

On the Complexity of Detecting Convexity over a Box

Amir Ali Ahmadi and Georgina Hall *

Abstract

It has recently been shown that the problem of testing global convexity of polynomials of degree four is strongly NP-hard, answering an open question of N.Z. Shor. This result is minimal in the degree of the polynomial when global convexity is of concern. In a number of applications however, one is interested in testing convexity only over a compact region, most commonly a box (i.e., hyper-rectangle). In this paper, we show that this problem is also strongly NP-hard, in fact for polynomials of degree as low as three. This result is minimal in the degree of the polynomial and in some sense justifies why convexity detection in nonlinear optimization solvers is limited to quadratic functions or functions with special structure. As a byproduct, our proof shows that the problem of testing whether all matrices in an interval family are positive semidefinite is strongly NP-hard. This problem, which was previously shown to be (weakly) NP-hard by Nemirovski, is of independent interest in the theory of robust control.

Keywords: Convexity detection, convex optimization, computational complexity, interval positive semidefiniteness.

1 Introduction

In a relatively recent paper [3], Ahmadi, Olshevsky, Parrilo, and Tsitsiklis have shown that the problem of testing whether a quartic (multivariate) polynomial is globally convex is NP-hard. This result answers a question of N.Z. Shor that had appeared in 1996 on a list of seven open problems in computational complexity for numerical optimization [20]. The significance of the question stems from the fact that in the theory of optimization, the divide between convex and nonconvex optimization has proved to be a fundamental one. Indeed, for many classes of optimization problems, stronger algorithmic, analytic, and geometric statements can be made when the functions involved in the description of the problem are convex compared to when they are not. Hence, it is natural to ask (as Shor did) whether one can test for convexity of functions in an efficient manner. Polynomial functions provide a convenient setting for a rigorous study of this question from a computational complexity viewpoint.

For a number of problems of applied nature, one is interested not in checking for convexity of a function globally, but only over a compact region. The most common region that arises in practical applications is a *box* (i.e., a hyper-rectangle). For example, the variables in many real-world optimization problems are lower and upper bounded a priori, allowing the algorithm designer to restrict attention to a box. Similarly, in branch-and-bound approaches to nonlinear and nonconvex optimization, one often recursively breaks the original problem into a number of

*The authors are with the department of Operations Research and Financial Engineering at Princeton University. Email: {a.a.a, gh4}@princeton.edu. This work was partially supported by the DARPA Young Faculty Award, the CAREER Award of the NSF, the Google Faculty Award, the Innovation Award of the School of Engineering and Applied Sciences at Princeton University, and the Sloan Fellowship.

subproblems over smaller boxes and aims to provide upper and lower bounds on the minimum of the objective function over these smaller boxes. In such a scenario, the subproblems for which the objective function is convex over the associated box can often be solved to global optimality.

The problem of detecting convexity over a box, or the related problem of *imposing* convexity over a box, appears in areas outside of optimization as well. In control theory for example, Lyapunov functions that certify properties of dynamical systems are sometimes required to be convex over the region of the space where the state variables lie [10, 2]. Hence, the need for algorithms that check convexity of a candidate Lyapunov function over a compact region—or those that find a Lyapunov function that meets the convexity requirement by construction—naturally arises. Similarly, in statistics, the well-known problem of *convex regression* [23, 14, 17] is about finding a convex function (within a certain function class) that best agrees with observed data. In this situation again, imposing convexity of the function globally is often too restrictive. Indeed, one generally has access to a feature domain where the explanatory variables of the problem are known to lie. This domain, more often than not, is a box.

Motivated by these considerations, our goal in this paper is to study the complexity of testing convexity of a given function over a box (which can only be simpler than the problem of imposing convexity over a box). As is done in [3], we restrict our attention to polynomial functions since they appear routinely in applications and are amenable to complexity theoretic investigations due to their finite parametrization. Our main result (Theorem 2.3) shows that the problem of testing convexity of a polynomial over a box is strongly NP-hard already for cubic polynomials. This result completely classifies the complexity of testing convexity of polynomials of any degree over a box (cf. Proposition 2.7). It also justifies, at least from a complexity viewpoint, why convexity detection in some of the most widely-used optimization packages such as BARON [21, 16], CVX [13], and Gurobi [1] is restricted to quadratic functions, or is replaced with the task of checking stronger but more tractable conditions. These stronger conditions require e.g. that the function in question be the output of certain convexity-preserving operations applied to an initial set of convex functions; see, e.g., [9, Section 3.2].

The proof of our main result has two parts, the first of which relies heavily on a clever gadget by Nemirovski [19]. Some of our modifications to his arguments are of potential interest to the field of robust control (cf. Corollary 2.6).

2 The Main Result

We begin with some basic definitions.

