X^\top DX - P \circ (X^\top DX))X^\top D$
- $\nabla(\Theta(X)) = 2[DX(R \circ (X^\top DX) - C) + (R \circ (X^\top DX) - C)X^\top D]$.

Proof. The function $\Phi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ defined above is differentiable and its Fréchet derivative along Z , $\mathcal{D}\Phi(X)(Z)$, is identified as the second term of the right hand side of the following equality (see, e.g., [1, Section A.5]):

$$\Phi(X + Z) = \Phi(X) + \mathcal{D}\Phi(X)(Z) + o(\|Z\|).$$

Since $\Phi(X + Z) = \Phi(X) + X^\top DZ + Z^\top DX + Z^\top DZ$, then $\mathcal{D}\Phi(X)(Z) = X^\top DZ + Z^\top DX$.

The function $\Psi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ defined above is differentiable and its gradient can be obtained as follows. Consider the auxiliary function $g : \mathbb{R} \rightarrow \mathbb{R}$, given by

$$g(t) = \Psi(X + tZ),$$

for any arbitrary matrix Z . From basic calculus we know that $g'(0) = \langle \nabla \Psi(X), Z \rangle_F$. After simple algebraic manipulations, using that $D^\top = D$, $P^\top = P$ and $P \circ P = P$, we obtain

$$\begin{aligned} g'(0) &= \langle X^\top DZ + Z^\top DX, X^\top DX - P \circ (X^\top DX) \rangle_F \\ &= \langle Z, DXW \rangle_F + \langle Z, WX^\top D \rangle_F, \end{aligned}$$

where $W = X^\top DX - P \circ (X^\top DX)$, and the expression of the gradient of $\Psi(X)$ is obtained. In a similar way, the expressions for $\nabla(\Omega(X))$ and $\nabla(\Theta(X))$ can also be established. \square

We note that the computational work for building the gradient at any iteration k , given the matrix X_k , requires the products DX_k and $X_k^\top D$, and from any of them we can obtain $X_k^\top DX_k$. Using those 3 matrices, that appear several times each, the entire gradient matrix is obtained, which is nothing but the summation of the last 3 terms in Lemma 2.1.

In order to guarantee the zero entries in the set of indices I_0 (given by the pattern matrix P), as well as the given specified nonzero entries in the set of indices I_{fix} (given by the matrix C and the pattern matrix R), we can incorporate a penalization strategy. First, let us recall the main ideas of the classical and straightforward penalty approach for solving a constrained optimization problem. Penalty methods transform constrained optimization problems into a sequence of unconstrained subproblems, whose solutions ideally converge to a minimizer of the original optimization problem. In our case, the required constraints,

$$P \circ (\Phi(X)) = \Phi(X) \quad \text{and} \quad R \circ (\Phi(X)) = C,$$

have been already included in the objective function (2.2). Hence, our penalization strategy consists in considering, instead of (2.2), the following sequence of functions to be minimized (over the set of matrices that satisfy $X^\top X = I$):

$$F_k(X) = -\|Q \circ (\Phi(X))\|_F^2 + \mu_k \|P \circ (\Phi(X)) - \Phi(X)\|_F^2 + \sigma_k \|R \circ (\Phi(X)) - C\|_F^2, \quad (2.3)$$

where $\mu_k > 0$ and $\sigma_k > 0$ are the penalty parameters that increase at every k to penalize constraint violations. Starting from $\mu_0 > 0$ and $\sigma_0 > 0$, under some mild assumptions and some specific choices of the sequences $\{\mu_k\}$ and $\{\sigma_k\}$, it has been established that the sequence of minimizers of (2.3), over the set $X^\top X = I$, converges to a solution of (2.2) when k goes to infinity; see, e.g., [20, Secc. 12.1].

3. Optimization over the orthogonal group. In this section, we describe some iterative methods for solving the orthogonality constrained optimization problem (2.1). In particular, we briefly review two specialized Riemannian gradient methods to solve optimization problems with orthogonal constraints. In the literature, the feasible set $\mathcal{O}(n) := \{X \in \mathbb{R}^{n \times n} : X^\top X = I\}$ is known as the *orthogonal group* [1, 6, 28]. In fact this set is a Lie group and its Lie algebra is the set of $n \times n$ skew-symmetric matrices [28]. Let $X \in \mathcal{O}(n)$ be an arbitrary matrix, the tangent space of the orthogonal group at X is given by

$$T_X \mathcal{O}(n) = \{Z = XW \in \mathbb{R}^{n \times n} : W^\top = -W\} = X \mathcal{S}_{\text{skew}}(n),$$

where $\mathcal{S}_{\text{skew}}(n)$ represents the set of all $n \times n$ skew-symmetric matrices, see [1, 11, 26].

On the other hand, it is well-known that if we equip each tangent space $T_X \mathcal{O}(n)$ with the standard matrix inner product $\langle A, B \rangle = \text{tr}[A^\top B]$, then the pair $(\mathcal{O}(n), \langle \cdot, \cdot \rangle)$

forms a Riemannian manifold [1, 6]. In this context, the inner product is so-called *Riemannian metric*, see [1, 6]. Under this specific Riemannian metric, the orthogonal projection operator over the tangent space $T_X \mathcal{O}(n)$ is given by

$$P_X(V) = X \text{skew}(X^\top V) = \frac{1}{2}(V - XV^\top X),$$

where $\text{skew}(A) := \frac{1}{2}(A - A^\top)$ is the skew-symmetric part of $A \in \mathbb{R}^{n \times n}$, for details see Example 3.6.2 in [1].

In view of $\mathcal{O}(n)$ is an embedded Riemannian sub-manifold of the Euclidean space $\mathbb{R}^{n \times n}$, the Riemannian gradient of the objective function $F : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ can be easily computed as follows

$$\nabla_{\mathcal{O}(n)} F(X) = P_X(\nabla F(X)). \quad (3.1)$$

Here we are using $\nabla_{\mathcal{O}(n)} F(X)$ to denote the Riemannian gradient of F and $\nabla F(X)$ to denote the Euclidean (classical) gradient of F . The justification for formula (3.1) appears in (3.37) in [1].

The Riemannian gradient provides us a search direction that we can use to minimize the function F over the domain $\mathcal{O}(n)$, see [1, 15, 23, 24]. However, we need an additional tool to preserve the feasibility of all the iterates. For this purpose we need a mapping called *retraction* [1, 6, 23]. Roughly speaking, a retraction is a smooth function that sends vectors from $T_X \mathcal{O}(n)$ to a point on the manifold $\mathcal{O}(n)$, with a local rigidity property that preserves gradients at X . See [1, 6, 23] for a rigorous definition of retraction.

For the orthogonal group there are several retractions. In this work, we only consider two specific retractions that to our knowledge are the ones that require the least computational cost to evaluate them. The first retraction is based on the QR factorization. Let $X \in \mathcal{O}(n)$ and $Z = XW \in T_X \mathcal{O}(n)$ the first retraction is defined as follows

$$R_X^{\text{qr}}(Z) = \text{qf}(X + Z),$$

where $\text{qf}(P)$ denotes the function that sends the matrix P to the Q factor of its QR factorization such that its R factor has strictly positive diagonal entries. We emphasize that this particular QR factorization is unique.

The second considered retraction is given by,

$$R_X^{\text{cayley}}(Z) = X \left(I - \frac{1}{2}W \right) \left(I + \frac{1}{2}W \right)^{-1},$$

which is based on the Cayley transform [15]. Notice that since $W \in \mathcal{S}_{\text{skew}}(n)$, the matrix $I + \frac{1}{2}W$ is non-singular. These two retractions are described in Example 4.1.2 in [1].

