On Constraint Qualifications for MPECs with Applications to Bilevel Hyperparameter Optimization for Machine Learning

Jiani Li[a], Qingna Li[a,b] and Alain Zemkoho[c]

[a]School of Mathematics and Statistics, Beijing Institute of Technology, Beijing, China;
[b]Beijing Key Laboratory on MCAACI/Key Laboratory of Mathematical Theory and Computation in Information Security, Beijing Institute of Technology, Beijing, China;
[c]School of Mathematical Sciences, University of Southampton, Southampton, SO17 1BJ, UK.

ARTICLE HISTORY
Compiled August 19, 2025

ABSTRACT
Constraint qualifications for a Mathematical Program with Equilibrium Constraints (MPEC) are essential for analyzing stationarity properties and establishing convergence results. In this paper, we explore several classical MPEC constraint qualifications and clarify the relationships among them. We subsequently examine the behavior of these constraint qualifications in the context of a specific MPEC derived from bilevel hyperparameter optimization (BHO) for L1-loss support vector classification. In particular, for such an MPEC, we provide a complete characterization of the well-known MPEC linear independence constraint qualification (MPEC-LICQ), therefore, establishing conditions under which it holds or fails for our BHO for support vector machines.

KEYWORDS
Bilevel optimization, Mathematical program with equilibrium constraints, Constraint qualifications, MPEC-linear constraint qualification, MPEC-Mangasarian-Fromovitz constraint qualification, Support vector classification, Hyperparameter selection

1. Introduction

A bilevel optimization problem (BLO) is a hierarchical optimization problem in which the feasible region of the upper-level problem depends on the solution set of the lower-level problem [1–3]. It has been extensively studied because of its significant applications across various fields of science and engineering [4–6], especially in the fast-developing area of machine learning [7–14].

In general, a BLO is highly complicated and computationally challenging to solve due to its nonconvexity and nondifferentiability [15]. A commonly used approach for solving a BLO is to replace the lower-level problem with its Karush–Kuhn–Tucker (KKT) conditions, thereby transforming it into a mathematical program with equilibrium constraints (MPEC) [16–18].

Email: lijiani_qaz@163.com
Email: qnl@bit.edu.cn. Corresponding author.
Email: a.b.zemkoho@soton.ac.uk

Due to the presence of complementarity constraints, many commonly used constraint qualifications (CQs) in nonlinear programming (NLP), such as the linear independence constraint qualification (LICQ) and Mangasarian-Fromovitz constraint qualification (MFCQ) fail to hold at all feasible points of an MPEC [19,20].

When CQs fail, the KKT conditions may not be valid at a minimizer, and the convergence assumptions for most standard methods for constrained optimization may not be satisfied [21].

For these reasons, establishing appropriate constraint qualifications tailored to the complementarity structure in MPECs is crucial, in part to characterize local optimal solutions. Specialized CQs that capture the structure of the problem have been widely studied in the literature; see, e.g., [16,22]. For example, we mention the Tightened NLP–based extension of MFCQ, which we label MPEC-MFCQ-T, proposed by Scheel and Scholtes [23], and the Relaxed NLP–based MFCQ, denoted MPEC-MFCQ-R, introduced by Ralph and Wright [24]. Ye [25] develops other extensions of MFCQ for MPECs. Flegel and Kanzow [20,26] explored MPEC-specific adaptations of the Abadie constraint qualification (MPEC-ACQ) and the Guignard constraint qualification (MPEC-GCQ). Further developments include the work by Henrion and Outrata [27], who study conditions for detecting the calmness property in constraint set-valued mappings.

Considering this wide range of CQs tailored to MPECs, the first aim of this paper is to take a close look at them and study the possible connections among them. In doing so, we identify those that are equivalent to each other (especially in the context of the multiple MFCQ–type CQs) and those that stand in strict implication relationships. Secondly, we study the behavior of these CQs for the bilevel hyperparameter optimization (BHO) problem in support vector machines.

Note that, in the context of BHO, there have been many recent publications, often with varied uses of CQs. For instance, Bennett et al. have conducted a series of studies on hyperparameter selection based on transforming the problem into an MPEC [35–39]. Recently, Li et al. [40] formulate the support vector classification (SVC) hyperparameter selection problem as a bilevel optimization problem and then transform it into an MPEC, which they solve with a relaxation algorithm that converges under an MPEC-MFCQ.

In [41], Qian et al. introduce an efficient linear-programming-based Newton global relaxation method to solve the MPEC derived from the bilevel model in [37], which incorporates multiple hyperparameters for feature selection. Furthermore, they show that, under appropriate conditions, the MPEC-MFCQ defined in [40] holds. In [42], the authors further analyze and characterize MPEC-LICQ for the MPEC arising from the bilevel model used for hyperparameter selection in the context of the kernel-based SVC hyperparameter optimization problem. Furthermore, in [43], Ward et al. improve solution methods for MPCCs with new convergence results under weaker assumptions and formulate SVM hyperparameter tuning as an MPCC to leverage derivative information for efficiency. Wang and Li [44] formulate hyperparameter selection for SVC with logistic loss as a bilevel optimization problem, transform it into a single-level NLP via KKT conditions, and propose a smoothing Newton method with superlinear convergence to efficiently obtain strict local minimizers.

After systematically studying MPEC-tailored CQs, analyzing their connections, and providing examples that illustrate cases where certain implications between them do not hold, we apply these CQs to the BHO problem, focusing in particular on identifying cases where MPEC-LICQ fails or is satisfied. This latter aspect of our analysis provides a way to check the satisfaction of MPEC-LICQ for similar MPECs derived from BHO for support vector machines.

The organization of the paper is as follows. In Section 2, we introduce various definitions of constraint qualifications for MPEC and study the relationships between them. In Section 3, we give a brief review about the bilevel model for hyperparameter selection and its resulting MPEC in [40]. Then we show the equivalence of MPEC-MFCQ-T and MPEC-LICQ for the resulting MPEC. Moreover, we analyse the fulfillment of MPEC-MFCQ-R and MPEC-LICQ. The final conclusions are made in Section 4.

Notations. We use $\|x\|_0$ to denote the number of nonzero elements in $x \in \mathbb{R}^n$, while $\|x\|_1$ and $\|x\|_2$ correspond to the l_1-norm and l_2-norm of x, respectively. We will use $x_+ = \left((x_1)_+, \cdots, (x_n)_+\right) \in \mathbb{R}^n$, where $(x_i)_+ = \max(x_i, 0)$. $|\Omega|$ denotes the number of elements in the set $\Omega \subset \mathbb{R}^n$. We use $[k]$ to denote $\{1, 2, ..., k\}$ for the positive integer k. We use $\mathbf{1}_k$ to denote a vector with elements all ones in $\mathbb{R}^k$. I_k is the identity matrix in $\mathbb{R}^{k\times k}$. The notation $\mathbf{0}_{k\times q}$ represents a zero matrix in $\mathbb{R}^{k\times q}$ and $\mathbf{0}_k$ stands for a zero vector in $\mathbb{R}^k$. On the other hand, $\mathbf{0}_{(\tau,\kappa)}$ will be used for a submatrix of the zero matrix, where τ is the index set of the rows and κ is the index set of the columns. Similarly to the case of zero matrix, $I_{(\tau,\tau)}$ corresponds to a submatrix of an identity matrix indexed by both rows and columns in the set τ. Finally, $\Theta_{(\tau,\cdot)}$ represents a submatrix of the matrix Θ, where τ is the index set of the rows, and x_τ is a subvector of the vector x corresponding to the index set τ.

2. Constraint Qualifications for MPECs

In this section, we will give the definitions of different CQs for MPECs, explain their theoretical foundations, and discuss the relationships between them.

2.1. Preliminaries for NLP

We start with the definition of constraint qualifications for a general nonlinear programming problem

$$\begin{array}{ll} \min\limits_{v\in\mathbb{R}^n} & f(v) \\ \text{s.t.} & g_i(v) \le 0, \ \forall\, i \in [m], \\ & h_j(v) = 0, \ \forall\, j \in [p], \end{array} \tag{NLP}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable, $i \in [m]$, $j \in [p]$. For a feasible vector v^* of (NLP), we consider the set

$$I_g = \{i \in [m] : g_i(v^*) = 0\},$$

which gathers the indices of constraints that are active at v^*. We now recall several standard constraint qualifications used in nonlinear programming (NLP).

Definition 2.1. A point v^* feasible for (NLP) satisfies the Linear Independence Constraint Qualification (LICQ) if and only if the set of gradient vectors $\{\nabla g_i(v^*) \mid i \in I_g\} \cup \{\nabla h_i(v^*) \mid i \in [p]\}$ is linearly independent.

To introduce MFCQ, we first recall the notion of positively linearly independence.

Definition 2.2. Let v^* be feasible for (NLP) and let $I_1 \subseteq I_g$, $I_2 \subseteq [p]$ be arbitrarily given. The set of gradients $\{\nabla g_i(v^*) \mid i \in I_1\} \cup \{\nabla h_i(v^*) \mid i \in I_2\}$ is called positively linearly independent if there do not exist scalars $\{\alpha_i\}_{i\in I_1}$ and $\{\beta_i\}_{i\in I_2}$, with $\alpha_i \geq 0$ for all $i \in I_1$, not all zero, such that $\sum_{i\in I_1} \alpha_i \nabla g_i(v^*) + \sum_{i\in I_2} \beta_i \nabla h_i(v^*) = 0$. Otherwise, we say that this set of gradient vectors is positively linearly dependent.

Definition 2.3. A point v^* feasible for (NLP) satisfies the Mangasarian–Fromovitz Constraint Qualification (MFCQ) if and only if the set of gradient vectors $\{\nabla g_i(v^*) \mid i \in I_g\} \cup \{\nabla h_i(v^*) \mid i \in [p]\}$ is positively linearly independent.

We also invoke the Abadie Constraint Qualification (ACQ), a geometric condition requiring the tangent cone to coincide with the linearized cone at the reference point.

Definition 2.4. A point v^* feasible for (NLP) satisfies the ACQ if and only if $\mathcal{T}(v^*) = \mathcal{T}^{\text{lin}}(v^*)$, where $\mathcal{T}(v^*)$ and $\mathcal{T}^{\text{lin}}(v^*)$ are the tangent cone and the linearized cone at v^*, respectively, defined as

$$\mathcal{T}(v^*) := \{d \in \mathbb{R}^n \mid \exists\ d_n \to d,\ t_n \downarrow 0,\ \text{s.t. } v^* + t_n d_n \text{ is a feasible point of (NLP)},\ \forall\ n \in \mathbb{N}\},$$
$$\mathcal{T}^{\text{lin}}(v^*) := \left\{d \in \mathbb{R}^n \mid \nabla g_i(v^*)^\top d \leq 0,\ \forall\ i \in I_g,\ \nabla h_i(v^*)^\top d = 0,\ \forall\ i \in [p]\right\}.$$

Finally, we recall the KKT conditions, which will link the above CQs to first-order optimality.

Definition 2.5. A feasible point v^* satisfies the *Karush–Kuhn–Tucker (KKT) conditions* if there exist Lagrange multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that

$$\begin{aligned}
&\nabla f(v^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(v^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(v^*) = 0,\\
&h(v^*) = 0,\\
&g(v^*) \leq 0, \quad \lambda \geq 0, \quad \lambda^\top g(v^*) = 0.
\end{aligned}$$

(v^*, λ, μ) is called a KKT point.

If v^* is a local minimum of problem (NLP) and a suitable constraint qualification, such as ACQ or any of the previously mentioned CQs, holds at v^*, then v^* is a stationary point of (NLP), i.e., $(v^*,\ \lambda,\ \mu)$ is a KKT point of (NLP). Thus, the point v^* corresponding to a KKT point is referred to as a stationary point.

2.2. Constraint Qualifications for MPECs

In this part, we introduce the essential constraint qualifications for MPECs and their implications for optimality. We begin with the model and index sets, proceed from standard to more general conditions, and close with stability concepts. The MPEC takes the following form:

$$\begin{aligned}
\min_{v\in\mathbb{R}^n}\quad & f(v)\\
\text{s.t.}\quad & g_i(v) \leq 0,\ \forall\ i \in [m],\\
& h_i(v) = 0,\ \forall\ i \in [p],\\
& G_i(v) \geq 0,\ H_i(v) \geq 0,\ G_i(v)H_i(v) = 0,\ \forall\ i \in [l],
\end{aligned} \tag{MPEC}$$

where $G_i : \mathbb{R}^n \to \mathbb{R}$ and $H_i : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable. Let v^* be a feasible point for (MPEC), and define the index sets

$$\begin{aligned} I_g &= \{i \in [m] : \ g_i(v^*) = 0\}, \\ I_G &= \{i \in [l] : \ G_i(v^*) = 0, \ H_i(v^*) > 0\}, \\ I_H &= \{i \in [l] : \ G_i(v^*) > 0, \ H_i(v^*) = 0\}, \\ I_{GH} &= \{i \in [l] : \ G_i(v^*) = 0, \ H_i(v^*) = 0\}. \end{aligned}$$

We start with MPEC-LICQ for (MPEC), then move to weaker constraint qualifications tailored to (MPEC).

Definition 2.6. [33, page 920] (MPEC-LICQ) A feasible point v^* for problem (MPEC) satisfies MPEC-LICQ if and only if the set of gradient vectors

$$\begin{aligned} &\{\nabla g_i(v^*) \mid i \in I_g\} \cup \big\{ \{\nabla h_i(v^*) \mid i \in [p]\} \\ &\cup \{\nabla G_i(v^*) \mid i \in I_G \cup I_{GH}\} \cup \{\nabla H_i(v^*) \mid i \in I_H \cup I_{GH}\} \big\} \end{aligned} \tag{1}$$

is linearly independent.

One type of MPEC-MFCQ (we refer to it as MPEC-MFCQ-T) is defined based on the so-called tightened nonlinear programming problem at v^*:

$$\begin{aligned} \min_{v \in \mathbb{R}^n} \quad & f(v) \\ \text{s.t.} \quad & g_i(v) \leq 0, \ \forall\, i \in [m], \\ & h_i(v) = 0, \ \forall\, i \in [p], \\ & G_i(v) = 0, \ H_i(v) \geq 0, \ \forall\, i \in I_G, \\ & G_i(v) \geq 0, \ H_i(v) = 0, \ \forall\, i \in I_H, \\ & G_i(v) = 0, \ H_i(v) = 0, \ \forall\, i \in I_{GH}. \end{aligned} \tag{TNLP(v^*)}$$

Definition 2.7. [45, page 263] (MPEC-MFCQ-T) A feasible point v^* for problem (MPEC) satisfies MPEC-MFCQ-T if standard MFCQ holds for the corresponding TNLP. Equivalently, MPEC-MFCQ-T holds at v^* if and only if the set of gradient vectors in (1) is *positively linearly* independent.

Remark 1. Recall that the set of gradient vectors in (1) is said to be positive-linearly *dependent* if there exist scalars $\{\delta_i\}_{i \in I_g}$, $\{\alpha_i\}_{i \in [p]}$, $\{\beta_i\}_{i \in I_G \cup I_{GH}}$ and $\{\gamma_i\}_{i \in I_H \cup I_{GH}}$ with $\delta_i \geq 0$ for $i \in I_g$, not all of them being zero, such that $\Sigma_{i \in I_g} \delta_i \nabla g_i(v^*) + \Sigma_{i \in [p]} \alpha_i \nabla h_i(v^*) + \Sigma_{i \in I_G \cup I_{GH}} \beta_i \nabla G_i(v^*) + \Sigma_{i \in I_H \cup I_{GH}} \gamma_i \nabla H_i(v^*) = 0$. Otherwise, we say that this set of gradient vectors is positive-linearly independent.

In comparison, another type of MPEC-MFCQ [24] (we refer to it as MPEC-MFCQ-R) is based on the following relaxed nonlinear programming problem at the feasible point v^*:

$$\begin{aligned} \min_{v \in \mathbb{R}^n} \quad & f(v) \\ \text{s.t.} \quad & g_i(v) \leq 0, && \forall\, i \in [m], \\ & h_i(v) = 0, && \forall\, i \in [p], \\ & G_i(v) = 0, && \forall\, i \in I_G, \\ & H_i(v) = 0, && \forall\, i \in I_H, \\ & G_i(v) \geq 0, \ H_i(v) \geq 0, && \forall\, i \in I_{GH}. \end{aligned} \tag{RNLP(v^*)}$$

Definition 2.8. (MPEC-MFCQ-R) MPEC-MFCQ-R holds at v^* if standard MFCQ holds for (RNLP(v^*)); that is, the set

$$\begin{aligned}&\big\{ \{\nabla g_i(v^*) \mid i \in I_g\} \cup \{\nabla G_i(v^*) \mid i \in I_{GH}\} \cup \{\nabla H_i(v^*) \mid i \in I_{GH}\} \big\} \\ &\cup \big\{ \{\nabla h_i(v^*) \mid i \in [p]\} \cup \{\nabla G_i(v^*) \mid i \in I_G\} \cup \{\nabla H_i(v^*) \mid i \in I_H\} \big\}\end{aligned}$$

is *positively linearly* independent.