Definition 2.1. *Given a set of scalars $l_1, \dots, l_n, u_1, \dots, u_n$, with $l_i \leq u_i$ for $i = 1, \dots, n$, a box $B \subseteq \mathbb{R}^n$ is a set of the form*

$$B = \{x \in \mathbb{R}^n \mid l_i \leq x_i \leq u_i, i = 1, \dots, n\}.$$

Definition 2.2. *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be convex over a set $C \subseteq \mathbb{R}^n$ if for every $x, y \in C$ and for any $\lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.*

We study the complexity of detecting convexity in the standard Turing model of computation (see, e.g., [22]) where the input to every problem instance must be defined by a finite number of bits. As a consequence, in the statement of Theorem 2.3 below, the input to the problem—which consists of the coefficients of the polynomial f and the scalars $l_1, \dots, l_n, u_1, \dots, u_n$ that define the box B —is taken to be rational.

We further remark that the problem in the statement of Theorem 2.3 is *strongly* NP-hard. This is as opposed to problems that are *weakly* NP-hard such as the classical problems of KNAPSACK or PARTITION [11]. This distinction only occurs in problems whose instances involve numerical data¹ and relates to how the numerical values that appear in an instance compare against the length of the instance. Roughly speaking, being NP-hard in the strong sense means that the problem remains NP-hard even when the numerical values in all problem instances are “small”. More rigorously, define a Max (resp. Length) function which maps any instance I of a decision problem to a nonnegative integer $\text{Max}(I)$ (resp. $\text{Length}(I)$) that represents the largest numerator or denominator in magnitude of all rational numbers appearing in I (resp. the number of bits required to write down I). A decision problem Π is said to be *strongly NP-hard* if there exists a subproblem $\hat{\Pi}$ which is NP-hard and for which one can find a polynomial p such that $\text{Max}(I)$ is upperbounded by $p(\text{Length}(I))$ for all instances $I \in \hat{\Pi}$.

The advantage of showing that a problem is strongly NP-hard (as opposed to weakly NP-hard) is that such a statement rules out, unless $P=NP$, the possibility of a *pseudo-polynomial time* algorithm. This is an algorithm whose running time is polynomial in the numerical value of the input but not necessarily in the bit length of the input; see [11, Section 4.2.2] for more details. More concretely, if we had only shown weak NP-hardness of testing convexity of an n -variate cubic polynomial over a box, it could be that the problem is polynomial-time solvable (even assuming $P \neq NP$) except for instances where some of the coefficients of the polynomial or the bounds of the box are of order 2^n . Indeed, for the KNAPSACK and PARTITION problems for example, it is well known that one can solve, in polynomial time, all instances which do not contain such large numbers, using dynamic programming [11, p.96].

Theorem 2.3 (Main result). *Given a cubic polynomial f and a box B , it is strongly NP-hard to test whether f is convex over B .*

The proof of this theorem is based on a reduction from the problem of testing whether a matrix $L(x)$ —whose entries are affine polynomials in x —is positive semidefinite for all x in a full-dimensional box B . This problem has already been shown to be NP-hard by Nemirovski [19]. The connection between this problem and the problem of testing convexity of a cubic polynomial $f := f(x)$ over a box B lies in the facts that (i) the Hessian $\nabla^2 f(x)$ of $f(x)$ is an affine polynomial matrix, and (ii) a twice continuously differentiable function f is convex over a convex set S with nonempty interior if and only if $\nabla^2 f(x) \succeq 0$, $\forall x \in S$; see, e.g., [7, Chapter 1]. (Here, the notation $M \succeq 0$ is used to denote that a symmetric matrix M is positive semidefinite, i.e., has nonnegative eigenvalues.) The proof of Theorem 2.3 is split into two parts which we outline below.

In the first part of the proof, we concern ourselves with strengthening the NP-hardness result from [19], which we use here, from weak to strong (Theorem 2.4). Indeed, the reduction provided in [19] is from the PARTITION problem which is only weakly NP-hard [11], and hence only proves weak NP-hardness of testing positive semidefiniteness of an affine polynomial matrix over a box. As a consequence, we give a new reduction, which still uses the insights that we gleaned from Nemirovski’s proof, but with two modifications. First, we give a reduction from a strongly NP-hard problem. Second, we bypass a step in the Nemirovski’s construction that involves matrix inversion. This is because it is possible for the inverse of a matrix to have entries that are exponential in the entries of the original matrix, thus causing the reduction to lose its strong NP-hardness implication. To bypass matrix inversion, we instead approximate the inverse by its first-order Taylor expansion and control the spectrum of the matrix in such a way that the higher-order terms in the Taylor expansion can be appropriately bounded (cf. Lemma 2.5).

