A Generalized Theorem of the Alternative for Certifiable Optimization with Redundant Constraints

Hanna Jiamei Zhang[1], Alan Papalia[2], Michael Everett[1,3], and David M. Rosen[1,3,4]

***Abstract*—Low-rank Semidefinite Programming via Burer–Monteiro (BM) factorization is a cornerstone of certifiably correct estimation in robotics. A key component is a theorem of the alternative that either certifies global optimality of a BM stationary point via positive semidefiniteness of a certificate matrix, or provides a second-order descent direction for continued search in a lifted dimension. However, current approaches require the linear independence constraint qualification (LICQ) to hold for the BM factorization, an assumption that necessarily fails when redundant constraints are added to tighten the semidefinite program (SDP) relaxation. Yet adding redundant constraints is a widely used technique for improving relaxation quality. We present a generalized theorem of the alternative that removes the LICQ requirement for BM factorized SDPs that does not depend on the LICQ. We show that global optimality can be certified via an eigenvalue maximization over the full multiplier set, and that this procedure preserves the certify-or-escape dichotomy: a nonnegative optimum certifies global optimality, while a negative optimum proves there are feasible points that reduce the objective.**

***Index Terms*—Burer-Monteiro Factorization, Riemannian Staircase, Semidefinite Programming, Certifiable Optimization, Redundant Constraints**

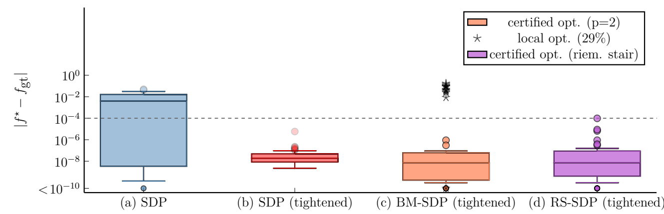

Fig. 1: Our generalized theorem of the alternative enables certified globally optimal solutions via the Riemannian Staircase even when redundant constraints violate the LICQ. We demonstrate this on a neural network verification problem [11], [12] across 50 random network instances. (a) The standard SDP leaves significant relaxation gaps. (b) Adding redundant constraints tightens the relaxation but causes the LICQ to fail. (c) The rank-2 BM factorization frequently recovers the ground truth; when it does not, our eigenvalue certificate (Eq. 2) reliably identifies local optima (marked $\star$, 29% of runs) with no false certifications. (d) Our proposed Riemannian Staircase (Algorithm 1) recovers the certified global optimum in all cases.

I. Introduction

SDP relaxations combined with low-rank BM factorization [1] have become a standard computational paradigm for certifiably correct estimation in robotic [2], [3], [4], [5], [6]. Central to this framework is a theorem of the alternative [7, Thm. 4], whose certify-or-escape dichotomy relies on the LICQ to guarantee unique Lagrange multipliers at stationary points: uniqueness is used both to construct the certificate matrix directly and to prove that a negative eigenvalue of the certificate matrix (i.e. non-PSDness) guarantees the solution is not a global optima (via Lagrangian duality) [7, Footnote 3]. Without unique multipliers, a certificate that fails the check is *inconclusive* rather than a proof of non- optimality [8].

In practice, however, introducing redundant constraints to tighten a relaxation necessarily violates the LICQ by construction, yet this technique is known to be necessary for SDP relaxations of certain important problems to produce non-trivial solutions. Matrix-weighted state-estimation problems lose tightness under anisotropic noise without redundant constraints [9], and the need for such constraints is pervasive enough that methods have been developed to generate them automatically [10]. Another instance is doubly nonnegative relaxations of completely positive programs, which inherit redundant entrywise nonnegativity constraints from the completely positive cone, constraints essential for tightness [11] yet necessarily violating LICQ.

Matrix-weighted state-estimation problems lose tightness under anisotropic noise without redundant constraints [9], and [10] develop methods to generate such "tightening" constraints automatically. Doubly nonnegative relaxations of completely positive programs, which inherit redundant entrywise nonnegativity constraints from the completely positive cone, have been proven effective in tightening of neural network verification problem formulations [11].

We present a *generalized theorem of the alternative* that relaxes the LICQ requirement by replacing the current state-of-the-art certification procedure (which requires uniqueness of Lagrange multipliers) with an eigenvalue maximization over a *set* of admissible Lagrange multipliers, preserving the certify-or-escape dichotomy of [7]. This extends the Riemannian Staircase (RS) framework to SDPs with redundant constraints, enabling certifiably correct low-rank optimization for a broader class of problems.

