

Exact Optimization: Part I

Li-Gang Lin^{a,1} and Yew-Wen Liang^{b,1}

^aDepartment of Mechanical Engineering, National Central University, Taoyuan, 32001 Taiwan and Department of Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, 39106 Magdeburg, Germany; ^bInstitute of Electrical Control Engineering, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan

This manuscript was compiled on May 31, 2022

The nonlinear programming is being explicitly analyzed via a novel method/perspective and from a bottom-up manner. The philosophy is based on the recent findings on the convex quadratic equation (CQE), which help clarify a geometric interpretation that relates CQE to the convex quadratic function (CQF). More specifically, regarding the solvability of CQE, its necessary and sufficient condition as well as a unified parameterization of all the solutions has just been analytically formulated in the recent literature. Moving forward, the understanding of CQE is utilized to describe the geometric structure of CQF and the CQE-CQF relation. All these results are shown closely related to a basis in the optimization literature, namely the quadratic programming (QP). For the first time from this viewpoint, the QPs subject to equality, inequality, equality-and-inequality, and extended constraints can be algebraically solved in derivative-free closed formulae, respectively. All the results are derived without knowing a feasible point, a priori and any time during the process. Moreover, a variety of practical validations, evaluations, and comparisons with the management literature and beyond are provided to demonstrate the superiority of the proposed method, such as the enhanced computational efficiency and a counterexample to the state of the art.

nonlinear/quadratic programming | matrix algebra | convex quadratic function and equation

??(1)

Recently, (2) provided more understanding of the mathematical fundamentals: convex quadratic equation and function (CQE and CQF). In particular, (2) has analytically and completely formulated an equivalent solvability condition of CQE and a parameterization of all its solutions. These results encourage new investigations into a spectrum of applications, such as the one to nonlinear optimal control (Sec. 4 in (2)). The findings not only facilitate preliminary optimality recovery using the state-dependent Riccati equation (SDRE) scheme and its differential variant (see (3) for a survey), but also enhance the computational performance of the other major application in this article. Specifically, we analyze the nonlinear programming (4)/convex optimization (5) from a new perspective based on (2), which particularly boasts impacts in management science, operations research, and control engineering, for example, the model predictive control (MPC; (6)). It is worth mentioning that (7) and (8) reveal more connection between the two applications, that is, nonlinear MPC using “state-dependent coefficient” (a design flexibility) in the SDRE scheme. Such a unified framework is envisioned/pioneered by (9) and also pursued by this article.

With regard to the convex optimization, a basis in the literature of nonlinear programming algorithms is the quadratic programming (QP; (4)), which includes the linear programming (LP) as a special case and is closely linked with the CQE and CQF (10, 11). Regarding the former reference, (10) utilize a solution subset of CQE, which is reformulated from

the explicitly considered CQF and in terms of the polyhedral characterization (12). Considering the latter, (11) make use of an important equality that is related with a CQF, as well as the existence and uniqueness of the corresponding optimum (point), with respect to a QP-like optimization problem. This supports the construction of a feasible sequence converging to the optimum. Generally speaking, more iterative/numerical QP solvers are available in the mainstream literature (13), which are mostly analyzed from the differential perspective and can be classified by three levels (5): i) equality-constrained QP; ii) linear equality-constrained optimization problem with twice-differentiable objective, as solved using Newton’s method by reducing it to a sequence of equality-constrained QPs; and, after further imposing inequality constraints, iii) interior-point methods, which reduce the problem to a sequence of ii). It is worth noting that the equality-constrained QP solver is the most fundamental to build up various algorithms, subject to the constraints at the three levels. In addition, one area of great creativity among existing algorithms, namely the conjugate direction methods, illustrates that *detailed analysis of QP leads to significant practical advances* (4), such as solution accuracy, computation time, and more reported in (14) (see Sec. 2.5 later for a comparison discussion). Moreover, the constrained QP can also be regarded as an optimization problem with the objective of constrained convex quadratic function (CCQF). These observations quite endorse the importance of QP and CQF, while motivate this article to provide further theoretical support (extended from (2)). *One highlight is the first closed-form QP solver subject*

Significance Statement

Optimization has played a major role throughout human civilization. We have observed its increasing importance in recent years in diverse fields including biology, engineering, finance, and physics. These fields pose challenging optimization problems for humans and/or machines, notably, computational issues that have been anticipated for enhancement. This article responds to the long-standing call and initiates a research direction toward *Exact Optimization* (rather than the common *Numerical Optimization*). This research starts with an examination of the core of optimization literature, quadratic programming, and presents exact/specific solutions to its general problem formulation and an extension of this formulation. The philosophy of this research builds on the latest understanding of convex quadratic equation, and demonstrates various advantages from both analytical and practical perspectives.

L.-G.L. completed this work while appreciates valuable discussions with Y.-W.L..

The authors declare no competing interest.

¹To whom correspondence should be addressed. E-mail: lglin@g.ncu.edu.tw.

to a variety of constraints. The analytical philosophy is inspired by (15) (vector space methods) as well as the computational interest among a variety of research fields (4, 16–18); but based on a novel understanding and perspective of the associated CQE, and interpretations of its relations to CQF and QP, respectively.

Following the previous findings of CQE and CQF (2), this article deals with another application to nonlinear programming and – in response to, for example, the open question in Sec. VI in (16) – the main result is exactly, autonomously, and algebraically solving the QPs that

- are subject to equality, inequality, equality-and-inequality, and extended constraints, resp.;
- refrain from any derivative (with only functional information);
- consist of finite steps, the maximum of which is explicitly pre-determined;
- and do not require any knowledge of a feasible point (any time before and during the process).

The derivation leverages the vector space methods such as the singular value decomposition (SVD), and investigates from a perspective based on (2) and geometric interpretations that clarify the structures of CQE, CQF, and CQE-CQF relation, respectively.

The rest of this article is arranged as follows. The notations and the formulation of main problem are provided in Sec. 1. The method proposed in (2) is applied to this focused application “nonlinear programming/convex optimization” in Sec. 2. At first, the analysis of extreme/minimal values of CQF is performed in Sec. 2.1, which is followed by a complete characterization of the explicit solution to the equality-constrained QP in Sec. 2.2. As a step forward, the two preliminary subsections are shown indispensable for the proposed closed-form QP solver in Sec. 2.3, and an extension subject to the constraints in Sec. 2.4. Therefore, discussions on differences and advantages with respect to existing solvers in the literature (management, operations research, and beyond) can be justified in Sec. 2.5. Finally, three representative examples are demonstrated to offer computational evidences and additional insights (verified using MATLAB®) in Sec. 3, while concluding remarks are given in Sec. 4.

1. Notation and Problem Formulation

Unless specified otherwise, we use the notational conventions: $\|\cdot\|$, $\mathcal{N}(\cdot)$, $\mathcal{R}(\cdot)$, $\lambda(\cdot)$, $(\cdot)^\dagger$, $(\cdot)^{\dagger/2}$, and $(\cdot)^T$ denote the Euclidean norm, null space, range space, spectrum/eigenvalues, Moore-Penrose generalized inverse (shortened as “pseudoinverse” afterward), square root of the pseudoinverse (19), and transpose of a vector or matrix, respectively. Additionally, we denote $(\cdot)^\perp$ as the orthogonal complement of a vector space, \oplus the direct sum of vector spaces, \mathbf{e}_1 the first standard basis vector in \mathbb{R}^n , and $\{\sigma_i\}_{i=1}^n$ the set of singular values of an $n \times n$ matrix with rank r ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ and $\sigma_r > 0$). Moreover, in accordance with (4), denote $P \succ 0$ (resp., $P \succeq 0$), if $P = P^T \in \mathbb{R}^{n \times n}$ is positive definite (resp., semidefinite).

Consider the following CQF $F_{\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}$ (4),

$$F_{\mathbf{x}}(\mathbf{x}) = \mathbf{x}^T P \mathbf{x} / 2 + \mathbf{q}^T \mathbf{x} + s, \quad [1]$$

where both $\mathbf{x}, \mathbf{q} \in \mathbb{R}^n$, $P = P^T \in \mathbb{R}^{n \times n}$, $P \succeq 0$, and $s \in \mathbb{R}$. This quadratic function is convex, if its Hessian matrix is $P \succeq 0$; and strictly convex, if $P \succ 0$. Note that the early findings in (2) start with the formulation of CQE, as in Eq. [2] later, to emphasize its dominance. Then, in this continuing work, the concept of CQF dominates, and thus we also give its formulation in the very beginning [1] and interpret its close relation to CQE later in detail. More specifically, *in this article, the main focus shifts to explicitly analyzing and solving the following problem.*

Problem 1.1. From a new viewpoint based on the solution to Problem 2.1 in (2) as well as the CQE-CQF relation, solve the QPs in an analytical, autonomous, and derivative-free manner and subject to equality, inequality, equality-and-inequality, and extended constraints, respectively.

2. Application to Convex Optimization (QP)

The main concepts behind all the derivations in this section are the geometrical interpretations of i) the value/difference “ $\mathbf{q}^T P^\dagger \mathbf{q} / 2 - s$ ” according to the solvability conditions in Eqs. [3] and [5] later and ii) the solution parameterization in Eqs. [4] and [7]. Notably, in i), $P^\dagger = P^{-1}$ when P is nonsingular. Moreover, the novelty can be revealed by that i) facilitates the derivations that exploit the (additional) *analysis perspective* from the image of the CQF [1]. As a step forward, the *hierarchical layers* in the parameterization, with respect to ii), exactly categorize the location(s) of the unconstrained optimum/optima. In the following, Sec. 2.1 considers the unconstrained optimization, which are preliminaries supporting the derivations of constrained ones in the remaining subsections. Following the QP formulation/definition in literature (4, 5), Sec. 2.4 enlarges the included types of constraints, without introducing any excessive variable but, actually, reducing to a lower-dimensional unconstrained equivalent problem. Finally, Sec. 2.5 discusses advantages or differences over existing solvers (13). Remarkably, we recall the following lemma because it is largely used in this article, where the notation is consistent with (2, 20) to avoid misleading/confusion.

Lemma 2.1. (Solvability and Solutions of CQE by Theorem 3.1 in (2))

Consider the CQE below (4):

$$\mathbf{z}^T M \mathbf{z} + \mathbf{k}^T \mathbf{z} + c = 0, \quad [2]$$

where both $\mathbf{z}, \mathbf{k} \in \mathbb{R}^n$, $M = M^T \in \mathbb{R}^{n \times n}$, $M \succeq 0$, and $c \in \mathbb{R}$.

A) If $\text{rank}(M) = n$, then CQE [2] is solvable, if and only if (iff)

$$\mathbf{k}^T M^{-1} \mathbf{k} \geq 4c. \quad [3]$$

Accordingly, the set of solutions are, and can be parameterized by,

$$\mathbf{z} = -M^{-1} \mathbf{k} / 2 + \sqrt{\mathbf{k}^T M^{-1} \mathbf{k} / 4 - c} \cdot M^{-1/2} \cdot \mathbf{v}, \quad [4]$$

where $\mathbf{v} \in \mathbb{R}^n$ and $\|\mathbf{v}\| = 1$.

B) Otherwise ($\text{rank}(M) < n$), it is solvable, iff [5] or [6], where

$$\mathbf{k} \in \mathcal{R}(M) \text{ and } \mathbf{k}^T M^\dagger \mathbf{k} \geq 4c, \quad [5]$$

$$\mathbf{k} \notin \mathcal{R}(M). \quad [6]$$

Accordingly, the sets of solutions are, and can be parameterized by, respectively,

a) for Condition [5],

$$\mathbf{z} = -M^\dagger \mathbf{k}/2 + \sqrt{\mathbf{k}^T M^\dagger \mathbf{k}/4 - c} \cdot M^{\dagger/2} \boldsymbol{\rho} + \boldsymbol{\varepsilon}, \quad [7]$$

where both $\boldsymbol{\rho}, \boldsymbol{\varepsilon} \in \mathbb{R}^n$, $\boldsymbol{\rho} \in \mathcal{R}(M)$, $\|\boldsymbol{\rho}\| = 1$, and $\boldsymbol{\varepsilon} \in \mathcal{N}(M)$;

b) for Condition [6], decompose $\mathbf{k} = \mathbf{k}_M + \mathbf{k}_{M^\perp}$, where $\mathbf{k}_M \in \mathcal{R}(M)$, $\mathbf{k}_{M^\perp} \in \mathcal{R}(M)^\perp$, and both $\mathbf{k}_M, \mathbf{k}_{M^\perp} \in \mathbb{R}^n$. Then,

$$\mathbf{z} = -(F_w/\|\mathbf{k}_{M^\perp}\|^2) \cdot \mathbf{k}_{M^\perp} + \boldsymbol{\varphi} + \boldsymbol{\tau}, \quad [8]$$

where the CQF $F_w : \mathcal{R}(M) \subset \mathbb{R}^n \rightarrow \mathbb{R}$,

$$F_w(\mathbf{w}) = \mathbf{w}^T M \mathbf{w} + \mathbf{k}_M^T \mathbf{w} + c, \quad [9]$$

all $\mathbf{w}, \boldsymbol{\varphi}, \boldsymbol{\tau} \in \mathbb{R}^n$, both $\mathbf{w}, \boldsymbol{\tau} \in \mathcal{R}(M)$, and $\boldsymbol{\varphi} \in \mathcal{N}(M) \cap \mathcal{N}(\mathbf{k}^T)$.

2.1. Unconstrained QP.

Theorem 2.1. (Solutions to Unconstrained QP)

Consider the optimization problem, minimize $F_{\mathbf{x}}$ [1] with respect to \mathbf{x} .

- i) The preimage of any level set of $F_{\mathbf{x}}$ [1] can be parameterized by Eqs. [4, 7-9], respectively, where $M = P/2$, $\mathbf{k} = \mathbf{q}$, $c = s - l$, and $l \in \mathbb{R}$ is any level set value of $F_{\mathbf{x}}$.
- ii) The optimal value is finite, iff $\mathbf{q} \in \mathcal{R}(P)$.
- iii) The finite optimal value is $l^* = s - \mathbf{q}^T P^\dagger \mathbf{q}/2$. And the corresponding unique optimum is, or optima are all parameterized by, $\mathbf{x}^* = \mathbf{x}_p^* + \boldsymbol{\xi}$, where $\mathbf{x}_p^* := -P^\dagger \mathbf{q}$, $\boldsymbol{\xi} \in \mathbb{R}^n$, and $\boldsymbol{\xi} \in \mathcal{N}(P)$.

Proof. See Appendix A for a complete and unified proof, which is new with regard to ii) and iii), while first for i). Along with the proof, a geometric interpretation for the novel concept of “critical shift” (resp., “hierarchical layers”) is given in Fig. 1 (resp., Remark A.1 and Fig. 2). \square

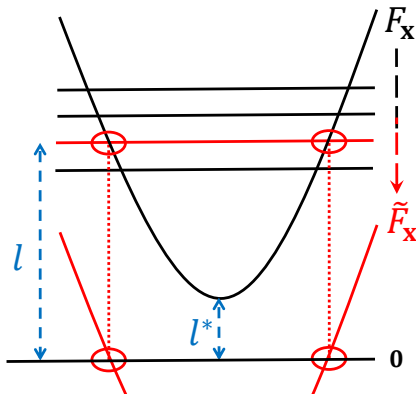


Fig. 1. Geometric interpretation of a “shift” l and the critical one “ l^* ” in Sec. 2.1.