¹Note that the instances of not all of NP-hard problems involve numerical data; consider, e.g., SATISFIABILITY [11]. For such problems, NP-hardness is by default in the strong sense.

In the second and core part of the proof, we give the reduction from the problem of testing positive semidefiniteness of an affine polynomial matrix over a box to the problem of testing convexity of a cubic polynomial over a box. The main obstacle that we need to overcome here is that not every affine (symmetric) polynomial matrix $L(x)$ is the Hessian of some cubic polynomial. For instance, if, for some integers $i, j, k \in \{1, \dots, n\}$, the equality

$$\frac{\partial L_{ij}(x)}{\partial x_k} = \frac{\partial L_{ik}(x)}{\partial x_j}$$

is violated, then $L(x)$ cannot be a valid Hessian matrix. This is because partial derivatives of polynomials must commute. Our goal is then to show that even with these additional constraints, the problem of checking positive semidefiniteness remains hard. To do this, we will introduce new variables y and carefully construct a polynomial $f(x, y)$ whose Hessian can be related back to $L(x)$.

Theorem 2.4 (see the proof of Proposition 2.1 in [19] for a related result). *Given a symmetric $n \times n$ matrix A with entries in $\{0, 1\}$ and a positive integer $k \leq n^2$, let*

$$C := \frac{4}{(n+1)^3} \left(I_n + \frac{1}{(n+1)^3} A \right), \text{ and } \mu := \frac{n(n+1)^3}{4} + k - 1 - \frac{1}{4} e^T A e, \quad (1)$$

where e is the $n \times 1$ vector of all ones and I_n is the $n \times n$ identity matrix. It is (strongly) NP-hard to test whether

$$L(x) := \begin{bmatrix} C & x \\ x^T & \mu + \frac{1}{4} \end{bmatrix} \succeq 0, \forall x \in \mathbb{R}^n \text{ with } \|x\|_\infty \leq 1. \quad (2)$$

Furthermore, for any matrix C and scalar μ thus defined, either (2) holds² or there exists $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$ such that $x^T C^{-1} x \geq \mu + \frac{3}{4}$.

To prove this theorem, we will make use of the following lemma.

Lemma 2.5. *Let A be a symmetric $n \times n$ matrix with entries in $\{0, 1\}$ and let C be defined as in (1). Then, for any $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$, we have*

$$\frac{1}{4} x^T ((n+1)^3 I_n - A) x - \frac{1}{4} \leq x^T C^{-1} x \leq \frac{1}{4} x^T ((n+1)^3 I_n - A) x + \frac{1}{4}.$$

Proof. First, note that C is by construction strictly diagonally dominant, and hence, by Gershgorin's circle theorem [12], positive definite. This implies that C^{-1} is well defined. Furthermore, we have

$$C^{-1} = \frac{(n+1)^3}{4} \left(I_n - \frac{-A}{(n+1)^3} \right)^{-1} = \frac{(n+1)^3}{4} \sum_{k=0}^{\infty} \frac{(-A)^k}{(n+1)^{3k}}, \quad (3)$$

where the second equality follows from the identity $(I_n - M)^{-1} = \sum_{k=0}^{\infty} M^k$, which holds when the spectral radius of M is less than one; see, e.g., [15, Section 5.6]. This assumption is met here as the spectral radius of $\frac{-A}{(n+1)^3}$ is equal to its 2-norm, which, by Gershgorin's circle theorem again, is bounded above by $\frac{n}{(n+1)^3} < 1$. Using (3), we have

$$x^T C^{-1} x = \frac{1}{4} x^T ((n+1)^3 I_n - A) x + \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{x^T (-A)^k x}{(n+1)^{3k}}.$$

²Note that by using the Schur complement, the constraint in (2) is equivalent to having $x^T C^{-1} x \leq \mu + \frac{1}{4}$ for all x with $\|x\|_\infty \leq 1$.

It remains to show that

$$\left| \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{x^T(-A)^k x}{(n+1)^{3k}} \right| \leq \frac{1}{4}$$

for any $x \in \mathbb{R}^n$ with $\|x\|_{\infty} \leq 1$ to conclude the proof. Letting $x \in \mathbb{R}^n$ with $\|x\|_{\infty} \leq 1$, we have

$$\begin{aligned} \left| \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{x^T(-A)^k x}{(n+1)^{3k}} \right| &\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{|x^T(-A)^k x|}{(n+1)^{3k}} \\ &\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{\|x\|_2 \cdot \|(-A)^k x\|_2}{(n+1)^{3k}} \\ &\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{\|x\|_2 \cdot \|A\|_2^k \cdot \|x\|_2}{(n+1)^{3k}} \\ &\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{n \|A\|_2^k}{(n+1)^{3k}} \\ &\leq \frac{n(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{1}{(n+1)^{2k}}, \end{aligned}$$