II. Background

Consider the following (convex) SDP:

$$f^\star_{\mathbb{S}_+} = \min_{X \in \mathbb{S}^n} \; f(X) \;\; \text{s.t. } \mathcal{A}(X) = b,\; \mathcal{B}(X) \leq u,\; X \succeq 0 \tag{SDP}$$

where $X \in \mathbb{S}^n_+$ is a positive semidefinite (PSD) matrix variable, $f : \mathbb{S}^n \to \mathbb{R}$ is a convex and twice-continuously-differentiable function, and $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^{m_1}$, and $\mathcal{B}$:

[1]Department of Computer Science, Northeastern University, Boston, MA.
[2]Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI.
[3]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA.
[4]Department of Mathematics, Northeastern University, Boston, MA.

$\mathbb{S}^n \to \mathbb{R}^{m_2}$ are linear operators defined by $\mathcal{A}(X)_i = \langle A_i, X \rangle, i \in [m_1]$ and $\mathcal{B}(X)_j = \langle B_j, X \rangle, j \in [m_2]$ for fixed sets of symmetric matrices $\{A_i\}_{i=1}^{m_1}$, $\{B_i\}_{j=1}^{m_2} \subset \mathbb{S}^n$, $b \in \mathbb{R}^{m_1}$, and $u \in \mathbb{R}^{m_2}$. For the convex problem (SDP), the Karush–Kuhn–Tucker (KKT) conditions (SDP-KKT) are always sufficient for global optimality [13]. If, additionally, Slater's condition holds for (SDP), then they are also necessary. The KKT conditions at a global minimizer $X \in \mathbb{S}^n$ with Lagrange multipliers $\lambda \in \mathbb{R}^{m_1}$, $\gamma \in \mathbb{R}_+^{m_2}$ (jointly $\mu = (\lambda, \gamma)$) are:

$$\begin{aligned} &\mathcal{A}(X) = b, \mathcal{B}(X) \leq u, && X \succeq 0 \\ &(\mathcal{B}(X) - u) \odot \gamma = 0,\ S(X,\mu)X = 0, && S(X,\mu) \succeq 0 \\ &S(X,\mu) \triangleq \nabla f(X) - \mathcal{A}^*(\lambda) - \mathcal{B}^*(\gamma), \end{aligned} \tag{SDP-KKT}$$

where $\mathcal{A}^* : \mathbb{R}^{m_1} \to \mathbb{S}^n$, and $\mathcal{B}^* : \mathbb{R}^{m_2} \to \mathbb{S}^n$ linear operators defined as $\mathcal{A}^*(\lambda) = \sum_{i=1}^{m_1} \lambda_i A_i$ and $\mathcal{B}^*(\gamma) = \sum_{j=1}^{m_2} \gamma_j B_j$.

Solving (SDP) directly is computationally expensive: the decision variable $X \in \mathbb{S}^n$ has $O(n^2)$ entries, and standard interior-point methods require $O(n^6)$ operations per iteration. Often the specific instances of (SDP) in certifiable estimation and robotics admit *low-rank* solutions $X^\star$ under realistic operating conditions. For such cases, Burer and Monteiro (BM) [1] proposed to algorithmically exploit this observation by reparameterizing X using an *assumed* rank-p factorization $X = YY^\mathsf{T}$ with *factor* $Y \in \mathbb{R}^{n \times p}$, $p \ll n$ with the aim of using a Y such that $\operatorname{rank}(X^\star) \leq p \ll n$. Substituting this parameterization into (SDP) yields the BM factorization, which enforces positive semi-definiteness by construction and reduces the number of decision variables from n^2 to np.

$$f^\star_{BM} = \min_{Y \in \mathbb{R}^{n\times p}} f(YY^\top) \ \text{ s.t. }\ \mathcal{A}(YY^\top) = b,\ \mathcal{B}(YY^\top) \leq u. \tag{BM-SDP}$$