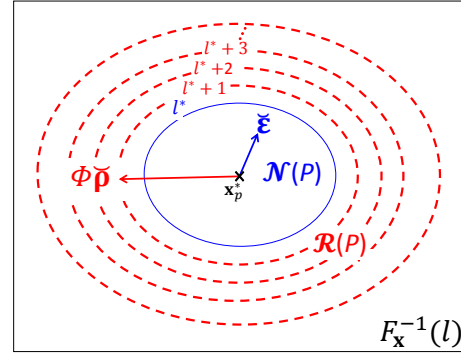


Fig. 2. Geometric interpretation of “hierarchical layers” in Sec. 2.1 (Remark A.1).

Remark 2.1. Based on Lemma 2.1, the results of ii) and iii) in Theorem 2.1 are consistent with the literature, for example, (5) as analyzed from the differential perspective, which thus and also endorses Lemma 2.1. All the results in Theorem 2.1, particularly i), will be utilized for the constrained and extended QP problems, later in Secs. 2.2 to 2.4. On the other hand, more potentials are expected toward the applications to least-squares approximation and regression analysis. Finally, according to the respective items in Theorem 2.1, it is worth remarking that:

- i) A geometric interpretation is illustrated in Fig. 1.
- ii) From its proof, typically Eq. [19], $F_{\mathbf{x}}$ is always unbounded above; bounded below, iff $\mathbf{q} \in \mathcal{R}(P)$.
- iii) Regarding the special but popular case of nonsingular P (4), then $P^\dagger = P^{-1}$, $\mathcal{N}(P) = \{\mathbf{0}\}$, the unique optimum $\mathbf{x}^* = -P^{-1} \mathbf{q}$, and $l^* = s - \mathbf{q}^T P^{-1} \mathbf{q}/2$.

Remark 2.2. After introducing the concept of “critical shift l^* ” in this section, as illustrated in Fig. 1, it suffices to complement the examination previously in Remark 3.2 in (2). By ii) and iii) of Theorem 2.1, the optimal/minimal value (l^*) is finite and equals zero (no shift) for $M \geq 0$, either singular or nonsingular. In other words, $\mathbf{z}^T M \mathbf{z} > 0$, for all $\mathbf{z} \notin \mathcal{N}(M)$. To sum up, combining Remark 3.2 in (2) and this Remark 2.2 endorse both Theorem 2.1 and Lemma 2.1 above, according to the basic mathematical properties of positive definiteness and semidefiniteness.

2.2. Equality-Constrained QP.

Theorem 2.2. (Solutions to Equality-Constrained QP)

Consider the optimization problem,

$$\text{minimize } F_{\mathbf{x}} \text{ subject to } A\mathbf{x} = \mathbf{b}, \quad [10]$$

where $F_{\mathbf{x}}$ is in Eq. [1], $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = m \in \{1, 2, \dots, n-1\}$, and $\mathbf{b} \in \mathbb{R}^m$. From the SVD of A , let a $V_2 \in \mathbb{R}^{n \times (n-m)}$ be given, which is associated with $\mathcal{R}(V_2) = \mathcal{N}(A)$. Denote $\bar{l} \in \mathbb{R}$ as any level set value of $F_{\mathbf{x}}$.

A) If $V_2^T P V_2 \neq 0 \Leftrightarrow \mathcal{R}(P) \cap \mathcal{N}(A) \neq \{\mathbf{0}\} \Leftrightarrow \mathcal{N}(A) \not\subseteq \mathcal{N}(P)$,

a) Problem [10] is equivalent to an unconstrained QP,

b) the preimage of any level set/value of " $F_{\mathbf{x}}$ subject to $A\mathbf{x} = \mathbf{b}$ " can be respectively parameterized by

$$\mathbf{x} = A^\dagger \mathbf{b} + V_2 \mathbf{y}, \quad [11]$$

where $\mathbf{y} \in \mathbb{R}^{n-m}$ is determined by Eqs. [4, 7-9], with $\mathbf{z} = \mathbf{y}$, $M = V_2^T P V_2 / 2$, $\mathbf{k} = V_2^T (\mathbf{q} + P A^\dagger \mathbf{b})$, and $c = s + [\mathbf{q}^T + \mathbf{b}^T (A^\dagger)^T P / 2] A^\dagger \mathbf{b} - \bar{l}$,

c) among all \bar{l} in Ab), the optimal one/minimum is finite, iff $V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \in \mathcal{R}(V_2^T P V_2)$,

d) following Ac), this finite optimal value is $\bar{l}^* \in \mathbb{R}$,

$$\begin{aligned} \bar{l}^* = & [\mathbf{b}^T (A^\dagger)^T P / 2 + \mathbf{q}^T] [I_n - V_2 (V_2^T P V_2)^\dagger V_2^T P] A^\dagger \mathbf{b} \\ & + s - \mathbf{q}^T V_2 (V_2^T P V_2)^\dagger V_2^T \mathbf{q} / 2, \end{aligned} \quad [12]$$

and the corresponding unique optimum is, or optima are all, parameterized by,

$$\mathbf{x}^* = \mathbf{x}_p^* + V_2 \boldsymbol{\varepsilon}^*, \quad [13]$$

where $\mathbf{x}_p^* := A^\dagger \mathbf{b} - V_2 (V_2^T P V_2)^\dagger V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \in \mathbb{R}^n$ denotes the particular solution of $\mathbf{x}^* \in \mathbb{R}^n$, $\boldsymbol{\varepsilon}^* \in \mathbb{R}^{n-m}$, and $\boldsymbol{\varepsilon}^* \in \mathcal{N}(V_2^T P V_2)$,

e) to further categorize Ad), the optimum is unique, iff $\mathcal{N}(A) \cap \mathcal{N}(P) = \{\mathbf{0}\}$;

B) else, if $V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \neq \mathbf{0}$,

a) Problem [10] is equivalent to an unconstrained LP, and thus unbounded,

b) the preimage of any level set/value of " $F_{\mathbf{x}}$ subject to $A\mathbf{x} = \mathbf{b}$ " can be respectively parameterized by Eq. [11], where \mathbf{y} is determined by Lemma 3.1 in (2), with $\mathbf{z} = \mathbf{y}$, $\boldsymbol{\zeta} = V_2^T (\mathbf{q} + P A^\dagger \mathbf{b})$, and $\nu = s + [\mathbf{q}^T + \mathbf{b}^T (A^\dagger)^T P / 2] A^\dagger \mathbf{b} - \bar{l}$;

C) else,

a) Problem [10] is equivalent to a constant function,

b) its value is " $s + [\mathbf{q}^T + \mathbf{b}^T (A^\dagger)^T P / 2] A^\dagger \mathbf{b}$ ",

c) the (only) preimage is parameterized by Eq. [11], for all \mathbf{y} .

Proof. See Appendix B for the first, complete, and unified proof. \square

Remark 2.3. Both Theorems 2.1 and 2.2 present the results within a unified framework. In particular, the followings summarize further analyses of the more significant, quadratic case A) in Theorem 2.2.

- Using the common terminology, such as Sec. 4.1.1 in (5), the considered problem is feasible since $\mathbf{b} \in \mathcal{R}(A)$. The preimage of any level set of " $F_{\mathbf{x}}$ subject to $A\mathbf{x} = \mathbf{b}$ " is the *feasible set*. The definitions of the optimum/optima and corresponding optimal value follow the general consensus (5) throughout Sec. 2, and thus omitted for brevity.

ii) Regarding the consistent constraint $A\mathbf{x} = \mathbf{b}$ (in other words, $\mathbf{b} \in \mathcal{R}(A)$ where A is of full rank) in this theorem, we omit the only remaining case (that is, $\text{rank}(A) = n$) for a more concise presentation. This corresponds to $\mathcal{N}(A) = \{\mathbf{0}\}$ and that there is only one feasible point $\mathbf{x}^* = A^\dagger \mathbf{b}$. However, by an extended definition of $V_2 := 0_n$, the results of Ac)-Ae) in this theorem still apply to this sub-QP problem.

iii) In the proof, the considered QP problem [10] is equivalently reformulated as an unconstrained QP: minimize $F_{\mathbf{y}}$ [23] with respect to \mathbf{y} . This is viable, iff $V_2^T P V_2 \neq 0$. With regard to the unconstrained QP, $F_{\mathbf{y}}$ (resp., the original QP Problem [10]) is always unbounded above; bounded below, iff $V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \in \mathcal{R}(V_2^T P V_2)$. Note that a geometric interpretation regarding the CQF $F_{\mathbf{y}}$ (resp., QP [10] for the optimal value) can be inferred from Fig. 1.

Remark 2.4. An example to endorse Theorem 2.2 comes from the minimum or least-norm problem (5, 21). More specifically, to find the optimum $\mathbf{x}_{l_n}^*$ that (satisfies the consistent under-determined constraint $A\mathbf{x} = \mathbf{b}$ and) is of least l_2 -norm, which (norm) is most commonly adopted. This is in the applicable form using this theorem, with $P = 2I_n \succ 0$, $\mathbf{q} = \mathbf{0}$, and $s = 0$. Accordingly, this is a QP given $\mathcal{R}(P) \cap \mathcal{N}(A) = \mathcal{N}(A)$, which (null space) is at least one-dimensional and thus nonzero. Moreover, the optimal value (denoted $\bar{l}_{l_n}^*$) is finite since $V_2^T A^\dagger \mathbf{b} = V_2^T V_1 \Gamma_1^{-1} W^T \mathbf{b} = \mathbf{0} \in \mathcal{R}(I_m) = \mathbb{R}^m$, where both $V_1 \in \mathbb{R}^{n \times m}$ being orthogonal to V_2 and $A^\dagger = V_1 \Gamma_1^{-1} W^T \in \mathbb{R}^{n \times m}$ are given in Eq. [22] in the proof. Specifically,

$$\begin{aligned} \bar{l}_{l_n}^* = & \mathbf{b}^T (A^\dagger)^T A^\dagger \mathbf{b} - \mathbf{b}^T (A^\dagger)^T V_2 (V_2^T V_2)^{-1} \underbrace{V_2^T A^\dagger \mathbf{b}}_{\mathbf{0}} \\ = & \|A^\dagger \mathbf{b}\|^2. \end{aligned}$$

Correspondingly and similarly, the unique optimum is

$$\begin{aligned} \mathbf{x}_{l_n}^* = & A^\dagger \mathbf{b} - 2V_2 (V_2^T P V_2)^\dagger V_2^T A^\dagger \mathbf{b} + V_2 \boldsymbol{\varepsilon}^* \\ = & A^\dagger \mathbf{b}, \end{aligned} \quad [14]$$

where $\boldsymbol{\varepsilon}^* \in \mathcal{N}(V_2^T V_2) = (\mathbb{R}^{n-m})^\perp = \{\mathbf{0}\}$. As expected, these results agree with the literature (for example, (5)).

Remark 2.5. This remark focuses upon the special case of $m = 1$, and more efficiently constructs an example of V_2 in Theorem 2.2. Without loss of generality, let $A = \mathbf{a}^T \in \mathbb{R}^{1 \times n}$ be normalized. Given the results in Theorem 4.2 in (2) and further analyses in Remark 4.4 therein, V_2 is readily available by computing the last $(n-1)$ columns of $H_{\mathbf{a}}^\mathbf{a} := I_n - 2\boldsymbol{\iota}_{\mathbf{a}} \boldsymbol{\iota}_{\mathbf{a}}^T$, where $\boldsymbol{\iota}_{\mathbf{a}} := (\mathbf{a} - \mathbf{e}_1) / (\|\mathbf{a} - \mathbf{e}_1\|)$ and $Y = I_{n-1}$. This remark also indicates a research direction, which aims at extending to general m in the equality-constrained QP [10] based on the generalization of Theorem 4.2 in (2), and thus involves the orthogonalization on the rows of A in the first place.

2.3. QP. Consider the optimization problem (4),

$$\begin{aligned} & \text{minimize } F_{\mathbf{x}} \\ & \text{subject to } A\mathbf{x} = \mathbf{b}, \\ & \quad \mathbf{c}_i^T \mathbf{x} \leq d_i, \quad i \in \mathcal{I} = \{1, 2, \dots, \kappa\}, \end{aligned} \quad [15]$$

where $F_{\mathbf{x}}$ is in Eq. [1], $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = m \in \{1, 2, \dots, n-1\}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c}_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, and there exists a point that satisfies all the constraints.

334 To this general QP Problem [15], an explicit solver –
 335 namely, without any iterations, approximations, or assump-
 336 tions as in the classical/numerical optimization (4, 13) – is
 337 available in Algorithm 1, whose proof is in Appendix C (to-
 338 gether with geometric demonstrations via Figs. 3 and 4).

