On the complexity of detecting convexity over a box

Amir Ali Ahmadi[1] · **Georgina Hall**[2]

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., a 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.

✉ Georgina Hall
  georgina.hall@insead.edu

  Amir Ali Ahmadi
  a_a_a@princeton.edu

1   Department of Operations Research and Financial Engineering, Princeton University, Princeton, USA

2   Decision Sciences, INSEAD, Boulevard de Constance, Fontainebleau, France

1 Introduction

In a relatively recent paper [2], 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 [25]. 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 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 [1,9]. 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* [17,21,28] 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 that 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 [2], 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 [20,26], CVX [14], Gurobi [15], and DR.AMPL/COCONUT [11] is restricted to quadratic functions, or is replaced with the task of checking some sufficient or necessary conditions for convexity that are more tractable. These sufficient conditions typically require that the function in question be the output of certain convexity-preserving operations applied to an initial set of convex functions; see, e.g., [8, Section 3.2] or [11, Section 3]. The necessary conditions, which can be used to disprove convexity, often involve a search for a direction of negative curvature at some sample points in the region of interest; see, e.g., [11, Section 5] or [10]. There are also algebraic sufficient conditions for convexity over a box that can be efficiently checked by sum of squares optimization; see, e.g., [22, Section 5], [16, Chapter 7]. These tests are mainly applicable to polynomial functions and to our knowledge are not currently implemented in popular software.

Assuming P $\neq$ NP, the implication of our Theorem 2.3 below for convexity detection is as follows. Suppose an optimization solver is endowed with a pseudo-polynomial time algorithm for proving convexity of a function over a box and a pseudo-polynomial time algorithm for disproving it. Then, regardless of how these algorithms are designed, there are cubic polynomials for which convexity can neither be proved nor disproved by this solver. In fact, as the number of variables grows, this solver will fail to prove or disprove convexity for a "significant portion" of cubic polynomials; see Corollary 2.2 of [18] for a precise complexity-theoretic statement.

The proof of Theorem 2.3 has two parts, the first of which relies heavily on a clever gadget by Nemirovski [24]. 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 $l_1, \ldots, l_n$, $u_1, \ldots, u_n$, with $l_i \leq u_i$ for $i = 1, \ldots, n$, a *box* $B \subseteq \mathbb{R}^n$ is a set of the form

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

Definition 2.2 A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex over a convex 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., [27]) 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 $l_1, \ldots, l_n, u_1, \ldots, u_n$ 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 [12]. This distinction only occurs in

problems whose instances involve numerical data[1] 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, following the conventions used in [12], define a Max (resp. Length) function which maps any instance I of a decision problem to a nonnegative integer Max(I) (resp. 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 Max(I) is upperbounded by p(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 [12, 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 [12, 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.*

Remark 2.1 In fact, testing convexity of a cubic polynomial f over a box B is a co-NP-complete problem. To see membership in co-NP, first observe that since the entires of $\nabla^2 f$ are affine, positive semidefiniteness of $\nabla^2 f$ at the extreme points of B implies its positive semidefiniteness over all of B. Hence, if f is not convex over B, one can always present one of the extreme points of B as a certificate of nonconvexity. This certificate has polynomial size since its entries consist of the rational numbers l_i, u_i which are part of the input. Moreover, once such a candidate extreme point is presented, one can conclude in polynomial time that the Hessian matrix evaluated at it is not positive semidefinite; see e.g. the arguments in the first paragraph of the proof of Proposition 2.7.

The proof of Theorem 2.3 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 (weakly) NP-hard by Nemirovski [24]. 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

---

[1] Note that the instances of not all of NP-hard problems involve numerical data; consider, e.g., SATISFIABILITY [12]. For such problems, NP-hardness is by default in the strong sense.

set S with nonempty interior if and only if $\nabla^2 f(x) \succeq 0$, $\forall x \in S$; see, e.g., [6, Chapter 1].[2] 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 [24], which we use here, from weak to strong (Theorem 2.4). Indeed, the reduction provided in [24] is from the PARTITION problem which is only weakly NP-hard [12], 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).

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 $i, j, k \in \{1, \ldots, n\}$, 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 $L(x)$.