where the first inequality uses the triangle inequality, the second the Cauchy-Schwarz inequality, the third the definition and the submultiplicativity property of the matrix 2-norm, the fourth the facts that $\|x\|_2 \leq \sqrt{n}\|x\|_{\infty}$ and $\|x\|_{\infty} \leq 1$, and the fifth the fact that $\|A\|_2 \leq n \leq n+1$. Now, using the geometric series identity

$$\sum_{k=2}^{\infty} \frac{1}{(n+1)^{2k}} = \frac{1}{(n+1)^4} \cdot \frac{1}{1 - \frac{1}{(n+1)^2}},$$

we conclude that

$$\begin{aligned} \left| \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{x^T(-A)^k x}{(n+1)^{3k}} \right| &\leq \frac{1}{4} \cdot \frac{n}{n+1} \cdot \frac{1}{1 - \frac{1}{(n+1)^2}} \\ &= \frac{1}{4} \cdot \frac{n+1}{n+2} \\ &\leq \frac{1}{4}. \end{aligned}$$

□

Proof of Theorem 2.4. We give a reduction from the SIMPLE MAX-CUT problem, which is the following decision problem: Given a simple³ graph $G = (V, E)$ and a positive integer $s \leq |V|^2$, decide whether there is a cut of size greater or equal to s in G , i.e., a partition of V into two disjoint sets V_1 and V_2 such that the number of edges from E that have one endpoint in V_1 and one endpoint in V_2 is greater or equal to s . This problem is known to be (strongly) NP-hard [11, p. 210].

Consider an instance of SIMPLE MAX-CUT given by a graph $G = (V, E)$ and a positive integer $s \leq |V|^2$. We construct an instance of the problem given in the statement of the theorem by taking A to be the adjacency matrix of G (i.e., a symmetric $|V| \times |V|$ matrix whose (i, j) -th entry is equal

³Recall that a graph is *simple* if it is unweighted, undirected, and has no self-loops or multiple edges.

to one if $\{i, j\} \in E$ and equal to zero otherwise) and $k = s$. Clearly this reduction is polynomial in length. Now, let C and μ be as defined in (1) with $n = |V|$.

We show that $L(x)$ as defined in (2) is positive semidefinite for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$ if and only if there is no cut in G of size greater or equal to s . This would show that the problem in the statement of the theorem is NP-hard. In fact, as explained previously, since

$$\text{Max}(I) = k \leq n^2 + \lceil \log_2(k) \rceil = \text{Length}(I)$$

for any instance I of the problem, the problem is automatically strongly NP-hard; see e.g. [11, p. 95].

Suppose first that there is no cut in G of size greater or equal to s . As $k = s$, this implies that the largest cut in G is of size less than or equal to $k - 1$. It is straightforward to verify that the size of the largest cut in G is given by $\max_{x \in \{-1, 1\}^n} \frac{1}{4} \sum_{i, j} A_{ij}(1 - x_i x_j)$. Hence, if

$$\begin{aligned} p^* &:= \min_{x \in \mathbb{R}^n} \frac{1}{4} x^T A x \\ \text{s.t.} \quad &x \in \{-1, 1\}^n, \end{aligned} \tag{4}$$

the size of the largest cut in G is less than or equal to $k - 1$ if and only if $\frac{1}{4} e^T A e - p^* \leq k - 1 \Leftrightarrow p^* \geq \frac{1}{4} e^T A e - k + 1$. As $x^T x = n$ when $x \in \{-1, 1\}^n$, this is equivalent (after basic algebra) to the optimal value of

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad &\frac{1}{4} x^T ((n+1)^3 I_n - A) x \\ \text{s.t.} \quad &x \in \{-1, 1\}^n \end{aligned} \tag{5}$$

being less than or equal to $\frac{n(n+1)^3}{4} + k - 1 - \frac{1}{4} e^T A e = \mu$. As the quadratic function

$$x \mapsto \frac{1}{4} x^T ((n+1)^3 I_n - A) x$$

is convex, and as the maximum of a convex function over a box is attained at an extreme point of the box (see, e.g., [6, Property 12]), the previous statement is equivalent to the optimal value of

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad &\frac{1}{4} x^T ((n+1)^3 I_n - A) x \\ \text{s.t.} \quad &x \in [-1, 1]^n \end{aligned} \tag{6}$$

being less than or equal to μ . Hence we have shown that if the size of the largest cut in G is less than or equal to $k - 1$, then $\frac{1}{4} x^T ((n+1)^3 I_n - A) x \leq \mu$ for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$. From Lemma 2.5, it follows that

$$x^T C^{-1} x \leq \mu + \frac{1}{4}$$

for any $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$. Using the Schur complement, this is equivalent to $L(x)$ being positive semidefinite for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$.