Algorithm 1 (Closed-Form QP Solver)

Require: $P, \mathbf{q}, s, A, \mathbf{b}, \{c_i, d_i\}_{i \in \mathcal{I}}$ as defined in [15]

Require: (Initial Conditions) $\tilde{\mathcal{L}} = \tilde{\mathcal{X}} = \tilde{\mathcal{I}} = \tilde{\mathcal{I}}^* = \emptyset$

Ensure: The optimal value \tilde{l}^* and the (resp., a subset of)
 corresponding optimum (resp., optima) \mathbf{x}^*

```

1: compute the SVD of  $A$  till  $\mathcal{R}(V_2) = \mathcal{N}(A)$  is obtained
2: if  $V_2^T P V_2 \neq 0$  then
3:    $\mathbf{x}_p^* = A^\dagger \mathbf{b} - V_2(V_2^T P V_2)^\dagger V_2^T (\mathbf{q} + P A^\dagger \mathbf{b})$ 
4:    $\tilde{l}^* = [\mathbf{b}^T (A^\dagger)^T P / 2 + \mathbf{q}^T] [I_n - V_2(V_2^T P V_2)^\dagger V_2^T P] A^\dagger \mathbf{b} +$ 
      $s - \mathbf{q}^T V_2(V_2^T P V_2)^\dagger V_2^T \mathbf{q} / 2$ 
5:   if " $V_2^T P V_2 \succ 0$ " and " $\mathbf{c}_i^T \mathbf{x}_p^* \leq d_i$  for all  $i \in \mathcal{I}$ " then
6:     return  $\tilde{l}^* = \tilde{l}^*, \mathbf{x}^* = \mathbf{x}_p^*$ , and  $\tilde{\mathcal{I}}^*$ 
7:   else if " $V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \in \mathcal{R}(V_2^T P V_2)$ " and " $\mathbf{c}_i^T \mathbf{x}_p^* \leq d_i$ "
     for all  $i \in \mathcal{I}$  then
8:      $(\tilde{l}^*, \mathbf{x}_p^*, \emptyset) \in \tilde{\mathcal{L}} \times \tilde{\mathcal{X}} \times \tilde{\mathcal{I}}$ 
9:     go to line 11
10: else
11:   for all  $\mathcal{I}_j := \{j_1, j_2, \dots, j_k\} \subseteq \mathcal{I}$ , where  $k \in \{1, 2, \dots, \kappa\}$ , do
12:     let  $\tilde{\mathbf{C}} = [\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \dots, \mathbf{c}_{j_k}]^T$ ,  $\tilde{\mathbf{A}} = [A^T, \tilde{\mathbf{C}}^T]^T$ , and
      $\tilde{\mathbf{b}} = [\mathbf{b}^T, d_{j_1}, d_{j_2}, \dots, d_{j_k}]^T$ 
13:     if  $\tilde{\mathbf{b}} \in \mathcal{R}(\tilde{\mathbf{A}})$  then
14:       if  $\text{rank}(\tilde{\mathbf{A}}) = m + k < n$  then
15:         compute the SVD of  $\tilde{\mathbf{A}}$  till  $\mathcal{R}(\tilde{\mathbf{V}}_2) = \mathcal{N}(\tilde{\mathbf{A}})$ 
         is obtained
16:         if " $\tilde{\mathbf{V}}_2^T P \tilde{\mathbf{V}}_2 \neq 0$ ", " $\tilde{\mathbf{V}}_2^T (\mathbf{q} + P \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{b}}) \in$ 
            $\mathcal{R}(\tilde{\mathbf{V}}_2^T P \tilde{\mathbf{V}}_2)$ ", and " $\mathbf{c}_i^T \tilde{\mathbf{x}}_p^* \leq d_i$  for all  $i \in \mathcal{I} \setminus \mathcal{I}_j$ , where
17:            $\tilde{\mathbf{x}}_p^* = \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{b}} - V_2(V_2^T P V_2)^\dagger V_2^T (\mathbf{q} + P \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{b}})$ " then
18:              $\tilde{l}^* = [\tilde{\mathbf{b}}^T (\tilde{\mathbf{A}}^\dagger)^T P / 2 + \mathbf{q}^T] [I_n -$ 
               $V_2(V_2^T P V_2)^\dagger V_2^T P] \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{b}} + s - \mathbf{q}^T V_2(V_2^T P V_2)^\dagger V_2^T \mathbf{q} / 2$ 
19:              $(\tilde{l}^*, \tilde{\mathbf{x}}_p^*, \mathcal{I}_j) \in \tilde{\mathcal{L}} \times \tilde{\mathcal{X}} \times \tilde{\mathcal{I}}$ 
20:           else if  $\text{rank}(\tilde{\mathbf{A}}) = n$  then
21:             compute  $\hat{\mathbf{x}} = \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{b}}$ 
22:             if  $\mathbf{c}_i^T \hat{\mathbf{x}} \leq d_i$ , for all  $i \in \mathcal{I} \setminus \mathcal{I}_j$  then
23:                $\hat{l} = F_{\mathbf{x}}(\hat{\mathbf{x}})$ , where  $F_{\mathbf{x}}$  is from [1]
24:                $(\hat{l}, \hat{\mathbf{x}}, \mathcal{I}_j) \in \tilde{\mathcal{L}} \times \tilde{\mathcal{X}} \times \tilde{\mathcal{I}}$ 
25:              $\tilde{l}^* = \min \tilde{\mathcal{L}}$ , while  $\mathbf{x}^*$  and  $\tilde{\mathcal{I}}^*$  are as associated
26:           return  $\tilde{l}^*, \mathbf{x}^*$ , and  $\tilde{\mathcal{I}}^*$ 

```

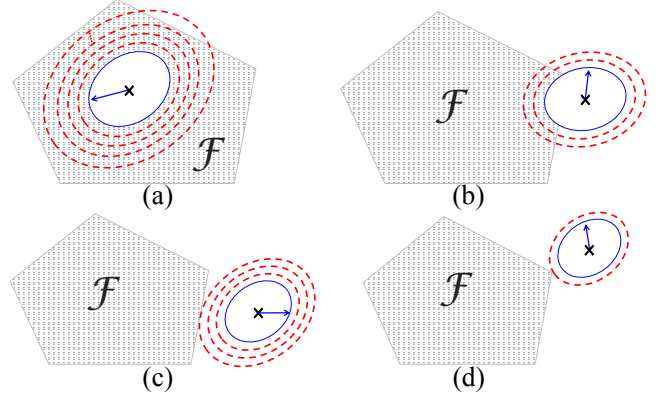


Fig. 3. Representative cases in terms of the location of \mathbf{x}^* [13] for Algorithm 1: (a) $\mathbf{x}_p^* \in \mathcal{F}$, (b) $\mathbf{x}_p^* \notin \mathcal{F}$ and $\mathbf{x}^* \in \mathcal{F}$ for some $\varepsilon^* \neq 0$, (c) $\mathbf{x}^* \notin \mathcal{F}$ while \mathbf{x}^* on an edge, and (d) $\mathbf{x}^* \notin \mathcal{F}$ while \mathbf{x}^* on a vertex.

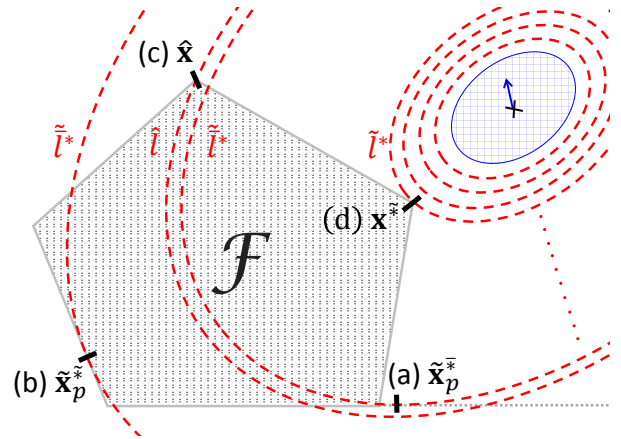


Fig. 4. A planar demonstration of "optimality candidates" in Appendix C.2 when $\mathbf{x}^* \notin \mathcal{F}$ according to Eq. [13], where (a) $\tilde{\mathbf{x}}_p^*$ is not a candidate, both (b) and (c) $\tilde{\mathbf{x}}_p^*$ are candidates on an edge and vertex, respectively, and the optimality occurs at another vertex/candidate (d) \mathbf{x}^* .

339 *Remark 2.6.* Regarding Algorithm 1,

- 340 i) it does not require to know a feasible point *at any time* in
 341 the solving process (in contrast to, for example, *a priori*
 342 to initiate the process according to (5) and (4)) since the
 343 core of the analytical philosophy is a novel categorization
 344 of QP Problem [15] in Appendix C, which particularly
 345 builds on new results in Lemma 2.1.
- 346 ii) in the for-all environment (lines 11-24), the processing
 347 order is arbitrary, as long as every nonempty subset of \mathcal{I}

is examined. Examples are given later in Sec. 3.

iii) an economic/expedited version of this algorithm is available with respect to Case (a) in Fig. 3, as motivated by the particular interest in the computational performance using existing QP solvers (to name a few, (17, 18)). Specifically, by replacing lines 8 and 9 with 6, the complete searching for all optima is economized since the optimal value and an optimum have already been obtained, without further effort for any possible, additional optimum. Accordingly, the algorithm is expedited because it terminates at line 8 without proceeding into the for-all environment (lines 11-24). More details and related discussions can be found in the end of Appendix C.1, as well as Remark C.4.

Corollary 2.1. With regard to Algorithm 1, consider

- A) a special case of the QP Problem [15], $P \succ 0$ as is commonly seen in literature (4). Algorithm 1 can be more efficiently compacted by removing lines 2, 7-9, 10, and 26, as well as “ $V_2^T P V_2 \succ 0$ ” at line 5, and both “ $\tilde{V}_2^T P \tilde{V}_2 \neq 0$ ” and “ $\tilde{V}_2^T (\mathbf{q} + P \tilde{A}^\dagger \tilde{\mathbf{b}}) \in \mathcal{R}(\tilde{V}_2^T P \tilde{V}_2)$ ” at line 17.
- B) an extension to the QP under inequality constraints only, that is, $A = 0$ and $\mathbf{b} = \mathbf{0}$ in the QP Problem [15]. Algorithm 1 applies to this case by replacing “both A, A^\dagger with 0”, “ V_2 with I_n ”, “ \mathbf{x}_p^* with \mathbf{x}_p^* ”, “ \mathbf{b} with $\mathbf{0}$ ”, and “ l^* with l^* ”.

Proof. A): Given $P \succ 0$, in Algorithm 1 we readily have $V_2^T P V_2 \succ 0$, $V_2^T (\mathbf{q} + P A^\dagger \mathbf{b}) \in \mathcal{R}(V_2^T P V_2) = \mathbb{R}^{n-m}$, $\tilde{V}_2^T P \tilde{V}_2 \succ 0$, and $\tilde{V}_2^T (\mathbf{q} + P \tilde{A}^\dagger \tilde{\mathbf{b}}) \in \mathcal{R}(\tilde{V}_2^T P \tilde{V}_2) = \mathbb{R}^{n-m-k}$. The result immediately follows.

B): At first, the concept is to replace the application of Theorem 2.2 by Theorem 2.1 in the beginning of Algorithm 1. More specifically, the processes before the for-all environment (lines 3-9) instead consider the unconstrained QP, and determine whether the associated unique optimum, or the particular solution of nonunique optima, satisfies the (inequality) constraints. Starting from line 11, the remaining processes in Algorithm 1 are slightly modified. The main difference is that, at line 12, the remaining processes are instead built on the “base” of unconstrained QP ($A = 0$ and $\mathbf{b} = \mathbf{0}$) and then taking each inequality constraint into consideration. All in all, from this novel perspective and design, this extension can be easily established by the simple parameter replacements. \square

Remark 2.7. With respect to Corollary 2.1,

- i) in A) of Corollary 2.1, the more compacted algorithm renders $\mathbf{x}^* = \mathbf{x}_p^*$ and $\tilde{\mathbf{x}}^* = \tilde{\mathbf{x}}_p^*$ for all $\mathcal{I}_j \subseteq \mathcal{I}$. That is, the associated solution freedom vanishes (because of $P \succ 0$). Even though Algorithm 1 is designed to *implicitly* while *exhaustively* include the optimality searching within such a freedom (more detailed arguments at Appendix C), this remark also reminds that the searching process, excessive in this case, automatically vanishes as well – as designed.
- ii) Example 3.1 in Sec. 3 demonstrates both Algorithm 1 and Corollary 2.1.

2.4. Extended QP.

Theorem 2.3. (Solutions to An Extended QP)

Consider the optimization problem,

$$\text{minimize } F_{\mathbf{x}} \text{ subject to both } \mathbf{x}, \mathbf{q} \in \mathcal{R}(P), \quad [16]$$

where $F_{\mathbf{x}}$ is in Eq. [1].

A) The preimage of any level set of the CCQF “ $F_{\mathbf{x}}$ subject to both $\mathbf{x}, \mathbf{q} \in \mathcal{R}(P)$ ” can be respectively parameterized by

$$\mathbf{x} = -P^\dagger \mathbf{q} + \sqrt{\mathbf{q}^T P^\dagger \mathbf{q} - 2s + 2\check{l} \cdot P^{\dagger/2} \cdot \check{\rho}}, \quad [17]$$

where $\check{l} \in \mathbb{R}$ is any level set value of the CCQF, $\check{\rho} \in \mathcal{R}(P)$, and $\|\check{\rho}\| = 1$.

B) The optimal value is finite, and equals $\check{l}^* = s - \mathbf{q}^T P^\dagger \mathbf{q} / 2$. The corresponding unique optimum is $\mathbf{x}^* = -P^\dagger \mathbf{q}$.

Proof. See Appendix D. \square

Corollary 2.2. Consider the specific but popular case, $P \succ 0$ (4). Theorem 2.3 can be specialized by replacing P^\dagger with P^{-1} , $P^{\dagger/2}$ with $P^{-1/2}$, and “ $\check{\rho} \in \mathcal{R}(P)$ ” with “ $\check{\rho} \in \mathbb{R}^n$ ”, respectively.

Proof. Given $P \succ 0$, such a case-specific result readily follows by relating the basic SVD properties in Eqs. (35-39) in (2) to this special consideration. \square

Remark 2.8. An application of Theorem 2.3 is associated with the CQF in Eq. [9], that is, to obtain its preimage \mathbf{w} at any level set/value. To be more detailed, by letting $P = 2M$ which is rank deficient, $\mathbf{q} = \mathbf{k}_M$, and $s = c$ in accordance with the CQF in Eq. [1], the preimage \mathbf{w} can be algebraically while completely parameterized by (using Theorem 2.3)

$$\mathbf{w} = -M^\dagger \mathbf{k}_M / 2 + \sqrt{\mathbf{k}_M^T M^\dagger \mathbf{k}_M / 4 - c + \check{F}_{\mathbf{w}} \cdot M^{\dagger/2} \cdot \check{\rho}}, [18]$$

where $\check{F}_{\mathbf{w}} \in \mathbb{R}$ is any (given) level set value, $\check{\rho} \in \mathcal{R}(M)$, and $\|\check{\rho}\| = 1$. As an example, this application-benefit is also demonstrated (with an illustration) in Sec. 5 in (2).

2.5. Comparisons and Impacts to Management Literature and Beyond. With respect to the QP problem [15], several representative and popular solvers in the literature or readily in the market (13) are selected to demonstrate this section. At first, according to (4), such a general QP (with inequality constraints) is usually, numerically solved using an “Active Set Method” (ASM), which is also valued in the fields of management science (22, 23) and control engineering (17), to name a few. It has been reported that a feasible point is required to initiate the solving process of ASM – including a simpler/common case of $P \succ 0$ (4) – till the latest MATLAB[®] implementation (namely, in the “quadprog” routine). Similar difficulty/inconvenience can also be found in other celebrated solvers: Newton’s method, the barrier method, and primal-dual interior-point methods (5, 17, 22). In contrast, the proposed solvers autonomously generate the results/solutions, notably Algorithm 1 and Theorem 2.2. The analytical

philosophy differentiates from the classical Lagrange/Primal-Dual method (4, 5, 13), but stems from Lemma 2.1 as well as Figs. 1 (“critical shift”) and 2 (“hierarchical layers”).

In addition, another advantage is revealed from an accuracy perspective, with regard to the other representative solver: the barrier method that is renowned for large-scale problems (17). The importance of such an interior-point solver is particularly focused upon in (5, Chapter 11). Accordingly, under reasonable assumptions, a minimum number of iteration steps are generally required for convergence, or a desired accuracy; whilst, the maximum number of steps (and computation time) are deemed challenging (17). As a comparison, the proposed results yield exact solutions for both the optimal point(s) and value. Specifically, considering the equality-constrained QP problem [10], the finite solutions (when existed, according to a comprehensive categorization in Theorem 2.2) are available in Eqs. [12] and [13]. Moreover, regarding the general (resp., extended) QP problem in [15] (resp., [16]), Algorithm 1 (resp., Theorem 2.3) also renders results in analytical, algebraic, and closed forms. Remarkably, it is worth mentioning that all the proposed results refrain from any convergence analysis, namely the estimation of convergence rate, which can cause computational difficulty (16) but is generally necessary till the up-to-date literature.