Before we proceed, we briefly contrast the proof of the main result of this paper with the proof of the main result (Theorem 2.1) in [2] as they have some commonalities. Both proofs involve a reduction where it is shown that an $m \times m$ symmetric polynomial matrix $P(x)$ in n variables is positive semidefinite over a region R if and only if the Hessian matrix of some polynomial f in a higher number of variables is positive semidefinite over a related region R'. In [2], $m = n$, the entries of $P(x)$ are homogeneous quadratic polynomials, R and R' are the entire Euclidean space, and the number of variables in f is $2n$. In our proof, $m = n + 1$, the entries of $P(x)$ are affine polynomials, R and R' are boxes, and the number of variables in f is $2n + 1$. In [2], the proof starts from the fact that testing positive semidefiniteness of a quadratic polynomial matrix on the unit sphere is strongly NP-hard. This fact is not helpful for us since quadratic polynomial Hessian matrices are second derivatives of quartic poly-

---

[2] Here, the notation $M \succeq 0$ is used to denote that a symmetric matrix M is positive semidefinite, i.e., has nonnegative eigenvalues. We will also use the notation $A \succeq B$ for two symmetric matrices A and B to denote that $A - B \succeq 0$.

nomials while our goal is to end up with cubic polynomials. Hence, we have to first establish strong NP-hardness of testing positive semidefiniteness of affine polynomial matrices (over boxes) as mentioned before. Another difference in the two proofs is in the construction of the polynomial f which as compared to [2] necessitates a more careful balancing of the coefficients of the terms involved. We also note that the reduction in [2] turns *any* homogeneous quadratic polynomial matrix into a valid Hessian matrix while preserving positive semidefiniteness. By contrast, our reduction exploits the specific structure of the affine polynomial matrix that emerges from Theorem 2.4 below; see $L(x)$ in (2).

Theorem 2.4 (see the proof of Proposition 2.1 in [24] 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*

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

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

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

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

Remark 2.2 As the matrix C is positive definite (see the proof of Lemma 2.5), by using the Schur complement [8, Appendix C.4], the constraint in (2) is equivalent to having $x^T C^{-1} x \leq \mu + \frac{1}{4}$ for all x with $||x||_\infty \leq 1$.

To prove Theorem 2.4, 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*

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

Proof First, note that C is by construction strictly diagonally dominant, and hence, by Gershgorin's circle theorem [13], positive definite. This implies that C^{-1} is well defined. Furthermore, we have

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

where the second equality follows from the identity $(I_n - M)^{-1} = \sum_{k=0}^{\infty} M^k$, which holds if for some matrix norm $|| \cdot ||$ we have $||M|| < 1$; see, e.g., [19, Section 5.6].

Here, for example, it is easy to see that the maximum absolute row sum norm of $\frac{-A}{(n+1)^3}$ is strictly less than one. Indeed, each row of A contains at most n ones, and so the maximum absolute row sum norm of $-A/(n+1)^3$ is less than $1/(n+1)^2 < 1$. Using (3), we have

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

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

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

$$\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{||x||_2 \cdot ||(-A)^k x||_2}{(n+1)^{3k}}$$

$$\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{||x||_2 \cdot ||A||_2^k \cdot ||x||_2}{(n+1)^{3k}}$$

$$\leq \frac{(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{n||A||_2^k}{(n+1)^{3k}}$$

$$\leq \frac{n(n+1)^3}{4} \sum_{k=2}^{\infty} \frac{1}{(n+1)^{2k}},$$

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

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

$$= \frac{1}{4} \cdot \frac{n+1}{n+2}$$

$$\le \frac{1}{4}.$$

□

Proof of Theorem 2.4 We give a reduction from the SIMPLE MAX-CUT problem, which is the following decision problem: Given a simple[3] graph $G = (V, E)$ and a positive integer $s \le |V|^2$, 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 strongly NP-hard [12, p. 210].

Consider an instance of SIMPLE MAX-CUT given by a graph $G = (V, E)$ and a positive integer $s \le |V|^2$. 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 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 $||x||_\infty \le 1$ 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

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

for any instance I of the problem, the problem is automatically strongly NP-hard; see e.g. [12, 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

$$p^* := \min_x \frac{1}{4} x^T A x$$
$$\text{s.t.} \quad x \in \{-1, 1\}^n,$$

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

$$\max_x \frac{1}{4} x^T ((n+1)^3 I_n - A) x$$
$$\text{s.t.} \quad x \in \{-1, 1\}^n$$

---

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

being less than or equal to $\frac{n(n+1)^3}{4} + k - 1 - \frac{1}{4}e^T Ae = \mu$. 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., [5, Property 12]), the previous statement is equivalent to the optimal value of

$$\max_x \frac{1}{4}x^T ((n+1)^3 I_n - A)x$$
$$\text{s.t.} \quad x \in [-1, 1]^n$$

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 $\frac{1}{4}x^T ((n+1)^3 I_n - Ax)x \leq \mu$ 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 (see Remark 2.2), 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 $\frac{1}{4}(e^T Ae - \hat{x}^T A\hat{x})$. Hence, as $k = s$, we have

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

which is equivalent to $\frac{1}{4}n(n+1)^3 - \frac{1}{4}\hat{x}^T A\hat{x} \geq \frac{1}{4}n(n+1)^3 - \frac{1}{4}e^T Ae + k = \mu + 1$. 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 $L(\hat{x}) \not\succeq 0$.

