Certification of Real Inequalities

Templates and Sums of Squares

Xavier Allamigeon · Stéphane Gaubert ·
Victor Magron · Benjamin Werner

Abstract We consider the problem of certifying lower bounds for real-valued multivariate transcendental functions. The functions we are dealing with are nonlinear and involve semialgebraic operations as well as some transcendental functions like cos, arctan, exp, etc. Our general framework is to use different approximation methods to relax the original problem into polynomial optimization problems, which we solve by sparse sums of squares relaxations. In particular, we combine the ideas of the maxplus approximations (originally introduced in optimal control) and of the linear templates (originally introduced in static analysis by abstract interpretation). The nonlinear templates control the complexity of the semialgebraic relaxations at the price of coarsening the maxplus approximations. In that way, we arrive at a new - template based - certified global optimization method, which exploits both the precision of sums of squares relaxations and the scalability of abstraction methods. We analyze the performance of the method on problems from the global optimization literature, as well as medium-size inequalities issued from the Flyspeck project.

X. Allamigeon
INRIA and CMAP École Polytechnique
Tel.: +33 (0)1 69 33 46 32
E-mail: xavier.allamigeon@inria.fr

S. Gaubert
INRIA and CMAP École Polytechnique
Tel.: +33 (0)1 69 33 46 13
E-mail: Stephane.Gaubert@inria.fr

V. Magron
INRIA and LIX/CMAP École Polytechnique
Tel.: +33 (0)1 69 35 69 85
E-mail: victor.magron@inria.fr

B. Werner
LIX École Polytechnique
Tel.: +33 (0)1 69 33 41 41
E-mail: benjamin.werner@polytechnique.edu

Quadratic Cuts · Maxplus Approximation · Template Method · Certified Global Optimization

1 Introduction

1.1 Certification of Nonlinear Inequalities

Numerous problems coming from different fields of mathematics (like combinatorics, geometry or group theory) have led to computer assisted proofs. One famous example is the proof of the Kepler conjecture, proved by Thomas Hales [17, 18]. Recent efforts have been made to complete the formal verification of this conjecture. In particular, extensive computation are required to certify hundreds of nonlinear inequalities. We will often refer to the following inequality taken from Hales' proof:

Example 1 (Lemma$_{9922699028}$ Flyspeck) Let K, $\Delta\mathbf{x}$, l, t and f be defined as follows:

$$
\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 ~, \\
\Delta\mathbf{x} &:= x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad - x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6 ~, \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad + 0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) ~, \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4 x_1 \Delta\mathbf{x}}} ~, \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) ~.
\end{aligned}
$$

Then, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geqslant 0$.

Note that the inequality of Example 1 would be much simpler to check if l was a constant (rather than a function of $\mathbf{x}$). Indeed, semialgebraic optimization methods would provide precise lower and upper bounds for the argument of arctan. Then we could conclude by monotonicity of arctan using interval arithmetic. Here, both l and t depend on $\mathbf{x}$. Hence, by using interval arithmetic addition (without any domain subdivision) on the sum $l + t$, which ignores the correlation between the argument of arctan and the function l, we only obtain a coarse lower bound (equal to -0.87, see Example 3 for details); too coarse to assert the inequality . A standard way to improve this bound consists in subdividing the initial box (i.e. the Cartesian product of closed intervals) K and performing interval arithmetic on smaller boxes. However, this approach suffers from the so called *curse of dimensionality*. Therefore, it is desirable to develop alternative certified global optimization methods, applicable to a wide class of problems involving semialgebraic and transcendental functions.

Moreover, the nonlinear inequalities of Flyspeck are challenging for numerical solvers for two reasons. First, they involve a medium-scale number of variables (6∼10). Then, they are essentially *tight*. For instance, the function f involved in Example 1 has a nonnegative infimum which is less than 10^{-3}. The tightness of the inequalities to be certified is actually a frequent feature in mathematical proofs. Hence, we will pay a special attention in the present work to *scalability* and *numerical precision* issues to provide certified bounds for global optimization. This is called

informal certification as one obtains numerical certificates (e.g. sums-of-squares certificates) that can be in turn *formally* checked inside a proof assistant, such as Coq. The practical difficulties related to formalization (e.g. polynomial arithmetic implementation, appropriate certificate data-structures, formal proofs of approximations for nonlinear functions) are specifically addressed in [7], as a further contribution of the authors.

1.2 Nonlinear Global Optimization Problems

Let $\langle \mathcal{D} \rangle^{\mathrm{sa}}$ be the set of functions obtained by composing (multivariate) semialgebraic functions with special functions taken from a *dictionary $\mathcal{D}$*. We will typically include in $\mathcal{D}$ the usual functions tan, arctan, cos, arccos, sin, arcsin, exp, log, $(\cdot)^r$ with $r \in \mathbb{R} \setminus \{0\}$. As we allow the composition with semialgebraic functions in our setting, elementary functions like $+, -, \times, /, |\cdot|, \sup(\cdot, \cdot), \inf(\cdot, \cdot)$ are of course covered. Actually, we shall see that some of the present results remain valid if the dictionary includes semiconvex[1] or semiconcave functions with effective lower and upper bounds on the Hessian. More details about semiconcave functions can be found in [10].

Given $f, f_1, \ldots, f_p \in \langle \mathcal{D} \rangle^{\mathrm{sa}}$, we will address the following global optimization problem:

$$\inf_{\mathbf{x} \in \mathbb{R}^n} \quad f(\mathbf{x}) \ , \tag{1.1}$$
$$\mathrm{s.t.} \quad f_1(\mathbf{x}) \geqslant 0, \ldots, f_p(\mathbf{x}) \geqslant 0 \ .$$

The inequalities issued from Flyspeck actually deal with special cases of computation of a certified lower bound for a real-valued multivariate function $f : \mathbb{R}^n \to \mathbb{R}$ over a compact semialgebraic set $K \subset \mathbb{R}^n$. Checking these inequalities boils down to automatically provide lower bounds for the following instance of Problem (1.1):

$$f^* := \inf_{\mathbf{x} \in K} f(\mathbf{x}) \ , \tag{1.2}$$

We shall also search for *certificates* to assess that:

$$\forall \mathbf{x} \in K, f(\mathbf{x}) \geqslant 0 \ . \tag{1.3}$$

A well studied case is when $\mathcal{D}$ is reduced to the identity map $\{Id\}$. Then, $f = f_{\mathrm{sa}}$ belongs to the algebra $\mathcal{A}$ of semialgebraic functions (extension of multivariate polynomials with arbitrary compositions of $(\cdot)^p, (\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), |\cdot|, +, -, \times, /, \sup(\cdot, \cdot), \inf(\cdot, \cdot)$, where $\mathbb{N}_0$ stands for the set of positive integers) and Problem (1.1) specializes to the semialgebraic optimization problem:

$$f_{\mathrm{sa}}^* := \inf_{\mathbf{x} \in K} f_{\mathrm{sa}}(\mathbf{x}) \ . \tag{1.4}$$

Another important sub-case is Polynomial Optimization Problems (POP), when $f = f_{\mathrm{pop}}$ is a multivariate polynomial and $K = K_{\mathrm{pop}}$ is given by finitely many polynomial inequalities. Thus, Problem (1.4) becomes:

$$f_{\mathrm{pop}}^* := \inf_{\mathbf{x} \in K_{\mathrm{pop}}} f_{\mathrm{pop}}(\mathbf{x}) \ . \tag{1.5}$$

---

[1] Recall that for $\gamma \geqslant 0$, a function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be γ-*semiconvex* if the function $\mathbf{x} \mapsto \phi(\mathbf{x}) + \frac{\gamma}{2} \|\mathbf{x}\|_2^2$ is convex.

We shall see that the presented methods also provide certified lower bounds (possibly coarse), for optimization problems which are hard to solve by traditional POP techniques. Such problems have a relatively large number of variables (10∼100) or are polynomial inequalities of a moderate degree. For illustration purposes, we consider the following running example coming from the global optimization literature.

Example 2 (Modified Schwefel Problem 43 from Appendix B in [4])

$$\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n-1} (x_i + \epsilon x_{i+1}) \sin(\sqrt{x_i}) \ ,$$

where ϵ is a fixed parameter in $\{0, 1\}$. In the original problem, $\epsilon = 0$, *i.e.* the objective function f is the sum of independent functions involving a single variable. This property may be exploited by a global optimization solver by reducing it to the problem $\min_{x \in [1,500]} x \sin(\sqrt{x})$. Hence, we also consider a modified version of this problem with $\epsilon = 1$.

1.3 Certified Global Optimization in the Literature

A common idea to handle Problem (1.2) is to first approximate f by multivariate polynomials and then obtain a lower bound of the resulting approximation by polynomial optimization techniques.

Computing lower bounds in constrained POP (see Problem(1.5)) is already a difficult problem, which has received much attention. Sums of squares (SOS) relaxation based methods, leading to the resolution of semidefinite programs (SDP) have been developed in [24, 34]. They can be applied to the more general class of semialgebraic problems [36]. Moreover, Kojima has developed a sparse refinement of the hierarchy of SOS relaxations (see [40]). This has been implemented in the SPARSEPOP solver. Checking the validity of the lower bound of POP implies being able to control and certify the numerical error, as SDP solvers are typically implemented using floating point arithmetic. Such techniques rely on hybrid symbolic-numeric certification methods, see Peyrl and Parrilo [35] and Kaltofen et al. [21]. They allow one to produce positivity certificates for such POP. Alternative approaches to SOS/SDP are based on Bernstein polynomials [41].

The task is obviously more difficult in presence of transcendental functions. Other methods of choice, not restricted to polynomial systems, include global optimization by interval methods (see e.g. [20]), branch and bound methods with Taylor models [11, 8]. Other methods involve rigorous Chebyshev approximations. An implementation of such approximations is available in the `Sollya` tool [12].

1.4 Contribution

In this paper, we develop a general certification framework, combining methods from semialgebraic programming (SOS certificates, SDP relaxations) and from approximation theory. This includes classical methods like best uniform polynomials and less classical ones like maxplus approximation (inspired by optimal control and static analysis by abstract interpretation).