The Riemannian gradient method [1, 6, 13, 25] is a line-search iterative scheme, which constructs a sequence of iterates $\{X_k\} \subset \mathcal{O}(n)$, starting at $X_0 \in \mathcal{O}(n)$, by using the following recursion

$$X_{k+1} = R_{X_k}(-\alpha_k \nabla_{\mathcal{O}(n)} F(X_k)),$$

where $\alpha_0 > 0$ is the step-size and $R_{X_k}(\cdot)$ is any retraction. Typically, the step-size α_k is computed using a backtracking strategy in such a way that it satisfies the following non-monotone globalization rule

$$F(X_{k+1}) \leq C_k - c_1 \alpha_k \|\nabla_{\mathcal{O}(n)} F(X_k)\|_F^2,$$

where $c_1 \in (0, 1)$ and $\{C_k\}$ corresponds to the Zhang-Hager non-monotone strategy [34], i.e.,

$$C_{k+1} = \frac{\eta q_k C_k + F(X_k)}{q_{k+1}}, q_{k+1} = \eta q_k + 1, \quad \eta \in (0, 1), \quad C_0 = F(X_0), \quad \text{and} \quad q_0 = 1.$$

In order to accelerate the convergence, the Barzilai-Borwein step-sizes [2, 14] are incorporated in this method; see, e.g., [13, 14, 22, 25].

4. Penalty alternating minimization methods. In this section we introduce a block alternating minimization algorithm to recover the matrix A . Using the combination of variable splitting and a penalty scheme, the proposed procedure solves the inverse eigenvalue problem by iteratively optimizing manifold problems and linearly constrained quadratic problems. In order to force the known values of A into the desired positions we reformulate the problem (2.1) as

$$\min_{X \in \mathbb{R}^{n \times n}} \bar{F}(X) \quad \text{s.t.} \quad X^\top X = I, \quad P_{I_{\text{knw}}}[\Phi(X) - A] = 0, \quad (4.1)$$

where $I_{\text{knw}} = I_0 \cup I_{\text{fix}}$ and $P_{I_{\text{knw}}}[M]$ is the projection onto the subspace of n -by- n sparse matrices with non-zeros values in the entries indexed by I_{knw} , that is,

$$P_{I_{\text{knw}}}[M] = \begin{cases} M_{ij} & \text{if } (i, j) \in I_{\text{knw}}, \\ 0 & \text{otherwise} \end{cases}$$

and

$$\bar{F}(X) = -\|Q \circ \Phi(X)\|_F^2,$$

is the cost function, which seeks to complete, with non-null values, the coordinates contained in I_{nz} . An important property of the projection operator $P_{I_{\text{knw}}}[\cdot]$ is

$$P_{I_{\text{knw}}}[M_1 + \beta M_2] = P_{I_{\text{knw}}}[M_1] + \beta P_{I_{\text{knw}}}[M_2], \quad \forall M_1, M_2 \in \mathbb{R}^{n \times n}, \quad \forall \beta \in \mathbb{R},$$

which indicates that $P_{I_{\text{knw}}}[\cdot]$ is a linear operator. In addition, observe that the proposed optimization model (4.1) is well-defined because of the compactness of the feasible set $\{X \in \mathcal{O}(n) : P_{I_{\text{knw}}}[\Phi(X) - A] = 0\}$ and the continuity of the objective function, which guarantees the existence of global minimizers (even maximizers) for problem (4.1).

On the other hand, by applying the variable splitting technique to Problem (4.1), we obtain

$$\min_{X, Y \in \mathbb{R}^{n \times n}} \bar{F}(X) \quad \text{s.t.} \quad X^\top X = I, \quad P_{I_{\text{knw}}}[Y - A] = 0, \quad Y = \Phi(X). \quad (4.2)$$

To solve the above problem, we can use the quadratic penalty method to deal with the non-linear constraint $Y = \Phi(X)$. It is well-known that, starting from (X_0, Y_0) , the classic quadratic penalization method solves

$$\min_{X, Y \in \mathbb{R}^{n \times n}} \bar{F}(X) + \frac{\mu_k}{2} \|Y - \Phi(X)\|_F^2 \quad \text{s.t.} \quad X^\top X = I, \quad P_{I_{\text{knw}}}[Y - A] = 0, \quad (4.3)$$

at the k -th iteration for (X_{k+1}, Y_{k+1}) , until convergence, where $\{\mu_k\}$ is an increasing sequence of positive real numbers such that $\mu_k \rightarrow \infty$. However, solving (4.3) for (X, Y)

simultaneously is challenging because of the complicated form of the constraints. Additionally, the sub-problems (4.3) cannot be solved analytically due to the non-linearity and non-convexity of the feasible domain.

Using the idea of the classical alternating minimization methods [3], we solve problem (4.3) with respect to each matrix variable X and Y separately, while fixing the other at its previous values, and then the penalty parameter is incremented. Particularly, starting from (X_0, Y_0) and $\mu_0 = 1$, we propose the following framework

$$X_{k+1} = \arg \min_X \bar{F}(X) + \frac{\mu_k}{2} \|\Phi(X) - Y_k\|_F^2 \quad \text{s.t.} \quad X^\top X = I \quad (4.4a)$$

$$Y_{k+1} = \arg \min_Y \frac{\mu_k}{2} \|Y - \Phi(X_{k+1})\|_F^2 \quad \text{s.t.} \quad P_{I_{\text{knw}}}[Y - A] = 0 \quad (4.4b)$$

$$\mu_{k+1} = \rho \mu_k, \quad (4.4c)$$

where $\rho > 1$ is a global parameter. The orthogonality constrained sub-problems have no closed solution, thus inner iterations are needed to solve (4.4a) approximately. In practice, we employ the Riemannian gradient method discussed in Section 3. Nevertheless, the advantage of the iterative process (4.4b)-(4.4a)-(4.4c) is that the second optimization sub-problem has an analytical solution (see Lemma 4.1) given by

$$Y_{k+1} = \Phi(X_{k+1}) + P_{I_{\text{knw}}}[A - \Phi(X_{k+1})]. \quad (4.5)$$

The following lemma provides the closed-form solution for the optimization problem (4.4b).

LEMMA 4.1. *Let $A, B \in \mathbb{R}^{n \times n}$ be two given square matrices. Then, the global solution of the problem*

$$\min_{Y \in \mathbb{R}^{n \times n}} \frac{1}{2} \|Y - B\|_F^2 \quad \text{s.t.} \quad P_{I_{\text{knw}}}[Y - A] = 0, \quad (4.6)$$

has the following closed-form

$$Y = B + P_{I_{\text{knw}}}[A - B].$$

Proof. By introducing a Lagrange multiplier $\hat{\Lambda} \in \mathbb{R}^{n \times n}$, so that $\hat{\Lambda} = P_{I_{\text{knw}}}[\hat{\Lambda}]$, the Lagrangian function associated with (4.6) is defined as

$$\mathcal{L}(Y, \hat{\Lambda}) = \frac{1}{2} \|Y - B\|_F^2 - \text{tr}(\hat{\Lambda}^\top P_{I_{\text{knw}}}[Y - A]).$$

Now, differentiating this function, we obtain the Karush-Kuhn-Tucker conditions for Problem (4.6):

$$Y - B - \hat{\Lambda} = 0 \quad (4.7)$$

and

$$P_{I_{\text{knw}}}[Y - A] = 0. \quad (4.8)$$

In view of (4.7)-(4.8) and $P_{I_{\text{knw}}}[\hat{\Lambda}] = \hat{\Lambda}$, we have

$$\hat{\Lambda} = P_{I_{\text{knw}}}[Y - B] = P_{I_{\text{knw}}}[Y] - P_{I_{\text{knw}}}[B] = P_{I_{\text{knw}}}[A] - P_{I_{\text{knw}}}[B] = P_{I_{\text{knw}}}[A - B].$$