Furthermore, the proposed exact optimization is beneficial in the light of its derivative-free feature. Such a property subject to special cases has a long history in literature while attracts renewed interest in recent years, for instance, (16) that highlights its necessity in the very beginning. In other words, the feature is a merit since it does not require any gradients. More specifically, it is desirable in settings in which exact gradient/first-order calculations are computationally expensive or impossible. Note that, in such cases, the classical techniques – including Kiefer-Wolfowitz-type procedures (24) – often resort to compromises that are based on approximations, which (observation) further strengthens the above mentioned advantage in accuracy while reminds a special consideration in Remark C.4. Moreover, in the fields of statistics and machine learning (see, for example, Refs. 3-7 in (16)), there has been a growing spectrum of optimization problems that entail functional information only, which also broadens the impacts of this research direction. Finally, it is worth reiterating that the proposed scheme exploits all the information of the function, which (unified framework) includes both the zero-order (16) and first-order values that have been widely leveraged in the classical/iterative optimization (4, 13).

All of the above advantages will be validated in Sec. 3 below, which provides explicit computational evidence. Promisingly, there exist more advantages, such as the one considered in Remark C.4 (Appendix C); and those of practical importance: computational speed, hardware complexity, and safety standard, since this research direction toward an exact solution aligns with, for example, (14) and its mathematical preliminaries (notably (9), which is with respect to a positive definite Hessian matrix). Last but not least, an additional impact to literature suggests the (bi)linear matrix inequalities-based design (to name one recent representative, (25)), which traces its core value/concept to the QP problem and thus motivates further research.

3. Three Illustrative Examples

The main objective is to demonstrate the *autonomy*, *generality*, and *functionality* of the proposed novel *closed-form* QP solvers (Theorem 2.2, Algorithm 1, and so on), particularly from the implementation perspective. More specifically, it is worth reiterating that all the proposed results do not require any knowledge of a feasible point – at any time before and during the solving process (namely, autonomy); deal with the general QP problem in accordance with, for example, (4), as well as an extension that enlarges the application coverage (generality); and exploit all the information of the function that includes the commonly utilized zero-order (16) and first-order functional values in the classical/numerical optimization ((13); functionality). Example 3.1 is directly adopted from (4), and used to demonstrate Theorems 2.1, 2.2, Algorithm 1, Corollary 2.1, Remarks 2.1, 2.3, 2.7, C.3, and Lemma 2.1. One step forward, Example 3.2 is extended and modified from Example 3.1, by adding an equality constraint, rendering the Hessian matrix singular, and increasing the system dimension. Therefore, it also demonstrates those mentioned above, except the special cases of Corollary 2.1 and Remark 2.7, and additionally Remark C.1 for the concept of *terminal optima* and related solution freedom. For compactness, the illustration for Example 3.1 can be either inferred from Fig. 4 or referenced to Fig. 14.1 in ((4), 2nd edn.); whereas, in Example 3.2, a geometric interpretation for Remark C.1 is given in Fig. 5. Last but not least, Example 3.3 leverages a benchmark problem as selected in (26) and compares the proposed approach with the built-in QP solver “quadprog”, which (among others) boasts a wide popularity in the optimization literature and beyond (27). Note that all computations in the three examples are performed and/or verified on the MATLAB® platform, and they are subject to a variety of hardware implementations (details in Table 1 later).

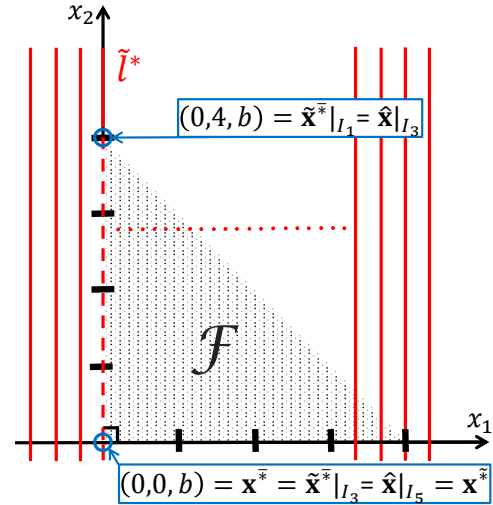


Fig. 5. A geometric interpretation of “terminal optima” in Example 3.2, where $x_3 = b$.

Example 3.1. Consider a QP problem of nonsingular Hessian matrix and under inequality constraints only, where $n = 2$, $P(1,:) = [4, 1]$, $P(2,:) = [1, 2]$, $\mathbf{q} = [-12, -10]^T$, $s = 0$, $\mathbf{c}_1 = [1, 1]^T$, $\mathbf{c}_2 = [-1, 0]^T$, $\mathbf{c}_3 = [0, -1]^T$, $d_1 = 4$, and $d_2 = d_3 = 0$ (4). Given $P \succ 0$, according to Corollary

2.1, iii) of Theorem 2.1, or line 3 of Algorithm 1, we have the unique optimum with respect to the unconstrained problem as $\mathbf{x}^* = -P^{-1}\mathbf{q} = [2, 4]^T$, which does not satisfy the inequality constraint associated with (\mathbf{c}_1, d_1) . That is, this unconstrained optimum does not reside in the feasible set. Therefore, following Algorithm 1, the next step is the for-all environment starting at line 11, where $\mathcal{I} = \{1, 2, 3\}$ ($\kappa = 3$). Note that the significant parameters and values in this step are summarized by Table 2 in Appendix E for completeness and easy comparison. Moreover, the examination order for all \mathcal{I}_j 's is arbitrary, as long as every nonempty subset of \mathcal{I} is examined, which is discussed in ii) of Remark 2.6.

In Table 2 (Appendix E), at first we note that only those \mathcal{I}_j 's with $\tilde{\mathbf{b}} \in \mathcal{R}(\tilde{A})$ in the corresponding augmented system are included, as filtered at line 13 in Algorithm 1; otherwise, it does not associate with the optimality since being inconsistent, and the only case is \mathcal{I}_7 , where, equivalently,

$$\det \begin{pmatrix} \mathbf{c}_1^T & d_1 \\ \mathbf{c}_2^T & d_2 \\ \mathbf{c}_3^T & d_3 \end{pmatrix} \neq 0.$$

Second, regarding \mathcal{I}_1 , the corresponding augmented matrix \tilde{A} is of full rank, and the augmented system $\tilde{A}\mathbf{x} = \tilde{\mathbf{b}}$ is underdetermined and has more than one solution. This falls into the processes at lines 15-19 of Algorithm 1. According to Theorem 2.2, an example of $\tilde{V}_2 = [1/\sqrt{2}, -1/\sqrt{2}]^T$ is by means of SVD, which is computed at line 15 of Algorithm 1. This yields the value " $\tilde{V}_2^T P \tilde{V}_2 = 2 \neq 0$ ", which implies that this case is equivalently an unconstrained QP. As a matter of fact, given $P \succ 0$, the following results also endorse Corollary 2.1:

- $\tilde{V}_2^T(\mathbf{q} + P\tilde{A}^\dagger\tilde{\mathbf{b}}) \in \mathcal{R}(\tilde{V}_2 P \tilde{V}_2) = \mathcal{R}(2) = \mathbb{R}$, which means that the equivalent QP problem is of finite value by Ac) of Theorem 2.2.
- $\mathcal{N}(P) = \{\mathbf{0}\}$, and thus $\mathcal{N}(\tilde{A}) \cap \mathcal{N}(P) = \{\mathbf{0}\}$. According to Ae) of Theorem 2.2 or i) of Remark 2.7, we have that the associated optimum for this case \mathcal{I}_1 is unique and denoted $\tilde{\mathbf{x}}^* = \tilde{\mathbf{x}}_p^* = [1.5, 2.5]^T$. This point is computed at line 17 in Algorithm 1.
- $\tilde{\mathbf{x}}^*$ is an optimality candidate for the original problem, as recognized at line 19, since it satisfies the other (inequality) constraints with respect to $\{(\mathbf{c}_i, d_i)\}_{i=2,3}$. In other words, it is (also) a feasible point with regard to the original problem/constraints.
- The associated optimal value ($\tilde{l}^* = -28.5$) for this case \mathcal{I}_1 is determined by Eq. [12], computed at line 18, and included at line 24 for the overall comparison later at line 25 of Algorithm 1.

Similar considerations apply to Cases $\{\mathcal{I}_j\}_{j=2,3}$, but note that the associated unique optimum for Case \mathcal{I}_2 ($\tilde{\mathbf{x}}^* = [0, 5]^T$) is not a candidate, because it violates the first inequality constraint " $x_1 + x_2 \leq 4$ ".

Third, Case \mathcal{I}_4 corresponds to a "vertex", which is zero-dimensional, as a result of $\text{rank}(\tilde{A}) = n = 2$. Specifically, according to ii) of Remark 2.3, this case is *instead* included at lines 20-24 of Algorithm 1, and there is only one feasible point $\tilde{\mathbf{x}} = [0, 4]^T$ as computed at line 21. Since $\tilde{\mathbf{x}}$ satisfies all the constraints, it is recognized as a candidate (line 24). The associated, only level set value $\hat{l} = -24$ is exclusively optimal

in this case, computed at line 23, and thus included for the comparison at line 25. Similar considerations apply to Cases $\{\mathcal{I}_j\}_{j=5,6}$, but omitted for conciseness.

Finally, all the values, either \tilde{l}^* or \hat{l} , are compared at line 25. This concludes that Case \mathcal{I}_1 corresponds to the optimality. To be more detailed, $\mathbf{x}^* = \tilde{\mathbf{x}}^* = [1.5, 2.5]^T$ for \mathcal{I}_1 is also the unique optimum for the original problem, with the optimal value $\tilde{l}^* = \hat{l}^* = -28.5$, and \mathbf{x}^* is located on the *edge* of \mathcal{F} that is associated with \mathcal{I}_1 , explicitly, $x_1 + x_2 = 4$.

Remark 3.1. The results agree with (4) that uses ASM, but without any prior knowledge of a feasible point to initiate the solving process. Besides, all computations are performed in terms of analytical, algebraic, closed-form representations. Moreover, there are only seven cases considered in the for-all environment of Algorithm 1 (lines 11-24). In consistency with Remark C.3, the cardinality of $\tilde{\mathcal{L}}$, which equals 5, is less than the upper bound " $2^\kappa = 8$ ".

Example 3.2. This example is extended and modified from Example 3.1. The main difference in the motivation is additionally to demonstrate the concept of *terminal optima*, as explained in Appendix C.2, and a potential to assist the analytical and comprehensive representation of all the nonunique optima. According to the problem formulation in QP [15], we let $n = 3$, singular $P = \text{diag}(1, 0, 0) \succeq 0$, $\mathbf{q} = \mathbf{0}$, $A = [0 \ 0 \ 1]$, $\mathbf{b} = b \in \mathbb{R}$, while the other parameters follow Example 3.1. For succinctness, only the significantly different analyses and results are presented below.

First (lines 1-9 in Algorithm 1), consider the QP under only the equality constraint ($A\mathbf{x} = b$). At line 1, V_2 is computed as $V_2(:, 1) = [0 \ 1 \ 0]^T$ and $V_2(:, 2) = [1 \ 0 \ 0]^T$. The associated matrix $V_2^T P V_2 = \text{diag}(0, 1) \neq 0$, which implies that this is equivalently an unconstrained QP by A) of Theorem 2.2. In addition, given the singular $V_2^T P V_2 \succeq 0$ as well as $V_2^T(\mathbf{q} + P A^\dagger \mathbf{b}) = \mathbf{0} \in \mathcal{R}(V_2^T P V_2)$, by Ac) of Theorem 2.2, the optimal value (that is, minimum) is finite. Moreover, since $\mathcal{N}(A) \cap \mathcal{N}(P) = \mathcal{R}([0 \ 1 \ 0]^T) \neq \{\mathbf{0}\}$, applying Ae) of Theorem 2.2 yields that the optima are nonunique. Preliminarily at this stage, Algorithm 1 only examines the particular solution of the optima (\mathbf{x}_p^* , as computed at line 3), while integrates the examination on its solution freedom ($V_2 \varepsilon^*$) into the for-all environment later (lines 11-24). The result is $\mathbf{x}_p^* = [0, 0, b]^T$, which resides in the feasible set \mathcal{F} of the original problem (grey/shaded area in Fig. 5). Hence it qualifies as an optimality candidate (included at line 8), associated with $\tilde{l}^* = 0$ (computed at line 4).

Second, the process goes to the for-all environment in Algorithm 1. Following the analyses in Example 3.1, the results are briefed and summarized by Table 3 in Appendix E for better presentation clarity. Note that:

- Considering \mathcal{I}_1 , this corresponds to the *terminal optimum* associated with the above mentioned solution freedom of \mathbf{x}^* (namely, $V_2 \varepsilon^*$) as a result of the additional constraint with respect to (\mathbf{c}_1, d_1) . This can also be inferred from Fig. 5, where the red/solid/vertical lines represent the level sets/values of the involved CQF; while the x_2 -axis is just the linear variety \mathbf{x}^* , and the terminal optima $(0, 4, b)$ is owing to the additional constraint (\mathbf{c}_1, d_1) . This demonstration agrees with the discussions in Appendix C.1.

Regarding \mathcal{I}_2 , this is equivalently an optimization problem but of constant value, according to B) and C) of Theorem 2.2. Specifically, it is owing to $\tilde{V}_2^T P \tilde{V}_2 = 0$, $\tilde{V}_2^T (\mathbf{q} + P \tilde{A}^\dagger \tilde{\mathbf{b}}) = 0$, and the constant value equals $s + [\mathbf{q}^T + \tilde{\mathbf{b}}^T (\tilde{A}^\dagger)^T P / 2] \tilde{A}^\dagger \tilde{\mathbf{b}} = 0$. Additionally, as addressed in iib) of Appendix C.2, the algorithm is designed to economize this examination by focusing on the associated terminal optima. Specifically, it is instead included in Cases $\{\mathcal{I}_j\}_{j=4,5}$, and associated with additional constraints $\{(\mathbf{c}_i, d_i)\}_{i=1,3}$, respectively. Notably, the associated optimal value remains. Finally, a geometric interpretation is also available in Fig. 5, where, on the x_2 -axis (that is, the border of $\mathbf{c}_2^T \mathbf{x} \leq d_2$), the equivalent problem/function bears the constant value. In its domain, the upper (resp., lower) terminal is due to the additional constraint “ (\mathbf{c}_1, d_1) ” (resp., “ (\mathbf{c}_3, d_3) ”), and the discussion/categorization goes to the case \mathcal{I}_4 (resp., \mathcal{I}_5).