Any constraint qualification (CQ) on (BM-SDP) ensures that Lagrange multipliers exist at any local minimizer, with stronger CQs such as the LICQ further guaranteeing their uniqueness. Let $Y^\star \in \mathbb{R}^{n \times p}$ be a first-order KKT point of (BM-SDP) with corresponding Lagrange multipliers $\mu = (\lambda, \gamma)$. The first-order (KKT) conditions at $Y^\star$ characterize first-order stationary points of the *nonconvex* problem (BM-SDP):

$$\begin{aligned} &\mathcal{A}(YY^\top) = b, \mathcal{B}(YY^\top) \leq u, \\ &(\mathcal{B}(YY^\top) - u) \odot \gamma = 0,\ 2S(Y,\mu)Y = 0 \\ &S(Y,\mu) \triangleq \nabla f(YY^\top) - \mathcal{A}^*(\lambda) - \mathcal{B}^*(\gamma), \end{aligned} \tag{BM-SDP-KKT}$$

Unlike the convex case, stationarity of (BM-SDP) is not sufficient to establish global optimality: local minimizers and saddle points of (BM-SDP) alike satisfy (BM-SDP-KKT).

We seek to recover global minimizers of the convex relaxation (SDP) by performing local optimization over low-rank factors Y via the nonconvex problem (BM-SDP). The central question is then: when does a stationary point of (BM-SDP) correspond to a global minimizer of (SDP)?

Comparing the KKT systems of the two problems provides a precise answer in the form of a theorem of the alternative [7, Thm. 4]. Given a KKT point $Y^\star \in \mathbb{R}^{n \times p}$ of (BM-SDP) with multipliers μ satisfying (BM-SDP-KKT), the matrix $X = Y^\star Y^{\star\top}$ automatically satisfies all conditions of (SDP-KKT) except possibly $S(Y,\mu) \succeq 0$ [7, Cor. 3]. If this remaining PSD condition holds, then X is a global minimizer of (SDP). Accordingly, $S(Y,\mu)$ is called the *certificate matrix* and the inclusion test $S(Y,\mu) \succeq 0$, the *certificate of optimality*.

The Riemannian Staircase [14] [7, Algo. 1] exploits this property: starting from a rank-p factorization, it iteratively recovers a KKT point $Y^\star$ of (BM-SDP) via local optimization, constructs the certificate matrix $S(Y,\mu)$, and either certifies global optimality of $X = Y^\star Y^{\star\top}$ for (SDP) when $S(Y,\mu) \succeq 0$, or escapes to rank $p{+}1$ along a second-order direction of descent that can be computed from $S(Y,\mu)$ [7, Thm. 4].

This certify-or-escape dichotomy relies on the LICQ holding at $Y^\star$, which guarantees that any Lagrange multipliers μ satisfying the KKT conditions of (BM-SDP) at $Y^\star$ are unique [15]. Uniqueness is imperative because it ensures that $S(Y,\mu) \not\succeq 0$ is conclusive: since no alternative multiplier exists, non-positive-semidefiniteness definitively proves non-optimality [7, Footnote 3]. Furthermore, under the LICQ it is possible to establish a valid second-order direction of descent to continue the search at higher rank-$(p+1)$ factorizations of (BM-SDP).

Without the LICQ, at a given stationary point $Y^\star$ there may exist more than one set of valid multipliers μ that satisfy (BM-SDP-KKT) for (BM-SDP), and a certificate $S(Y^\star,\mu^\star) \not\succeq 0$ constructed from the particular multipliers returned by the local non-linear program (NLP) solver is therefore inconclusive. Certification thus requires a *search* over the set of admissible Lagrange multipliers for one that renders $S(Y,\mu)$ positive semidefinite. Equivalently, this means finding a point in the intersection of the affine subspace defined by the KKT conditions (BM-SDP-KKT) and the PSD cone.

In this paper, we show that the theorem of the alternative proposed in [7], which underpins the certification and saddle escape procedures used in current state of the art implementations of BM-based estimators, can be extended to the case in which the LICQ is *not* satisfied.

III. A Generalized Theorem of the Alternative

A key limitation of the standard certification procedure is that when the LICQ does not hold for (BM-SDP), a certificate $S(Y^\star,\mu^\star) \not\succeq 0$ constructed from a particular multiplier set $\mu^\star$ is *inconclusive*, since alternative multipliers may yet yield $S(Y^\star,\mu) \succeq 0$.

We show that the search for an optimality certificate over the space of valid Lagrange multipliers can naturally can be expressed as an eigenvalue maximization problem. Furthermore, to establish a rigorous mechanism for properly identifying non-optimality of points of (BM-SDP) and continuing the search for the global minimizer of (SDP), we invoke the appropriate second order conditions.