Suppose now that there is a cut in G of size greater or equal to s . Let $\hat{x} \in \{-1, 1\}^n$ be the indicator vector of this cut; i.e., $\hat{x}_i = 1$ if node i belongs to V_1 and $\hat{x}_i = -1$ if node i belongs to V_2 . It is easy to check that the number of edges with one endpoint in V_1 and one endpoint in V_2 is given by $\frac{1}{4}(e^T A e - \hat{x}^T A \hat{x})$. Hence, as $k = s$, we have

$$\frac{1}{4}(e^T A e - \hat{x}^T A \hat{x}) \geq k,$$

which is equivalent to $\frac{1}{4}n(n+1)^3 - \frac{1}{4}\hat{x}^T A \hat{x} \geq \frac{1}{4}n(n+1)^3 - \frac{1}{4}e^T A e + k = \mu + 1$. As $\hat{x}^T \hat{x} = n$, it follows that

$$\frac{1}{4}\hat{x}^T ((n+1)^3 I_n - A) \hat{x} \geq \mu + 1.$$

Using Lemma 2.5, we get

$$\hat{x}^T C^{-1} \hat{x} \geq \mu + \frac{3}{4} > \mu + \frac{1}{4}.$$

By the Schur complement, we conclude that $L(\hat{x}) \not\geq 0$.

Note that the final claim of the theorem is already implied by the arguments given thus far. Indeed, given C and μ as defined in (1), one can construct a graph whose adjacency matrix is A and take $s = k$. Either this graph has no cut of size greater or equal to s , in which case we have shown that $x^T C^{-1} x \leq \mu + \frac{1}{4}$ for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$, or it has a cut of size greater or equal to s , in which case there exists \hat{x} such that $\|\hat{x}\|_\infty \leq 1$ and $\hat{x}^T C^{-1} \hat{x} \geq \mu + \frac{3}{4}$. \square

We briefly provide an immediate corollary of this theorem, which we believe can be of independent interest as it relates to problems in robust control [5, 8, 4]. This statement was first proven in [19], except that the result there shows NP-hardness in the weak sense.

Corollary 2.6 (Strong NP-hardness of testing interval positive semidefiniteness). *Given rational numbers $\hat{m}_{ij}, \bar{m}_{ij}, i = 1, \dots, N, j = 1, \dots, N$, with $\hat{m}_{ij} \leq \bar{m}_{ij}$, $\hat{m}_{ij} = \hat{m}_{ji}$, and $\bar{m}_{ij} = \bar{m}_{ji}$ for $i = 1, \dots, N, j = 1, \dots, N$, it is strongly NP-hard to test whether all $N \times N$ symmetric matrices M with entries $m_{ij} \in [\hat{m}_{ij}, \bar{m}_{ij}]$ for $i = 1, \dots, N, j = 1, \dots, N$, are positive semidefinite.*

Proof. The claim follows straightforwardly by a reduction from the problem in the statement of Theorem 2.4. Let A be a symmetric $n \times n$ matrix with entries in $\{0, 1\}$ and $k \leq n^2$ be a positive integer. Construct the symmetric matrix C and the scalar μ as in (1). Let $\hat{m}_{ij} = \bar{m}_{ij} = C_{ij}$ for $i = 1, \dots, n, j = 1, \dots, n$, $\hat{m}_{i(n+1)} = \hat{m}_{(n+1)i} = -1$ and $\bar{m}_{i(n+1)} = \bar{m}_{(n+1)i} = 1$ for $i = 1, \dots, n$, and $\hat{m}_{(n+1)(n+1)} = \bar{m}_{(n+1)(n+1)} = \mu + \frac{1}{4}$. It follows that all $(n+1) \times (n+1)$ symmetric matrices M with entries $m_{ij} \in [\hat{m}_{ij}, \bar{m}_{ij}]$ for $i = 1, \dots, n+1, j = 1, \dots, n+1$, are positive semidefinite if and only if the matrix $L(x)$ defined in (2) is positive semidefinite for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$.

Note that the reduction we have given leads to a strong NP-hardness result. Indeed, all instances I of the interval positive semidefiniteness problem that arise from this construction have $\text{Length}(I)$ of order n^2 and $\text{Max}(I)$ of order n^6 , and hence $\text{Max}(I)$ is upper bounded by a polynomial in $\text{Length}(I)$. \square

We now move on to the proof of the main theorem of this paper.