The present approach exploits both the accuracy of SOS relaxations and the scalability of the approximation and abstraction procedure. This leads to a new method in global optimization, the nonlinear template method. Namely, we alternate steps of semialgebraic approximation for some constituents of the objective function f and semialgebraic optimization. The resulting constrained polynomial optimization problems are solved with sums of squares relaxation from Lasserre hierarchy, by calling a semidefinite solver. In this way, each iteration of the algorithms refines the following inequalities:

$$f^* \geqslant f_{sa}^* \geqslant f_{pop}^* \ , \tag{1.6}$$

where f^* is the optimal value of the original problem, f_{sa}^* the optimal value of its current semialgebraic approximation and f_{pop}^* the optimal value of the SOS relaxation which we solve. Under certain moderate assumptions, the lower bound f_{pop}^* does converge to f^* (see Corollary 2).

The present nonlinear template method is an improved version of the maxplus approximation method originally presented in [6]. By comparison, the new ingredient is the introduction of the template technique (approximating projections of the feasible sets), leading to an increase in scalability. This technique is an abstraction method, which is inspired by the linear template of Sankaranarayanan, Sipma and Manna in static analysis [38], their nonlinear extensions by Adjé et al. [1]. As discussed below, it is closely related to the maxplus basis methods, although the methods differ in the way they propagate approximations.

In the present application, templates are used both to approximate transcendental functions, and to produce coarser but still tractable relaxations when the standard SOS relaxation of the semialgebraic problem is too complex to be handled. As a matter of fact, SOS relaxations are a powerful tool to get tight certified lower bound for semialgebraic optimization problems, but applying them is currently limited to small or medium size problems: their execution time grows exponentially with the relaxation order, which itself grows with the degree of the polynomials involved in the semialgebraic relaxations. The template method allows to reduce these degrees, by approximating certain projections of the feasible set by a moderate number of nonlinear inequalities.

In this article, we present the following approximation schemes:

– **Semialgebraic maxplus templates for multivariate transcendental functions** This method uses maxplus approximation of semiconvex transcendental functions by quadratic functions. The idea of maxplus approximation comes from optimal control: it was originally introduced by Fleming and McEneaney [13] and developed by several authors [3, 30, 29, 39, 14], to represent the value function by a "maxplus linear combination", which is a supremum of certain basis functions, like quadratic polynomials. When applied to the present context, this idea leads to approximate from above and from below every transcendental function appearing in the description of the problem by infima and suprema of finitely many quadratic polynomials. In that way, we are reduced to a converging sequence of semialgebraic problems. A geometrical way to interpret the method is to think of it in terms of "quadratic cuts" quadratic inequalities are successively added to approximate the graph of a transcendental function (Sect. 4.1).
– **Non-convex quadratic templates** Sub-components of the objective function f (resp. its semialgebraic approximations) are replaced by suprema of quadratic polynomials (Sect. 4.2.1).

– **Polynomial under-approximations for semialgebraic functions** Given a degree d and a semialgebraic sub-component f_{sa} of f that involves a large number of lifting variables, we build a hierarchy of polynomial approximations, that converge to the best (for the L_1 norm) degree-d polynomial under-approximation of f_{sa} (Sect. 4.2.2).

The paper is organized as follows. In Sect. 2, we recall the definition and properties of Lasserre relaxations of polynomial problems, together with reformulations by Lasserre and Putinar of semialgebraic problems classes. The maxplus approximation and the nonlinear templates are presented in Sect. 3. In Sect. 4, we describe the nonlinear template optimization algorithm together with the convergence study of the method. The main numerical results are presented in Sect. 5.

2 Application of SOS to Semialgebraic Optimization

Let $\mathbb{R}_d[\mathbf{x}]$ be the vector space of real forms in n variables of degree d and $\mathbb{R}[\mathbf{x}]$ the set of multivariate polynomials in n variables. We also define the cone $\Sigma_d[\mathbf{x}]$ of sums of squares of degree at most $2d$.

2.1 Constrained Polynomial Optimization Problems and SOS

We consider the general constrained polynomial optimization problem (POP):

$$f^*_{\mathrm{pop}} := \inf_{\mathbf{x} \in K_{\mathrm{pop}}} f_{\mathrm{pop}}(\mathbf{x}) \ , \tag{2.1}$$

where $f_{\mathrm{pop}} : \mathbb{R}^n \to \mathbb{R}$ is a d-degree multivariate polynomial, K_{pop} is a compact set defined by polynomials inequalities $g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0$ with $g_j(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ being a real-valued polynomial of degree $w_j, j = 1, \ldots, m$. We call K_{pop} the feasible set of Problem (2.1). Let $g_0 := 1$. We introduce the k-truncated quadratic module $QM_k(K_{\mathrm{pop}}) \subset \mathbb{R}_{2k}[\mathbf{x}]$ associated with $g_1, \cdots, g_m$:

$$QM_k(K_{\mathrm{pop}}) = \Big\{ \sum_{j=0}^m \sigma_j(\mathbf{x}) g_j(\mathbf{x}) : \sigma_j \in \Sigma_{k-\lceil w_j/2 \rceil}[\mathbf{x}] \Big\} \ ,$$

and define the quadratic module $QM(K_{\mathrm{pop}}) := \bigcup_{k \in \mathbb{N}} QM_k(K_{\mathrm{pop}})$.

Definition 1 A quadratic module M is called *archimedean* if $N - \|\mathbf{x}\|_2^2 \in M$ for some $N \in \mathbb{N}_0$.

Let $k \geqslant k_0 := \max(\lceil d/2 \rceil, \max_{0 \leqslant j \leqslant m}\{\lceil w_j/2 \rceil\})$ and consider the hierarchy of semidefinite relaxations $Q_k : \sup\{\mu : f_{\mathrm{pop}}(x) - \mu \in QM_k(K_{\mathrm{pop}}), \ \mu \in \mathbb{R}\}$, with optimal value denoted by $\sup(Q_k)$. The integer k refers to the SOS relaxation order.

Theorem 1 *The sequence of optimal values* $(\sup(Q_k))_{k \geqslant k_0}$ *is non-decreasing. If the quadratic module* $QM(K_{pop})$ *is archimedean, then this sequence converges to* f^*_{pop}.

Proof The proof follows from [24, Theorem 4.2 (a)] as [24, Assumption 4.1] is satisfied when $QM(K_{\mathrm{pop}})$ is archimedean. □

The non-linear inequalities to be proved in the Flyspeck project typically involve a variable $\mathbf{x}$ lying in a box. Thus, Theorem 1 applies when one adds to the definition of K_{pop} the redundant constraint $g(\mathbf{x}) := N - \|\mathbf{x}\|_2^2 \geqslant 0$ for some large enough $N \in \mathbb{N}_0$.

2.2 Semialgebraic Optimization

In this section, we recall how the previous approach can be extended to semialgebraic optimization problems by introducing lifting variables. The set $\mathcal{A}$ of semialgebraic functions $f_{\mathrm{sa}} : K_{\mathrm{sa}} \mapsto \mathbb{R}$ is the algebra generated by finite composition of the operations $+, -, \times, /, \sup, \inf, |\cdot|, (\cdot)^{\frac{1}{p}}\,(p \in \mathbb{N}_0)$ on polynomials, whenever these operations are well-defined (e.g. division by zero never occurs). Let consider the problem

$$f_{\mathrm{sa}}^* = \inf_{\mathbf{x} \in K_{\mathrm{sa}}} f_{\mathrm{sa}}(\mathbf{x}) \;, \tag{2.2}$$

where $K_{\mathrm{sa}} := \{\mathbf{x} \in \mathbb{R}^n \,:\, g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0\}$ is a basic semialgebraic set.

Definition 2 (Basic Semialgebraic Lifting) A semialgebraic function f_{sa} is said to have a basic semialgebraic lifting if there exist $p, s \in \mathbb{N}$, polynomials $h_1, \ldots, h_s \in \mathbb{R}[\mathbf{x}, z_1, \ldots, z_p]$ and a basic semialgebraic set K_{pop} defined by:

$$K_{\mathrm{pop}} := \{(\mathbf{x}, z_1, \ldots, z_p) \in \mathbb{R}^{n+p} : \mathbf{x} \in K_{\mathrm{sa}}, h_1(\mathbf{x}, \mathbf{z}) \geqslant 0, \ldots, h_s(\mathbf{x}, \mathbf{z}) \geqslant 0\} \;,$$

such that the graph of f_{sa} (denoted $\Psi_{f_{\mathrm{sa}}}$) satisfies:

$$\Psi_{f_{\mathrm{sa}}} := \{(\mathbf{x}, f_{\mathrm{sa}}(\mathbf{x})) : \mathbf{x} \in K_{\mathrm{sa}}\} = \{(\mathbf{x}, z_p) : (\mathbf{x}, \mathbf{z}) \in K_{\mathrm{pop}}\} \;.$$

By [25, Lemma 3], every function $f_{\mathrm{sa}} \in \mathcal{A}^2$ has a basic semialgebraic lifting. To ensure that the Archimedean condition is preserved, we add bound constraints over the lifting variables. These bounds are computed by solving semialgebraic optimization sub-problems. All the semialgebraic functions involved in Flyspeck inequalities have a basic semialgebraic lifting.

Example 3 (from Lemma$_{9922699028}$ Flyspeck) Continuing Example 1, we consider the function $f_{\mathrm{sa}} := \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}$ and the set $K_{\mathrm{sa}} := [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2$. The latter can be equivalently rewritten as

$$K_{\mathrm{sa}} := \{\mathbf{x} \in \mathbb{R}^6 : g_1(\mathbf{x}) \geqslant 0, \ldots, g_{12}(\mathbf{x}) \geqslant 0\} \;,$$

where $g_1(\mathbf{x}) := x_1 - 4, g_2(\mathbf{x}) := 6.3504 - x_1, \ldots, g_{11}(\mathbf{x}) := x_6 - 4, g_{12}(\mathbf{x}) := 6.3504 - x_6$.