Obviously, (TNLP(v^*)) and (RNLP(v^*)) depend on the chosen point v^*. Note that (TNLP(v^*)) is "tightened" because its feasible region is a subset of that of (MPEC). However, there is no inclusion relationship between the feasible regions of (RNLP(v^*)) and (MPEC). Moreover, if v^* is a local minimizer of (RNLP(v^*)), then it is also a local minimizer of (MPEC); and if v^* is a local minimizer of (MPEC), then it is a local minimizer of (TNLP(v^*)). The reverse implications generally hold only under strict complementarity at v^*.

While classical CQs (such as MPEC-MFCQ-T and MPEC-LICQ) provide fundamental feasibility conditions for optimality in MPECs, they can be restrictive in practice. To address this, generalized CQs have been introduced that relax certain assumptions and broaden applicability. A flexible variant was proposed by Ye [25], termed the MPEC generalized MFCQ (MPEC-GMFCQ), which allows a finer classification of equilibrium constraints.

Definition 2.9. (MPEC-GMFCQ) Let v^* be a feasible point of MPEC where all functions are continuously differentiable at v^*. We say that MPEC generalized MFCQ (MPEC-GMFCQ) holds at v^* if the following two conditions hold:

(*i*) For every partition of I_{GH} into sets P, Q, R with $R \neq \emptyset$, there exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{cases} \nabla g_i(v^*)^\top d \le 0, & \forall\, i \in I_g, \\ \nabla h_i(v^*)^\top d = 0, & \forall\, i \in [p], \\ \nabla G_i(v^*)^\top d = 0, & \forall\, i \in I_G \cup Q, \\ \nabla H_i(v^*)^\top d = 0, & \forall\, i \in I_H \cup P, \\ \nabla G_i(v^*)^\top d \ge 0,\ \nabla H_i(v^*)^\top d \ge 0, & \forall\, i \in R, \end{cases} \tag{2}$$

and for some $i \in R$, either $\nabla G_i(v^*)^\top d > 0$ or $\nabla H_i(v^*)^\top d > 0$.

(*ii*) For every partition of I_{GH} into sets P and Q, the gradient vectors

$$\nabla h_i(v^*),\ i \in [p], \quad \nabla G_i(v^*),\ \forall\, i \in I_G \cup Q, \quad \nabla H_i(v^*),\ \forall\, i \in I_H \cup P, \tag{3}$$

are linearly independent, and there exists $d \in \mathbb{R}^n$ such that

$$\begin{cases} \nabla g_i(v^*)^\top d < 0, & \forall\, i \in I_g, \\ \nabla h_i(v^*)^\top d = 0, & \forall\, i \in [p], \\ \nabla G_i(v^*)^\top d = 0, & \forall\, i \in I_G \cup Q, \\ \nabla H_i(v^*)^\top d = 0, & \forall\, i \in I_H \cup P. \end{cases} \tag{4}$$

Another useful CQ is the No Nonzero Abnormal Multiplier Constraint Qualification (NNAMCQ) [46].

Definition 2.10. (NNAMCQ) Let v^* be a feasible point of (MPEC) where all functions are continuously differentiable at v^*. NNAMCQ holds at v^* if there is no nonzero vector $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{m+p+2l}$ such that

$$\begin{cases} 0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(v^*) + \sum_{i \in [p]} \lambda_i^h \nabla h_i(v^*) - \sum_{i=1}^{l} (\lambda_i^G \nabla G_i(v^*) + \lambda_i^H \nabla H_i(v^*)), \\ \lambda_{I_g}^g \geq 0,\ \lambda_{I_H}^G = 0,\ \lambda_{I_G}^H = 0, \\ \text{either } \lambda_i^G > 0,\ \ \lambda_i^H > 0,\ \text{ or } \lambda_i^G \lambda_i^H = 0, \forall\ i \in I_{GH}. \end{cases} \tag{5}$$

MPEC-ACQ is a relatively relaxed CQ [20].

Definition 2.11. (MPEC-ACQ) Let v^* be a feasible point of (MPEC). MPEC-ACQ holds at v^* if

$$\mathcal{T}_{MPEC}^{\text{lin}}(v^*) = \mathcal{T}(v^*),$$

where $\mathcal{T}_{MPEC}^{\text{lin}}(v^*)$ is the MPEC-linearized tangent cone of (MPEC) at v^*, defined by

$$\mathcal{T}_{MPEC}^{\text{lin}}(v^*) = \left\{ d \in \mathbb{R}^n : \begin{array}{l} \nabla g_i(v^*)^\top d \leq 0,\ i \in I_g, \\ \nabla h_i(v^*)^\top d = 0,\ i \in [p], \\ \nabla G_i(v^*)^\top d = 0,\ i \in I_G, \\ \nabla H_i(v^*)^\top d = 0,\ i \in I_H, \\ \nabla G_i(v^*)^\top d \geq 0,\ i \in I_{GH}, \\ \nabla H_i(v^*)^\top d \geq 0,\ i \in I_{GH}, \\ \big(\nabla G_i(v^*)^\top d\big)\big(\nabla H_i(v^*)^\top d\big) \geq 0,\ i \in I_{GH} \end{array} \right\}.$$

Beyond traditional CQs, the calmness and Aubin property provide stability characterizations for the constraint system. Consider the function $F(v) = \big(g(v),\ \ h(v),\ \ \big(G(v),\ \ H(v)\big)\big)$ and the set $\Pi = \mathbb{R}_-^m \times \{0\}^p \times \mathcal{C}^l$ with $\mathcal{C} := \big\{(a, b) \in \mathbb{R}^2 \mid a \geq 0,\ b \geq 0,\ ab = 0\big\}$.

Problem (MPEC) is equivalent to

$$\min f(v) \quad \text{ s.t. } \quad F(v) \in \Pi. \tag{6}$$

Define the multifunction $\Phi : \mathbb{R}^{m+p+2l} \rightrightarrows \mathbb{R}^n$ associated with the constraint system of (6) by

$$\Phi(u) := \{z \in \mathbb{R}^n \mid F(z) + u \in \Pi\}. \tag{7}$$

Definition 2.12 (Calmness)**.** The multifunction $\Phi : \mathbb{R}^{m+p+2l} \rightrightarrows \mathbb{R}^n$ has a closed graph. At a point $(u, v) \in \operatorname{gph} \Phi := \{(u, v) \in \mathbb{R}^{m+p+2l} \times \mathbb{R}^n \mid v \in \Phi(u)\}$, Φ is calm provided there exist neighborhoods $\mathcal{U}$ of u, $\mathcal{V}$ of v, and a modulus $L \geq 0$ such that

$$\Phi\left(u'\right) \cap \mathcal{V} \subset \Phi(u) + L\|u' - u\|\mathbb{B}, \quad \forall\ u' \in \mathcal{U},$$

where $\mathbb{B} = \{w \in \mathbb{R}^n \mid \|w\| \leq 1\}$ is the closed unit ball.

Definition 2.13 (Aubin Property)**.** The multifunction $\Phi : \mathbb{R}^{m+p+2l} \rightrightarrows \mathbb{R}^n$ has the Aubin property at $(u, v) \in \operatorname{gph} \Phi$ if there exist neighborhoods U of u, V of v, and a modulus $M \geq 0$ such that

$$\Phi(u_1) \cap V \subset \Phi(u_2) + M\|u_1 - u_2\|\mathbb{B}, \quad \forall\, u_1, u_2 \in U.$$

2.3. Relationships Between MPEC Constraint Qualifications

Having introduced various constraint qualifications, we now examine their interconnections and hierarchy. There are two immediate implication results.

Proposition 2.14. [25] *At a feasible point v^* of* (MPEC)*, if MPEC-LICQ holds at v^*, then MPEC-MFCQ-T holds at v^*.*

Proposition 2.15. [25] *At a feasible point v^* of* (MPEC)*, if MPEC-MFCQ-T holds at v^*, then MPEC-GMFCQ holds at v^*.*

Next, comparing the tightened and relaxed formulations (Definitions 2.7 and 2.8) yields the following.

Proposition 2.16. *At a feasible point v^* of* (MPEC)*, it holds that:*

- (i) *If MPEC-MFCQ-T holds at v^*, then MPEC-MFCQ-R holds at v^*.*
- (ii) *If strict complementarity condition holds at v^*, MPEC-MFCQ-T is equivalent to MPEC-MFCQ-R.*

Proof. (i) is trivial. For (ii), note that if strict complementarity condition holds at v^*, $I_{GH} = \emptyset$. In that case, MPEC-MFCQ-R holds if the following set of vectors

$$\big\{\nabla g_i(v^*) \mid i \in I_g\big\} \cup \big\{\, \{\nabla h_i(v^*) \mid i \in [p]\} \cup \{\nabla G_i(v^*) \mid i \in I_G\} \cup \{\nabla H_i(v^*) \mid i \in I_H\}\,\big\}$$

is positively linearly independent. MPEC-MFCQ-T holds if the above set of vectors is positively linearly independent as well. Therefore, in this case, MPEC-MFCQ-T is equivalent to MPEC-MFCQ-R. □

We next relate MPEC-GMFCQ to NNAMCQ and to MPEC-MFCQ-R.

Proposition 2.17. [25, Proposition 2.1] *NNAMCQ is equivalent to MPEC-GMFCQ.*

The connection between MPEC-GMFCQ and MPEC-MFCQ-R is given below.

Proposition 2.18. *For a feasible point v^* of* (MPEC)*, it holds that:*

- (i) *If MPEC-GMFCQ holds at v^*, then MPEC-MFCQ-R holds at v^*.*
- (ii) *In particular, if strict complementarity condition holds at v^*, the two are equivalent.*

Proof. (i). It is not easy to show (i) from the definition of MPEC-GMFCQ in [25] and MPEC-MFCQ-R in Definition 2.8. Note that by Proposition 2.17, MPEC-GMFCQ is equivalent to NNAMCQ for (MPEC), we show (i) from the dual point of view; i.e., we show that NNAMCQ for (MPEC) implies MPEC-MFCQ-R. By definition, NNAMCQ holds at v^* is there if no nonzero vector $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{m+p+2l}$ such that

$$\begin{cases} 0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(v^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(v^*) - \sum_{i=1}^{l} (\lambda_i^G \nabla G_i(v^*) + \lambda_i^H \nabla H_i(v^*)), \\ \lambda_{I_g}^g \geq 0, \ \lambda_{I_H}^G = 0, \ \lambda_{I_G}^H = 0, \\ \text{either } \lambda_i^G > 0, \quad \lambda_i^H > 0, \ \text{or } \lambda_i^G \lambda_i^H = 0, \forall \ i \in I_{GH}. \end{cases} \tag{8}$$

Now we make a further partition of I_{GH} by defining the following sets.

$$\begin{aligned} P &:= \left\{ i \in I_{GH}, \ \lambda_i^G > 0, \ \lambda_i^H > 0 \right\}, \qquad & Q &:= \left\{ i \in I_{GH}, \ \lambda_i^G \neq 0, \ \lambda_i^H = 0 \right\}, \\ R &:= \left\{ i \in I_{GH}, \ \lambda_i^G = 0, \ \lambda_i^H \neq 0 \right\}, \qquad & S &:= \left\{ i \in I_{GH}, \ \lambda_i^G = 0, \ \lambda_i^H = 0 \right\}. \end{aligned}$$

Then NNAMCQ holds at v^*, if for every partition $P, \ Q, \ R, \ S$ of I_{GH}, there is no nonzero vector $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{m+p+2l}$ such that

$$\begin{cases} 0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(v^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(v^*) - \sum_{i \in P} (\lambda_i^G \nabla G_i(v^*) + \lambda_i^H \nabla H_i(v^*)) \\ \quad - \sum_{i \in Q} \lambda_i^G \nabla G_i(v^*) - \sum_{i \in R} \lambda_i^H \nabla H_i(v^*) - \sum_{i \in I_G} \lambda_i^G \nabla G_i(v^*) - \sum_{i \in I_H} \lambda_i^H \nabla H_i(v^*), \\ \lambda_{I_g}^g \geq 0, \quad \lambda_i^G > 0, \ \lambda_i^H > 0, \ i \in P, \quad \lambda_i^H = 0, \ i \in Q, \quad \lambda_i^G = 0, \ i \in R, \\ \lambda_i^G = 0, \ \lambda_i^H = 0, \ i \in S. \end{cases} \tag{9}$$

(9) reduces to the following conditions

$$\begin{cases} 0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(v^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(v^*) - \sum_{i \in P} (\lambda_i^G \nabla G_i(v^*) + \lambda_i^H \nabla H_i(v^*)) \\ \quad - \sum_{i \in Q} \lambda_i^G \nabla G_i(v^*) - \sum_{i \in R} \lambda_i^H \nabla H_i(v^*) - \sum_{i \in I_G} \lambda_i^G \nabla G_i(v^*) - \sum_{i \in I_H} \lambda_i^H \nabla H_i(v^*), \\ \lambda_{I_g}^g \geq 0, \quad \lambda_i^G > 0, \ \lambda_i^H > 0, \ i \in P. \end{cases}$$

On the other hand, recall that MPEC-MFCQ-R holds at v^* if MPEC-MFCQ-T holds at v^* for the relaxed nonlinear programming (RNLP(v^*)), MPEC-MFCQ-R holds at v^* if there is no nonzero vector $(\widehat{\lambda}^g, \widehat{\lambda}^h, \widehat{\lambda}^G, \widehat{\lambda}^H) \in \mathbb{R}^{m+p+2l}$ such that

$$\begin{cases} 0 = \sum_{i \in I_g} \widehat{\lambda}_i^g \nabla g_i(v^*) + \sum_{i=1}^{p} \widehat{\lambda}_i^h \nabla h_i(v^*) - \sum_{i \in I_{GH}} (\widehat{\lambda}_i^G \nabla G_i(v^*) + \widehat{\lambda}_i^H \nabla H_i(v^*)) \\ \quad - \sum_{i \in I_G} \widehat{\lambda}_i^G \nabla G_i(v^*) - \sum_{i \in I_H} \widehat{\lambda}_i^H \nabla H_i(v^*), \\ \widehat{\lambda}_{I_g}^g \geq 0, \ \widehat{\lambda}_i^G \geq 0, \quad \widehat{\lambda}_i^H \geq 0, \ i \in I_{GH}. \end{cases} \tag{10}$$

Similarly, we make a partition for I_{GH} as follows

$$\widehat{P} := \left\{ i \in I_{GH},\ \widehat{\lambda}_i^G > 0,\ \widehat{\lambda}_i^H > 0 \right\}, \qquad \widehat{Q} := \left\{ i \in I_{GH},\ \widehat{\lambda}_i^G \neq 0,\ \widehat{\lambda}_i^H = 0 \right\},$$
$$\widehat{R} := \left\{ i \in I_{GH},\ \widehat{\lambda}_i^G = 0,\ \widehat{\lambda}_i^H \neq 0 \right\}, \qquad \widehat{S} := \left\{ i \in I_{GH},\ \widehat{\lambda}_i^G = 0,\ \widehat{\lambda}_i^H = 0 \right\}.$$

Then MPEC-MFCQ-R holds at v^*, if for every partition $\widehat{P},\ \widehat{Q},\ \widehat{R},\ \widehat{S}$ of I_{GH}, there is no nonzero vector $(\widehat{\lambda}^q, \widehat{\lambda}^h, \widehat{\lambda}^G, \widehat{\lambda}^H) \in \mathbb{R}^{m+p+2l}$ that solves the following

$$\begin{cases} 0 = \sum_{i \in I_g} \widehat{\lambda}_i^g \nabla g_i(v^*) + \sum_{i=1}^{p} \widehat{\lambda}_i^h \nabla h_i(v^*) - \sum_{i \in \widehat{P}} (\widehat{\lambda}_i^G \nabla G_i(v^*) + \widehat{\lambda}_i^H \nabla H_i(v^*)) \\ \quad - \sum_{i \in \widehat{Q}} \widehat{\lambda}_i^G \nabla G_i(v^*) - \sum_{i \in \widehat{R}} \widehat{\lambda}_i^H \nabla H_i(v^*) - \sum_{i \in I_G} \widehat{\lambda}_i^G \nabla G_i(v^*) - \sum_{i \in I_H} \widehat{\lambda}_i^H \nabla H_i(v^*), \\ \widehat{\lambda}_{I_g}^g \geq 0,\ \ \widehat{\lambda}_i^G > 0,\ \widehat{\lambda}_i^H > 0,\ i \in P,\ \ \widehat{\lambda}_i^H = 0,\ i \in Q,\ \ \widehat{\lambda}_i^G = 0,\ i \in R, \\ \widehat{\lambda}_i^G = 0,\ \widehat{\lambda}_i^H = 0,\ i \in S. \end{cases} \tag{11}$$

Therefore, it can be seen that if for every partition $\widehat{P},\ \widehat{Q},\ \widehat{R},\ \widehat{S}$ of I_{GH} of I_{GH}, no nonzero vector $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{m+p+2l}$ satisfies (9), then for every partition $\widehat{P},\ \widehat{Q},\ \widehat{R},\ \widehat{S}$ of I_{GH}, no nonzero vector $(\widehat{\lambda}^q, \widehat{\lambda}^h, \widehat{\lambda}^G, \widehat{\lambda}^H) \in \mathbb{R}^{m+p+2l}$ satisfies (11). In other words, if NNAMCQ holds at v^*, MPEC-MFCQ-R holds at v^*. (i) is proved.