Proof of Theorem 2.3. We prove the claim via a reduction from the NP-hard problem stated in Theorem 2.4. Let A be an $n \times n$ symmetric matrix with entries in $\{0, 1\}$ and $k \leq n^2$ be a positive integer. Let C and μ be as in (1) and define $L(x)$ to be the following $(n+1) \times (n+1)$ symmetric matrix

$$L(x) := \begin{bmatrix} C & x \\ x^T & \mu + \frac{1}{4} \end{bmatrix}.$$

Let $y := (y_1, \dots, y_{n+1})^T$ be a new vector of variables and $H(y)$ be the following $n \times (n+1)$ matrix of mixed partial derivatives of the cubic polynomial $y^T L(x) y$:

$$H_{ij}(y) = \frac{\partial^2 y^T L(x) y}{\partial x_i \partial y_j}, \text{ for } i = 1, \dots, n, \text{ and } j = 1, \dots, n+1.$$

Using the structure of $L(x)$, after some basic algebra, we obtain:

$$H(y) = \begin{bmatrix} 2y_{n+1} & 0 & \dots & 0 & 2y_1 \\ 0 & \ddots & 0 & 0 & 2y_2 \\ 0 & 0 & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 2y_{n+1} & 2y_n \end{bmatrix}. \quad (7)$$

Let

$$\begin{aligned} \alpha &:= 16 \cdot n(1 + 16n^7), \\ \eta &:= \frac{1}{4} \cdot \frac{1}{1 + 16n^7}. \end{aligned} \quad (8)$$

Consider the following cubic polynomial in $2n + 1$ variables $x_1, \dots, x_n, y_1, \dots, y_{n+1}$

$$f(x, y) = \frac{1}{2}y^T L(x)y + \frac{\alpha}{2}x^T x + \frac{\eta}{2}y^T y,$$

and the box $B = [-1, 1]^{2n+1}$. We claim that $f(x, y)$ is convex over B if and only if $L(x) \succeq 0$ for all $x \in \mathbb{R}^n$ with $\|x\|_\infty \leq 1$. This would imply the desired NP-hardness result as the construction of f and B from A and k can be carried out in polynomial time. Moreover, NP-hardness here is in the strong sense as all instances I of the problem in the statement of Theorem 2.3 that arise from this construction have $\text{Length}(I)$ of order n^3 (the number of coefficients of f) and $\text{Max}(I)$ of order n^8 (which is the magnitude of α , which dominates all other coefficients). Hence, $\text{Max}(I)$ is upper bounded by a polynomial in $\text{Length}(I)$.

To prove the claim, we begin by observing that the Hessian of f has the following structure:

$$\nabla^2 f(x, y) = \begin{bmatrix} \alpha I_n & \frac{1}{2}H(y) \\ \frac{1}{2}H(y)^T & L(x) + \eta I_{n+1} \end{bmatrix}.$$

Suppose first that $L(x)$ is not positive semidefinite over $[-1, 1]^n$. Then, from the second claim in Theorem 2.4, there exists $\bar{x} \in [-1, 1]^n$ such that $\bar{x}^T C^{-1} \bar{x} \geq \mu + \frac{3}{4}$. Let $\bar{y} = 0_{(n+1) \times 1}$, i.e., the vector of zeros of length $n + 1$, and observe that $(\bar{x}, \bar{y}) \in [-1, 1]^{2n+1} = B$. Let

$$z = \frac{1}{\|(-C^{-1}\bar{x}, 1)^T\|_2} \begin{pmatrix} 0_{n \times 1} \\ -C^{-1}\bar{x} \\ 1 \end{pmatrix}.$$

We will show that $z^T \nabla^2 f(\bar{x}, \bar{y}) z < 0$. Indeed, we have

$$\begin{aligned} z^T \nabla^2 f(\bar{x}, \bar{y}) z &= \frac{1}{1 + \|C^{-1}\bar{x}\|_2^2} \cdot \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}^T (L(\bar{x}) + \eta I_{n+1}) \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix} \\ &= \frac{1}{1 + \|C^{-1}\bar{x}\|_2^2} \cdot \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}^T \begin{bmatrix} C + \eta I_n & \bar{x} \\ \bar{x}^T & \mu + \frac{1}{4} + \eta \end{bmatrix} \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix} \\ &= \frac{1}{1 + \|C^{-1}\bar{x}\|_2^2} \left(\mu + \frac{1}{4} - \bar{x}^T C^{-1} \bar{x} + \eta(1 + \|C^{-1}\bar{x}\|_2^2) \right) \\ &= \frac{\mu + \frac{1}{4} - \bar{x}^T C^{-1} \bar{x}}{1 + \|C^{-1}\bar{x}\|_2^2} + \eta. \end{aligned} \quad (9)$$