    We introduce two lifting variables z_1 and z_2, respectively representing the terms $\sqrt{4x_1 \Delta \mathbf{x}}$ and $\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}$.

    We also use a lower bound m_1 of $\inf_{\mathbf{x} \in K_{\mathrm{sa}}} \sqrt{4x_1 \Delta \mathbf{x}}$ and an upper bound M_1 of $\sup_{\mathbf{x} \in K_{\mathrm{sa}}} \sqrt{4x_1 \Delta \mathbf{x}}$ which can be both computed by solving auxiliary subproblems.

    Now the basic semialgebraic set K_{pop} and the graph $\Psi_{f_{\mathrm{sa}}}$ of f_{sa} can be defined as follows:

$$K_{\mathrm{pop}} := \{(\mathbf{x}, z_1, z_2) \in \mathbb{R}^{6+2} : \mathbf{x} \in K_{\mathrm{sa}}, h_j(\mathbf{x}, z_1, z_2) \geqslant 0, j = 1, \ldots, 6\} \;,$$
$$\Psi_{f_{\mathrm{sa}}} := \{(\mathbf{x}, f_{\mathrm{sa}}(\mathbf{x})) : \mathbf{x} \in K_{\mathrm{sa}}\} = \{(\mathbf{x}, z_2) : (\mathbf{x}, z_1, z_2) \in K_{\mathrm{pop}}\} \;,$$

where the multivariate polynomials h_j are defined by:

$$
\begin{aligned}
h_1(\mathbf{x}, \mathbf{z}) &:= z_1 - m_1 \;, & h_4(\mathbf{x}, \mathbf{z}) &:= -z_1^2 + 4x_1 \Delta \mathbf{x} \;, \\
h_2(\mathbf{x}, \mathbf{z}) &:= M_1 - z_1 \;, & h_5(\mathbf{x}, \mathbf{z}) &:= z_2 z_1 - \partial_4 \Delta \mathbf{x} \;, \\
h_3(\mathbf{x}, \mathbf{z}) &:= z_1^2 - 4x_1 \Delta \mathbf{x} \;, & h_6(\mathbf{x}, \mathbf{z}) &:= -z_2 z_1 + \partial_4 \Delta \mathbf{x} \;.
\end{aligned}
$$

---

^{2} We presume that in [25, Lemma 3], "well-defined function f" stands for the fact that f can be evaluated in a non-ambiguous way on the considered domain.

Let $h_0 := 1, \omega_l := \deg h_l$ $(0 \leqslant l \leqslant 6)$. Consider the following semidefinite relaxations:

$$Q_k^{sa} : \begin{cases} \max_{\mu, \sigma_j, \theta_l} \mu \\ \text{s.t.} \quad z_2 - \mu = \sum_{j=1}^{12} \sigma_j(\mathbf{x}, \mathbf{z}) g_j(\mathbf{x}) + \sum_{l=0}^{6} \theta_l(\mathbf{x}, \mathbf{z}) h_l(\mathbf{x}, \mathbf{z}), \;\; \forall(\mathbf{x}, \mathbf{z}) \;\;, \\ \sigma_j \in \Sigma_{k-1}[\mathbf{x}, \mathbf{z}], \;\; 1 \leqslant j \leqslant 12 \;\;, \\ \theta_l \in \Sigma_{k-\lceil \omega_l/2 \rceil}[\mathbf{x}, \mathbf{z}], 0 \leqslant l \leqslant 7 \;\;. \end{cases}$$

When $k \geqslant k_0 := \max_{1 \leqslant j \leqslant 6}\{\lceil \omega_j/2 \rceil\} = 2$, then as a special case of Theorem 1, the sequence $(\inf(Q_k^{sa}))_{k \geqslant 2}$ is monotonically non-decreasing and converges to f_{sa}^*. The lower bound $m_2 = -0.618$ computed at the Q_2^{sa} relaxation is too coarse. A tighter lower bound $m_3 = -0.445$ is obtained at the third relaxation, but it consumes more CPU time.

3 Maxplus Approximations and Nonlinear Templates

3.1 The Basis of Maxplus Functions

Let $\mathcal{B}$ be a set of functions $\mathbb{R}^n \to \mathbb{R}$, whose elements will be called *maxplus basis functions*. Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we look for a representation of f as a linear combination of basis functions in the maxplus sense, i.e.,

$$f = \sup_{w \in \mathcal{B}} (a(w) + w) \;\;, \tag{3.1}$$

where $(a(w))_{w \in \mathcal{B}}$ is a family of elements of $\mathbb{R} \cup \{-\infty\}$ (the "coefficients"). The correspondence between the function $x \mapsto f(x)$ and the coefficient function $w \mapsto a(w)$ is a well studied problem, which has appeared in various guises (Moreau conjugacies, generalized Fenchel transforms, Galois correspondences, see [2] for more background).

The idea of maxplus approximation [13, 28, 3] is to choose a space of functions f and a corresponding set $\mathcal{B}$ of basis functions w and to approximate from below a given f in this space by a finite maxplus linear combination, $f \simeq \sup_{w \in \mathcal{F}}(a(w) + w)$, where $\mathcal{F} \subset \mathcal{B}$ is a finite subset. Note that $\sup_{w \in \mathcal{F}}(a(w) + w)$ is not only an approximation but a valid lower bound of f. This is reminiscent of classical linear approximation methods and in particular of the finite element methods, in which a function in an finite dimensional space is approximated by a linear combination of prescribed elementary functions. Note that the term "basis" is abusive in the maxplus setting, as the family of functions $w \in \mathcal{F}$ is generally not free in the tropical sense.

A convenient choice of maxplus basis functions is the following [13, 3]. For each constant $\gamma \in \mathbb{R}$, we shall consider the family of quadratic functions $\mathcal{B} = \{w_{\mathbf{y}} \mid \mathbf{y} \in \mathbb{R}^n\}$, where

$$w_{\mathbf{y}}(\mathbf{x}) := -\frac{\gamma}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 \;\;. \tag{3.2}$$

Whereas in classical approximation problems, the ambient function spaces of interest are Sobolev spaces H^k, or spaces $\mathcal{C}^k$ of k times differentiable functions, in the tropical settings, the appropriate spaces, consistent with the choice of quadratic maxplus basis functions, turn out to consist of *semiconvex functions*, which we next examine.

3.2 Maxplus Approximation for Semiconvex Functions

The following definition is standard in variational analysis.

Definition 3 (Semiconvex function) Let γ denote a nonnegative constant. A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be γ-*semiconvex* if the function $\mathbf{x} \mapsto \phi(\mathbf{x}) + \frac{\gamma}{2}\|\mathbf{x}\|_2^2$ is convex.

Proposition 1 *Let $\mathcal{B}$ denote the set of quadratic functions $w_{\mathbf{y}}$ of the form (3.2) with $\mathbf{y} \in \mathbb{R}^n$. Then, the set of functions f which can be written as a maxplus linear combination (3.1) for some function $a : \mathcal{B} \to \mathbb{R} \cup \{-\infty\}$ is precisely the set of lower semicontinuous γ-semiconvex functions.*

Proof Let us note $h^\star : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ the Legendre-Fenchel transform of a function $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, so that $h^\star(p) := \sup_{x\in\mathbb{R}^n} \langle p, x\rangle - h(x)$. A known fact is that a convex lower semicontinuous function $g : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is the supremum of the affine functions that it dominates [37, Th. 12.1]. Actually, it is shown there that $g(x) = g^{\star\star}(x) = \sup_{p\in\mathbb{R}^n} \langle p, x\rangle - g^\star(p)$. By applying this result to the function $g(x) = f(x) + \frac{\gamma}{2}\|x\|_2^2$, we deduce that $f(x) = \sup_{p\in\mathbb{R}^n} \langle p, x\rangle - \frac{\gamma}{2}\|x\|_2^2 - g^\star(p) = \sup_{p\in\mathbb{R}^n} -\frac{\gamma}{2}\|x - \frac{1}{\gamma}p\|_2^2 - g^\star(p) + \frac{1}{2\gamma}\|p\|_2^2$. which is of the form (3.2).

Conversely, since an arbitrary supremum of γ-semiconvex and lower semicontinuous is also γ-semiconvex and lower semicontinuous, the supremum in (3.2) defines a γ-semiconvex and lower semicontinuous function. $\qquad\square$

The transcendental functions which we consider here are twice continuously differentiable. Hence, their restriction to any bounded convex set is γ-semiconvex for a sufficiently large γ, so that they can be approximated by finite suprema of the form $\sup_{w\in\mathcal{F}}(a(w) + w)$ with $\mathcal{F} \subset \mathcal{B}$.

The following result is derived in [14, Theorem 3.2] using methods and results of Grüber [16], who studied the best approximation of a convex body by a polytope. It shows that if $N = |\mathcal{F}|$ basis functions are used, then the best approximation error is precisely of order $1/N^{2/n}$ (the error is the sup-norm, over any compact set), provided that the function to be approximated is of class $\mathcal{C}^2$. We call $\mathcal{D}^2(\phi)(\mathbf{x})$ the Hessian matrix of ϕ at $\mathbf{x}$ and suppose that we approximate the function ϕ by the finite supremum of N γ-semiconvex functions parametrized by $p_i (i = 1, \ldots, N)$ and $a_i (i = 1, \ldots, N)$:

$$\phi \simeq \tilde{\phi}_N := \max_{1\leqslant i\leqslant N} \left\{ \frac{\gamma}{2}\|\mathbf{x}\|_2^2 + p_i^T\mathbf{x} + a(p_i) \right\} \ .$$

Theorem 2 (sup approximation error, [14, Theorem 3.2]) *Let $\gamma \in \mathbb{R}$, $\epsilon > 0$ and let $K \subset \mathbb{R}^n$ denote any full dimensional compact convex subset. If $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ is $(\gamma - \epsilon)$-semiconvex of class $\mathcal{C}^2$, then there exists a positive constant α depending only on n such that:*

$$\|\phi - \tilde{\phi}_N\|_\infty \sim \frac{\alpha}{N^{2/n}} \Big(\int_K [\det(\mathcal{D}^2(\phi)(\mathbf{x}) + \gamma I_n)]^{\frac{1}{2}} d\mathbf{x} \Big)^{\frac{2}{n}} \ as \ N \to \infty \ .$$