Noting the additional column of $\mathcal{N}(P) \cap \mathcal{N}(\tilde{A})$ in Table 3, which is omitted in Table 2 (Appendix E) because of the more convenient condition “ $P \succ 0$ ” instead/previously in Example 3.1, the case \mathcal{I}_1 (resp., \mathcal{I}_3) renders this intersection only at the origin, and thus the corresponding optimum is unique, $\tilde{\mathbf{x}}_p^* = \tilde{\mathbf{x}}^*$, according to Ae) of Theorem 2.2. This can also be inferred from Fig. 5, where there is only one intersection between the optimal (level set) value, denoted \tilde{l}^* , and the border of $\mathbf{c}_1^T \mathbf{x} \leq d_1$ (resp., $\mathbf{c}_3^T \mathbf{x} \leq d_3$).

Finally, corresponding to the optimal value \tilde{l}^* , at line 25 in this algorithm, it concludes with the five optimality pairs/points residing in $\tilde{\mathcal{L}} \times \tilde{\mathcal{X}} \times \tilde{\mathcal{I}}$: $([0, 0, b]^T, 0, \emptyset)$, $([0, 4, b]^T, 0, \mathcal{I}_1)$, $([0, 0, b]^T, 0, \mathcal{I}_3)$, $([0, 4, b]^T, 0, \mathcal{I}_4)$, and $([0, 0, b]^T, 0, \mathcal{I}_5)$. That is, Algorithm 1 gives the optimal value, $\tilde{l}^* = 0$, with respect to the original problem. All the *terminal optima* are $[0, 0, b]^T$ and $[0, 4, b]^T$, as illustrated in Fig. 5. Notably, the former i) is the particular solution with regard to the QP under only the equality constraint, and also feasible to the original QP; and ii) resides on an edge associated with the inequality constraint “ $\mathbf{c}_3^T \mathbf{x} \leq d_3$ ”, while also on a vertex with the additional constraint “ $\mathbf{c}_2^T \mathbf{x} \leq d_2$ ”. On the other hand, the latter terminal optimum resides on an edge associated with the constraint “ $\mathbf{c}_1^T \mathbf{x} \leq d_1$ ”, while also on a vertex with the additional “ $\mathbf{c}_2^T \mathbf{x} \leq d_2$ ”. It is worth remarking that Algorithm 1 ably gives all/both the *terminal optima*. Based on this, it can be analyzed that all the *optima* are just *intermediate*, denoted as $\mathcal{X}^* = \{[0, t, b]^T \mid 0 \leq t \leq 4, t \in \mathbb{R}\}$, which can also be revealed from Fig. 5 (specifically, the red-dashed line segment).

Remark 3.2. As expected and similar to Example 3.1 (Remark 3.1), Algorithm 1 autonomously yields the results, with only functional information (derivative-free) and without any (prior) knowledge of a feasible (starting) point. Also, in this setting, all computations are performed in terms of algebraic and explicit representations. Moreover, there are still only few cases (seven, to be exact) considered in the for-all environment of Algorithm 1 and, being consistent with Remark C.3, the cardinality of $\tilde{\mathcal{L}}$ (6) is indeed less than the upper bound “ $2^\kappa = 8$ ”.

Example 3.3. The extensive QP literature highlights a universal and popular solver “quadprog” in MATLAB® (27), and this example leverages a benchmark problem as demonstrated in (26) to offer more computational evidences of

the superiority of the proposed approach. Specifically, let $P(1, :) = [1, -1, 1]$, $P(2, :) = [-1, 2, -2]$, $P(3, :) = [1, -2, 4]$, $\mathbf{q} = [-7, -12, -15]^T$, $s = 0$, $A = [1, 1, 1]$, and $b = 3$ in Problem/Eq. [10], and this Example 3.3 compares Theorem 2.2 with “quadprog”. Among various implementations for Theorem 2.2, we demonstrate an intuitive and preliminary one using MATLAB® for the sake of brevity, whose primary steps are sequentially summarized:

```

1 A_dagger=pinv(A);
2 V_2=null(A);
3 x_star_particular=A_dagger*b-V_2*pinv(...
4   V_2'*P*V_2)*(V_2'*q+V_2'*P*A_dagger*b);
5 l_star_equilibrium=(b'*A_dagger'*P/2...
6   +q')*(eye(n)-V_2*pinv(V_2'*P*V_2)...
7   *V_2'*P)*A_dagger*b+s...
8   -q'*V_2*pinv(V_2'*P*V_2)*V_2'*q/2; (i)

```

Note that lines 3 and 4 in Implementation (i) correspond to Eq. [13] while 5 and 6 to [12]. In addition, this example falls into Case A) of Theorem 2.2 because of $\mathcal{R}(P) \cap \mathcal{N}(A) = \mathcal{N}(A) \neq \{\mathbf{0}\}$, which can be easily verified – using $V_2^T P V_2 \neq 0$ – and thus omitted in the summary of primary steps (i).

On the MATLAB® platform, the proposed solver yields the optimal value $\tilde{l}^* = -47.1786$ and the optimum $\mathbf{x}^* = [-3.5714, 2.9286, 3.6429]^T$, where the uniqueness of \tilde{l}^* (resp., \mathbf{x}^*) is owing to the nonsingular $V_2^T P V_2$, namely the nulled flexibility $\epsilon^* = \mathbf{0}$ in Eq. [13]; whereas, the built-in “quadprog” routine renders the same \tilde{l}^* and \mathbf{x}^* but *not* the associated uniqueness, respectively (26). Moreover, the performance in computation time also quite differ, which is evaluated – for better comprehensiveness – subject to a variety of software and hardware in Table 1. Specifically, Table 1 also more favors Theorem 2.2 than “quadprog”, where the evaluation is in terms of the ratio of computation time (Theorem 2.2 over “quadprog”) while among several Intel Core® CPUs and MATLAB® versions till the state of the art. Accordingly, the maximum (resp., minimum) advantage amounts to 2.3% (resp., 9.6%) of the computation time using “quadprog”; in other words, the proposed solver reduces the computational effort by 90.4%–97.7%. On average, the numerics still significantly account for a computational reduction of 94.5% by virtue of Theorem 2.2. It is worth reiterating that the Implementation (i) is only preliminary, whose efficiency potentials can be more exploited by, for instance, “coalescing” lines 1 and 2, namely, by performing the SVD operation on A only once. This requires modifications for the “svd” routine in MATLAB®, and thus serves as a future research to remain focused of this article.

Table 1. Advantage in computation time^a in Example 3.3

	i3, 2.75GHz	i5, 3.4GHz	i7, 2.2GHz
MATLAB® 2010b	2.5%	2.4%	2.3%
MATLAB® 2020b	8.9%	7.4%	9.6%

^a Ratio of time using Theorem 2.2 over “quadprog” (8GB RAM).

Remark 3.3. Regarding the up-to-date universal solver “quadprog”, (26) highlights a *necessary* condition “ $P \succ 0$ ” that associates with a finite optimal value, even if the QP problem

is without any constraint (Sec. 2.1). Actually, there exists a counterexample to the condition/necessity: $P = \text{diag}(1, 0)$, $\mathbf{q} = \mathbf{0}$, and $s = 0$ in Eq. [1]. That is, the QP problem is to minimize $F_{\mathbf{x}}(\mathbf{x}) = x_1^2$, which has a/the finite minimum (zero) whereas the Hessian matrix is not positive definite. As a comparison, ii) of Theorem 2.1 provides a *necessary and sufficient* condition, which readily guarantees *the* finite minimum in the above counterexample – owing to $\mathbf{q} = \mathbf{0} \in \mathcal{R}(P)$. Remarkably, Theorem 2.1 fundamentally helps develop the remaining results (with constraints) in Sec. 2, such as another *equivalent* condition for *the* finite optimal value in Ac) of Theorem 2.2.

4. Conclusion

A motivation of this application – to the field of nonlinear programming based on the recent observations of CQE – responds to an open question in (16); and an expectation in the general optimization literature: computational enhancement for practical advances (4). The philosophy aims at analytical and exact optimization, as a comparison to the common numerical algorithms (13). Specifically, the focus of this article is on the QP, which is a constituent in various approaches such as the Newton’s method. At first, we re-examine the unconstrained QP based on, and from the perspective of, the new analyses of CQE (in the recent literature), CQF, and CQE-CQF relation. This preliminary finding facilitates a complete and analytical characterization of the equality-constrained QP, which actually can be categorized into three equivalent problems in a unified framework. In addition, all the above results are consistent with the literature, when specialized to specific considerations; while promising in terms of the computational performance, which is partially owing to the shared motivation and beneficial results in the recent contribution (2). Another highlight in this presentation is the proposed QP solver, in accordance with the general problem formulation. The analytical philosophy also benefits from the above mentioned categorization of equality-constrained QP, and yields algebraic zero-order closed-form results/solutions in a guaranteed, pre-determined, finite, and explicit number of steps; and without any knowledge of a feasible point, a priori and any time in the process. All these characteristics/merits are shared throughout this article. Moreover, to further exploit its computational capabilities, this flexible QP solver includes problem-specific (and more efficient) variants below:

- an economic/expedited version for a subset of QP problems;
- a special but popular case “the objective function whose Hessian matrix is positive definite”, which corresponds to a thinned version of the algorithm;
- QP problems subject to inequality constraints only, being applicable (that is, solvable) after a simple replacement of parameters in the algorithm.

Note that, regarding a subset of QP problems, we preliminarily and explicitly present the associated terminal optimum/optima. This indicates a further research direction, because it requires an analytical representation of all the elements bounded by the inequality constraints. Moreover, we extend the general formulation on the constraints to a different branch, as inspired by the analyses of CQE and CQF, which further broadens the spectrum of QP applications. Last but not least, two representative solvers in literature (the ASM

and barrier method) help justify advantages/potentials of the proposed results; while three examples demonstrate the results with illustrations, computational evidences, and further insights (notably, a counterexample to the state of the art) – analytically as well as numerically on the MATLAB® platform. Under positive definite and semidefinite Hessian matrices in the objective functions, respectively, these demonstrations endorse the effectiveness, autonomy, efficiency, functionality, and exactness of the proposed solvers.

A. Proof of Theorem 2.1 (Unconstrained QP)

For brevity, only the results that are significantly required in this proof are given, whereas most of the shared similarities are referenced to Lemma 2.1 and its proof.

i) To obtain the preimage of any level set, $F_{\mathbf{x}} = l$, is equivalent to that of the zero level set of $\tilde{F}_{\mathbf{x}}$, where the CQF $\tilde{F}_{\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\tilde{F}_{\mathbf{x}} = F_{\mathbf{x}} - l$, as “(vertically) shifted” from $F_{\mathbf{x}}$ by l . Furthermore, this is equivalent to solving the CQE, $\mathbf{x}^T P \mathbf{x} + \mathbf{q}^T \mathbf{x} + s - l = 0$, which is always solvable since the original process is to find the preimage. The result directly follows by applying Lemma 2.1 to this CQE. Note that the above mentioned “shift” is widely used throughout Sec. 2, and thus exemplified/illustrated in Fig. 1 for presentation clarity.

ii) Divide the derivations, following Lemma 2.1, into whether P is of full rank. A) If $\text{rank}(P) = n$, then we reformulate the CQF $F_{\mathbf{x}}$ in Eq. [1] equivalently as

$$F_{\mathbf{x}}(\mathbf{x}) = \|\mathbf{P}^{1/2} \mathbf{x} + \mathbf{P}^{-1/2} \mathbf{q}\|^2 / 2 + s - \mathbf{q}^T \mathbf{P}^{-1} \mathbf{q} / 2, \quad [19]$$

which is similar to Eq. (34) in (2). Obviously, this function [19] is bounded below by the value “ $(s - \mathbf{q}^T \mathbf{P}^{-1} \mathbf{q} / 2)$ ”, which (value) is finite and important in the derivations afterward. Note that, in this case, $\mathbf{q} \in \mathcal{R}(P) = \mathbb{R}^n$. On the other hand, B) if $\text{rank}(P) = r < n$, then reformulate $F_{\mathbf{x}}$ [1] equivalently as

$$F_{\mathbf{x}}(\mathbf{x}) = \mathbf{x}_1^T \check{\Sigma}_1 \mathbf{x}_1 + \mathbf{q}_1^T \mathbf{x}_1 + s + \mathbf{q}_2^T \mathbf{x}_2, \quad [20]$$

which is similar to Eq. (42) in (2). Obviously, if $\mathbf{q}_2 = \mathbf{0}$, that is, $\mathbf{q} \in \mathcal{R}(P)$, then the domain of $F_{\mathbf{x}}$ *shrinks* and corresponds to the \mathbf{x}_1 -freedom in \mathbf{x} . Similar to A) above, $F_{\mathbf{x}}$, or equivalently the CQF $F_{\mathbf{x}_1} : \mathbb{R}^r \rightarrow \mathbb{R}$,

$$F_{\mathbf{x}_1}(\mathbf{x}_1) = \mathbf{x}_1^T \check{\Sigma}_1 \mathbf{x}_1 + \mathbf{q}_1^T \mathbf{x}_1 + s, \quad [21]$$

is bounded below by $\check{l} := s - \mathbf{q}_1^T \check{\Sigma}_1 \mathbf{q}_1 / 2 = s - \mathbf{q}^T P^\dagger \mathbf{q} / 2 \in \mathbb{R}$. However, if $\mathbf{q}_2 \neq \mathbf{0} \Leftrightarrow \mathbf{q} \notin \mathcal{R}(P)$, then $F_{\mathbf{x}}$ in Eq. [20] is unbounded because of the \mathbf{x}_2 -freedom, which is decoupled from \mathbf{x} .

iii) Given $\mathbf{q} \in \mathcal{R}(P)$, the derivations differ only in the singularity of P according to Conditions/Eqs. [3] and [5], which are divided into Cases A) and B) below.

A) If $\text{rank}(P) = n$, by i) above and Eq. [4], the unique optimum happens exclusively when the only solution freedom (\mathbf{v}) is canceled. In other words, the term in the square-root operator (“ $-l^*$ ” as in this theorem) vanishes. Geometrically from Fig. 1, this can also be explained by the *critical shift* of l^* , which allows the only one-point intersection with the zero plane/line. It is worth mentioning that this shift can be “ascending” (resp., “descending”), if $l^* < 0$ (resp., $l^* > 0$), that is, the critical “distance” to *null* the square-root operation

(resp., make it *consistent*). For succinctness, another explanation using Eq. [19] is omitted. To summarize, the preimage of the level set at the optimal value l^* is the singleton $\{\mathbf{x}^*\}$.

B) If $\text{rank}(P) = r < n$, we at first introduce the concept of *hierarchical layers* in the preimage of CQF in Eq. [1], according to i) above and Eq. [7]. There are three layers: I) the unique point, $\mathbf{x}_p^* = -P^\dagger \mathbf{q}$, II) the ε -freedom in $\mathcal{N}(P)$, and III) the ρ -freedom in $\mathcal{R}(P)$. This I-III) order follows the size of included preimage elements, and I) being the smallest. Geometrically, the I) and II) layers correspond to the optima, where II) mainly reflects the singularity of P . Then, introducing III) equips with the full freedom, which thus corresponds to the complete preimage. For brevity, other explanations by means of, for instance, a) the geometrical perspective, which interprets the critical shift of optimal value l^* similarly as in A); b) Eq. [21], which can be reformulated similar to [19]; and c) the support from Theorem 3.2 (and its proof) in (2) are omitted.