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

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 [3,4,7]. This statement was first proven in [24], 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 $\hat{m}_{ij}, \bar{m}_{ij}, i = 1, \ldots, N, j = 1, \ldots, N$, with $\hat{m}_{ij} \leq \bar{m}_{ij}, \hat{m}_{ij} = \hat{m}_{ji}$, and $\bar{m}_{ij} = \bar{m}_{ji}$ for $i = 1, \ldots, N, j = 1, \ldots, N$, 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 $i = 1, \ldots, N, j = 1, \ldots, N$, 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 $i = 1, \ldots, n, j = 1, \ldots, n, \hat{m}_{i(n+1)} = \hat{m}_{(n+1)i} = -1$ and $\bar{m}_{i(n+1)} = \bar{m}_{(n+1)i} = 1$ for $i = 1, \ldots, n$, 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 $i = 1, \ldots, n + 1, j = 1, \ldots, n + 1$, 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 Length(I) of order n^2 and Max(I) of order n^6, and hence Max(I) is upper bounded by a polynomial in Length(I). $\qquad\square$

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, for $x \in \mathbb{R}^n$, 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 $y := (y_1, \ldots, y_{n+1})^T$ 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 $y^T L(x) y$:

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

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

$$H(y) = \begin{bmatrix} 2y_{n+1} & 0 & \ldots & 0 & 2y_1 \\ 0 & \ddots & 0 & 0 & 2y_2 \\ 0 & 0 & \ddots & 0 & \vdots \\ 0 & & \ldots & 0 & 2y_{n+1} & 2y_n \end{bmatrix}. \qquad (4)$$

Let

$$\alpha := 16 \cdot n(1 + 16n^7),$$

$$\text{and } \eta := \frac{1}{4} \cdot \frac{1}{1 + 16n^7}. \tag{5}$$

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

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

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 Length(I) of order n^3 (the number of coefficients of f) and Max(I) of order n^8 (which is the magnitude of α, which dominates all other coefficients). Hence, Max(I) is upper bounded by a polynomial in 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 for all $x \in [-1, 1]^n$. 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

$$z^T \nabla^2 f(\bar{x}, \bar{y}) z = \frac{1}{1 + ||C^{-1}\bar{x}||_2^2} \cdot \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}^T (L(\bar{x}) + \eta I_{n+1}) \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}$$

$$= \frac{1}{1 + ||C^{-1}\bar{x}||_2^2} \cdot \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}^T \begin{bmatrix} C + \eta I_n & \bar{x} \\ \bar{x}^T & \mu + \frac{1}{4} + \eta \end{bmatrix} \begin{bmatrix} -C^{-1}\bar{x} \\ 1 \end{bmatrix}$$

$$= \frac{1}{1 + ||C^{-1}\bar{x}||_2^2} \left(\mu + \frac{1}{4} - \bar{x}^T C^{-1}\bar{x} + \eta(1 + ||C^{-1}\bar{x}||_2^2) \right)$$

$$= \frac{\mu + \frac{1}{4} - \bar{x}^T C^{-1}\bar{x}}{1 + ||C^{-1}\bar{x}||_2^2} + \eta. \tag{6}$$

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:

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

Combining this with (6) 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 (5), 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 (see Remark 2.2), 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 $\eta I_{n+1} - \frac{1}{\alpha} H(y)^T H(y) \succeq 0$ 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 (4), 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 [13], for any $y \in [-1, 1]^{n+1}$, we can upper bound the largest eigenvalue of this matrix as follows:

$$\lambda_{\max}(H(y)^T H(y)) \leq 4 \max \left\{ \max_{i=1,\dots,n} \{y_{n+1}^2 + |y_i y_{n+1}|\}, \sum_{i=1}^n y_i^2 + \sum_{i=1}^n |y_{n+1} y_i| \right\}$$

$$\leq 4 \max\{2, 2n\} = 8n.$$

Using this bound and replacing α and η by their expressions in (5), we conclude that

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

$\square$

Remark 2.3 As observed by a referee, a careful analysis of the proof of Theorem 2.3 shows that it is also strongly NP-hard to check convexity of a cubic polynomial over an open hyper-rectangle (i.e., a set of the type $\{x \in \mathbb{R}^n \mid l_i < x_i < u_i, i = 1, \dots, n\}$).

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 $\nabla^2 f$ [23], or by computing the characteristic polynomial of $\nabla^2 f$ exactly and then checking that the signs of its coefficients alternate [19, 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-dimensional[4] box $\tilde{B} \subseteq \mathbb{R}^n$. Given such g, $\tilde{B}$, let

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

---

[4] Note 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.