Thus, the best approximation satisfies

$$\|\phi - \tilde{\phi}_N\|_\infty \simeq \frac{C(\phi)}{N^{2/n}} \ , \tag{3.3}$$

where the constant $C(\phi)$ is explicit (it depends of $\det(\mathcal{D}^2(\phi) + \gamma I_n)$ and is bounded away from 0 when ϵ is fixed). This approximation indicates that some curse of dimensionality is unavoidable: to get a uniform error of order ϵ, one needs a number of basis functions of order $1/\epsilon^{n/2}$. Equivalently, the approximation error is of order $O(h^{\frac{2}{n}})$ where h is a space discretization step. The assumption that $\tilde{\phi}_N$ is of class $\mathcal{C}^2$ in Theorem 2 is needed to obtain a tight asymptotics of the approximation error. However, the max-plus approximation error is known to be of order $O(N^{2/n})$ under milder assumptions, requiring only semi-convexity type condition, see Proposition 64 of [22], and also Lemma 16 of [3] for a coarser approximation in $O(N^{1/n})$ valid in more general circumstances. This is due to the asymmetrical character of the maxplus approximation (a "one-sided" regularity, captured by the semiconvexity condition, is involved). Thus, unlike Taylor models, max-plus approximation does not require a $\mathcal{C}^k$ type regularity. For instance, a nonsmooth function like $|x| - x^2/2 = \max(x - x^2/2, -x - x^2/2)$ can be perfectly represented by two quadratic max-plus basis functions. In what follows, we shall always apply the approximation to small dimensional constituents of the optimization problems.

In this way, starting from a transcendental univariate elementary function $f \in \mathcal{D}$, such as arctan, exp, *etc* , defined on a real bounded interval I, we arrive at a semialgebraic lower bound of f, which is nothing but a supremum of a finite number of quadratic functions.

Example 4 Consider the function $f = \arctan$ on an interval $I := [m, M]$. For every point $a \in I$, we can find a constant γ such that

$$\arctan(x) \geqslant \mathrm{par}_a^-(x) := -\frac{\gamma}{2}(x - a)^2 + f'(a)(x - a) + f(a) \ .$$

Choosing $\gamma = \sup_{x \in I} -f''(x)$ always work. However, it will be convenient to allow γ to depend on the choice of a to get tighter lower bounds. Choosing a finite subset $A \subset I$, we arrive at an approximation

$$\forall x \in I, \arctan(x) \geqslant \max_{a \in A} \mathrm{par}_a^-(x) \ . \tag{3.4}$$

Semialgebraic over-approximations $x \mapsto \min_{a \in A} \mathrm{par}_a^+(x)$ can be defined in a similar way. Examples of such under-approximations and over-approximations are depicted in Fig. 1.
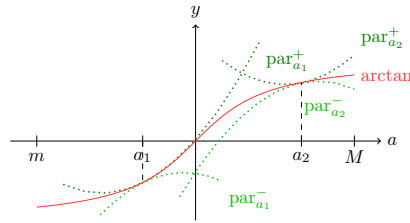

Fig. 1 Semialgebraic under-approximations and over-approximations for arctan

Example 5 Consider the bivariate function $g : (x_1, x_2) \mapsto \sin(x_1 + x_2)$, defined on $K := [-1.5, 4] \times [-3, 3]$, which is a component of the objective function from Problem MC (see Appendix A). As in the previous example, we can build underapproximations for the sine function. Choosing $\gamma = 1$, for every $(x_1, x_2) \in K$ and every $a \in [-4.5, 7]$, one has $\sin(x_1 + x_2) \geqslant -\frac{1}{2}(x_1 + x_2 - a)^2 + \cos(a)(x_1 + x_2 - a) + \sin(a)$.

3.3 Nonlinear Templates

The non-linear template method is a refinement of polyhedral based methods in static analysis [38]. It can also be closely related to the non-linear extension [1] of the template method and to the class of affine relaxation methods [31].

Templates allow one to determine invariants of programs by considering parametric families of subsets of $\mathbb{R}^n$ of the form $S(\alpha) = \{\mathbf{x} \mid w_i(\mathbf{x}) \leqslant \alpha_i, \ 1 \leqslant i \leqslant p\}$, where the vector $\alpha \in \mathbb{R}^p$ is the parameter, and $w_1, \ldots, w_p$ (the template) are fixed possibly non-linear functions, tailored to the program characteristics.

The nonlinear template method yields a tradeoff between the coarse bounds of interval calculus and the tighter bounds obtained with high-degree polynomial approximation (see Remark 1). On the one hand, templates take into account the correlations between the different variables. On the other hand, instead of increasing the degree of the approximation, one may increase the number of functions in the template.

Remark 1 Notice that by taking a trivial template (bound constraints, *i.e.* , functions of the form $\pm x_i$), the template method specializes to a version of interval calculus, in which bounds are derived by SOS techniques. The standard Taylor (resp. Chebyshev) approximations of transcendental functions can also be retrieved by instantiating some of the w_i to degree-d Taylor polynomials (resp. best uniform degree-d polynomials).

The max-plus basis method introduced in Sect. 3.1 is equivalent to the approximation of the epigraph of a function by a set $S(\alpha)$. This method involves the approximation from below of a function f in n variables by a supremum $f \gtrsim g := \sup_{1 \leqslant i \leqslant p} \lambda_i + w_i$. The functions w_i are fixed in advance, or dynamically adapted by exploiting the problem structure. The parameters λ_i are degrees of freedom.

The template method consists in propagating approximations of the set of reachable values of the variables of a program by sets of the form $S(\alpha)$. The non-linear template and max-plus approximation methods are somehow related. Indeed, the 0-level set of g, $\{\mathbf{x} \mid g(\mathbf{x}) \leqslant 0\}$, is nothing but $S(-\lambda)$, so templates can be recovered from max-plus approximations and vice versa. The functions w_i are usually required to be quadratic polynomials, $w_i(\mathbf{x}) = q_i^T \mathbf{x} + \frac{1}{2}\mathbf{x}^T A_i \mathbf{x}$, where $q_i \in \mathbb{R}^n$ and A_i is a symmetric matrix. A basic choice is $A_i = -\gamma I_n$, where γ is a fixed constant. Then, the parameters q remain the only degrees of freedom.

4 The Nonlinear Template Optimization Algorithm

Here we explain how to combine semialgebraic optimization techniques with approximation tools for univariate or semialgebraic functions. Let us consider an instance

of Problem (1.2). We assimilate the objective function f with its abstract syntax tree t. We assume that the leaves of t are semialgebraic functions in the set $\mathcal{A}$ and other nodes are univariate transcendental functions (arctan, *etc*) or basic operations ($+$, $\times$, $-$, $/$). For the sake of the simplicity, we suppose that each univariate transcendental function is monotonic.

4.1 A Semialgebraic Template Approximation Algorithm

The auxiliary algorithm `template_approx` is presented in Fig. 2.

Given an abstract syntax tree t, a semialgebraic set $K := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0\}$, an SOS relaxation order k and a precision p which can be either a finite sequence s of points $\mathbf{x}_1, \ldots, \mathbf{x}_p \in K$ or a polynomial approximation degree d, the algorithm `template_approx` computes a lower bound m (resp. upper bound M) of t over K and an under-approximation t^- (resp. an over-approximation t^+) of t by means of semialgebraic functions. We assume that the semialgebraic set $K := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0\}$ is contained in a box of $\mathbb{R}^n$.

When $t \in \mathcal{A}$ (Line 1), it suffices to set $t^- = t^+ := t$.

When the root of t is a binary operation whose arguments are two children c_1 and c_2, we apply recursively `template_approx` to each child and get semialgebraic under-approximations c_1^-, c_2^- and over-approximations c_1^+, c_2^+. Then, we obtain semialgebraic approximations of t by using the semialgebraic arithmetic procedure `compose_bop` (the rules are analogous with interval calculus).

When t corresponds to the composition of a transcendental (unary) function r with a child c, lower and upper bounds m_c and M_c are recursively obtained (Line 7), as well as semialgebraic approximations c^- and c^+. Then we define $I := [m_c, M_c]$ and apply the function `unary_approx` to get approximations r^- and r^+ of r over I. The parameters of `unary_approx` are the univariate function r, the precision p of the approximation and the closed interval I where r must be approximated. We shall need to consider various schemes for `unary_approx`:

1. A classical one is the approximation of univariate functions with best uniform polynomials of increasing degrees through Remez algorithm. In practice, we use the function `remez` available in the `Sollya` tool [12]. In this case, the precision is the degree-d of the minimax polynomial approximation. When the algorithm converges and returns a degree-d polynomial f_d, then a numerical approximation of the infinity norm of the error function ($r - f_d$) on the interval I can be obtained (`infnorm` routine from `Sollya`).
2. An alternative approach is to compute maxplus approximation using the semi-convexity properties of r on the interval I (see Sect. 3.2). Doing so, one bounds r from below with a function r^- being a supremum of parabola as well as from above with a function r^+ being a infimum of parabola. In this case, the precision is determined by certain sets s of points, which also parametrize the approximations r^- and r^+ (see e.g. the right hand side of (3.4) for an example of under-approximation r^-).

The approximations r^- and r^+ are composed with c^- and c^+ (`compose_approx` function at Line 10) to obtain an under-approximation t^- (resp. over-approximation t^+) of t. Notice that the behavior of `compose_approx` depends on the monotonicity properties of r.

