Noname manuscript No.
(will be inserted by the editor)

Stratification for Nonlinear Semidefinite Programming: Geometry and Perturbation Analysis

Chenglong Bao · Chao Ding · Fuxiaoyue Feng · Jingyu Li

This version: August 29, 2026

Abstract Classical perturbation analysis imposes uniform requirements on all nearby perturbations and therefore does not distinguish whether they preserve or change the local geometric structure. We develop a geometry-resolved perturbation analysis for nonlinear semidefinite programming (NLSDP) based on a stratification of $\mathbb{S}^n$. We show that the positive semidefinite projection is C^∞ on every stratum and derive explicit manifold differentials for the smooth restrictions of the Karush–Kuhn–Tucker (KKT) mapping. We construct an auxiliary NLSDP and, based on it, two nested manifold optimization models that provide geometric interpretations of the regularity conditions: the constraint qualifications correspond to transversality conditions, while the second-order conditions correspond to manifold second-order conditions. At a KKT pair, this geometry relates injectivity of the manifold differential on its lifted stratum to the weak second order condition (W-SOC) and the weak strict Robinson constraint qualification (W-SRCQ). Across strata, these weak conditions have complementary inertia-dependent stability properties, while the local validity of the W-SRCQ and the local uniform validity of the W-SOC recover constraint nondegeneracy and the strong second order sufficient condition, respectively. This recovery of the classical strong conditions from weak regularity across nearby strata, in turn, provides a new and intuitive route to the known equivalence between the Aubin property of the inverse KKT mapping and strong regularity.

Key words: nonlinear semidefinite programming; stratification; variational analysis; perturbation analysis; strong metric regularity

---

Authors are listed alphabetically by surname.

C.L. Bao
Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, P.R. China; Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing 101408, P.R. China. The research of this author is supported in part by National Natural Science Foundation of China (12271291, 12625119) and New Cornerstone Investigator Program (100001127).
E-mail: clbao@tsinghua.edu.cn

C. Ding
State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P.R. China; School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P.R. China; Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P.R. China. The research of this author is supported in part by National Key R&D Program of China (No. 2021YFA1000300, No. 2021YFA1000301), National Natural Science Foundation of China (No. 12531014), and CAS Project for Young Scientists in Basic Research (No. YSBR-034).
E-mail: dingchao@amss.ac.cn

F.X.Y. Feng
Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, P.R. China; School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P.R. China.
E-mail: fengfuxiaoyue@amss.ac.cn

J.Y. Li
Qiuzhen College, Tsinghua University, Beijing 100084, P.R. China.
E-mail: lijingyu23@mails.tsinghua.edu.cn

1 Introduction

Perturbation analysis of optimization problems studies how solutions and multipliers respond to perturbations of an optimization problem or its optimality system. It underlies local stability theory and the local convergence analysis of numerical methods (see, e.g., [21,7]). Classical stability notions usually require regularity to hold uniformly under all sufficiently small perturbations in the ambient space. They do not separate perturbations that preserve a fixed local geometric structure from those that move across different structures. Consequently, ambient regularity may fail even though stratified regularity remains valid within the fixed structure. This paper takes a first step toward a perturbation analysis that makes this geometric distinction explicit.

For a general optimization problem, identifying and exploiting the local geometry preserved by perturbations can be difficult. Matrix optimization problems such as nonlinear semidefinite programming (NLSDP), however, provide a setting in which this geometry is explicit. In NLSDP, the semidefinite constraint takes values in $\mathbb{S}^n$, which admits a finite inertia stratification: symmetric matrices with the same numbers of positive and negative eigenvalues form a smooth manifold, called an *inertia stratum* (see Section 3 for details). The projection onto the positive semidefinite cone is smooth on each stratum, even where it is nonsmooth in the ambient space. This stratification organizes perturbations according to the inertia information they preserve: both inertia indices, only one index, or neither. We refer to an analysis that resolves perturbations in this way as *geometry-resolved perturbation analysis*. It allows stability to be studied relative to the geometry preserved by a perturbation, under conditions that can be weaker than those required for arbitrary ambient perturbations. Although we develop this analysis through the inertia stratification, the viewpoint may also extend to other families of geometric structures preserved by perturbations.

Against this background, we develop geometry-resolved perturbation analysis for NLSDP. This problem class is important in its own right, with applications in robust control, passive reduced-order modeling, and structural robust optimization [24,26,30]. We focus on the sensitivity of KKT pairs because the associated KKT mapping is a central computational object for numerical methods. Its local structure enters the convergence analysis of sequential semidefinite programming and semismooth Newton methods for NLSDP [26,54,25]. Related KKT linearizations are also used by Newton–CG augmented Lagrangian solvers for large-scale semidefinite programs, where degeneracy is a documented source of numerical difficulty [55]. Understanding the local structure of this mapping when classical ambient regularity fails is therefore important for both perturbation theory and convergence analysis.

Specifically, consider the following NLSDP problem

$$\begin{aligned} \min_{x\in\mathbb{X}} \quad & f(x) \\ \text{s.t.} \quad & g(x) \in \mathbb{S}^n_+, \end{aligned} \tag{1}$$

where $\mathbb{X}$ is a Euclidean space, the objective function $f : \mathbb{X} \to \mathbb{R}$ and the constraint mapping $g : \mathbb{X} \to \mathbb{S}^n$ are twice continuously differentiable, and $\mathbb{S}^n_+$ denotes the closed convex cone of symmetric positive semidefinite matrices. The NLSDP (1) naturally includes classical nonlinear programming, nonlinear second-order cone programming, and linear semidefinite programming as special cases.

A primal-dual pair $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ is called a Karush–Kuhn–Tucker (KKT) pair for (1) if it solves

$$F(z) := \begin{bmatrix} \nabla f(x) + \nabla g(x)y \\ -g(x) + \Pi_{\mathbb{S}^n_+}\big(G(z)\big) \end{bmatrix} = 0, \tag{2}$$

where $\Pi_{\mathbb{S}^n_+} : \mathbb{S}^n \to \mathbb{S}^n$ denotes the metric projection onto the positive semidefinite cone $\mathbb{S}^n_+$, the symbol ∇ denotes the gradient (i.e., the adjoint of the differential), and $G(z) := g(x)+y$. We write $L(x, y) := f(x)+\langle y, g(x)\rangle$ for the Lagrangian and $L(z) := L(x, y)$ when $z = (x, y)$. We call F the *KKT mapping* of (1). Its only nonsmooth term is $\Pi_{\mathbb{S}^n_+}$: although globally Lipschitz continuous and directionally differentiable, it fails to be Fréchet differentiable at any matrix with a zero eigenvalue (see, e.g., [14]). We thus know that F is continuously differentiable on an ambient neighborhood of a KKT pair $\bar{z}$ if and only if $G(\bar{z})$ is nonsingular, that is, strict complementarity holds at $\bar{z}$ (Definition 1). Classical regularity conditions may fail even when strict complementarity holds, for example when the Jacobian of F is singular, although F is still continuously differentiable near $\bar{z}$.

For the KKT system (2), a principal stability benchmark is strong regularity, introduced originally by Robinson [44, Section 2] for generalized equations. At a KKT pair, strong metric regularity (Definition 11) of F is equivalent to strong regularity of the KKT system. In nonlinear programming, strong regularity is characterized by the linear independence constraint qualification and the strong second order sufficient condition (S-SOSC) [44, 34,8,20]. For NLSDP, Sun [48, Theorem 4.1] proved that, at a local minimizer satisfying the Robinson constraint qualification (RCQ, Definition 2), strong regularity is equivalent to constraint nondegeneracy (Definition 4) and the S-SOSC (Definition 7). Under the same setting, Ding, Sun and Zhang [19, Theorem 24] characterized strong metric subregularity by the strict Robinson constraint qualification (SRCQ, Definition 3) and the second order sufficient condition (SOSC, Definition 6). At a KKT pair associated with a local minimizer, recent work further established the equivalence between the Aubin property of the inverse KKT mapping and strong metric regularity of the KKT mapping for nonlinear second-order cone programming, NLSDP, and, more generally, C^2-cone reducible conic programs [13,12,38]. These results give the classical characterizations of KKT stability with respect to arbitrary nearby ambient perturbations.

The strong conditions in these characterizations may fail at a *degenerate* KKT pair, that is, a pair at which strong regularity fails. As noted above, such a failure does not by itself imply that no useful regularity remains. It may instead arise from treating geometrically different perturbations. Conditions capturing this remaining regularity already exist. The weak strict Robinson constraint qualification (W-SRCQ, Definition 9) and the weak second order condition (W-SOC, Definition 10) were introduced by Feng, Ding and Li [25] in their analysis of semismooth Newton methods [43,42]. The W-SOC together with constraint nondegeneracy ensures the nonsingularity of the selected Bouligand element $\mathcal{U}_0$ of the KKT mapping F, while the S-SOSC together with the W-SRCQ ensures that of $\mathcal{U}_I$. Viewed through the inertia stratification, this raises three questions. What smooth geometry underlies the W-SRCQ and the W-SOC? What stability properties do these weak conditions exhibit? How do they relate to their classical strong counterparts? Our results show that the weak conditions are not merely algebraic relaxations of the classical ones. They describe the smooth manifold regularity associated with individual inertia strata, while the local validity of the W-SRCQ and the local uniform validity of the W-SOC connect this local geometry to the classical ambient theory.

Our results proceed in three stages. The first stage uses the inertia stratification to obtain smooth restrictions of the KKT mapping and extends the regularity conditions to arbitrary primal–dual pairs. The second one provides geometric interpretations of these conditions through two nested manifold optimization models. The third one analyzes regularity along and across strata, relating the stratified weak conditions to the classical ambient theory. Figure 1 gives a conceptual roadmap of the three main sections and highlights a small number of principal waypoints.

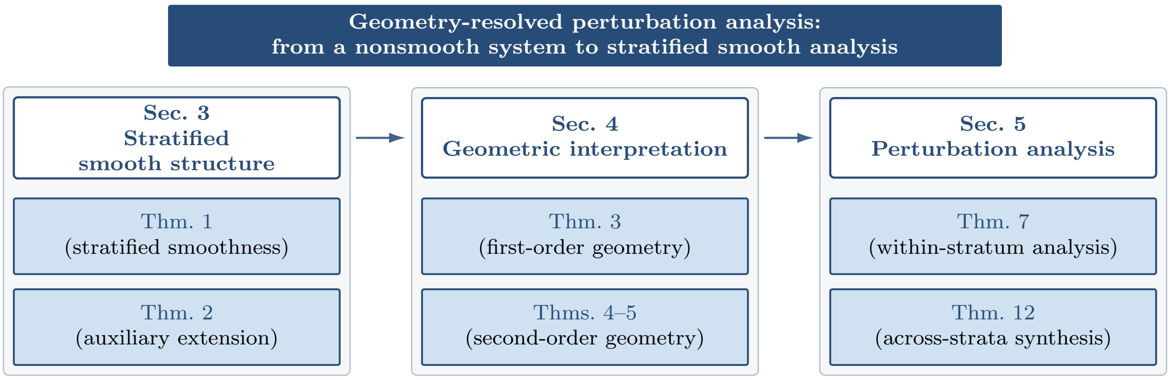

Fig. 1: Conceptual roadmap of the paper.

The construction starts from the inertia stratification of $\mathbb{S}^n$. We show that the projection onto the positive semidefinite cone is C^∞ on every stratum and derive its manifold differential, strengthening the C^1 differentiability available from [37, Example 4.14] and [23, Proposition 9.7]. Lifting this stratification of $\mathbb{S}^n$ through G yields a stratification of the primal–dual space, and the restriction of F to each lifted stratum is smooth, again with explicit differentials. Since comparisons across strata involve nearby pairs that need not satisfy the original KKT system, we extend the regularity conditions to arbitrary primal–dual pairs, using a non-vanishing formulation of the W-SOC away from the KKT set, and introduce an auxiliary problem $\mathcal{P}_z$ to interpret these conditions at such pairs. This auxiliary problem preserves the data entering these conditions and coincides with the original NLSDP when z is a KKT pair. We also construct two nested smooth manifold models for $\mathcal{P}_z$: $\mathcal{A}_z$ represents the local fixed-rank part of the semidefinite constraint, while $\mathcal{B}_z$ locally extends $\mathcal{A}_z$ in directions normal to the translated stratum containing it.

These two models answer the first question above. Constraint nondegeneracy and the W-SRCQ characterize transversality to $\mathcal{A}_z$ and $\mathcal{B}_z$, respectively, while the W-SOC and the strong second order condition (S-SOC) are the corresponding manifold second-order conditions. At a KKT pair, we further provide an outline suggesting how the selected Bouligand generalized Jacobian elements used in semismooth Newton analysis may arise as Jacobians of smooth KKT systems associated with the two models, up to an invertible change of primal–dual coordinates [25].

We next analyze regularity along a fixed lifted stratum and introduce *stratum-restricted strong metric regularity*, which requires the restricted KKT mapping to have a locally single-valued Lipschitz inverse on its local image. At a KKT pair, this property is equivalent to injectivity of the manifold differential. The W-SRCQ and the W-SOC together imply this injectivity. When the SRCQ and the second order necessary condition (SONC) also hold, the converse follows, so the weak pair characterizes regularity on the stratum. This characterization has a direct consequence for Newton linearizations. If the manifold differential is not injective, every element of the Clarke generalized Jacobian has a common nonzero kernel direction tangent to the stratum. Failure of stratum-restricted strong metric regularity at a KKT pair therefore rules out any nonsingular Clarke-based semismooth Newton linearization.

Across strata, the two weak conditions have complementary inertia-dependent stability properties. The W-SRCQ is stable under perturbations that preserve the negative inertia index, whereas the W-SOC is stable under those that preserve the positive inertia index. The local validity of the W-SRCQ and the local uniform validity of the W-SOC recover the classical strong conditions: constraint nondegeneracy is equivalent to the W-SRCQ holding at every nearby pair, and the S-SOSC is equivalent to the W-SOC holding uniformly on a neighborhood.

These results allow us to characterize strong regularity in terms of these stratified conditions, with the local uniform injectivity of the manifold differentials serving as the bridge. Since the Aubin property yields this local uniform injectivity, the known equivalence between the Aubin property and strong regularity follows as a by-product [12, Theorem 4.1].

The role of the stratification becomes clearer by comparison with methods that reduce the problem to a single smooth manifold. One classical approach is low-rank optimization. When the rank is prescribed, the problem can be reformulated either on a fixed-rank matrix manifold or through a low-rank factorization, allowing Riemannian optimization techniques to exploit the resulting smooth geometry and reduced-dimensional parametrization [1, 35, 50, 10, 11]. Such a fixed-rank formulation, however, requires the rank, and hence the relevant stratum, to be selected in advance. Another approach is to use partial smoothness to identify an active manifold [37, 18, 53]. For the semidefinite cone, however, the standard identification mechanism requires strict complementarity. More importantly for the present perturbation analysis, a single fixed manifold, whether prescribed or identified, does not by itself describe how regularity changes when perturbations move between neighboring strata. The analysis developed above works instead with the full inertia stratification: it characterizes regularity on each stratum and describes its stability across neighboring strata in terms of their inertia indices.

The remainder of this paper is organized as follows. Section 2 collects the preliminary material. Section 3 develops the inertia stratification, extends the regularity conditions to arbitrary primal–dual pairs, and introduces the auxiliary NLSDP. Section 4 constructs $\mathcal{A}_z$ and $\mathcal{B}_z$ and gives the geometric interpretations of the extended conditions. Section 5 characterizes stratum-restricted strong metric regularity through the weak conditions, establishes the inertia-dependent stability of the weak conditions across strata, and shows how their validity across all nearby strata recovers the classical strong conditions. Section 6 concludes with remarks on the scope and further directions.

2 Preliminaries

Let $\mathbb{X}$ be a Euclidean space. We denote the space of $n \times n$ symmetric matrices by $\mathbb{S}^n$, the cone of symmetric positive semidefinite matrices by $\mathbb{S}^n_+$, and set $\mathbb{S}^n_- := -\mathbb{S}^n_+$. We equip $\mathbb{S}^n$ with the trace inner product and use the corresponding product inner product on product spaces. Unsubscripted norms are the corresponding induced norms, while $\|\cdot\|_2$ denotes the spectral norm. The metric projection onto $\mathbb{S}^n_+$ is denoted by $\Pi_{\mathbb{S}^n_+}$. For a smooth mapping F between Euclidean spaces, $F'(x)$ denotes its Fréchet derivative (Jacobian), and $\nabla F(x)$ denotes its adjoint (or gradient, if F is real-valued). For a directionally differentiable function f, $f'(x; d)$ denotes its directional derivative at x along the direction d. Additional notation is summarized as follows:

- A_{ij}: the (i, j)-th entry of a matrix $A \in \mathbb{R}^{m\times n}$.
- A_i: the i-th column of $A \in \mathbb{R}^{m\times n}$.
- $A_{\alpha\beta}$: the submatrix of A consisting of the rows and columns indexed by the sets α and β, respectively.
- $\circ$: the Hadamard (elementwise) product of matrices. For $A, B \in \mathbb{R}^{m\times n}$, $(A \circ B)_{ij} = A_{ij}B_{ij}$.
- $\mathrm{Diag}(a)$: the diagonal matrix generated by a vector a; more generally, $\mathrm{Diag}(A_1, \ldots, A_k)$ denotes the block-diagonal matrix with square diagonal blocks $A_1, \ldots, A_k$.
- $\mathrm{diag}(A)$: the vector consisting of the main diagonal entries of A.
- $C + D := \{c + d \mid c \in C,\ d \in D\}$: the Minkowski sum of subsets C, D of the same Euclidean space.
- $a + C = C + a := \{a + c \mid c \in C\}$: the translation of a set C by a point a in the same Euclidean space.
- $LC := \{Lc \mid c \in C\}$: the image of a set C under a linear mapping L, where C is contained in the domain of L.

- $\mathrm{lin}(C)$: the linearity space of a closed convex set $C \subseteq \mathbb{X}$ (i.e., the largest linear subspace contained in C).
- $\mathrm{aff}(C)$: the affine hull of C (i.e., the smallest affine set containing C).
- $\mathcal{M}_{p,q}$ (or $\mathcal{M}$): the manifold of symmetric matrices in $\mathbb{S}^n$ with exactly p positive and q negative eigenvalues.
- $\widetilde{\mathcal{M}}_{p,q}$ (or $\widetilde{\mathcal{M}}$): the preimage $G^{-1}(\mathcal{M}_{p,q})$ (see (11)).
- $\mathrm{Im}(\cdot)$: the image of a map.
- $\oplus$: the direct sum of vector spaces.
- $A|_W$: the restriction of a self-adjoint operator or quadratic form A to a linear subspace W, identified with the associated self-adjoint operator on W.
- $\sigma_{\min}(A)$: the minimum absolute eigenvalue of a self-adjoint operator A; on the zero space, it is $+\infty$.
- $\succeq$ ($\preceq$) and $\succ$ ($\prec$): the Loewner partial order. $A \succeq B$ ($A \preceq B$) means $A-B$ is positive (negative) semidefinite, and $A \succ B$ ($A \prec B$) means $A - B$ is positive (negative) definite.

2.1 Variational properties of $\Pi_{\mathbb{S}^n_+}$

This subsection reviews fundamental properties of the metric projection $\Pi_{\mathbb{S}^n_+}$. For a given matrix $A \in \mathbb{S}^n$, let its eigenvalues (counted with multiplicities) be ordered as

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A).$$

Let $\lambda(A)$ be the vector of these ordered eigenvalues, and define $\Lambda(A) := \mathrm{Diag}(\lambda(A))$. We partition the index set $\{1, \ldots, n\}$ into $\alpha(A)$, $\beta(A)$, and $\gamma(A)$ according to the signs of the eigenvalues:

$$\alpha(A) := \{i \mid \lambda_i(A) > 0\}, \quad \beta(A) := \{i \mid \lambda_i(A) = 0\} \quad \text{and} \quad \gamma(A) := \{i \mid \lambda_i(A) < 0\}.$$

When the context is unambiguous, we suppress the explicit dependence on A and simply write α, β, and γ. The positive and negative inertia indices of A are denoted by $p := |\alpha|$ and $q := |\gamma|$, respectively. The eigenvalue decomposition of A takes the block form

$$A = \begin{bmatrix} P_\alpha & P_\beta & P_\gamma \end{bmatrix} \begin{bmatrix} \Lambda(A)_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda(A)_{\gamma\gamma} \end{bmatrix} \begin{bmatrix} P_\alpha^\top \\ P_\beta^\top \\ P_\gamma^\top \end{bmatrix}, \tag{3}$$

where the orthogonal matrix $P \in \mathcal{O}^n$ is partitioned accordingly. Let $\mathcal{O}^n(A)$ denote the set of all such orthogonal matrices P satisfying (3). We refer to the tuple $(\alpha, \beta, \gamma, p, q, P, \lambda)$ as an *indexed eigenvalue decomposition (IED)* of A.

Remark 1 We refer to *an* IED rather than *the* IED because the orthogonal matrix $P \in \mathcal{O}^n(A)$ is generally not unique. Nevertheless, our subsequent constructions and analytical results are invariant to the specific choice of P.

The metric projection of A onto $\mathbb{S}^n_+$ is explicitly given by $\Pi_{\mathbb{S}^n_+}(A) = P_\alpha \Lambda(A)_{\alpha\alpha} P_\alpha^\top$. Furthermore, $\Pi_{\mathbb{S}^n_+}$ is globally Lipschitz continuous and directionally differentiable (i.e., B-differentiable [5,6]) and strongly semismooth [49], although it is generally not Fréchet differentiable. Specifically, the directional derivative of $\Pi_{\mathbb{S}^n_+}$ at A along the direction $H \in \mathbb{S}^n$ is formulated as

$$\Pi'_{\mathbb{S}^n_+}(A; H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\alpha\beta}^\top & \Pi_{\mathbb{S}^{|\beta|}_+}(\widetilde{H}_{\beta\beta}) & 0 \\ \Xi_{\alpha\gamma}^\top \circ \widetilde{H}_{\alpha\gamma}^\top & 0 & 0 \end{bmatrix} P^\top, \tag{4}$$

where $P \in \mathcal{O}^n(A)$, $\widetilde{H} := P^\top H P$, and the symmetric weight matrix $\Xi \in \mathbb{S}^n$ has entries defined by

$$\Xi_{ij} := \frac{\max\{\lambda_i(A), 0\} - \max\{\lambda_j(A), 0\}}{\lambda_i(A) - \lambda_j(A)}, \quad i, j = 1, \ldots, n,$$

adopting the convention that $0/0 := 1$.

2.2 Regularity conditions

In this subsection, we collect several variational notions and regularity conditions that will be used throughout the paper. A comprehensive account of the background material can be found, e.g., in [45,7].

We begin with the tangent cone. For a closed convex set $\mathcal{K} \subseteq \mathbb{S}^n$, the tangent cone to $\mathcal{K}$ at $y \in \mathcal{K}$ is defined by

$$T_{\mathcal{K}}(y) := \big\{ d \in \mathbb{S}^n \mid \exists\, t_k \downarrow 0 \text{ such that } \operatorname{dist}(y + t_k d, \mathcal{K}) = o(t_k) \big\}.$$

When $\mathcal{K}$ is a smooth embedded submanifold $\mathcal{M}$ of a Euclidean space $\mathbb{X}$, the variational tangent cone $T_{\mathcal{M}}(x)$ coincides with the classical tangent space $T_x\mathcal{M}$. To keep track of whether we are working in the variational or differential-geometric convention, we retain both notations: we write $T_{\mathcal{K}}(y)$ for the tangent cone to a set $\mathcal{K}$ at y, and $T_A\mathcal{M}$ for the tangent space to a manifold $\mathcal{M}$ at A. In particular, switching the positions of the set and the base point will always indicate the intended geometric context.

We next recall strict complementarity for a KKT pair.

Definition 1 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair for (1). We say that *strict complementarity* holds at $\bar{z}$ if

$$\operatorname{rank}(g(\bar{x})) + \operatorname{rank}(\bar{y}) = n.$$

Equivalently, $G(\bar{z})$ is nonsingular, or, in terms of an IED of $G(\bar{z})$, $\beta = \emptyset$.

Let x be a local minimizer of (1). As shown in [7, Theorem 3.9 and Proposition 3.17], the Robinson constraint qualification (RCQ) stated below is equivalent to the nonemptiness and boundedness of the associated Lagrange multiplier set $M(x)$.

Definition 2 The *Robinson constraint qualification (RCQ)* holds at a feasible point x if

$$g'(x)\mathbb{X} + T_{\mathbb{S}^n_+}(g(x)) = \mathbb{S}^n.$$

We next recall the classical SRCQ and constraint nondegeneracy. Their extensions to arbitrary primal–dual pairs are given in Section 3.2.

Definition 3 Let $\bar{x}$ be a feasible point of (1), and let $\bar{y} \in M(\bar{x})$. The *strict Robinson constraint qualification (SRCQ)* holds at $(\bar{x}, \bar{y})$ if

$$g'(\bar{x})\mathbb{X} + T_{\mathbb{S}^n_+}(g(\bar{x})) \cap \bar{y}^{\perp} = \mathbb{S}^n.$$

Under the classical SRCQ, the multiplier set $M(\bar{x})$ is a singleton [7, Proposition 4.50].

Constraint nondegeneracy is stated in the standard form of [7, Definition 5.70].

Definition 4 Let $\bar{x}$ be a feasible point of (1). We say that *constraint nondegeneracy* holds at $\bar{x}$ if

$$g'(\bar{x})\mathbb{X} + \operatorname{lin}\big(T_{\mathbb{S}^n_+}(g(\bar{x}))\big) = \mathbb{S}^n.$$

We now state classical second order sufficient conditions at a KKT pair $\bar{z} = (\bar{x}, \bar{y})$. Define the critical cone

$$\mathcal{C}(\bar{z}) := \{ v_x \in \mathbb{X} \mid \nabla g(\bar{x})^* v_x \in T_{\mathbb{S}^n_+}(g(\bar{x})),\ \langle \nabla f(\bar{x}), v_x \rangle = 0 \}.$$

Following Sun [48, (38)–(40)], we use $\operatorname{app}(\bar{z})$ as a computable surrogate for $\operatorname{aff}(\mathcal{C}(\bar{z}))$. If $(\alpha, \beta, \gamma, p, q, P, \lambda)$ is an IED of $G(\bar{z})$, then

$$\mathcal{C}(\bar{z}) = \left\{ v_x \in \mathbb{X} \mid [P^\top(\nabla g(\bar{x})^* v_x)P]_{\beta\beta} \succeq 0,\ [P^\top(\nabla g(\bar{x})^* v_x)P]_{\beta\gamma} = 0,\ [P^\top(\nabla g(\bar{x})^* v_x)P]_{\gamma\gamma} = 0 \right\},$$

and

$$\operatorname{app}(\bar{z}) = \left\{ v_x \in \mathbb{X} \mid [P^\top(\nabla g(\bar{x})^* v_x)P]_{\beta\gamma} = 0,\ [P^\top(\nabla g(\bar{x})^* v_x)P]_{\gamma\gamma} = 0 \right\}.$$

In general, $\operatorname{aff}(\mathcal{C}(\bar{z})) \subseteq \operatorname{app}(\bar{z})$, and the two sets need not coincide.

The second-order conditions below use the following second-order tangent set from [7, Definition 3.28].

Definition 5 For a closed set $C \subseteq \mathbb{S}^n$, a point $\bar{A} \in C$, and a direction $H \in T_C(\bar{A})$, the *second-order tangent set* of C at $\bar{A}$ in direction H is

$$T_C^2(\bar{A}, H) := \big\{ W \in \mathbb{S}^n : \exists\, t_k \downarrow 0,\ W_k \to W \text{ with } \bar{A} + t_k H + \tfrac{1}{2} t_k^2 W_k \in C \big\},$$

The following SOSC is adapted from [7, (3.276)].

Definition 6 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of (1). The *second order sufficient condition (SOSC)* holds at $\bar{z}$ if

$$\langle v_x, \nabla^2_{xx} L(\bar{x}, \bar{y}) v_x \rangle - \sigma\left(\bar{y}, T^2_{\mathbb{S}^n_+}(g(\bar{x}), g'(\bar{x})v_x)\right) > 0 \quad \forall\, v_x \in \mathcal{C}(\bar{z}) \setminus \{0\}.$$

Strengthening the direction set in the SOSC gives the strong SOSC (S-SOSC) of [48, Definition 3.2].

Definition 7 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of (1). The *strong second order sufficient condition (S-SOSC)* holds at $\bar{z}$ if

$$\langle v_x, \nabla^2_{xx} L(\bar{x}, \bar{y}) v_x \rangle - \sigma\left(\bar{y}, T^2_{\mathbb{S}^n_+}(g(\bar{x}), g'(\bar{x})v_x)\right) > 0 \quad \forall\, v_x \in \operatorname{app}(\bar{z}) \setminus \{0\}.$$

The corresponding classical second order necessary condition for NLSDP is given in [46, Theorem 8].