Remark A.1. As an example, Fig. 2 illustrates the concept of “hierarchical layers”. Regarding the singular P , the three layers are represented by the I) black/crossed point: this particular solution is denoted by $\mathbf{x}_p^* = -P^{-1} \mathbf{q}$ at the center of the figure, II) blue/gridded closed region, and III) red/dashed region, respectively. All the layers are within the preimage of $F_{\mathbf{x}}$, denoted by $F_{\mathbf{x}}^{-1} := F_{\mathbf{x}}^{-1}(l) : \mathbb{R} \rightarrow \mathbb{R}^n$ with respect to any level set value l . I) and II) layers are in the same level set at (the optimal value) l^* , while each red/dashed ellipse in the III) layer corresponds to a level set at a higher value. Moreover, II) corresponds to $\mathcal{N}(P)$ and each element/point is referenced by the vector $\check{\epsilon}$; while III) to a point $\Phi \cdot \check{\rho} \in \mathcal{R}(P)$, where $\Phi := \Phi(l) = \sqrt{2l + \mathbf{q}^T P^\dagger \mathbf{q} - 2s} \cdot P^{\dagger/2} \in \mathbb{R}^{n \times n}$ while $\check{\rho} \in \mathcal{R}(P) \subseteq \mathbb{R}^n$ is of unit length. Note that $\mathcal{R}(\Phi) = \mathcal{R}(P^{\dagger/2}) = \mathcal{R}(P)$ and $\|\Phi\| = \sqrt{(2l + \mathbf{q}^T P^\dagger \mathbf{q} - 2s)/\sigma_r}$ (21), where $\sigma_r > 0$ denotes the smallest (positive) singular value of P . On the other hand, considering the case of invertible P , the only differences are

- the lack of II) layer, because of $\varepsilon \in \mathcal{N}(P) = \{\mathbf{0}\}$ and $P^\dagger = P^{-1}$;
- $\Phi = \sqrt{2l + \mathbf{q}^T P^{-1} \mathbf{q} - 2s} \cdot P^{-1/2}$ and $\|\Phi\| = \sqrt{(2l + \mathbf{q}^T P^{-1} \mathbf{q} - 2s)/\sigma_n}$, where σ_n denotes the smallest (positive) singular value or, equivalently here, eigenvalue of P .

Finally, this concept of “hierarchical layers” is essential for the subsequent derivations and results in Sec. 2, particularly the constrained optimization as analytically solved from a novel perspective (Secs. 2.2 and 2.3).

B. Proof of Theorem 2.2 (Equality Constraints)

Let the SVD of A be given by

$$A = W \cdot \begin{bmatrix} \Gamma_1 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

with its unique pseudoinverse $A^\dagger \in \mathbb{R}^{n \times m}$ (21),

$$\begin{aligned} A^\dagger &= [V_1 \ V_2] \begin{bmatrix} \Gamma_1^{-1} \\ 0 \end{bmatrix} \cdot W^T \\ &= V_1 \Gamma_1^{-1} W^T \text{ (the thin version),} \end{aligned} \quad [22]$$

where $W \in \mathbb{R}^{m \times m}$, $V_1 \in \mathbb{R}^{n \times m}$, $V_2 \in \mathbb{R}^{n \times (n-m)}$, $\mathcal{R}(V_1) = \mathcal{R}(A^T)$, $\mathcal{R}(V_2) = \mathcal{N}(A)$, $\Gamma_1 \in \mathbb{R}^{m \times m}$ is nonsingular, and more properties are widely available in literature (21). The idea is to equivalently transform the equality-constrained QP into an unconstrained optimization problem, by down-casting to the linear variety/constraint as described by $A\mathbf{x} = \mathbf{b}$. Specifically, the first step is to parameterize all the feasible points, namely $A\mathbf{x} = \mathbf{b} \Leftrightarrow$ Eq. [11], where $\mathbf{y} \in \mathbb{R}^{n-m}$. Then, reformulate the considered constrained optimization problem, “minimize $F_{\mathbf{x}}$ [1] subject to $A\mathbf{x} = \mathbf{b}$ ”, equivalently as the following one: minimize $F_{\mathbf{y}}$ with respect to \mathbf{y} , where $F_{\mathbf{y}} : \mathbb{R}^{n-m} \rightarrow \mathbb{R}$,

$$\begin{aligned} F_{\mathbf{y}}(\mathbf{y}) &= \mathbf{y}^T V_2^T P V_2 \mathbf{y} / 2 + [\mathbf{b}^T (A^\dagger)^T P + \mathbf{q}^T] V_2 \mathbf{y} \\ &\quad + s + [\mathbf{q}^T + \mathbf{b}^T (A^\dagger)^T P / 2] A^\dagger \mathbf{b}. \end{aligned} \quad [23]$$

Therefore, the QP Problem [10] is equivalently categorized into the three cases A)-C), according to the coefficients of quadratic, linear, and constant terms in [23], respectively. The system dimension is expectedly reduced, from \mathbb{R}^n to \mathbb{R}^{n-m} . The following derivations focus on the case of (equivalently unconstrained) QP, while it is straightforward to derive the other cases and thus omitted for compactness.

To start with, we further analyze the equivalence conditions such that $F_{\mathbf{y}}$ in Eq. [23] is quadratic, intentionally/preferably in terms of the original, given parameters. Let the SVD of (non)singular $P \succeq 0$ be

$$P = \check{U}_1 \check{\Sigma}_1 \check{U}_1^T \text{ (full/thin version),} \quad [24]$$

where $\check{U}_1 \in \mathbb{R}^{n \times \check{r}}$ is with orthonormal column(s), $\check{r} := \text{rank}(P)$, and $\check{\Sigma}_1 \in \mathbb{R}^{\check{r} \times \check{r}}$ is nonsingular. With respect to the case of singular P , more properties and details regarding this SVD similarly follow Eqs. (35-37) in (2), which can be easily extended to the nonsingular/other case. It is worth emphasizing that the following derivations *coalesce* both cases of P . On the one hand, this is a QP, namely, $F_{\mathbf{y}}$ in Eq. [23] is a CQF, iff $V_2^T P V_2 \neq 0 \Leftrightarrow V_2^T \check{U}_1 \check{\Sigma}_1 \check{U}_1^T V_2 \neq 0 \Leftrightarrow \check{U}_1^T V_2 \neq 0 \Leftrightarrow \mathcal{R}(P) \cap \mathcal{N}(A) \neq \{\mathbf{0}\}$. Similarly on the other, this is *not* a QP, iff $\check{U}_1^T V_2 = 0 \Leftrightarrow \mathcal{N}(A) \subseteq (\mathcal{R}(P))^\perp = \mathcal{N}(P)$, where the equality holds when $\mathcal{R}(P) \oplus \mathcal{N}(A) = \mathbb{R}^n$.

Notably, to conform to the applicable form using Theorem 2.1, any level set of $F_{\mathbf{y}}$ [23] is regarded as a CQE [2] in the unknown variable $\mathbf{z} = \mathbf{y}$, with the parameters given in Ab) of this theorem. Applying Theorem 2.1 yields the results in A), except Ae), while omitting the lengthy but somewhat straightforward calculations. Regarding the exceptional/last one Ae), based on the above derivations, the optimum is unique, iff the Hessian matrix of the CQF $F_{\mathbf{y}}$ in Eq. [23] is positive definite. Rewrite this equivalence condition as $V_2^T P V_2 = V_2^T \check{U}_1 \check{\Sigma}_1 \check{U}_1^T V_2 \succ 0$, according to the thin version of Eq. [24]. Note that the full version corresponds to the case of nonsingular P , and the result similarly follows. Let any $\tilde{\mathbf{y}} \in \mathbb{R}^{n-m} \setminus \{\mathbf{0}\}$ be given, this condition is further equivalent to

$$\begin{aligned} &\|\check{\Sigma}_1^{1/2} \check{U}_1^T V_2 \tilde{\mathbf{y}}\|^2 > 0 \\ &\Leftrightarrow \check{U}_1^T \cdot V_2 \tilde{\mathbf{y}} \neq \mathbf{0} \\ &\Leftrightarrow \mathcal{N}(P) \cap \mathcal{N}(A) = \{\mathbf{0}\}, \end{aligned}$$

where $\check{\Sigma}_1^{1/2} := \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \dots, \sqrt{\sigma_{\check{r}}}) \succ 0$ and $\{\sigma_i\}_{i=1}^{\check{r}}$ denotes the set of nonzero singular values or, equivalently in this case, eigenvalues of P .

Remark B.1. Regarding the SVD of A , the columns of V_2 are not unique. However, the effect by this nonuniqueness is completely coalesced by virtue of Eq. [11], in the very beginning of this proof. In other words, the results in Theorem 2.2 are independent of the nonunique choices of V_2 . From the implementation perspective, in the literature there exist many and various algorithms to construct an example, such as the Golub-Reinsch SVD till the second step (21). Note that, in the MATLAB[®] platform, the “null” function ably gives an example of $V_2 \in \mathcal{N}(A)$, which is implemented by computing the SVD of A .

Remark B.2. For computational purposes, we select the SVD process to obtain a basis of $\mathcal{N}(A)$, that is, the right singular vectors of A that are associated with the zero singular value. As expected, the derivations only utilize the property of $\mathcal{R}(V_2) = \mathcal{N}(A)$ where $V_2 \in \mathbb{R}^{n \times (n-m)}$, and can be easily extended to the general case: any matrix in $\mathbb{R}^{n \times (n-m)}$ with its range spanning $\mathcal{N}(A)$. However, owing to many computational benefits (21) while noting the largely used parameter A^\dagger (and P^\dagger) in Sec. 2, the SVD process is more reasonably preferred to construct an example. Besides, regarding the extension in Sec. 2.3, the proposed algorithm is based on this theorem, and thus shares such a consideration.

C. Proof of Algorithm 1 (Exact QP Solver)

Outline: The proof builds on the concept “hierarchical layers” as illustrated in Fig. 2, which clarifies the geometric structure of the solution to CQE [2] that is associated with the optimality/minimality of CQF [1]; and case-by-case arguments in accordance with Fig. 3, which exactly extract the optimality for most of the cases, except a subset that requires further investigation (parameterization of inequality-constrained feasible points) to complementarily locate all the optima. Moreover, the derivations largely utilize Theorem 2.2, which are divided into: Appendix C.1 corresponding to lines 2-9 in Algorithm 1; and C.2 to lines 10-25, which augments/coalesces the inequality constraints with respect to each nonempty subset (line 11). Note that the design of categorizations examines every “optimality candidate” (a demonstration in Fig. 4) for the ultimate comparison at line 25, and benefits efficient extensions to QP variants such as those in Corollary 2.1.

To start with, Fig. 3 illustrates the main categorization for the derivations in this proof. With regard to the QP Problem [15] under only equality constraints, (a) considers the case that the unique optimum, or the particular solution \mathbf{x}_p^* associated with the optima, resides in the feasible set \mathcal{F} (shaded/grey area, a polyhedron according to (4) and (5)). If the considered point in (a) is outside of \mathcal{F} , but the degree of freedom in [13] associated with the case of nonunique optima ($V_2 \varepsilon^*$) intersects \mathcal{F} , then this is included in (b). On the other hand ($\mathbf{x}^* \notin \mathcal{F}$), the derivations go to either (c) or (d), where (c) addresses the case that the optimality occurs on an “edge/space of nonzero dimension” while (d) at a “vertex/point”. Note that $\mathcal{F} \neq \emptyset$, according to the QP problem formulation in [15]. Moreover, denote $\tilde{\mathcal{I}}^* \subseteq \mathcal{I}$ as the *optimality subset of involved inequality constraints*, which associates with the additional equality constraint(s) that is/are the *border(s)* of involved inequality constraint(s). This subset indicates the “edge(s)” or “vertex/vertices” that the optimum/optima reside(s). A special case is $\tilde{\mathcal{I}}^* = \emptyset$, which corresponds to the case when QP Problems [10] and [15] yield the identical solution. Finally, the first line (fundamental step) of this algorithm is to com-

pute an orthonormal basis of $\mathcal{N}(A)$, denoted V_2 , preferably using SVD due to its many computational advantages (21).

C.1. Case (a) $\mathbf{x}_p^* \in \mathcal{F}$. According to Theorem 2.2, this case corresponds to A) where $V_2^T P V_2 \neq 0$, as filtered at line 2. Note that the case of $V_2^T P V_2 = 0$ corresponds to either an LP or constant function. If, due to additional inequality constraints, the optimality of QP Problem [15] is associated with the former (resp., latter), then the optimality occurs (resp., also) at the “terminal optimum/optima” that will be detailed later in Appendix C.2 while determined/computed at lines 10-25 of this algorithm.

By Ae) of Theorem 2.2, the optimum of equality-constrained QP is unique, iff $\mathcal{N}(A) \cap \mathcal{N}(P) = \{\mathbf{0}\}$, or equivalently $V_2^T P V_2 \succ 0$ in Eq. [13]. Accordingly, further divide this case into whether $V_2^T P V_2$ is nonsingular.

i) If $V_2^T P V_2 \succ 0$, then the unique optimum $\mathbf{x}^* = \mathbf{x}_p^*$ is given by Eq. [13], where the freedom $\varepsilon^* \in \mathcal{N}(V_2^T P V_2) = \{\mathbf{0}\}$ is nulled, and included in line 3 of this algorithm. The associated optimal value \bar{l}^* is given by Eq. [12], and included in line 4. However, only if this unique optimum resides in the feasible set \mathcal{F} , which is verified at line 5, the algorithm completes solving QP Problem [15], since this optimality (\mathbf{x}^*, \bar{l}^*) under only equality constraints also solves the general QP Problem [15]. It is worth mentioning that the pair of “ $V_2^T P V_2 \succ 0$ and the additional, returned parameter $\tilde{\mathcal{I}}^* = \emptyset$ ” serves as a unique identifier for this case, but this excludes the extreme case of singleton $\mathcal{F} \neq \emptyset$ (details at Remark C.4 later).

ii) If $V_2^T P V_2$ is singular, then at first we need to check whether the optimal value is finite, whose equivalence condition is formulated in Ac) of Theorem 2.2 and included in line 7. If infinite, then the optimality of QP Problem [15] can only possibly occur at the “terminal optimum/optima”, owing to additional inequality constraints, which will be determined later in Appendix C.2. On the other hand, consider the finite optimal value as given by Eq. [12], with the associated optima by [13]. These two have also been computed in lines 3 and 4 for a concise presentation. Note that, at this stage, only the particular solution of Eq. [13] is checked to be feasible or not; while the freedom arising from $V_2 \varepsilon^*$, where $\varepsilon^* \in \mathcal{N}(V_2^T P V_2) \neq \{\mathbf{0}\}$, will be checked later in the for-all environment (at lines 11-24). This argument similarly follows iid) in Appendix C.2 below. It is worth remarking that an expedited alternative is available by economizing the searching of any possible further optimum, as determined by some $\varepsilon^* \neq \{\mathbf{0}\}$. Specifically, replacing lines 8-9 by 6.