Input: tree t, semialgebraic K, semidefinite relaxation order k, precision p
Output: lower bound m, upper bound M, lower semialgebraic approximation t_2^-, upper semi-
   algebraic approximation t_2^+
 1: **if** $t \in \mathcal{A}$ **then** $t^- := t$, $t^+ := t$
 2: **else if** $\texttt{bop} := \texttt{root}(t)$ is a binary operation with children c_1 and c_2 **then**
 3:     $m_i, M_i, c_i^-, c_i^+ := \texttt{template\_approx}(c_i, K, k, p)$ for $i \in \{1, 2\}$
 4:     $I_2 := [m_2, M_2]$
 5:     $t^-, t^+ := \texttt{compose\_bop}(c_1^-, c_1^+, c_2^-, c_2^+, \texttt{bop}, I_2)$
 6: **else if** $r := \texttt{root}(t) \in \mathcal{D}$ with child c **then**
 7:     $m_c, M_c, c^-, c^+ := \texttt{template\_approx}(c, K, k, p)$
 8:     $I := [m_c, M_c]$
 9:     $r^-, r^+ := \texttt{unary\_approx}(r, I, c, p)$
10:     $t^-, t^+ := \texttt{compose\_approx}(r, r^-, r^+, I, c^-, c^+)$
11: **end**
12: $t_2^- := \texttt{reduce\_lift}(t, K, k, p, t^-)$, $t_2^+ := -\texttt{reduce\_lift}(t, K, k, p, -t^+)$
13: **return** $\texttt{min\_sa}(t_2^-, K, k), \texttt{max\_sa}(t_2^+, K, k), t_2^-, t_2^+$

Fig. 2 $\texttt{template\_approx}$

4.2 Reducing the Complexity of Semialgebraic approximations

The semialgebraic approximations previously computed are used to determine lower and upper bounds of the function associated with the tree t, at each step of the induction. The bounds are obtained by calling the functions $\texttt{min\_sa}$ and $\texttt{max\_sa}$ respectively, which reduce the semialgebraic optimization problems to polynomial optimization problems by introducing extra lifting variables (see Section 2.2). However, the complexity of solving the SOS relaxations can grow significantly because of the number n_{lifting} of lifting variables. If k denotes the relaxation order, the corresponding SOS problem Q_k indeed involves linear matrix inequalities of size $\binom{n + n_{\text{lifting}} + k}{k}$ over $\binom{n + n_{\text{lifting}} + 2k}{2k}$ variables. The complexity of the semialgebraic approximations is controlled with the function $\texttt{reduce\_lift}$ (Line 12), when the number of lifting variables exceeds a user-defined threshold value $n_{\text{lifting}}^{\max}$. Consequently, this is crucial to control the number of lifting variables, or equivalently, the complexity of the semialgebraic approximations. For this purpose, we introduce two approximation schemes. The first one is presented in Sect. 4.2.1. It allows to compute approximations for some sub-components of the tree t (or its under-approximation t^-) by means of suprema/infima of quadratic functions. An alternative approach is to approximate these sub-components with degree-d polynomial under-approximations, using the semidefinite relaxation described in Sect. 4.2.2.

4.2.1 Multivariate Maxplus Quadratic Templates

Let $K \subset \mathbb{R}^n$ be a compact semialgebraic set and $f : K \to \mathbb{R}$ be a multivariate non-linear function. We consider the vector space $\mathcal{S}_n$ of real symmetric $n \times n$ matrices. Given a matrix $M \in \mathcal{S}_n$, let $\lambda_{\max}(M)$ (resp. $\lambda_{\min}(M)$) be the maximum (resp. minimum) eigenvalue of M. In the sequel, we will often refer to the quadratic polynomial defined below.

Definition 4 Let $\mathbf{x}_c \in K$. The quadratic polynomial $f_{\mathbf{x}_c, \lambda'}$ is given by:

$$
\begin{aligned}
f_{\mathbf{x}_c, \lambda'} : K &\longrightarrow \mathbb{R} \\
x &\longmapsto f(\mathbf{x}_c) + \mathcal{D}(f)(\mathbf{x}_c)\,(\mathbf{x} - \mathbf{x}_c) \\
&\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_c)^T \mathcal{D}^2(f)(\mathbf{x}_c)(\mathbf{x} - \mathbf{x}_c) \\
&\quad + \frac{1}{2}\lambda' \|\mathbf{x} - \mathbf{x}_c\|_2^2 \ ,
\end{aligned}
\tag{4.1}
$$

with,

$$
\lambda' \leqslant \lambda := \min_{\mathbf{x} \in K}\{\lambda_{\min}(\mathcal{D}^2(f)(\mathbf{x}) - \mathcal{D}^2(f)(\mathbf{x}_c))\} \ .
\tag{4.2}
$$

The quadratic polynomial $f_{\mathbf{x}_c, \lambda'}$ is an under-approximation of f on the set K:

Lemma 1 $\forall \mathbf{x} \in K, \ f(\mathbf{x}) \geqslant f_{\mathbf{x}_c, \lambda'}$.

Proof It comes from the first order Taylor expansion with the integral form for the remainder and the definition of the minimal eigenvalue. □

Definition 5 Given a symmetric real-valued matrix $M \in \mathcal{S}_n$, the spectral radius of M is given by $\rho(M) := \max(\lambda_{\max}(M), -\lambda_{\min}(M))$.

In the sequel, we use the following inequality:

$$
\rho(M) \leqslant \|M\|_1 := \max_{x \neq 0} \frac{\|Mx\|_1}{\|x\|_1} \ .
\tag{4.3}
$$

Now, we explain how to approximate f from below over K with quadratic polynomials as in (4.1). To approximate from below the value of λ, we determine an interval matrix $\widetilde{\mathcal{D}^2(f)} := ([\underline{d_{ij}}, \overline{d_{ij}}])_{1 \leqslant i,j \leqslant n}$, containing coarse bounds of the Hessian difference entries, using interval arithmetic. Let consider the following interval matrix minimal eigenvalue problem:

$$
\lambda' := \lambda_{\min}(\widetilde{\mathcal{D}^2(f)}) \ .
\tag{4.4}
$$

Different approximations of λ can be considered.

Tight lower bound of λ

For each interval $[\underline{d_{ij}}, \overline{d_{ij}}]$, we define the symmetric matrix B:

$$
B_{ij} := \max\{| \underline{d_{ij}} |, | \overline{d_{ij}} |\}, \ 1 \leqslant i, j \leqslant n \ .
$$

Let $\mathcal{S}^n$ be the set of diagonal matrices of sign:

$$
\mathcal{S}^n := \{\mathrm{diag}\,(s_1, \ldots, s_n), \ s_1 = \pm 1, \ldots s_n = \pm 1\} \ .
$$

Next, one specializes the result in [9, Theorem 2.1] for robust optimization procedure with reduced vertex set.

Lemma 2 *The robust interval SDP Problem* (4.4) *is equivalent to the following single variable SDP:*

$$\begin{cases} \min_{t} & -t \;, \\ s.t. & -tI - SBS \succcurlyeq 0 \;, \\ & S = \begin{pmatrix} 1 & 0 \\ 0 & S' \end{pmatrix}, \quad \forall S' \in \mathcal{S}^{n-1} \;. \end{cases}$$

Let λ_1' be the solution of this single variable SDP. Then, $\lambda_1' \leqslant \lambda$.

However, solving the semidefinite program given in Lemma 2 introduces a subset of sign matrices of cardinal 2^{n-1}, thus reduces the problem to a manageable size only if n is small.

Coarse lower bound of λ

Here, one writes $\widetilde{\mathcal{D}^2(f)} := X + Y$, where X and Y are defined as follows:

$$X_{ij} := \Big[\frac{d_{ij} + \overline{d_{ij}}}{2}, \frac{d_{ij} + \overline{d_{ij}}}{2}\Big], \quad Y_{ij} := \Big[-\frac{\overline{d_{ij}} - d_{ij}}{2}, \frac{\overline{d_{ij}} - d_{ij}}{2}\Big] \;.$$

Proposition 2 *Define $\lambda_2' := \lambda_{\min}(X) - \max_{1 \leqslant i \leqslant n}\Big\{\sum_{j=1}^{n} \dfrac{\overline{d_{ij}} - d_{ij}}{2}\Big\}$. Then, $\lambda_2' \leqslant \lambda$.*

Proof By concavity and homogeneity of the $\lambda_{\min}$ function, one has:

$$\lambda_{\min}(X + Y) \geqslant \lambda_{\min}(X) + \lambda_{\min}(Y) = \lambda_{\min}(X) - \lambda_{\max}(-Y) \;. \qquad (4.5)$$

Using Proposition 4.3 yields $\lambda_{\max}(-Y) \leqslant \max_{1 \leqslant i \leqslant n}\Big\{\sum_{j=1}^{n} \dfrac{\overline{d_{ij}} - d_{ij}}{2}\Big\}$. $\qquad \square$

The matrix X is real valued and symmetric matrix, thus one can compute its minimal eigenvalue with the classical semidefinite program: $\min\{-t : X \succcurlyeq tI\}$. Finally, we can compute a coarse certified lower bound λ_2' of λ with a procedure which is polynomial in n.

Now we describe the procedure `reduce_lift` in case of using the multivariate maxplus quadratic templates defined in this section, assuming that the precision p is determined by a sequence of points s. For each point $\mathbf{x}_c$ of the sequence s, a sub-routine returns the polynomial defined in (4.1). In particular, one has $f_{\mathbf{x}_c,1} := \texttt{build\_quadratic\_form}(f, \mathbf{x}_c, \lambda_1')$, which is built with the tight eigenvalue approximation λ_1'. Similarly, $f_{\mathbf{x}_c,2}$ is defined with the coarse eigenvalue approximation λ_2'. Since each $f_{\mathbf{x}_c,i}$ does not necessarily approximate f from below, we determine a lower bound m_c of the function $t^- - f_{\mathbf{x}_c,i}$, which ensures that $t_2^- := \max_{\mathbf{x}_c \in s}\{f_{\mathbf{x}_c,i} + m_c\}$ is a valid lower approximation of f.