To show that this expression is negative, we upper bound the numerator and the denominator of its first summand. Using (3), we can write

$$\begin{aligned}
\|C^{-1}\bar{x}\|_2 &\leq \frac{(n+1)^3}{4} \sum_{k=0}^{\infty} \frac{\|A^k\bar{x}\|_2}{(n+1)^{3k}} \\
&\leq \frac{(n+1)^3}{4} \sum_{k=0}^{\infty} \frac{\|A\|_2^k \cdot \|\bar{x}\|_2}{(n+1)^{3k}} \\
&\leq \frac{\sqrt{n}(n+1)^3}{4} \sum_{k=0}^{\infty} \frac{1}{(n+1)^{2k}} \\
&\leq \frac{(n+1)^4}{4\sqrt{n}}
\end{aligned}$$

where, much as was done in the proof of Lemma 2.5, the first inequality uses the triangle inequality, the second the definition and the submultiplicativity property of the matrix 2-norm, the third the facts that $\|\bar{x}\|_2 \leq \sqrt{n}\|\bar{x}\|_{\infty} \leq \sqrt{n}$ and $\|A\|_2 \leq n+1$, and the last the formula for the sum of a geometric series and the fact that $\frac{n+1}{n+2} \leq 1$. Hence:

$$1 + \|C^{-1}\bar{x}\|_2^2 = 1 + \frac{(n+1)^8}{16n} \leq 1 + \frac{(2n)^8}{16n} \leq 1 + 16n^7.$$

Combining this with (9) and the fact that

$$\mu + \frac{1}{4} - \bar{x}^T C^{-1} \bar{x} \leq -\frac{1}{2},$$

we get

$$z^T \nabla^2 f(\bar{x}, \bar{y}) z \leq -\frac{1}{2(1+16n^7)} + \eta.$$

Replacing η by its expression in (8), we see that $z^T \nabla^2 f(\bar{x}, \bar{y}) z \leq -\frac{1}{4} \cdot \frac{1}{1+16n^7} < 0$. Hence $\nabla^2 f(x, y)$ is not positive semidefinite over B , and therefore $f(x, y)$ is not convex over B .

Suppose now that $L(x) \succeq 0$ for all $x \in [-1, 1]^n$. We prove that $f(x, y)$ is convex over B by showing that $\nabla^2 f(x, y) \succeq 0$ for all $(x, y) \in B$. As $\alpha > 0$, this is equivalent to showing, using the Schur complement, that

$$L(x) + \eta I_{n+1} - \frac{1}{4\alpha} H(y)^T H(y) \succeq 0, \text{ for all } (x, y) \in B.$$

We prove that $\eta I_{n+1} - \frac{1}{4\alpha} H(y)^T H(y) \succeq 0$ for all $y \in [-1, 1]^{n+1}$, and as $L(x) \succeq 0$ for any $x \in [-1, 1]^n$, the claim would follow. From (7), we deduce the following expression for the symmetric $(n+1) \times (n+1)$ matrix $H(y)^T H(y)$:

$$H(y)^T H(y) = 4 \begin{bmatrix} y_{n+1}^2 I_n & y_{n+1} \cdot y \\ y_{n+1} \cdot y^T & \sum_{i=1}^n y_i^2 \end{bmatrix}.$$

By Gershgorin's circle theorem [12], for all $y \in [-1, 1]^{n+1}$, we can upper bound the largest eigenvalue of this matrix as follows:

$$\begin{aligned}
\lambda_{\max}(H(y)^T H(y)) &\leq 4 \max_{y \in [-1, 1]^{n+1}} \left\{ \max_{i=1, \dots, n} \{y_{n+1}^2 + |y_i y_{n+1}|\}, \sum_{i=1}^n y_i^2 + \sum_{i=1}^n |y_{n+1} y_i| \right\} \\
&\leq 4 \max\{2, 2n\} = 8n.
\end{aligned}$$

Using this bound and replacing α and η by their expressions in (8), we conclude that

$$\eta I_{n+1} - \frac{1}{4\alpha} H(y)^T H(y) \succeq \left(\eta - \frac{2n}{\alpha} \right) I_{n+1} = \frac{1}{8(1 + 16n^7)} I_{n+1} \succeq 0.$$

□

We end with the observation that Theorem 2.3 immediately classifies the complexity of detecting convexity of a polynomial of degree d over a box, for any integer $d \geq 1$.

Proposition 2.7. *The problem of testing whether a polynomial f of degree d is convex over a box B can be solved in polynomial time for $d = 1$ and $d = 2$ and is strongly NP-hard for any fixed integer $d \geq 3$.*

Proof. When $d = 1$, f is always globally convex. Hence the problem of testing whether f is convex over a box B can trivially be solved in polynomial time (by answering ‘yes’ to all instances). When $d = 2$, the Hessian of f is constant, and so testing convexity of f over B is equivalent to testing whether its (constant) Hessian matrix $\nabla^2 f$ is positive semidefinite. This can be done in polynomial time, e.g. by performing Gaussian pivot steps along the main diagonal of $\nabla^2 f$ [18], or by computing the characteristic polynomial of $\nabla^2 f$ exactly and then checking that the signs of its coefficients alternate [15, p. 403].