C.2. Cases (b-d) $\mathbf{x}_p^* \notin \mathcal{F}$. As illustrated in Figs. 3 (b)-(d), the remaining derivations correspond to lines 10-25. They examine all possible cases that the optimality occurs, corresponding to the boundary as determined by additional inequality constraint(s). It is worth noting that such a boundary can be either “edge” or “vertex”, depending on its dimension. Specifically, at line 11, we examine all the possible, $(2^k - 1)$ ’s subsets/cases $\mathcal{I}_k \subseteq \mathcal{I}$. The feasibility will be checked during the process, and each one is associated with the “augmented” equality-constrained QP problem with respect to $(\tilde{A}, \tilde{\mathbf{b}})$ at line 12, where $\tilde{C} \in \mathbb{R}^{k \times n}$, $\tilde{A} \in \mathbb{R}^{(m+k) \times n}$, and $\tilde{\mathbf{b}} \in \mathbb{R}^{m+k}$.

In the very beginning, line 13 filters out the inconsistent case(s): augmented QP(s) imposed by inconsistent equality constraint(s), or not an “edge/vertex” of the polyhedron \mathcal{F} as shown in Fig. 3. Obviously, the optimality does not occur

in such case(s). The next stage, the if-else setting at lines 14-24, filters in only the cases of full-rank \tilde{A} , according to the formulation in [10]; while the other(s) has/have already been considered before (this case/consideration), corresponding to the reduced, full-rank, equivalent counterpart. Also, the if-else setting distinguishes between augmented systems of many solutions (commonly an underdetermined system) and of the only one. Specifically, lines 14-19 mimic the processes at lines 1-8 but with respect to the “augmented” equality constraints $(\tilde{A}, \tilde{\mathbf{b}})$, where $\tilde{V}_2 \in \mathbb{R}^{n \times (n-m-k)}$ is an orthogonal matrix as computed at line 15; while lines 20-24 consider the full-column-rank \tilde{A} , and the unique feasible point/optimum candidate $\hat{\mathbf{x}}$ (computed at line 21), which is associated with some \mathcal{I}_j and the only level set (corresponding to the optimal value for this augmented system). Note that:

- i) At lines 17 and 22, it is sufficient to check all those inequality constraints that are not “augmented/involved” in $(\tilde{A}, \tilde{\mathbf{b}})$ corresponding to an \mathcal{I}_j . This is beneficial from a computational perspective, because excessive computations are avoided.
- ii) If the conditions at line 17 are all satisfied, then lines 18 and 19 gather the optimality candidate, as well as the involved/associated inequality constraints, in $\tilde{\mathcal{L}} \times \tilde{\mathcal{X}} \times \tilde{\mathcal{I}}$. The last element provides the information on which “edge” the unique optimum (or the particular solution of optima) resides, with respect to the augmented equality-constrained QP. A geometric interpretation of this *optimality candidate* is illustrated in Fig. 4, with more details in Remark C.2. On the other hand (when unsatisfied at line 17), according to Theorem 2.2, the consideration instead goes to one of the followings:
 - a) An unconstrained LP, as in B) of Theorem 2.2, the optimality can be determined in another case, say $\mathcal{I}_j \supset \mathcal{I}_j$, if the optimality is actually associated with the \mathcal{I}_j in this case. In other words, given the unbounded LP, the optimal value only happens by imposing further (inequality) constraints, and will be at the “terminal(s)”. Accordingly, the associated optimum/optima are further termed *terminal*.
 - b) A constant problem, as in C) of Theorem 2.2. Obviously, its optimal value is just the only level set value. However, instead of checking whether the domain intersects \mathcal{F} , a more efficient way is to reserve this case to another one in the for-all environment, since there are only two possibilities: one is that this case will never be further constrained, which obviously does not solve QP Problem [15]; regarding the other, the optimal value still remains and *also* occurs at the “terminal optimum/optima”, which is/are included at line 19 and later compared with all the other candidate(s) at line 25. This case is demonstrated by Example 3.2 (Sec. 3).
 - c) An (equality-constrained) QP but unbounded below, as determined by Ac) of Theorem 2.2. This derivation is similar to iia) above and thus omitted.
 - d) An (equality-constrained) QP of finite optimal value, with its unique optimum, or the particular solution of optima, outside of the polyhedron \mathcal{F} ; in other words, not satisfying the other inequality constraint(s) that

is/are not involved in this augmented system. Notably, there is still only one possibility that the optimality of QP Problem [15] is associated with this case, which corresponds to the nonuniqueness of optima. Specifically, this possibility is owing to the further imposed equality constraint(s), or the boundary of inequality constraint(s) from $\mathcal{I} \setminus \mathcal{I}_j$, on the degree of freedom of the optima with respect to the augmented QP under equality constraints and additional ones from \mathcal{I}_j . This will be elsewhere included/considered by virtue of the for-all environment in Algorithm 1.

- iii) Regarding the full-column-rank \tilde{A} , at lines 20-24, the only feasible point/optimum with respect to the augmented equality-constrained QP is computed a priori at line 21. If this point is also feasible with respect to the original/general QP Problem [15], as determined at line 22, then the corresponding, only, feasible/optimal value is later computed at line 23, and included as an optimality candidate at line 24 (the overall comparison in the end of this algorithm).

Finally, at line 25, the result concludes by comparing all the candidates in $\tilde{\mathcal{L}}$ of finite elements. As a matter of fact, the cardinality of $\tilde{\mathcal{L}}$ is at most 2^k , which is detailed later in Remark C.3. The minimum/optimal value \tilde{l}^* is associated with the corresponding *terminal optimum/optima* in $\tilde{\mathcal{X}}$, which can be either the unique optimum, or the particular/terminal solution of nonunique optima, located on the “edge(s)” or “vertex/vertices” as determined by $\tilde{\mathcal{I}}^*$.

Remark C.1. In the end of proof/algorithm, such an optimality pair $(\tilde{l}^*, \mathbf{x}^*, \tilde{\mathcal{I}}^*)$ can be nonunique. A more in-depth interpretation is that the optimality can occur at the different/nonunique optima, or different “edges/vertices”. The information regarding the latter is completely gathered by means of this algorithm. As for the former, if the optimum is unique with respect to all possible $\tilde{\mathcal{I}}^*(\cdot)$ s, then this information is also complete. On the other hand, if the optima (with respect to some $\tilde{\mathcal{I}}^*$) are nonunique, then, to remain focused of this presentation, preliminarily a subset is formulated using Algorithm 1. This subset includes all the *terminal optima*. That is, Algorithm 1 preliminarily gives closed-form solutions for all the optimum/optima at the terminals; while the intermediate points require specific attention due to the flexibility of inequality constraints. Although this indicates an interesting further research direction, Example 3.2 in Sec. 3 demonstrates a potential of the proposed results for such a comprehensive generalization. That is, a closed-form solution to the considered QP, including an explicit representation for all the (nonunique) optima.

Remark C.2. Fig. 4 demonstrates a fundamental concept “optimality candidates” for the design of Algorithm 1, in terms of the case where all $\mathbf{x}^* \notin \mathcal{F}$. It is worth noting that, according to Eq. [13], if there exists an ϵ^* such that the corresponding $\mathbf{x}^* \in \mathcal{F}$, then $\tilde{l}^* = \bar{l}^*$, and follow similar discussions in the remaining of this remark. Regarding this planar example of Fig. 4, the optimality occurs at (d) \mathbf{x}^* , and there are three representative cases/points (a)-(c) that are considered in the for-all design at lines 11-24 of Algorithm 1:

- a) The corresponding augmented pair $(\tilde{A}, \tilde{\mathbf{b}})$ satisfies all the conditions at lines 14 and 17, except that its unique opti-

1225 mum is outside of the feasible set \mathcal{F} . Thus, this point is
1226 not an optimality candidate.

1227 b) The pair $(\tilde{A}, \tilde{\mathbf{b}})$ satisfies all the conditions at lines 14 and
1228 17, where its unique optimum is located on an edge of \mathcal{F} .
1229 Hence, this optimality candidate $\tilde{\mathbf{x}}_p^* = \tilde{\mathbf{x}}^*$ is included at
1230 line 19, for the overall comparison later at line 25.

1231 c) Since \tilde{A} is of full column rank and its associated unique
1232 feasible point (optimum) $\tilde{\mathbf{x}}$ is also feasible with regard to
1233 the other inequality constraints, $\tilde{\mathbf{x}}$ is a candidate and also
1234 a vertex of \mathcal{F} , which is included at line 24 for the overall
1235 comparison (line 25).

1236 However, the optimality of this demonstration occurs at the
1237 vertex/candidate (d) \mathbf{x}^* , whose arguments similarly follow the
1238 above (c). After comparing the associated, finitely many opti-
1239 mal values among all the candidates (line 25), including the
1240 ones at (b)-(d), the corresponding unique optimum \mathbf{x}^* , opti-
1241 mal (level set) value \tilde{l}^* , and information for the location of \mathbf{x}^*
1242 (\tilde{T}^*) are exactly determined. Note that the complete and ex-
1243 plicit formulation of the QP example in Fig. 4 is not required
1244 for the discussions in this remark, particularly regarding the
1245 inequality constraints, since it is compactly two-dimensional
1246 for ease of reading.

1247 *Remark C.3.* This is to determine the largest possible cardi-
1248 nality of $\tilde{\mathcal{L}}$. To start with, note that there are 2^{κ} 's different
1249 subsets of \mathcal{I} . The case of empty set is considered in lines 3-9 to
1250 check the feasibility of \mathbf{x}_p^* , which is associated with $\tilde{\mathcal{L}} = \emptyset$ (as
1251 in the designed initial condition); while its possible freedom
1252 ($V_2\epsilon^*$) is designed to be checked within the for-all environ-
1253 ment (lines 11-24), if the solutions to QP Problems [10] and
1254 [15] intersect, such as the demonstrations (a) and (b) in Fig.
1255 3. On the other hand, the for-all design/coverage includes
1256 all the other subsets (note that $k \geq 1$ at line 11), which is
1257 associated with, at most, $(2^{\kappa} - 1)$'s elements for $\tilde{\mathcal{L}}$.

1258 *Remark C.4.* An extreme example considers the singleton
1259 $\mathcal{F} \neq \emptyset$ (noting that the subset \emptyset contains no point), and
1260 this remark explains how Algorithm 1 solves it. Similarly,
1261 divide the derivations into whether $\{\mathbf{x}_p^*\} = \mathcal{F}$. Regarding
1262 the equivalent/simpler case, then the solving process can be
1263 more efficiently completed without excessively going into the
1264 for-all environment (lines 11-24), while returning and ensur-
1265 ing the correct optimality. On the other hand, the process at
1266 lines 14-19 yields the optimality, only if $\tilde{\mathbf{x}}_p^* \in \mathcal{F}$ for some \mathcal{I}_j ;
1267 while that at 20-24 “exhaustively” includes this case, a ver-
1268 tex of \mathcal{F} . Although in the former (two) cases, this algorithm
1269 only concentrates on the particular solution $\mathbf{x}_p^*/\tilde{\mathbf{x}}_p^*$; nonethe-
1270 less, this example also demonstrates that the consideration on
1271 the freedom ($V_2\epsilon^*/\tilde{V}_2\tilde{\epsilon}^*$) can be more than necessary, where,
1272 for brevity, the notation $\tilde{\epsilon}^*$ similarly follows ϵ^* but with re-
1273 spect to the augmented pair $(\tilde{A}, \tilde{\mathbf{b}})$. As a comparison, it is
1274 not straightforward to apply most of the existing QP solvers
1275 (5, 13) to this singleton example, which are commonly/classi-
1276 cally designed from the differential perspective while initiated
1277 by (knowing) a feasible point, such as the popular ASM (4).

1278 D. Proof of Theorem 2.3 (Extended QP)

1279 A) By i) of Theorem 2.1, rewrite equivalently the CCQF at
1280 any level set value of \tilde{l} as the following constrained convex

quadratic equation (CCQE):

$$(1/2)\mathbf{x}^T P \mathbf{x} + \mathbf{q}^T \mathbf{x} + s - \tilde{l} = 0$$

$$\text{subject to both } \mathbf{x}, \mathbf{q} \in \mathcal{R}(P). \quad [25]$$

According to Lemma 2.1, a necessary condition that renders
the above CCQE in [25] solvable is

$$\begin{cases} \mathbf{q}^T P^{-1} \mathbf{q} - 2s + 2\tilde{l} \geq 0, & \text{if } \text{rank}(P) = n; [26a] \\ \mathbf{q}^T P^\dagger \mathbf{q} - 2s + 2\tilde{l} \geq 0 \text{ and } \mathbf{x} \in \mathcal{R}(P), & \text{if } \text{rank}(P) < n. [26b] \end{cases}$$

Note that the constraint “ $\mathbf{q} \in \mathcal{R}(P)$ ” has been (implicitly)
taken into consideration. The remaining part of this proof
starts with the derivations for the sufficiency of Condition [26].
Accordingly, divide the derivations into whether the Hessian
matrix is of full rank or not, as A-1) and A-2) below.

A-1) $\text{rank}(P) = n$. In this case, we have i) P is nonsingular,
ii) $\mathcal{R}(P) = \mathbb{R}^n$, and thus iii) the constraint “both $\mathbf{x}, \mathbf{q} \in \mathcal{R}(P)$ ”
is automatically lifted, which ensures the sufficiency of
Condition [26a]. By virtue of Eq./Parameterization [4] under
Condition [26a], the result thus follows.