4.2.2 Polynomial under-approximations for Semialgebraic Functions

Given a box $K \subset \mathbb{R}^n$, we consider a semialgebraic sub-component $f_{\mathrm{sa}} : K \to \mathbb{R}$ of the abstract syntax tree of f. A common way to represent f_{sa} is to use its semialgebraic lifting, which leads to solve semialgebraic optimization problems with a possibly large number of lifting variables n_{lifting}. One way to reduce this number is to approximate f_{sa} from below with a degree-d polynomial h_d, which should involve less variables than n_{lifting}. This section describes how to obtain such an h_d, which has the property to minimize the L_1 norm of the difference $(f_{\mathrm{sa}} - h)$, over all degree-d polynomial under-approximations h of f_{sa}. We exploit a technique of Lasserre and Thanh [23], who showed how to obtain convex under-approximations of polynomials. Here, we derive a similar hierarchy of SOS relaxations, whose optimal solutions are the best (for the L_1 norm) degree-d (but possibly non convex) polynomial under-approximations of t on K. We assume without loss of generality that K is the unit ball $[0,1]^n$. By comparison with [23], the main difference is that the input is a semialgebraic function, rather than a polynomial.

Best polynomial under-approximations of semialgebraic functions for the L_1 norm.
Let $f_{\mathrm{sa}} : [0,1]^n \to \mathbb{R}$ be a semialgebraic component of f and λ_n be the standard Lebesgue measure on $\mathbb{R}^n$, which is normalized so that $\lambda_n([0,1]^n) = 1$. Define $g_1 := x_1(1 - x_1), \ldots, g_n := x_n(1 - x_n)$. The function f_{sa} has a basic semialgebraic lifting, thus there exist $p, s \in \mathbb{N}$, polynomials $g_{n+1}, \ldots, g_{n+s} \in \mathbb{R}[\mathbf{x}, z_1, \ldots, z_p]$ and a basic semialgebraic set K_{pop} defined by:

$$K_{\mathrm{pop}} := \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^{n+p} : g_1(\mathbf{x}, \mathbf{z}) \geqslant 0, \ldots, g_m(\mathbf{x}, \mathbf{z}) \geqslant 0, g_{m+1}(\mathbf{x}, \mathbf{z}) \geqslant 0\} \ ,$$

such that the graph $\Psi_{f_{\mathrm{sa}}}$ satisfies:

$$\Psi_{f_{\mathrm{sa}}} := \{(\mathbf{x}, f_{\mathrm{sa}}(\mathbf{x})) : \mathbf{x} \in K\} = \{(\mathbf{x}, z_p) : (\mathbf{x}, \mathbf{z}) \in K_{\mathrm{pop}}\} \ ,$$

with $m := n + s$ and $g_{m+1} := M - \|\mathbf{z}\|_2^2$, for some positive constant M obtained by adding bound constraints over the lifting variables $\mathbf{z}$ (to ensure that the module $QM(K_{\mathrm{pop}})$ is Archimedean). Define the polynomial $f_{\mathrm{pop}}(\mathbf{x}, \mathbf{z}) := z_p$ and the total number of variables $n_{\mathrm{pop}} := n + p$.

Consider the following optimization problem with optimal value m_d:

$$(P^{\mathrm{sa}}) \begin{cases} \min\limits_{h \in \mathbb{R}_d[\mathbf{x}]} \int_K (f_{\mathrm{sa}} - h) d\lambda_n \\ \text{s.t.} \quad f_{\mathrm{sa}} - h \geqslant 0 \text{ on } K \ . \end{cases}$$

Lemma 3 *Problem (P^{sa}) has a degree-d polynomial minimizer h_d.*

For a proof, see Appendix B.2. Now, define $QM(K_{\mathrm{pop}})$ to be the quadratic module associated with $g_1, \ldots, g_{m+1}$. As a consequence of Putinar's Positivstellensatz for Archimedean quadratic modules [36], the optimal solution h_d of (P^{sa}) is a maximizer of the following problem:

$$(P_d) \begin{cases} \max\limits_{h \in \mathbb{R}_d[\mathbf{x}]} \int_{[0,1]^n} h \, d\lambda_n \\ \text{s.t.} \quad (f_{\mathrm{pop}} - h) \in QM(K_{\mathrm{pop}}) \ . \end{cases}$$

Let μ_d be the optimal value of (P_d). Then, one has $m_d = \int_K f_{\mathrm{sa}} \, d\lambda - \mu_d$.

Convergent hierarchy of SOS relaxations. We write $h = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_d^n} h_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$, with $\mathbb{N}_d^n :=$ $\{\boldsymbol{\alpha} \in \mathbb{N}^n : \sum_{i=1}^n \alpha_i \leqslant d\}$. Let $\tilde{\omega}_0 := \lceil (\deg g_0)/2 \rceil, \ldots, \tilde{\omega}_{m+1} := \lceil (\deg g_{m+1})/2 \rceil$ and define $k_0 := \max\{\lceil d/2 \rceil, \lceil (\deg f_{\text{pop}})/2 \rceil, \tilde{\omega}_0, \ldots, \tilde{\omega}_{m+1}\}$. Now, consider the following SOS relaxation (P_{dk}) of (P_d), with optimal value μ_{dk}:

$$(P_{dk}) \begin{cases} \displaystyle\max_{h \in \mathbb{R}_d[\mathbf{x}], \sigma_j} \sum_{\boldsymbol{\alpha} \in \mathbb{N}_d^n} h_{\boldsymbol{\alpha}} \gamma_{\boldsymbol{\alpha}} \\ \text{s.t.} \qquad f_{\text{pop}}(\mathbf{x}, \mathbf{z}) = h(\mathbf{x}) + \displaystyle\sum_{j=0}^{m+1} \sigma_j(\mathbf{x}, \mathbf{z}) g_j(\mathbf{x}, \mathbf{z}), \ \forall (\mathbf{x}, \mathbf{z}) \ , \\ \qquad\qquad \sigma_j \in \Sigma_{k - \tilde{\omega}_j}[\mathbf{x}, \mathbf{z}], \ 0 \leqslant j \leqslant m + 1 \ , \end{cases}$$

with $k \geqslant k_0$ and $\gamma_{\boldsymbol{\alpha}} := \int_{[0,1]^n} \mathbf{x}^{\boldsymbol{\alpha}} d\mathbf{x}$ for all $\boldsymbol{\alpha} \in \mathbb{N}_d^n$. This problem is an SOS program with variables $(h_d, \sigma_0, \ldots, \sigma_{m+1})$. Let m_d be the optimal value of Problem (P^{sa}). As in [23], the optimal value of the SOS relaxation (P_{dk}) can become as close as desired to $m_d - f_{\text{sa}}^*$.

Theorem 3 *The sequence $(\int_K f_{sa} d\lambda - \mu_{dk})_{k \geqslant k_0}$ is non-increasing and converges to m_d. Moreover, if h_{dk} is a maximizer of (P_{dk}), then the sequence $(\|f_{sa} - h_{dk}\|_1)_{k \geqslant k_0}$ is non-increasing and converges to m_d. Furthermore, any accumulation point of the sequence $(h_{dk})_{k \geqslant k_0}$ is an optimal solution of Problem (P^{sa}).*

Proof The proof is analogous with [23, Theorem 3.3]. □

Numerical experiments. We present the numerical results obtained when computing the best degree-d polynomial under-approximations of semialgebraic functions for the L_1 norm, using the techniques presented in Sect. 4.2.2. The sequence of lower bounds (μ_{dk}) is computed by solving the SOS relaxations (P_{dk}). The "tightness" score $\|f_{\text{sa}} - h_{dk}\|_1$ evaluates the quality of the approximation h_{dk}, together with its lower bound μ_{dk}.

Example 6 In Example 3, we obtained lower bounds for the semialgebraic function $f_{\text{sa}} := \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}}$, using two lifting variables. However, when solving inequalities involving f_{sa}, one would like to solve POP that do not necessarily include these two lifting variables and the associated constraints. Table 1 displays the tightness scores and

Table 1 Comparing the tightness score $\|f_{\text{sa}} - h_{dk}\|_1$ and μ_{dk} for various values of d and k

d	k	Upper bound of $\|f_{\text{sa}} - h_{dk}\|_1$	μ_{dk}
2	2	0.8024	-1.171
	3	0.3709	-0.4479
4	2	1.617	-1.056
	3	0.1766	-0.4493
6	3	0.08826	-0.4471

the lower bounds of the approximations obtained for various values of the approximation degree d and the relaxation order k. Notice that μ_{dk} only bounds from below the actual infimum h_{dk}^* of the under-approximation h_{dk}. It requires a few seconds to compute approximations at $k = 2$ against 10 minutes at $k = 3$, but one shall consider to take advantage of replacing f_{sa} by its approximation h_{63} to solve more complex POP.

Input: abstract syntax tree t, semialgebraic set K, $iter_{\max}$ (optional argument), precision p
Output: lower bound m
1: $s := [\texttt{argmin}(\texttt{randeval}(t))]$ $\triangleright \, s \in K$
2: $m := -\infty$
3: $iter := 0$
4: **while** $iter \leqslant iter_{\max}$ **do**
5:      Choose an SOS relaxation order $k \geqslant k_0$
6:      $m, M, t^-, t^+ := \texttt{template\_approx}(t, K, k, p)$
7:      $\mathbf{x}_{opt} := \texttt{guess\_argmin}(t^-)$ $\triangleright \, t^-(\mathbf{x}_{opt}) \simeq m$
8:      $s := s \cup \{\mathbf{x}_{opt}\}$
9:      $p := \texttt{update\_precision}(p, \mathbf{x}_{opt})$
10:     $iter := iter + 1$
11: **done**
12: **return** $m, \mathbf{x}_{opt}$

Fig. 3 `template_optim` : Template Optimization Algorithm

As an alternative to the method proposed in Sect. 4.2.1, the procedure `reduce_lift` can return the polynomial under-approximation h_{dk} while solving the SOS relaxation (P_{dk}).

4.3 A Semialgebraic Template Optimization Algorithm

Our main optimization algorithm `template_optim` is an iterative procedure which relies on `template_approx`. At each iteration step, the global precision parameter $p \in \mathcal{P}$ is updated dynamically. A convenient way to express the refinement of the precision, for the general nonlinear template approximation scheme (see Fig. 2), is to use the vocabulary of nets. We recall the following definitions, using [33]:

Definition 6 A directed set is a set D with a relation $\leqslant$ which is reflexive, transitive and directed, *i.e.* for each $a, b \in D$, there exists some $c \in D$ such that $a \leqslant c$ and $b \leqslant c$.