A. Practical Search for Optimality Certificate

Without satisfaction of the LICQ at a KKT point $Y^\star$ of (BM-SDP), corresponding Lagrange multipliers $\mu^\star$ may not be unique. As per (BM-SDP-KKT) the set of valid multipliers associated with $Y^\star$ is defined as:

$$\Lambda(Y^\star) \triangleq \{\mu = (\lambda, \gamma) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}_+ : S(Y^\star, \mu)\, Y^\star = 0,\ \ (\mathcal{B}(Y^\star Y^{\star\top}) - u) \odot \gamma = 0\} \tag{1}$$

Note that this set is non-empty under Slater's CQ. Furthermore, if the local NLP method applied to (BM-SDP) converges to a KKT point, the returned multipliers $\mu^\star$ constitute one such element.

We formalize the search for a certificate as the following optimization problem: since any $\mu \in \Lambda$ yielding $S(Y^\star, \mu) \succeq 0$ suffices for optimality certification, we maximize the minimum eigenvalue of the certificate matrix $S(Y^\star, \mu)$ over the space of feasible multipliers $\mu \in \Lambda$ of (BM-SDP):

$$\begin{aligned} \max_{\mu=\lambda,\gamma} \quad & \text{eigmin} \circ S(Y^\star, \mu) \\ \text{s.t.} \quad & S(Y^\star, \mu)\, Y^\star = 0,\ c \odot \gamma = 0,\ \gamma \geq 0. \end{aligned} \tag{2}$$

where $Y^\star$ is the stationary point returned by the NLP solver and $c \triangleq (\mathcal{B}(Y^\star Y^{\star\top}) - u)$ is the constraint slack and eigmin $\circ$ (M) is an operation that extracts the minimum eigenvalue of square matrix M. The constraints of (2) enforce $\mu \in \Lambda$ (1).

If the optimal value of (2) is non-negative, the corresponding $\mu^\star$ yields $S(Y^\star, \mu^\star) \succeq 0$, certifying that $X = Y^\star Y^{\star\top}$ is a global minimizer of (SDP). If it is negative, then no valid multiplier μ renders $S(Y^\star, \mu)$ positive semidefinite. Under the relaxed Slater CQ for the convex problem (SDP) [7], the KKT conditions are necessary and sufficient for global optimality; the nonexistence of any $\mu \in \Lambda$ with $S(Y^\star, \mu) \succeq 0$ therefore certifies that $X = Y^\star Y^{\star\top}$ is *not* a global minimizer of (SDP).

B. Generalized Theorem

In the absence of the LICQ, the eigenvalue maximization over the multiplier set (2) determines whether a stationary point of (BM-SDP) corresponds to a global minimizer of (SDP). When it does not, we require a mechanism to escape the suboptimal point in order to continue the *local* optimization for a minimal low-rank factor in the Riemannian Staircase. Following [7], we embed the current iterate into a rank-$(p{+}1)$ instance of (BM-SDP) and show, via the second order necessary condition (SONC), that the lifted point is *not* a local optimum, implying the existence of a direction(s) of descent, enabling the continued search for the global optimum.

Definition 1 (Linearized cone): Let Y be a feasible point of (BM-SDP), and let $\mathcal{J}(Y) \triangleq \{j : \mathcal{B}_j(YY^\top) = u_j\}$ denote the *active* inequality constraints. The *linearized cone* at Y is the set of directions

$$L(Y) \triangleq \{V \in \mathbb{R}^{n \times p} \mid D(\mathcal{A}(YY^\top))[V] = 0,\ D(\mathcal{B}_j(YY^\top))[V] \leq 0\ \ \forall j \in \mathcal{J}(Y)\}. \tag{3}$$

Definition 2 (Tangent cone): The *tangent cone* $T(Y)$ at some feasible point Y is the set of directions $V \in \mathbb{R}^{n\times p}$ such that $V = 0$ or $V = \lim \dfrac{(Y^k - Y)}{||Y^k - Y||}$ for some feasible sequence $Y^k \to Y$.

The tangent cone (Def. 2) characterizes feasible directions based on set constraints, allowing optimization algorithms to stay within the constraint boundary.