(ii). If $I_{GH} = \emptyset$, the definition of MPEC-GMFCQ reduces to the following, the gradient vectors

$$\nabla h_i(v^*),\ i \in [p],\ \nabla G_i(v^*),\ \forall\ i \in I_G,\ \nabla H_i(v^*),\ \forall\ i \in I_H,$$

are linearly independent and there exists $d \in \mathbb{R}^n$ such that

$$\begin{cases} \nabla g_i(v^*)^\top d < 0, & \forall\ i \in I_g, \\ \nabla h_i(v^*)^\top d = 0, & \forall\ i \in [p], \\ \nabla G_i(v^*)^\top d = 0, & \forall\ i \in I_G, \\ \nabla H_i(v^*)^\top d = 0, & \forall\ i \in I_H, \end{cases}$$

which is the definition of MPEC-MFCQ-R for $I_{GH} = \emptyset$. Therefore, if $I_{GH} = \emptyset$, MPEC-GMFCQ is equivalent to MPEC-MFCQ-R. □

We finally connect these CQs to stability properties of the associated feasibility mapping.

Proposition 2.19. [47] *At a feasible point v^* of* (MPEC)*, if MPEC-GMFCQ holds at v^*, then Φ is calm at $(0, v^*)$.*

Proposition 2.20. [48, Corollary 4.4] *MPEC-GMFCQ implies the Aubin property of Φ around $(0, v^*)$.*

Proposition 2.21. [49] *If the multifunction Φ has the Aubin property around $(0, v^*)$, then Φ is calm at $(0, v^*)$.*

Proposition 2.22. [47] *If the multifunction Φ is calm at $(0, v^*)$, then MPEC-ACQ holds at v^*.*

Next, we provide illustrative counterexamples that clarify how the main constraint qualifications for MPECs relate to each other.
E1. Consider the MPEC

$$\begin{array}{ll} \min\limits_{v \in \mathbb{R}^3} & f(v) = v_2 + v_3 \\ \text{s.t.} & g_1(v) = v_1 \leq 0, \\ & g_2(v) = v_1 + v_2 \leq 0, \\ & G(v) = v_2 \geq 0, \\ & H(v) = v_3 \geq 0, \\ & G(v) \cdot H(v) = v_2 v_3 = 0. \end{array}$$

At $v^* = (0,0,0)$, we have $I_g = \{1,2\}$, $I_G = I_H = \emptyset$, and $I_{GH} = \{1\}$. The corresponding tightened nonlinear program TNLP at v^* is

$$\begin{array}{ll} \min\limits_{v \in \mathbb{R}^3} & v_2 + v_3 \\ \text{s.t.} & g_1(v) = v_1 \leq 0, \\ & g_2(v) = v_1 + v_2 \leq 0, \\ & G(v) = v_2 = 0, \\ & H(v) = v_3 = 0. \end{array}$$

At $v^* = (0,0,0)$, the gradients of the active constraints are $\nabla g_1(v^*)^\top = (1,0,0)$, $\nabla g_2(v^*)^\top = (1,1,0)$, $\nabla G(v^*)^\top = (0,1,0)$, and $\nabla H(v^*)^\top = (0,0,1)$. The set $\{\nabla g_1(v^*), \nabla g_2(v^*)\} \cup \{\nabla G(v^*), \nabla H(v^*)\}$ is *positively linearly independent* but *linearly dependent*. Therefore, MFCQ holds for the TNLP at $v^* = (0,0,0)$, while LICQ does not. Consequently, MPEC-MFCQ-T holds at $v^* = (0,0,0)$, and MPEC-LICQ does not.

E2. Consider the MPEC

$$\begin{array}{ll} \min\limits_{v \in \mathbb{R}^2} & f(v) = v_1 + v_2 \\ \text{s.t.} & g_1(v) = -v_1 - v_2 \leq 0, \\ & G(v) = v_1 \geq 0, \\ & H(v) = v_2 \geq 0, \\ & G(v)H(v) = v_1 v_2 = 0. \end{array}$$

At $v^* = (0,0)$, we have $I_g = \{1\}$, $I_G = \emptyset$, $I_H = \emptyset$, and $I_{GH} = \{1\}$. The gradients are $\nabla g_1(v^*)^\top = (-1,-1)$, $\nabla G(v^*)^\top = (1,0)$, and $\nabla H(v^*)^\top = (0,1)$. By Definition 2.10, there is no nonzero vector $(\lambda^g, \lambda^G, \lambda^H) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ such that

$$\begin{cases} 0 = \lambda^g \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \lambda^G \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda^H \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\lambda^g - \lambda^G \\ -\lambda^g - \lambda^H \end{bmatrix}, \\ \lambda^g \geq 0, \\ \text{either } \lambda^G > 0,\ \lambda^H > 0,\ \text{or } \lambda^G \lambda^H = 0. \end{cases}$$

Therefore, NNAMCQ is satisfied—and equivalently, MPEC-GMFCQ holds. However, the set $\{\nabla g_1(v^*)\} \cup \{\nabla G(v^*), \nabla H(v^*)\}$ is *positively linearly dependent*, so MPEC-

MFCQ-T is not satisfied.

E3. Consider the MPEC

$$\begin{array}{ll} \min\limits_{v\in\mathbb{R}^3} & f(v) = v_1 + v_2 + v_3 \\ \text{s.t.} & G_1(v) = v_1 \geq 0, \\ & H_1(v) = v_2 \geq 0, \\ & G_2(v) = v_1 - v_2^2 \geq 0, \\ & H_2(v) = v_3 \geq 0, \\ & G(v)^\top H(v) = 0. \end{array}$$

At $v^* = (0,0,0)$, we have $I_{GH} = \{1,2\}$ and $I_G = I_H = \emptyset$. The corresponding RNLP at v^* is

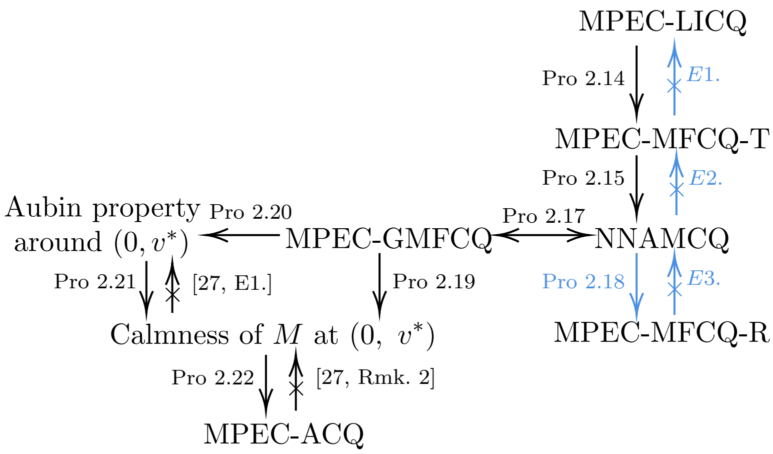
$$\begin{array}{ll} \min\limits_{v\in\mathbb{R}^3} & f(v) = v_1 + v_2 + v_3 \\ \text{s.t.} & G_1(v) = v_1 \geq 0, \quad H_1(v) = v_2 \geq 0, \\ & G_2(v) = v_1 - v_2^2 \geq 0, \quad H_2(v) = v_3 \geq 0. \end{array}$$

At $v^* = (0,0,0)$, the active gradients are $\nabla G_1(v^*)^\top = (1,0,0)$, $\nabla H_1(v^*)^\top = (0,1,0)$, $\nabla G_2(v^*)^\top = (1,0,0)$, and $\nabla H_2(v^*)^\top = (0,0,1)$. The set $\{\nabla G_1(v^*), \nabla H_1(v^*), \nabla G_2(v^*), \nabla H_2(v^*)\}$ is *positively linearly independent*, so MPEC-MFCQ-R holds. However, there exists a nonzero vector $(\lambda_1^G, \lambda_1^H, \lambda_2^G, \lambda_2^H) = (1,0,-1,0)$ satisfying

$$\begin{cases} 0 = -\lambda_1^G \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \lambda_1^H \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \lambda_2^G \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \lambda_2^H \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\lambda_1^G - \lambda_2^G \\ -\lambda_1^H \\ -\lambda_2^H \end{bmatrix}, \\ \text{either } \lambda_i^G > 0, \; \lambda_i^H > 0, \; \text{or } \lambda_i^G \lambda_i^H = 0, \quad i \in I_{GH}, \end{cases}$$

hence NNAMCQ is not satisfied.

Based on these properties and counterexamples, we can summarize the implications among the CQs at a feasible point v^* of (MPEC) in a relationship diagram:



Unlike classical NLP, where suitable CQs make KKT conditions necessary, MPECs demand refined stationarity concepts to accommodate equilibrium constraints.

Definition 2.23. Let v^* be feasible for (MPEC). Then v^* is

(a) *weakly stationary* if there exist multipliers $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$, $\gamma,\ \nu \in \mathbb{R}^l$ such that

$$\nabla f(v^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(v^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(v^*) - \sum_{i=1}^{l} \gamma_i \nabla G_i(v^*) - \sum_{i=1}^{l} \nu_i \nabla H_i(v^*) = 0,$$

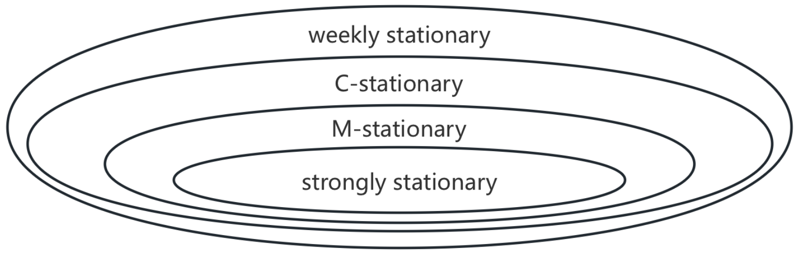
with $\lambda_i \geq 0$, $\lambda_i g_i(v^*) = 0$ for $i = 1, \ldots, m$, and $\gamma_i = 0$ for $i \in I_H$, $\nu_i = 0$ for $i \in I_G$;

(b) *C-stationary* if it is weakly stationary and $\gamma_i \nu_i \geq 0$ for all $i \in I_{GH}$;

(c) *M-stationary* if it is weakly stationary and, for all $i \in I_{GH}$, either $\gamma_i > 0$ and $\nu_i > 0$, or $\gamma_i \nu_i = 0$;

(d) *strongly stationary* if it is weakly stationary and $\gamma_i \geq 0$, $\nu_i \geq 0$ for all $i \in I_{GH}$.

It has been proven that strong stationarity is equivalent to the standard KKT conditions of an MPEC. Moreover, strong stationarity implies M-stationarity, M-stationarity implies C-stationarity, and C-stationarity further implies weak stationarity [28]. However, even for very simple MPECs, strong stationarity may fail to hold at a global minimum [23]. Therefore, under suitable assumptions, M-stationarity is generally the strongest stationarity concept that a local minimum can satisfy. Based on the above introduction, the relationships among various stable points are summarized as follows:



Theorem 2.24. [23] *Let v^* be a local solution of MPEC. Suppose that MPEC-LICQ is satisfied at v^*. Then v^* is strongly stationary.*

Theorem 2.25. [25, Theorem 3.1] *Let v^* be a local solution of MPEC. If one of the MPEC constraint qualifications, such as MPEC-ACQ, NNAMCQ, or MPEC-MFCQ-T, holds at v^*, then v^* is M-stationary.*

Theorem 2.26. [32, Theorem 3] *Let v^* be a local minimizer of MPEC. If the multifunction Φ is calm at $(0, v^*)$, then v^* is M-stationary.*

3. Applications to Bilevel Hyperparameter Optimization

In this section, we briefly review the BHO for L1-loss SVC and its corresponding MPEC reformulation [40], and then analyze the relevant CQs for the resulting model.

The BHO for L1-loss SVC is

$$
\begin{aligned}
\min_{C\in\mathbb{R},\ w^t\in\mathbb{R}^p,\ t=1,\ldots,T} \quad & \frac{1}{T}\sum_{t=1}^{T}\frac{1}{m_1}\sum_{i\in\mathcal{N}_t}\left\|\left(-y_i\, x_i^\top w^t\right)_+\right\|_0 \\
\text{s.t.} \quad & C\geq 0, \\
& \text{and for each } t=1,\ldots,T: \\
& \quad w^t\in\operatorname*{argmin}_{w\in\mathbb{R}^p}\left\{\frac{1}{2}\|w\|_2^2 + C\sum_{i\in\overline{\mathcal{N}}_t}\left(1-y_i\, x_i^\top w\right)_+\right\}.
\end{aligned}
\tag{12}
$$

Here, the expression $\sum_{i\in\mathcal{N}_t}\|\left(-y_i\left(x_i^\top w^t\right)\right)_+\|_0$ basically counts the number of data points that are misclassified in the validation set Ω_t, while the outer summation (i.e., the objective function in (12)) averages the misclassification error over all the folds. It is evident that $\|(\cdot)_+\|_0$ is both discontinuous and nonconvex. However, as demonstrated by Mangasarian [50], this function can be expressed as the minimum of the sum of all elements in the solution to the following linear optimization problem

$$
\left\|r_+\right\|_0 = \left\{\min\sum_{i=1}^{m_1}\zeta_i \ : \ \zeta = \operatorname*{argmin}_{u}\left\{-u^\top r \ : \ \mathbf{0}\leq u\leq\mathbf{1}\right\}\right\}. \tag{13}
$$

In the special case where $r_i=0$, the function $\|(r_i)_+\|_0$ evaluates to 0. However, the solution to the corresponding linear programming problem is not unique and can take any value within the interval $[0,1]$. To maintain the robustness and consistency of formulation (13), we assume throughout this paper that the SVC model assigns a distinct classification to all validation data points. More precisely, for each validation data point x_i, where $i\in[m_1]$, we impose the condition $x_i^\top w^t\neq 0$, where w^t is the solution for the lower level problem in (12), $t\in[T]$.