Definition 8 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of (1). The *second order necessary condition (SONC)* holds at $\bar{z}$ if

$$\langle v_x, \nabla^2_{xx} L(\bar{x}, \bar{y}) v_x \rangle - \sigma\left(\bar{y}, T^2_{\mathbb{S}^n_+}(g(\bar{x}), g'(\bar{x})v_x)\right) \geq 0 \quad \forall\, v_x \in \mathcal{C}(\bar{z}).$$

We next recall the W-SRCQ and the W-SOC of Feng, Ding and Li [25].
Following [25, Definition 8 and (20)], define

$$\operatorname{appl}(\bar{z}) := \left\{ v_x \in \mathbb{X} \mid g'(\bar{x})v_x \in \operatorname{lin}\!\left(T_{\mathbb{S}^n_+}(g(\bar{x})) \cap \bar{y}^{\perp}\right) \right\}.$$

Definition 9 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of (1) with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(\bar{z})$. The *weak strict Robinson constraint qualification (W-SRCQ)* holds at $\bar{z}$, as in [25, Definition 7], if

$$g'(\bar{x})\mathbb{X} + \left\{ PBP^{\top} \mid B \in \mathbb{S}^n,\ B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0 \right\} = \mathbb{S}^n.$$

Definition 10 Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of (1) with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(\bar{z})$. The *weak second order condition (W-SOC)* holds at $\bar{z}$, as in [25, Definition 8 and (19)], if

$$\langle v_x, \nabla^2_{xx} L(\bar{x}, \bar{y}) v_x \rangle - \sigma\left(\bar{y}, T^2_{\mathbb{S}^n_+}(g(\bar{x}), g'(\bar{x})v_x)\right) > 0 \quad \forall\, v_x \in \operatorname{appl}(\bar{z}) \setminus \{0\}.$$

The conditions above satisfy the following implication relations:

$$\text{constraint nondegeneracy} \implies \text{SRCQ} \implies \text{W-SRCQ}, \qquad \text{SRCQ} \implies \text{RCQ}$$

and

$$\text{S-SOSC} \implies \text{SOSC} \implies \text{W-SOC}.$$

More specifically, Feng, Ding and Li [25] relate the W-SRCQ and the W-SOC to the nonsingularity of special generalized-Jacobian elements of the KKT mapping and reveal a primal–dual relationship between the two conditions. In Section 3.2, we extend them from KKT pairs to arbitrary primal–dual pairs for the analysis along lifted strata.

The constraint qualifications and second-order conditions introduced in Definitions 2–10 will be referred to as *problem-level* regularity conditions, in the sense that they are formulated directly in terms of the problem data (derivatives and spectral information) and can, at least in principle, be checked without solving auxiliary variational systems. In contrast, we call the conditions introduced next *solution-level* regularity conditions, since they are typically expressed through properties of set-valued mappings and often involve solving equations and/or evaluating distances, which can be substantially more demanding computationally. These solution-level notions coincide with the conventional meaning of regularity in variational analysis. Nonetheless, we also view the above constraint qualifications and second-order conditions as regularity conditions, because they capture the geometric structure underlying a well-behaved local solution mapping. Among solution-level conditions, the most prominent one is *strong metric regularity*, which we recall next.

Definition 11 We say a set-valued mapping $\Phi : \mathbb{X} \rightrightarrows \mathbb{Y}$ is *strongly metrically regular* at $(\overline{x}, \overline{y})$ if $\overline{y} \in \Phi(\overline{x})$ and there exist constants $\kappa > 0$ and neighborhoods $\mathcal{U}$ of $\overline{x}$, $\mathcal{V}$ of $\overline{y}$, such that for all $x \in \mathcal{U}$ and $y \in \mathcal{V}$,

$$\operatorname{dist}(x, \Phi^{-1}(y)) \leq \kappa \cdot \operatorname{dist}(y, \Phi(x)),$$

and $\Phi^{-1}(y) \cap \mathcal{U}$ is a singleton for every $y \in \mathcal{V}$.

It is well known that, at a KKT pair of (1), the strong metric regularity of F is equivalent to the strong regularity [44] of the KKT system formulated as a generalized equation [21]. Moreover, under the RCQ at a local optimum, this strong regularity is in turn equivalent to the simultaneous satisfaction of the S-SOSC (Definition 7) and constraint nondegeneracy (Definition 4) [48]. Under the same assumptions, strong metric subregularity of F is equivalent to the conjunction of the SRCQ and the SOSC [19, Theorem 24].

Metric regularity is a weaker property, as it does not require the inverse mapping to be single-valued.

Definition 12 We say a set-valued mapping $\Phi : \mathbb{X} \rightrightarrows \mathbb{Y}$ is *metrically regular* at $(\overline{x}, \overline{y})$ if $\overline{y} \in \Phi(\overline{x})$ and there exist a constant $\kappa > 0$ and neighborhoods $\mathcal{U}$ of $\overline{x}$, $\mathcal{V}$ of $\overline{y}$, such that for all $x \in \mathcal{U}$ and $y \in \mathcal{V}$,

$$\operatorname{dist}(x, \Phi^{-1}(y)) \leq \kappa \cdot \operatorname{dist}(y, \Phi(x)).$$

Metric regularity is generally weaker than strong metric regularity. For KKT mappings arising from nonlinear second-order cone programs, NLSDP, and more generally C^2-cone reducible conic programs, the two properties nevertheless coincide [13, 12, 38]. For NLSDP, this equivalence will also follow in Theorem 12 as a consequence of the stratified perturbation results.

Another closely related notion is the *local error bound*. Various error bound formulations appear in the literature, and a systematic discussion is beyond the scope of this paper. We record one standard version below, noting that it follows directly from metric regularity.

Definition 13 Let $F : \mathbb{X} \to \mathbb{Y}$ be a single-valued mapping, and let $\overline{x} \in \mathbb{X}$ satisfy $F(\overline{x}) = 0$. We say that the equation $F(x) = 0$ admits a *local error bound* at $\overline{x}$ if there exist a constant $\kappa > 0$ and a neighborhood $\mathcal{U}$ of $\overline{x}$ such that for all $x \in \mathcal{U}$,

$$\operatorname{dist}\big(x, F^{-1}(0)\big) \leq \kappa \cdot \|F(x)\|.$$

We also recall the generalized Jacobians used below [16, Section 2.6].

Definition 14 Let $H : \mathbb{X} \to \mathbb{Y}$ be locally Lipschitz continuous, and let D_H denote the set of points at which H is Fréchet differentiable. The *Bouligand generalized Jacobian* of H at $\bar{x}$ is

$$\partial_B H(\bar{x}) := \left\{ \mathcal{V} \in \mathcal{L}(\mathbb{X}, \mathbb{Y}) \;\middle|\; \begin{array}{l} \text{there exists } x^k \in D_H \text{ with } x^k \to \bar{x} \\ \text{and } H'(x^k) \to \mathcal{V} \end{array} \right\}.$$

The *Clarke generalized Jacobian* of H at $\bar{x}$ is

$$\partial_C H(\bar{x}) := \operatorname{conv} \partial_B H(\bar{x}).$$

2.3 Manifolds and stratification

For standard definitions and notation concerning tangent spaces $T_x\mathcal{M}$, differentials dh_x, and exponential maps $\exp_x$, we refer to [36, 1].

We first clarify a potential notational ambiguity. Let $\mathcal{M}$ be a manifold embedded in a Euclidean space $\mathbb{Y}$, and let $p \in \mathcal{M}$. Under the canonical identification $T_p\mathbb{Y} \cong \mathbb{Y}$, we identify any tangent vector $v \in T_p\mathbb{Y}$ with an element of $\mathbb{Y}$. This natural identification permits the vector sum $p + v$ within the ambient space $\mathbb{Y}$. To distinguish the underlying geometries, we use dh_x to denote the differential of a mapping between manifolds, and reserve $h'(x)$ for the Fréchet derivative of mappings defined on Euclidean spaces. In accordance with standard optimization conventions, the adjoint of the linear operator $h'(x)$ is denoted by $\nabla h(x) := (h'(x))^*$.

Next, we recall the concept of transversality from differential topology (see, e.g., [32, Chapter 3.2]). This topological concept, which slightly generalizes the definition typically employed in variational analysis [7], formalizes the geometric intuition of a smooth map intersecting a submanifold in a "regular" manner.

Definition 15 ([32, Chapter 3.2]) Suppose $h : \mathcal{M} \to \mathcal{N}$ is a smooth mapping between smooth manifolds $\mathcal{M}$ and $\mathcal{N}$, where $\mathcal{N}' \subseteq \mathcal{N}$ is a submanifold of $\mathcal{N}$. We say h *intersects with* $\mathcal{N}'$ *transversally* at x in $\mathcal{N}$, or simply h *is transverse to* $\mathcal{N}'$ at x in $\mathcal{N}$, denoted by $h \pitchfork_x \mathcal{N}'$ in $\mathcal{N}$, whenever $h(x) \in \mathcal{N}'$,

$$dh_x(T_x\mathcal{M}) + T_{h(x)}\mathcal{N}' = T_{h(x)}\mathcal{N}, \tag{5}$$

where dh_x is the differential of h at x and T denotes the tangent space. Moreover, for a given closed subset $\mathcal{L}$ in $\mathcal{M}$, if (5) holds whenever $x \in \mathcal{L}$ and $h(x) \in \mathcal{N}'$, then we say h *intersects with* $\mathcal{N}'$ *transversally along* $\mathcal{L}$ in $\mathcal{N}$, denoted by $h \pitchfork_{\mathcal{L}} \mathcal{N}'$ in $\mathcal{N}$. We also simply denote $h \pitchfork \mathcal{N}'$ in $\mathcal{N}$ if $\mathcal{L} \equiv \mathcal{M}$.

Stratification is a fundamental notion in topology and singularity theory. It is well known that smooth manifolds provide a flexible framework that covers many nonlinear spaces, such as orthogonal groups and Grassmannians. However, sets with singularities lie outside the class of manifolds. The concept of a stratified space is designed as a natural extension of the manifold framework to accommodate singular sets, such as real algebraic varieties [51,52], which arise frequently in polynomial optimization and related areas [33]. For background on (Whitney) stratifications, we refer the reader to [51,52,39,40]. In the present work, the following basic definition will suffice for our subsequent developments.

Definition 16 Let $\mathbb{X}$ be a Euclidean space and $\mathcal{X} \subseteq \mathbb{X}$ a closed subset. A *stratification* of $\mathcal{X}$ is a locally finite partition $S = \{\mathcal{S}_i\}_{i\in I}$ of $\mathcal{X}$ into connected embedded C^k submanifolds (typically $k \geq 1$) $\mathcal{S}_i$ (called *strata*) such that

1. *Frontier condition*: For any two strata $\mathcal{S}_i, \mathcal{S}_j \in S$, if $\mathcal{S}_i \cap \overline{\mathcal{S}_j} \neq \emptyset$, then $\mathcal{S}_i \subseteq \overline{\mathcal{S}_j}$, where $\overline{\mathcal{S}_j}$ means the closure of $\mathcal{S}_j$. This implies the boundary $\overline{\mathcal{S}_j} \setminus \mathcal{S}_j$ is a union of strata.
2. *Local finiteness*: Every point $x \in \mathcal{X}$ has a neighborhood in $\mathbb{X}$ intersecting finitely many strata.

A topological space with a stratification structure is called a *stratified space*.

Roughly speaking, a stratification may be viewed as a decomposition of a given set into smooth pieces together with their well-organized boundaries. In Section 3, we shall introduce a stratification of the Euclidean space $\mathbb{S}^n$ tailored to our purposes, even though $\mathbb{S}^n$ is itself a linear space and hence a smooth manifold without singularities.

2.4 Manifold optimization and second-order constructions

We first fix the terminology for optimization over an embedded manifold.

The Riemannian gradient and stationary points are reviewed in [1, Sections 3.6 and 4.1]. The Riemannian Hessian is defined in [1, Definition 5.5.1]. It is self-adjoint by [1, Proposition 5.5.3], and [1, Proposition 5.5.4] identifies its quadratic form with the second derivative along exponential curves. Self-adjointness and the polarization identity ensure that this quadratic form uniquely determines the Hessian operator.

Definition 17 Let $\mathcal{M}$ be a smooth embedded submanifold of a Euclidean space $\mathbb{X}$ endowed with the induced Riemannian metric, and let $f : \mathbb{X} \to \mathbb{R}$ be twice continuously differentiable. Consider the manifold optimization problem

$$\min_{u\in\mathcal{M}} f(u). \tag{6}$$

(a) The *Riemannian gradient* of $f|_{\mathcal{M}}$ at $x \in \mathcal{M}$ is the unique vector $\mathrm{grad}(f|_{\mathcal{M}})(x) \in T_x\mathcal{M}$ satisfying

$$\langle \mathrm{grad}(f|_{\mathcal{M}})(x), d\rangle = f'(x)d \qquad \forall d \in T_x\mathcal{M}.$$

A point $\bar{x} \in \mathcal{M}$ is *stationary* for (6) if

$$\mathrm{grad}(f|_{\mathcal{M}})(\bar{x}) = 0, \quad \text{equivalently,} \quad f'(\bar{x})d = 0 \quad \forall d \in T_{\bar{x}}\mathcal{M}.$$

(b) The *Riemannian Hessian* of $f|_{\mathcal{M}}$ at $x \in \mathcal{M}$ is the unique self-adjoint linear operator $\mathrm{Hess}(f|_{\mathcal{M}})(x) : T_x\mathcal{M} \to T_x\mathcal{M}$ satisfying

$$\langle d, \mathrm{Hess}(f|_{\mathcal{M}})(x)d\rangle = \frac{d^2}{dt^2} f\big(\exp_x(td)\big)\Big|_{t=0} \qquad \forall d \in T_x\mathcal{M},$$

where exp is the Riemannian exponential map on $\mathcal{M}$.

(c) The *manifold second-order condition* (manifold SOC) holds at a stationary point $\bar{x}$ if

$$\langle d, \mathrm{Hess}(f|_{\mathcal{M}})(\bar{x})d\rangle \neq 0 \qquad \forall\, 0 \neq d \in T_{\bar{x}}\mathcal{M}.$$

Remark 2 In manifold optimization, the standard manifold second-order sufficient condition (manifold SOSC) requires strict positivity. Here we instead consider a manifold SOC requiring only non-vanishing, because the arbitrary primal–dual pairs studied in Section 4.3 need not correspond to local minimizers. When the manifold SOC holds, continuity implies that the Riemannian Hessian quadratic form has a constant sign on $T_{\bar{x}}\mathcal{M} \setminus \{0\}$. Hence $\mathrm{Hess}(f|_{\mathcal{M}})(\bar{x})$ is positive or negative definite, and in particular nonsingular.

3 Inertia stratification and extended conditions

This section develops the stratification framework used in the subsequent analysis. Starting from the inertia stratification of $\mathbb{S}^n$, we present a lifted stratification of the primal–dual space. The lifted strata separate the local matrix geometries encountered by F and provide the setting for its analysis on a fixed stratum.

We then formulate the regularity conditions at general primal–dual pairs and construct an auxiliary NLSDP for this purpose. This construction realizes each such pair as a KKT pair of an associated problem and thereby allows the extended conditions to be analyzed through the standard KKT framework.

3.1 The inertia stratification

For given integers $p, q \in \mathbb{Z}_+$ with $p + q \leq n$, we define the set

$$\mathcal{M}_{p,q} := \big\{A \in \mathbb{S}^n \mid |\alpha(A)| = p,\ |\gamma(A)| = q\big\}.$$

When the context is clear, we simply write $\mathcal{M}$ in place of $\mathcal{M}_{p,q}$. As shown in [31], each $\mathcal{M}_{p,q}$ is a smooth embedded submanifold of $\mathbb{S}^n$ with dimension $\dim(\mathcal{M}_{p,q}) = n(p+q) - \frac{1}{2}(p+q)(p+q-1)$. The family $\{\mathcal{M}_{p,q}\}_{0 \leq p+q \leq n}$ yields a finite disjoint decomposition of $\mathbb{S}^n$:

$$\mathbb{S}^n = \bigcup_{0 \leq p+q \leq n} \mathcal{M}_{p,q}.$$

This decomposition satisfies the frontier condition and local finiteness required by Definition 16. Hence, it forms a stratification of $\mathbb{S}^n$, which we call the *inertia stratification*. Each $\mathcal{M}_{p,q}$ is termed an *inertia stratum* of $\mathbb{S}^n$.

Figure 2 illustrates the stratification of $\mathbb{S}^2 \simeq \mathbb{R}^3$. This space decomposes into three 3-dimensional strata, two 2-dimensional strata, and the trivial zero-dimensional stratum $\{0\}$. The closure of $\mathcal{M}_{p,q}$ is the union of the strata $\mathcal{M}_{p',q'}$ satisfying $p' \leq p$ and $q' \leq q$.

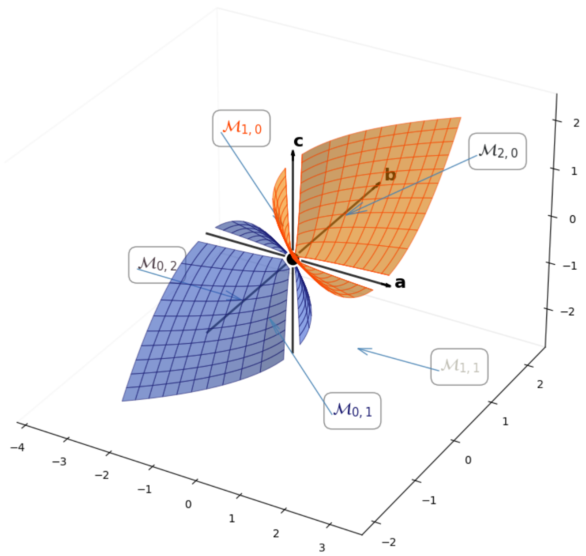

Fig. 2: The inertia stratification of $\mathbb{S}^2$.

Remark 3 The inertia stratification framework and our subsequent analyses extend seamlessly to NLSDPs with multiple SDP constraints $g_i(x) \in \mathbb{S}^{n_i}_+$ for $i = 1, \ldots, \ell$ (which inherently encompass polyhedral constraints of the form $h_E(x) = 0$ and $h_I(x) \geq 0$). Indeed, the product space $\mathbb{X} \times \mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_\ell}$ inherits a stratification whose strata take the form

$$\widetilde{\mathcal{M}}_{p_1,\ldots,p_\ell,q_1,\ldots,q_\ell} = \big\{(x, y_1, \ldots, y_\ell) \mid g_i(x) + y_i \in \mathcal{M}_{p_i,q_i},\ i = 1, \ldots, \ell\big\}.$$

Because this set is the preimage of the product manifold $\mathcal{M}_{p_1,q_1} \times \cdots \times \mathcal{M}_{p_\ell,q_\ell}$ under the smooth submersion $(x, y_1, \ldots, y_\ell) \mapsto (g_1(x) + y_1, \ldots, g_\ell(x) + y_\ell)$, it is a smooth embedded submanifold. Geometric objects such as tangent spaces and manifold derivatives can be computed componentwise. To streamline the presentation, we focus exclusively on (1).

Before examining further properties of the inertia stratification, we recall a convenient representation of the tangent space $T_A\mathcal{M}_{p,q}$ at a point $A \in \mathcal{M}_{p,q}$. It is well established (see, e.g., [31]) that

$$T_A\mathcal{M}_{p,q} = \left\{H = X^\top A + AX \in \mathbb{S}^n \mid X \in \mathbb{R}^{n\times n}\right\}. \tag{7}$$

The next proposition characterizes the tangent and normal spaces in terms of an IED as in (3).

Proposition 1 *Let $(\alpha, \beta, \gamma, p, q, P, \lambda)$ be an IED of $A \in \mathcal{M}_{p,q}$. Then*

$$T_A\mathcal{M}_{p,q} = \left\{H \in \mathbb{S}^n \mid (P^\top HP)_{\beta\beta} = 0\right\}. \tag{8}$$

Moreover,

$$N_A\mathcal{M}_{p,q} = \left\{H \in \mathbb{S}^n \mid (P^\top HP)_{ij} = 0 \text{ for all } (i,j) \notin \beta \times \beta\right\}. \tag{9}$$

Proof Let $\mathcal{R}$ denote the right-hand side of (8). For any given $H \in T_A\mathcal{M}_{p,q}$, it follows from (7) that $H = X^\top A + AX$ for some $X \in \mathbb{R}^{n\times n}$. Defining $\widetilde{H} := P^\top HP$, we obtain

$$\widetilde{H} = P^\top X^\top P \begin{bmatrix} \Lambda_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_{\gamma\gamma} \end{bmatrix} + \begin{bmatrix} \Lambda_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_{\gamma\gamma} \end{bmatrix} P^\top XP.$$

Since $\Lambda_{\beta\beta} = 0$, this yields $\widetilde{H}_{\beta\beta} = 0$, implying $T_A\mathcal{M}_{p,q} \subseteq \mathcal{R}$. Conversely, the dimension of $\mathcal{R}$ is clearly $\dim(\mathcal{R}) = n(p+q) - \frac{1}{2}(p+q)(p+q-1)$. Because $\dim(T_A\mathcal{M}_{p,q})$ precisely matches this value [31, Proposition 2.1], we conclude that $T_A\mathcal{M}_{p,q} = \mathcal{R}$. Formula (9) follows by taking the orthogonal complement of (8). ☐

Having established the stratified structure and characterized its tangent spaces, we analyze the differentiability of the metric projection $\Pi_{\mathbb{S}^n_+}$ along the strata. Specifically, when restricted to a fixed stratum $\mathcal{M}_{p,q}$, the mapping $\Pi_{\mathbb{S}^n_+} : \mathcal{M}_{p,q} \to \mathcal{M}_{p,0}$ is not only differentiable with an explicit closed-form differential, but is C^∞-smooth, strengthening previously known C^1-differentiability results (see Remark 5). These regularity properties fundamentally underpin the variational analysis in the sequel.

Theorem 1 *For any stratum $\mathcal{M} = \mathcal{M}_{p,q}$, the restriction of $\Pi_{\mathbb{S}^n_+}$ to $\mathcal{M}$, denoted by $\Pi_{\mathbb{S}^n_+}|_{\mathcal{M}} : \mathcal{M} \to \mathbb{S}^n$, is a C^∞-smooth map. Moreover, for an arbitrary $A \in \mathcal{M}$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$, the manifold differential $\xi_A := d(\Pi_{\mathbb{S}^n_+}|_{\mathcal{M}})_A : T_A\mathcal{M} \to T_{\Pi_{\mathbb{S}^n_+}(A)}\mathbb{S}^n$ at $H \in T_A\mathcal{M}$ is explicitly given by*

$$\xi_A(H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}^\top_{\alpha\beta} & 0 & 0 \\ \widetilde{H}^\top_{\alpha\gamma} \circ \Xi^\top_{\alpha\gamma} & 0 & 0 \end{bmatrix} P^\top, \tag{10}$$

where $\widetilde{H} := P^\top HP$.

Remark 4 It is straightforward to verify that $\Pi_{\mathbb{S}^n_+}$ maps $\mathcal{M}_{p,q}$ onto $\mathcal{M}_{p,0}$, which is smoothly embedded in $\mathbb{S}^n$. Because $\Pi_{\mathbb{S}^n_+} : \mathcal{M}_{p,q} \to \mathbb{S}^n$ is a C^∞ map, standard properties of smooth embeddings ensure that the corestriction $\Pi_{\mathbb{S}^n_+} : \mathcal{M}_{p,q} \to \mathcal{M}_{p,0}$ is also C^∞-smooth. Hereafter, we suppress the explicit restriction notation $|_{\mathcal{M}_{p,q}}$ and simply write $\Pi_{\mathbb{S}^n_+}$ depending on the context.

The following example shows that the ambient nonsmoothness of the projection arises from directions that change the inertia stratum.

Example 1 Let $A = \mathrm{Diag}(2, 0, -1) \in \mathcal{M}_{1,1} \subseteq \mathbb{S}^3$. With $P = I$, we have $\alpha = \{1\}$, $\beta = \{2\}$, and $\gamma = \{3\}$. Proposition 1 gives

$$T_A\mathcal{M}_{1,1} = \{H \in \mathbb{S}^3 \mid H_{22} = 0\}.$$

Since $\Xi_{13} = 2/(2-(-1)) = 2/3$, Theorem 1 specializes to

$$\xi_A(H) = \begin{bmatrix} H_{11} & H_{12} & \frac{2}{3}H_{13} \\ H_{12} & 0 & 0 \\ \frac{2}{3}H_{13} & 0 & 0 \end{bmatrix}, \qquad H \in T_A\mathcal{M}_{1,1}.$$

In contrast, the transverse direction E_{22} does not belong to $T_A\mathcal{M}_{1,1}$, and

$$\Pi_{\mathbb{S}^3_+}(A + tE_{22}) = \mathrm{Diag}(2, \max\{t, 0\}, 0),$$

which is not differentiable at $t = 0$.

Remark 5 The C^∞-smoothness established in Theorem 1, together with the closed-form differential (10), refines existing results for the metric projection $\Pi_{\mathbb{S}^n_+}$ along the strata. The C^1-differentiability of $\Pi_{\mathbb{S}^n_+}$ relative to $\mathcal{M}_{p,q}$ can be deduced via Löwner functional calculus or the theory of C^1-partly smooth mappings; see, e.g., [37, Example 4.14] and [23, Proposition 9.7]. To our knowledge, the full C^∞-smoothness on each inertia stratum together with the explicit manifold differential formula has not been recorded previously.

The proof of Theorem 1 relies on analytic functional calculus for symmetric matrices. We briefly recall the relevant facts. For background we refer to [17, Chapter VII] or [47, Sections 2.2–2.3].

In complex analysis, the classical Cauchy integral formula provides an integral representation of a scalar analytic function $f(\zeta)$ via the contour integral $f(\zeta) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-\zeta}\,dz$, where the contour Γ encloses ζ. Analogously, for a symmetric matrix $A \in \mathbb{S}^n$, analytic functional calculus defines $f(A)$ via

$$f(A) = \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}\,dz,$$

where the contour $\Gamma \subset \mathbb{C}$ strictly encloses the spectrum of A.

Lemma 1 ([17, Chapter VII]) *Let f be an analytic function on an open set $\Omega \subseteq \mathbb{C}$ containing the spectrum of $A \in \mathbb{S}^n$, and let the boundary $\Gamma = \partial\Omega$ enclose this spectrum. Then the matrix-valued function*

$$f(A) = \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}\,dz$$

depends analytically on A. In particular, the mapping $A \mapsto f(A)$ is C^∞-smooth and locally expressible as a convergent power series on the domain of symmetric matrices whose spectra lie within Ω.

Proof (Proof of Theorem 1) Fix $A \in \mathcal{M} \subseteq \mathbb{S}^n$. Since the inertia indices $|\alpha(A)|$, $|\beta(A)|$, and $|\gamma(A)|$ are structurally constant on the stratum $\mathcal{M}$, we select three mutually disjoint open disks $D_1, D_2, D_3 \subset \mathbb{C}$ enclosing the positive eigenvalues, the negative eigenvalues, and the origin, respectively. Set $D := D_1 \cup D_2 \cup D_3$. By the continuity of eigenvalues, there exists a sufficiently small neighborhood of A in $\mathcal{M}$ such that the positive, negative, and zero parts of the spectrum for any matrix in this neighborhood remain within D_1, D_2, and D_3, respectively. Define the following scalar functions on D:

$$f_+(z) := \begin{cases} z, & z \in D_1, \\ 0, & z \in D_2 \cup D_3, \end{cases} \qquad f_\dagger(z) := \begin{cases} 1/z, & z \in D_1 \cup D_2, \\ 0, & z \in D_3, \end{cases} \qquad p_0(z) := \begin{cases} 0, & z \in D_1 \cup D_2, \\ 1, & z \in D_3. \end{cases}$$

Since D_1, D_2, and D_3 are disjoint, these piecewise definitions are analytic on the open set D. By Lemma 1, the induced matrix functions $f_+(A)$, $f_\dagger(A)$, and $p_0(A)$ are well-defined and depend analytically on A within this local neighborhood.