Merging this last result with (4.7), we arrive at

$$Y = B + P_{I_{\text{knw}}}[A - B].$$

Therefore, the pair $(Y, \hat{\Lambda}) = (B + P_{I_{\text{knw}}}[A - B], P_{I_{\text{knw}}}[A - B])$ is a critical point of the Lagrangian function associated with the linearly constrained optimization problem (4.6). Finally, since (4.6) is convex, this critical point is also the global solution. \square

In numerical optimization, proximal-type algorithms [27, 26] are derived by incorporating a proximal term in the objective function to enforce that the new iterate does not stray too far from the previous iterate, such a property will be fundamental in our convergence analysis. Inspired by this strategy, we propose the following modification of the method (4.4b)-(4.4a)-(4.4c),

$$X_{k+1} = \arg \min_X \bar{F}(X) + \frac{\mu_k}{2} \|\Phi(X) - Y_k\|_F^2 \quad \text{s.t.} \quad X^\top X = I \quad (4.9a)$$

$$Y_{k+1} = \arg \min_Y \frac{\mu_k}{2} \|Y - \Phi(X_{k+1})\|_F^2 + \frac{\alpha_k}{2} \|Y - Y_k\|_F^2 \quad \text{s.t.} \quad P_{I_{\text{knw}}}[Y - A] = 0 \quad (4.9b)$$

$$\mu_{k+1} = \rho \mu_k, \quad (4.9c)$$

where $\{\alpha_k\} \subset \mathbb{R}^+$, is a sequence of proximal parameters. The sub-problem (4.9b) is equivalent to

$$Y_{k+1} = \arg \min_Y \frac{1}{2} \|Y - (\tau_k \Phi(X_{k+1}) + (1 - \tau_k)Y_k)\|_F^2,$$

where $\tau_k := \frac{\mu_k}{\mu_k + \alpha_k} \in (0, 1]$. Therefore, it follows from Lemma 4.1 that the global solution of (4.9b) is given by

$$\begin{aligned} Y_{k+1} &= (\tau_k \Phi(X_{k+1}) + (1 - \tau_k)Y_k) + P_{I_{\text{knw}}}[A - (\tau_k \Phi(X_{k+1}) + (1 - \tau_k)Y_k)] \\ &= \tau_k (\Phi(X_{k+1}) + P_{I_{\text{knw}}}[A - \Phi(X_{k+1})]) + (1 - \tau_k)Y_k. \end{aligned}$$

Notice that Y_{k+1} is the convex combination between the previous iterate and the Y_{k+1} matrix of the non-altered sub-problem in variable Y (4.4b). Later we will see that for suitable choices of α_k , the sequence $\{Y_k\}$ generated by the proximal variant satisfies that $\lim_{k \rightarrow \infty} \|Y_{k+1} - Y_k\|_F = 0$, this property is crucial in our analysis, and it is not fulfilled for the original iterative scheme (4.4b)-(4.4a)-(4.4c). It is precisely this property that motivates the modification of the original alternating proposal.

5. Convergence analysis. This section begins by establishing the necessary first-order optimality conditions of Problem (4.2). The Lagrangian function associated with this problem is defined by

$$\mathcal{L}(X, Y, \Lambda, \hat{\Lambda}, \tilde{\Lambda}) = \bar{F}(X) - \frac{1}{2} \langle \Lambda, X^\top X - I \rangle - \langle \hat{\Lambda}, P_{I_{\text{knw}}}[Y - A] \rangle - \langle \tilde{\Lambda}, Y - \Phi(X) \rangle, \quad (5.1)$$

where the matrices Λ , $\hat{\Lambda}$, $\tilde{\Lambda}$ represent the dual variables and satisfy $\Lambda^\top = \Lambda$ and $\hat{\Lambda} = P_{I_{\text{knw}}}[\hat{\Lambda}]$, due to the symmetry of the constraint $X^\top X = I$ and the structure of the restriction $P_{I_{\text{knw}}}[Y - A] = 0$. By differentiating the mapping (5.1) with respect to all its block variables, we derive the Karush-Kuhn-Tucker (KKT) conditions for problem (4.2),

$$\nabla \bar{F}(X) - X\Lambda + DX(\tilde{\Lambda}^\top + \tilde{\Lambda}) = 0, \quad (5.2a)$$

$$\hat{\Lambda} + \tilde{\Lambda} = 0, \quad (5.2b)$$

$$X^\top X - I = 0, \quad (5.2c)$$

$$P_{I_{\text{knw}}}[Y - A] = 0, \quad (5.2d)$$

$$Y - \Phi(X) = 0. \quad (5.2e)$$

Let us recall that D denotes the diagonal matrix formed by the pre-established eigenvalues of A .

PROPOSITION 5.1. *The sequences $\{X_k\}$, $\{Y_k\}$ generated by the alternating algorithm are bounded.*

Proof. Since $X_k^\top X_k = I$ for all $k \geq 0$, we have directly that $\|X_k\|_F = \sqrt{n}$. Thus, the sequence $\{X_k\}$ is bounded. In addition, from the updating formula (4.5), for all k , we get

$$\begin{aligned} \|Y_{k+1}\|_F &= \|\Phi(X_{k+1}) + P_{I_{\text{know}}}[A - \Phi(X_{k+1})]\|_F \\ &\leq \|\Phi(X_{k+1})\|_F + \|P_{I_{\text{know}}}[A - \Phi(X_{k+1})]\|_F \\ &= \|D\|_F + \|P_{I_{\text{know}}}[A - \Phi(X_{k+1})]\|_F \\ &\leq \|D\|_F + \|A - \Phi(X_{k+1})\|_F \\ &\leq \|D\|_F + \|A\|_F + \|\Phi(X_{k+1})\|_F \\ &= 3\|D\|_F. \end{aligned}$$

Therefore, $\{Y_k\}$ is also bounded. \square

Now, let $\{(X_k, Y_k)\}$ be any sequence generated by the alternating algorithm described in section 4. From the optimality of the iterates X_k and Y_k , for each $k \geq 0$ there must exist matrices Θ_k and $\hat{\Theta}_k$ (Lagrange multipliers) verifying the following equations

$$\nabla \bar{F}(X_k) - X_k \Theta_k + 2\mu_{k-1} D X_k (\Phi(X_k) - Y_{k-1}) = 0, \quad (5.3)$$

and

$$\mu_{k-1} (Y_k - \Phi(X_k)) - \hat{\Theta}_k = 0, \quad (5.4)$$

respectively.

Now we are ready to present our main convergence result.

THEOREM 5.2. *Let $\{(X_k, Y_k)\}$ be any sequence constructed by the alternating algorithm. Let us suppose that $\{\Theta_k\}$ is a bounded sequence and that $\lim_{k \rightarrow \infty} \|Y_{k+1} - Y_k\|_F = 0$. Let (X_*, Y_*, Θ_*) be a limit point of the sequence $\{(X_k, Y_k, \Theta_k)\}$. Then, the pair (X_*, Y_*) is a KKT point for the problem (4.2).*

Proof. It follows from equation (5.3) that

$$\nabla \bar{F}(X_{k+1}) - X_{k+1} \Theta_{k+1} + 2\mu_k D X_{k+1} (\Phi(X_{k+1}) - Y_k) = 0, \quad (5.5)$$

or equivalently

$$2\mu_k D X_{k+1} [(\Phi(X_{k+1}) - Y_{k+1}) + (Y_{k+1} - Y_k)] = X_{k+1} \Theta_{k+1} - \nabla \bar{F}(X_{k+1}).$$

Then, using the fact that $X_k \in \mathcal{O}(n)$ for each $k \geq 0$ and recalling that D is non-singular, we get

$$\Phi(X_{k+1}) - Y_{k+1} = \frac{1}{2\mu_k} X_{k+1}^\top D^{-1} X_{k+1} (\Theta_{k+1} - \nabla \bar{F}(X_{k+1})) + (Y_k - Y_{k+1}).$$