Let $Q_u=\{1,2,\ldots,Tm_1\}$ and $Q_l=\{1,2,\ldots,Tm_2\}$ be the stacked index sets for the validation and training samples, respectively (with T folds and per-fold sizes m_1 and m_2). Replacing the lower level with its KKT conditions and using the reformulation of $\|\cdot\|_0$, problem (12) can be reformulated as follows

$$
\begin{aligned}
\min_{\substack{C\in\mathbb{R}\\ \zeta\in\mathbb{R}^{Tm_1},\ z\in\mathbb{R}^{Tm_1}\\ \alpha\in\mathbb{R}^{Tm_2},\ \xi\in\mathbb{R}^{Tm_2}}} \quad & \frac{1}{Tm_1}\mathbf{1}^\top\zeta \\
\text{s.t.} \quad & \mathbf{0}\leq\zeta \ \perp\ AB^\top\alpha+z\ \geq\mathbf{0}, \\
& \mathbf{0}\leq z\ \perp\ \mathbf{1}-\zeta\ \geq\mathbf{0}, \\
& \mathbf{0}\leq\alpha\ \perp\ BB^\top\alpha-\mathbf{1}+\xi\ \geq\mathbf{0}, \\
& \mathbf{0}\leq\xi\ \perp\ C\mathbf{1}-\alpha\ \geq\mathbf{0},
\end{aligned}
\tag{14}
$$

where $\zeta\in\mathbb{R}^{Tm_1}$, $z\in\mathbb{R}^{Tm_1}$, $A\in\mathbb{R}^{Tm_1\times Tp}$, $\alpha\in\mathbb{R}^{Tm_2}$, $\xi\in\mathbb{R}^{Tm_2}$, $w\in\mathbb{R}^{Tp}$, and

$B \in \mathbb{R}^{Tm_2 \times Tp}$ are given by

$$\zeta := \begin{bmatrix} \zeta^1 \\ \zeta^2 \\ \vdots \\ \zeta^T \end{bmatrix}, \quad z := \begin{bmatrix} z^1 \\ z^2 \\ \vdots \\ z^T \end{bmatrix}, \quad A := \begin{bmatrix} A^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & A^2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & A^T \end{bmatrix},$$

$$\alpha := \begin{bmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^T \end{bmatrix}, \quad \xi := \begin{bmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^T \end{bmatrix}, \quad w := \begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^T \end{bmatrix}, \quad B := \begin{bmatrix} B^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B^2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B^T \end{bmatrix},$$

with

$$A^t = \begin{bmatrix} y_{t_1} x_{t_1}^\top \\ \vdots \\ y_{t_{m_1}} x_{t_{m_1}}^\top \end{bmatrix} \in \mathbb{R}^{m_1 \times p}, \quad (x_{t_k}, y_{t_k}) \in \Omega_t, \qquad B^t = \begin{bmatrix} y_{t_{m_1+1}} x_{t_{m_1+1}}^\top \\ \vdots \\ y_{t_{l_1}} x_{t_{l_1}}^\top \end{bmatrix} \in \mathbb{R}^{m_2 \times p}, \quad (x_{t_k}, y_{t_k}) \in \overline{\Omega}_t.$$

Stacking variables leads to the compact MPEC

$$\begin{aligned} &\min_{v \in \mathbb{R}^n} f(v) \\ &\text{s.t.} \quad \mathbf{0} \leq H(v) \perp G(v) \geq \mathbf{0}, \end{aligned} \tag{BHO-SVM}$$

where $v = \left[C,\ \zeta^\top,\ z^\top,\ \alpha^\top,\ \xi^\top\right]^\top \in \mathbb{R}^n$ with $n = 2T(m_1 + m_2) + 1$, and

$$f(v) = c^\top v, \qquad H(v) = \mathcal{Q}v, \qquad G(v) = \mathcal{P}v + a, \tag{15}$$

with

$$c = \frac{1}{Tm_1} \begin{bmatrix} 0 \\ \mathbf{1}_{Tm_1} \\ \mathbf{0}_{Tm_1} \\ \mathbf{0}_{Tm_2} \\ \mathbf{0}_{Tm_2} \end{bmatrix} \in \mathbb{R}^n, \quad a = \begin{bmatrix} \mathbf{0}_{Tm_1} \\ \mathbf{1}_{Tm_1} \\ -\mathbf{1}_{Tm_2} \\ \mathbf{0}_{Tm_2} \end{bmatrix} \in \mathbb{R}^{n-1}, \quad \mathcal{Q} = \begin{bmatrix} \mathbf{0}_{n-1} & I_{n-1} \end{bmatrix} \in \mathbb{R}^{(n-1) \times n},$$

$$\mathcal{P} = \begin{bmatrix} \mathbf{0}_{Tm_1} & \mathbf{0}_{Tm_1 \times Tm_1} & I_{Tm_1} & AB^\top & \mathbf{0}_{Tm_1 \times Tm_2} \\ \mathbf{0}_{Tm_1} & -I_{Tm_1} & \mathbf{0}_{Tm_1 \times Tm_1} & \mathbf{0}_{Tm_1 \times Tm_2} & \mathbf{0}_{Tm_1 \times Tm_2} \\ \mathbf{0}_{Tm_2} & \mathbf{0}_{Tm_2 \times Tm_1} & \mathbf{0}_{Tm_2 \times Tm_1} & BB^\top & I_{Tm_2} \\ \mathbf{1}_{Tm_2} & \mathbf{0}_{Tm_2 \times Tm_1} & \mathbf{0}_{Tm_2 \times Tm_1} & -I_{Tm_2} & \mathbf{0}_{Tm_2 \times Tm_2} \end{bmatrix} \in \mathbb{R}^{(n-1) \times n}.$$

Compared with (MPEC), the MPEC in (BHO-SVM) has the following properties.

Theorem 3.1. *For* (BHO-SVM)*,*

(i) MPEC-ACQ holds at all feasible points.
(ii) MPEC-MFCQ-T holds at a feasible point v^* if and only if MPEC-LICQ holds at v^*.

Proof.

(i) By Theorem in [20], MPEC-ACQ, holds for all feasible points for (BHO-SVM), since all fuctions in (BHO-SVM) are affine linear.

(ii) Since (BHO-SVM) does not have inequality constraints g, according to Definition 2.7 and Definition 2.6, MPEC-LICQ and MPEC-MFCQ-T are equivalent for (BHO-SVM) at a feasible point v.

□

Moreover, by Theorem 3.1, we can obtain that MPEC-MFCQ-T and the strongest constraint qualification, MPEC-LICQ, are equivalent for the problem (12). Next, we will mainly discuss the fulfillment of MPEC-MFCQ-R and MPEC-LICQ for the MPEC (BHO-SVM).

Theorem 3.2. *Let v^* be a feasible point of problem* (BHO-SVM). *If $\big(BB^\top\big)_{(\Lambda_1\cup\Lambda_3,\Lambda_1\cup\Lambda_3)}$ is positive definite at v^*, then MPEC-MFCQ-R automatically holds at v^*, where Λ_1 and Λ_3 are defined by*

$$\begin{aligned}
\Lambda_1 &:= \Big\{i \in Q_l \ \mid \ \alpha_i = 0,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i = 0\Big\},\\
\Lambda_3 &:= \Big\{i \in Q_l \ \mid \ 0 < \alpha_i \le C,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i = 0\Big\}.
\end{aligned}$$

To state the MPEC-LICQ result, we introduce the index sets

$$\begin{aligned}
I_{GH_3} &:= \Big\{i \in Q_l \ \mid \ \alpha_i = 0,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0\Big\},\\
I_{GH_4} &:= \{i \in Q_l \ \mid \ \xi_i = 0,\ C - \alpha_i = 0\},\\
\Lambda_3^+ &:= \Big\{i \in Q_l \ \mid \ 0 < \alpha_i < C,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i = 0\Big\},\\
\Lambda_3^c &:= \Big\{i \in Q_l \ \mid \ \alpha_i = C,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i = 0\Big\},\\
\Lambda_u &:= \Big\{i \in Q_l \ \mid \ \alpha_i = C,\ (BB^\top\alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i > 0\Big\}.
\end{aligned}$$

Theorem 3.3. *Let v^* be a feasible point of* (BHO-SVM).

(i) *If $|I_{GH}| > 1$, then MPEC-LICQ fails at v^*.*

(ii) *If $|I_{GH}| = 0$, then MPEC-LICQ holds at v^*.*

(iii) *If $|I_{GH_3}| = 1$, $|I_{GH_4}| = 0$, $\widehat{a}_3 \neq 0$, and $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ is positive definite at v^**[1]*, then MPEC-LICQ holds at v^*, where*

$$\widehat{a}_3 = (BB^\top)_{(I_{GH_3},\Lambda_u)}\mathbf{1}_{|\Lambda_u|} - A_3(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}, \quad A_3 = (BB^\top)_{(I_{GH_3},\Lambda_3^+)}(BB^\top)^{-1}_{(\Lambda_3^+,\Lambda_3^+)}.$$

(iv) *If $|I_{GH_4}| = 1$, $|I_{GH_3}| = 0$, $\widehat{a}_4 \neq 0$, and $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ is positive definite at v^**[2]*, then MPEC-LICQ holds at v^*, where*

$$\widehat{a}_4 = (BB^\top)_{(\Lambda_3^c,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|} - A_4(BB^\top)_{(\Lambda_3^+,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|}, \quad A_4 = (BB^\top)_{(\Lambda_3^c,\Lambda_3^+)}(BB^\top)^{-1}_{(\Lambda_3^+,\Lambda_3^+)}.$$

(v) *Otherwise, MPEC-LICQ fails at v^*.*

---

[1] If $\Lambda_3^+ = \emptyset$, then $A_3 = 0$.

[2] We follow a convention for A_4 similar to A_3.

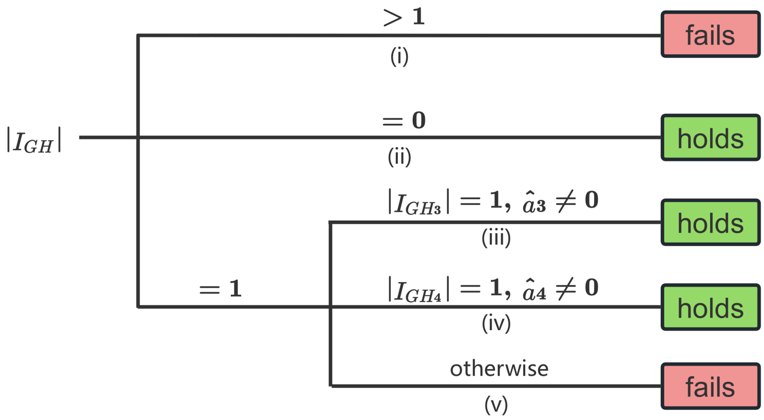


Figure 1.: Summary of the fulfillment of MPEC-LICQ.

Remark 2. Note that $|I_{GH}| = 0$ indicates that strict complementarity conditions hold at v^*. From the numerical point of view, given a specific feasible point v^* of (BHO-SVM), it may be very often that strict complementarity conditions hold at v^*, since it may be very rare that numerically two terms $H_i(v)$ and $G_i(v)$ are zero at the same time.

4. Conclusions

In this paper, we introduce several specific formulations of constraint qualifications (such as MPEC-LICQ, MPEC-MFCQ-T, MPEC-MFCQ-R, etc.) developed for MPECs, and provide a thorough theoretical analysis of their interrelationships. Then we address the hyperparameter selection problem for support vector classification, reformulating it as a mathematical program with equilibrium constraints (MPEC) and analyzing the relevant constraint qualifications (CQs). The study highlights the role of positively linearly independence and demonstrates the equivalence of MPEC-MFCQ-T and MPEC-LICQ for this problem. Furthermore, we establish refined conditions under which MPEC-LICQ holds and identify specific cases where it may fail. The core contribution lies in providing a more robust theoretical foundation for complex bilevel optimization models, especially in the context of hyperparameter optimization in modern machine learning. Future research could explore the extension of the applicability of MPEC-related CQs and further investigate their properties in more complex models.

References

[1] Bard JF. Practical bilevel optimization: algorithms and applications. Springer Science & Business Media, 2013.

[2] Dempe S. Bilevel optimization: theory, algorithms and applications. Freiberg, Germany: TU Bergakademie Freiberg, Fakultät für Mathematik und Informatik; 2018.

[3] Dempe S, Zemkoho A. Bilevel optimization. In: Springer Optimization and Its Applications. Vol. 161. Springer; 2020.

[4] Dempe S, Kalashnikov V, Pérez-Valdés GA, et al. Bilevel programming problems. Energy Systems. Springer, Berlin, 2015; 10(978-3): 53-56.
[5] Sinha A, Malo P, Deb K. A review on bilevel optimization: From classical to evolutionary approaches and applications. IEEE transactions on evolutionary computation. 2017; 22(2): 276-295.
[6] Wogrin S, Pineda S, Tejada-Arango DA. Applications of bilevel optimization in energy and electricity markets. Bilevel Optimization: Advances and Next Challenges. 2020; 139-168.
[7] MacKay M, Vicol P, Lorraine J, et al. Self-tuning networks: Bilevel optimization of hyperparameters using structured best-response functions. 2019; arXiv preprint arXiv:1903.03088.
[8] Okuno T, Takeda A, Kawana A, et al. On ℓ_p-hyperparameter learning via bilevel nonsmooth optimization. Journal of Machine Learning Research. 2021; 22(245): 1-47.
[9] Liu H, Simonyan K, Yang Y. Darts: Differentiable architecture search. 2018; arXiv preprint arXiv:1806.09055.
[10] Lian D, Zheng Y, Xu Y, et al. Towards fast adaptation of neural architectures with meta learning. International Conference on Learning Representations. 2020.
[11] Zhang H, Chen W, Huang Z, et al. Bi-level actor-critic for multi-agent coordination. Proceedings of the AAAI Conference on Artificial Intelligence. 2020; 34(05): 7325-7332.
[12] Yang Z, Chen Y, Hong M, et al. Provably global convergence of actor-critic: A case for linear quadratic regulator with ergodic cost. Advances in neural information processing systems. 2019; 32.
[13] Liu R, Li Z, Zhang Y, et al. Bi-level probabilistic feature learning for deformable image registration. Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence. 2021; 723-730.
[14] Ma L, Jin D, An N, et al. Bilevel fast scene adaptation for low-light image enhancement. International Journal of Computer Vision. 2023; 1-19.
[15] Dempe S, Franke S. On the solution of convex bilevel optimization problems. Computational Optimization and Applications. 2016; 63: 685-703.
[16] Luo ZQ, Pang JS, Ralph D. Mathematical programs with equilibrium constraints. Cambridge University Press. 1996.
[17] Dempe S, Zemkoho A. The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. Mathematical Programming. 2013; 138: 447-473.
[18] Dempe S, Zemkoho A. On the Karush–Kuhn–Tucker reformulation of the bilevel optimization problem. Nonlinear Analysis: Theory, Methods & Applications. 2012; 75(3): 1202-1218.
[19] Ye JJ, Zhu DL, Zhu QJ. Exact penalization and necessary optimality conditions for generalized bilevel programming problems. SIAM Journal on optimization. 1997; 7(2): 481-507.
[20] Flegel ML, Kanzow C. Abadie-type constraint qualification for mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications. 2005; 124(3): 595-614.
[21] Ramos A. Mathematical programs with equilibrium constraints: a sequential optimality condition, new constraint qualifications and algorithmic consequences. Optimization Methods and Software. 2021; 36(1): 45-81.
[22] Pang JS and Fukushima M. Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. Computational Optimization and Applications. 1999; 13: 111-136.
[23] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. Mathematics of Operations Research, 2000, 25(1): 1-22.
[24] Ralph* D, Wright SJ. Some properties of regularization and penalization schemes for MPECs. Optimization Methods and Software. 2004; 19(5): 527-556.
[25] Jane JY. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. Journal of Mathematical Analysis and Applications. 2005; 307(1): 350-369.
[26] Flegel ML, Kanzow C. A direct proof for M-stationarity under MPEC-GCQ for mathe-

matical programs with equilibrium constraints. Optimization with Multivalued Mappings: Theory, Applications, and Algorithms. Boston, MA: Springer US, 2006; 111-122.

[27] Henrion R, Outrata JV. Calmness of constraint systems with applications. Mathematical Programming. 2005; 104(2): 437-464.

[28] Flegel ML, Kanzow C. On the Guignard constraint qualification for mathematical programs with equilibrium constraints. Optimization. 2005; 54(6): 517-534.

[29] Outrata JV. Optimality conditions for a class of mathematical programs with equilibrium constraints. Mathematics of operations research. 1999; 24(3): 627-644.

[30] Outrata JV. A generalized mathematical program with equilibrium constraints. SIAM Journal on Control and Optimization. 2000; 38(5): 1623-1638.

[31] Flegel ML, Kanzow C. On M-stationary points for mathematical programs with equilibrium constraints. Journal of Mathematical Analysis and Applications. 2005; 310(1): 286-302.

[32] Flegel ML, Kanzow C, Outrata JV. Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. Set-Valued Analysis. 2007; 15(2): 139-162.

[33] Scholtes S. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM Journal on Optimization. 2001; 11(4): 918-936.

[34] Fukushima M, Pang J S. Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. Ill-posed Variational Problems and Regularization Techniques: Proceedings of the "Workshop on Ill-Posed Variational Problems and Regulation Techniques" held at the University of Trier, September 3–5, 1998. Springer Berlin Heidelberg, 1999; 99-110.

[35] Bennett KP, Hu J, Ji X, et al. Model selection via bilevel optimization. The 2006 IEEE International Joint Conference on Neural Network Proceedings. IEEE, 2006; 1922-1929.

[36] Bennett KP, Kunapuli G, Hu J, et al. Bilevel optimization and machine learning. IEEE world congress on computational intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008; 25-47.

[37] Kunapuli G, Bennett KP, Hu J, et al. Classification model selection via bilevel programming. Optimization Methods & Software. 2008; 23(4): 475-489.

[38] Kunapuli G, Bennett KP, Hu J, et al. Bilevel model selection for support vector machines. Data Mining and Mathematical Programming. 2008; 45: 129–158.

[39] Li Z, Qian Y, Li Q. A unified framework and a case study for hyperparameter selection in machine learning via bilevel optimization. 5th International Conference on Data Science and Information Technology (DSIT), 2022, pp. 1-8.

[40] Li Q, Li Z, Zemkoho A. Bilevel hyperparameter optimization for support vector classification: theoretical analysis and a solution method. Mathematical Methods of Operations Research. 2022; 96(3): 315-350.

[41] Qian Y, Li Q, Zemkoho A. Global relaxation-based LP-Newton method for multiple hyperparameter selection in support vector classification with feature selection. arXiv preprint arXiv:2312.10848, 2023.

[42] Coniglio S, Dunn S, Li Q, et al. Bilevel hyperparameter optimization for nonlinear support vector machines. Optimization Online, 2023, pp. 1-78.

[43] Ward S, Zemkoho A, Ahipasaoglu S. Mathematical programs with complementarity constraints and application to hyperparameter tuning for nonlinear support vector machines, manuscript, 2025.

[44] Wang Y, Li Q. A fast smoothing Newton method for bilevel hyperparameter optimization for SVC with Logistic loss. Optimization, 2024, DOI: 10.1080/02331934.2024.2394612.