For any $A \in \mathcal{M} \subseteq \mathbb{S}^n$, consider the local exponential chart $\exp_A : \mathcal{U} \to \mathcal{U}'$ with respect to the induced Riemannian metric on the embedded submanifold $\mathcal{M}$ in $\mathbb{S}^n$, which maps from a neighborhood $\mathcal{U} \subseteq T_A\mathcal{M}$ of the origin to a neighborhood $\mathcal{U}' \subseteq \mathcal{M}$ of A. Since the induced matrix-valued function $M \mapsto f_+(M)$ is analytic, and hence C^∞-smooth, on an open subset of $\mathbb{S}^n$ containing $\mathcal{U}'$, the composition $v \mapsto f_+(\exp_A(v))$ smoothly maps $\mathcal{U}$ into $\mathbb{S}^n$. For any matrix $\widetilde{A} \in \mathcal{U}' \cap \mathcal{M}$, we have $\Pi_{\mathbb{S}^n_+}(\widetilde{A}) = f_+(\widetilde{A})$, since both operations isolate the positive eigenspace of the matrix from the others. Consequently, the local coordinate representation $v \mapsto \Pi_{\mathbb{S}^n_+}(\exp_A(v))$ is smooth, which guarantees that the restriction $\Pi_{\mathbb{S}^n_+}|_{\mathcal{M}} : \mathcal{M} \to \mathbb{S}^n$ is C^∞-smooth.

Next, we derive the explicit manifold differential formula (10). For any $A \in \mathcal{M}$ and tangent vector $H \in T_A\mathcal{M}$, the directional derivative formula (4) gives

$$\Pi'_{\mathbb{S}^n_+}(A; H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\alpha\beta}^\top & \Pi_{\mathbb{S}^{|\beta|}_+}(0) & 0 \\ \widetilde{H}_{\alpha\gamma}^\top \circ \Xi_{\alpha\gamma}^\top & 0 & 0 \end{bmatrix} P^\top = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\alpha\beta}^\top & 0 & 0 \\ \widetilde{H}_{\alpha\gamma}^\top \circ \Xi_{\alpha\gamma}^\top & 0 & 0 \end{bmatrix} P^\top,$$

where $\widetilde{H} = P^\top H P$. The second equality uses $\widetilde{H}_{\beta\beta} = 0$, which holds for tangent vectors by Proposition 1, so that $\Pi_{\mathbb{S}^{|\beta|}_+}(0) = 0$. Because the restriction of $\Pi_{\mathbb{S}^n_+}$ to the smooth manifold $\mathcal{M}$ is smooth, its directional derivative coincides with its Fréchet differential on the manifold, yielding $d(\Pi_{\mathbb{S}^n_+}|_{\mathcal{M}})_A(H) = \Pi'_{\mathbb{S}^n_+}(A; H)$. $\square$

The nonsmoothness of the KKT mapping $F(z)$ in (2) stems exclusively from the projection term $\Pi_{\mathbb{S}^n_+}(G(z))$. To isolate and exploit the hidden smooth structure, we lift the inertia stratification of $\mathbb{S}^n$ to the primal–dual space $\mathbb{X} \times \mathbb{S}^n$ via the mapping G. For each pair (p, q) with $0 \le p+q \le n$, we define

$$\widetilde{\mathcal{M}}_{p,q} := G^{-1}(\mathcal{M}_{p,q}) = \{z = (x, y) \in \mathbb{X} \times \mathbb{S}^n \mid G(z) = g(x) + y \in \mathcal{M}_{p,q}\}. \tag{11}$$

Since G is a smooth submersion (see, e.g., [36, Chapter 4]), it follows from [36, Corollary 6.31] that each $\widetilde{\mathcal{M}}_{p,q}$ is a smoothly embedded submanifold of $\mathbb{X} \times \mathbb{S}^n$. Consequently, the finite disjoint union

$$\mathbb{X} \times \mathbb{S}^n = \bigcup_{0 \le p+q \le n} \widetilde{\mathcal{M}}_{p,q}$$

constitutes a stratification of $\mathbb{X} \times \mathbb{S}^n$, which we continue to call the *inertia stratification*. When the indices p and q are clear from context, we simply write $\widetilde{\mathcal{M}}$ for $\widetilde{\mathcal{M}}_{p,q}$.

By Theorem 1, while F is globally nonsmooth on $\mathbb{X} \times \mathbb{S}^n$, its restriction to any fixed stratum $\widetilde{\mathcal{M}}_{p,q}$ is completely C^∞-smooth. The following lemma characterizes the tangent space of $\widetilde{\mathcal{M}}_{p,q}$ and facilitates the computation of the manifold differential of F.

Lemma 2 *Fix (p, q) with $0 \le p+q \le n$, and let $\mathcal{M} := \mathcal{M}_{p,q}$ and $\widetilde{\mathcal{M}} := \widetilde{\mathcal{M}}_{p,q} = G^{-1}(\mathcal{M})$. For any $z = (x, y) \in \widetilde{\mathcal{M}}$, the tangent space $T_z\widetilde{\mathcal{M}}$ can be represented as*

$$T_z\widetilde{\mathcal{M}} = \{(v_x, v_y) \in \mathbb{X} \times \mathbb{S}^n \mid \nabla g(x)^* v_x + v_y \in T_{G(z)}\mathcal{M}\}. \tag{12}$$

In particular, define the linear isomorphism

$$\phi_z : \mathbb{X} \times \mathbb{S}^n \to \mathbb{X} \times \mathbb{S}^n, \qquad \phi_z(v_x, v_y) := (v_x, \nabla g(x)^* v_x + v_y), \tag{13}$$

with inverse

$$\phi_z^{-1}(v_x, H) := (v_x, H - \nabla g(x)^* v_x). \tag{14}$$

Then

$$\phi_z(T_z\widetilde{\mathcal{M}}) = \mathbb{X} \times T_{G(z)}\mathcal{M},$$

so its restriction is a linear isomorphism between these two spaces. We use the same notation for the ambient map and this restriction. Moreover, the assignments $z \mapsto \phi_z$ and $z \mapsto \phi_z^{-1}$ are continuously differentiable.

Proof Since G is a submersion, the preimage theorem [36, Corollary 6.31] gives $T_z\widetilde{\mathcal{M}} = (dG_z)^{-1}(T_{G(z)}\mathcal{M})$. Computing $dG_z(v_x, v_y) = \nabla g(x)^* v_x + v_y$ proves (12). The formulas in (13) and (14) define mutually inverse ambient linear maps. Equation (12) then gives $\phi_z(T_z\widetilde{\mathcal{M}}) = \mathbb{X} \times T_{G(z)}\mathcal{M}$. Their continuous differentiability in z follows from $g \in C^2$. □

Proposition 2 *For $z = (x, y) \in \widetilde{\mathcal{M}}$, in the coordinates $(v_x, H) \in \mathbb{X} \times T_{G(z)}\mathcal{M}$ induced by ϕ_z (13),*

$$(dF_z \circ \phi_z^{-1})(v_x, H) = \begin{bmatrix} \nabla^2_{xx} L(z) - \nabla g(x) \nabla g(x)^* & \nabla g(x) \\ -\nabla g(x)^* & \xi_{G(z)} \end{bmatrix} \begin{bmatrix} v_x \\ H \end{bmatrix}. \tag{15}$$

Proof By (14), $\phi_z^{-1}(v_x, H) = (v_x, H - \nabla g(x)^* v_x)$. Substituting this expression into dF_z gives

$$(dF_z \circ \phi_z^{-1})(v_x, H) = \begin{bmatrix} \nabla^2_{xx} L(z) v_x + \nabla g(x)(H - \nabla g(x)^* v_x) \\ -\nabla g(x)^* v_x + \xi_{G(z)}(H) \end{bmatrix} = \begin{bmatrix} \nabla^2_{xx} L(z) - \nabla g(x) \nabla g(x)^* & \nabla g(x) \\ -\nabla g(x)^* & \xi_{G(z)} \end{bmatrix} \begin{bmatrix} v_x \\ H \end{bmatrix},$$

□

The following lemma provides a smoothly varying orthonormal frame for the invariant subspace associated with the near-zero eigenvalues. It is used in the proofs of Lemma 4 and Lemma 8.

Lemma 3 *Let $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{p}, \bar{q}, \bar{P}, \bar{\lambda})$ be an IED of $\bar{A} \in \mathbb{S}^n$, set $k := |\bar{\beta}|$, and let $\delta > 0$ be smaller than the modulus of every nonzero eigenvalue of $\bar{A}$. Then there exist a neighborhood $\mathcal{V}$ of $\bar{A}$ in $\mathbb{S}^n$ and a C^∞ map $U_\delta : \mathcal{V} \to \mathbb{R}^{n \times k}$ with $U_\delta(\bar{A}) = \bar{P}_{\bar{\beta}}$ such that, for every $A \in \mathcal{V}$, the columns of $U_\delta(A)$ form an orthonormal basis of the invariant subspace of A associated with its eigenvalues in $(-\delta, \delta)$.*

Proof Let $D \subset \mathbb{C}$ be the open disk of radius δ centered at the origin, and let D' be a finite union of open disks with $\overline{D} \cap D' = \emptyset$ that contains all nonzero eigenvalues of $\bar{A}$. By the continuity of eigenvalues, there is a neighborhood $\mathcal{V}$ of $\bar{A}$ such that every $A \in \mathcal{V}$ has exactly $|\bar{\beta}|$ eigenvalues in D, counted with multiplicities, and all remaining eigenvalues in D'. Since the eigenvalues of A are real, the eigenvalues in D are exactly those in $(-\delta, \delta)$. Let p_D be the scalar function equal to 1 on D and to 0 on D'. Since p_D is analytic, by Lemma 1 the induced matrix function $p_D(A)$ depends analytically on $A \in \mathcal{V}$. Since A is symmetric, $p_D(A)$ is the orthogonal projector onto the invariant subspace of A associated with its eigenvalues in $(-\delta, \delta)$. Let $\bar{P}_{\bar{\beta}}$ collect the $\bar{\beta}$-columns of $\bar{P}$. Since the columns of $\bar{P}_{\bar{\beta}}$ form an orthonormal basis of $\ker \bar{A}$, we have $p_D(\bar{A})\bar{P}_{\bar{\beta}} = \bar{P}_{\bar{\beta}}$. Hence, after shrinking $\mathcal{V}$, the columns of $p_D(A)\bar{P}_{\bar{\beta}}$ are linearly independent and span this invariant subspace, and the Gram–Schmidt procedure yields the desired orthonormal frame $U_\delta(A)$, which depends analytically on A and satisfies $U_\delta(\bar{A}) = \bar{P}_{\bar{\beta}}$. □

The construction in this proof is that of [7, Exercise 3.140], where it is carried out at the smallest eigenvalues to reduce the cone $\mathbb{S}^n_+$.

The stratified perturbation analysis in Section 5.2.3 requires paths within a fixed stratum that remain arbitrarily close to a point of its closure. The following lemma provides precisely this local connectivity.

Lemma 4 *Let $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{p}, \bar{q}, \bar{P}, \bar{\lambda})$ be an IED of $\bar{A} \in \mathbb{S}^n$ and $k := |\bar{\beta}|$. Every neighborhood of $\bar{A}$ contains an open neighborhood $\mathcal{V}$ of $\bar{A}$ such that, for every $a = 0, \ldots, k$, the set*

$$\mathcal{M}_{\bar{p}+a,\bar{q}+k-a} \cap \mathcal{V}$$

is path-connected.

Proof When $k = 0$, the assertion is immediate. Hence suppose $k \geq 1$. Let $\bar{P}_{\bar{\beta}}$ be the matrix of $\bar{\beta}$-columns of $\bar{P}$. For A near $\bar{A}$, the spectrum splits into $\bar{p}$ eigenvalues near the $\bar{\lambda}_i$ with $i \in \bar{\alpha}$, $\bar{q}$ eigenvalues near the $\bar{\lambda}_i$ with $i \in \bar{\gamma}$, and k eigenvalues near 0. Apply Lemma 3 with $\delta > 0$ smaller than the modulus of every nonzero eigenvalue of $\bar{A}$, and let $U(A) := U_\delta(A)$. Then U is smooth, $U(\bar{A}) = \bar{P}_{\bar{\beta}}$, and the columns of $U(A)$ form an orthonormal basis of the invariant subspace of the k middle eigenvalues. The eigenvalues of $b(A) := U(A)^\top A U(A) \in \mathbb{S}^k$ are then the k middle eigenvalues of A, so for $\mathcal{V}$ small and every $A \in \mathcal{V}$,

$$A \in \mathcal{M}_{\bar{p}+a,\bar{q}+k-a} \iff b(A) \in \Sigma_{a,k-a} := \big\{B \in \mathbb{S}^k : \ B \text{ has inertia } (a, k-a)\big\}.$$

Moreover, $Db(\bar{A})H = \bar{P}_{\bar{\beta}}^\top H \bar{P}_{\bar{\beta}}$, where the terms involving $DU(\bar{A})$ vanish because $\bar{A}U(\bar{A}) = \bar{A}\bar{P}_{\bar{\beta}} = 0$. This derivative is surjective: for every $M \in \mathbb{S}^k$, taking $H = \bar{P}_{\bar{\beta}} M \bar{P}_{\bar{\beta}}^\top$ gives $Db(\bar{A})H = M$. Let $\mathcal{K} := \ker Db(\bar{A})$, let $\Pi_{\mathcal{K}}$ denote the orthogonal projection onto $\mathcal{K}$, and define

$$\Psi(A) := \big(b(A), \Pi_{\mathcal{K}}(A - \bar{A})\big).$$

Then $D\Psi(\bar{A})H = \big(Db(\bar{A})H, \Pi_{\mathcal{K}}H\big)$ is an isomorphism from $\mathbb{S}^n$ onto $\mathbb{S}^k \times \mathcal{K}$. Indeed, if its value at H is zero, then $H \in \mathcal{K}$ and $\Pi_{\mathcal{K}}H = H = 0$; bijectivity follows because the domain and codomain have the same dimension. The inverse function theorem therefore gives, after shrinking the neighborhood, a diffeomorphism $\Psi : \mathcal{V} \to V_0 \times W_0$, where $V_0 \subseteq \mathbb{S}^k$ is a ball centered at $0 = b(\bar{A})$ and W_0 is a ball in $\mathcal{K}$. The first component of Ψ is b, and hence

$$\mathcal{M}_{\bar{p}+a,\bar{q}+k-a} \cap \mathcal{V} = \Psi^{-1}\big((\Sigma_{a,k-a} \cap V_0) \times W_0\big).$$

By [31, Proposition 2.1], $\Sigma_{a,k-a}$ is a connected component of the manifold of nonsingular symmetric $k \times k$ matrices and hence is path-connected. Since $\Sigma_{a,k-a}$ is invariant under positive scaling and V_0 is a ball centered at 0, the set $\Sigma_{a,k-a} \cap V_0$ is path-connected. Therefore $\Psi^{-1}\big((\Sigma_{a,k-a} \cap V_0) \times W_0\big)$ is path-connected. □

3.2 Extended regularity conditions

We now define the extended conditions used at arbitrary primal–dual pairs.

We first extend the SRCQ, constraint nondegeneracy, and the W-SRCQ (Definitions 3, 4, and 9) to arbitrary primal–dual pairs.

Definition 18 For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *strict Robinson constraint qualification (SRCQ)* holds at z if

$$g'(x)\mathbb{X} + \left\{PBP^\top \in \mathbb{S}^n \,\middle|\, B_{\beta\beta} \succeq 0, \ B_{\beta\gamma} = 0, \ B_{\gamma\gamma} = 0\right\} = \mathbb{S}^n.$$

Definition 19 For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *constraint nondegeneracy* holds at z if

$$g'(x)\mathbb{X} + \left\{PBP^\top \mid B \in \mathbb{S}^n,\ B_{\beta\beta} = 0,\ B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0\right\} = \mathbb{S}^n. \tag{16}$$

Definition 20 For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *weak strict Robinson constraint qualification (W-SRCQ)* holds at z if

$$g'(x)\mathbb{X} + \left\{PBP^\top \mid B \in \mathbb{S}^n,\ B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0\right\} = \mathbb{S}^n. \tag{17}$$

At a KKT pair $z = (x, y)$, $g(x) = \Pi_{\mathbb{S}^n_+}(G(z))$ and $y = \Pi_{\mathbb{S}^n_-}(G(z))$, so Definitions 18, 19, and 20 reduce to Definitions 3, 4, and 9, respectively.

For $z = (x, y)$, define

$$X_z := \Pi_{\mathbb{S}^n_+}(G(z)), \qquad Y_z := \Pi_{\mathbb{S}^n_-}(G(z)). \tag{18}$$

For an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, define the quadratic form

$$\mathcal{Q}_z(v_x) := \langle v_x, \nabla^2_{xx} L(x, y) v_x\rangle + 2 \sum_{i\in\alpha}\sum_{j\in\gamma} \frac{-\lambda_j}{\lambda_i} [P^\top(\nabla g(x)^* v_x)P]^2_{ij}. \tag{19}$$

The extended critical cone at z is

$$\mathcal{C}(z) := \left\{v_x \in \mathbb{X} \mid g'(x)v_x \in T_{\mathbb{S}^n_+}(X_z) \cap Y_z^\perp\right\},$$

where $Y_z^\perp := \{H \in \mathbb{S}^n \mid \langle Y_z, H\rangle = 0\}$. Following [48, (38)–(40)] and [25, Definition 8 and (20)], we use $\operatorname{app}(z)$ and $\operatorname{appl}(z)$ as computable surrogates for $\operatorname{aff}(\mathcal{C}(z))$ and $\operatorname{lin}(\mathcal{C}(z))$, respectively:

$$\operatorname{app}(z) := \left\{v_x \in \mathbb{X} \mid g'(x)v_x \in \operatorname{aff}(T_{\mathbb{S}^n_+}(X_z) \cap Y_z^\perp)\right\},$$
$$\operatorname{appl}(z) := \left\{v_x \in \mathbb{X} \mid g'(x)v_x \in \operatorname{lin}(T_{\mathbb{S}^n_+}(X_z) \cap Y_z^\perp)\right\}.$$

In the above IED, these three sets are given by

$$\mathcal{C}(z) := \left\{v_x \in \mathbb{X} \mid [P^\top(\nabla g(x)^* v_x)P]_{\beta\beta} \succeq 0,\ [P^\top(\nabla g(x)^* v_x)P]_{\beta\gamma} = 0,\ [P^\top(\nabla g(x)^* v_x)P]_{\gamma\gamma} = 0\right\}, \tag{20}$$

$$\operatorname{app}(z) := \left\{v_x \in \mathbb{X} \mid [P^\top(\nabla g(x)^* v_x)P]_{\beta\gamma} = 0,\ [P^\top(\nabla g(x)^* v_x)P]_{\gamma\gamma} = 0\right\}, \tag{21}$$

$$\operatorname{appl}(z) := \left\{v_x \in \mathbb{X} \mid [P^\top(\nabla g(x)^* v_x)P]_{\beta\beta} = 0,\ [P^\top(\nabla g(x)^* v_x)P]_{\beta\gamma} = 0,\ [P^\top(\nabla g(x)^* v_x)P]_{\gamma\gamma} = 0\right\}. \tag{22}$$

Lemma 5 *If the SRCQ holds at z, then*

$$\operatorname{app}(z) = \operatorname{aff}\bigl(\mathcal{C}(z)\bigr) = \operatorname{span}\mathcal{C}(z).$$

Proof Set $\mathcal{T}_z := T_{\mathbb{S}^n_+}(X_z) \cap Y_z^\perp$. Then $\mathcal{T}_z$ is a convex cone. Definition 18 and (20) give $g'(x)\mathbb{X} + \mathcal{T}_z = \mathbb{S}^n$ and $\mathcal{C}(z) = \{d \in \mathbb{X} \mid g'(x)d \in \mathcal{T}_z\}$. Moreover, since $\mathcal{T}_z$ is a cone,

$$\operatorname{app}(z) = \{d \in \mathbb{X} \mid g'(x)d \in \operatorname{span}\mathcal{T}_z\}.$$

The inclusion $\operatorname{span}\mathcal{C}(z) \subseteq \operatorname{app}(z)$ follows from $\mathcal{C}(z) \subseteq \operatorname{app}(z)$.

Conversely, take any $u \in \operatorname{app}(z)$. Since $g'(x)u \in \operatorname{span}\mathcal{T}_z = \mathcal{T}_z - \mathcal{T}_z$, there exist $H_1, H_2 \in \mathcal{T}_z$ such that $g'(x)u = H_1 - H_2$. By the SRCQ, there exist $w \in \mathbb{X}$ and $H_3 \in \mathcal{T}_z$ such that $-H_2 = g'(x)w + H_3$. Hence

$$g'(x)(u - w) = H_1 + H_3 \in \mathcal{T}_z, \qquad g'(x)(-w) = H_2 + H_3 \in \mathcal{T}_z.$$

Thus $u - w \in \mathcal{C}(z)$ and $-w \in \mathcal{C}(z)$, so $u = (u - w) - (-w) \in \operatorname{span}\mathcal{C}(z)$. This proves $\operatorname{app}(z) \subseteq \operatorname{span}\mathcal{C}(z)$. Finally, because $\mathcal{C}(z)$ is a cone, $\operatorname{aff}(\mathcal{C}(z)) = \operatorname{span}\mathcal{C}(z)$. $\square$

We next extend the SOSC and S-SOSC in Definitions 6 and 7 to arbitrary primal–dual pairs.

Definition 21 For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *second order sufficient condition (SOSC)* holds at z if

$$\mathcal{Q}_z(v_x) > 0 \quad \forall\, v_x \in \mathcal{C}(z) \setminus \{0\}.$$

Definition 22 For $z = (x, y) \in \mathbb{X}\times\mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *strong second order sufficient condition (S-SOSC)* holds at z if

$$\mathcal{Q}_z(v_x) > 0 \quad \forall\, v_x \in \operatorname{app}(z) \setminus \{0\}.$$

Definition 23 For $z = (x, y) \in \mathbb{X}\times\mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *strong second order condition (S-SOC)* holds at z if

$$\mathcal{Q}_z(v_x) \neq 0 \quad \forall\, v_x \in \mathrm{app}(z) \setminus \{0\},$$

where $\mathrm{app}(z)$ is given by (21).

Remark 6 The S-SOSC implies the S-SOC. Unlike the S-SOSC, the S-SOC also allows $\mathcal{Q}_z$ to be negative definite on $\mathrm{app}(z)$.

For arbitrary primal–dual pairs, we use the following non-vanishing version of the W-SOC in Definition 10.

Definition 24 For $z = (x, y) \in \mathbb{X}\times\mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *weak second order condition (W-SOC)* holds at z if

$$\mathcal{Q}_z(v_x) \neq 0 \quad \forall\, v_x \in \mathrm{appl}(z) \setminus \{0\},$$

where $\mathrm{appl}(z)$ is defined in (22).

Remark 7 The W-SOC in [25] is formulated with strict positivity because it is considered at a KKT pair associated with a local minimizer. Here we work at arbitrary primal–dual pairs and use non-vanishing, which allows $\mathcal{Q}_z$ to be either positive or negative definite on $\mathrm{appl}(z)$. At a KKT pair satisfying the W-SONC, the negative case is excluded, and our definition reduces to that in [25].

The extended conditions satisfy

$$\text{constraint nondegeneracy} \quad\Longrightarrow\quad \text{SRCQ} \quad\Longrightarrow\quad \text{W-SRCQ}$$

and

$$\text{S-SOSC} \quad\Longrightarrow\quad \text{SOSC} \quad\Longrightarrow\quad \text{W-SOC}, \qquad \text{S-SOSC} \quad\Longrightarrow\quad \text{S-SOC} \quad\Longrightarrow\quad \text{W-SOC}.$$

We also extend the SONC in Definition 8 by requiring $\mathcal{Q}_z$ to be nonnegative on $\mathcal{C}(z)$.

Definition 25 For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$, the *second order necessary condition (SONC)* holds at z if

$$\mathcal{Q}_z(v_x) \geq 0 \quad \forall\, v_x \in \mathcal{C}(z).$$

The weak second order necessary condition in [27, Definition 1] also admits a primal–dual formulation: compared with the SONC above, it imposes the same nonnegativity on the weak-direction space $\mathrm{appl}(z)$.

Definition 26 Let $z = (x, y) \in \mathbb{X}\times\mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$. The *weak second order necessary condition (W-SONC)* holds at z if

$$\mathcal{Q}_z(v_x) \geq 0 \quad \forall\, v_x \in \mathrm{appl}(z),$$

where $\mathrm{appl}(z)$ is defined in (22).

Remark 8 In [27], the "W-SOC" refers to the *weak second order necessary* condition. In the present paper, however, we reserve "W-SOC" for the condition in Definition 24, which is introduced for our regularity analysis and should not be confused with a second order sufficient condition. To avoid ambiguity, we therefore refer to the weak second order necessary condition in Definition 26 as *W-SONC*. The two definitions also use different direction spaces. The condition in [27, Definition 1] is imposed on $\mathrm{lin}(\mathcal{C}(x))$, whereas Definition 26 uses the outer approximation $\mathrm{appl}(z) \supseteq \mathrm{lin}(\mathcal{C}(z))$. We refer to [27] for a detailed discussion of the W-SONC at KKT pairs, which is not central to our development. We only emphasize one key point: even in nonlinear programming, barrier methods [29] and augmented Lagrangian methods [3] may produce limit points that violate the classical second order necessary condition, whereas such limit points are still guaranteed to satisfy the W-SONC (see [2]).

At a KKT pair $z = (x, y)$, we have $X_z = g(x)$ and $Y_z = y$, so $\mathcal{C}(z)$, $\mathrm{app}(z)$, and $\mathrm{appl}(z)$ coincide with the corresponding sets in Section 2. Moreover, [48, (28)] gives the corresponding reduction of the second order expression,

$$\mathcal{Q}_z(v_x) = \langle v_x, \nabla^2_{xx} L(x, y) v_x\rangle - \sigma\left(y, T^2_{\mathbb{S}^n_+}(g(x), g'(x)v_x)\right). \tag{23}$$

Consequently, the extended SONC, SOSC, and S-SOSC agree at KKT pairs with Definitions 8, 6, and 7, respectively. Under the W-SONC, the extended W-SOC also agrees with Definition 10.

Remark 9 Suppose that problem (1) is convex, i.e., f is convex and g is matrix-concave, meaning that

$$g(tx_1 + (1-t)x_2) \succeq tg(x_1) + (1-t)g(x_2) \quad \forall\, x_1, x_2 \in \mathbb{X},\ t \in [0,1].$$

Then, for any $x \in \mathbb{X}$ and any $y \in \mathbb{S}^n_-$, we have $\nabla^2_{xx} L(x,y) \succeq 0$. Consequently, the SONC and the W-SONC hold automatically at every $z = (x,y) \in \mathbb{X} \times \mathbb{S}^n_-$.

Remark 10 It is worth emphasizing that the extended regularity conditions in Definitions 18–26 are defined at an arbitrary primal–dual pair z via an IED of $G(z)$. At non-KKT pairs, they should not be viewed as standalone optimality conditions. In particular, by themselves they do not imply primal feasibility, complementarity, stationarity, or the identity $F(z) = 0$. This level of generality is essential for the analysis in Sections 4–5. Indeed, the stratified variational analysis is carried out on the lifted strata, and therefore necessarily involves arbitrary primal–dual pairs, not only KKT pairs. The following elementary example shows that the S-SOSC from Definition 22 and the constraint nondegeneracy from Definition 19 may hold simultaneously at a non-KKT pair. Consequently, the weaker conditions SONC, SOSC, W-SOC, W-SONC, SRCQ, and W-SRCQ may also hold at a non-KKT pair. Consider the scalar problem

$$\min_{x \in \mathbb{R}} \ \frac{1}{2}(x-2)^2 \quad \text{s.t.} \quad g(x) := x \in \mathbb{S}^1_+.$$

For any primal–dual pair $z = (x,y) \in \mathbb{R} \times \mathbb{S}^1$, since $g'(x) \equiv 1$ is surjective, constraint nondegeneracy holds automatically at z. Moreover, $Q_z(v_x) = \langle v_x, \nabla^2_{xx} L(x,y) v_x \rangle = v_x^2 > 0$ for every $v_x \neq 0$, so the S-SOSC also holds at z. Hence, both the S-SOSC and constraint nondegeneracy are satisfied at every primal–dual pair z, including non-KKT pairs. In conclusion, these extended regularity conditions alone are not intended to characterize optimality or stationarity. Rather, they provide the appropriate geometric and variational framework for the stratified analysis developed in the sequel.

From this point onward, unless stated otherwise, SRCQ, constraint nondegeneracy, W-SRCQ, SONC, SOSC, S-SOSC, S-SOC, W-SOC, and W-SONC refer to their extended definitions above.

3.3 The auxiliary problem

The extended regularity conditions in Section 3.2 are defined at arbitrary primal–dual pairs, although such a pair need not satisfy the KKT system. The construction below assigns to every $z = (x,y)$ an auxiliary NLSDP for which (x, Y_z) is a KKT pair. It preserves the IED, the constraint derivative $g'(x)$, and the Lagrangian Hessian, which are the data entering the extended conditions. Section 4 will use this KKT realization to construct two smooth manifold models.