Applying the Frobenius norm on both sides of the last result, we obtain

$$\begin{aligned}
\|\Phi(X_{k+1}) - Y_{k+1}\|_F &= \left\| \frac{1}{2\mu_k} X_{k+1}^\top D^{-1} X_{k+1} (\Theta_{k+1} - \nabla \bar{F}(X_{k+1})) + (Y_k - Y_{k+1}) \right\|_F \\
&\leq \left\| \frac{1}{2\mu_k} X_{k+1}^\top D^{-1} X_{k+1} (\Theta_{k+1} - \nabla \bar{F}(X_{k+1})) \right\|_F + \|Y_k - Y_{k+1}\|_F \\
&= \frac{1}{2\mu_k} \|D^{-1} X_{k+1} (\Theta_{k+1} - \nabla \bar{F}(X_{k+1}))\|_F + \|Y_k - Y_{k+1}\|_F \\
&\leq \frac{1}{2\mu_k} \|D^{-1}\|_F \|\Theta_{k+1} - \nabla \bar{F}(X_{k+1})\|_F + \|Y_k - Y_{k+1}\|_F, \\
&\leq \frac{1}{2\mu_k} \|D^{-1}\|_F (\|\Theta_{k+1}\|_F + \|\nabla \bar{F}(X_{k+1})\|_F) + \|Y_k - Y_{k+1}\|_F. \quad (5.6)
\end{aligned}$$

Since $\bar{F}(\cdot)$ is continuous and $\mathcal{O}(n)$ is compact then there exists $C_1 > 0$ such that $\|\bar{F}(X_k)\|_F \leq C_1$, for all k . In addition, from the hypothesis there also exists a constant $C_2 > 0$ such that $\|\Theta_k\|_F \leq C_2$, for any k . Substituting these two inequalities in (5.6) we arrive at

$$\|\Phi(X_{k+1}) - Y_{k+1}\|_F \leq \frac{(C_1 + C_2)}{2\mu_k} \|D^{-1}\|_F + \|Y_k - Y_{k+1}\|_F.$$

Applying limits in this last relation, we conclude that

$$\lim_{k \rightarrow \infty} \|\Phi(X_{k+1}) - Y_{k+1}\|_F = 0.$$

Notice that this result directly implies that $\Phi(X_*) - Y_* = 0$, which indicates that the pair (X_*, Y_*) satisfies the KKT condition (5.2e).

On the other hand, let us introduce the notations $\tilde{\Lambda}_{k+1} := \mu_k(\Phi(X_{k+1}) - Y_k)$ and $\hat{\Lambda}_{k+1} := \hat{\theta}_{k+1} + \mu_k(Y_k - Y_{k+1})$, where $\hat{\theta}_k$ is the matrix that appears in (5.4). Thus, equation (5.3) can be rewritten as

$$\nabla \bar{F}(X_{k+1}) - X_{k+1} \Theta_{k+1} + 2D X_{k+1} \tilde{\Lambda}_{k+1} = 0,$$

which implies that

$$\tilde{\Lambda}_{k+1} = \frac{1}{2} X_{k+1}^\top D^{-1} (X_{k+1} \Theta_{k+1} - \nabla \bar{F}(X_{k+1})).$$

Hence by taking the limit as $k \in \mathcal{K}$ goes to ∞ , we find that

$$\tilde{\Lambda}_* := \lim_{k \in \mathcal{K}} \tilde{\Lambda}_{k+1} = \frac{1}{2} X_*^\top D^{-1} (X_* \Theta_* - \nabla \bar{F}(X_*)).$$

Additionally, observe that the equation (5.4) can be reformulated as

$$\begin{aligned}
0 &= \mu_k(Y_{k+1} - \Phi(X_{k+1})) - \hat{\Theta}_{k+1} \\
&= \mu_k(Y_k - \Phi(X_{k+1}) + Y_{k+1} - Y_k) - \hat{\Theta}_{k+1} \\
&= -\mu_k(\Phi(X_{k+1}) - Y_k) - (\hat{\Theta}_{k+1} + \mu_k(Y_k - Y_{k+1})) \\
&= -\tilde{\Lambda}_{k+1} - \hat{\Lambda}_{k+1},
\end{aligned}$$

which leads us to

$$\hat{\Lambda}_{k+1} = -\tilde{\Lambda}_{k+1}, \quad \forall k \geq 0. \quad (5.7)$$

By taking limits in (5.7), we conclude that

$$\hat{\Lambda}_* := \lim_{k \in \mathcal{K}} \hat{\Lambda}_{k+1} = -\tilde{\Lambda}_*,$$

and consequently

$$\hat{\Lambda}_* + \tilde{\Lambda}_* = 0$$

which establishes the existence of two matrices satisfying the KKT condition (5.2b).

Similarly, since $\tilde{\Lambda}_{k+1}$ is a symmetric matrix for all $k \geq 0$, the equation (5.5) can be posed as

$$\nabla \bar{F}(X_{k+1}) - X_{k+1} \Theta_{k+1} + DX_{k+1}(\tilde{\Lambda}_{k+1}^\top + \tilde{\Lambda}_{k+1}) = 0,$$

where we are using the notation $\Lambda_{k+1} := \Theta_{k+1}$. Here, it is important to note that $\Lambda_* := \lim_{k \in \mathcal{K}} \Lambda_{k+1} = \Theta_*$. Hence, applying limits in the previous equality we have

$$\nabla \bar{F}(X_*) - X_* \Lambda_* + DX_*(\tilde{\Lambda}_*^\top + \tilde{\Lambda}_*) = 0.$$

This indicates that the KKT condition (5.2a) is also satisfied in the limit. Finally, notice that the KKT condition (5.2c) is easily verified at X^* because $\{X_k\} \subset \mathcal{O}(n)$ and $\mathcal{O}(n)$ is a closed set; and the condition KKT (5.2d) is fulfilled trivially by applying the limit, since $P_{I_{kru}}[Y_k - A] = 0$, for all $k \geq 0$. Therefore the theorem is proved. \square

Now we will analyze a modification of the previous alternating algorithm by incorporating a proximal term. For the proximal alternating minimization approach there must exists, at the k -th iteration, a matrix $\hat{\Theta}_k$ (Lagrangian multiplier) such that

$$\mu_{k-1}(Y_k - \Phi(X_k)) + \alpha_{k-1}(Y_k - Y_{k-1}) - \hat{\Theta}_k = 0,$$

for all k .

PROPOSITION 5.3. *Let $\{(X_k, Y_k)\}$ be a sequence generated by the iterative process defined by (4.9a)-(4.9b). Assume that the sequence $\{Y_k\}$ is bounded and that the proximal parameter is given by $\alpha_k = \sigma \mu_k^2$, for some $\sigma > 0$. Then,*

$$\lim_{k \rightarrow \infty} \|Y_{k+1} - Y_k\|_F = 0.$$

Proof. It follows from the definition of X_{k+1} in (4.9a) that

$$\bar{F}(X_{k+1}) + \frac{\mu_k}{2} \|Y_k - \Phi(X_{k+1})\|_F^2 \leq \bar{F}(X_k) + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2. \quad (5.8)$$

Similarly, from the definition of Y_{k+1} in (4.9b), we have

$$\frac{\mu_k}{2} \|Y_{k+1} - \Phi(X_{k+1})\|_F^2 + \frac{\alpha_k}{2} \|Y_{k+1} - Y_k\|_F^2 \leq \frac{\mu_k}{2} \|Y_k - \Phi(X_{k+1})\|_F^2. \quad (5.9)$$

Combining the inequalities (5.8) and (5.9) we get

$$\bar{F}(X_{k+1}) + \frac{\mu_k}{2} \|Y_{k+1} - \Phi(X_{k+1})\|_F^2 \leq \bar{F}(X_k) + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2 - \frac{\alpha_k}{2} \|Y_{k+1} - Y_k\|_F^2. \quad (5.10)$$