For $d = 3$, Theorem 2.3 establishes the claim. For $d \geq 4$, we give a reduction from the problem of testing convexity of a cubic polynomial $g(x) := g(x_1, \dots, x_n)$ over a full-dimensional⁴ box $\tilde{B} \subseteq \mathbb{R}^n$. Given such g, \tilde{B} , let

$$f(x_1, \dots, x_n, x_{n+1}) := g(x_1, \dots, x_n) + x_{n+1}^d,$$

and $B := \tilde{B} \times [0, 1]$. We have

$$\nabla^2 f(x, x_{n+1}) = \begin{bmatrix} \nabla^2 g(x) & 0 \\ 0 & d(d-1) \cdot x_{n+1}^{d-2} \end{bmatrix},$$

and hence $\nabla^2 f(x, x_{n+1}) \succeq 0$ over B if and only if $\nabla^2 g(x) \succeq 0$ over \tilde{B} (as $d(d-1)x_{n+1}^{d-2}$ is nonnegative over $[0, 1]$). It follows that testing convexity of f over B is strongly NP-hard. □

To conclude, we would like to emphasize that our result should not discourage researchers from seeking algorithms for testing convexity of polynomials over a box. It should simply inform them that a search for a pseudo-polynomial time algorithm that works on all cubic polynomials is hopeless (unless P=NP). It is possible however that such an algorithm exists if one restricts one’s attention to carefully-chosen subsets of cubic polynomials.

References

- [1] Gurobi optimizer reference manual. URL: <http://www.gurobi.com>, 2012.
- [2] A. A. Ahmadi and R. M. Jungers. SOS-convex Lyapunov functions and stability of difference inclusions. Available at *arXiv:1803.02070*, 2018.
- [3] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical Programming*, 137(1-2):453–476, 2013.

⁴Note that the proof of Theorem 2.3 established strong NP-hardness of this problem as the box that arose in the proof was the unit hypercube, which is full-dimensional.

- [4] A. Ben-Tal and A. Nemirovski. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM Journal on Optimization*, 12(3):811–833, 2002.
- [5] A. Ben-Tal, A. Nemirovski, and C. Roos. Extended matrix cube theorems with applications to μ -theory in control. *Mathematics of Operations Research*, 28(3):497–523, 2003.
- [6] H. Benson. Concave minimization: theory, applications and algorithms. In R. Horst and P. Pardalos, editors, *Handbook of Global Optimization*, pages 43–148. Springer, Boston, MA, 1995.
- [7] D. P. Bertsekas. *Convex Optimization Theory*. Athena Scientific Belmont, 2009.
- [8] V. Blondel and J. N. Tsitsiklis. NP-hardness of some linear control design problems. *SIAM Journal on Control and Optimization*, 35(6):2118–2127, 1997.
- [9] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [10] G. Chesi and Y. S. Hung. Establishing convexity of polynomial Lyapunov functions and their sublevel sets. *IEEE Transactions on Automatic Control*, 53(10):2431–2436, 2008.
- [11] M. R. Garey and D. S. Johnson. *Computers and Intractability*. W. H. Freeman and Co., San Francisco, 1979.
- [12] S. A. Gershgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na*, (6):749–754, 1931.
- [13] M. Grant and S. Boyd. CVX: MATLAB software for disciplined convex programming, version 1.21. <http://cvxr.com/cvx>, May 2010.
- [14] L. A. Hannah and D. B. Dunson. Multivariate convex regression with adaptive partitioning. *The Journal of Machine Learning Research*, 14(1):3261–3294, 2013.
- [15] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1995.
- [16] A. Khajavirad and N. V. Sahinidis. A hybrid LP/NLP paradigm for global optimization relaxations. *Mathematical Programming Computation*, pages 1–39, 2017.
- [17] E. Lim and P. W. Glynn. Consistency of multidimensional convex regression. *Operations Research*, 60(1):196–208, 2012.
- [18] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.
- [19] A. Nemirovskii. Several NP-hard problems arising in robust stability analysis. *Mathematics of Control, Signals, and Systems*, 6(2):99–105, 1993.
- [20] P. M. Pardalos and S. A. Vavasis. Open questions in complexity theory for numerical optimization. *Mathematical Programming*, 57(1-3):337–339, 1992.
- [21] N. Sahinidis. BARON: A general purpose global optimization software package. *Journal of Global Optimization*, 8(2):201–205, 1996.
- [22] M. Sipser. *Introduction to the Theory of Computation*, volume 2. Thomson Course Technology Boston, 2006.
- [23] M. Xu, M. Chen, and J. Lafferty. Faithful variable screening for high-dimensional convex regression. *The Annals of Statistics*, 44(6):2624–2660, 2016.