Recall from (18) that $X_z + Y_z = G(z)$ and $\langle X_z, Y_z \rangle = 0$. Since $G(z) = g(x) + y$, we also have $X_z - g(x) = y - Y_z$. Define

$$g_z(u) := g(u) - g(x) + X_z = g(u) + y - Y_z \tag{24}$$

and

$$f_z(u) := f(u) + \big\langle y - Y_z,\ g(u) - g(x) \big\rangle - \big\langle \nabla_x L(x,y),\ u - x \big\rangle.$$

The *auxiliary problem* associated with z is

$$(\mathcal{P}_z) \qquad \min_{u \in \mathbb{X}} \ f_z(u) \quad \text{s.t.} \quad g_z(u) \in \mathbb{S}^n_+. \tag{25}$$

The following theorem shows that (x, Y_z) is a KKT pair of (25) and that the extended regularity conditions at z for (1) coincide with the corresponding conditions at (x, Y_z) for (25).

Remark 11 The perturbations defining $\mathcal{P}_z$ differ from two common forms of canonical perturbations for NLSDP. One adds a linear perturbation to the objective and a constant perturbation to the constraint, while the other allows both to be perturbed by functions satisfying suitable continuity requirements. Both the former and our perturbations are special cases of the latter. The special design of our perturbations allows the extended regularity conditions at an arbitrary primal–dual pair to be realized as standard KKT regularity conditions of an auxiliary NLSDP, and thereby provides a basis for their geometric interpretation through manifold optimization.

Theorem 2 *Let $z = (x,y) \in \mathbb{X} \times \mathbb{S}^n$, and let $(\alpha, \beta, \gamma, p, q, P, \lambda)$ be an IED of $G(z)$. The following statements hold for the auxiliary problem* (25).

(a) *$(x,\, Y_z)$ is a KKT pair of* (25).

(b) *The same tuple* $(\alpha, \beta, \gamma, p, q, P, \lambda)$ *is an IED of* $g_z(x) + Y_z = G(z)$*, and the Lagrangian* $L_z(u, v) := f_z(u) + \langle v,\, g_z(u)\rangle$ *satisfies* $\nabla^2_{uu} L_z(x, Y_z) = \nabla^2_{xx} L(x, y)$*.*

(c) *The conditions SRCQ, constraint nondegeneracy, and W-SRCQ (Definitions 18, 19, and 20) at* z *for* (1) *coincide with the corresponding conditions at* (x, Y_z) *for* (25)*.*

(d) *The critical cone* $\mathcal{C}(z)$ *and the sets* $\operatorname{app}(z)$ *and* $\operatorname{appl}(z)$ *coincide with the corresponding objects of* (25) *at* (x, Y_z)*. Consequently, the conditions SONC, SOSC, S-SOSC, S-SOC, W-SOC, and W-SONC (Definitions 25, 21, 22, 23, 24, and 26) at* z *for* (1) *coincide with the corresponding conditions at* (x, Y_z) *for* (25)*.*

Proof (a) Since $g_z(x) + Y_z = X_z + Y_z = G(z)$,

$$\begin{aligned}\nabla f_z(x) + \nabla g_z(x)\, Y_z &= \big[\nabla f(x) + \nabla g(x)(y - Y_z) - \nabla_x L(x, y)\big] + \nabla g(x)\, Y_z \\ &= \nabla_x L(x, y) - \nabla_x L(x, y) = 0,\end{aligned}$$

where we used $\nabla_x L(x, y) = \nabla f(x) + \nabla g(x) y$ and $\nabla g_z(x) = \nabla g(x)$. Moreover, $g_z(x) = X_z = \Pi_{\mathbb{S}^n_+}(g_z(x) + Y_z)$. Hence (x, Y_z) is a KKT pair of (25).

(b) The identity $g_z(x) + Y_z = G(z)$ is immediate from (24) and (18). For the Lagrangian Hessian, collecting the second-order terms of $f_z(u) + \langle Y_z, g_z(u)\rangle$ at $u = x$:

$$\begin{aligned}\nabla^2_{uu} L_z(x, Y_z) &= \nabla^2 f(x) + \nabla^2_{uu}\langle y - Y_z,\, g(u)\rangle\big|_x + \nabla^2_{uu}\langle Y_z,\, g(u)\rangle\big|_x \\ &= \nabla^2_{xx} L(x, y),\end{aligned}$$

where the term $-\langle \nabla_x L(x, y),\, u - x\rangle$ is linear and contributes zero.

(c) and (d) follow directly from (b). Every condition in Definitions 18–26 depends only on $\nabla g(x)$, the IED of $G(z)$, and $\nabla^2_{xx} L(x, y)$. Since these three objects are the same for both problems at their respective evaluation points, all listed conditions coincide. $\square$

The auxiliary problem reduces to the original when z is already a KKT pair.

Proposition 3 *Let* $z = (x, y)$ *be a KKT pair of* (1)*. Then* $X_z = g(x)$*,* $Y_z = y$*,* $g_z = g$*, and* $f_z = f$*. Consequently, the auxiliary problem* (25) *reduces to* (1)*.*

Proof Since $F(z) = 0$, the second component of (2) gives $g(x) = \Pi_{\mathbb{S}^n_+}(G(z)) = X_z$, and hence $Y_z = G(z) - X_z = y$. It follows that $g_z(u) = g(u)$. Stationarity gives $\nabla_x L(x, y) = 0$, so $f_z(u) = f(u)$. $\square$

The role of (25) is therefore to realize the data attached to an arbitrary primal–dual pair as KKT data without changing the regularity and second-order conditions considered here. Section 4 constructs two smooth manifold models of this auxiliary problem. By Proposition 3, these become models of the original problem whenever z is already a KKT pair.

Figure 3 summarizes these constructions and equivalences.

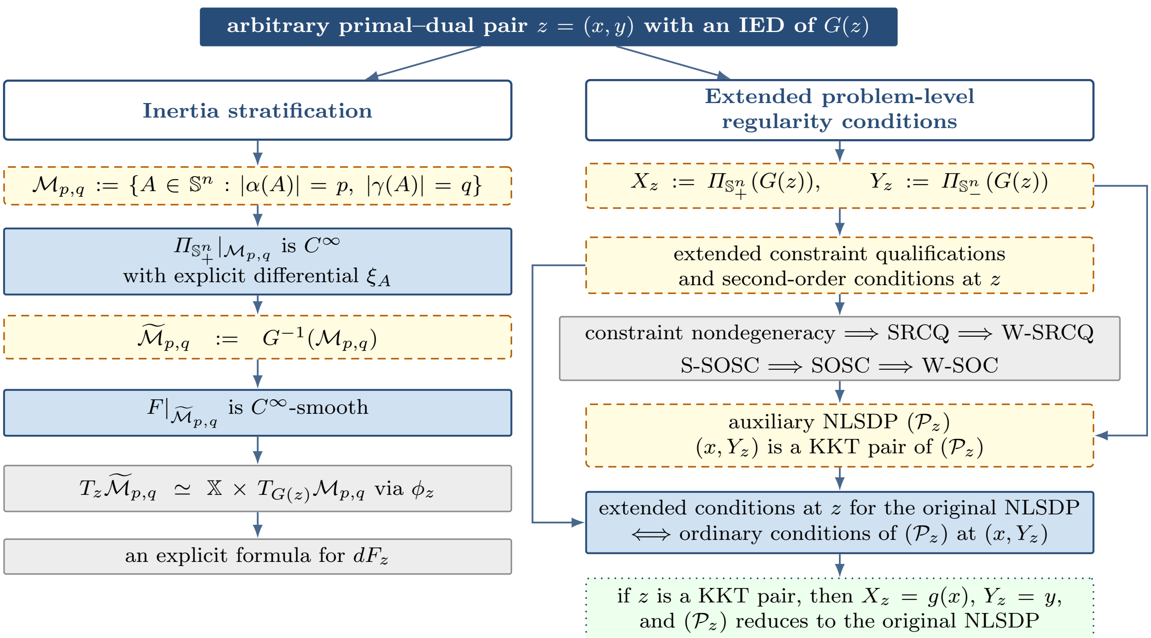

Fig. 3: Inertia stratification, extended conditions, and the auxiliary problem.

4 Geometric interpretations of regularity conditions from a stratified perspective

This section constructs two nested local submanifolds of $\mathbb{S}^n$ from the inertia stratification and combines them with the auxiliary NLSDP to form two smooth manifold optimization problems. On these models, the extended constraint qualifications become transversality conditions and the extended second-order conditions become manifold second-order conditions. These interpretations supply the geometric structure for the perturbation analysis of Section 5.

4.1 Construction of $\mathcal{A}_z$ and $\mathcal{B}_z$

Let $(\alpha, \beta, \gamma, p, q, P, \lambda)$ be an IED of $G(z)$. If $G(z) = 0$, choose any $\delta(z) > 0$. Otherwise, define the minimal non-vanishing eigenvalue modulus of $G(z)$ by

$$\delta(z) := \min\bigl\{|\lambda_i(G(z))| \mid \lambda_i(G(z)) \neq 0\bigr\}.$$

Set $\mathcal{B}(A, r) := \{A' \in \mathbb{S}^n \mid \|A - A'\|_2 < r\}$. Note that $X_z = \Pi_{\mathbb{S}^n_+}(G(z)) \in \mathcal{M}_{p,0}$, so the ball $\mathcal{B}(X_z, \delta(z)/2)$ is centered at the positive semidefinite part of $G(z)$.

Lemma 6 *For $z \in \widetilde{\mathcal{M}}_{p,q}$, the set*

$$\bigl(\mathcal{M}_{p,0} \cap \mathcal{B}(X_z, \delta(z)/2)\bigr) + Y_z$$

is smoothly embedded in $\mathcal{M}_{p,q}$.

Proof Let $A'_+ \in \mathcal{M}_{p,0} \cap \mathcal{B}(X_z, \delta(z)/2)$, so $\|A'_+ - X_z\|_2 < \delta(z)/2$. Then $\|(A'_+ + Y_z) - G(z)\|_2 = \|A'_+ - X_z\|_2 < \delta(z)/2$. By [4, Corollary III.2.6], corresponding eigenvalues of $A'_+ + Y_z$ and $G(z)$ differ by less than $\delta(z)/2$. Since $G(z) \in \mathcal{M}_{p,q}$ has p eigenvalues $\geq \delta(z)$ and q eigenvalues $\leq -\delta(z)$, it follows that $A'_+ + Y_z$ has at least p positive eigenvalues and at least q negative eigenvalues.

For the upper bounds on the inertia indices: the null space of $A'_+ \succeq 0$ has dimension $n - p$, and for every v in it, $\langle v, (A'_+ + Y_z)v\rangle = \langle v, Y_z v\rangle \leq 0$. By [4, Theorem III.1.2], the matrix $A'_+ + Y_z$ has at most p positive eigenvalues. Symmetrically, the null space of $Y_z \preceq 0$ has dimension $n - q$, and $\langle v, (A'_+ + Y_z)v\rangle = \langle v, A'_+ v\rangle \geq 0$ for all v in it, so $A'_+ + Y_z$ has at most q negative eigenvalues.

Thus $A'_+ + Y_z$ has exactly p positive eigenvalues and exactly q negative eigenvalues, so $A'_+ + Y_z \in \mathcal{M}_{p,q}$. The smooth embedding follows from the fact that translation by the fixed matrix Y_z is a diffeomorphism of $\mathbb{S}^n$. $\square$

The following two manifolds provide the smooth models underlying our geometric interpretation of the extended constraint qualifications and second-order conditions.

Definition 27 For $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$, define

$$\mathcal{A}_z := \Bigl(\mathcal{M}_{p,0} \cap \mathcal{B}\Bigl(X_z, \tfrac{\delta(z)}{2}\Bigr)\Bigr) + Y_z - y. \tag{26}$$

Define the normal bundle of $\mathcal{M}_{p,q} - y$ restricted to $\mathcal{A}_z$ by

$$\mathcal{E}_z := \mathcal{N}(\mathcal{M}_{p,q} - y)|_{\mathcal{A}_z} = \bigl\{(A, N) \mid A \in \mathcal{A}_z,\ N \in N_{A+y}\mathcal{M}_{p,q}\bigr\}. \tag{27}$$

Define the normal-addition map on $\mathcal{E}_z$ by

$$E : \mathcal{E}_z \to \mathbb{S}^n, \qquad E(A, N) := A + N.$$

Set

$$\mathcal{W}_z := \bigl\{(A, N) \in \mathcal{E}_z \mid \|N\|_2 < \tfrac{\delta(z)}{4}\bigr\}. \tag{28}$$

Finally, define

$$\mathcal{B}_z := E(\mathcal{W}_z). \tag{29}$$

When $G(z) = 0$, we have $p = q = 0$ and $X_z = Y_z = 0$. Consequently,

$$\mathcal{A}_z = \{g(x)\}, \qquad \mathcal{B}_z = \bigl\{g(x) + N \mid \|N\|_2 < \delta(z)/4\bigr\}.$$

Thus, in this case, $\mathcal{A}_z$ is a singleton and $\mathcal{B}_z$ is an open neighborhood of $g(x)$; the choice of $\delta(z) > 0$ only affects the size of this neighborhood.

Lemma 6 and (26) give $\mathcal{A}_z \subset \mathcal{M}_{p,q} - y$, so $\mathcal{E}_z$ is well defined. Moreover, (26) gives $g(x) \in \mathcal{A}_z$, while $(A, 0) \in \mathcal{W}_z$ and $E(A, 0) = A$ for every $A \in \mathcal{A}_z$. Hence $g(x) \in \mathcal{A}_z \subset \mathcal{B}_z$.

Proposition 4 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$ with IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$. The following hold.*

(a) *$\mathcal{A}_z$ is a smooth embedded manifold near $g(x)$ with tangent space*

$$T_{g(x)}\mathcal{A}_z = T_{X_z}\mathcal{M}_{p,0} = \{PBP^\top \mid B \in \mathbb{S}^n,\ B_{\beta\beta} = 0,\ B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0\}. \tag{30}$$

(b) *$\mathcal{B}_z$ is a smooth embedded manifold near $g(x)$, the map $E|_{\mathcal{W}_z} : \mathcal{W}_z \to \mathcal{B}_z$ is a diffeomorphism, and its tangent space at $g(x)$ is*

$$T_{g(x)}\mathcal{B}_z = T_{g(x)}\mathcal{A}_z + N_{g(x)}(\mathcal{M}_{p,q} - y) = \{PBP^\top \mid B \in \mathbb{S}^n,\ B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0\}. \tag{31}$$

Proof Part (a): $\mathcal{A}_z$ is a smooth embedded submanifold of $\mathcal{M}_{p,q} - y$ by Lemma 6 and (26). Since translation by $Y_z - y$ is an isometry, the tangent space at $g(x)$ equals $T_{X_z}\mathcal{M}_{p,0}$. Applying Proposition 1 to $\mathcal{M}_{p,0}$ at X_z (where the zero block spans $\mathbb{R}^{|\beta|+|\gamma|}$) gives (30).

Part (b): Since $\mathcal{A}_z$ is a smooth embedded submanifold of $\mathcal{M}_{p,q} - y$ by Part (a), $\mathcal{E}_z$ is a smooth submanifold of the normal bundle $\mathcal{N}(\mathcal{M}_{p,q} - y)$.

We prove that $E|_{\mathcal{W}_z}$ is a smooth embedding. Fix $(A, N) \in \mathcal{W}_z$. By (26), $A+y \in \mathcal{M}_{p,q}$ and $\|(A+y)-G(z)\|_2 < \delta(z)/2$. Hence, by [4, Corollary III.2.6], the nonzero eigenvalues of $A + y$ have modulus greater than $\delta(z)/2$. Moreover, $N \in N_{A+y}\mathcal{M}_{p,q}$, so (9) gives $(A + y)N = N(A + y) = 0$. Since $\|N\|_2 < \delta(z)/4$, the small spectrum of N is separated from the nonzero spectrum of $A + y$ in $E(A, N) + y$.

Define a real-analytic function χ_z on the open set

$$\left(-\tfrac{\delta(z)}{3}, \tfrac{\delta(z)}{3}\right) \cup \left((-\infty, -\tfrac{5\delta(z)}{12}) \cup (\tfrac{5\delta(z)}{12}, \infty)\right)$$

by setting $\chi_z(t) = 0$ on the central interval and $\chi_z(t) = t$ on the two outer intervals. By the spectral separation above, the spectrum of $E(A, N) + y$ lies in this set. The spectral calculus for real symmetric matrices gives

$$A + y = \chi_z(E(A, N) + y), \qquad A = \chi_z(E(A, N) + y) - y, \qquad N = E(A, N) + y - \chi_z(E(A, N) + y).$$

Thus A and N are recovered uniquely from the single matrix $E(A, N)$. Hence $E|_{\mathcal{W}_z}$ is injective and has an inverse obtained by restricting a smooth ambient map. It is therefore a smooth embedding, and $\mathcal{B}_z = E(\mathcal{W}_z)$ is a smooth embedded submanifold of $\mathbb{S}^n$.

Since $\mathcal{W}_z$ is open in $\mathcal{E}_z$ and $E|_{\mathcal{W}_z}$ is a diffeomorphism,

$$T_{g(x)}\mathcal{B}_z = dE_{(g(x),0)}T_{(g(x),0)}\mathcal{E}_z.$$

At the zero section, $T_{(g(x),0)}\mathcal{E}_z = T_{g(x)}\mathcal{A}_z \oplus N_{g(x)}(\mathcal{M}_{p,q} - y)$, and $dE_{(g(x),0)}(H, K) = H + K$. Hence

$$T_{g(x)}\mathcal{B}_z = T_{g(x)}\mathcal{A}_z + N_{g(x)}(\mathcal{M}_{p,q} - y).$$

Substituting (30) and (9) gives (31). □

The following lemma records the corresponding local normal decomposition of curves in $\mathcal{B}_z$.

Lemma 7 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$. Every smooth curve $B(t)$ in $\mathcal{B}_z$ with $B(0) = g(x)$ admits, after shrinking the interval if necessary, unique smooth curves $A(t)$ and $N(t)$ such that*

$$B(t) = A(t) + N(t), \qquad A(t) \in \mathcal{A}_z, \qquad N(t) \in N_{A(t)+y}\mathcal{M}_{p,q}, \qquad A(0) = g(x), \quad N(0) = 0.$$

Proof By Proposition 4(b), $E|_{\mathcal{W}_z} : \mathcal{W}_z \to \mathcal{B}_z$ is a diffeomorphism. Applying its inverse to $B(t)$ gives the unique smooth curve $(A(t), N(t)) \in \mathcal{W}_z$. The stated properties follow from (27) and (28). Since $B(0) = g(x) = E(g(x), 0)$, uniqueness gives $A(0) = g(x)$ and $N(0) = 0$. □

Remark 12 The manifold $\mathcal{A}_z$ is the translation by $Y_z - y$ of a neighborhood of X_z in the fixed-rank positive semidefinite stratum $\mathcal{M}_{p,0}$. At a KKT pair, $Y_z = y$, so $\mathcal{A}_z$ is the fixed-rank part of the original semidefinite constraint near $g(x)$.

The manifold $\mathcal{B}_z$ enlarges $\mathcal{A}_z$ along the normal directions to $\mathcal{M}_{p,q} - y$. At a KKT pair, $Y_z = y$ is the multiplier, and the β-subspace is the common null space of the primal matrix and the multiplier. Thus, passing from $\mathcal{A}_z$ to $\mathcal{B}_z$ amounts to allowing unrestricted symmetric perturbations on this subspace.

4.2 Constraint qualifications under stratification

We characterize the constraint nondegeneracy and the W-SRCQ as transversality conditions of g against $\mathcal{A}_z$ and $\mathcal{B}_z$, respectively. It is well known [7, (4.182)] that constraint nondegeneracy at a KKT pair is equivalent to $g \pitchfork_x \mathcal{M}_{p,0}$ in $\mathbb{S}^n$. The following proposition extends this to general primal–dual pairs via $\mathcal{A}_z$.

Proposition 5 *Let $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$. The following are equivalent:*
(a) *$g \pitchfork_x \mathcal{A}_z$ in $\mathbb{S}^n$;*
(b) *the constraint nondegeneracy holds at z.*

Proof Proposition 4(a) gives the tangent space $T_{g(x)}\mathcal{A}_z$. Substituting into the transversality condition $g'(x)\mathbb{X} + T_{g(x)}\mathcal{A}_z = \mathbb{S}^n$ yields exactly Definition 19. □

The next theorem establishes the analogous characterization for the W-SRCQ via $\mathcal{B}_z$.

Theorem 3 *Let $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$ with an IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$. The following are equivalent:*
(a) *$g \pitchfork_x \mathcal{B}_z$ in $\mathbb{S}^n$;*
(b) *the W-SRCQ holds at z.*

Proof The transversality condition $g \pitchfork_x \mathcal{B}_z$ asserts $g'(x)\mathbb{X} + T_{g(x)}\mathcal{B}_z = \mathbb{S}^n$. Substituting the tangent space from Proposition 4(b) gives exactly Definition 20. □

Figure 4 illustrates the geometry. The light blue surface represents the translated stratum $\mathcal{M}_{p,q} - y$, with the local model $\mathcal{A}_z$ shown as the dark blue curve. The bounded yellow ribbon represents $\mathcal{B}_z$. The red and green planes represent the affine tangent images $g_i(x) + g_i'(x)\mathbb{X}$. In the left subfigure, g_1 is transverse to $\mathcal{A}_z$, so both constraint nondegeneracy and W-SRCQ hold. In the right subfigure, g_2 is transverse to $\mathcal{B}_z$ but not to $\mathcal{A}_z$: W-SRCQ holds while constraint nondegeneracy fails.

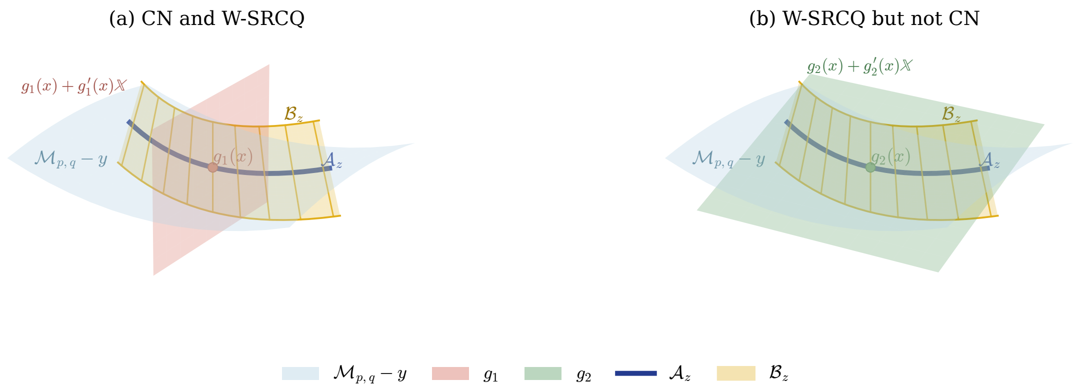

Fig. 4: Transversality interpretation: W-SRCQ vs. constraint nondegeneracy.

Corollary 1 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$ with $p + q = n$. Then the W-SRCQ holds at z if and only if the constraint nondegeneracy holds at z.*

Proof When $p + q = n$, the stratum $\mathcal{M}_{p,q}$ is open in $\mathbb{S}^n$, so $N_{g(x)}(\mathcal{M}_{p,q} - y) = \{0\}$ and $\mathcal{B}_z = \mathcal{A}_z$. Apply Proposition 5 and Theorem 3. □

Remark 13 Geometrically, the transversality characterization gives the local stability of the W-SRCQ along each lifted stratum. More precisely, if the W-SRCQ holds at $z \in \widetilde{\mathcal{M}}_{p,q}$, then there exists a neighborhood $\widetilde{\mathcal{U}}$ of z relative to $\widetilde{\mathcal{M}}_{p,q}$ such that the W-SRCQ holds at every $z' \in \widetilde{\mathcal{U}}$. However, this statement follows directly from Proposition 9, which establishes the stronger stability under perturbations that preserve the negative inertia index of $G(z)$. For brevity, we omit a separate proof here.

Remark 14 Under affine-linear perturbations consisting of a linear term added to the objective and an affine term added to the constraint mapping, the transversality constraint qualification at feasible points is generic [22, Theorem 8 and Remark 9]. At every complementary primal–dual pair (x, y), this transversality condition at x is equivalent to the extended constraint nondegeneracy at (x, y), which in turn implies the W-SRCQ. Hence the W-SRCQ is generic on the set of complementary primal–dual pairs in the same sense.

4.3 Second-order conditions under stratification

We now construct and study the following two manifold optimization problems associated with the auxiliary NLSDP (25):

$$(\mathcal{P}_z^A) \qquad \min_{g(u)\in\mathcal{A}_z} f_z(u),$$

and

$$(\mathcal{P}_z^B) \qquad \min_{g(u)\in\mathcal{B}_z} f_z(u).$$

By (24), their constraints can equivalently be written as

$$g_z(u) \in \mathcal{A}_z + y - Y_z \quad \text{and} \quad g_z(u) \in \mathcal{B}_z + y - Y_z,$$

respectively. Moreover,

$$\mathcal{A}_z + y - Y_z = \mathcal{M}_{p,0} \cap \mathcal{B}\Big(X_z, \tfrac{\delta(z)}{2}\Big) \subseteq \mathbb{S}^n_+.$$

Hence $(\mathcal{P}_z^A)$ is a local fixed-rank restriction of (25), whereas $\mathcal{B}_z + y - Y_z$ is not generally contained in $\mathbb{S}^n_+$.

Proposition 6 *The following statements hold.*

(a) *If constraint nondegeneracy holds at z, then $g^{-1}(\mathcal{A}_z)$ is a smooth manifold near x and*

$$T_x g^{-1}(\mathcal{A}_z) = \{v_x \in \mathbb{X} \mid \nabla g(x)^* v_x \in T_{g(x)}\mathcal{A}_z\} = \mathrm{appl}(z).$$

(b) *If the W-SRCQ holds at z, then $g^{-1}(\mathcal{B}_z)$ is a smooth manifold near x and*

$$T_x g^{-1}(\mathcal{B}_z) = \{v_x \in \mathbb{X} \mid \nabla g(x)^* v_x \in T_{g(x)}\mathcal{B}_z\} = \mathrm{app}(z).$$

Proof By Proposition 5, constraint nondegeneracy is equivalent to $g \pitchfork_x \mathcal{A}_z$, and by Theorem 3, the W-SRCQ is equivalent to $g \pitchfork_x \mathcal{B}_z$. The transverse preimage theorem [32, Theorem 3.3.4] then gives the smoothness of the corresponding inverse image near x and the tangent formula

$$T_x g^{-1}(\mathcal{N}) = (g'(x))^{-1}(T_{g(x)}\mathcal{N}) = \{v_x \in \mathbb{X} \mid \nabla g(x)^* v_x \in T_{g(x)}\mathcal{N}\},$$

with $\mathcal{N} = \mathcal{A}_z$ or $\mathcal{B}_z$. Substituting (30) gives exactly (22), and substituting (31) gives exactly (21). The proof is then completed. □

With the tangent spaces identified, we next compute the common curvature term that determines the restricted Hessians of the two manifold models.

Lemma 8 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$ with IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$. Let either $A(t)$ be a smooth curve in $\mathcal{A}_z$ with $A(0) = g(x)$, or $B(t)$ be a smooth curve in $\mathcal{B}_z$ with $B(0) = g(x)$. In each case, set H to be the initial velocity of the chosen curve and $\widetilde{H} := P^\top HP$. Then, respectively,*

$$\begin{cases} \langle Y_z, \ddot{A}(0)\rangle = 2\displaystyle\sum_{i\in\alpha}\sum_{j\in\gamma}\frac{\lambda_j}{\lambda_i}\,\widetilde{H}_{ij}^2, \\ \langle Y_z, \ddot{B}(0)\rangle = 2\displaystyle\sum_{i\in\alpha}\sum_{j\in\gamma}\frac{\lambda_j}{\lambda_i}\,\widetilde{H}_{ij}^2. \end{cases} \tag{32}$$

Proof We first treat the cases in which one of the index sets is empty. If $\gamma = \emptyset$, then the sum on the right-hand side of (32) is empty and $Y_z = P\,\mathrm{Diag}(0_\alpha, 0_\beta)P^\top = 0$. Hence both identities in (32) hold.

Suppose next that $\alpha = \emptyset$. Then $p = 0$, $X_z = 0$, and $\mathcal{M}_{0,0} = \{0\}$, so $\mathcal{A}_z = \{Y_z - y\} = \{g(x)\}$. Hence every curve $A(t) \in \mathcal{A}_z$ is constant and $\langle Y_z, \ddot{A}(0)\rangle = 0$.