By rearranging the inequality (5.10) we obtain

$$\begin{aligned}
\frac{\alpha_k}{2} \|Y_{k+1} - Y_k\|_F^2 &\leq \bar{F}(X_k) - \bar{F}(X_{k+1}) + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2 - \frac{\mu_k}{2} \|Y_{k+1} - \Phi(X_{k+1})\|_F^2 \\
&\leq \bar{F}(X_k) - \bar{F}(X_{k+1}) + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2 \\
&\leq -\bar{F}(X_{k+1}) + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2 \\
&\leq \|Q \circ \Phi(X_{k+1})\|_F^2 + \frac{\mu_k}{2} \|Y_k - \Phi(X_k)\|_F^2.
\end{aligned}$$

As a consequence,

$$\frac{\alpha_k}{2\mu_k^2} \|Y_{k+1} - Y_k\|_F^2 \leq \frac{1}{\mu_k^2} \|Q \circ \Phi(X_{k+1})\|_F^2 + \frac{1}{2\mu_k} \|Y_k - \Phi(X_k)\|_F^2.$$

Since $\{Y_k\}$ is bounded, there exists a positive real number $M > 0$ such that $\|Y_k\| \leq M$, for all k . Using this bound and the formula $\alpha_k = \sigma\mu_k^2$ provided by the hypothesis, we obtain

$$\begin{aligned}
\|Y_{k+1} - Y_k\|_F^2 &\leq \frac{2}{\sigma} \left(\frac{1}{\mu_k^2} \|Q \circ \Phi(X_{k+1})\|_F^2 + \frac{1}{2\mu_k} \|Y_k - \Phi(X_k)\|_F^2 \right) \\
&\leq \frac{2}{\sigma} \left(\frac{1}{\mu_k^2} \|\Phi(X_{k+1})\|_F^2 + \frac{1}{2\mu_k} (\|Y_k\|_F + \|\Phi(X_k)\|_F^2) \right) \\
&\leq \frac{2}{\sigma} \left(\frac{1}{\mu_k^2} \|D\|_F^2 + \frac{1}{2\mu_k} (M + \|D\|_F^2) \right) \\
&\leq \frac{1}{\sigma\mu_k} \left(\frac{2 + \mu_k}{\mu_k} \|D\|_F^2 + M \right)
\end{aligned}$$

By taking limits in both sides of the above relation, and using (4.4c) we conclude that

$$\lim_{k \rightarrow \infty} \|Y_{k+1} - Y_k\|_F^2 = 0.$$

□

6. A manifold ADMM. In this section we design a Riemannian alternating direction method of multipliers (RADMM) as an alternative to recover the matrix A . Let us recall the optimization model (4.2). Instead of penalizing the equality constraint $Y = \Phi(X)$ to obtain the model (4.3), we now consider the Lagrangian function associated with (4.2), given by

$$\begin{aligned}
\mathcal{L}_\mu(X, Y, \tilde{\Theta}) &= \bar{F}(X) + \langle \tilde{\Theta}, Y - \Phi(X) \rangle + \frac{\mu}{2} \|Y - \Phi(X)\|_F^2 \\
&= \bar{F}(X) + \frac{\mu}{2} \|Y - \Phi(X) + \mu^{-1} \tilde{\Theta}\|_F^2 - \frac{1}{2\mu} \|\tilde{\Theta}\|_F^2,
\end{aligned}$$

where $\tilde{\Theta} \in \mathbb{R}^{n \times n}$ represents the dual variable related to the constraint $Y = \Phi(X)$. Starting from $\tilde{\Theta}_0 = 0$ and $\mu_0 = 1$, the classic augmented Lagrangian method (ALM) [4, 5] generates a sequence $\{(X_k, Y_k, \tilde{\Theta}_k)\}$ by

$$(X_k, Y_k) = \arg \min_{X, Y} \mathcal{L}_{\mu_k}(X, Y, \tilde{\Theta}_k) \quad \text{s.t.} \quad X^\top X = I, \quad P_{I_{knw}}[Y - A] = 0, \quad (6.1)$$

then updates the multiplier matrix $\tilde{\Theta}$ by the recursive formula

$$\tilde{\Theta}_{k+1} = \tilde{\Theta}_k + \mu_k(Y_{k+1} - \Phi(X_{k+1})),$$

and set $\mu_{k+1} = \rho\mu_k$, for some global parameter $\rho > 1$ selected by the user. An efficient approach to solve the subproblem (6.1) is the well-known alternating direction method of multipliers ADMM [9, 10, 12]. Specifically, we propose the following iterative process:

$$X_{k+1} = \arg \min_X \mathcal{L}_{\mu_k}(X, Y_k, \tilde{\Theta}_k) \quad \text{s.t.} \quad X^\top X = I \quad (6.2a)$$

$$Y_{k+1} = \arg \min_Y \mathcal{L}_{\mu_k}(X_{k+1}, Y, \tilde{\Theta}_k) \quad \text{s.t.} \quad P_{I_{\text{knw}}}[Y - A] = 0 \quad (6.2b)$$

$$\tilde{\Theta}_{k+1} = \tilde{\Theta}_k + \mu_k(Y_{k+1} - \Phi(X_{k+1})) \quad (6.2c)$$

$$\mu_{k+1} = \rho\mu_k, \quad (6.2d)$$

Here, the subproblem (6.2a) can be solved by using the Riemannian gradient method described in Section 3, while using Lemma 4.1 the subproblem (6.2b) has an exact solution given by

$$Y_{k+1} = \Phi(X_{k+1}) - \mu_k^{-1}\tilde{\Theta}_k + P_{I_{\text{knw}}}[A - (\Phi(X_{k+1}) - \mu_k^{-1}\tilde{\Theta}_k)]. \quad (6.3)$$

We close this section with some remarks. The ALM is an effective tool for solving constrained optimization problems, which has been well studied and enjoys global convergence under certain assumptions, see, e.g., [4, 5] as two standard references. The alternating direction method of multipliers (ADMM) given by (6.2) can be seen as an approximated ALM specially designed to solve the non-convex problem (4.2). In fact, at every cycle, our proposed ADMM scheme (6.2) needs to solve two optimization subproblems: (6.2a) and (6.2b). Problem (6.2b) is solved exactly using (6.3), and (6.2a) minimizes a continuous function on a (non-convex) compact set, and so it has global solutions. We note that the ADMM has been successfully adapted to solve several non-convex problems; see, e.g., matrix completion [29, 30], noisy color image restoration [18], and clustering in pattern recognition [21]. In addition, the global convergence of ADMM for a class of non-convex and non-smooth optimization problems has been established; see, e.g., [9, 10, 12]. Moreover, the ADMM strategy has been extended to solve optimization problems on a manifold, see [17, 19].

7. Numerical Results. To give further insight into the inverse structured symmetric matrix problem with given eigenvalues, and to illustrate the practical performance of the proposed algorithms, we present the results of some numerical experiments. We compare several proposed methods: the Riemannian gradient method combined with the QR retraction (R-QR), Riemannian gradient method combined with the Cayley retraction (R-Cayley), the alternating scheme based on a penalization strategy described in Section 4 (Alternating-P), and the alternating scheme based on an augmented Lagrangian strategy described in Section 6 (Alternating-AL). All considered methods were implemented in Matlab. The experiments were executed in a laptop computer with CPU Intel core i7, 4.7 GHz, RAM memory of 16 GB.