Definition 3 (Abadie Constraint Qualification): We say a feasible point Y fulfills the Abadie constraint qualification (ACQ) when the tangent cone (Def. 2) equals the linearized cone (Def. 1), i.e., $T(Y) = L(Y)$. In other words, the ACQ guarantees that the *first order algebra* (i.e. the *linearizations* of the constraints at Y) correctly captures the set of first-order feasible directions (*geometry*) of the underlying feasible set.
The ACQ (Def. 3) is among the weakest conditions that guarantee the existence of Lagrange multipliers at local minimizers [16]. Stronger conditions such as the LICQ additionally ensure uniqueness of multipliers and simplify algorithmic design, but are not satisfied in settings with redundant constraints.

Definition 4 (Critical cone): Let $Y^\star$ be a KKT point of (BM-SDP) with multipliers $\mu = (\lambda, \gamma)$, and let $\mathcal{J}(Y^\star)$ denote the active inequality constraints as in Definition 1. The *strong critical cone* at $(Y^\star, \mu)$ is the set of directions

$$T_c(Y^\star) := \{V \in \mathbb{R}^{n\times p} \mid Df(YY^\top)[V] \leq 0,\ D\mathcal{A}(YY^\top)[V] = 0,\ D(\mathcal{B}_j(YY^\top))[V] \leq 0\ \ \forall j \in \mathcal{J}(Y) \tag{4}$$

while the smaller, *weak critical cone* $\tau_c(Y^\star)$ is the same but with the inequalities changed to equalities.
The critical cones (Def. 4) refine the linearized cone (Def. 1) by restricting to directions that do not increase the objective, making them the appropriate objects for first-order optimality conditions.

Theorem 1 (Weak Second-Order Necessary Optimality Conditions): [17, Thm. 3.1] Let $Y^\star$ be a local minimizer of (BM-SDP) at which the ACQ (Def. 3) holds. Then for every $\mu \in \Lambda(Y^\star)$:

$$\forall V \in \tau(Y^\star), \quad \langle V, \nabla^2_Y \mathcal{L}(Y^\star, \mu)[V]\rangle \geq 0 \tag{WSOC}$$

The Lagrangian of (BM-SDP) and can be simplified using the adjoint operators $\mathcal{A}^*$ and $\mathcal{B}^*$ as

$$\mathcal{L}(Y, \lambda, \gamma) \tag{5}$$
$$\triangleq f(YY^\top) + \langle \mathcal{A}(YY^\top) - b, \lambda\rangle + \langle \mathcal{B}(YY^\top) - u, \gamma\rangle \tag{6}$$
$$= f(YY^\top) + \langle \mathcal{A}^*(\lambda) + \mathcal{B}^*(\gamma), YY^\top\rangle - \langle \lambda, b\rangle - \langle \gamma, u\rangle \tag{7}$$

The gradient and hessian of the Lagrangian reduce to

$$\nabla_Y \mathcal{L}(Y, \lambda, \gamma) = 2S(Y, \mu)Y \tag{8}$$
$$\nabla^2_Y \mathcal{L}(Y, \lambda, \gamma) = 2S(Y, \mu) \tag{9}$$

We proceed by evaluating (WSOC) at the augmented stationary point Y_+ along a *restricted* class of augmented descent directions V, embedded within a higher-rank instance of the relaxation (BM-SDP). Let $v \in \mathbb{R}^n$ be any vector.

$$Y_+ \triangleq \begin{bmatrix} Y^\star & 0 \end{bmatrix} \in \mathbb{R}^{n\times(p+1)}, \quad V_+ \triangleq \begin{bmatrix} 0 & v \end{bmatrix} \in \mathbb{R}^{n\times(p+1)}. \tag{10}$$

To use (WSOC), the proposed augmented direction V_+ must belong to the weak critical cone $\tau(Y_+)$. By the chain rule, the gradient of any differentiable function $h(YY^\top)$ evaluated at the augmented point $Y_+ = [Y^\star \quad 0]$ takes the block form $\nabla_Y h = [2\nabla h(YY^\top)Y \quad 0]$. Because V_+ is non-zero only in the appended columns, its Frobenius inner product with any such gradient vanishes identically. Consequently, V_+ is strictly orthogonal to the gradients of the objective and all active constraints, fulfilling the algebraic requirements of the weak critical cone, $V_+ \in \tau(Y_+)$. Assuming the ACQ holds at Y_+ as per (WSOC) we can write the SONC as

$$\langle V_+, 2S(Y_+, \mu)V_+\rangle \geq 0, \quad \forall v \in \mathbb{R}^n,\ \forall \mu \in \Lambda \tag{11}$$