A-2) $\text{rank}(P) = r < n$. Let the SVD of P be given by

$$P = [\tilde{U}_1 \ \tilde{U}_2] \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{U}_1^T \\ \tilde{U}_2^T \end{bmatrix}$$

$$= \tilde{U}_1 \tilde{\Sigma}_1 \tilde{U}_1^T, \quad [27]$$

where $\tilde{U}_1 \in \mathbb{R}^{n \times r}$, $\tilde{U}_2 \in \mathbb{R}^{n \times (n-r)}$, and $\tilde{\Sigma}_1 \in \mathbb{R}^{r \times r}$. For con-
ciseness, the properties of this SVD [27] for derivations after-
ward are similar to, and thus referenced at, Eqs. (35-37) in (2).
Likewise, define $P^\dagger := \tilde{U}_1 \tilde{\Sigma}_1^{-1} \tilde{U}_1^T$ and $P^{\dagger/2} := \tilde{U}_1 \tilde{\Sigma}_1^{-1/2} \tilde{U}_1^T$.
As a result, the CCQE in [25] is equivalent to the following
CQE:

$$(1/2)\mathbf{x}^T \tilde{U}_1 \tilde{\Sigma}_1 \tilde{U}_1^T \mathbf{x} + \mathbf{q}^T (\tilde{U}_1 \tilde{U}_1^T + \tilde{U}_2 \tilde{U}_2^T) \mathbf{x} + s - \tilde{l} = 0$$

$$\Leftrightarrow (1/2)\tilde{\mathbf{x}}^T \tilde{\Sigma}_1 \tilde{\mathbf{x}} + \tilde{\mathbf{q}}^T \tilde{\mathbf{x}} + s - \tilde{l} = 0, \quad [28]$$

where $\tilde{U}_1 \tilde{U}_1^T + \tilde{U}_2 \tilde{U}_2^T = I_n$ is inserted purposely, and $\tilde{\mathbf{x}} = \tilde{U}_1^T \mathbf{x} \in \mathbb{R}^r$ (resp., $\tilde{\mathbf{q}} = \tilde{U}_1^T \mathbf{q} \in \mathbb{R}^r$) is the coordinate vector of
 $\mathbf{x} \in \mathcal{R}(P)$ (resp., $\mathbf{q} \in \mathcal{R}(P)$) with respect to the \tilde{U}_1 -basis.
Note that the original CCQF has been equivalently trans-
formed into the CQE [28] with the full-rank Hessian matrix
 $\tilde{\Sigma}_1$, and the dimension is reduced from n to r (because of the
constraint “both $\mathbf{x}, \mathbf{q} \in \mathcal{R}(P)$ ”). In addition, the solvability
of the CQE [28] has already been ensured. To see this, by A)
of Lemma 2.1, the CQE [28] is solvable, iff

$$(1/4)\tilde{\mathbf{q}}^T (\tilde{\Sigma}_1/2)^{-1} \tilde{\mathbf{q}} \geq s - \tilde{l}$$

$$\Leftrightarrow \mathbf{q}^T \tilde{U}_1 \tilde{\Sigma}_1^{-1} \tilde{U}_1^T \mathbf{q} - 2s + 2\tilde{l} \geq 0. \quad [29]$$

This has already been given/guaranteed according to Condi-
tion/Eq. [26b], and also ensures its sufficiency that renders
the above CCQE [25] solvable.

Therefore, by virtue of Parameterization/Eq. [4], the solu-
tion set to the CQE in [28], where $\tilde{\Sigma}_1$ is of full rank (r), is
given as

$$\tilde{\mathbf{x}} = -\tilde{\Sigma}_1^{-1} \tilde{\mathbf{q}} + \sqrt{\tilde{\mathbf{q}}^T \tilde{\Sigma}_1^{-1} \tilde{\mathbf{q}}/2 - s + \tilde{l}} \cdot (\tilde{\Sigma}_1/2)^{-1/2} \hat{\mathbf{v}}, \quad [29]$$

where $\hat{\mathbf{v}} \in \mathbb{R}^r$ and $\|\hat{\mathbf{v}}\| = 1$. In terms of the original coordi-
nate, finally we obtain all the constrained solutions (namely,

elements):

$$\begin{aligned} \mathbf{x} &= \tilde{U}_1 \tilde{\mathbf{x}} \\ &= \sqrt{\mathbf{q}^T \tilde{U}_1 \tilde{\Sigma}_1^{-1} \tilde{U}_1^T \mathbf{q} / 2 - s + \check{l} \cdot \sqrt{2} \cdot \tilde{U}_1 \tilde{\Sigma}_1^{-1/2} \cdot \tilde{U}_1^T \tilde{U}_1 \cdot \hat{\mathbf{v}}} \\ &\quad - \tilde{U}_1 \tilde{\Sigma}_1^{-1} \tilde{U}_1^T \mathbf{q} \\ &= \text{Eq. [17]}, \end{aligned}$$

where $\tilde{U}_1^T \tilde{U}_1 = I_r$ is inserted purposely, $\check{\rho} = \tilde{U}_1 \hat{\mathbf{v}} \in \mathcal{R}(P)$, and $\|\check{\rho}\| = \|\tilde{U}_1 \hat{\mathbf{v}}\| = \|\hat{\mathbf{v}}\| = 1$. Remarkably, the constraint $\mathbf{x} \in \mathcal{R}(P)$ is satisfied since the two vectors in Eq. [17], $P^\dagger \mathbf{q}$ and $P^{\dagger/2} \check{\rho}$, reside in $\mathcal{R}(P)$.

B) The optimal value is finite, whose arguments similarly follow the proof for ii) of Theorem 2.1, and is mainly due to the constraint “ $\mathbf{q} \in \mathcal{R}(P)$ ”. Moreover, according to the above derivations for A), the necessary condition in Eq. [26] is also sufficient to render the CCQE [25] solvable. Thus, the optimal value \check{l}^* readily follows from the Condition [26]. Substituting \check{l}^* into Eq. [17] yields the associated unique optimum \mathbf{x}^* , which completes the proof.

Remark D.1. A geometric interpretation of the preimage, as addressed in A) of Theorem 2.3, can be similarly inferred from Fig. 2, where the solution freedom residing in $\mathcal{N}(P)$ is not available/effective. This is essentially due to the constraint “ $\mathbf{x} \in \mathcal{R}(P)$ ” in Problem [16], while the other “ $\mathbf{q} \in \mathcal{R}(P)$ ” mainly renders this optimization problem with a finite optimal value. From a different viewpoint, this remark is also consistent with Eq. [7], whose derivations are similar and hence omitted for conciseness.

E. Tables for all \mathcal{I}_j 's in Examples 3.1 and 3.2 (Sec. 3)

This appendix is designed to facilitate comparisons between Examples 3.1 and 3.2. Table 2 (resp., 3) is associated with Example 3.1 (resp., 3.2), which demonstrates the solving process on a *positive definite Hessian matrix* given in (4) (resp., *positive semidefinite and singular Hessian matrix*), to summarize the details/steps when examining each case \mathcal{I}_j according to Algorithm 1, specifically, the for-all environment at lines 11-24. Note that the case of subset/empty set “ \emptyset ” (in other words, no inequality constraint imposed) has been more efficiently considered a priori, in the beginning of the two examples (lines 2-9 in Algorithm 1). Moreover, $\mathcal{I}_7 = \{1, 2, 3\}$ imposed by all the three inequality constraints is excluded in both tables/examples, for compactness, because it is obviously not associated with the optimality ($\tilde{\mathbf{b}} \notin \mathcal{R}(\tilde{A})$).

Table 2. A summary of significant parameters and values for all \mathcal{I}_j 's in Example 3.1

\mathcal{I}_j ($\tilde{\mathbf{b}} \in \mathcal{R}(\tilde{A})$)	\tilde{A}	$\tilde{\mathbf{b}}$	\tilde{V}_2	$\tilde{V}_2^T P \tilde{V}_2$	Categorization	$\tilde{\mathbf{x}}^*/\hat{\mathbf{x}}?$	Candidate?	$\tilde{l}^*/\hat{l}?$
$\mathcal{I}_1 = \{1\}$	$\begin{bmatrix} 1 & 1 \end{bmatrix}$	4	$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T$	2	QP	$\tilde{\mathbf{x}}^* = \begin{bmatrix} 1.5 & 2.5 \end{bmatrix}^T$	yes	$\tilde{l}^* = -28.5$
$\mathcal{I}_2 = \{2\}$	$\begin{bmatrix} -1 & 0 \end{bmatrix}$	0	$\begin{bmatrix} 0 & 1 \end{bmatrix}^T$	2	QP	$\tilde{\mathbf{x}}^* = \begin{bmatrix} 0 & 5 \end{bmatrix}^T$	no	n/a
$\mathcal{I}_3 = \{3\}$	$\begin{bmatrix} 0 & -1 \end{bmatrix}$	0	$\begin{bmatrix} -1 & 0 \end{bmatrix}^T$	4	QP	$\tilde{\mathbf{x}}^* = \begin{bmatrix} 3 & 0 \end{bmatrix}^T$	yes	$\tilde{l}^* = -18$
$\mathcal{I}_4 = \{1, 2\}$	$\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	$\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$	yes	$\hat{l} = -24$
$\mathcal{I}_5 = \{2, 3\}$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	$\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	yes	$\hat{l} = 0$
$\mathcal{I}_6 = \{1, 3\}$	$\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	$\hat{\mathbf{x}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$	yes	$\hat{l} = -16$

Table 3. A summary of significant parameters and values for all \mathcal{I}_j 's in Example 3.2

\mathcal{I}_j ($\tilde{\mathbf{b}} \in \mathcal{R}(\tilde{A})$)	\tilde{A}	$\tilde{\mathbf{b}}$	\tilde{V}_2	$\tilde{V}_2^T P \tilde{V}_2$	Categorization	$\mathcal{N}(P) \cap \mathcal{N}(\tilde{A})$	$\tilde{\mathbf{x}}^*/\hat{\mathbf{x}}?$	Candidate?	$\tilde{l}^*/\hat{l}?$
$\mathcal{I}_1 = \{1\}$	$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 4 \end{bmatrix}$	$\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$	$\frac{1}{2}$	QP	$\{0\}$	$\tilde{\mathbf{x}}^* = \begin{bmatrix} 0 \\ 4 \\ b \end{bmatrix}$	yes	$\tilde{l}^* = 0$
$\mathcal{I}_2 = \{2\}$	$\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} b \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$	0	constant ^a	n/a	n/a	n/a ^b	n/a
$\mathcal{I}_3 = \{3\}$	$\begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} b \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	1	QP	$\{0\}$	$\tilde{\mathbf{x}}^* = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix}$	yes	$\tilde{l}^* = 0$
$\mathcal{I}_4 = \{1, 2\}$	$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} b \\ 4 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	n/a	$\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 4 \\ b \end{bmatrix}$	yes	$\hat{l} = 0$
$\mathcal{I}_5 = \{2, 3\}$	$\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	n/a	$\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix}$	yes	$\hat{l} = 0$
$\mathcal{I}_6 = \{1, 3\}$	$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} b \\ 4 \\ 0 \end{bmatrix}$	n/a	n/a	vertex	n/a	$\hat{\mathbf{x}} = \begin{bmatrix} 4 \\ 0 \\ b \end{bmatrix}$	yes	$\hat{l} = 16$

^a $\tilde{V}_2^T \mathbf{q} + \tilde{V}_2^T P \tilde{A}^\dagger \tilde{\mathbf{b}} = 0$ (not an LP), and the constant value is $s + \mathbf{q}^T \tilde{A}^\dagger \tilde{\mathbf{b}} + \tilde{\mathbf{b}}^T (\tilde{A}^\dagger)^T P \tilde{A}^\dagger \tilde{\mathbf{b}}/2 = 0$.

^b The associated *terminal optima* are explicitly formulated after further imposing constraints (\mathbf{c}_1, d_1) and (\mathbf{c}_3, d_3) , as included in Cases $\{\mathcal{I}_4\}$ and $\{\mathcal{I}_5\}$, respectively.

ACKNOWLEDGMENTS. The authors thank Julian Mellor and Scott Read for their assistance in the native use of this language. This work was supported by the Max Planck Society, Germany; and the Ministry of Science and Technology, Taiwan.

1. S. K. Vaddamani, T. P. Xiao, and E. Yablonovitch. Physics successfully implements Lagrange multiplier optimization. *Proc. Natl. Acad. Sci. U. S. A.*, 117(43):26639–26650, 2020.
2. L.-G. Lin, Y.-W. Liang, and W.-Y. Hsieh. Convex quadratic equation. *J. Optim. Theory Appl.*, 186(3):1006–1028, Sep. 2020.
3. T. Çimen. Systematic and effective design of nonlinear feedback controllers via the state-dependent Riccati equation (SDRE) method. *Annu. Rev. Control*, 34(1):32–51, 2010.
4. D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Switzerland: Springer, 2016. 2nd and 4th/latest editions.
5. S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004. 23rd/latest printing with corrections, 2018, and its solution manual.
6. D. F. Cioacan and V. F. Farias. Model predictive control for dynamic resource allocation. *Math. Oper. Res.*, 37(3):501–525, 2012.
7. S. Kang, J. Wang, G. Li, J. Shan, and I. R. Petersen. Optimal cooperative guidance law for salvo attack: An MPC-based consensus perspective. *IEEE Trans. Aerosp. Electron. Syst.*, 54(5):2397–2410, 2018.
8. P. Quillen and K. Subbarao. Minimum control effort-based path planning and nonlinear guidance for autonomous mobile robots. *Int. J. Adv. Robot. Syst.*, 15(6):1–23, 2018.
9. A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, 2002.
10. P. Patrinos and H. Sarimveis. A new algorithm for solving convex parametric quadratic programs based on graphical derivatives of solution mappings. *Automatica*, 46(9):1405–1418, 2010.
11. J. B. Rawlings, D. Bonné, J. B. Jørgensen, A. N. Venkat, and S. B. Jørgensen. Unreachable setpoints in model predictive control. *IEEE Trans. Automat. Control*, 53(9):2209–2215, 2008.
12. O. L. Mangasarian. A simple characterization of solution sets of convex programs. *Oper. Res. Lett.*, 7(1):21–26, 1988.
13. J. Nocedal and S. J. Wright. In T. V. Mikosch, S. I. Resnick, and S. M. Robinson, editors, *Numerical Optimization*, Springer Ser. Oper. Res. Financ. Eng. USA: Springer, 2nd/latest edition, 2006.
14. T. A. Johansen, T. I. Fossen, and P. Tøndel. Efficient optimal constrained control allocation via multiparametric programming. *J. Guid. Control Dyn.*, 28(3):506–515, 2005.
15. D. G. Luenberger. *Optimization by Vector Space Methods*. New York, USA: Wiley, 1969.
16. J. C. Duchi, M. I. Jordan, M. J. Wainwright, and A. Wibisono. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Trans. Inf. Theory*, 61(5):2788–2806, 2015.
17. T. A. Johansen and T. I. Fossen. Control allocation—A survey. *Automatica*, 49(5):1087–1103, 2013.
18. S. V. Raković and W. S. Levine, editors. *Handbook of Model Predictive Control*. Cham, Switzerland: Birkhäuser, 2019.
19. J. A. Urigüen, T. Blu, and P. L. Dragotti. FRI sampling with arbitrary kernels. *IEEE Trans. Signal Process.*, 61(21):5310–5323, 2013.
20. L.-G. Lin, Y.-W. Liang, and W.-Y. Hsieh. Convex quadratic equations and functions. *arXiv*, 2019. URL arxiv.org/abs/1906.00177. arxiv.org/abs/1906.00177.
21. G. H. Golub and C. F. Van Loan. *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 4th/latest edition, 2013.
22. M. Heydari. Optimal inspection policy for a second-hand product with a two-dimensional warranty. *IMA J. Manag. Math.*, 8 2021. ISSN 1471-678X. dpab030.
23. H. Konno and H. Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Sci.*, 37(5):519–531, 1991.
24. H. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. USA: Springer-Verlag, 2nd edition, 2003.
25. A. Mittal and G. A. Hanasusanto. Finding minimum volume circumscribing ellipsoids using generalized copositive programming. *Oper. Res.*, 2021. .
26. MATLAB® Documentation, 2022. (“quadprog” function, accessed in Mar. 2022). The Mathworks, Inc.
27. Y.-K. Tsai and R. Malak. A methodology for designing a nonlinear feedback controller via parametric optimization: State-parameterized nonlinear programming control. *Proc. ASME International Design Engineering Technical Conf. and Computers and Information in Engineering Conf.*, 2021. V03AT03A011.