Definition 7 A net in a set X is a map $\lambda : D \to X$. If X is a topological space, we say that the net λ converges to $x \in X$ and write $\lambda \to x$ if and only if for every neighborhood U of x, there exists some tail $\Lambda := \{\lambda(c) : d \leqslant c \in D\}$ such that $\Lambda \subseteq U$.

We represent the precision p by an element of a directed set $\mathcal{P}$. When using minimax polynomial approximations to approximate an univariate function on a given interval I, the sequence of approximation degrees defines the net. For the maxplus approximations, the net is the set of finite subsets of I.

Let $c_1, \ldots, c_l$ be the components of the tree t, on which one calls approximation algorithms with respective precisions $p_1 \in \mathcal{P}_1, \ldots, p_l \in \mathcal{P}_l$. Let $\mathcal{P} = \mathcal{P}_1 \times \cdots \times \mathcal{P}_l$ be the set of precisions, ordered with the product order.

Our main optimization algorithm `template_optim`, relies on `template_approx` and updates the global precision parameter $p \in \mathcal{P}$ dynamically at each step of an iteration procedure (Line 1).

Now we describe our main semialgebraic optimization algorithm `optim` (see Figure 3). Given an abstract syntax tree t and a compact semialgebraic set K this algorithm returns a lower bound m of t using semialgebraic minimax approximations

computed recursively with `template_approx`. The relaxation order k (Line 5) is a parameter of the semialgebraic optimization functions `min_sa` (as well as `max_sa`) and `reduce_lift`.

Assuming that K is described by polynomial inequalities $g_1(\mathbf{x}) \geqslant 0, \ldots, g_m(\mathbf{x}) \geqslant 0$. Then, the relaxation order must be at least $k_0 := \max_{1 \leqslant j \leqslant m} \{\lceil \deg(g_j)/2 \rceil)\}$. In practice, we solve semialgebraic optimization problems with the second or third SOS Lasserre's relaxation and take $k = k_0$. At the beginning, the set of points consists of a single point of the box K. This point is chosen so that it minimizes the value of the function associated to the tree t among a set of random points (Line 1). Then, at each iteration of the loop from Lines 4 to 11, the auxiliary function `template_approx` is called to compute a lower bound m of the function t (Line 6), using the approximations t^- and t^+. At Line 7, a minimizer candidate $\mathbf{x}_{opt}$ of the under-approximation tree t^- is computed. It is obtained by projecting a solution $\mathbf{x}_{sdp}$ of the SOS relaxation Q_k of Section 2 on the coordinates representing the first order moments, following [24, Theorem 4.2]. However, the projection may not belong to K when the relaxation order k is not large enough. This is why tools like SPARSEPOP use local optimization solver in a post-processing step to provide a point in K which may not be a global minimizer. In any case, $\mathbf{x}_{opt}$ is then added to the set of points (Line 8). Alternatively, if we are only interested in determining whether the infimum of t over K is nonnegative (Problem (1.3)), the loop can be stopped as soon as $m \geqslant 0$.

By comparison, when using minimax approximations, the stopping criterion is the maximal precision corresponding to a minimax polynomial approximation degree. This maximal degree $d_{\max}$ shall be selected after consideration of the computational power available since one may need to solve SOS relaxations involving $O(d_{\max}^n)$ variables with matrices of size $O(\lceil d_{\max}/2 \rceil^n)$.

Example 7 (Lemma$_{9922699028}$ Flyspeck) We continue Example 3. Since we computed lower and upper bounds (m and M) for $f_{\text{sa}} := \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}}$, we know that the f_{sa} argument of arctan lies in $I := [m, M]$. We describe three iterations of the algorithm `template_optim` while using maxplus approximation for `unary_approx`.

0. Multiple evaluations of f return a set of values and we obtain a first minimizer guess $\mathbf{x}_1 := \texttt{argmin}(\texttt{randeval}(f))$ corresponding to the minimal value of the set. One has $\mathbf{x}_1 := (4.8684, 4.0987, 4.0987, 7.8859, 4.0987, 4.0987)$.
1. We compute $a_1 := f_{\text{sa}}(\mathbf{x}_1) = 0.3962$ and define $\text{par}_{a_1}^-(x) := -\frac{1}{2}(x - a_1)^2 + \frac{1}{1+a_1^2}(x - a_1) + \arctan(a_1)$. Finally, we obtain $m_1 \leqslant \min_{\mathbf{x} \in K}\{l(\mathbf{x}) + \text{par}_{a_1}^-(f_{\text{sa}}(\mathbf{x}))\}$. For $k = 2$, one has $m_1 = -0.2816 < 0$ and a new minimizer $\mathbf{x}_2$.
2. We get $a_2 := f_{\text{sa}}(\mathbf{x}_2)$ and $m_2 \leqslant \min_{\mathbf{x} \in K}\{l(\mathbf{x}) + \max_{1 \leqslant i \leqslant 2}\{\text{par}_{a_i}^-(f_{\text{sa}}(\mathbf{x}))\}\}$. For $k = 2$, we get $m_2 = -0.0442 < 0$ and $\mathbf{x}_3$.
3. We get $a_3 := f_{\text{sa}}(\mathbf{x}_3)$, $\text{par}_{a_3}^-$ and $m_3 \leqslant \min_{\mathbf{x} \in K}\{l(\mathbf{x}) + \max_{1 \leqslant i \leqslant 3}\{\text{par}_{a_i}^-(f_{\text{sa}}(\mathbf{x}))\}\}$. For $k = 2$, we obtain $m_3 = -0.0337 < 0$ and get a new minimizer $\mathbf{x}_4$.

4.4 Convergence of the Nonlinear Template Method

Given an accuracy $\epsilon > 0$, we prove that the objective function f can be uniformly ϵ-approximated over the semialgebraic set K with the algorithm `template_approx` under certain assumptions.

Assumption 4 *The Archimedean condition holds for the quadratic modules that we consider when solving SOS relaxations.*

For the sake of simplicity, we assume that the function `reduce_lift` calls the procedure that returns the sequence of best (for the L_1 norm) polynomial under-approximations for semialgebraic functions (see Sect. 4.2.2). Let the relaxation order k be fixed and t_p^- (resp. t_p^+) be the under-approximation (resp. over-approximation) of t on K obtained with the `template_approx` function at precision p. The limit of a net indexed by $p \in \mathcal{P}$ is obtained by increasing the precision of each elementary approximation algorithm (either `unary_approx` or `reduce_lift`) applied to the components of t.

Proposition 3 (Convergence of `template_approx`) *Under Assumption 4, the nets $(t_p^-)_p$ and $(t_p^+)_p$ uniformly converge to t on K.*

For a proof, see Appendix B.3. Given a precision p, define $m_p^* := \inf_{\mathbf{x} \in K} t_p^-$ to be the optimal value of the under-approximation t_p^- on K. Notice that under Assumption 4, we can theoretically obtain this optimal value, using Theorem 3.

Corollary 1 (Convergence of the approximations optimal values) *Under Assumption 4, the net $(m_p^*)_p$ converges to the infimum f^*.*

Proof Let $\mathbf{x}_p^*$ be a minimizer of t_p^- on K and note $\mathbf{x}^*$ one minimizer of t on K, then one has $t(\mathbf{x}^*) = f^*, t_p^-(\mathbf{x}_p^*) = m_p^*$. By definition, the following inequalities hold:

$$t_p^-(\mathbf{x}_p^*) \leqslant t_p^-(\mathbf{x}^*) \leqslant t(\mathbf{x}^*) \leqslant t(\mathbf{x}_p^*) \ . \tag{4.6}$$

Let $\epsilon > 0$ be given. From Proposition 3, there exists a precision d_0 such that for all $d \geqslant d_0$, one has: $t(\mathbf{x}^*) - t_p^-(\mathbf{x}^*) < \epsilon/2$ and $t(\mathbf{x}_p^*) - t_p^-(\mathbf{x}_p^*) < \epsilon/2$. Thus, applying (4.6) yields $t(\mathbf{x}^*) - t_p^-(\mathbf{x}_p^*) < \epsilon$, the desired result.

$\square$

Corollary 2 (Convergence of `template_optim`) *Under Assumption 4, each limit point of the net of minimizers $(\mathbf{x}_p^*)_p$ is a global minimizer of t over K.*

For a proof, see Appendix B.4.2.

5 Numerical Results of the Nonlinear Template Method

We now present some numerical test results by applying the semialgebraic minimax optimization method to examples from the global optimization literature (see Appendix A), as well as inequalities from the Flyspeck project. The nonlinear template method is implemented as a software package, called `NLCertify`, written in OCaml and interfaced with the `Sollya` tool. For more details about this tool, we refer to the software web-page[3] as well as to the dedicated publication [26].

For each problem presented in Table 2, our aim is to certify a lower bound m of a function f on a box K. The semialgebraic optimization problems are solved at the SOS relaxation order k. When the relaxation gap is too high to certify the requested bound, then we perform a domain subdivision in order to get tighter bounds: we divide the maximal width interval of K in two halves to get two sub-boxes K_1 and K_2 such that $K = K_1 \cup K_2$. We repeat this subdivision procedure, by applying

---

[3] `http://nl-certify.forge.ocamlcore.org/`

`template_optim` on a finite set of sub-boxes, until we succeed to certify that m is a lower bound of f. We note #boxes the total number of sub-boxes generated by the algorithm.

The algorithm `template_optim` returns more precise bounds by successive updates of the precision p. For each univariate component $u \in \mathcal{D}$ of the objective f, we note $\#s_u$ the number of points for the maxplus approximations of u and d_u the degree of the minimax approximation of u.

A template-free SOS method coincides with the particular case in which $d_u = 0$ (or $\#s_u = 0$) for each univariate component $u \in \mathcal{D}$ and $n_{\text{lifting}} = 0$. We mentioned in [5] that this method already outperforms the interval arithmetic solvers. However, it can only be used for problems with a moderate number of variables. The algorithm `template_optim` allows us to overcome this restriction, while keeping a similar performance (or occasionally improving this performance) on medium-size examples.