For a curve $B(t) \in \mathcal{B}_z$, (28)–(29) give $B(t) - g(x) \in N_{g(x)+y}\mathcal{M}_{0,q}$. By (9), $P^\top(B(t) - g(x))P$, and hence $P^\top \ddot{B}(0)P$, has only a $\beta\beta$ block. Since $P^\top Y_z P = \mathrm{Diag}(0_\beta, \Lambda_\gamma)$, we have $\langle Y_z, \ddot{B}(0)\rangle = 0$. The sum in (32) is empty, so both identities follow.

We may therefore assume that both α and γ are nonempty.

First suppose $A(t) \in \mathcal{A}_z$. Write $A(t) = X(t) + Y_z - y$ with $X(t) \in \mathcal{M}_{p,0}$ and $\dot{X}(0) = H$. Let $\widetilde{X}(t) := P^\top X(t)P$. Since $X(t)$ has rank p and $\widetilde{X}_{\alpha\alpha}(0) = \Lambda_\alpha$ is nonsingular, the Schur complement identity gives, for t close to zero,

$$\widetilde{X}_{\alpha^c\alpha^c}(t) = \widetilde{X}_{\alpha^c\alpha}(t)\widetilde{X}_{\alpha\alpha}(t)^{-1}\widetilde{X}_{\alpha\alpha^c}(t).$$

Differentiating twice at $t=0$, using $\widetilde{X}_{\alpha^c\alpha^c}(0)=0$ and $\widetilde{X}_{\alpha^c\alpha}(0)=0$, yields

$$\ddot{\widetilde{X}}_{\alpha^c\alpha^c}(0)=2\,\widetilde{H}_{\alpha^c\alpha}\Lambda_\alpha^{-1}\widetilde{H}_{\alpha\alpha^c}.$$

Since $Y_z=P\,\mathrm{Diag}(0_\alpha,0_\beta,\Lambda_\gamma)P^\top$, only the $\gamma\gamma$ block pairs with Y_z, and hence

$$\begin{aligned}\langle Y_z,\ddot{A}(0)\rangle &= \left\langle \Lambda_\gamma,\ddot{\widetilde{X}}_{\gamma\gamma}(0)\right\rangle\\ &=2\left\langle \Lambda_\gamma,\widetilde{H}_{\gamma\alpha}\Lambda_\alpha^{-1}\widetilde{H}_{\alpha\gamma}\right\rangle\\ &=2\sum_{i\in\alpha}\sum_{j\in\gamma}\frac{\lambda_j}{\lambda_i}\,\widetilde{H}_{ij}^2.\end{aligned}$$

Now suppose $B(t)\in\mathcal{B}_z$. By Lemma 7, write $B(t)=A(t)+N(t)$, where $A(t)\in\mathcal{A}_z$, $N(t)\in N_{A(t)+y}\mathcal{M}_{p,q}$, $A(0)=g(x)$, and $N(0)=0$. If $\beta=\emptyset$, then (9) gives $N(t)=0$, so $B(t)=A(t)$ and the second identity follows from the first. Hence suppose $\beta\neq\emptyset$.

Since $A(t)+y\in\mathcal{M}_{p,q}$, its kernel has the constant dimension $|\beta|$, and for t close to zero the nonzero eigenvalues of $A(t)+y$ remain bounded away from zero. Apply Lemma 3 at $G(z)$ with $\delta:=\delta(z)/2$. After shrinking the interval, $A(t)+y$ lies in the neighborhood from that lemma and its eigenvalues in $(-\delta(z)/2,\delta(z)/2)$ are exactly its $|\beta|$ zero eigenvalues. Hence $U(t):=U_\delta(A(t)+y)$ is smooth in t, $U(0)=P_\beta$, and its columns form an orthonormal basis of $\ker(A(t)+y)$. Define the smooth curve

$$C(t):=U(t)^\top N(t)U(t)\in\mathbb{S}^{|\beta|}.$$

Since $N(t)\in N_{A(t)+y}\mathcal{M}_{p,q}$ and the columns of $U(t)$ form an orthonormal basis of $\ker(A(t)+y)$, formula (9) gives $N(t)=U(t)C(t)U(t)^\top$. Since $N(0)=0$, we have $C(0)=0$, and hence $C(t)=O(t)$. Moreover, $U(0)=P_\beta$ and the smoothness of U give $P_\alpha^\top U(t)=O(t)$ and $P_\gamma^\top U(t)=O(t)$. Hence the $\gamma\gamma$ and $\alpha\gamma$ blocks of $P^\top N(t)P$ are $O(t^3)$. Because $P^\top Y_z P$ has only a $\gamma\gamma$ block, $\langle Y_z,N(t)\rangle=O(t^3)$, and hence $\langle Y_z,\ddot{N}(0)\rangle=0$. Moreover, $N(t)=B(t)-A(t)$ has zero $\alpha\gamma$ block to first order at $t=0$, so $(P^\top\dot{A}(0)P)_{ij}=\widetilde{H}_{ij}$ for $i\in\alpha$ and $j\in\gamma$. Applying the first part to $A(t)$ and using $\ddot{B}(0)=\ddot{A}(0)+\ddot{N}(0)$ gives

$$\langle Y_z,\ddot{B}(0)\rangle=\langle Y_z,\ddot{A}(0)\rangle=2\sum_{i\in\alpha}\sum_{j\in\gamma}\frac{\lambda_j}{\lambda_i}\,\widetilde{H}_{ij}^2.$$

□

We next use this curvature formula to identify $\mathcal{Q}_z$ with the Riemannian Hessian quadratic form of each pullback manifold.

Proposition 7 *The following identities hold.*

(a) *If constraint nondegeneracy holds at z, then*

$$\left\langle v_x,\mathrm{Hess}\big(f_z|_{g^{-1}(\mathcal{A}_z)}\big)(x)v_x\right\rangle=\mathcal{Q}_z(v_x)\qquad\forall\,v_x\in\mathrm{appl}(z).$$

(b) *If the W-SRCQ holds at z, then*

$$\left\langle v_x,\mathrm{Hess}\big(f_z|_{g^{-1}(\mathcal{B}_z)}\big)(x)v_x\right\rangle=\mathcal{Q}_z(v_x)\qquad\forall\,v_x\in\mathrm{app}(z).$$

Proof We prove part (a). The proof of part (b) is identical with $\mathcal{A}_z$ and $\mathrm{appl}(z)$ replaced by $\mathcal{B}_z$ and $\mathrm{app}(z)$, respectively. By Proposition 6, $g^{-1}(\mathcal{A}_z)$ is a smooth manifold near x with tangent space $\mathrm{appl}(z)$. For $v_x\in\mathrm{appl}(z)$, set $u(t):=\exp_x(tv_x)$ on this manifold and $A(t):=g(u(t))$. Then $A(t)\in\mathcal{A}_z$ and $\dot{A}(0)=g'(x)v_x$. By Definition 17, the chain rule, and $\nabla f_z(x)+\nabla g(x)Y_z=0$, we have

$$\begin{aligned}\left\langle v_x,\mathrm{Hess}\big(f_z|_{g^{-1}(\mathcal{A}_z)}\big)(x)v_x\right\rangle &=\frac{d^2}{dt^2}f_z(u(t))\Big|_{t=0}=\langle v_x,\nabla^2 f_z(x)v_x\rangle+\langle\nabla f_z(x),\ddot{u}(0)\rangle\\ &=\langle v_x,\nabla^2 f_z(x)v_x\rangle-\langle Y_z,g'(x)\ddot{u}(0)\rangle\\ &=\langle v_x,\nabla^2 f_z(x)v_x\rangle-\langle Y_z,\ddot{A}(0)\rangle+\langle Y_z,g''(x)(v_x,v_x)\rangle\\ &=\langle v_x,\nabla^2_{xx}L(x,y)v_x\rangle-\langle Y_z,\ddot{A}(0)\rangle.\end{aligned}$$

Applying Lemma 8 and (19) proves both identities. □

Remark 15 By the Gauss formula for Riemannian submanifolds [9, Lemma 5.50], the curvature terms in Lemma 8 are the second fundamental forms of $\mathcal{A}_z$ and $\mathcal{B}_z$ paired with Y_z. Although $\mathcal{A}_z$ is strictly contained in $\mathcal{B}_z$ when $\beta \neq \emptyset$, the additional normal-bundle directions in $\mathcal{B}_z$ make no contribution after pairing the second fundamental form with Y_z. Hence the two paired second fundamental forms have the same quadratic representation in H.

Together with the Riemannian submanifold Hessian formula [9, Corollary 5.47], Proposition 7 identifies $\mathcal{Q}_z$ as the Riemannian Hessian quadratic form of both problems. Passing from $\mathcal{A}_z$ to $\mathcal{B}_z$ enlarges the tangent domain from $\mathrm{appl}(z)$ to $\mathrm{app}(z)$ without changing $\mathcal{Q}_z$. At a KKT pair $\bar{z} = (\bar{x}, \bar{y})$, the common curvature expression with $H = g'(\bar{x})v_x$ is the classical sigma term by (23).

We now identify the W-SOC and the S-SOC with the manifold SOCs of $(\mathcal{P}_z^A)$ and $(\mathcal{P}_z^B)$, respectively.

Theorem 4 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$, and suppose that constraint nondegeneracy holds at z. Then x is a stationary point of $(\mathcal{P}_z^A)$ with the unique multiplier Y_z. Moreover, the W-SOC holds at z if and only if the manifold SOC holds at x for $(\mathcal{P}_z^A)$.*

Proof By Proposition 5, constraint nondegeneracy is equivalent to $g \pitchfork_x \mathcal{A}_z$. Moreover, $\nabla f_z(x) + \nabla g(x)Y_z = 0$, and Proposition 4(a) gives $Y_z \in N_{g(x)}\mathcal{A}_z$. Thus x is stationary. If two multipliers existed, their difference would be orthogonal to both $g'(x)\mathbb{X}$ and $T_{g(x)}\mathcal{A}_z$, and hence would vanish by transversality. The multiplier is therefore unique.

By Proposition 6, the tangent space of the inverse-image manifold is $\mathrm{appl}(z)$. Consequently, Proposition 7(a) and Definition 24 show that the manifold SOC at x for $(\mathcal{P}_z^A)$ is equivalent to the W-SOC at z. □

Theorem 5 *Let $z = (x, y) \in \widetilde{\mathcal{M}}_{p,q}$, and suppose that the W-SRCQ holds at z. Then x is a stationary point of $(\mathcal{P}_z^B)$ with the unique multiplier Y_z. Moreover, the S-SOC holds at z if and only if the manifold SOC holds at x for $(\mathcal{P}_z^B)$.*

Proof By Theorem 3, the W-SRCQ is equivalent to $g \pitchfork_x \mathcal{B}_z$. Moreover, $\nabla f_z(x) + \nabla g(x)Y_z = 0$, and Proposition 4(b) gives $Y_z \in N_{g(x)}\mathcal{B}_z$. Thus x is stationary, and the same transversality argument as in Theorem 4 proves uniqueness of the multiplier.

By Proposition 6, the tangent space of the inverse-image manifold is $\mathrm{app}(z)$. Consequently, Proposition 7(b) and Definition 23 show that the manifold SOC at x for $(\mathcal{P}_z^B)$ is equivalent to the S-SOC at z. □

At a KKT pair $z = (x, y)$ of the original NLSDP, Proposition 3 gives $X_z = g(x)$, $Y_z = y$, $g_z = g$, and $f_z = f$. Hence $(\mathcal{P}_z^A)$ and $(\mathcal{P}_z^B)$ use the original objective f and constraint mapping g.

The preceding results also clarify how our geometric interpretations relate to the classical ones.

Remark 16 Constraint nondegeneracy already admits the classical interpretation as transversality to the active stratum [7, (4.182)], so Proposition 5 extends this interpretation to arbitrary primal–dual pairs rather than introducing a fundamentally new one. By contrast, Theorem 5 provides a new geometric interpretation of the second-order conditions without assuming strict complementarity. For comparison, Shapiro's interpretation of the S-SOSC as a manifold second-order sufficient condition is established under constraint nondegeneracy and strict complementarity [46, Theorem 9, condition (46), and its proof]. Under constraint nondegeneracy, Proposition 7 shows that the S-SOSC is likewise equivalent to the manifold second-order sufficient condition for $(\mathcal{P}_z^B)$, without requiring strict complementarity. When strict complementarity holds, $\mathcal{B}_z = \mathcal{A}_z$ locally, and the present interpretation reduces to Shapiro's.

Remark 17 At a KKT pair, the two manifold models suggest a geometric explanation for the selected Bouligand generalized Jacobian elements $\mathcal{U}_0$ and $\mathcal{U}_I$ and their associated nonsingularity results in [25]. A complete proof requires establishing the well-definedness of local nearest-point projections onto $\mathcal{A}_z$ and $\mathcal{B}_z$ along their normal bundles and deriving the relevant derivative identities. Since this requires a lengthy auxiliary development and is not needed in the subsequent analysis, we omit the details and only outline the argument. Given the well-definedness of these local projections, the KKT systems of $(\mathcal{P}_z^A)$ and $(\mathcal{P}_z^B)$ can be expressed as smooth equations. Up to an invertible change of primal–dual coordinates, the Jacobians of these equations formally correspond to $\mathcal{U}_0$ and $\mathcal{U}_I$, respectively. Thus, the manifold geometry developed here suggests a smooth origin for these objects arising in nonsmooth analysis and relates their nonsingularity to transversality and the corresponding manifold second-order conditions.

Figure 5 summarizes the two smooth manifold models and their geometric interpretations.

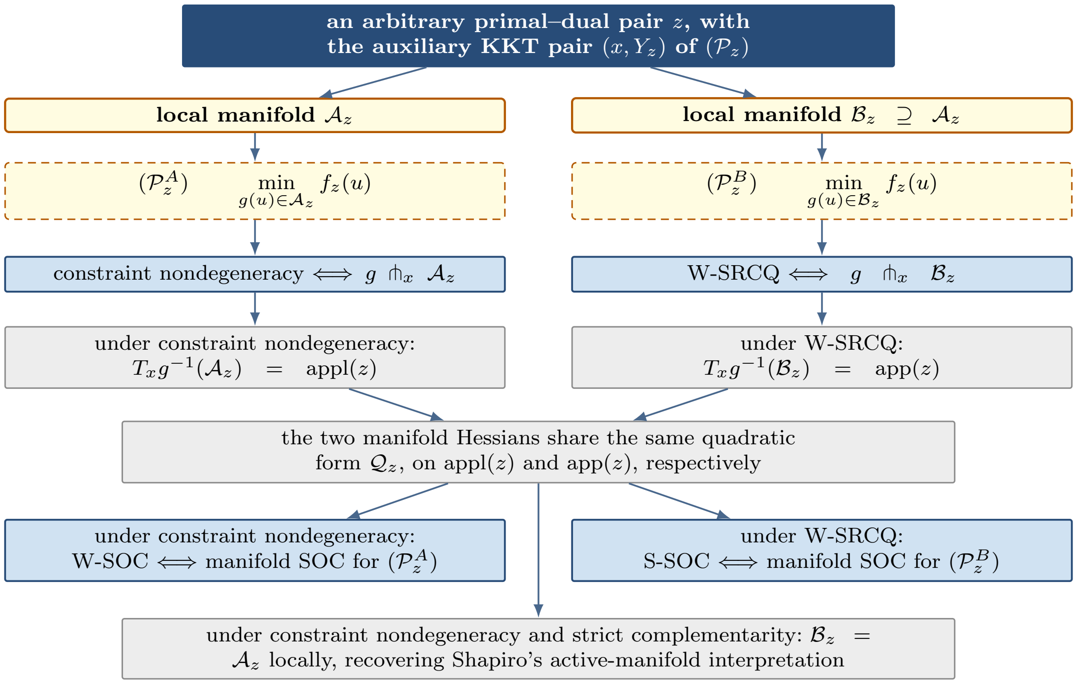

Fig. 5: Geometric interpretations associated with $\mathcal{A}_z$ and $\mathcal{B}_z$.

5 Stratum-restricted strong metric regularity and stratified perturbation analysis

This section analyzes F along individual lifted strata and studies its behavior under perturbations across strata. On a fixed stratum, we connect the weak conditions with injectivity of dF_z and stratum-restricted strong metric regularity. Across strata, we show the stability of the weak conditions under perturbations preserving the relevant inertia index and recover the corresponding strong conditions under arbitrary perturbations. Together, these results characterize strong regularity through the local and local uniform validity of the stratified conditions.

5.1 Analysis along a stratum

This subsection analyzes the restriction of the KKT mapping to a fixed lifted stratum, on which both inertia indices of $G(z)$ are preserved.

Definition 28 A set-valued mapping $\Phi : \mathbb{X} \rightrightarrows \mathbb{Y}$ is said to be *stratum-restricted strongly metrically regular* on a stratum $\mathcal{M} \subseteq \mathbb{X}$ at $(\bar{x}, \bar{y})$ if $\bar{x} \in \mathcal{M}$, $\bar{y} \in \Phi(\bar{x})$, and there exist a constant $\kappa > 0$ and neighborhoods $\mathcal{U}$ of $\bar{x}$ and $\mathcal{V}$ of $\bar{y}$ such that for all $x \in \mathcal{U} \cap \mathcal{M}$ and $y \in \mathcal{V} \cap \Phi(\mathcal{U} \cap \mathcal{M})$,

$$\mathrm{dist}\big(x, \Phi^{-1}(y) \cap \mathcal{M} \cap \mathcal{U}\big) \leq \kappa \cdot \mathrm{dist}\big(y, \Phi(x)\big),$$

and $(\Phi^{-1}(y) \cap \mathcal{M}) \cap \mathcal{U}$ is a singleton for every $y \in \mathcal{V} \cap \Phi(\mathcal{U} \cap \mathcal{M})$.

Remark 18 In the present setting F is single-valued and smooth on $\widetilde{\mathcal{M}}_{p,q}$, so stratum-restricted strong metric regularity of F on $\widetilde{\mathcal{M}}_{p,q}$ at $(\bar{z}, F(\bar{z}))$ implies that the restriction $F|_{\mathcal{U}}$ is a Lipschitz homeomorphism from a neighborhood $\mathcal{U} \subseteq \widetilde{\mathcal{M}}_{p,q}$ of $\bar{z}$ onto its local image $F(\mathcal{U})$.

Using the explicit block formula for dF_z, the following theorem gives the key differential criterion for the analysis along a fixed stratum. It relates the injectivity of dF_z to the W-SRCQ and the W-SOC, with the converse obtained under the SRCQ and the SONC. At a KKT pair, Theorem 7 further turns this injectivity into stratum-restricted strong metric regularity.

Theorem 6 *For $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$, the following statements hold.*

(a) *If the W-SOC and the W-SRCQ hold at z, then the manifold differential $dF_z : T_z\widetilde{\mathcal{M}} \to \mathbb{X} \times \mathbb{S}^n$ given by (15) is injective.*
(b) *If dF_z is injective, then the W-SRCQ holds at z.*

(c) *If dF_z is injective and the SRCQ and the SONC hold at z, then the W-SOC holds at z.*

Proof Fix an IED $(\alpha,\beta,\gamma,p,q,P,\lambda)$ of $A := G(z)$. By Proposition 1 and (10), the kernel of $\xi_A : T_A\mathcal{M}_{p,q} \to \mathbb{S}^n$ consists precisely of the matrices H for which $P^\top HP$ is supported on the $\beta\gamma$, $\gamma\beta$, and $\gamma\gamma$ blocks. Taking the orthogonal complement in $\mathbb{S}^n$, $(\ker\xi_A)^\perp$ is therefore the matrix space in (17). Taking orthogonal complements there, and using (21), gives

$$\begin{aligned} \text{W-SRCQ at } z \quad &\Longleftrightarrow \quad (\ker\xi_A) \cap \ker\nabla g(x) = \{0\}, \\ \mathrm{app}(z) &= \big(\nabla g(x)(\ker\xi_A)\big)^\perp. \end{aligned} \tag{33}$$

Let $\mathfrak{q}_z$ be the symmetric bilinear form associated with $\mathcal{Q}_z$:

$$\mathfrak{q}_z(u,v) := \langle u, \nabla^2_{xx}L(z)v\rangle + 2\sum_{i\in\alpha}\sum_{j\in\gamma}\frac{-\lambda_j}{\lambda_i}[P^\top(\nabla g(x)^*u)P]_{ij}[P^\top(\nabla g(x)^*v)P]_{ij}.$$

Thus $\mathcal{Q}_z(v) = \mathfrak{q}_z(v,v)$. For $v \in \mathrm{appl}(z)$, define $H_v \in T_A\mathcal{M}_{p,q}$ by

$$P^\top H_v P = \begin{bmatrix} [P^\top(\nabla g(x)^*v)P]_{\alpha\alpha} & [P^\top(\nabla g(x)^*v)P]_{\alpha\beta} & \widehat{\Xi}_{\alpha\gamma}\circ[P^\top(\nabla g(x)^*v)P]_{\alpha\gamma} \\ [P^\top(\nabla g(x)^*v)P]_{\beta\alpha} & 0 & 0 \\ \big(\widehat{\Xi}_{\alpha\gamma}\circ[P^\top(\nabla g(x)^*v)P]_{\alpha\gamma}\big)^\top & 0 & 0 \end{bmatrix},$$

where $(\widehat{\Xi}_{\alpha\gamma})_{ij} := \Xi_{ij}^{-1}$ for $i\in\alpha$ and $j\in\gamma$. Formula (10) gives $\xi_A(H_v) = \nabla g(x)^*v$. Moreover, $\Xi_{ij}^{-1} - 1 = -\lambda_j/\lambda_i$ for $i\in\alpha$ and $j\in\gamma$. Consequently, the definition of $\mathfrak{q}_z$ yields

$$\big\langle d, \nabla^2_{xx}L(z)v + \nabla g(x)\big(H_v - \nabla g(x)^*v\big)\big\rangle = \mathfrak{q}_z(d,v) \qquad \forall\, d\in\mathbb{X}. \tag{34}$$

(a) Let $(v,H)\in\mathbb{X}\times T_A\mathcal{M}_{p,q}$ belong to the kernel of $dF_z\circ\phi_z^{-1}$. By (15),

$$\nabla^2_{xx}L(z)v + \nabla g(x)\big(H - \nabla g(x)^*v\big) = 0, \qquad \xi_A(H) = \nabla g(x)^*v.$$

The second equation and (10) imply $v\in\mathrm{appl}(z)$. Since $H, H_v \in T_A\mathcal{M}_{p,q}$ and $\xi_A(H) = \nabla g(x)^*v = \xi_A(H_v)$, we have $K := H - H_v \in \ker\xi_A$. Then $H = H_v + K$, and the first equation becomes

$$\nabla^2_{xx}L(z)v + \nabla g(x)\big(H_v - \nabla g(x)^*v\big) + \nabla g(x)K = 0.$$

Pairing it with v, and using (33), (34), and $\mathrm{appl}(z)\subseteq\mathrm{app}(z)$, gives $\mathcal{Q}_z(v) = 0$. The W-SOC gives $v = 0$. It follows that $H \in (\ker\xi_A)\cap\ker\nabla g(x)$, so the W-SRCQ and (33) give $H = 0$. Thus dF_z is injective.

(b) If the W-SRCQ failed, (33) would give a nonzero $H\in(\ker\xi_A)\cap\ker\nabla g(x)$. Then (15) would yield $(dF_z\circ\phi_z^{-1})(0,H) = 0$, contradicting injectivity. Hence the W-SRCQ holds.

(c) Suppose, to the contrary, that the W-SOC fails. Then some $0\neq v\in\mathrm{appl}(z)$ satisfies $\mathcal{Q}_z(v) = 0$. For every $d\in\mathcal{C}(z)$ and $t\in\mathbb{R}$, (20) and (22) give $d + tv\in\mathcal{C}(z)$. The SONC therefore implies

$$0 \le \mathcal{Q}_z(d+tv) = \mathcal{Q}_z(d) + 2t\,\mathfrak{q}_z(d,v) \qquad \forall\, t\in\mathbb{R},$$

and hence $\mathfrak{q}_z(d,v) = 0$ for every $d\in\mathcal{C}(z)$. By the SRCQ and Lemma 5, this equality extends to every $d\in\mathrm{app}(z)$. Equations (33) and (34) now give

$$\nabla^2_{xx}L(z)v + \nabla g(x)\big(H_v - \nabla g(x)^*v\big) \in \nabla g(x)(\ker\xi_A).$$

Choose $R\in\ker\xi_A$ such that

$$\nabla g(x)R = -\nabla^2_{xx}L(z)v - \nabla g(x)\big(H_v - \nabla g(x)^*v\big),$$

and set $H := H_v + R$. Then

$$\xi_A(H) = \nabla g(x)^*v, \qquad \nabla^2_{xx}L(z)v + \nabla g(x)\big(H - \nabla g(x)^*v\big) = 0.$$

Thus (15) gives $(dF_z\circ\phi_z^{-1})(v,H) = 0$. Since $v\neq 0$, this contradicts the injectivity of dF_z. Therefore the W-SOC holds. □

Remark 19 The injectivity of dF_z and the W-SRCQ are nondegeneracy conditions. The W-SOC is a definiteness condition: $\mathcal{Q}_z \neq 0$ on the subspace $\mathrm{appl}(z)\setminus\{0\}$ means that $\mathcal{Q}_z$ is positive or negative definite there. In part (c) the SONC excludes the negative-definite alternative. The SRCQ guarantees that the critical cone $\mathcal{C}(z)$ spans $\mathrm{app}(z)$ (Lemma 5). Without the SRCQ, the SONC controls $\mathcal{Q}_z$ only on $\mathcal{C}(z)$, which may span a proper subspace of $\mathrm{app}(z)$, leaving the remaining directions required by the converse implication unconstrained. Examples 2 and 3 show that neither assumption can be dropped.

Example 2 Let $f(x) = x_1x_2$ and $g(x) = x_1 + 1 \in \mathbb{S}^1$ for $x \in \mathbb{R}^2$. Then $z = (0,0) \in \mathbb{R}^2 \times \mathbb{S}^1$ is a KKT pair with $G(z) = 1$ and $\beta = \gamma = \emptyset$, so $\mathcal{C}(z) = \mathrm{appl}(z) = \mathbb{R}^2$ and $\mathcal{Q}_z(v) = 2v_1v_2$. Since $X_z = 1$ is positive definite, $T_{\mathbb{S}^1_+}(X_z) = \mathbb{S}^1$ and the SRCQ holds. Since $\mathcal{Q}_z(1,-1) < 0$, the SONC fails. Near z we have $\Pi_{\mathbb{S}^1_+}(G) = G$ and hence $F(x,y) = (\nabla_x L(x,y),\, y)$, whose Jacobian at z is block triangular with nonsingular diagonal blocks $\nabla^2_{xx}L(z) = \nabla^2 f(0)$ and 1, so on the open stratum $\widetilde{\mathcal{M}}_{1,0}$ the differential $dF_z = \nabla F(z)$ is injective. Since $\mathcal{Q}_z(1,0) = 0$, the W-SOC fails at z.

Example 3 Let $f(x) = x_1x_2$ and $g(x) = \mathrm{Diag}(x_1, -x_1) \in \mathbb{S}^2$ for $x \in \mathbb{R}^2$. Then $z = (0,0) \in \mathbb{R}^2 \times \mathbb{S}^2$ is a KKT pair with $G(z) = 0$ and $\beta = \{1,2\}$, so, taking $P = I_2$, $\mathcal{C}(z) = \mathrm{appl}(z) = \{v \in \mathbb{R}^2 \mid v_1 = 0\}$ and $\mathcal{Q}_z(v) = 2v_1v_2$. Since $\mathcal{Q}_z = 0$ on $\mathcal{C}(z)$, the SONC holds and the W-SOC fails. By $X_z = Y_z = 0$, the SRCQ requires $g'(x)\mathbb{R}^2 + \mathbb{S}^2_+ = \mathbb{S}^2$. Every matrix in this sum has nonnegative trace, so the SRCQ fails. On $\widetilde{\mathcal{M}}_{0,0} = \{(x,y) \mid y = -g(x)\}$ we have $G \equiv 0$ and $F(x,y) = (\nabla_x L(x,y),\, -g(x))$, so

$$dF_z\big(v,\, -g'(x)v\big) = \big((v_2 - 2v_1,\, v_1),\, -\mathrm{Diag}(v_1, -v_1)\big),$$

and dF_z is injective.

For $p + q = n$ the stratum $\widetilde{\mathcal{M}}_{p,q}$ is open in $\mathbb{X} \times \mathbb{S}^n$, and at its points $T_z\widetilde{\mathcal{M}}_{p,q} = \mathbb{X} \times \mathbb{S}^n$, $dF_z = \nabla F(z)$, and $\mathrm{app}(z) = \mathrm{appl}(z)$. In this setting the kernel analysis underlying Theorem 6 admits the following quantitative counterpart.