[45] Hoheisel T, Kanzow C, Schwartz A. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Mathematical Programming. 2013; 137: 257-288.

[46] Ye JJ. Multiplier rules under mixed assumptions of differentiability and Lipschitz continuity. SIAM Journal on Control and Optimization, 2000, 39(5): 1441-1460.

[47] Kanzow C, Schwartz A. Mathematical programs with equilibrium constraints: Enhanced Fritz John conditions, new constraint qualifications, and improved exact penalty results.

SIAM Journal on Optimization. 2010; 20(5): 2730-2753.
[48] Mordukhovich B. Lipschitzian stability of constraint systems and generalized equations. Nonlinear Analysis: Theory, Methods & Applications. 1994; 22(2): 173-206.
[49] Henrion R. Calmness as a constraint qualification for M-stationarity conditions in MPECs. Generalized Nash Equilibrium Problems, Bilevel Programming and MPEC. Singapore: Springer Singapore, 2018: 21-41.
[50] Mangasarian OL. Misclassification minimization. Journal of Global Optimization. 1994; 5(4): 309-323.

Disclosure statement: There are no relevant financial or non-financial competing interests to report.

Appendix A. Proof of Theorem 3.2

Recall $I_H := \overset{4}{\underset{k=1}{\cup}} I_{H_k}$, $I_G := \overset{4}{\underset{k=1}{\cup}} I_{G_k}$ and $I_{GH} := \overset{4}{\underset{k=1}{\cup}} I_{GH_k}$ in [40], where

$$I_{H_1} := \left\{ i \in Q_u \mid \zeta_i = 0,\ (AB^\top \alpha + z)_i > 0 \right\},\ I_{H_2} := \left\{ i \in Q_u \mid z_i = 0,\ 1 - \zeta_i > 0 \right\},$$
$$I_{H_3} := \left\{ i \in Q_l \mid \alpha_i = 0,\ (BB^\top \alpha - \mathbf{1} + \xi)_i > 0 \right\},\ I_{H_4} := \left\{ i \in Q_l \mid \xi_i = 0,\ C - \alpha_i > 0 \right\},$$
$$I_{G_1} := \left\{ i \in Q_u \mid \zeta_i > 0,\ (AB^\top \alpha + z)_i = 0 \right\},\ I_{G_2} := \left\{ i \in Q_u \mid z_i > 0,\ 1 - \zeta_i = 0 \right\},$$
$$I_{G_3} := \left\{ i \in Q_l \mid \alpha_i > 0,\ (BB^\top \alpha - \mathbf{1} + \xi)_i = 0 \right\},\ I_{G_4} := \left\{ i \in Q_l \mid \xi_i > 0,\ C - \alpha_i = 0 \right\},$$
$$I_{GH_1} := \left\{ i \in Q_u \mid \zeta_i = 0, \left(AB^\top \alpha + z \right)_i = 0 \right\},\ I_{GH_2} := \left\{ i \in Q_u \mid z_i = 0, 1 - \zeta_i = 0 \right\},$$
$$I_{GH_3} := \left\{ i \in Q_l \mid \alpha_i = 0, \left(BB^\top \alpha - \mathbf{1} + \xi \right)_i = 0 \right\},\ I_{GH_4} := \left\{ i \in Q_l \mid \xi_i = 0, C - \alpha_i = 0 \right\}.$$

To provide a clearer description of the assumptions, we need the definitions for the following index sets.

Definition A.1.

$$\begin{aligned}
\Lambda_2 &:= \left\{ i \in Q_l \mid \alpha_i = 0,\ (BB^\top \alpha - \mathbf{1} + \xi)_i > 0,\ \xi_i = 0 \right\}, \\
\Lambda_3 &:= \left\{ i \in Q_l \mid 0 < \alpha_i \le C,\ (BB^\top \alpha - \mathbf{1} + \xi)_i = 0,\ \xi_i = 0 \right\}, \\
\Psi_2 &:= \left\{ i \in Q_u \mid \zeta_i = 0,\ (AB^\top \alpha + z)_i > 0,\ z_i = 0 \right\}, \\
\Psi_3 &:= \left\{ i \in Q_u \mid \zeta_i = 1,\ (AB^\top \alpha + z)_i = 0,\ z_i > 0 \right\}.
\end{aligned}$$

Proposition A.2. *The relationship between index sets*
(a) $I_{H_1} = \Psi_2,\ I_{G_1} = \Psi_3,\ I_{GH_1} = \emptyset.$
(b) $I_{H_2} = \Psi_2,\ I_{G_2} = \Psi_3,\ I_{GH_2} = \emptyset.$
(c) $I_{H_3} = \Lambda_2,\ I_{G_3} = \Lambda_3 \cup \Lambda_u,\ I_{GH_3} = \Lambda_1.$
(d) $I_{H_4} = \Lambda_1 \cup \Lambda_2 \cup \Lambda_3^+,\ I_{G_4} = \Lambda_u,\ I_{GH_4} = \Lambda_3^c.$

Proof. Based on our discussion after (13), $x_i^\top w^t \neq 0$ for each validation data point x_i, where $i \in [m_1]$, $t \in [T]$. It implies that index set

$\{i \in Q_u \mid 0 \leq \zeta_i < 1, (AB^\top \alpha + z)_i = 0, z_i = 0\} = \emptyset$. Likewise, the index set Ψ_1 in [40] is also empty. By incorporating the content of [40, Proposition 5], the proof is complete. □

To get the conclusions about MPEC-MFCQ-R and MPEC-LICQ at a feasible point v^* for the (BHO-SVM) we need analyze the properties of the following set of gradient vectors

$$\{\nabla G_i(v) \mid i \in I_G \cup I_{GH}\} \cup \{\nabla H_i(v) \mid i \in I_H \cup I_{GH}\}. \tag{A2}$$

Proposition A.3. *[40, Proposition 6] The set of gradient vectors in* (A2) *at a feasible point v^* for the* (BHO-SVM) *can be written in the matrix form*

$$\Gamma = \begin{bmatrix} \mathbf{0}_{(I_{G_1},L_1)} & \mathbf{0}_{(I_{G_1},L_2)} & \Gamma_a^3 & (AB^\top)_{(I_{G_1},\cdot)} & \mathbf{0}_{(I_{G_1},L_5)} \\ \mathbf{0}_{(I_{H_1},L_1)} & \Gamma_b^2 & \mathbf{0}_{(I_{H_1},L_3)} & \mathbf{0}_{(I_{H_1},L_4)} & \mathbf{0}_{(I_{H_1},L_5)} \\ \mathbf{0}_{(I_{G_2},L_1)} & \Gamma_c^2 & \mathbf{0}_{(I_{G_2},L_3)} & \mathbf{0}_{(I_{G_2},L_4)} & \mathbf{0}_{(I_{G_2},L_5)} \\ \mathbf{0}_{(I_{H_2},L_1)} & \mathbf{0}_{(I_{H_2},L_2)} & \Gamma_d^3 & \mathbf{0}_{(I_{H_2},L_4)} & \mathbf{0}_{(I_{H_2},L_5)} \\ \mathbf{0}_{(I_{G_3},L_1)} & \mathbf{0}_{(I_{G_3},L_2)} & \mathbf{0}_{(I_{G_3},L_3)} & (BB^\top)_{(I_{G_3},\cdot)} & \Gamma_e^5 \\ \mathbf{0}_{(I_{GH_3},L_1)} & \mathbf{0}_{(I_{GH_3},L_2)} & \mathbf{0}_{(I_{GH_3},L_3)} & (BB^\top)_{(I_{GH_3},\cdot)} & \Gamma_f^5 \\ \mathbf{0}_{(I_{GH_3},L_1)} & \mathbf{0}_{(I_{GH_3},L_2)} & \mathbf{0}_{(I_{GH_3},L_3)} & \Gamma_g^4 & \mathbf{0}_{(I_{GH_3},L_5)} \\ \mathbf{0}_{(I_{H_3},L_1)} & \mathbf{0}_{(I_{H_3},L_2)} & \mathbf{0}_{(I_{H_3},L_3)} & \Gamma_h^4 & \mathbf{0}_{(I_{H_3},L_5)} \\ \mathbf{1}_{(I_{G_4},L_1)} & \mathbf{0}_{(I_{G_4},L_2)} & \mathbf{0}_{(I_{G_4},L_3)} & \Gamma_i^4 & \mathbf{0}_{(I_{G_4},L_5)} \\ \mathbf{1}_{(I_{GH_4},L_1)} & \mathbf{0}_{(I_{GH_4},L_2)} & \mathbf{0}_{(I_{GH_4},L_3)} & \Gamma_j^4 & \mathbf{0}_{(I_{GH_4},L_5)} \\ \mathbf{0}_{(I_{GH_4},L_1)} & \mathbf{0}_{(I_{GH_4},L_2)} & \mathbf{0}_{(I_{GH_4},L_3)} & \mathbf{0}_{(I_{GH_4},L_4)} & \Gamma_k^5 \\ \mathbf{0}_{(I_{H_4},L_1)} & \mathbf{0}_{(I_{H_4},L_2)} & \mathbf{0}_{(I_{H_4},L_3)} & \mathbf{0}_{(I_{H_4},L_4)} & \Gamma_l^5 \end{bmatrix}, \tag{A3}$$

where L_q, $q = 1, ..., 5$ are the index sets of columns corresponding to the variables C, ζ, z, α and ξ, respectively, and

$$\begin{aligned} \Gamma_a^3 &:= \begin{bmatrix} \mathbf{0}_{(I_{G_1},\Psi_2)} & I_{(I_{G_1},\Psi_3)} \end{bmatrix}, & \Gamma_b^2 &:= \begin{bmatrix} \mathbf{0}_{(I_{H_1},\Psi_3)} & I_{(I_{H_1},\Psi_2)} \end{bmatrix}, \\ \Gamma_c^2 &:= \begin{bmatrix} \mathbf{0}_{(I_{G_2},\Psi_2)} & -I_{(I_{G_2},\Psi_3)} \end{bmatrix}, & \Gamma_d^3 &:= \begin{bmatrix} I_{(I_{H_2},\Psi_2)} & \mathbf{0}_{(I_{H_2},\Psi_3)} \end{bmatrix}, \\ \Gamma_e^5 &:= \begin{bmatrix} \mathbf{0}_{(I_{G_3},\Lambda_1\cup\Lambda_2)} & I_{(I_{G_3},\Lambda_3\cup\Lambda_u)} \end{bmatrix}, & \Gamma_f^5 &:= \begin{bmatrix} I_{(I_{GH_3},\Lambda_1)} & \mathbf{0}_{(I_{GH_3},\Lambda_2\cup\Lambda_3\cup\Lambda_u)} \end{bmatrix}, \\ \Gamma_g^4 &:= \begin{bmatrix} I_{(I_{GH_3},\Lambda_1)} & \mathbf{0}_{(I_{GH_3},\Lambda_2\cup\Lambda_3\cup\Lambda_u)} \end{bmatrix}, & \Gamma_h^4 &:= \begin{bmatrix} \mathbf{0}_{(I_{H_3},\Lambda_1\cup\Lambda_3\cup\Lambda_u)} & I_{(I_{H_3},\Lambda_2)} \end{bmatrix}, \\ \Gamma_i^4 &:= \begin{bmatrix} \mathbf{0}_{(I_{G_4},\Lambda_1\cup\Lambda_2\cup\Lambda_3)} & -I_{(I_{G_4},\Lambda_u)} \end{bmatrix}, & \Gamma_j^4 &:= \begin{bmatrix} \mathbf{0}_{(I_{GH_4},\Lambda_1\cup\Lambda_2\cup\Lambda_3^+\cup\Lambda_u)} & -I_{(I_{GH_4},\Lambda_3^c)} \end{bmatrix}, \\ \Gamma_k^5 &:= \begin{bmatrix} \mathbf{0}_{(I_{GH_4},\Lambda_1\cup\Lambda_2\cup\Lambda_3^+\cup\Lambda_u)} & I_{(I_{GH_4},\Lambda_3^c)} \end{bmatrix}, & \Gamma_l^5 &:= \begin{bmatrix} I_{(I_{H_4},\Lambda_1\cup\Lambda_2\cup\Lambda_3^+)} & \mathbf{0}_{(I_{H_4},\Lambda_3^c\cup\Lambda_u)} \end{bmatrix}. \end{aligned} \tag{A4}$$

Proof of Theorem 3.2

To prove the MPEC-MFCQ-R for (BHO-SVM) at a feasible point v^*, assume there exists nonzero column vector ρ defined by $\rho = [\ \rho_a^\top, \rho_b^\top, \rho_c^\top, \rho_d^\top, \rho_e^\top, \rho_f^\top, \rho_g^\top, \rho_h^\top, \rho_i^\top, \rho_j^\top, \rho_k^\top, \rho_l^\top\ ]^\top$, where $\rho_f,\ \rho_g,\ \rho_j,\ \rho_k \geq 0$ such that

$\rho^\top \Gamma = 0$. It gives that

$$0 = \rho^\top \Gamma := \left[\begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \right],$$

where

$$\begin{aligned}
S_1 =& \rho_i^\top \mathbf{1}_{(\Lambda_u, L_1)} + \rho_j^\top \mathbf{1}_{(\Lambda_3^c, L_1)} = 0, \\
S_2 =& \rho_b^\top \Gamma_b^2 + \rho_c^\top \Gamma_c^2 = 0, && \text{(A5)} \\
S_3 =& \rho_a^\top \Gamma_a^3 + \rho_d^\top \Gamma_d^3, && \text{(A6)} \\
S_4 =& \rho_a^\top (AB^\top)_{(I_{G_1}, \cdot)} + \rho_e^\top (BB^\top)_{(I_{G_3}, \cdot)} + \rho_f^\top (BB^\top)_{(I_{GH_3}, \cdot)} \\
& + \rho_g^\top \Gamma_g^4 + \rho_h^\top \Gamma_h^4 + \rho_i^\top \Gamma_i^4 + \rho_j^\top \Gamma_j^4 = 0, \\
S_5 =& \rho_e^\top \Gamma_e^5 + \rho_f^\top \Gamma_f^5 + \rho_k^\top \Gamma_k^5 + \rho_l^\top \Gamma_l^5 = 0. && \text{(A7)}
\end{aligned}$$

By [40, Lemma 1] and (A5), $\rho_b = 0$ and $\rho_c = 0$. By (A6) and Proposition A.3, it holds that

$$S_3 = \rho_a^\top \left[\mathbf{0}_{(I_{G_1}, \Psi_2)} \quad I_{(I_{G_1}, \Psi_3)} \right] + \rho_d^\top \left[I_{(I_{H_2}, \Psi_2)} \quad \mathbf{0}_{(I_{H_2}, \Psi_3)} \right] = \left[\begin{array}{cc} \rho_d^\top & \rho_a^\top \end{array} \right] = 0.$$

By (A7) and Proposition A.3, it holds that

$$\begin{aligned}
S_5 =& \rho_e^\top \left[\mathbf{0}_{(I_{G_3}, \Lambda_1 \cup \Lambda_2)} \quad I_{(I_{G_3}, \Lambda_3 \cup \Lambda_u)} \right] + \rho_f^\top \left[I_{(I_{GH_3}, \Lambda_1)} \quad \mathbf{0}_{(I_{GH_3}, \Lambda_2 \cup \Lambda_3 \cup \Lambda_u)} \right] \\
& + \rho_k^\top \left[\mathbf{0}_{(I_{GH_4}, \Lambda_1 \cup \Lambda_2 \cup \Lambda_3^+ \cup \Lambda_u)} \quad I_{(I_{GH_4}, \Lambda_3^c)} \right] + \rho_l^\top \left[I_{(I_{H_4}, \Lambda_1 \cup \Lambda_2 \cup \Lambda_3^+)} \quad \mathbf{0}_{(I_{H_4}, \Lambda_3^c \cup \Lambda_u)} \right] \\
=& \left[\begin{array}{ccccc} \rho_f^\top + \rho_{l_{\Lambda_1}}^\top & \rho_{l_{\Lambda_2}}^\top & \rho_{e_{\Lambda_3^+}}^\top + \rho_{l_{\Lambda_3^+}}^\top & \rho_{e_{\Lambda_3^c}}^\top + \rho_k^\top & \rho_{e_{\Lambda_u}}^\top \end{array} \right] = 0.
\end{aligned}$$