All runs are terminated as soon as one of the following three conditions holds: a maximum number of iterations (*maxiter*) is reached, the difference in norm of two consecutive iterates is less than or equal to a given tolerance (*tol* > 0), or the Riemannian gradient norm is less than or equal to *tol*. To guarantee that the initial

matrix X_0 belongs to the Stiefel manifold, we generate a random symmetric matrix W such that $w_{i,j} = 0$ for all pairs $(i, j) \in I_0$, and set X_0 to be the orthogonal factor Q of the QR factorization of W . The required input parameters in our algorithms are fixed as follows: $maxiter = 500$, $tol = 10^{-5}$, the growing factor (ρ) for the penalty parameters is 1.5, and any of the penalty parameters stop growing when they reach the value 10^{20} . The entries of the matrices $\tilde{\Theta}_k$, in (6.2c), are kept between the bounds -10^{20} and 10^{20} . In our experiments, we report the number of required iterations (Iter), the objective function value (given by (2.3) for R-QR and R-Cayley or by (4.2) for the Alternating schemes) at the obtained solution (feval), and the final value of the penalty parameters μ_k and σ_k . If a method reaches $maxiter$ iterations without obtaining a solution, a failure is reported using the symbol (**) under feval.

For our first set of experiments, we fix the diagonal entries of D to be the eigenvalues of the tridiagonal symmetric positive definite $n \times n$ matrix obtained by discretizing the second-order derivative, over a closed interval in one variable, with zero values at the extreme points. In other words, the given λ_i s are the eigenvalues of the matrix $tridiag(-1, 2, -1)$, which can be seen as a special case of the so-called Laplacian matrices that have an impact in chemical graph theory [7, 32]. It is well-known that, for any $n \geq 2$, the eigenvalues of $tridiag(-1, 2, -1)$ are given by $\lambda_i = 2 - 2 \cos(i \pi / (n + 1))$ for $1 \leq i \leq n$. The considered structure is the following: The operator P is defined to guarantee that out of the three main diagonals all entries are zero; i.e., $P = tridiag(1, 1, 1)$. As for additional structural requirements, the entries in the main diagonal are all totally free variables (indicated with the symbol x); there are no nonzero preset values, so $R = 0$, $C = 0$ and $\sigma_k = 0$ for all k ; and all the entries in the main sub-diagonals are free variables that must be nonzero (indicated with the symbol nz), thus the matrix Q has 1 at the entries of the two main sub-diagonals and zero elsewhere. For example, the given matrix pattern for $n = 6$ is given by:

$$\begin{bmatrix} x & nz & 0 & 0 & 0 & 0 \\ nz & x & nz & 0 & 0 & 0 \\ 0 & nz & x & nz & 0 & 0 \\ 0 & 0 & nz & x & nz & 0 \\ 0 & 0 & 0 & nz & x & nz \\ 0 & 0 & 0 & 0 & nz & x \end{bmatrix},$$

which is identified with a simple undirected graph of 6 vertices [16, pp. 5-8]. For this tridiagonal matrix and for several values of n , in Table 7.1 we report the performance of the 4 considered methods: R-QR, R-Cayley, Alternating-P, and Alternating-AL. We note that, except for the R-QR method when $n = 10$, the eigenvalues of the obtained matrix (although each method obtains a different matrix) are the eigenvalues of $tridiag(-1, 2, -1)$. From Table 7.1 we also see that, even when the penalty parameter μ_k reaches the upper bound 10^{20} or a near-by value, the corresponding method works but it requires many more iterations; see, e.g., the behavior of the R-QR and the Alternating-P schemes for $n = 10$. Finally, as expected, for the Alternating-AL scheme the parameter μ_k reaches a much smaller value than the one reached by the Alternating-P method.

TABLE 7.1
Performance of the 4 considered methods for the inverse structured tridiagonal matrix case.

n	Method	Iter	feval	μ_k
4	R-QR	45	-4.61	9.7×10^5
	R-Cayley	44	-4.51	9.2×10^5
	Alternating-P	42	-5.88	5.6×10^5
	Alternating-AL	35	-4.79	4.4×10^3
7	R-QR	59	-6.28	2.1×10^9
	R-Cayley	57	-6.3	4.8×10^9
	Alternating-P	90	-4.84	7.3×10^{14}
	Alternating-AL	53	-8.61	1.4×10^6
10	R-QR	293	-3.16	1.0×10^{20}
	R-Cayley	68	-3.92	7.5×10^9
	Alternating-P	103	-15.7	2.7×10^{17}
	Alternating-AL	48	-16.1	6.4×10^5

For $n = 4$ we now present the obtained matrices A_4^{Alt-P} (left) and A_4^{Alt-AL} (right), produced by the Alternating-P and the Alternating-AL methods, respectively. We note that the 4 obtained matrices are different, but they all have the same 4 eigenvalues: 0.382, 1.382, 2.618, and 3.618.

$$\begin{bmatrix} 1.85 & -1.21 & 0 & 0 \\ -1.21 & 1.9 & -0.89 & 0 \\ 0 & -0.89 & 2.33 & -0.82 \\ 0 & 0 & -0.82 & 1.92 \end{bmatrix}, \quad \begin{bmatrix} 1.96 & -1.45 & 0 & 0 \\ -1.45 & 1.94 & -0.64 & 0 \\ 0 & -0.64 & 2.46 & -0.56 \\ 0 & 0 & -0.56 & 1.64 \end{bmatrix}.$$

For the second experiment, we consider the 7×7 symmetric Laplacian matrix described in [32, Table II, p. 369]. These matrices have a potential for use in chemical graph theory. We fix the diagonal entries of D to be the eigenvalues of the matrix displayed in [32], i.e., the given eigenvalues are: 0, 0.3672, 1.0571, 2.3745, 4.4681, 7.0357, and 8.6973. The required structural pattern is the following:

$$\begin{bmatrix} x & nz & 0 & 0 & 0 & 0 & 0 \\ nz & x & nz & 0 & 0 & 0 & -3 \\ 0 & nz & x & nz & 0 & -3 & 0 \\ 0 & 0 & nz & x & nz & 0 & 0 \\ 0 & 0 & 0 & nz & x & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & x & 0 \\ 0 & -3 & 0 & 0 & 0 & 0 & x \end{bmatrix},$$

which is identified with the undirected weighted graph (a linear tree in this case) of the 2,3-Dimethyl pentane, shown in Figure 7.1. In Table 7.2 we report the behavior of the 4 considered methods for this 7×7 Laplacian matrix. In all cases, the structural requirements are satisfied and the 7 eigenvalues are the given ones.

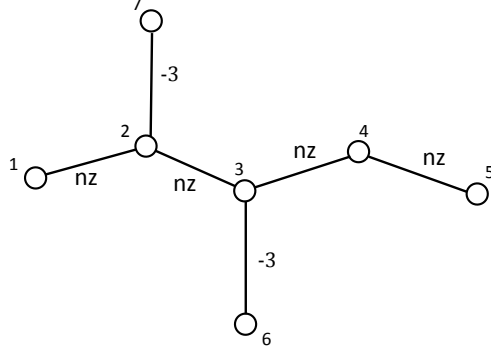


FIG. 7.1. Structural pattern graph of the symmetric Laplacian matrix (2,3-Dimethyl pentane).

TABLE 7.2
Performance of the 4 considered methods for the 7×7 symmetric Laplacian case.