Lemma 9 *Let $z = (x,y) \in \widetilde{\mathcal{M}}_{p,q}$ with $p + q = n$, and suppose that $\nabla F(z)$ is nonsingular. Then*

$$\sigma_{\min}\big(\mathcal{Q}_z|_{\mathrm{appl}(z)}\big) \;\geq\; \kappa_z^{-1}, \qquad \kappa_z := \|\nabla F(z)^{-1}\|\,\big(1 + \|\nabla g(x)\|\big).$$

Proof Let $(\alpha, \emptyset, \gamma, p, q, P, \lambda)$ be an IED of $G(z)$ and $\xi := \xi_{G(z)}$. Fix a unit $v \in \mathrm{appl}(z)$. It suffices to construct $H \in \mathbb{S}^n$ with

$$\big(\nabla F(z) \circ \phi_z^{-1}\big)(v, H) = \big(\mathcal{Q}_z|_{\mathrm{appl}(z)}\, v,\; 0\big), \tag{35}$$

where ϕ_z is the isomorphism of (13): indeed, $\|\phi_z\| \leq 1 + \|\nabla g(x)\|$, so (35) yields

$$\big\|\mathcal{Q}_z|_{\mathrm{appl}(z)} v\big\| \;\geq\; \frac{\|\phi_z^{-1}(v,H)\|}{\|\nabla F(z)^{-1}\|} \;\geq\; \frac{\|(v,H)\|}{\|\nabla F(z)^{-1}\|\,\|\phi_z\|} \;\geq\; \kappa_z^{-1}\|v\| \;=\; \kappa_z^{-1}.$$

Since $p + q = n$, $dF_z = \nabla F(z)$, and Proposition 2 reads

$$\big(\nabla F(z) \circ \phi_z^{-1}\big)(v, H) = \Big(\nabla^2_{xx}L(z)v + \nabla g(x)\big(H - \nabla g(x)^* v\big),\;\; -\nabla g(x)^* v + \xi(H)\Big). \tag{36}$$

By (10) and (22) with $\beta = \emptyset$, there is a unique $H_0 \in \mathbb{S}^n$ with $[P^\top H_0 P]_{\gamma\gamma} = 0$ and $\xi(H_0) = \nabla g(x)^* v$, and it satisfies

$$[P^\top(H_0 - \nabla g(x)^* v)P]_{ij} = \Big(\frac{1}{\Xi_{ij}} - 1\Big)[P^\top(\nabla g(x)^* v)P]_{ij} = \frac{-\lambda_j}{\lambda_i}\,[P^\top(\nabla g(x)^* v)P]_{ij} \quad (i \in \alpha,\; j \in \gamma), \tag{37}$$

while all other blocks of $P^\top(H_0 - \nabla g(x)^* v)P$ vanish. Moreover, $\ker\xi$ consists of the H' with $P^\top H' P$ supported in the $\gamma\gamma$ block, so by (22) with $\beta = \emptyset$, the subspaces $\nabla g(x)\ker\xi$ and $\mathrm{appl}(z)$ are orthogonal complements of each other in $\mathbb{X}$.

Set $H := H_0 + H'$, where $H' \in \ker\xi$ is chosen with $\nabla g(x)H' = -\Pi_{\nabla g(x)\ker\xi}\, r$ for $r := \nabla^2_{xx}L(z)v + \nabla g(x)(H_0 - \nabla g(x)^* v)$, with $\Pi_{\nabla g(x)\ker\xi}$ the orthogonal projection onto $\nabla g(x)\ker\xi$. Then the two components of (36) equal $\Pi_{\mathrm{appl}(z)} r$ and 0.

It remains to verify $\Pi_{\mathrm{appl}(z)} r = \mathcal{Q}_z|_{\mathrm{appl}(z)} v$. For every $v' \in \mathrm{appl}(z)$,

$$\begin{aligned}
\langle v', r\rangle &= \langle v', \nabla^2_{xx}L(z)v\rangle + \big\langle \nabla g(x)^* v',\; H_0 - \nabla g(x)^* v\big\rangle \\
&= \langle v', \nabla^2_{xx}L(z)v\rangle + \big\langle P^\top(\nabla g(x)^* v')P,\; P^\top\big(H_0 - \nabla g(x)^* v\big)P\big\rangle \\
&= \langle v', \nabla^2_{xx}L(z)v\rangle + 2\sum_{i\in\alpha}\sum_{j\in\gamma} \frac{-\lambda_j}{\lambda_i}[P^\top(\nabla g(x)^* v')P]_{ij}\,[P^\top(\nabla g(x)^* v)P]_{ij} \\
&= \big\langle v',\; \mathcal{Q}_z|_{\mathrm{appl}(z)} v\big\rangle,
\end{aligned}$$

where the third equality uses (37), and the fourth follows from (19) and the definition of $\mathcal{Q}_z|_{\mathrm{appl}(z)}$. Since $\Pi_{\mathrm{appl}(z)} r$ and $\mathcal{Q}_z|_{\mathrm{appl}(z)} v$ both lie in $\mathrm{appl}(z)$, they coincide, and (35) holds with $H = H_0 + H'$. $\square$

When $\bar{z} = (\bar{x}, \bar{y})$ is a KKT pair, the injectivity of $dF_{\bar{z}}$ acquires an additional solution-level interpretation. At such a pair, the extended SRCQ, SONC, and W-SRCQ coincide with their KKT-point counterparts by the reductions in Section 3.2. Under the SONC, the non-vanishing W-SOC in Definition 24 likewise coincides with the strict-positivity condition in Definition 10; see (23).

Theorem 7 *Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of* (1) *with an IED $(\alpha, \beta, \gamma, p, q, \bar{P}, \bar{\lambda})$ of $G(\bar{z})$. The following two statements are equivalent:*

(a) *the manifold differential $dF_{\bar{z}}$ is injective;*
(b) *the KKT mapping F is stratum-restricted strongly metrically regular on $\widetilde{\mathcal{M}}_{p,q}$ at $(\bar{z}, F(\bar{z}))$.*

Moreover, if the SRCQ and the SONC hold at $\bar{z}$, statements (a) *and* (b) *are further equivalent to:*

(c) *the W-SOC and the W-SRCQ hold simultaneously at $\bar{z}$.*

Proof **(a)⇒(b).** By [36, Proposition 5.18], there exists a neighborhood $\mathcal{U} \subseteq \widetilde{\mathcal{M}}_{p,q}$ of $\bar{z}$ and a submanifold $\mathcal{U}'$ containing $F(\bar{z})$, embedded in $\mathbb{X} \times \mathbb{S}^n$, such that F is a diffeomorphism from $\mathcal{U}$ onto $\mathcal{U}'$. Denote the restriction of F to $\mathcal{U}$ by $F_{\mathcal{U}}$. The restriction $F_{\mathcal{U}}^{-1}$ is therefore single-valued and smooth on $F(\mathcal{U})$. After shrinking $\mathcal{U}$ and choosing $\mathcal{V}$ accordingly, this inverse is Lipschitz on the local image, so Definition 28 holds.

(b)⇒(a). By Remark 18, statement (b) implies the existence of a neighborhood $\mathcal{U} \subseteq \widetilde{\mathcal{M}}_{p,q}$ of $\bar{z}$ such that $F|_{\mathcal{U}}$ is bijective with a Lipschitz inverse. Suppose for contradiction that $dF_{\bar{z}}$ is not injective, so there exists $0 \neq v \in T_{\bar{z}}\widetilde{\mathcal{M}}_{p,q}$ with $dF_{\bar{z}}(v) = 0$. Choose a smooth curve $c : (-\varepsilon, \varepsilon) \to \mathcal{U}$ with $c(0) = \bar{z}$ and $c'(0) = v$. Since $F|_{\widetilde{\mathcal{M}}_{p,q}}$ is smooth, $F(c(t)) - F(\bar{z}) = t\, dF_{\bar{z}}(v) + o(t) = o(t)$, so

$$\|c(t) - \bar{z}\| = \|F_{\mathcal{U}}^{-1}(F(c(t))) - F_{\mathcal{U}}^{-1}(F(\bar{z}))\| \leq L\|F(c(t)) - F(\bar{z})\| = o(t),$$

contradicting $c(t) - \bar{z} = tv + o(t)$ with $v \neq 0$.

The equivalence (a)⇔(c) under the SRCQ and the SONC follows directly from Theorem 6. □

The following example contrasts stratum-restricted and ambient strong metric regularity.

Example 4 Consider

$$\min_{x \in \mathbb{R}} \frac{1}{2}x^2 \qquad \text{subject to} \qquad g(x) = \begin{bmatrix} 1 & 0 \\ 0 & x^2 \end{bmatrix} \succeq 0.$$

The pair $\bar{z} = (0, 0)$ is a KKT pair and $G(\bar{z}) = \mathrm{Diag}(1, 0)$. Since $\gamma = \emptyset$, the W-SRCQ holds at $\bar{z}$. Moreover, $\mathrm{appl}(\bar{z}) = \mathbb{R}$ and $\mathcal{Q}_{\bar{z}}(v_x) = v_x^2$, so the W-SOC also holds. A direct calculation gives

$$T_{\bar{z}}\widetilde{\mathcal{M}}_{1,0} = \big\{(v_x, v_y) \mid (v_y)_{22} = 0\big\}, \qquad dF_{\bar{z}}(v_x, v_y) = (v_x, v_y).$$

Hence, by Theorem 7, F is stratum-restricted strongly metrically regular at $\bar{z}$.

Let $E_{22} := \mathrm{Diag}(0, 1)$ and $z^t = (0, tE_{22})$. Since $F(z^t) = F(\bar{z}) = 0$ for all $t < 0$, $F^{-1}(0)$ is not locally single-valued and ambient strong metric regularity fails.

The Clarke generalized Jacobian from Definition 14 plays a central role in classical semismooth Newton methods. We next show that every element of the ambient Clarke generalized Jacobian acts on vectors tangent to the stratum as the manifold differential dF_z. In particular, if dF_z is not injective, then every element of $\partial_C F(z)$ is singular, with a common nonzero kernel direction tangent to the stratum.

Proposition 8 *For every $z = (x, y) \in \mathbb{X} \times \mathbb{S}^n$,*

$$\mathcal{U}v = dF_z v \qquad \textit{for all } \mathcal{U} \in \partial_C F(z) \textit{ and } v \in T_z\widetilde{\mathcal{M}}.$$

In particular, if the manifold differential dF_z is not injective, then every element of $\partial_C F(z)$ is singular.

Proof Let $\mathcal{U} \in \partial_C F(z)$ and $v = (v_x, v_y) \in T_z\widetilde{\mathcal{M}}$, and set $H := \nabla g(x)^* v_x + v_y \in T_{G(z)}\mathcal{M}_{p,q}$. The structural characterizations in [41, Lemma 11] and [48, Lemma 2.1] give $\mathcal{V} \in \partial_C \Pi_{\mathbb{S}^n_+}(G(z))$ and $\mathcal{W} \in \partial_C \Pi_{\mathbb{S}^{|\beta|}_+}(0)$ such that

$$\mathcal{U}v = \begin{bmatrix} \nabla^2_{xx} L(x, y)v_x + \nabla g(x)v_y \\ -\nabla g(x)^* v_x + \mathcal{V}(H) \end{bmatrix}$$

and, writing $\widetilde{H} = P^\top H P$,

$$\mathcal{V}(H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\beta\alpha} & \mathcal{W}(\widetilde{H}_{\beta\beta}) & 0 \\ \Xi_{\gamma\alpha} \circ \widetilde{H}_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^\top.$$

Because $H \in T_{G(z)}\mathcal{M}_{p,q}$, Proposition 1 gives $\widetilde{H}_{\beta\beta} = 0$. Hence $\mathcal{W}(\widetilde{H}_{\beta\beta}) = 0$ and $\mathcal{V}(H) = \xi_{G(z)}(H)$. The formula for dF_z in (15) now gives $\mathcal{U}v = dF_z v$. If dF_z is not injective, a nonzero $v \in T_z\widetilde{\mathcal{M}}$ with $dF_z v = 0$ therefore lies in the kernel of every element of $\partial_C F(z)$. □

Together with the injectivity characterization above, Proposition 8 connects the Clarke generalized Jacobian with the weak second-order conditions. Under the SRCQ and the SONC, the W-SOC used here, which is formulated by non-vanishing, is equivalent to strict positivity on $\mathrm{appl}(\bar{z}) \setminus \{0\}$. This yields the following equivalence.

Corollary 2 *Let $\bar{z} = (\bar{x}, \bar{y})$ be a KKT pair of* (1)*. Suppose that the SRCQ and the SONC hold at $\bar{z}$. Then the following statements are equivalent:*

(a) *the W-SOC and the W-SRCQ hold at $\bar{z}$;*
(b) *the manifold differential $dF_{\bar{z}}$ is injective;*
(c) *the KKT mapping F is stratum-restricted strongly metrically regular on $\widetilde{\mathcal{M}}_{p,q}$ at $(\bar{z}, F(\bar{z}))$;*
(d) *there exists a nonsingular element of $\partial_C F(\bar{z})$.*

In particular, the equivalence holds for convex instances of (1) *satisfying the SRCQ.*

Proof The implication (a)⇒(b) follows from Theorem 6(a), while (b)⇒(a) follows from Theorem 6(b)–(c). The equivalence (b)⇔(c) is Theorem 7, and Proposition 8 gives (d)⇒(b).

For every $v_x \in \mathrm{appl}(\bar{z})$, one has $tv_x \in \mathcal{C}(\bar{z})$ for all $t \in \mathbb{R}$. Hence the SONC implies the W-SONC. Consequently, the W-SOC is equivalent to strict positivity on $\mathrm{appl}(\bar{z}) \setminus \{0\}$, and (a)⇒(d) follows from [25, Proposition 4]. For convex instances, the SONC is automatic by Remark 9. □

Remark 20 Proposition 8 and Corollary 2 have a direct consequence for Newton linearizations. Let $\bar{z}$ be a KKT pair at which the SRCQ and the SONC hold. If the W-SOC and the W-SRCQ do not hold simultaneously, then stratum-restricted strong metric regularity fails and every element of $\partial_C F(\bar{z})$ is singular. Consequently, every semismooth Newton system formed from an element of $\partial_C F(\bar{z})$ is singular. Thus, in this setting, stratum-restricted strong metric regularity is necessary for the existence of a nonsingular semismooth Newton linearization.

5.2 Analysis across strata

This subsection turns to perturbations across neighboring strata and relates the weak conditions to their strong counterparts.

5.2.1 Constraint qualifications

The following proposition shows that the W-SRCQ is stable under perturbations that preserve the negative inertia index of $G(z)$.

Proposition 9 *Let $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{S}^n$, and let q be the negative inertia index of $G(\bar{z})$. If the W-SRCQ holds at $\bar{z}$, then there exists a neighborhood $\mathcal{U}$ of $\bar{z}$ such that the W-SRCQ holds at every $z \in \mathcal{U}$ for which $G(z)$ has negative inertia index q.*

Proof Let $(\alpha, \beta, \gamma, p, q, P, \lambda)$ be an IED of $G(\bar{z})$. Suppose for contradiction that the assertion fails. Then there exists a sequence $z^\nu = (x^\nu, y^\nu) \to \bar{z}$ such that the negative inertia index of $G(z^\nu)$ is q and the W-SRCQ fails at each z^ν. Passing to a subsequence and relabeling the IEDs if necessary, we write an IED of $G(z^\nu)$ as $(\alpha', \beta', \gamma, p', q, P^\nu, \lambda^\nu)$ with fixed index sets $(\alpha', \beta', \gamma)$. By [15, Lemma 3], passing to a further subsequence, $P^\nu \to P'$ for some $P' \in \mathcal{O}^n(G(\bar{z}))$. Eigenvalue continuity gives $\beta' \subseteq \beta$.

Since the W-SRCQ fails at z^ν, there exists $0 \neq W^\nu \in \mathbb{S}^n$ such that

$$\langle W^\nu,\, g'(x^\nu)u\rangle = 0 \quad \forall\, u \in \mathbb{X} \quad \text{and} \quad \langle W^\nu,\, P^\nu B(P^\nu)^\top\rangle = 0 \quad \forall\, B \in \mathbb{S}^n :\, B_{\beta'\gamma} = 0,\, B_{\gamma\gamma} = 0.$$

Normalizing $\|W^\nu\| = 1$ and passing to a further subsequence $W^\nu \to W \neq 0$, letting $\nu \to \infty$ yields

$$\langle W,\, g'(\bar{x})u\rangle = 0 \quad \forall\, u \in \mathbb{X} \quad \text{and} \quad g'(\bar{x})\mathbb{X} + \{P' B(P')^\top \mid B_{\beta'\gamma} = 0,\, B_{\gamma\gamma} = 0\} \neq \mathbb{S}^n.$$

Since $P' \in \mathcal{O}^n(G(\bar{z}))$ and $\beta' \subseteq \beta$, this contradicts the W-SRCQ at $\bar{z}$. □

Constraint nondegeneracy enjoys an unconditional openness property with no restriction on the inertia indices.

Corollary 3 *Let $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{S}^n$. If the constraint nondegeneracy holds at $\bar{z}$, then there exists a neighborhood $\mathcal{U}$ of $\bar{z}$ such that the constraint nondegeneracy holds at every $z \in \mathcal{U}$.*

Proof Let $(\alpha, \beta, \gamma, \bar{p}, \bar{q}, P, \lambda)$ be an IED of $G(\bar{z})$. Suppose for contradiction that there exists a sequence $z^\nu = (x^\nu, y^\nu) \to \bar{z}$ at which constraint nondegeneracy fails. Passing to a subsequence and relabeling the IEDs if necessary, we write an IED of $G(z^\nu)$ as $(\alpha', \beta', \gamma', p', q', P^\nu, \lambda^\nu)$ with fixed index sets $(\alpha', \beta', \gamma')$. By [15, Lemma 3], passing to a further subsequence, $P^\nu \to P'$ for some $P' \in \mathcal{O}^n(G(\bar{z}))$. Eigenvalue continuity gives $\alpha' \supseteq \alpha$, $\gamma' \supseteq \gamma$, and $\beta' \subseteq \beta$.

Failure of (16) at z^ν yields $0 \neq W^\nu \in \mathbb{S}^n$ with

$$\langle W^\nu,\, g'(x^\nu)u\rangle = 0 \quad \forall\, u \in \mathbb{X} \quad \text{and} \quad \langle W^\nu,\, P^\nu B (P^\nu)^\top\rangle = 0 \quad \forall\, B :\, B_{\beta'\beta'} = 0,\, B_{\beta'\gamma'} = 0,\, B_{\gamma'\gamma'} = 0.$$

Normalizing and passing to a limit $W^\nu \to W \neq 0$ gives

$$g'(\bar{x})\mathbb{X} + \{P' B (P')^\top \mid B_{\beta'\beta'} = 0,\, B_{\beta'\gamma'} = 0,\, B_{\gamma'\gamma'} = 0\} \neq \mathbb{S}^n,$$

The inclusions $\beta' \subseteq \beta$ and $\gamma' \supseteq \gamma$ show that the subspace in the last display contains the left-hand side of (16) at $\bar{z}$ with eigenbasis P'. This contradicts constraint nondegeneracy at $\bar{z}$. □

The preceding W-SRCQ openness result shows stability under preservation of the negative inertia index. We now show that the classical strong-form condition, constraint nondegeneracy, is precisely the condition that enforces this stability uniformly across all nearby strata.

Theorem 8 *Let $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{S}^n$. The following three statements are equivalent:*

(a) *the constraint nondegeneracy holds at $\bar{z}$;*
(b) *the W-SRCQ holds at every pair (x, y) sufficiently close to $(\bar{x}, \bar{y})$;*
(c) *the W-SRCQ holds at every pair $(\bar{x}, y)$ with y sufficiently close to $\bar{y}$.*

Proof **(a)⇒(b).** By Corollary 3, constraint nondegeneracy holds for all (x, y) sufficiently close to $(\bar{x}, \bar{y})$, which directly implies the W-SRCQ via Definition 20.

(b)⇒(c). Trivial.

(c)⇒(a). Let $(\alpha, \beta, \gamma, p, q, P, \lambda)$ be an IED of $G(\bar{z})$, and define

$$y^t = \bar{y} - P \begin{bmatrix} 0 & 0 & 0 \\ 0 & tI & 0 \\ 0 & 0 & 0 \end{bmatrix} P^\top, \quad t > 0 \text{ small.}$$

By assumption, the W-SRCQ holds at $(\bar{x}, y^t)$. Applying Definition 20 at $(\bar{x}, y^t)$ immediately recovers constraint nondegeneracy at $(\bar{x}, \bar{y})$. □

Example 4 also shows why preservation of the negative inertia index is essential in Proposition 9, and how the failure of constraint nondegeneracy obstructs the local validity in Theorem 8. Indeed, $g'(0) = 0$ and

$$g'(0)\mathbb{R} + \{B \in \mathbb{S}^2 \mid B_{22} = 0\} \neq \mathbb{S}^2,$$

so constraint nondegeneracy fails at $\bar{z}$. For the family $z^t = (0, tE_{22})$ defined in Example 4, the negative inertia index remains zero and the W-SRCQ holds when $t \geq 0$. When $t < 0$, the negative inertia index becomes one and the W-SRCQ reduces to the proper subspace displayed above, so it fails.

The preceding theorem directly characterizes the classical constraint nondegeneracy in Definition 4 in terms of the local validity of the W-SRCQ.

Corollary 4 *Let $\bar{x}$ be a feasible point of (1) and $\bar{y} \in M(\bar{x})$. Then constraint nondegeneracy holds at $\bar{x}$ if and only if the W-SRCQ holds at every pair (x, y) sufficiently close to $(\bar{x}, \bar{y})$. It is also equivalent to the existence of a matrix y strictly complementary to $g(\bar{x})$ such that the W-SRCQ holds at $(\bar{x}, y)$.*

Next we record the stratified closedness of the W-SRCQ: if the W-SRCQ holds throughout a neighborhood intersected with a stratum $\widetilde{\mathcal{M}}_{p,q}$, then it automatically holds on all lower-q strata in that neighborhood.

Proposition 10 *Let $\mathcal{U}$ be an open set in $\mathbb{X} \times \mathbb{S}^n$. If the W-SRCQ holds at every $z \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q}$ for some p, q, then the W-SRCQ holds at every $z \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q'}$ for all $0 \leq q' \leq q$.*

Proof Fix $q' \le q$ and $z = (x, y) \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q'}$ with IED $(\alpha, \beta', \gamma', p, q', P, \lambda)$ of $G(z)$. Define

$$\Delta = P \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} P^\top,$$

where the blocks have dimensions p, $n - p - q$, $q - q'$, q'. For $t > 0$ the point $z^t = (x, y + t\Delta)$ lies in $\widetilde{\mathcal{M}}_{p,q}$, and $z^t \to z$ as $t \downarrow 0$. For small t, $z^t \in \mathcal{U}$, so by hypothesis the W-SRCQ holds at z^t:

$$g'(x)\mathbb{X} + \{PBP^\top \mid B_{\beta\gamma} = 0,\ B_{\gamma\gamma} = 0\} = \mathbb{S}^n,$$

where $\beta = \{p+1, \dots, n-q\}$ and $\gamma = \{n-q+1, \dots, n\}$. Since $\beta \subseteq \beta'$ and $\gamma \supseteq \gamma'$, the subspace appearing above for z^t is contained in the analogous W-SRCQ subspace at z. Hence the span remains $\mathbb{S}^n$ and the W-SRCQ holds at z. □

5.2.2 Second-order conditions

The following proposition shows that the W-SOC is stable under perturbations that preserve the positive inertia index of $G(z)$.

Proposition 11 *Let $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{S}^n$, and let $\bar{p}$ be the positive inertia index of $G(\bar{z})$. If the W-SOC holds at $\bar{z}$, then there exists a neighborhood $\mathcal{U}$ of $\bar{z}$ such that the W-SOC holds at every $z \in \mathcal{U}$ for which $G(z)$ has positive inertia index $\bar{p}$.*

Proof Let $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{p}, \bar{q}, \bar{P}, \bar{\lambda})$ be an IED of $G(\bar{z})$. Suppose for contradiction that the assertion fails. There exists a sequence $z^\nu = (x^\nu, y^\nu) \to \bar{z}$ such that the positive inertia index of $G(z^\nu)$ is $\bar{p}$ and the W-SOC fails at z^ν. Passing to a subsequence and relabeling the IEDs if necessary, we write an IED of $G(z^\nu)$ as $(\bar{\alpha}, \beta', \gamma', \bar{p}, q', P^\nu, \lambda^\nu)$ with fixed index sets $(\bar{\alpha}, \beta', \gamma')$. By [15, Lemma 3], passing to a further subsequence, $P^\nu \to P'$ for some $P' \in \mathcal{O}^n(G(\bar{z}))$. Eigenvalue continuity gives $\gamma' \supseteq \bar{\gamma}$.

The failure of the W-SOC at z^ν yields $d^\nu \in \mathbb{X}$ with $\|d^\nu\| = 1$, $d^\nu \in \operatorname{appl}(z^\nu)$, and

$$\langle d^\nu, \nabla^2_{xx} L(z^\nu) d^\nu \rangle + 2 \sum_{i \in \bar{\alpha}} \sum_{j \in \gamma'} \frac{-\lambda^\nu_j}{\lambda^\nu_i} [(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]^2_{ij} = 0,$$

together with $[(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]_{\beta'\beta'} = 0$, $[(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]_{\beta'\gamma'} = 0$, and $[(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]_{\gamma'\gamma'} = 0$. Passing to a subsequence $d^\nu \to d^\infty$ with $\|d^\infty\| = 1$ and taking $\nu \to \infty$, the three zero-block conditions on d^ν pass to the limit: since $P^\nu \to P'$ and $z^\nu \to \bar{z}$, we get $[(P')^\top (\nabla g(\bar{x})^* d^\infty) P']_{\beta'\beta'} = 0$, $[(P')^\top (\nabla g(\bar{x})^* d^\infty) P']_{\beta'\gamma'} = 0$, and $[(P')^\top (\nabla g(\bar{x})^* d^\infty) P']_{\gamma'\gamma'} = 0$. Because $\bar{\alpha} = \{1, \dots, \bar{p}\}$ is fixed, $\beta' \cup \gamma' = \bar{\beta} \cup \bar{\gamma} = \{\bar{p}+1, \dots, n\}$, so these three conditions together say the entire $\bar{\beta} \cup \bar{\gamma}$ block of $(P')^\top (\nabla g(\bar{x})^* d^\infty) P'$ vanishes, which is exactly $d^\infty \in \operatorname{appl}(\bar{z})$. For the quadratic form, the terms with $j \in \bar{\gamma}$ converge normally to give $\mathcal{Q}_{\bar{z}}(d^\infty)$. The terms with $j \in \gamma' \setminus \bar{\gamma}$ vanish in the limit because $\bar{\lambda}_j = 0$ while the block entries remain bounded. Hence $\mathcal{Q}_{\bar{z}}(d^\infty) = 0$ with $d^\infty \in \operatorname{appl}(\bar{z}) \setminus \{0\}$, contradicting the W-SOC at $\bar{z}$. □

The preceding W-SOC openness result shows stability under preservation of the positive inertia index. We now show that the classical strong-form condition, the S-SOSC, is precisely the condition that enforces this stability uniformly across all nearby strata. The dual picture requires a notion of *uniform* validity of the W-SOC on a neighborhood.

Definition 29 Let $\mathcal{Z} \subseteq \mathbb{X} \times \mathbb{S}^n$. The W-SOC is said to hold *uniformly on $\mathcal{Z}$* if there exists a constant $c > 0$ such that, for every $z = (x, y) \in \mathcal{Z}$ with IED $(\alpha, \beta, \gamma, p, q, P, \lambda)$ of $G(z)$,

$$\mathcal{Q}_z(v_x) \ge c\|v_x\|^2 \quad \forall\, v_x \in \operatorname{appl}(z). \tag{38}$$

The reason why uniformity strengthens the pointwise non-vanishing condition to a strict positive lower bound is discussed in the subsequent remark.

Remark 21 The W-SOC requires only non-vanishing and does not prescribe a sign. Under the W-SONC, however, this non-vanishing condition is equivalent to strict positivity. Requiring a common positive modulus on a neighborhood is therefore the corresponding uniform condition. This could be termed the uniform W-SOSC, but since the W-SOSC would not be used separately, we avoid introducing additional notation and retain the term *uniform W-SOC*.

Theorem 9 *Let $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{X} \times \mathbb{S}^n$. The following three statements are equivalent:*

(a) *the S-SOSC holds at $\bar{z}$;*
(b) *the W-SOC holds uniformly on some neighborhood of $\bar{z}$;*
(c) *the W-SOC holds uniformly on $\{(\bar{x}, y) \mid y \in \mathcal{V}\}$ for some neighborhood $\mathcal{V}$ of $\bar{y}$.*

Proof Let $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{p}, \bar{q}, \bar{P}, \bar{\lambda})$ be an IED of $G(\bar{z})$.