Substituting V_+ reduces the above to

$$2v^\top S(Y^\star, \mu)v \geq 0, \quad \forall v \in \mathbb{R}^n,\ \forall \mu \in \Lambda \tag{12}$$

If Y_+ is a local minimizer of (BM-SDP), the SONC (WSOC) reduces to (12). By the variational characterization of eigenvalues, $v^\top S(Y^\star, \mu)v \geq 0$ for all $v \in \mathbb{R}^n$ if and only if $S(Y^\star, \mu) \succeq 0$. Therefore, (12) holds if and only if $S(Y^\star, \mu) \succeq 0$ for *all* valid multipliers $\mu \in \Lambda$. By the contrapositive, if the necessary condition (12) fails for any single $\mu \in \Lambda$, this certifies that Y_+ is *not* a local minimizer of (BM-SDP) and implies the existence of points arbitrarily close to Y_+ with a strictly lower objective value. We summarize this as the following:

Theorem 2 (Generalized theorem of the alternative): Let $Y^\star \in \mathbb{R}^{n\times p}$ be a KKT point of (BM-SDP) at which the ACQ holds with Lagrange multiplier set Λ (1), and (SDP) satisfies Slater's CQ. Define the search for a valid optimality certificate as (2) yielding optimal value $\theta^\star$. Then exactly one of the following two cases hold:

(a) $\theta^\star \geq 0$: there exists $\mu^\star \in \Lambda$ with $S(Y^\star, \mu^\star) \succeq 0$, and $X = Y^\star Y^{\star\top}$ is a global minimizer of (SDP).
(b) $\theta^\star < 0$: $X = Y^\star Y^{\star\top}$ is not a global minimizer of (SDP), and $Y_+ = [Y^\star \quad 0]$ is not a local minimizer of (BM-SDP).

Theorem (2) immediately implies a RS-style approach described in Algorithm 1. We note, the violation of SONC (Theorem 2 b.) proves the existence of nearby points with strictly lower objective value, but does not immediately provide a feasible curve along which to descend.

Following [7, Algo. 1], we present the minimum eigenvector of $S(Y^\star, \mu^\star)$ corresponding to the minimum eigenvalue $\theta^\star_{\text{min}}$ obtained in our certificate search procedure (Eq. (2)) as a natural candidate descent direction and present evidence that this approach is effective in practice.

IV. Experiments

We validate Thm. 2 and Algo. 1 on the neural network verification problem [11], [12], using the doubly nonnegative (DNN) relaxation as a concrete instance of (BM-SDP) with redundant constraints (element-wise non-negativity constraints) that generically violate the LICQ. Full experimental details (network generation, solver configuration, and benchmark setup) can be found in our open-source Julia/JuMP implementation[1].

[1] https://github.com/hanjzh/dnn-sq

Algorithm 1 Riemannian Staircase without the LICQ

Input: Initial feasible point $Y \in \mathbb{R}^{n\times p}$ for p-dimensional instance of the Burer-Monteiro factorization (BM-SDP).
Output: Symmetric factor $Y^\star$ for a minimizer $X^\star = Y^\star Y^{\star\mathsf{T}}$ of SDP (SDP), optimal value $f^\star_{\mathbb{S}_+}$.

function RIEMANNIAN STAIRCASE(Y)
  loop
    // Find first-order KKT point of (BM-SDP)
    $(Y^\star, \mu^\star) \leftarrow$ LOCALOPTIMIZATION(Y).
      // Search for $S \succeq 0$ with (2).
    $(\theta^\star,\ S(\mu^\star)) \leftarrow$ CERTIFICATESEARCH($Y^\star$).
    if $\theta^\star \geq 0$ **then**
      // $Z^\star = Y^\star Y^{\star\mathsf{T}}$ minimizes (SDP) by Thm. 2(a)
      $f^\star_{\mathbb{S}_+} \leftarrow \langle Q, Y^\star Y^{\star\mathsf{T}}\rangle$.
      return $\{Y^\star,\ f^\star_{\mathbb{S}_+}\}$
    end if
      // Construct lifted embedding of $Y^\star$
    $Y_+ \leftarrow (Y^\star\ 0)$
      // Construct second-order descent direction from the minimum eigenvector v of $S(\mu^\star)$
    $V_+ \leftarrow (0\ v)$
      // Backtracking line search for saddle escape
    $Y \leftarrow$ LINESEARCH(Y_+, V_+)
  end loop
end function