It implies that $\rho_a = 0$, $\rho_d = 0$, $\rho_{l_{\Lambda_2}} = 0$ and $\rho_{e_{\Lambda_u}} = 0$. Therefore, proving that the matrix Γ demonstrates positively linearly independence is reduced to examining the positively linearly independence of the following matrix

$$\widetilde{\Gamma} = \left[\begin{array}{ccc}
\mathbf{0}_{(I_{G_3}, L_1)} & (BB^\top)_{(I_{G_3}, \cdot)} & \widetilde{\Gamma}_e^5 \\
\mathbf{0}_{(I_{GH_3}, L_1)} & (BB^\top)_{(I_{GH_3}, \cdot)} & \Gamma_f^5 \\
\mathbf{0}_{(I_{GH_3}, L_1)} & \Gamma_g^4 & \mathbf{0}_{(I_{GH_3}, L_5)} \\
\mathbf{0}_{(I_{H_3}, L_1)} & \Gamma_h^4 & \mathbf{0}_{(I_{H_3}, L_5)} \\
\mathbf{1}_{(I_{G_4}, L_1)} & \Gamma_i^4 & \mathbf{0}_{(I_{G_4}, L_5)} \\
\mathbf{1}_{(I_{GH_4}, L_1)} & \Gamma_j^4 & \mathbf{0}_{(I_{GH_4}, L_5)} \\
\mathbf{0}_{(I_{GH_4}, L_1)} & \mathbf{0}_{(I_{GH_4}, L_4)} & \Gamma_k^5 \\
\mathbf{0}_{(I_{H_4}, L_1)} & \mathbf{0}_{(I_{H_4}, L_4)} & \widetilde{\Gamma}_l^5
\end{array} \right], \tag{A8}$$

where

$$\widetilde{\Gamma}_e^5 := \left[\mathbf{0}_{(\Lambda_3, \Lambda_1 \cup \Lambda_2 \cup \Lambda_u)} \quad I_{(\Lambda_3, \Lambda_3)} \right], \quad \widetilde{\Gamma}_l^5 := \left[I_{(\Lambda_1 \cup \Lambda_3^+, \Lambda_1 \cup \Lambda_3^+)} \quad \mathbf{0}_{(\Lambda_1 \cup \Lambda_3^+, \Lambda_2 \cup \Lambda_3^c \cup \Lambda_u)} \right].$$

Note that $\widetilde{\Gamma}$ is the submatrix of $\widetilde{\Gamma}$ in [41, Theorem 1] corresponding to the second row block to the last row block. Therefore, using the same technique of proof in [41, Theorem 1], if $\left(BB^\top\right)_{(\Lambda_1\cup\Lambda_3,\Lambda_1\cup\Lambda_3)}$ is positive definite at v^*, $\rho = \left[\ \rho_a^\top, \rho_b^\top, \rho_c^\top, \rho_d^\top, \rho_e^\top, \rho_f^\top, \rho_g^\top, \rho_h^\top, \rho_i^\top, \rho_j^\top, \rho_k^\top, \rho_l^\top\ \right]^\top = 0$. Then MPEC-MFCQ-R holds at v^*. □

Appendix B. Proof of Theorem 3.3

Proposition B.1. *In view of Proposition A.2, Γ in* (A3) *takes the following form*

$$\Gamma = \begin{bmatrix}
0_{(\Psi_3,L_1)} & 0_{(\Psi_3,L_2)} & \Gamma_a^3 & (AB^\top)_{(\Psi_3,\cdot)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & \Gamma_b^2 & 0_{(\Psi_2,L_3)} & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Psi_3,L_1)} & \Gamma_c^2 & 0_{(\Psi_3,L_3)} & 0_{(\Psi_3,L_4)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & 0_{(\Psi_2,L_2)} & \Gamma_d^3 & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma_{e_1}^5 \\
0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & (BB^\top)_{(\Lambda_3^c,\cdot)} & \Gamma_{e_2}^5 \\
0_{(\Lambda_u,L_1)} & 0_{(\Lambda u,L_2)} & 0_{(\Lambda_u,L_3)} & (BB^\top)_{(\Lambda_u,\cdot)} & \Gamma_{e_3}^5 \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & (BB^\top)_{(\Lambda_1,\cdot)} & \Gamma_f^5 \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & \Gamma_g^4 & 0_{(\Lambda_1,L_5)} \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\
\mathbf{1}_{(\Lambda_u,L_1)} & 0_{(\Lambda_u,L_2)} & 0_{(\Lambda u,L_3)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\
\mathbf{1}_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & \Gamma_j^4 & 0_{(\Lambda_3^c,L_5)} \\
0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & 0_{(\Lambda_3^c,L_4)} & \Gamma_k^5 \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & 0_{(\Lambda_1,L_4)} & \Gamma_{l_1}^5 \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & 0_{(\Lambda_2,L_4)} & \Gamma_{l_2}^5 \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5
\end{bmatrix}, \tag{B1}$$

where

$$\begin{aligned}
\Gamma_{e_1}^5 &:= \left[\ \mathbf{0}_{(\Lambda_3^+,\Lambda_1)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} \ \ I_{(\Lambda_3^+,\Lambda_3^+)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_3^c)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_u)}\ \right], \\
\Gamma_{e_2}^5 &:= \left[\ \mathbf{0}_{(\Lambda_3^c,\Lambda_1)} \ \ \mathbf{0}_{(\Lambda_3^c,\Lambda_2)} \ \ \mathbf{0}_{(\Lambda_3^c,\Lambda_3^+)} \ \ I_{(\Lambda_3^c,\Lambda_3^c)} \ \ \mathbf{0}_{(\Lambda_3^c,\Lambda_u)}\ \right], \\
\Gamma_{e_3}^5 &:= \left[\ \mathbf{0}_{(\Lambda_u,\Lambda_1)} \ \ \mathbf{0}_{(\Lambda_u,\Lambda_2)} \ \ \mathbf{0}_{(\Lambda_u,\Lambda_3^+)} \ \ \mathbf{0}_{(\Lambda_u,\Lambda_3^c)} \ \ I_{(\Lambda_u,\Lambda_u)}\ \right], \\
\Gamma_{l_1}^5 &:= \left[\ I_{(\Lambda_1,\Lambda_1)} \ \ \mathbf{0}_{(\Lambda_1,\Lambda_2)} \ \ \mathbf{0}_{(\Lambda_1,\Lambda_3^+)} \ \ \mathbf{0}_{(\Lambda_1,\Lambda_3^c)} \ \ \mathbf{0}_{(\Lambda_1,\Lambda_u)}\ \right], \\
\Gamma_{l_2}^5 &:= \left[\ \mathbf{0}_{(\Lambda_2,\Lambda_1)} \ \ I_{(\Lambda_2,\Lambda_2)} \ \ \mathbf{0}_{(\Lambda_2,\Lambda_3^+)} \ \ \mathbf{0}_{(\Lambda_2,\Lambda_3^c)} \ \ \mathbf{0}_{(\Lambda_2,\Lambda_u)}\ \right], \\
\Gamma_{l_3}^5 &:= \left[\ \mathbf{0}_{(\Lambda_3^+,\Lambda_1)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} \ \ I_{(\Lambda_3^+,\Lambda_3^+)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_3^c)} \ \ \mathbf{0}_{(\Lambda_3^+,\Lambda_u)}\ \right].
\end{aligned}$$

Proposition B.2. *By the definition of L_q, $q = 1, ..., 5$, there are $2n-1$ columns in Γ. The number of rows is $2n - 2 + |\Lambda_1| + |\Lambda_3^c|$ by the Γ in* (B1)*.*

Related Lemmas and Proofs

Define

$$\Gamma_{sub1} = \begin{bmatrix}
0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma_{e_1}^5 \\
0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\
\mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5
\end{bmatrix}. \tag{B2}$$

Lemma B.3. *Let v^* be a feasible point of* (BHO-SVM) *and satisfy Assumption 1. If $|I_{GH}| = 0$, Γ_{sub1} has full row rank at v^*.*

Proof. We conduct the following basic row transformation to Γ_{sub1}. Subtracting the second forth block from the first row block, we get the following matrix

$$\Gamma^0 = \begin{bmatrix} 0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & 0_{(\Lambda_3^+,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix}.$$

Next, we conduct the following row transformation.

Multiplying the second (j = 2) row block by $-(BB^\top)_{(\Lambda_3^+,\Lambda_2)}$ from the left and adding it to the first (i = 1) row block, we can obtain the third row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_2)}$ replaced by zero. We denote this process as B-(1, 2).

Multiplying the third (j =3) row block by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ from the left and adding it to the first (i = 1) row block, we can obtain the second row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ replaced by zero. Meanwhile, $0_{(\Lambda_3^+,L_1)}$ is replaced by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$. We denote this process as B+(1, 3).

Then we reach the following matrix

$$\widehat{\Gamma} = \begin{bmatrix} (BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} & Q_1 & 0_{(\Lambda_3^+,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix},$$

where

$$Q_1 = \begin{bmatrix} \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} & (BB^\top)_{(\Lambda_3^+,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_u)} \end{bmatrix}.$$

Assume there exists nonzero column vector ρ defined by $\rho = \left[\rho_1^\top, \rho_2^\top, \rho_3^\top, \rho_4^\top\right]^\top$ such that $\rho^\top\overline{\Gamma} = 0$. It gives that

$$0 = \rho^\top\overline{\Gamma} := \begin{bmatrix} S_1 & S_2 & S_3 \end{bmatrix},$$

where

$$S_1 = \rho_1^\top(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} + \rho_3^\top\mathbf{1}_{(\Lambda_u,L_1)} = 0, \tag{B3}$$

$$S_2 = \rho_1^\top Q_1 + \rho_2^\top\Gamma_h^5 + \rho_3^\top\Gamma_i^4 = \begin{bmatrix} \rho_2^\top & \rho_1^\top(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)} & \rho_3^\top \end{bmatrix} = 0, \tag{B4}$$

$$S_3 = \rho_4^\top\Gamma_{l_3}^5 = 0. \tag{B5}$$

By (B4), (B5) and Assumption 1, it holds that $\rho_1 = 0$, $\rho_2 = 0$, $\rho_3 = 0$ and $\rho_4 = 0$. Overall, $\widehat{\Gamma}$ has full row rank. □

Then, define

$$\Gamma_{sub2} = \begin{bmatrix} 0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma^5_{e_1} \\ 0_{(\Lambda_1,L_1)} & (BB^\top)_{(\Lambda_1,\cdot)} & \Gamma^5_{f} \\ 0_{(\Lambda_1,L_1)} & \Gamma^4_{g} & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma^4_{h} & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma^4_{i} & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_4)} & \Gamma^5_{l_1} \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma^5_{l_3} \end{bmatrix}. \tag{B6}$$

Lemma B.4. *Let v^* be a feasible point of* (BHO-SVM) *and satisfy Assumption 1. If $|I_{GH_3}| = 1$ and $\widehat{a}_3 \neq 0$, Γ_{sub2} has full row rank at v^*. Otherwise, Γ_{sub2} does not have full row rank at v^*.*

Proof. We conduct the following basic row transformation to Γ_{sub2}. Subtracting the seventh row block from the first row block, and Subtracting the sixth row block from the second block, we reach the following matrix

$$\Gamma^0 = \begin{bmatrix} 0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & 0_{(\Lambda_3^+,L_5)} \\ 0_{(\Lambda_1,L_1)} & (BB^\top)_{(\Lambda_1,\cdot)} & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_1,L_1)} & \Gamma^4_{g} & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma^4_{h} & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma^4_{i} & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_4)} & \Gamma^5_{l_1} \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma^5_{l_3} \end{bmatrix}.$$

Next, we conduct the following row transformation.

Multiplying the third (j = 3) row block by $-(BB^\top)_{(\Lambda_3^+,\Lambda_1)}$ from the left and adding it to the first (i = 1) row block, we can obtain the first row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_1)}$ replaced by zero. We denote this process as B-(1, 3). We carry on to conduct B-(1, 4) and we obtain the fourth row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_2)}$ replaced by zero.

Multiplying the fifth (j = 5) row block by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ from the left and adding it to the first (i = 1) row block, we can obtain the first row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ replaced by zero. Meanwhile, $0_{(\Lambda_3^+,L_1)}$ is replaced by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$. We denote this process as B+(1, 5).

We carry on to conduct row transformation B-(2, 3) to obtain the second row block with $(BB^\top)_{(\Lambda_1,\Lambda_1)}$ replaced by zero. Conduct B-(2, 4) to obtain the second row block with $(BB^\top)_{(\Lambda_1,\Lambda_2)}$ replaced by zero. And conduct B+(2, 5) to obtain the second row block with $(BB^\top)_{(\Lambda_1,\Lambda_u)}$ replaced by zero. Meanwhile, $0_{(\Lambda_1,L_1)}$ is replaced by $(BB^\top)_{(\Lambda_1,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$. We reach the following matrix

$$\widehat{\Gamma} = \begin{bmatrix} (BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} & Q_3^1 & 0_{(\Lambda_3^+,L_5)} \\ (BB^\top)_{(\Lambda_1,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} & Q_3^2 & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_1,L_1)} & \Gamma_g^4 & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_4)} & \Gamma_{l_1}^5 \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix},$$

where

$$\begin{aligned} Q_3^1 =& [\mathbf{0}_{(\Lambda_3^+,\Lambda_1)} \quad \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} \quad (BB^\top)_{(\Lambda_3^+,\Lambda_3^+)} \quad \mathbf{0}_{(\Lambda_3^+,\Lambda_u)}], \\ Q_3^2 =& [\mathbf{0}_{(\Lambda_1,\Lambda_1)} \quad \mathbf{0}_{(\Lambda_1,\Lambda_2)} \quad (BB^\top)_{(\Lambda_1,\Lambda_3^+)} \quad \mathbf{0}_{(\Lambda_1,\Lambda_u)}]. \end{aligned}$$

We discuss the following two cases.
(i) By Assumption 1, if $\Lambda_3^+ \neq \emptyset$, $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ is positive definite. Therefore, the rows in $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ can fully express the rows in $(BB^\top)_{(\Lambda_1,\Lambda_3^+)}$. In other words, by conducting basic row transformation, we can make Q_3^2 replaced by zeros. Specifically, there exists $A_3 \in \mathbb{R}^{|\Lambda_1|\times|\Lambda_3^+|}$

$$(BB^\top)_{(\Lambda_1,\Lambda_3^+)} = A_3(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}.$$

By multiplying the first row block by $-A_3$ from left and add them to the second row block, we obtain the following matrix

$$\overline{\Gamma} = \begin{bmatrix} (BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} & Q_3^1 & 0_{(\Lambda_3^+,L_5)} \\ a_3 & 0_{(\Lambda_1,L_3)} & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_1,L_1)} & \Gamma_g^4 & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_4)} & \Gamma_{l_1}^5 \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix},$$

where

$$a_3 = (BB^\top)_{(\Lambda_1,\Lambda_u)}\mathbf{1}_{|\Lambda_u|} - A_3(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}.$$

Assume there exists nonzero column vector ρ defined by $\rho = \left[\rho_1^\top, \rho_2^\top, \dots, \rho_7^\top\right]^\top$ such that $\rho^\top\overline{\Gamma} = 0$. It gives that

$$0 = \rho^\top\overline{\Gamma} := \left[\ S_1 \quad S_2 \quad S_3\ \right],$$

where

$$S_1 = \rho_1^\top (BB^\top)_{(\Lambda_3^+,\Lambda_u)} \mathbf{1}_{|\Lambda_u|} + \rho_2^\top a_3 + \rho_5^\top \mathbf{1}_{(\Lambda_u,L_1)} = 0, \tag{B7}$$

$$S_2 = \rho_1^\top Q_3^1 + \rho_3^\top \Gamma_g^4 + \rho_4^\top \Gamma_h^4 + \rho_5^\top \Gamma_i^4 = \begin{bmatrix} \rho_3^\top & \rho_4^\top & \rho_1^\top (BB^\top)_{(\Lambda_3^+,\Lambda_3^+)} & \rho_5^\top \end{bmatrix} = 0, \tag{B8}$$

$$S_3 = \rho_6^\top \Gamma_{l_1}^5 + \rho_7^\top \Gamma_{l_3}^5 = 0. \tag{B9}$$

By (B8), (B9) and Assumption 1, it holds that $\rho_1 = 0$, $\rho_3 = 0$, $\rho_4 = 0$, $\rho_5 = 0$, $\rho_6 = 0$ and $\rho_7 = 0$. Then S_1 in (B7) reduces to the following form

$$S_1 = \rho_2^\top a_3 = 0.$$

Recall that $|\Lambda_1| = 1$. If $a_3 \neq 0$, S_1 reduces to $\rho_2 a_3 = 0$, implying that $\rho_2 = 0$. Therefore, $\overline{\Gamma}$ has full row rank. Otherwise, $\overline{\Gamma}$ does not have full row rank.