(a)⇒(b). Suppose that (a) holds but (b) fails. Then there exist $z^\nu = (x^\nu, y^\nu) \to \bar{z}$ and $c^\nu \downarrow 0$ such that the inequality in (38) with $c = c^\nu$ fails at z^ν. Passing to a subsequence and relabeling the IEDs if necessary, we write an IED of $G(z^\nu)$ as $(\alpha, \beta, \gamma, p, q, P^\nu, \lambda^\nu)$ with fixed index sets (α, β, γ). The conditions at $\bar{z}$ are independent of the reference IED. Hence, by [15, Lemma 3], after passing to a further subsequence we may choose $\bar{P}$ so that $P^\nu \to \bar{P} \in \mathcal{O}^n(G(\bar{z}))$. Eigenvalue continuity gives $\alpha \supseteq \bar{\alpha}$, $\beta \subseteq \bar{\beta}$, and $\gamma \supseteq \bar{\gamma}$.

The failure of (38) at z^ν yields $d^\nu \in \mathrm{appl}(z^\nu)$ with $\|d^\nu\| = 1$ and

$$\langle d^\nu, \nabla^2_{xx} L(z^\nu) d^\nu \rangle + 2 \sum_{i\in\alpha} \sum_{j\in\gamma} \frac{-\lambda^\nu_j}{\lambda^\nu_i} [(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]^2_{ij} < c^\nu. \tag{39}$$

Since the Hessian term in (39) is bounded and all spectral terms are nonnegative, the spectral sum is uniformly bounded above. For $i \in \alpha \setminus \bar{\alpha}$ and $j \in \bar{\gamma}$ we have $\lambda^\nu_i \to 0$ and $\lambda^\nu_j \to \bar{\lambda}_j < 0$, so $-\lambda^\nu_j / \lambda^\nu_i \to +\infty$. Hence

$$[(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]_{ij} \to 0 \quad \forall\, i \in \alpha \setminus \bar{\alpha},\ j \in \bar{\gamma}.$$

Passing to a subsequence, suppose that $d^\nu \to d^\infty$ with $\|d^\infty\| = 1$. We first show that $d^\infty \in \mathrm{app}(\bar{z})$. The condition $d^\nu \in \mathrm{appl}(z^\nu)$ gives the vanishing of the $\beta\gamma$ and $\gamma\gamma$ blocks. Since $\gamma \supseteq \bar{\gamma}$, the $\bar{\gamma}\bar{\gamma}$ block condition follows in the limit. For the $\bar{\beta}\bar{\gamma}$ block, the row indices in $\bar{\beta}$ split into $\alpha \setminus \bar{\alpha}$, β, and $\gamma \setminus \bar{\gamma}$. The parts with row indices in β and $\gamma \setminus \bar{\gamma}$ vanish by the $\beta\gamma$ and $\gamma\gamma$ block conditions for d^ν, respectively. The part with row indices in $\alpha \setminus \bar{\alpha}$ vanishes by the preceding blow-up argument. Thus the two block conditions in (21) hold for d^∞ at $\bar{z}$.

We now pass to the second-order expression. Since $\bar{\alpha} \subseteq \alpha$ and $\bar{\gamma} \subseteq \gamma$, and since all omitted spectral terms are nonnegative, (39) implies

$$\langle d^\nu, \nabla^2_{xx} L(z^\nu) d^\nu \rangle + 2 \sum_{i\in\bar{\alpha}} \sum_{j\in\bar{\gamma}} \frac{-\lambda^\nu_j}{\lambda^\nu_i} [(P^\nu)^\top (\nabla g(x^\nu)^* d^\nu) P^\nu]^2_{ij} < c^\nu.$$

Taking $\nu \to \infty$ yields

$$\langle d^\infty, \nabla^2_{xx} L(\bar{z}) d^\infty \rangle + 2 \sum_{i\in\bar{\alpha}} \sum_{j\in\bar{\gamma}} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} [\bar{P}^\top (\nabla g(\bar{x})^* d^\infty) \bar{P}]^2_{ij} \leq 0.$$

This contradicts the S-SOSC at $\bar{z}$.

(b)⇒(c). Trivial.

(c)⇒(a). Define $\varDelta = \bar{P}\, \mathrm{Diag}(0_{\bar{p}}, I_{\bar{\beta}}, 0_{\bar{q}}) \bar{P}^\top$, where $I_{\bar{\beta}}$ is the identity of order $n - \bar{p} - \bar{q}$. For small $t > 0$, $z^t = (\bar{x}, \bar{y} + t\varDelta) \to \bar{z}$ and $G(z^t)$ has the same $\bar{P}$-eigenbasis with $\bar{\alpha}$- and $\bar{\gamma}$-eigenvalues unchanged and $\bar{\beta}$-eigenvalues shifted to $t > 0$. By assumption there exists $c > 0$ such that

$$\langle v_x, \nabla^2_{xx} L(z^t) v_x \rangle + 2 \sum_{i\in\bar{\alpha}} \sum_{j\in\bar{\gamma}} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} [\bar{P}^\top (\nabla g(\bar{x})^* v_x) \bar{P}]^2_{ij} + 2 \sum_{i=\bar{p}+1}^{n-\bar{q}} \sum_{j\in\bar{\gamma}} \frac{-\bar{\lambda}_j}{t} [\bar{P}^\top (\nabla g(\bar{x})^* v_x) \bar{P}]^2_{ij} \geq c\|v_x\|^2 \tag{40}$$

for all $v_x \in \mathrm{appl}(z^t) = \{v_x \mid [\bar{P}^\top (\nabla g(\bar{x})^* v_x) \bar{P}]_{\bar{\gamma}\bar{\gamma}} = 0\}$. If the S-SOSC fails at $\bar{z}$, there exists $d \in \mathrm{app}(\bar{z}) \setminus \{0\}$ with $\mathcal{Q}_{\bar{z}}(d) \leq 0$. Because $d \in \mathrm{app}(\bar{z})$ implies $d \in \mathrm{appl}(z^t)$ and the cross terms involving $1/t$ vanish on $\mathrm{app}(\bar{z})$, substituting $v_x = d$ into (40) and letting $t \downarrow 0$ gives $\mathcal{Q}_{\bar{z}}(d) \geq c\|d\|^2 > 0$, a contradiction. □

The following example shows that the W-SOC may hold at a reference pair even though the S-SOSC fails there and, consequently, the W-SOC does not hold uniformly on any neighborhood of that pair.

Example 5 Consider

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \frac{1}{2} x_1^2 \qquad \text{subject to} \qquad g(x_1, x_2) = \begin{bmatrix} 1 & 0 \\ 0 & x_2 \end{bmatrix} \succeq 0,$$

and the KKT pair $\bar{z} = ((0,0), 0)$, for which $G(\bar{z}) = \mathrm{Diag}(1, 0)$, with $p = 1$, $q = 0$, $\beta = \{2\}$, and $\gamma = \emptyset$. Directly from (22), $\mathrm{appl}(\bar{z}) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_2 = 0\}$, and $\mathcal{Q}_{\bar{z}}(v_1, v_2) = v_1^2$. Hence the W-SOC holds at $\bar{z}$. On the other hand, $\mathrm{app}(\bar{z}) = \mathbb{R}^2$ and $\mathcal{Q}_{\bar{z}}(0, 1) = 0$, so the S-SOSC fails at $\bar{z}$. For $t > 0$, let $z^t = ((0, t), 0)$. Then $G(z^t) = \mathrm{Diag}(1, t)$ is positive definite, so $\mathrm{appl}(z^t) = \mathbb{R}^2$ and $\mathcal{Q}_{z^t}(v_1, v_2) = v_1^2$. Thus the W-SOC fails at every z^t in the direction $(0, 1)$, while $z^t \to \bar{z}$ as $t \downarrow 0$.

Corollary 5 *Let $\bar z = (\bar x, \bar y) \in \mathbb{X} \times \mathbb{S}^n$. The following uniform S-SOSC is equivalent to the three statements in Theorem 9: there exist a neighborhood $\mathcal{U}$ of $\bar z$ and a constant $c > 0$ such that*

$$\mathcal{Q}_z(v_x) \geq c\|v_x\|^2 \quad \forall\, v_x \in \operatorname{app}(z), \quad \forall\, z \in \mathcal{U}.$$

Proof By Theorem 9, it suffices to prove that the S-SOSC at $\bar z$ is equivalent to the displayed uniform S-SOSC. The uniform condition immediately implies the S-SOSC by taking $z = \bar z$. Conversely, suppose that the S-SOSC holds at $\bar z$ and that the uniform condition above fails. Then there exist $z^\nu \to \bar z$, $c^\nu \downarrow 0$, and $d^\nu \in \operatorname{app}(z^\nu)$ with $\|d^\nu\| = 1$ and $\mathcal{Q}_{z^\nu}(d^\nu) < c^\nu$. Repeating the proof of Theorem 9(a)⇒(b), with $\operatorname{app}(z^\nu)$ in place of $\operatorname{appl}(z^\nu)$, gives a subsequence $d^\nu \to d^\infty$ such that $d^\infty \in \operatorname{app}(\bar z)$, $\|d^\infty\| = 1$, and $\mathcal{Q}_{\bar z}(d^\infty) \leq 0$, contradicting the S-SOSC at $\bar z$. □

The following stratified closedness result is the dual analogue of Proposition 10: uniform W-SOC on a stratum extends to all lower-p strata.

Proposition 12 *Let $\mathcal{U}$ be an open set in $\mathbb{X} \times \mathbb{S}^n$. If the W-SOC holds uniformly on $\mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q}$ for some p, q, then it holds uniformly on $\mathcal{U} \cap \widetilde{\mathcal{M}}_{p',q}$ for every $0 \leq p' \leq p$.*

Proof Let $c > 0$ be a uniform constant for the W-SOC on $\mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q}$. Fix $p' \leq p$ and $z = (x, y) \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p',q}$ with IED $(\alpha', \beta', \gamma, p', q, P, \lambda)$ of $G(z)$. Define

$$\Delta = P \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} P^\top,$$

where the blocks have dimensions p', $p - p'$, $n - p - q$, and q. For all sufficiently small $t > 0$, $z^t = (x, y + t\Delta) \in \widetilde{\mathcal{M}}_{p,q} \cap \mathcal{U}$ with IED $(\alpha, \beta, \gamma, p, q, P, \lambda + t\operatorname{diag}(P^\top \Delta P))$, where $\alpha \supseteq \alpha'$ and $\beta \subseteq \beta'$. By hypothesis,

$$\langle v_x, \nabla^2_{xx} L(z^t) v_x\rangle + 2\sum_{i\in\alpha'}\sum_{j\in\gamma} \frac{-\lambda_j}{\lambda_i}[P^\top(\nabla g(x)^* v_x)P]_{ij}^2 + 2\sum_{i\in\alpha\setminus\alpha'}\sum_{j\in\gamma} \frac{-\lambda_j}{t}[P^\top(\nabla g(x)^* v_x)P]_{ij}^2 \geq c\|v_x\|^2$$

for all $v_x \in \operatorname{appl}(z^t)$. Since $\beta \subseteq \beta'$, $\operatorname{appl}(z) \subseteq \operatorname{appl}(z^t)$. If the bound with the constant c failed at z, there would exist nonzero $d \in \operatorname{appl}(z)$ such that $\mathcal{Q}_z(d) < c\|d\|^2$. But $d \in \operatorname{appl}(z^t)$, and the cross terms vanish on $\operatorname{appl}(z)$. Substituting d and letting $t \downarrow 0$ yields a contradiction. Since $p' \leq p$ and z were arbitrary, the same constant c proves the result. □

5.2.3 Local uniform validity of the weak second order condition

The preceding subsubsections studied the perturbation behavior of the constraint qualification and of the second order condition. This subsubsection establishes a stronger local consequence: under an additional uniform injectivity assumption, the SONC at the reference pair propagates to the local uniform validity of the W-SOC. The following assumption is central to this subsubsection. Its local uniform injectivity condition is closely related to the Aubin property of F^{-1}.

Assumption 10 *The SONC holds at $\bar z = (\bar x, \bar y) \in \mathbb{X} \times \mathbb{S}^n$, and there exist $\kappa > 0$ and a neighborhood $\mathcal{U}$ of $\bar z$ such that, for every (p, q),*

$$\|dF_z v\| \;\geq\; \kappa^{-1}\|v\| \qquad \forall\, z \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q}, \quad \forall\, v \in T_z\widetilde{\mathcal{M}}_{p,q}. \tag{41}$$

Throughout this subsubsection, $(\bar\alpha, \bar\beta, \bar\gamma, \bar p, \bar q, \bar P, \bar\lambda)$ is a fixed IED of $G(\bar z)$ and $k := |\bar\beta|$. After shrinking $\mathcal{U}$ if necessary, set

$$\kappa' := \kappa\,\big(1 + \sup_{(x,y)\in\mathcal{U}} \|\nabla g(x)\|\big) < \infty. \tag{42}$$

For $M \in \mathbb{S}^k$ define the *multiplier slice*

$$z_M := \big(\bar x,\; \bar y + \bar P_{\bar\beta} M \bar P_{\bar\beta}^\top\big),$$

where $\bar P_{\bar\beta}$ collects the $\bar\beta$-columns of $\bar P$. Then $G(z_M) = G(\bar z) + \bar P_{\bar\beta} M \bar P_{\bar\beta}^\top$ admits an IED whose $\bar\alpha$- and $\bar\gamma$-data agree with those of $G(\bar z)$ and whose $\bar\beta$-part carries the eigenvalues of M, with eigenvectors rotated inside the $\bar\beta$ block by an eigenframe of M; in particular, z_M lies on a stratum with $p + q = n$ iff M is nonsingular. Fix $t_1 > 0$ such that $z_M \in \mathcal{U}$ whenever $\|M\| \leq 2t_1$. Combining (41) with Lemma 9 gives

$$\sigma_{\min}\big(\mathcal{Q}_z|_{\operatorname{appl}(z)}\big) \;\geq\; 1/\kappa' \qquad \text{for every } z \in \mathcal{U} \cap \widetilde{\mathcal{M}}_{p,q} \text{ with } p + q = n; \tag{43}$$

in particular, this applies to z_M for every nonsingular $M \in \mathbb{S}^k$ with $\|M\| \le 2t_1$.

We first record the immediate regularity consequences of the uniform injectivity assumption, which will supply the base conditions for the subsequent inertia argument.

Proposition 13 *Under Assumption 10:*

(a) *the W-SRCQ holds at every $z \in \mathcal{U}$;*
(b) *the constraint nondegeneracy, and thus the SRCQ, holds at every $z \in \mathcal{U}$;*
(c) *the W-SOC and the W-SONC hold at $\bar z$.*

Proof (a) By (41), dF_z is injective for every $z \in \mathcal{U}$. Hence Theorem 6(b) yields the W-SRCQ at every $z \in \mathcal{U}$.

(b) Since $\mathcal{U}$ is open, statement (a) provides the W-SRCQ on an ambient neighborhood of each $z \in \mathcal{U}$. Hence Theorem 8 gives the constraint nondegeneracy at every $z \in \mathcal{U}$, which clearly implies the SRCQ.

(c) Theorem 6(c) gives the W-SOC at $\bar z$ and the SONC gives the W-SONC. □

We next isolate the linear-algebraic device used to pass to singular limits. It shows that, in the presence of a diverging positive semidefinite term, the limiting negative inertia is determined by the compression to the kernel of that term. For a self-adjoint operator or quadratic form A, $n_-(A)$ denotes the number of negative eigenvalues of A, counted with multiplicity. On the zero-dimensional space, we use the conventions $n_-(A) = 0$ and $\sigma_{\min}(A) = +\infty$.

Lemma 10 *Let $V = W \oplus W'$ be an orthogonal decomposition of a finite-dimensional inner-product space, and for $s \in (0, s_0]$ let $A_s = Q_s + s^{-1}B$ be self-adjoint operators on V with $B \succeq 0$, $\ker B = W$, $Q_s \to Q_0$ as $s \downarrow 0$, and $\sigma_{\min}(A_s) \ge c > 0$ for all s. Then the compression $Q_0|_W$ is nondegenerate with $\sigma_{\min}(Q_0|_W) \ge c$, and for all small s,*

$$n_-\big(A_s\big) = n_-\big(Q_0|_W\big).$$

Proof If $W' = \{0\}$, then $B = 0$ and $A_s = Q_s \to Q_0$, so the conclusion follows directly from the uniform lower bound on $\sigma_{\min}(A_s)$. Hence suppose $W' \ne \{0\}$. Since $B \succeq 0$ and $\ker B = W$, the compression $B|_{W'}$ is positive definite, so for all small s, $A_s|_{W'} \succeq \big(s^{-1}\sigma_{\min}(B|_{W'}) - \|Q_s\|\big)I \succ 0$. Eliminating the W'-block by block Gaussian elimination transforms A_s, by congruence, into the block-diagonal operator formed by $A_s|_{W'}$ and its Schur complement S_s on W. Congruence preserves inertia, and the positive definite block $A_s|_{W'}$ has no negative eigenvalues. Hence $n_-(A_s) = n_-(S_s)$. For a unit $w \in W$, the invertibility of the W'-block yields $u \in W'$ with $A_s(w+u) = (S_s w, 0) \in W \oplus W'$, so $\|S_s w\| \ge c\|w+u\| \ge c$, and therefore $\sigma_{\min}(S_s) \ge c$. Since B is self-adjoint with $\ker B = W$, its range is W', so the W-diagonal and off-diagonal blocks of A_s are those of Q_s, while the inverse of the W'-block is $O(s)$. Hence $S_s = Q_s|_W + O(s) \to Q_0|_W$. Two nondegenerate operators with a common modulus c at distance $\to 0$ share their inertia, so $n_-(S_s) = n_-(Q_0|_W)$ for small s, and $\sigma_{\min} \ge c$ passes to the limit. □

For orthonormal vectors $u_1, \ldots, u_\ell \in \mathbb{R}^k$ $(0 \le \ell \le k)$ set

$$A(u_1, \ldots, u_\ell) := \big\{v \in \mathrm{app}(\bar z) \mid \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta}\, u_j = 0,\ j = 1, \ldots, \ell\big\}.$$

Thus $A(\,) = \mathrm{app}(\bar z)$, while $A(u_1, \ldots, u_k) = \mathrm{appl}(\bar z)$ for any orthonormal basis $u_1, \ldots, u_k$. For $\ell \le k-1$ and a unit vector $u_0 \perp \mathrm{span}\{u_1, \ldots, u_\ell\}$, set

$$E_{u_0} := A(u_1, \ldots, u_\ell) \cap \big\{v \mid u_0^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta}\, u_0 = 0\big\},$$

which imposes only the scalar constraint in the direction u_0, so that $A(u_1, \ldots, u_\ell) \supseteq E_{u_0} \supseteq A(u_1, \ldots, u_\ell, u_0)$.

These definitions are made for the proof of Theorem 11, which descends from $\mathrm{appl}(\bar z)$, where the positivity of $\mathcal{Q}_{\bar z}$ is known, to $\mathrm{app}(\bar z)$, along the filtration $A(u_1, \ldots, u_\ell)$, one direction at a time. The intermediate space E_{u_0} is needed because a negative direction w of $\mathcal{Q}_{\bar z}$ on $A(u_1, \ldots, u_\ell)$ yields, in general, only a unit u_0 with the scalar property $u_0^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* w\big) \bar P_{\bar\beta} u_0 = 0$, that is, membership $w \in E_{u_0}$ rather than $w \in A(u_1, \ldots, u_\ell, u_0)$. The next proposition uses nonsingular multiplier slices and the preceding inertia lemma to show that, for the purpose of counting negative directions, the two spaces are interchangeable. This is the induction step needed in the proof of Theorem 11.

Proposition 14 *Suppose that Assumption 10 holds. For every $0 \le \ell \le k-1$, every orthonormal $u_1, \ldots, u_\ell \in \mathbb{R}^k$, and every unit $u_0 \perp \mathrm{span}\{u_1, \ldots, u_\ell\}$:*

(i) $n_-\big(\mathcal{Q}_{\bar z}|_{E_{u_0}}\big) = n_-\big(\mathcal{Q}_{\bar z}|_{A(u_1,\ldots,u_\ell,u_0)}\big)$;
(ii) *every neighborhood of 0 in $\mathbb{S}^k$ contains a nonsingular M with exactly $k-\ell-1$ positive eigenvalues such that $n_-\big(\mathcal{Q}_{z_M}|_{\mathrm{appl}(z_M)}\big) = n_-\big(\mathcal{Q}_{\bar z}|_{A(u_1,\ldots,u_\ell,u_0)}\big)$, and a $t > 0$ with $n_-\big(\mathcal{Q}_{z_{tI_k}}|_{\mathrm{appl}(z_{tI_k})}\big) = n_-\big(\mathcal{Q}_{\bar z}|_{A(\,)}\big)$.*

Proof Choose an orthonormal basis $w_1, \ldots, w_{k-\ell-1}$ of

$$\operatorname{span}\{u_1, \ldots, u_\ell, u_0\}^\perp.$$

If $\ell = k-1$, then $E_{u_0} = A(u_1, \ldots, u_\ell, u_0)$, so (i) is immediate and all sums below are empty.

Choose $\delta > 0$ sufficiently small and set $M_0 := -\delta \sum_{j \le \ell} u_j u_j^\top$. For $\omega = (\omega_1, \ldots, \omega_{k-\ell-1}) \in (0,\infty)^{k-\ell-1}$ and $s \downarrow 0$, define

$$M_s := M_0 - s u_0 u_0^\top + \sum_{i \le k-\ell-1} (s/\omega_i) w_i w_i^\top.$$

Then M_s is nonsingular and has exactly $k-\ell-1$ positive eigenvalues. Taking s and δ small ensures $\|M_s\| \le 2t_1$, so (43) applies to z_{M_s}. For $s \le \delta^2$, within the $\bar\beta$-block, the positive eigenpairs of M_s are $(w_i, s/\omega_i)$, while its negative eigenpairs are $(u_0, -s)$ and $(u_j, -\delta)$, $j = 1, \ldots, \ell$. Accordingly, in the IED of $G(z_{M_s})$, the index set $\bar\beta$ splits into the three groups $\bar\beta_w$, $\bar\beta_{u_0}$, and $\bar\beta_u$, indexing the eigenvectors $\bar P_{\bar\beta} w_i$, $\bar P_{\bar\beta} u_0$, and $\bar P_{\bar\beta} u_j$, respectively. Substituting these eigenpairs into (19), for every $v \in \operatorname{appl}(z_{M_s})$ we obtain

$$\begin{aligned}
\mathcal{Q}_{z_{M_s}}(v) = \ & \langle v, \nabla^2_{xx} L(z_{M_s}) v\rangle + 2 \sum_{i\in\bar\alpha}\sum_{j\in\bar\gamma} \frac{-\bar\lambda_j}{\bar\lambda_i} \left[\bar P^\top \big(\nabla g(\bar x)^* v\big) \bar P\right]_{ij}^2 \\
& + 2s \sum_{i\in\bar\alpha} \frac{1}{\bar\lambda_i} \left[\bar P^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_0\right]_i^2 \\
& + 2\delta \sum_{i\in\bar\alpha}\sum_{j\le\ell} \frac{1}{\bar\lambda_i} \left[\bar P^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_j\right]_i^2 \\
& + \frac{2}{s} \sum_{i\le k-\ell-1} \omega_i \sum_{j\in\bar\gamma} (-\bar\lambda_j) \left[w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\gamma}\right]_j^2 \\
& + \frac{2\delta}{s} \sum_{i\le k-\ell-1}\sum_{j\le\ell} \omega_i \left(w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_j\right)^2 \\
& + 2 \sum_{i\le k-\ell-1} \omega_i \left(w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_0\right)^2.
\end{aligned}$$

Since the negative eigenspace of $G(z_{M_s})$ is independent of both s and ω, (22) shows that $\operatorname{appl}(z_{M_s})$ is also independent of s and ω. The nonnegative terms in the preceding expansion of $\mathcal{Q}_{z_{M_s}}$ indexed by $(\bar\beta_w, \bar\gamma)$ and $(\bar\beta_w, \bar\beta_u)$ have coefficients proportional to s^{-1}. For $v \in \operatorname{appl}(z_{M_s})$, they vanish simultaneously exactly when the following equalities hold for all $i \le k-\ell-1$ and $j \le \ell$:

$$w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\gamma} = 0, \qquad w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_j = 0.$$

Together with the constraints defining $\operatorname{appl}(z_{M_s})$, these are precisely the constraints defining E_{u_0}. Define the quadratic form $R_\delta : \mathbb{X} \to \mathbb{R}$ by

$$\begin{aligned}
R_\delta(v) := \ & \left\langle v, \left[\nabla^2_{xx} L(z_{M_0}) - \nabla^2_{xx} L(\bar z)\right] v\right\rangle \\
& + 2\delta \sum_{i\in\bar\alpha}\sum_{j\le\ell} \frac{1}{\bar\lambda_i} \left[\bar P^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_j\right]_i^2.
\end{aligned}$$

Since $z_{M_0} - \bar z = O(\delta)$, there is a constant $C_0 > 0$ such that $\|R_\delta\| \le C_0 \delta$. Then for $v \in E_{u_0}$, $\mathcal{Q}_{z_{M_s}}(v)$ converges, as $s \downarrow 0$, to $q_\omega(v)$, where the quadratic form $q_\omega : E_{u_0} \to \mathbb{R}$ is defined by

$$q_\omega(v) := \mathcal{Q}_{\bar z}(v) + R_\delta(v) + 2 \sum_{i\le k-\ell-1} \omega_i \left(w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_0\right)^2.$$

To apply Lemma 10, write the sum of the $(\bar\beta_w, \bar\gamma)$ and $(\bar\beta_w, \bar\beta_u)$ terms as $s^{-1}B$. For fixed ω and δ, the quadratic form B is independent of s and positive semidefinite. The equalities above show that its kernel in $\operatorname{appl}(z_{M_s})$ is E_{u_0}. The remaining terms have coefficients converging as $s \downarrow 0$, so they converge as quadratic forms on the whole fixed space $\operatorname{appl}(z_{M_s})$, and the compression of their limit to E_{u_0} is q_ω. Moreover, (43) gives

$$\sigma_{\min}\big(\mathcal{Q}_{z_{M_s}}|_{\operatorname{appl}(z_{M_s})}\big) \ge 1/\kappa'.$$

Since the modulus bound in (43) is independent of M_s, for every ω the lemma shows that q_ω is nondegenerate with the same bound $\sigma_{\min}(q_\omega) \ge 1/\kappa'$ and, for all sufficiently small s,

$$n_-(q_\omega) = n_-\big(\mathcal{Q}_{z_{M_s}}|_{\operatorname{appl}(z_{M_s})}\big). \tag{44}$$

Suppose first that $\ell < k-1$. The parameter set $(0,\infty)^{k-\ell-1}$ is connected, and the uniformly nondegenerate forms q_ω depend continuously on ω. Hence $n_-(q_\omega)$ is constant in ω. Along the ray $\omega_1 = \cdots = \omega_{k-\ell-1} = t$, $t > 0$, letting $t \downarrow 0$ gives the compression of $\mathcal{Q}_{\bar z} + R_\delta$ to E_{u_0}. Letting $t \to \infty$ and applying Lemma 10 once more gives its compression to

$$\big\{v \in E_{u_0} \mid w_i^\top \bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* v\big) \bar P_{\bar\beta} u_0 = 0, i = 1, \dots, k-\ell-1\big\} = A(u_1, \dots, u_\ell, u_0).$$

Thus these two compressions have the same n_-. They moreover retain the modulus $1/\kappa'$. At $t \downarrow 0$ the forms converge on the fixed space E_{u_0}, so their eigenvalues, lying outside $(-1/\kappa', 1/\kappa')$, keep this property in the limit. At $t \to \infty$, Lemma 10 delivers the bound $\sigma_{\min} \ge 1/\kappa'$ directly. The same conclusion is immediate when $\ell = k-1$. Now choose δ with $C_0\delta < 1/(2\kappa')$. Since $\|R_\delta\| \le C_0\delta$, removing R_δ moves the eigenvalues of both compressions by at most $C_0\delta$, so no eigenvalue reaches zero and neither inertia changes. This proves (i).

For any fixed ω, equality (44) identifies $n_-\big(\mathcal{Q}_{z_{M_s}}|_{\mathrm{appl}(z_{M_s})}\big)$, for all small s, with the common value just obtained. Since δ may be chosen arbitrarily small and then s may be chosen arbitrarily small, the resulting nonsingular M_s, with exactly $k-\ell-1$ positive eigenvalues, can be placed in any prescribed neighborhood of 0. This proves the first assertion in (ii).