Method	Iter	feval	μ_k	σ_k
R-QR	137	-3.72	1.0×10^{20}	1.0×10^{20}
R-Cayley	134	-19.1	5.2×10^{19}	1.0×10^{20}
Alternating-P	205	-2.66	1.0×10^{20}	—
Alternating-AL	60	-9.7	5.5×10^7	—

We notice that, for the 4 methods, the structural requirements are satisfied and the 7 eigenvalues are the given ones. As an example, we now present the matrix A_7^{Alt-AL} obtained by the Alternating-AL method:

$$\begin{bmatrix} 2.3852 & -0.9914 & 0 & 0 & 0 & 0 & 0 \\ -0.9914 & 6.1461 & -1.2177 & 0 & 0 & 0 & -3 \\ 0 & -1.2177 & 5.0706 & 1.4586 & 0 & -3 & 0 \\ 0 & 0 & 1.4586 & 4.2547 & 1.3785 & 0 & 0 \\ 0 & 0 & 0 & 1.3785 & 0.6945 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 3.1763 & 0 \\ 0 & -3 & 0 & 0 & 0 & 0 & 2.2726 \end{bmatrix}.$$

For our third experiment, let us consider the 4×4 symmetric matrix indicated in [31, p. 265], in which the only fixed value is π at the (2,2) entry. In this case, we fix the given eigenvalues to be 1, 2, 3, and 4. For completeness we present the required structural pattern:

$$\begin{bmatrix} x & x & nz & 0 \\ x & \pi & nz & x \\ nz & nz & x & 0 \\ 0 & x & 0 & nz \end{bmatrix},$$

which is identified with an undirected graph. We note that in this case it is not a tree since it has a cycle (vertices 1, 2, and 3). In Table 7.3 we report the behavior of the 4 considered methods for this 4×4 matrix. In all cases, all the structural requirements are satisfied and the 4 obtained eigenvalues are the given ones. As an example, we

present the obtained matrices A_4^{Alt-P} (left) and A_4^{Alt-AL} (right), produced by the Alternating-P and the Alternating-AL methods, respectively:

$$\begin{bmatrix} 2.276 & -0.489 & 0.1364 & 0 \\ -0.489 & 3.1416 & -0.625 & -0.483 \\ 0.1364 & 0.625 & 3.45 & 0 \\ 0 & -0.483 & 0 & 1.132 \end{bmatrix}, \quad \begin{bmatrix} 1.488 & -0.0034 & -0.5 & 0 \\ -0.0034 & 3.1416 & -0.0024 & 0.3486 \\ -0.5 & -0.0024 & 1.512 & 0 \\ 0 & 0.3486 & 0 & 3.858 \end{bmatrix}.$$

TABLE 7.3

Performance of the 4 considered methods for the 4×4 symmetric matrix indicated in [31, p. 265].

Method	Iter	feval	μ_k	σ_k
R-QR	42	-3.4	8.5×10^6	2.1×10^8
R-Cayley	40	-2.1	1.3×10^6	1.4×10^8
Alternating-P	101	-2.85	1.2×10^{12}	—
Alternating-AL	58	-15.2	1.6×10^7	—

For the fourth and last experiment, we consider the 10×10 symmetric matrix described in [16, p. 150], which is identified with the undirected graph shown in Figure 7.2. This graph is the smallest possible nonlinear tree; see [16, Section 5]. Following the description in [16, p. 150], we fix the given eigenvalues to be -3, -2, -2, 0, 0, 0, 0, 2, 2, and 3. These 5 eigenvalues together with their multiplicity guarantee that there are possible solutions to the inverse structural matrix problem; see [16, Section 6]. The required structural pattern is the following:

$$\begin{bmatrix} 0 & 0 & \text{nz} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \text{nz} & \sqrt{2} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \sqrt{2} & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & \text{nz} & \text{nz} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{nz} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{nz} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & \text{nz} & \text{nz} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{nz} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{nz} & 0 & 0 \end{bmatrix}.$$

In Table 7.4 we report the behavior of the 4 considered methods for this 10×10 matrix. It can be observed that the methods R-QR and R-Cayley could not find a solution after 500 iterations, while the Alternating scheme was able to find a solution for which all the structural requirements are satisfied and the 10 obtained eigenvalues are the given ones. We now present the matrix A_{10}^{Alt-AL} obtained by the Alternating-AL method:

$$\begin{bmatrix} 0 & 0 & 1.4142 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.4142 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.4142 & 1.4142 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1.4142 & 0 & 0 & 1.4142 & 0 & 0 \\ 0 & 0 & 0 & 1.4142 & 0 & 0.0073 & 2.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0073 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.4142 & 0 & 0 & 0 & 0 & -0.7099 & 1.8698 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.7099 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.8698 & 0 & 0 & 0 \end{bmatrix}.$$

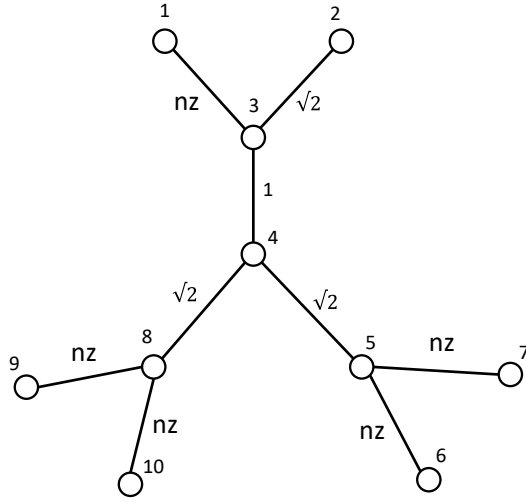


FIG. 7.2. Structural pattern graph of the symmetric matrix described in [31, p. 150].

TABLE 7.4

Performance of the 4 considered methods for the 10×10 symmetric matrix described in [31, p. 150].

Method	Iter	feval	μ_k	σ_k
R-QR	500	**	1.0×10^{20}	1.0×10^{20}
R-Cayley	500	**	1.0×10^{20}	1.0×10^{20}
Alternating-P	45	-19.9	5.5×10^7	—
Alternating-AL	52	-19.5	8.5×10^4	—

8. Concluding Remarks. We have focused on the inverse structured symmetric matrix problem for which certain entries have preassigned values, including a sparse pattern, and such that the desired matrix eigenvalues are given in advance. In principle, the variables of the problem are the rest of the entries, but some of them cannot be zero and others are completely free. For this problem, we have presented an optimization model in the space of square matrices, which attempts to achieve all the stated objectives only using the orthogonal eigenvector matrix of the solution as a variable. It should be noted that with only this matrix as an unknown a solution can be achieved, even though this matrix alone has no explicit value once the problem is solved.

To solve these problems, we have developed several iterative methods that keep the iterations on the Stiefel manifold of orthogonal matrices. In particular, we have combined a Riemmanian gradient-type method with a non-monotone line search globalization strategy and with two types of retraction operators to stay on the manifold: QR and Cayley. The proposed Riemmanian gradient schemes have been applied as internal methods of a penalty machinery that enforces the structural requirements by imposing the orthogonality of the variable matrix as the only constraint, producing the R-QR and the R-Cayley methods. In our numerical experiments we have observed the standard tendency of the penalty parameters to grow to very high values to guar-

antee convergence. As a remedy to this numerical difficulty, we have also proposed a block alternating scheme based on a suitable variable splitting. One of the alternating steps is combined with either a penalty or an augmented Lagrangian strategy, and the other alternating step is solved by a conveniently obtained closed formula.

In our numerical section we add understanding to the topic by illustrating the connection between the desired symmetric matrix and the associated undirected graph. In particular, we take advantage of knowledge recently developed in this line of research known as a spectral graph theory. In our experiments, we have observed that the Alternating-P and the Alternating-AL schemes present better practical performance compared to the R-QR and R-Cayley methods, especially when the size of the matrices increases. Furthermore, in the case of the nonlinear tree with $n = 10$, only the alternating methods were able to find a solution with few iterations, the other two reached the limit of 500 iterations without achieving convergence. In general, we have observed that the alternating schemes are effective up to $n = 20$, which includes all the graphs analyzed in [16, Appendix A], and are no longer effective when $n > 20$. Clearly, the large-scale case is an open challenge, which in our opinion will require an adequate adaptation of the ideas contained in this work, or the development of another type of suitable optimization schemes.

Acknowledgements. The third author would like to thank the Laboratory LAMFA at Université de Picardie Jules Verne, Amiens, France, for the hospitality during a three-week invitation in May 2023.

Funding.