(ii) If $\Lambda_3^+ = \emptyset$, $\widehat{\Gamma}$ reduces to the following form

$$\widehat{\Gamma} = \begin{bmatrix} (BB^\top)_{(\Lambda_1,\Lambda_u)} \mathbf{1}_{|\Lambda_u|} & 0_{(\Lambda_1,L_4)} & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_1,L_1)} & \Gamma_g^4 & 0_{(\Lambda_1,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_4)} & \Gamma_{l_1}^5 \end{bmatrix}.$$

Assume there exists nonzero column vector ρ defined by $\rho = \left[\rho_1^\top, \rho_2^\top, \ldots, \rho_5^\top\right]^\top$ such that $\rho^\top \widehat{\Gamma} = 0$. It gives that

$$0 = \rho^\top \widehat{\Gamma} := \begin{bmatrix} S_1 & S_2 & S_3 \end{bmatrix},$$

where

$$S_1 = \rho_1^\top (BB^\top)_{(\Lambda_1,\Lambda_u)} \mathbf{1}_{|\Lambda_u|} + \rho_4^\top \mathbf{1}_{(\Lambda_u,L_1)} = 0, \tag{B10}$$

$$S_2 = \rho_2^\top \Gamma_g^4 + \rho_3^\top \Gamma_h^4 + \rho_4^\top \Gamma_i^4 = \begin{bmatrix} \rho_2^\top & \rho_3^\top & \rho_4^\top \end{bmatrix} = 0, \tag{B11}$$

$$S_3 = \rho_5^\top \Gamma_{l_1}^5 = 0. \tag{B12}$$

By (B11) and (B12) , it holds that $\rho_2 = 0$, $\rho_3 = 0$, $\rho_4 = 0$ and $\rho_5 = 0$. Then S_1 in (B10) reduces to the following form

$$S_1 = \rho_1^\top (BB^\top)_{(\Lambda_1,\Lambda_u)} \mathbf{1}_{|\Lambda_u|} = 0.$$

Recall that $|\Lambda_1| = 1$. If $(BB^\top)_{(\Lambda_1,\Lambda_u)} \mathbf{1}_{|\Lambda_u|} \neq 0$, it gives that $\rho_1 = 0$, implying that $\widehat{\Gamma}$ has full row rank. Otherwise, $\widehat{\Gamma}$ does not have full row rank.

Overall, by the definition of $\widehat{a}_3$ in Theorem 3.3-(iii), we obtain the result. Since the row transformation does not change the rank, therefore, the same result applies to Γ_{sub2}. □

Finally, define

$$\Gamma_{sub3} = \begin{bmatrix} 0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma^5_{e_1} \\ 0_{(\Lambda_3^c,L_1)} & (BB^\top)_{(\Lambda_3^c,\cdot)} & \Gamma^5_{e_2} \\ 0_{(\Lambda_2,L_1)} & \Gamma^4_h & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma^4_i & 0_{(\Lambda_u,L_5)} \\ \mathbf{1}_{(\Lambda_3^c,L_1)} & \Gamma^4_j & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_4)} & \Gamma^5_k \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma^5_{l_3} \end{bmatrix}. \tag{B13}$$

Lemma B.5. *Let v^* be a feasible point of* (BHO-SVM) *and satisfy Assumption 1. If $|I_{GH_4}| = 1$ and $\widehat{a}_4 \neq 0$, Γ_{sub3} has full row rank at v^*. Otherwise, Γ_{sub3} does not have full row rank at v^*.*

Proof. We conduct the following basic row transformation to Γ_{sub3}. Subtracting the seventh row block from the first row block, and Subtracting the sixth row block from the second row block, we reach the following matrix

$$\Gamma^0 = \begin{bmatrix} 0_{(\Lambda_3^+,L_1)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & 0_{(\Lambda_3^+,L_5)} \\ 0_{(\Lambda_3^c,L_1)} & (BB^\top)_{(\Lambda_3^c,\cdot)} & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma^4_h & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma^4_i & 0_{(\Lambda_u,L_5)} \\ \mathbf{1}_{(\Lambda_3^c,L_1)} & \Gamma^4_j & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_4)} & \Gamma^5_k \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma^5_{l_3} \end{bmatrix}.$$

Next, we conduct the following row transformation.

Multiplying the third (j = 3) row block by $-(BB^\top)_{(\Lambda_3^+,\Lambda_2)}$ from the left and adding it to the first (i = 1) row block, we can obtain the first row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_2)}$ replaced by zero. We denote this process as B-(1, 3).

Multiplying the forth (j = 4) row block by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ from the left and adding it to the first (i = 1) row block, we can obtain the first row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}$ replaced by zero. Meanwhile, $0_{(\Lambda_3^+,L_1)}$ is replaced by $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$. We denote this process as B+(1, 4).

We carry on to conduct B+(1, 5) and we obtain the first row block with $(BB^\top)_{(\Lambda_3^+,\Lambda_3^c)}$ replaced by zero. Meanwhile, $(BB^\top)_{(\Lambda_3^+,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$ is replaced by $(BB^\top)_{(\Lambda_3^+,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|}$.

We carry on to conduct row transformation B-(2, 3) to obtain the second row block with $(BB^\top)_{(\Lambda_3^c,\Lambda_2)}$ replaced by zero. Conduct B+(2, 4) to obtain the second row block with $(BB^\top)_{(\Lambda_3^c,\Lambda_u)}$ replaced by zero. Meanwhile, $0_{(\Lambda_3^c,L_1)}$ is replaced by $(BB^\top)_{(\Lambda_3^c,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$. And conduct B+(2, 5) to obtain the second row block with $(BB^\top)_{(\Psi_1^+,\Lambda_3^c)}$ replaced by zero. Meanwhile, $(BB^\top)_{(\Lambda_3^c,\Lambda_u)}\mathbf{1}_{|\Lambda_u|}$ is replaced by

$(BB^\top)_{(\Lambda_3^c,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|}$. We reach the following matrix

$$\widehat{\Gamma}=\begin{bmatrix} a_4^1 & Q_4^1 & 0_{(\Lambda_3^+,L_5)} \\ a_4^2 & Q_4^2 & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ \mathbf{1}_{(\Lambda_3^c,L_1)} & \Gamma_j^4 & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_4)} & \Gamma_k^5 \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix},$$

where

$$\begin{aligned} Q_4^1 &= [\mathbf{0}_{(\Lambda_3^+,\Lambda_2)},(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)},\mathbf{0}_{(\Lambda_3^+,\Lambda_3^c)},\mathbf{0}_{(\Lambda_3^+,\Lambda_u)}], \\ Q_4^2 &= [\mathbf{0}_{(\Lambda_3^c,\Lambda_2)},(BB^\top)_{(\Lambda_3^c,\Lambda_3^+)},\mathbf{0}_{(\Lambda_3^c,\Lambda_3^c)},\mathbf{0}_{(\Lambda_3^c,\Lambda_u)}], \\ a_4^1 &= (BB^\top)_{(\Lambda_3^+,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|}, \quad a_4^2 = (BB^\top)_{(\Lambda_3^c,\Lambda_3^c\cup\Lambda_u)}\mathbf{1}_{|\Lambda_3^c\cup\Lambda_u|}. \end{aligned}$$

We discuss the following two cases.
(i)If $\Lambda_3^+\neq\emptyset$ and $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ is positive definite, the rows in $(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}$ can fully express the rows in $(BB^\top)_{(\Lambda_3^c,\Lambda_3^+)}$. In other words, by conducting basic row transformation, we can make Q_4^2 replaced by zeros. Specifically, there exists $A_4\in\mathbb{R}^{|\Lambda_3^c|\times|\Lambda_3^+|}$ satisfying

$$(BB^\top)_{(\Lambda_3^c,\Lambda_3^+)} = A_4(BB^\top)_{(\Lambda_3^+,\Lambda_3^+)}.$$

By multiplying the third row block by $-A_4^1$ $(-A_4^2)$ from left and add them to the first (forth) row block, we obtain the following matrix

$$\overline{\Gamma}=\begin{bmatrix} a_4^2 & Q_4^1 & 0_{(\Lambda_3^+,L_5)} \\ a_4^3 & 0_{(\Lambda_3^c,L_4)} & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_2,L_1)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ \mathbf{1}_{(\Lambda_3^c,L_1)} & \Gamma_j^4 & 0_{(\Lambda_3^c,L_5)} \\ 0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_4)} & \Gamma_k^5 \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix},$$

where

$$a_4^3 = a_4^2 - A_4a_4^1.$$

Assume there exists nonzero column vector ρ defined by $\rho=\left[\rho_1^\top,\rho_2^\top,\ldots,\rho_7^\top\right]^\top$ such that $\rho^\top\overline{\Gamma}=0$. It gives that

$$0=\rho^\top\overline{\Gamma} := \left[\begin{array}{ccc} S_1 & S_2 & S_3 \end{array}\right],$$

where

$$S_1 = \rho_1^\top a_4^2 + \rho_2^\top a_4^3 + \rho_4^\top \mathbf{1}_{(\Lambda_u, L_1)} + \rho_5^\top \mathbf{1}_{(\Lambda_3^c, L_1)} = 0, \tag{B14}$$

$$S_2 = \rho_1^\top Q_4^1 + \rho_3^\top \Gamma_h^4 + \rho_4^\top \Gamma_i^4 + \rho_5^\top \Gamma_j^4 = \begin{bmatrix} \rho_3^\top & \rho_1^\top (BB^\top)_{(\Lambda_3^+, \Lambda_3^+)} & \rho_5^\top & \rho_4^\top \end{bmatrix} = 0, \tag{B15}$$

$$S_3 = \rho_6^\top \Gamma_k^5 + \rho_7^\top \Gamma_{l_3}^5 = 0. \tag{B16}$$

By (B15), (B16) and Assumption 1, it holds that $\rho_1 = 0$, $\rho_3 = 0$, $\rho_4 = 0$, $\rho_5 = 0$, $\rho_6 = 0$ and $\rho_7 = 0$. Then S_1 in (B14) reduces to the following form

$$S_1 = \rho_2^\top a_4^3 = 0.$$

Recall that $|\Lambda_3^c| = 1$. If $a_4^3 \neq 0$, S_1 reduces to $\rho_2^\top a_4^3 = 0$, implying that $\rho_2 = 0$. Therefore, $\overline{\Gamma}$ has full row rank. Otherwise, $\overline{\Gamma}$ does not have full row rank.
(ii) If $\Lambda_3^+ = \emptyset$. $\widehat{\Gamma}$ reduces to the following form

$$\widehat{\Gamma} = \begin{bmatrix} a_4^1 & 0_{(\Lambda_3^c, L_4)} & 0_{(\Lambda_3^c, L_5)} \\ 0_{(\Lambda_2, L_1)} & \Gamma_h^4 & 0_{(\Lambda_2, L_5)} \\ \mathbf{1}_{(\Lambda_u, L_1)} & \Gamma_i^4 & 0_{(\Lambda_u, L_5)} \\ \mathbf{1}_{(\Lambda_3^c, L_1)} & \Gamma_j^4 & 0_{(\Lambda_3^c, L_5)} \\ 0_{(\Lambda_3^c, L_1)} & 0_{(\Lambda_3^c, L_4)} & \Gamma_k^5 \end{bmatrix}.$$

Assume there exists nonzero column vector ρ defined by $\rho = \left[\rho_1^\top, \rho_2^\top, \ldots, \rho_5^\top\right]^\top$ such that $\rho^\top \widehat{\Gamma} = 0$. It gives that

$$0 = \rho^\top \widehat{\Gamma} := \begin{bmatrix} S_1 & S_2 & S_3 \end{bmatrix},$$

where

$$S_1 = \rho_1^\top a_4^1 + \rho_3^\top \mathbf{1}_{(\Lambda_u, L_1)} + \rho_4^\top \mathbf{1}_{(\Lambda_u, L_1)} = 0, \tag{B17}$$

$$S_2 = \rho_2^\top \Gamma_h^4 + \rho_3^\top \Gamma_i^4 + \rho_4^\top \Gamma_j^4 = \begin{bmatrix} \rho_2^\top & \rho_4^\top & \rho_3^\top \end{bmatrix} = 0, \tag{B18}$$

$$S_3 = \rho_5^\top \Gamma_k^5 = 0. \tag{B19}$$

By (B18) and (B19) , it holds that $\rho_2 = 0$, $\rho_3 = 0$, $\rho_4 = 0$ and $\rho_5 = 0$. Then S_1 in (B17) in reduces to the following form

$$S_1 = \rho_1^\top a_4^1 = 0.$$

Recall that $|\Lambda_3^c| = 1$. If $a_4^1 \neq 0$, it gives that $\rho_1 = 0$, implying that $\widehat{\Gamma}$ has full row rank. Otherwise, $\widehat{\Gamma}$ dose not have full row rank.

Overall, by the definition of $\widehat{a}_4$ in Theorem 3.3-(iv), we obtain the result. Since the row transformation does not change the rank, therefore, the same result applies to Γ_{sub3}.

□

Proof of Theorem 3.3

(i) $|I_{GH}| > 1$.

If $|I_{GH}| > 1$, the number of row vectors is greater than the number of elements in each vector in Γ. The row vectors in Γ are linearly dependent. Therefore, MPEC-LICQ fails at v^*.

(ii) $|I_{GH}| = 0$

By Proposition $A.3$ (b), if $|I_{GH}| = 0$, Γ in (B1) reduces to the following form

$$\Gamma_1 := \begin{bmatrix} 0_{(\Psi_3,L_1)} & 0_{(\Psi_3,L_2)} & \Gamma_a^3 & (AB^\top)_{(\Psi_3,\cdot)} & 0_{(\Psi_3,L_5)} \\ 0_{(\Psi_2,L_1)} & \Gamma_b^2 & 0_{(\Psi_2,L_3)} & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\ 0_{(\Psi_3,L_1)} & \Gamma_c^2 & 0_{(\Psi_3,L_3)} & 0_{(\Psi_3,L_4)} & 0_{(\Psi_3,L_5)} \\ 0_{(\Psi_2,L_1)} & 0_{(\Psi_2,L_2)} & \Gamma_d^3 & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma_{e_1}^5 \\ 0_{(\Lambda_u,L_1)} & 0_{(\Lambda u,L_2)} & 0_{(\Lambda_u,L_3)} & (BB^\top)_{(\Lambda_u,\cdot)} & \Gamma_{e_3}^5 \\ 0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\ \mathbf{1}_{(\Lambda_u,L_1)} & 0_{(\Lambda_u,L_2)} & 0_{(\Lambda u,L_3)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\ 0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & 0_{(\Lambda_2,L_4)} & \Gamma_{l_2}^5 \\ 0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5 \end{bmatrix}.$$

Assume there exists nonzero column vector ρ defined by $\rho = [\rho_a^\top, \rho_b^\top, \rho_c^\top, \rho_d^\top, \rho_{e_1}^\top, \rho_{e_3}^\top, \rho_h^\top, \rho_i^\top, \rho_{l_2}^\top, \rho_{l_3}^\top]^\top$ such that $\rho^\top\Gamma_1 = 0$. It gives that

$$0 = \rho^\top\Gamma_1 := \begin{bmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \end{bmatrix},$$

where

$$\begin{aligned} S_1 =& \rho_i^\top \mathbf{1}_{(\Lambda_u,L_1)} = 0, \\ S_2 =& \rho_b^\top\Gamma_b^2 + \rho_c^\top\Gamma_c^2 = 0, && \text{(B20)} \\ S_3 =& \rho_a^\top\Gamma_a^3 + \rho_d^\top\Gamma_d^3 = 0, && \text{(B21)} \\ S_4 =& \rho_a^\top(AB^\top)_{(\Psi_3,\cdot)} + \rho_{e_1}^\top(BB^\top)_{(\Lambda_3^+,\cdot)} + \rho_{e_3}^\top(BB^\top)_{(\Lambda_u,\cdot)} + \rho_h^\top\Gamma_h^4 + \rho_i^\top\Gamma_i^4 = 0, \\ S_5 =& \rho_{e_1}^\top\Gamma_{e_1}^5 + \rho_{e_3}^\top\Gamma_{e_3}^5 + \rho_{l_2}^\top\Gamma_{l_2}^5 + \rho_{l_3}^\top\Gamma_{l_3}^5 = 0. && \text{(B22)} \end{aligned}$$

By [40, Lemma 1] and (B20), $\rho_b = 0$ and $\rho_c = 0$. By (B21), it holds that

$$S_3 = \rho_a^\top \begin{bmatrix} \mathbf{0}_{(\Psi_3,\Psi_2)} & I_{(\Psi_3,\Psi_3)} \end{bmatrix} + \rho_d^\top \begin{bmatrix} I_{(\Psi_2,\Psi_2)} & \mathbf{0}_{(\Psi_2,\Psi_3)} \end{bmatrix} \begin{bmatrix} \rho_d^\top & \rho_a^\top \end{bmatrix} = 0. \tag{B23}$$

By (B22), it holds that

$$\begin{aligned} S_5 = & \rho_{e_1}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} & I_{(\Lambda_3^+,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_u)} \end{bmatrix} \\ & + \rho_{e_3}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_u,\Lambda_2)} & \mathbf{0}_{(\Lambda_u,\Lambda_3^+)} & I_{(\Lambda_u,\Lambda_u)} \end{bmatrix} \\ & + \rho_{l_2}^\top \begin{bmatrix} I_{(\Lambda_2,\Lambda_2)} & \mathbf{0}_{(\Lambda_2,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_2,\Lambda_u)} \end{bmatrix} \\ & + \rho_{l_3}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} & I_{(\Lambda_3^+,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_u)} \end{bmatrix} \\ & = \begin{bmatrix} \rho_{l_2}^\top & \rho_{e_1}^\top + \rho_{l_3}^\top & \rho_{e_3}^\top \end{bmatrix} = 0. \end{aligned} \tag{B24}$$

It implies that $\rho_d = 0$, $\rho_a = 0$, $\rho_{l_2} = 0$ and $\rho_{e_3} = 0$.