Finally, consider the slices z_{tI_k} with $t \downarrow 0$. The space $\mathrm{appl}(z_{tI_k})$ does not depend on t, and in the expansion of $\mathcal{Q}_{z_{tI_k}}$ according to (19), the $(\bar\beta,\bar\gamma)$-terms have coefficients $-\bar\lambda_j/t$. Apply Lemma 10 with B now the sum of these terms, whose kernel in $\mathrm{appl}(z_{tI_k})$ is $A(\,) = \mathrm{app}(\bar z)$, and with the remaining terms converging to $\mathcal{Q}_{\bar z}$. Together with (43), this yields

$$n_-\big(\mathcal{Q}_{z_{tI_k}}|_{\mathrm{appl}(z_{tI_k})}\big) = n_-\big(\mathcal{Q}_{\bar z}|_{A(\,)}\big)$$

for all sufficiently small $t > 0$, completing (ii). □

Theorem 11 *Under Assumption 10, there exists a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of $\bar z$ such that*

$$\mathcal{Q}_z(v_x) \ge \frac{1}{\kappa'}\|v_x\|^2 \qquad \forall\, z \in \mathcal{U}', \quad \forall\, v_x \in \mathrm{appl}(z),$$

where κ' is defined in (42). *Consequently, the W-SOC holds uniformly on $\mathcal{U}'$ in the sense of Definition 29; in particular, the W-SONC holds at every $z \in \mathcal{U}'$.*

Proof If $k = 0$, then $\bar\beta = \emptyset$ and $\mathrm{appl}(\bar z) = \mathrm{app}(\bar z)$. By Proposition 13(c), the W-SOC and the W-SONC hold at $\bar z$, so the S-SOSC holds there. After shrinking the neighborhood, Corollary 5 gives positivity of $\mathcal{Q}_z|_{\mathrm{appl}(z)}$, while (43) gives the modulus $1/\kappa'$. Hence the desired bound follows. We may therefore suppose that $k \ge 1$.

We first prove, by downward induction on $\ell = k, k-1, \dots, 0$:

$$(P_\ell): \qquad n_-\big(\mathcal{Q}_{\bar z}|_{A(u_1,\dots,u_\ell)}\big) = 0 \quad \text{for every orthonormal } u_1, \dots, u_\ell.$$

For $\ell = k$, $A(u_1, \dots, u_k) = \mathrm{appl}(\bar z)$. By the W-SONC at $\bar z$ (Proposition 13(c)), $\mathcal{Q}_{\bar z} \ge 0$ on this space, so (P_k) holds. Let $\ell < k$, assume $(P_{\ell+1})$, and fix $A := A(u_1, \dots, u_\ell)$. Suppose, for contradiction, that $n_-(\mathcal{Q}_{\bar z}|_A) \ge 1$. Then there is a unit eigenvector $w \in A$ of $\mathcal{Q}_{\bar z}|_A$ with $\mathcal{Q}_{\bar z}(w) < 0$. Since $w \in A$, the symmetric matrix $\bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* w\big) \bar P_{\bar\beta}$ contains $u_1, \dots, u_\ell$ in its kernel. Let $\widehat B$ be its compression to $K = \mathrm{span}\{u_j\}^\perp$.

- If $\widehat B \succeq 0$ or $\widehat B \preceq 0$, then $\bar P_{\bar\beta}^\top \big(\nabla g(\bar x)^* w\big) \bar P_{\bar\beta} \succeq 0$ or $\preceq 0$. Since $w \in \mathrm{app}(\bar z)$, the $\bar\beta\bar\gamma$- and $\bar\gamma\bar\gamma$-blocks of $\bar P^\top (\nabla g(\bar x)^* w) \bar P$ vanish, so w or $-w$ lies in $\mathcal{C}(\bar z)$ by (20). By the SONC and the evenness of $\mathcal{Q}_{\bar z}$, $\mathcal{Q}_{\bar z}(w) \ge 0$, a contradiction.
- Otherwise $\widehat B$ has eigenvalues of both signs, and hence admits a unit vector $u_0 \in K$ with $u_0^\top \widehat B u_0 = 0$. Then $w \in E_{u_0}$ and $\mathcal{Q}_{\bar z}(w) < 0$, so $n_-\big(\mathcal{Q}_{\bar z}|_{E_{u_0}}\big) \ge 1$. By Proposition 14(i), $n_-\big(\mathcal{Q}_{\bar z}|_{A(u_1,\dots,u_\ell,u_0)}\big) \ge 1$, which contradicts $(P_{\ell+1})$.

Hence $n_-(\mathcal{Q}_{\bar z}|_A) = 0$, completing the induction.

Combining (P_ℓ) with Proposition 14(ii), for every $0 \le a \le k$ and every $\rho > 0$ there is a nonsingular $M^{(a)} \in \mathbb{S}^k$ with exactly a positive eigenvalues, $\|M^{(a)}\| \le \rho$, and $n_-\big(\mathcal{Q}_{z_{M^{(a)}}}|_{\mathrm{appl}(z_{M^{(a)}})}\big) = 0$. For $a \le k-1$, apply Proposition 14(ii) with $\ell = k-a-1$. For $a = k$, take $M^{(k)} = tI_k$ with t small.

Now let $\mathcal{V}$ be a neighborhood of $G(\bar z)$ as in Lemma 4 for $\bar A := G(\bar z)$, contained in $\mathcal{B}(G(\bar z), t_1)$, and choose $r_1 > 0$ small enough that

$$\mathcal{U}' := \big\{z = (x,y) \mid \|x - \bar x\| < r_1,\ G(z) \in \mathcal{V}\big\} \ \subseteq\ \mathcal{U}.$$

Fix $z = (x,y) \in \mathcal{U}'$ with $G(z)$ nonsingular, of inertia $(\bar p + a, \bar q + k - a)$, and fix $M^{(a)}$ as constructed above with ρ so small that $G(z_{M^{(a)}}) \in \mathcal{V}$. We connect z to $z_{M^{(a)}}$ by a path within $\mathcal{U}'$ that stays on the stratum of z. First, the shear path $z(t) := \big((1-t)x + t\bar x,\ G(z) - g((1-t)x + t\bar x)\big)$, $t \in [0,1]$, keeps $G(z(t)) = G(z)$ constant, remains in $\mathcal{U}'$, and ends at the pair $\big(\bar{x},\, G(z)-g(\bar{x})\big)$. Next, since $\mathcal{M}_{\bar{p}+a,\bar{q}+k-a}\cap\mathcal{V}$ is path-connected by the choice of $\mathcal{V}$, there is a path $\Gamma(\tau)$ in this set from $G(z)$ to $G(z_{M^{(a)}})$. The pairs $(\bar{x},\, \Gamma(\tau)-g(\bar{x}))$ then form a path in $\mathcal{U}'$, on the same stratum, from $\big(\bar{x},\, G(z)-g(\bar{x})\big)$ to $z_{M^{(a)}}$.

Along the concatenated path no eigenvalue of G vanishes, so the negative spectral subspace of G varies continuously. Since $\beta=\emptyset$ along the path, by (22) the space $\operatorname{appl}(z)$ is the kernel of the map sending v to the compression $P_\gamma^\top\big(\nabla g(x)^* v\big)P_\gamma$, where the columns of P_γ span the negative spectral subspace of $G(z)$. This map varies continuously and is surjective by Proposition 13(a) and (17), so the subspaces $\operatorname{appl}(\cdot)$ have constant dimension and vary continuously. Since $\mathcal{Q}_z$ does not depend on the choice of IED and its coefficients depend continuously on the spectral data along a family with no vanishing eigenvalues, the compressions $\mathcal{Q}_\cdot|_{\operatorname{appl}(\cdot)}$ also vary continuously. By (43), the eigenvalues of these compressions avoid $(-1/\kappa', 1/\kappa')$, so their n_- is constant along the path. This value vanishes at $z_{M^{(a)}}$ by the choice of $M^{(a)}$, hence at z. Combined with (43), this shows that

$$\mathcal{Q}_z(v_x) \ge \kappa'^{-1}\|v_x\|^2 \qquad \forall\, v_x\in\operatorname{appl}(z)$$

for every $z\in\mathcal{U}'$ with $G(z)$ nonsingular.

It remains to treat the pairs $z=(x,y)\in\mathcal{U}'$ with $G(z)$ singular. Let $(\alpha,\beta,\gamma,p,q,P,\lambda)$ be an IED of $G(z)$ with $\beta\neq\emptyset$, and let Π_β be the orthogonal projector onto the kernel of $G(z)$. For $\varepsilon>0$ set $z_\varepsilon := (x,\, y-\varepsilon\Pi_\beta)$. Since $G(z_\varepsilon)=G(z)-\varepsilon\Pi_\beta$ has the same eigenframe as $G(z)$ with the β-eigenvalues moved to $-\varepsilon$, and $\mathcal{V}$ is an open set containing $G(z)$, for all sufficiently small $\varepsilon>0$ the pair z_ε lies in $\mathcal{U}'$ and on the stratum $\widetilde{\mathcal{M}}_{p,\,q+|\beta|}$. Comparing blocks via (22) gives $\operatorname{appl}(z_\varepsilon)=\operatorname{appl}(z)$. On this common subspace, the spectral terms indexed by β carry coefficients $\varepsilon/\lambda_i\to 0$ with bounded entries, the remaining spectral terms do not depend on ε, and $\nabla^2_{xx}L(z_\varepsilon)-\nabla^2_{xx}L(z)=O(\varepsilon)$. Hence $\mathcal{Q}_{z_\varepsilon}\to\mathcal{Q}_z$ on this subspace as $\varepsilon\downarrow 0$. Since $z_\varepsilon\in\mathcal{U}'$ and $G(z_\varepsilon)$ is nonsingular, $\mathcal{Q}_{z_\varepsilon}(v_x)\ge\kappa'^{-1}\|v_x\|^2$ for every v_x in this subspace, and the bound passes to the limit. Thus the same lower bound holds when $G(z)$ is singular. Together with the nonsingular case, this proves that the W-SOC holds uniformly on $\mathcal{U}'$ with constant $1/\kappa'$. □

5.2.4 Regularity

We can now combine the results of this section into a single theorem showing that the local and local uniform validity of the stratified conditions can be used to recover strong regularity, with metric regularity included as an equivalent consequence.

Theorem 12 *Let $\bar{x}$ be a local minimizer of* (1) *satisfying the RCQ, and let $\bar{y}\in M(\bar{x})$. The following statements are equivalent:*

(a) *the S-SOSC and the constraint nondegeneracy hold at $(\bar{x},\bar{y})$;*
(b) *every element of the Clarke generalized Jacobian $\partial_C F(\bar{x},\bar{y})$ is nonsingular;*
(c) *the KKT mapping F is strongly metrically regular at $((\bar{x},\bar{y}),0)$;*
(d) *there exists a neighborhood $\mathcal{U}$ of $(\bar{x},\bar{y})$ on which the W-SOC holds uniformly in the sense of Definition 29 and the W-SRCQ holds at every $z\in\mathcal{U}$;*
(e) *there exist $\kappa>0$ and a neighborhood $\mathcal{U}$ of $(\bar{x},\bar{y})$ such that, for every (p,q),*

$$\|dF_z v\| \;\ge\; \kappa^{-1}\|v\| \qquad \forall\, z\in\mathcal{U}\cap\widetilde{\mathcal{M}}_{p,q},\quad \forall\, v\in T_z\widetilde{\mathcal{M}}_{p,q};$$

(f) *the KKT mapping F is metrically regular at $((\bar{x},\bar{y}),0)$; equivalently, F^{-1} has the Aubin property at $(0,(\bar{x},\bar{y}))$.*

Proof By the KKT reductions in Section 3.2, the S-SOSC and constraint nondegeneracy in (a) coincide with their classical counterparts at $(\bar{x},\bar{y})$. Hence, (a)⇔(b)⇔(c) follows from [48, Theorem 4.1], while (a)⇔(d) follows from Theorems 8 and 9. Since strong metric regularity implies metric regularity, (c)⇒(f). It remains to prove (f)⇒(e)⇒(d).

(f)⇒(e). By [28, Theorem 2.1 and the argument of Corollary 2.3], there exist $\alpha>0$ and a neighborhood $\mathcal{U}$ of $(\bar{x},\bar{y})$ such that

$$\|F'(z)v\|\ge\alpha^{-1}\|v\|$$

at every differentiability point $z\in\mathcal{U}$. For any $V\in\partial_B F(z)$, take differentiability points $z^k\to z$ such that $F'(z^k)\to V$. Since $z^k\in\mathcal{U}$ eventually, passing to the limit gives

$$\|Vv\|\ge\alpha^{-1}\|v\| \qquad \forall\, v.$$

For $z\in\mathcal{U}\cap\widetilde{\mathcal{M}}_{p,q}$ and $v\in T_z\widetilde{\mathcal{M}}_{p,q}$, choose $V\in\partial_B F(z)$. Proposition 8 gives $dF_z v = Vv$, proving (e).

(e)⇒(d). By (e), dF_z is injective for every $z\in\mathcal{U}$, so Theorem 6(b) gives the W-SRCQ at every $z\in\mathcal{U}$. Since $\mathcal{U}$ is open, Theorem 8 applies at $(\bar{x},\bar{y})$ and gives the constraint nondegeneracy there. By constraint nondegeneracy,

$M(\bar{x}) = \{\bar{y}\}$. Together with the fact that $\bar{x}$ is a local minimizer satisfying the RCQ, this multiplier uniqueness implies, by [46, Theorem 8], that the SONC holds at $(\bar{x}, \bar{y})$. Hence Assumption 10 is satisfied at $(\bar{x}, \bar{y})$. By Theorem 11, there is a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ on which the W-SOC holds uniformly. Since the W-SRCQ already holds throughout $\mathcal{U}$, it also holds throughout $\mathcal{U}'$. This proves (d). □

Figure 6 summarizes the results of the analysis along and across the lifted strata.

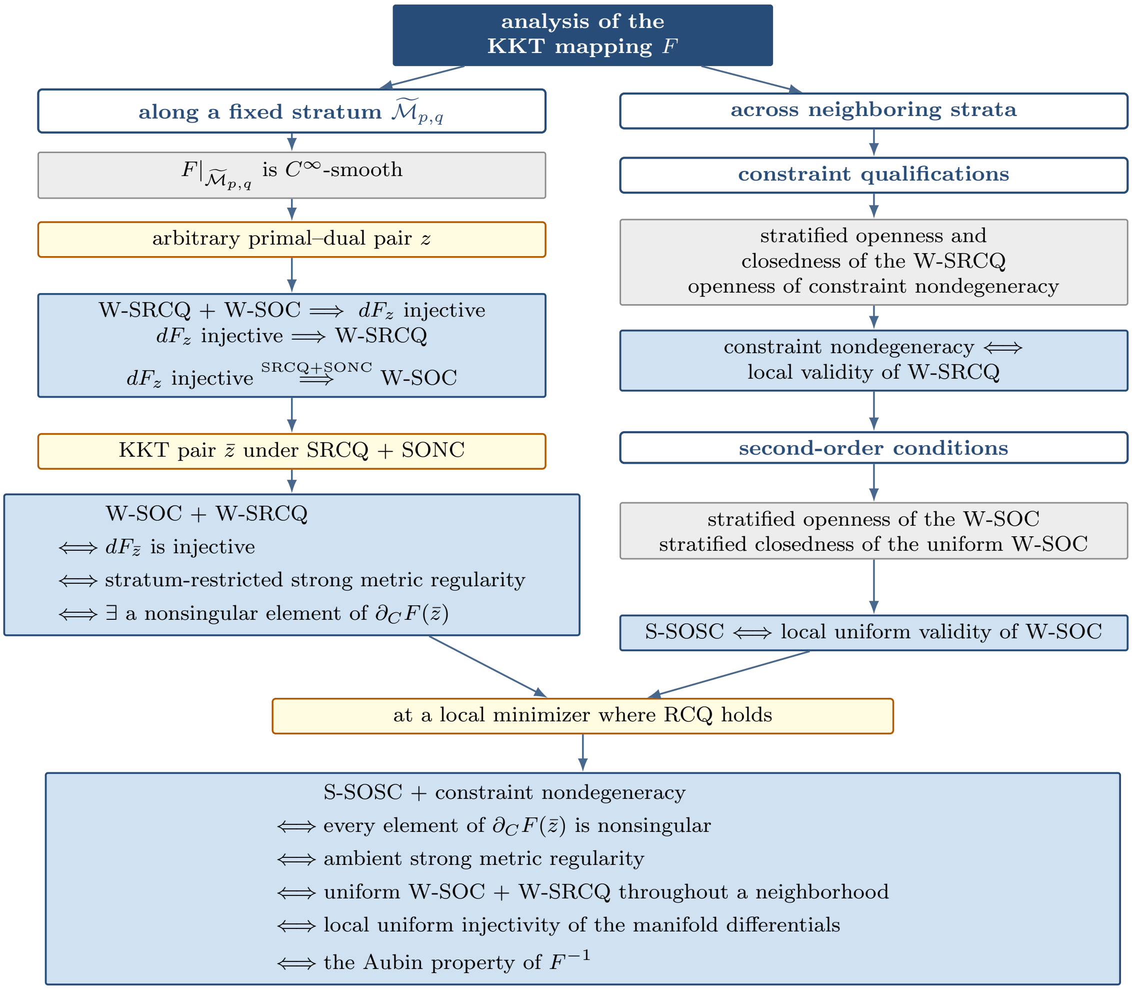

Fig. 6: Analysis along and across the lifted strata.

Remark 22 As a by-product of the stratified perturbation analysis, Theorem 12 recovers the known equivalence between the Aubin property and strong regularity for the NLSDP KKT mapping established in [12]; see [13] for nonlinear second order cone programming and [38] for general C^2-cone reducible programs. The present proof makes the role of the lifted inertia stratification explicit: the Aubin property yields the local uniform injectivity of the manifold differentials in (e), and the analysis along and across nearby lifted strata promotes (e) to the uniform weak conditions in (d), which in turn recover strong metric regularity.

Remark 23 The inertia stratification may seem related to partial smoothness because F is smooth on each lifted stratum. This connection, however, is limited to restricted smoothness. Partial smoothness concerns a function or set relative to an active manifold and requires restricted smoothness, regularity, normal sharpness, and subgradient continuity [37, Definition 2.7]. For example, at a KKT pair the squared residual $\frac{1}{2}\|F(z)\|^2$ has zero directional derivative in every direction. Hence normal sharpness fails. In the NLSDP setting, identification based on partial smoothness requires strict complementarity [18, 53].

Accordingly, we work with the full inertia stratification rather than a single stratum. This allows us to compare regularity under changes of inertia and suggests a possible direction for stratified algorithms.

6 Concluding remarks

In this paper, we developed a geometry-resolved perturbation analysis for the KKT mapping of NLSDP (1) based on the inertia stratification of $\mathbb{S}^n$. The C^∞-smoothness of the positive semidefinite projection on each inertia stratum yields smooth restrictions of the KKT mapping with explicit manifold differentials. The auxiliary NLSDP extends the geometric interpretation to arbitrary primal–dual pairs, while the two nested manifold models identify constraint nondegeneracy and the W-SRCQ with transversality and relate the W-SOC and the S-SOC to manifold second-order conditions under the corresponding qualification assumptions. Along a fixed lifted stratum, the W-SOC and the W-SRCQ imply injectivity of the manifold differential, which at a KKT pair characterizes the proposed stratum-restricted strong metric regularity.

The analysis across strata reveals how these stratified conditions assemble into the classical ambient theory. The W-SRCQ and the W-SOC have complementary inertia-dependent stability properties governed by the negative and positive inertia indices, respectively, while the local validity of the W-SRCQ and the local uniform validity of the W-SOC recover constraint nondegeneracy and the S-SOSC, respectively. At a KKT pair associated with a local minimizer satisfying the RCQ, local uniform injectivity of the manifold differentials provides the link between these stratified conditions and strong regularity. The known equivalence between the Aubin property of the inverse KKT mapping and strong metric regularity is then recovered as a by-product.

The stratified viewpoint developed here also motivates the design of algorithms for NLSDP at degenerate KKT pairs. Their development and analysis will be pursued in a separate work.

References

1. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization algorithms on matrix manifolds. Princeton University Press (2008)
2. Andreani, R., Haeser, G., Ramos, A., Silva, P.J.: A second-order sequential optimality condition associated to the convergence of optimization algorithms. IMA Journal of Numerical Analysis **37**(4), 1902–1929 (2017)
3. Andreani, R., Secchin, L.D.: A note on the convergence of an augmented Lagrangian algorithm to second-order stationary points. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics **6**(1) (2018)
4. Bhatia, R.: Matrix Analysis, *Graduate Texts in Mathematics*, vol. 169. Springer (1997)
5. Bonnans, J.F., Cominetti, R., Shapiro, A.: Sensitivity analysis of optimization problems under second order regular constraints. Mathematics of Operations Research **23**(4), 806–831 (1998)
6. Bonnans, J.F., Cominetti, R., Shapiro, A.: Second order optimality conditions based on parabolic second order tangent sets. SIAM Journal on Optimization **9**(2), 466–492 (1999)
7. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Verlag (2000)
8. Bonnans, J.F., Sulem, A.: Pseudopower expansion of solutions of generalized equations and constrained optimization problems. Mathematical Programming **70**(1-3), 123 – 148 (1995). DOI 10.1007/bf01585932
9. Boumal, N.: An introduction to optimization on smooth manifolds. Cambridge University Press (2023)
10. Burer, S., Monteiro, R.D.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Mathematical programming **95**(2), 329–357 (2003)
11. Burer, S., Monteiro, R.D.: Local minima and convergence in low-rank semidefinite programming. Mathematical programming **103**(3), 427–444 (2005)
12. Chen, L., Chen, R., Sun, D., Zhang, L.: Characterizations of the Aubin property of the solution mapping for nonlinear semidefinite programming. Mathematical Programming **215**(1–2), 637–668 (2026). DOI 10.1007/s10107-025-02231-2. URL `https://link.springer.com/10.1007/s10107-025-02231-2`
13. Chen, L., Chen, R., Sun, D., Zhu, J.: Aubin property and strong regularity are equivalent for nonlinear second-order cone programming. SIAM Journal on Optimization **35**(2), 712–738 (2025). DOI 10.1137/24M1670676. URL `https://doi.org/10.1137/24M1670676`
14. Chen, X., Qi, H., Tseng, P.: Analysis of nonsmooth symmetric-matrix-valued functions with applications to semidefinite complementarity problems. SIAM Journal on Optimization **13**(4), 960–985 (2003)
15. Chen, X., Tseng, P.: Non-interior continuation methods for solving semidefinite complementarity problems. Mathematical Programming **95**(3), 431–474 (2003)
16. Clarke, F.H.: Optimization and nonsmooth analysis. SIAM (1990)
17. Conway, J.B.: A course in functional analysis, vol. 96. Springer Science & Business Media (1994)
18. Davis, D., Drusvyatskiy, D., Jiang, L.: Active manifolds, stratifications, and convergence to local minima in nonsmooth optimization. Foundations of Computational Mathematics **26**(2), 779–861 (2026). DOI 10.1007/s10208-025-09691-0
19. Ding, C., Sun, D., Zhang, L.: Characterization of the robust isolated calmness for a class of conic programming problems. SIAM Journal on Optimization **27**(1), 67–90 (2017)
20. Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM Journal on Optimization **6**(4), 1087 – 1105 (1996). DOI 10.1137/s1052623495284029
21. Dontchev, A.L., Rockafellar, R.T.: Implicit functions and solution mappings, vol. 543. Springer (2009)
22. Dorsch, D., Gómez, W., Shikhman, V.: Sufficient optimality conditions hold for almost all nonlinear semidefinite programs. Mathematical Programming **158**(1–2), 77–97 (2016). DOI 10.1007/s10107-015-0915-0
23. Drusvyatskiy, D., Lewis, A.S.: Optimality, identifiability, and sensitivity. Mathematical Programming **147**(1), 467–498 (2014)
24. Fares, B., Noll, D., Apkarian, P.: Robust control via sequential semidefinite programming. SIAM Journal on Control and Optimization **40**(6), 1791–1820 (2002). DOI 10.1137/S0363012900373483
25. Feng, F., Ding, C., Li, X.: A quadratically convergent semismooth Newton method for nonlinear semidefinite programming without generalized Jacobian regularity. Mathematical Programming **214**, 643–683 (2025). DOI 10.1007/s10107-025-02198-0
26. Freund, R.W., Jarre, F., Vogelbusch, C.H.: Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling. Mathematical Programming **109**(2–3), 581–611 (2007). DOI 10.1007/s10107-006-0028-x

27. Fukuda, E.H., Haeser, G., Mito, L.M.: On the weak second-order optimality condition for nonlinear semidefinite and second-order cone programming. Set-Valued and Variational Analysis **31**(2), 15 (2023)
28. Fusek, P.: On metric regularity for weakly almost piecewise smooth functions and some applications in nonlinear semidefinite programming. SIAM Journal on Optimization **23**(2), 1041–1061 (2013). DOI 10.1137/120887035
29. Gould, N., Toint, P.: A note on the convergence of barrier algorithms to second-order necessary points. Mathematical programming **85**(2), 433–438 (1999)
30. Guo, X., Du, J., Gao, X.: Confidence structural robust optimization by non-linear semidefinite programming-based single-level formulation. International Journal for Numerical Methods in Engineering **86**(8), 953–974 (2011). DOI 10.1002/nme.3083
31. Helmke, U., Shayman, M.A.: Critical points of matrix least squares distance functions. Linear Algebra and its Applications **215**, 1–19 (1995). DOI https://doi.org/10.1016/0024-3795(93)00070-G. URL `https://www.sciencedirect.com/science/article/pii/002437959300070G`
32. Hirsch, M.W.: Differential topology, vol. 33. Springer Science & Business Media (2012)
33. Ioffe, A.D.: An invitation to tame optimization. SIAM Journal on Optimization **19**(4), 1894–1917 (2009)
34. Jongen, H.T., Möbert, T., Rückmann, J., Tammer, K.: On inertia and schur complement in optimization. Linear Algebra and its Applications **95**, 97–109 (1987)
35. Journée, M., Bach, F., Absil, P.A., Sepulchre, R.: Low-rank optimization on the cone of positive semidefinite matrices. SIAM Journal on Optimization **20**(5), 2327–2351 (2010)
36. Lee, J.: Introduction to Smooth Manifolds, vol. 218. Springer Science & Business Media (2012)
37. Lewis, A.S.: Active sets, nonsmoothness, and sensitivity. SIAM Journal on Optimization **13**(3), 702–725 (2002). DOI 10.1137/S1052623401387623
38. Ma, J., Sun, D.: The Aubin property for generalized equations over C^2-cone reducible sets. arXiv preprint arXiv:2509.14194 (2025)
39. Mather, J.: Notes on topological stability. Bulletin of the American Mathematical Society **49**(4), 475–506 (2012). DOI 10.1090/S0273-0979-2012-01383-6
40. Molina, J.L.C., Lê, D.T., Seade, J.: Handbook of Geometry and Topology of Singularities I. Springer (2020)
41. Pang, J.S., Sun, D., Sun, J.: Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems. Mathematics of Operations Research **28**(1), 39–63 (2003)
42. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Mathematics of operations research **18**(1), 227–244 (1993)
43. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Mathematical programming **58**(1), 353–367 (1993)
44. Robinson, S.M.: Strongly regular generalized equations. Mathematics of Operations Research **5**, 43–62 (1980)
45. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis, *Springer Berlin Heidelberg*, vol. 317. Springer Berlin Heidelberg (1998). DOI 10.1007/978-3-642-02431-3. URL `http://link.springer.com/10.1007/978-3-642-02431-3`
46. Shapiro, A.: First and second order analysis of nonlinear semidefinite programs. Mathematical Programming **77**(1), 301–320 (1997)
47. Simon, B.: Operator theory, vol. 4. American Mathematical Soc. (2015)
48. Sun, D.: The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. Mathematics of Operations Research **31**(4), 761–776 (2006)
49. Sun, D., Sun, J.: Semismooth matrix-valued functions. Mathematics of Operations Research **27**(1), 150–169 (2002)
50. Vandereycken, B.: Low-rank matrix completion by Riemannian optimization. SIAM Journal on Optimization **23**(2), 1214–1236 (2013)
51. Whitney, H.: Local properties of analytic varieties. Differential and Combinatorial Topology pp. 205–244 (1965)
52. Whitney, H.: Tangents to an analytic variety. The Annals of Mathematics **81**(3), 496 (1965)
53. Wu, H., Xie, Y.: Identifiability and error bound: Metric and geometric perspectives. arXiv preprint arXiv:2605.02754 (2026)
54. Yamakawa, Y., Okuno, T.: A stabilized sequential quadratic semidefinite programming method for degenerate nonlinear semidefinite programs. Computational Optimization and Applications **83**(3), 1027–1064 (2022). DOI 10.1007/s10589-022-00402-x
55. Yang, L., Sun, D., Toh, K.C.: SDPNAL+: A majorized semismooth Newton–CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. Mathematical Programming Computation **7**, 331–366 (2015). DOI 10.1007/s12532-015-0082-6