The second author was financially supported by the Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Chile, through the FES startup package for scientific research. The third author was financially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) under the scope of projects UIDB/MAT/00297/2020 and UIDP/MAT/00297/2020 (Centro de Matemática e Aplicações).

Data availability. The codes and data sets generated during the current study are available from the corresponding author on reasonable request.

Disclosure statement. No potential conflict of interest was reported by the authors.

REFERENCES

- [1] P. A. Absil, R. Mahony, and R. Sepulchre [2008], *Optimization algorithms on matrix manifolds*, Princeton University Press. DOI: <https://doi.org/10.1515/9781400830244>
- [2] J. Barzilai and J. M. Borwein [1988], *Two-point step size gradient methods*, IMA J. Numer. Anal., 8: 141–148. DOI: <https://doi.org/10.1093/imanum/8.1.141>
- [3] A. Beck [2015], *On the convergence of alternating minimization for convex programming with applications to iteratively reweighted least squares and decomposition schemes*, SIAM J. Optim., 25(1): 185–209. DOI: <https://doi.org/10.1137/13094829X>
- [4] D. Bertsekas [1982], *Constrained optimization and Lagrange multiplier methods*, Academic Press, New York.

- [5] E. Birgin and J. M. Martínez [2014], *Practical augmented Lagrangian methods for constrained optimization*, SIAM, Philadelphia.
- [6] N. Boumal [2023], *An introduction to optimization on smooth manifolds*, Cambridge University Press, Cambridge. DOI: <https://doi.org/10.1017/9781009166164>
- [7] D. Cvetković, P. Rowlinson, and S. L. Simić [2009], *An Introduction to the Theory of Graph Spectra*, Cambridge University Press, Cambridge. DOI: <https://doi.org/10.1017/CB09780511801518>
- [8] M. T. Chu [1998], *Inverse eigenvalue problems*, SIAM Rev., 40(1): 1–39. DOI: <https://doi.org/10.1137/S0036144596303984>
- [9] W. Deng and W. Yin [2016], *On the global and linear convergence of the generalized alternating direction method of multipliers.*, J. Sci. Comput., 66: 889–916. DOI: <https://doi.org/10.1007/s10915-015-0048-x>
- [10] J. Eckstein and W. Yao [2015], *Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives.*, Pac. J. Optim., 11(4): 619–644.
- [11] A. Edelman, T. Arias and S. Smith [1998], *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20(2): 303–353. DOI: <https://doi.org/10.1137/S0895479895290954>
- [12] M. Hong and Z. Luo [2017], *On the linear convergence of the alternating direction method of multipliers.*, Math. Program., 162(1-2): 165–199. DOI: <https://doi.org/10.1007/s10107-016-1034-2>
- [13] J. Hu, X. Liu, Z. Wen and Y.-X. Yuan [2020], *A brief introduction to manifold optimization*, Journal of the Operations Research Society of China, 8: 199–248. DOI: <https://doi.org/10.1007/s40305-020-00295-9>
- [14] B. Iannazzo and M. Porcelli [2018], *The Riemannian Barzilai-Borwein method with nonmonotone line search and the matrix geometric mean computation*, IMA Journal of Numerical Analysis, 38(1): 495–517. DOI: <https://doi.org/10.1093/imanum/drx015>
- [15] B. Jiang and Y.-H. Dai [2015], *A framework of constraint preserving update schemes for optimization on Stiefel manifold*, Mathematical Programming, 153(2): 535–575. DOI: <https://doi.org/10.1007/s10107-014-0816-7>
- [16] C. R. Johnson and C. M. Saigo [2018], *Eigenvalues, multiplicities and graphs*, volume 211 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge. DOI: <https://doi.org/10.1017/9781316155158>
- [17] A. Kovnatsky, K. Glashoff and M. Bronstein [2016], *MADMM: a generic algorithm for non-smooth optimization on manifolds.*, In: Computer Vision-ECCV 2016, 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V14 (pp. 680–696).
- [18] R. Lai and S. Osher [2014], *A splitting method for orthogonality constrained problems.* J. Sci. Comput., 58: 431–449. DOI: <https://doi.org/10.1007/s10915-013-9740-x>
- [19] J. Li, S. Ma and T. Srivastava [2022], *A Riemannian ADMM*, arXiv preprint arXiv:2211.02163.
- [20] D.G. Luenberger [1984], *Linear and Nonlinear Programming*, Addison-Wesley, Menlo Park, CA. DOI: <https://doi.org/10.1007/978-0-387-74503-9>
- [21] C. Lu, J. Feng, Z. Lin and S. Yan [2018], *Nonconvex sparse spectral clustering by alternating direction method of multipliers and its convergence analysis.* In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1). DOI: <https://doi.org/10.1609/aaai.v32i1.11658>
- [22] H. Oviedo, H. Lara, and O. Dalmau [2019], *A non-monotone linear search algorithm with mixed direction on Stiefel manifold*, Optimization Methods and Software, 34(2): 437–457. DOI: <https://doi.org/10.1080/10556788.2017.1415337>
- [23] H. Oviedo and H. Lara [2021], *Spectral residual method for nonlinear equations on Riemannian manifolds*, Computational and Applied Mathematics, 40(238): 1–23. DOI: <https://doi.org/10.1007/s40314-021-01630-3>
- [24] H. Oviedo, H. Lara and O. Dalmau [2021], *Two adaptive scaled gradient projection methods for Stiefel manifold constrained optimization*, Numerical Algorithms, 87: 1107–1127. DOI: <https://doi.org/10.1007/s11075-020-01001-9>
- [25] H. Oviedo [2022], *Global convergence of Riemannian line search methods with a Zhang-Hager-type condition*, Numerical Algorithms, 91(3): 1183–1203. DOI: <https://doi.org/10.1007/s11075-022-01298-8>
- [26] H. Oviedo [2023], *Proximal Point Algorithm with Euclidean Distance on the Stiefel Manifold*, Mathematics, 11(11): 2414. DOI: <https://doi.org/10.3390/math11112414>
- [27] N. Parikh and S. Boyd [2014], *Proximal Algorithms*, Foundations and Trends in Optimization, 1(3): 127–239. DOI: <http://dx.doi.org/10.1561/2400000003>

- [28] M. Plumbley [2004], *Lie group methods for optimization with orthogonality constraints*, In: Independent Component Analysis and Blind Signal Separation: Fifth International Conference, ICA 2004, Granada, Spain, September 22–24, 2004. Proceedings 5 (pp. 1245–1252). Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/978-3-540-30110-3_157
- [29] Y. Shen, Z. Wen and Y. Zhang [2014], *Augmented Lagrangian alternating direction method for matrix separation based on low-rank factorization*. Optim. Methods Softw., 29(2): 239–263. DOI: <https://doi.org/10.1080/10556788.2012.700713>
- [30] D. L. Sun and C. Fevotte [2014], *Alternating direction method of multipliers for non-negative matrix factorization with the beta-divergence*. In 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 6201–6205).
- [31] B. Sutton [2019], *Numerical construction of structured matrices with given eigenvalues*, Spec. Matrices, 7: 263–271. DOI: <https://doi.org/10.1515/spma-2019-0020>
- [32] N. Trinajstić, D. Babić, S. Nikolić, D. Plavšić, D. Amić, and Z. Mihalić [1994], *The Laplacian matrix in chemistry*, J. Chem. Znf. Comput. Sci., 34: 386–376. DOI: <https://doi.org/10.1021/ci00018a023>
- [33] S. Wasserman and K. Kaust [1994], *Social Networks Analysis– Methods and Applications*, Cambridge University Press, Cambridge. DOI: <https://doi.org/10.1017/CB09780511815478>
- [34] H. Zhang and W. W. Hager [2004], *A nonmonotone line search technique and its application to unconstrained optimization*, SIAM J. Optim., 14: 1043–1056. DOI: <https://doi.org/10.1137/S1052623403428208>