Recall the definition of Γ_{sub1} in (B2). Then $\rho^\top \Gamma_1$ reduces to

$$\begin{bmatrix} \rho_{e_1}^\top & \rho_h^\top & \rho_i^\top & \rho_{l_3}^\top \end{bmatrix} \Gamma_{sub1} = 0. \tag{B25}$$

By Lemma $B.3$, Γ_{sub1} has full row rank. Therefore, (B25) implies that $\rho_{e_1} = 0$, $\rho_h = 0$, $\rho_i = 0$, $\rho_{l_3} = 0$. In other words, the row vectors in Γ_1 are linearly independent, giving that MPEC-LICQ holds at v^*.

(iii) $|I_{GH_3}| = 1$ and $|I_{GH_4}| = 0$

By Proposition $A.3$ (b), if $|I_{GH_3}| = 1$ and $|I_{GH_4}| = 0$, Γ in (B1) reduces to the following form

$$\Gamma_3 := \begin{bmatrix}
0_{(\Psi_3,L_1)} & 0_{(\Psi_3,L_2)} & \Gamma_a^3 & (AB^\top)_{(\Psi_3,\cdot)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & \Gamma_b^2 & 0_{(\Psi_2,L_3)} & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Psi_3,L_1)} & \Gamma_c^2 & 0_{(\Psi_3,L_3)} & 0_{(\Psi_3,L_4)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & 0_{(\Psi_2,L_2)} & \Gamma_d^3 & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma_{e_1}^5 \\
0_{(\Lambda_u,L_1)} & 0_{(\Lambda u,L_2)} & 0_{(\Lambda_u,L_3)} & (BB^\top)_{(\Lambda_u,\cdot)} & \Gamma_{e_3}^5 \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & (BB^\top)_{(\Lambda_1,\cdot)} & \Gamma_f^5 \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & \Gamma_g^4 & 0_{(\Lambda_1,L_5)} \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\
\mathbf{1}_{(\Lambda_u,L_1)} & 0_{(\Lambda_u,L_2)} & 0_{(\Lambda u,L_3)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\
0_{(\Lambda_1,L_1)} & 0_{(\Lambda_1,L_2)} & 0_{(\Lambda_1,L_3)} & 0_{(\Lambda_1,L_4)} & \Gamma_{l_1}^5 \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & 0_{(\Lambda_2,L_4)} & \Gamma_{l_2}^5 \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5
\end{bmatrix}.$$

Assume there exists nonzero column vector ρ defined by
$\rho = \left[\rho_a^\top, \rho_b^\top, \rho_c^\top, \rho_d^\top, \rho_{e_1}^\top, \rho_{e_3}^\top, \rho_f^\top, \rho_g^\top, \rho_h^\top, \rho_i^\top, \rho_{l_1}^\top, \rho_{l_2}^\top, \rho_{l_3}^\top \right]^\top$ such that $\rho^\top \Gamma_3 = 0$. It gives that

$$0 = \rho^\top \Gamma_3 := \begin{bmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \end{bmatrix}.$$

where

$$\begin{aligned}
S_1 =& \rho_i^\top \mathbf{1}_{(\Lambda_u,L_1)} = 0, \\
S_2 =& \rho_b^\top \Gamma_b^2 + \rho_c^\top \Gamma_c^2 = 0, && \text{(B26)} \\
S_3 =& \rho_a^\top \Gamma_a^3 + \rho_d^\top \Gamma_d^3 = 0, && \text{(B27)} \\
S_4 =& \rho_a^\top (AB^\top)_{(\Psi_3,\cdot)} + \rho_{e_1}^\top (BB^\top)_{(\Lambda_3^+,\cdot)} + \rho_{e_3}^\top (BB^\top)_{(\Lambda_u,\cdot)} \\
& + \rho_f^\top (BB^\top)_{(\Lambda_1,\cdot)} + \rho_g^\top \Gamma_g^4 + \rho_h^\top \Gamma_h^4 + \rho_i^\top \Gamma_i^4 = 0, \\
S_5 =& \rho_{e_1}^\top \Gamma_{e_1}^5 + \rho_{e_3}^\top \Gamma_{e_3}^5 + \rho_f^\top \Gamma_f^5 + \rho_{l_1}^\top \Gamma_{l_1}^5 + \rho_{l_2}^\top \Gamma_{l_2}^5 + \rho_{l_3}^\top \Gamma_{l_3}^5 = 0. && \text{(B28)}
\end{aligned}$$

By [40, Lemma 1] and (B26), $\rho_b = 0$ and $\rho_c = 0$. By (B27), it holds that

$$S_3 = \rho_a^\top \begin{bmatrix} \mathbf{0}_{(\Psi_3,\Psi_2)} & I_{(\Psi_3,\Psi_3)} \end{bmatrix} + \rho_d^\top \begin{bmatrix} I_{(\Psi_2,\Psi_2)} & \mathbf{0}_{(\Psi_2,\Psi_3)} \end{bmatrix} = \begin{bmatrix} \rho_d^\top & \rho_a^\top \end{bmatrix} = 0. \tag{B29}$$

By (B28), it holds that

$$
\begin{aligned}
S_5 = &\rho_{e_1}^\top \left[\begin{array}{cccc} \mathbf{0}_{(\Lambda_3^+,\Lambda_1)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} & I_{(\Lambda_3^+,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_u)} \end{array} \right] \\
&+ \rho_{e_3}^\top \left[\begin{array}{cccc} \mathbf{0}_{(\Lambda_u,\Lambda_1)} & \mathbf{0}_{(\Lambda_u,\Lambda_2)} & \mathbf{0}_{(\Lambda_u,\Lambda_3^+)} & I_{(\Lambda_u,\Lambda_u)} \end{array} \right] \\
&+ \rho_f^\top \left[\begin{array}{cccc} I_{(\Lambda_1,\Lambda_1)} & \mathbf{0}_{(\Lambda_1,\Lambda_2)} & \mathbf{0}_{(\Lambda_1,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_1,\Lambda_u)} \end{array} \right] \\
&+ \rho_{l_1}^\top \left[\begin{array}{cccc} I_{(\Lambda_1,\Lambda_1)} & \mathbf{0}_{(\Lambda_1,\Lambda_2)} & \mathbf{0}_{(\Lambda_1,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_1,\Lambda_u)} \end{array} \right] \\
&+ \rho_{l_2}^\top \left[\begin{array}{cccc} \mathbf{0}_{(\Lambda_2,\Lambda_1)} & I_{(\Lambda_2,\Lambda_2)} & \mathbf{0}_{(\Lambda_2,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_2,\Lambda_u)} \end{array} \right] \\
&+ \rho_{l_3}^\top \left[\begin{array}{cccc} \mathbf{0}_{(\Lambda_3^+,\Lambda_1)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_2)} & I_{(\Lambda_3^+,\Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+,\Lambda_u)} \end{array} \right] \\
&= \left[\begin{array}{cccc} \rho_f^\top + \rho_{l_1}^\top & \rho_{l_2}^\top & \rho_{e_1}^\top + \rho_{l_3}^\top & \rho_{e_3}^\top \end{array} \right] = 0.
\end{aligned}
\tag{B30}
$$

It implies that $\rho_a = 0$, $\rho_d = 0$, $\rho_{l_2} = 0$ and $\rho_{e_3} = 0$.

Recall the definition of Γ_{sub2} in (B6). Then $\rho^\top \Gamma_3$ reduces to

$$
\left[\rho_{e_1}^\top \quad \rho_f^\top \quad \rho_g^\top \quad \rho_h^\top \quad \rho_i^\top \quad \rho_{l_1}^\top \quad \rho_{l_3}^\top\right] \Gamma_{sub2} = 0. \tag{B31}
$$

By Lemma $B.4$, if $\widehat{a}_3 \neq 0$, Γ_{sub2} has full row rank. Therefore, (B31) implies that $\rho_{e_1} = 0$, $\rho_f = 0$, $\rho_g = 0$, $\rho_h = 0$, $\rho_i = 0$, $\rho_{l_1} = 0$, $\rho_{l_3} = 0$. In other words, the row vectors in Γ_3 are linearly independent, giving that MPEC-LICQ holds at v^*. Otherwise, MPEC-LICQ fails at v^*.

(iv) $|I_{GH_4}| = 1$ and $|I_{GH_3}| = 0$

By Proposition $A.3$ (b), if $|I_{GH_4}| = 1$ and $|I_{GH_3}| = 0$, Γ in (B1) reduces to the following form

$$
\Gamma_4 := \begin{bmatrix}
0_{(\Psi_3,L_1)} & 0_{(\Psi_3,L_2)} & \Gamma_a^3 & (AB^\top)_{(\Psi_3,\cdot)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & \Gamma_b^2 & 0_{(\Psi_2,L_3)} & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Psi_3,L_1)} & \Gamma_c^2 & 0_{(\Psi_3,L_3)} & 0_{(\Psi_3,L_4)} & 0_{(\Psi_3,L_5)} \\
0_{(\Psi_2,L_1)} & 0_{(\Psi_2,L_2)} & \Gamma_d^3 & 0_{(\Psi_2,L_4)} & 0_{(\Psi_2,L_5)} \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & (BB^\top)_{(\Lambda_3^+,\cdot)} & \Gamma_{e_1}^5 \\
0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & (BB^\top)_{(\Lambda_3^c,\cdot)} & \Gamma_{e_2}^5 \\
0_{(\Lambda_u,L_1)} & 0_{(\Lambda u,L_2)} & 0_{(\Lambda_u,L_3)} & (BB^\top)_{(\Lambda_u,\cdot)} & \Gamma_{e_3}^5 \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & \Gamma_h^4 & 0_{(\Lambda_2,L_5)} \\
\mathbf{1}_{(\Lambda_u,L_1)} & 0_{(\Lambda_u,L_2)} & 0_{(\Lambda u,L_3)} & \Gamma_i^4 & 0_{(\Lambda_u,L_5)} \\
\mathbf{1}_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & \Gamma_j^4 & 0_{(\Lambda_3^c,L_5)} \\
0_{(\Lambda_3^c,L_1)} & 0_{(\Lambda_3^c,L_2)} & 0_{(\Lambda_3^c,L_3)} & 0_{(\Lambda_3^c,L_4)} & \Gamma_k^5 \\
0_{(\Lambda_2,L_1)} & 0_{(\Lambda_2,L_2)} & 0_{(\Lambda_2,L_3)} & 0_{(\Lambda_2,L_4)} & \Gamma_{l_2}^5 \\
0_{(\Lambda_3^+,L_1)} & 0_{(\Lambda_3^+,L_2)} & 0_{(\Lambda_3^+,L_3)} & 0_{(\Lambda_3^+,L_4)} & \Gamma_{l_3}^5
\end{bmatrix}.
$$

Assume there exists nonzero column vector ρ defined by $\rho = \left[\rho_a^\top, \rho_b^\top, \rho_c^\top, \rho_d^\top, \rho_{e_1}^\top, \rho_{e_2}^\top, \rho_{e_3}^\top, \rho_h^\top, \rho_i^\top, \rho_j^\top, \rho_k^\top, \rho_{l_2}^\top, \rho_{l_3}^\top \right]^\top$ such that $\rho^\top \Gamma_4 = 0$. It gives that

$$
0 = \rho^\top \Gamma_4 := \left[\begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \right],
$$

where

$$
\begin{aligned}
S_1 =& \rho_i^\top \mathbf{1}_{(\Lambda_u, L_1)} + \rho_j^\top \mathbf{1}_{(\Lambda_3^c, L_1)} = 0, \\
S_2 =& \rho_b^\top \Gamma_b^2 + \rho_c^\top \Gamma_c^2 = 0, && \text{(B32)} \\
S_3 =& \rho_a^\top \Gamma_a^3 + \rho_d^\top \Gamma_d^3 = 0, && \text{(B33)} \\
S_4 =& \rho_a^\top (AB^\top)_{(\Psi_3, \cdot)} + \rho_{e_1}^\top (BB^\top)_{(\Lambda_3^+, \cdot)} + \rho_{e_2}^\top (BB^\top)_{(\Lambda_3^c, \cdot)} \\
& + \rho_{e_3}^\top (BB^\top)_{(\Lambda_u, \cdot)} + \rho_h^\top \Gamma_h^4 + \rho_i^\top \Gamma_i^4 + \rho_j^\top \Gamma_j^4 = 0, \\
S_5 =& \rho_{e_1}^\top \Gamma_{e_1}^5 + \rho_{e_2}^\top \Gamma_{e_2}^5 + \rho_{e_3}^\top \Gamma_{e_3}^5 + \rho_k^\top \Gamma_k^5 + \rho_{l_2}^\top \Gamma_{l_2}^5 + \rho_{l_3}^\top \Gamma_{l_3}^5 = 0. && \text{(B34)}
\end{aligned}
$$

By [40, Lemma 1] and (B32), $\rho_b = 0$ and $\rho_c = 0$. By (B33), it holds that

$$
S_3 = \rho_a^\top \begin{bmatrix} \mathbf{0}_{(\Psi_3, \Psi_2)} & I_{(\Psi_3, \Psi_3)} \end{bmatrix} + \rho_d^\top \begin{bmatrix} I_{(\Psi_2, \Psi_2)} & \mathbf{0}_{(\Psi_2, \Psi_3)} \end{bmatrix} = \begin{bmatrix} \rho_d^\top & \rho_a^\top \end{bmatrix} = 0. \tag{B35}
$$

By (B34), it holds that

$$
\begin{aligned}
S_5 = & \rho_{e_1}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^+, \Lambda_2)} & I_{(\Lambda_3^+, \Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+, \Lambda_3^c)} & \mathbf{0}_{(\Lambda_3^+, \Lambda_u)} \end{bmatrix} \\
& + \rho_{e_2}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^c, \Lambda_2)} & \mathbf{0}_{(\Lambda_3^c, \Lambda_3^+)} & I_{(\Lambda_3^c, \Lambda_3^c)} & \mathbf{0}_{(\Lambda_3^c, \Lambda_u)} \end{bmatrix} \\
& + \rho_{e_3}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_u, \Lambda_2)} & \mathbf{0}_{(\Lambda_u, \Lambda_3^+)} & \mathbf{0}_{(\Lambda_u, \Lambda_3^c)} & I_{(\Lambda_u, \Lambda_u)} \end{bmatrix} \\
& + \rho_k^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^c, \Lambda_2)} & \mathbf{0}_{(\Lambda_3^c, \Lambda_3^+)} & I_{(\Lambda_3^c, \Lambda_3^c)} & \mathbf{0}_{(\Lambda_3^c, \Lambda_u)} \end{bmatrix} \\
& + \rho_{l_2}^\top \begin{bmatrix} I_{(\Lambda_2, \Lambda_2)} & \mathbf{0}_{(\Lambda_2, \Lambda_3^+)} & \mathbf{0}_{(\Lambda_2, \Lambda_3^c)} & \mathbf{0}_{(\Lambda_2, \Lambda_u)} \end{bmatrix} \\
& + \rho_{l_3}^\top \begin{bmatrix} \mathbf{0}_{(\Lambda_3^+, \Lambda_2)} & I_{(\Lambda_3^+, \Lambda_3^+)} & \mathbf{0}_{(\Lambda_3^+, \Lambda_3^c)} & \mathbf{0}_{(\Lambda_3^+, \Lambda_u)} \end{bmatrix} \\
& = \begin{bmatrix} \rho_{l_2}^\top & \rho_{e_1}^\top + \rho_{l_3}^\top & \rho_{e_2}^\top + \rho_k^\top & \rho_{e_3}^\top \end{bmatrix} = 0.
\end{aligned} \tag{B36}
$$

It implies that $\rho_a = 0$, $\rho_f = 0$, $\rho_{l_2} = 0$ and $\rho_{e_3} = 0$.

Recall the definition of Γ_{sub3} in (B13). Then $\rho^\top \Gamma_4$ reduces to

$$
\begin{bmatrix} \rho_{e_1}^\top & \rho_{e_2}^\top & \rho_h^\top & \rho_i^\top & \rho_j^\top & \rho_k^\top & \rho_{l_3}^\top \end{bmatrix} \Gamma_{sub3} = 0. \tag{B37}
$$

By Lemma $B.5$, if $\widehat{a}_4 \neq 0$, Γ_{sub3} has full row rank. Therefore, (B37) implies that $\rho_{e_1} = 0$, $\rho_{e_2} = 0$, $\rho_h = 0$, $\rho_i = 0$, $\rho_j = 0$, $\rho_k = 0$, $\rho_{l_3} = 0$. In other words, the row vectors in Γ_4 are linearly independent, giving that MPEC-LICQ holds at v^*. Otherwise, MPEC-LICQ fails at v^*.

□