The central experimental question is whether the eigenvalue maximization certificate search (Eq. 2) correctly distinguishes global from local optima when the LICQ fails, and whether the resulting minimum eigenvector works as an effective descent direction for the Riemannian Staircase. Fig. 1 confirms both across 50 random ReLU network instances: the certificate produces no false certifications, correctly identifying all local optima of the rank-2 BM factorization (29% of problems), and the minimum eigenvector-guided saddle escape of Algo. 1 recovers the certified global optimum in every case in support of the efficacy of the minimum eigenvector of $S(\mu^\star)$ as a descent direction.

V. Conclusion

We presented a generalized theorem of the alternative that extends the certify-or-escape framework of [7] to settings where the LICQ fails, replacing the single-multiplier certificate check with an eigenvalue maximization over the full multiplier set. This extends the Riemannian Staircase to SDPs tightened through redundant constraints.

Acknowledgment

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Northeastern University Institute for Experiential Robotics Postdoctoral Fellowship, and Army Research Lab awards W911NF-24-2-006 and W911NF-24-2-0017.

REFERENCES

[1] S. Burer and R. D. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical programming*, vol. 95, no. 2, pp. 329–357, 2003.

[2] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.

[3] F. Dellaert, D. M. Rosen, J. Wu, R. Mahony, and L. Carlone, "Shonan rotation averaging: Global optimality by surfing so (p) n," in *European Conference on Computer Vision*. Springer, 2020, pp. 292–308.

[4] A. Papalia, A. Fishberg, B. W. O'Neill, J. P. How, D. M. Rosen, and J. J. Leonard, "Certifiably correct range-aided slam," *IEEE Transactions on Robotics*, vol. 40, pp. 4265–4283, 2024.

[5] Y. Tian, K. Khosoussi, D. M. Rosen, and J. P. How, "Distributed certifiably correct pose-graph optimization," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 2137–2156, 2021.

[6] L. Wu and R. Tron, "Certifiably optimal estimation and calibration in robotics via trace-constrained semi-definite programming," *arXiv preprint arXiv:2509.23656*, 2025.

[7] D. M. Rosen, "Scalable low-rank semidefinite programming for certifiably correct machine perception," in *International Workshop on the Algorithmic Foundations of Robotics*. Springer, 2020, pp. 551–566.

[8] A. Papalia, Y. Tian, D. M. Rosen, J. P. How, and J. J. Leonard, "An overview of the burer-monteiro method for certifiable robot perception," 2025.

[9] C. Holmes, F. Dümbgen, and T. Barfoot, "On semidefinite relaxations for matrix-weighted state-estimation problems in robotics," *IEEE Transactions on Robotics*, vol. 40, pp. 4805–4824, 2024.

[10] F. Dümbgen, C. Holmes, B. Agro, and T. Barfoot, "Toward globally optimal state estimation using automatically tightened semidefinite relaxations," *IEEE Transactions on Robotics*, vol. 40, pp. 4338–4358, 2024.

[11] R. A. Brown, E. Schmerling, N. Azizan, and M. Pavone, "A unified view of sdp-based neural network verification through completely positive programming," in *International conference on artificial intelligence and statistics*. PMLR, 2022, pp. 9334–9355.

[12] A. Raghunathan, J. Steinhardt, and P. S. Liang, "Semidefinite relaxations for certifying robustness to adversarial examples," *Advances in neural information processing systems*, vol. 31, 2018.

[13] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[14] N. Boumal, "A riemannian low-rank method for optimization over semidefinite matrices with block-diagonal constraints," *arXiv preprint arXiv:1506.00575*, 2015.

[15] G. Wachsmuth, "On licq and the uniqueness of lagrange multipliers," *Operations Research Letters*, vol. 41, no. 1, pp. 78–80, 2013.

[16] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 2006.

[17] R. Andreani, R. Behling, G. Haeser, and P. J. Silva, "On second-order optimality conditions in nonlinear optimization," *Optimization Methods and Software*, vol. 32, no. 1, pp. 22–38, 2017.