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. $\qquad \square$

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. While we have shown that a search for a pseudo-polynomial time algorithm that works on all cubic polynomials is hopeless unless $P = NP$, it remains valuable to characterize interesting classes of polynomials on which convexity detection can be done efficiently.

References

1. Ahmadi, A.A., Jungers, R.M.: SOS-convex Lyapunov functions and stability of difference inclusions. arXiv:1803.02070 (2018)
2. Ahmadi, A.A., Olshevsky, A., Parrilo, P.A., Tsitsiklis, J.N.: NP-hardness of deciding convexity of quartic polynomials and related problems. Math. Program. **137**(1–2), 453–476 (2013)
3. Ben-Tal, A., Nemirovski, A.: On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. SIAM J. Optim. **12**(3), 811–833 (2002)
4. Ben-Tal, A., Nemirovski, A., Roos, C.: Extended matrix cube theorems with applications to μ-theory in control. Math. Oper. Res. **28**(3), 497–523 (2003)
5. Benson, H.: Concave minimization: theory, applications and algorithms. In: Horst, R., Pardalos, P. (eds.) Handbook of Global Optimization, pp. 43–148. Springer, Boston (1995)
6. Bertsekas, D.P.: Convex Optimization Theory. Athena Scientific Belmont, Belmont (2009)
7. Blondel, V., Tsitsiklis, J.N.: NP-hardness of some linear control design problems. SIAM J. Control Optim. **35**(6), 2118–2127 (1997)
8. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
9. Chesi, G., Hung, Y.S.: Establishing convexity of polynomial Lyapunov functions and their sublevel sets. IEEE Trans. Autom. Control **53**(10), 2431–2436 (2008)
10. Chinneck, J.W.: Analyzing mathematical programs using MProbe. Ann. Oper. Res. **104**(1–4), 33–48 (2001)
11. Fourer, R., Maheshwari, C., Neumaier, A., Orban, D., Schichl, H.: Convexity and concavity detection in computational graphs: tree walks for convexity assessment. INFORMS J. Comput. **22**(1), 26–43 (2010)
12. Garey, M.R., Johnson, D.S.: Computers and Intractability. W. H. Freeman and Co., San Francisco (1979)
13. Gershgorin, S.A.: Uber die Abgrenzung der Eigenwerte einer Matrix. Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathematiques et naturelles **6**, 749–754 (1931)
14. Grant, M., Boyd, S.: CVX: MATLAB software for disciplined convex programming, version 1.21. http://cvxr.com/cvx (2010)
15. Gurobi optimizer reference manual. http://www.gurobi.com (2012)
16. Hall, G.: Optimization over nonnegative and convex polynomials with and without semidefinite programming. Ph.D. thesis, Princeton University. arXiv:1806.06996 (2018)

17. Hannah, L.A., Dunson, D.B.: Multivariate convex regression with adaptive partitioning. J. Mach. Learn. Res. **14**(1), 3261–3294 (2013)
18. Hemaspaandra, L.A., Williams, R.: SIGACT News-Complexity Theory Column 76: an atypical survey of typical-case heuristic algorithms. ACM SIGACT News **43**(4), 70–89 (2012)
19. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1995)
20. Khajavirad, A., Sahinidis, N.V.: A hybrid LP/NLP paradigm for global optimization relaxations. Math. Program. Comput. **10**, 1–39 (2017)
21. Lim, E., Glynn, P.W.: Consistency of multidimensional convex regression. Oper. Res. **60**(1), 196–208 (2012)
22. Magnani, A., Lall, S., Boyd, S.: Tractable fitting with convex polynomials via sum-of-squares. In: Proceedings of the 44th IEEE Conference on Decision and Control, pp. 1672–1677 (2005)
23. Murty, K.G., Kabadi, S.N.: Some NP-complete problems in quadratic and nonlinear programming. Math. Program. **39**, 117–129 (1987)
24. Nemirovski, A.: Several NP-hard problems arising in robust stability analysis. Math. Control Signals Syst. **6**(2), 99–105 (1993)
25. Pardalos, P.M., Vavasis, S.A.: Open questions in complexity theory for numerical optimization. Math. Program. **57**(1–3), 337–339 (1992)
26. Sahinidis, N.: BARON: a general purpose global optimization software package. J. Glob. Optim. **8**(2), 201–205 (1996)
27. Sipser, M.: Introduction to the Theory of Computation, vol. 2. Thomson Course Technology, Boston (2006)
28. Xu, M., Chen, M., Lafferty, J.: Faithful variable screening for high-dimensional convex regression. Ann. Stat. **44**(6), 2624–2660 (2016)