The minimax approximation based method is eventually faster than the maxplus based method for moderate instances. For the example *H3* (resp. *H6*), the speed-up factor is 2 when the function exp is approximated by a quartic (resp. quadratic) minimax polynomial. On the other hand, notice that reducing the number of lifting variables allows us to provide more quickly coarse bounds for large-scale instances of the Schwefel problem. We discuss the results appearing in the two last lines of Table 2. Without any box subdivision, we can certify a better lower bound $m = -967n$ with $n_{\text{lifting}} = 2n$ since our semialgebraic approximation is more precise. However the last lower bound $m = -968n$ can be computed twice faster by considering only n lifting variables, thus reducing the size of the POP described in Example 2. This indicates that the method is able to avoid the explosion for certain hard sub-classes of problems where a standard (full lifting) POP formulation would involve a large number of lifting variables.

In Table 3, we present some test results for several non-linear Flyspeck inequalities. The integer $n_{\mathcal{D}}$ represents the number of transcendental univariate nodes in the corresponding abstract syntax trees. These inequalities are known to be tight and involve sum of arctan of correlated functions in many variables, whence we keep high the number of lifting variables to get precise semialgebraic approximations. However, some inequalities (e.g. 9922699028) are easier to solve by using coarser semialgebraic approximations. The first line ($n_{\text{lifting}} = 9$) corresponds to the algorithm described in [6]. The second and third line illustrate our improved template method. For the former ($n_{\text{lifting}} = 3$), we use no lifting variables to represent square roots of univariate functions. For the latter ($n_{\text{lifting}} = 1$), we use the `reduce_lift` variant of Sect. 4.2.2 to approximate from below the semialgebraic function $\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}$ with the approximation h_{42} (see Example 6), so that we save two more lifting variables. Note that this `reduce_lift` variant provides a systematic way to generate polynomial template approximations for semialgebraic functions but it is difficult in practice to generate precise approximations, due to the high computational cost of solving SOS relaxation (P_{dk}) for large d and k.

In Table 4, we compared our algorithm with the MATLAB toolbox `intsolver` [32] (based on the Newton interval method [19]) for random inequalities involving two transcendental functions. The functions that we consider are of the form $\mathbf{x} \mapsto \arctan(p(\mathbf{x})) + \arctan(q(\mathbf{x}))$, where p is a four-degree polynomial and q is a quadratic form. All variables lie in $[0, 1]$. Both p and q have random coefficients (taken in $[0, 1]$)

Table 2 Numerical results for global optimization examples using `template_optim`

Pb	n	m	p	n_{lifting}	k	#boxes	time
H3	3	-3.863	$\#s_{\exp} = 3$	4	2	99	$101\,s$
			$d_{\exp} = 0$	0	1	1096	$247\,s$
			$d_{\exp} = 2$	4	1	53	$132\,s$
			$d_{\exp} = 4$	4	2	19	$57\,s$
			$d_{\exp} = 6$	4	3	12	$101\,s$
H6	3	-3.33	$\#s_{\exp} = 1$	6	2	113	$102\,s$
			$d_{\exp} = 0$	0	1	113	$45\,s$
			$d_{\exp} = 2$	4	2	53	$51\,s$
MC	2	-1.92	$\#s_{\sin} = 2$	4	2	17	$1.8\,s$
			$d_{\sin} = 0$	0	1	92	$7.6\,s$
			$d_{\sin} = 2$	0	1	8	$6.3\,s$
			$d_{\sin} = 4$	0	2	4	$3.2\,s$
			$d_{\sin} = 6$	0	3	2	$3\,s$
			$d_{\sin} = 8$	0	4	1	$1.9\,s$
ML	10	-0.966	$\#s_{\cos} = 1$	5	1	5	$8.2\,s$
			$d_{\cos} = 0$	0	1	8	$6.6\,s$
			$d_{\cos} = 2$	5	1	1	$6.4\,s$
			$d_{\cos} = 4$	5	2	1	$8.1\,s$
SWF ($\epsilon = 0$)	10	$-430n$	$\#s_{\sin} = 6$	$2n$	2	16	$40\,s$
			$d_{\sin} = 0$	$2n$	1	3830	$129\,s$
			$d_{\sin} = 2$	$2n$	1	512	$2280\,s$
	10^2	$-440n$	$\#s_{\sin} = 6$	$2n$	2	274	$6840\,s$
			$d_{\sin} = 0$	0	1	$> 10^4$	$> 10^4\,s$
	10^3	$-486n$	$\#s_{\sin} = 4$	$2n$	2	1	$450\,s$
		$-488n$	$\#s_{\sin} = 4$	n	2	1	$250\,s$
SWF ($\epsilon = 1$)	10^3	$-967n$	$\#s_{\sin} = 2$	$2n$	2	1	$543\,s$
		$-968n$		n	2	1	$272\,s$

Table 3 Results for Flyspeck inequalities using `template_optim` with $n = 6$, $k = 2$, $m = 0$ and $\#s = \#s_{\text{arctan}}$

Inequality id	$n_{\mathcal{D}}$	p	n_{lifting}	#boxes	time
9922699028	1	$\#s = 4$	9	47	$241\,s$
	1	$\#s = 4, d_{\text{sqrt}} = 4$	3	39	$190\,s$
	1	$\#s = 1, d_{\text{sqrt}} = 4$	1	170	$1080\,s$
	1	$d_{\text{arctan}} = 4, d_{\text{sqrt}} = 4$	2	14	$244\,s$
3318775219	1	$\#s = 2$	9	338	$1560\,s$
	1	$d_{\text{arctan}} = 4, d_{\text{sqrt}} = 4$	2	266	$4423\,s$
7726998381	3	$\#s = 4$	15	70	$2580\,s$
7394240696	3	$\#s = 2$	15	351	$6480\,s$
4652969746_1	6	$\#s = 4$	15	81	$4680\,s$
OXLZLEZ 6346351218_2_0	6	$\#s = 4$	24	200	$20520\,s$

Table 4 Comparison results for random examples using either `intsolver` or `template_optim` with $k = 3$ and $\#s = \#s_{\text{arctan}} = 3$

n	m	`template_optim` (time)	`intsolver` (time)
3	0.4581	$3.8\,s$	$15.5\,s$
4	0.4157	$12.9\,s$	$172.1\,s$
5	0.4746	$60\,s$	$612\,s$
6	0.4476	$276\,s$	$12240\,s$

and are sparse. The results indicate that for such random examples, our method also outperforms interval arithmetic.

6 Conclusion

The present nonlinear template method computes certified lower bounds for global optimization problems. It can provide tight minimax or maxplus semialgebraic approximations to certify non-linear inequalities involving transcendental multivariate functions. Our algorithms can solve both small and intermediate size inequalities of the Flyspeck project as well as global optimization problems issued from the literature, with a moderate order of SOS relaxation.

The proposed approach bears some similarity with the "cutting planes" proofs in combinatorial optimization, the cutting planes being now replaced by nonlinear inequalities. It also allows one to limit the growth of the number of lifting variables as well as of polynomial constraints to be handled in the POP relaxations, at the price of a coarser approximation. Thus, our method is helpful when the size of optimization problems increases. Indeed, the coarse lower bounds obtained (even with a low SOS relaxation order) are better than those obtained with interval arithmetic or high-degree polynomial approximation.

Further research would be to apply nonlinear templates to discrete-time optimal control problems, e.g. the Mayer problem. In this case, each set of reachable vectors is abstracted by a template and computed in a forward fashion, by exploiting the dynamics, whereas the templates are refined in a backward fashion. Hence, the template method includes as a special case a set theoretical version of the familiar state/co-state optimality conditions in control.

References

1. Adje A, Gaubert S, Goubault E (2012) Coupling policy iteration with semi-definite relaxation to compute accurate numerical invariants in static analysis. Logical methods in computer science 8(1):1–32, DOI 10.2168/LMCS-8(1:1)2012
2. Akian M, Gaubert S, Kolokoltsov V (2005) Set coverings and invertibility of Functional Galois Connections. In: Litvinov G, Maslov V (eds) Idempotent Mathematics and Mathematical Physics, Contemporary Mathematics, vol 377, American Mathematical Society, pp 19–51
3. Akian M, Gaubert S, Lakhoua A (2008) The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis. SIAM J Control Optim 47(2):817–848, DOI 10.1137/060655286
4. Ali MM, Khompatraporn C, Zabinsky ZB (2005) A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems. J of Global Optimization 31(4):635–672, DOI 10.1007/s10898-004-9972-2
5. Allamigeon X, Gaubert S, Magron V, Werner B (2013) Certification of bounds of nonlinear functions : the templates method. To appear in the Proceedings of Conferences on Intelligent Computer Mathematics, CICM 2013 Calculemus, Bath
6. Allamigeon X, Gaubert S, Magron V, Werner B (2013) Certification of inequalities involving transcendental functions: combining sdp and max-plus approximation. To appear in the Proceedings of the European Control Conference, ECC'13, Zurich
7. Allamigeon X, Gaubert S, Magron V, Werner B (2014) Formal Proofs for Nonlinear Optimization. ArXiv e-prints 1404.7282
8. Berz M, Makino K (2009) Rigorous global search using taylor models. In: Proceedings of the 2009 conference on Symbolic numeric computation, ACM, New York, NY, USA, SNC '09, pp 11–20, DOI 10.1145/1577190.1577198

9. Calafiore G, Dabbene F (2008) Reduced vertex set result for interval semidefinite optimization problems. Journal of Optimization Theory and Applications 139:17–33, 10.1007/s10957-008-9423-1
10. Cannarsa P, Sinestrari C (2004) Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser Boston, URL `http://books.google.fr/books?id=kr-8FpVY2ooC`
11. Cartis C, Gould NIM, Toint PL (2011) Adaptive cubic regularisation methods for unconstrained optimization. part i: motivation, convergence and numerical results. Math Program 127(2):245–295
12. Chevillard S, Joldes M, Lauter C (2010) Sollya: An environment for the development of numerical codes. In: Fukuda K, van der Hoeven J, Joswig M, Takayama N (eds) Mathematical Software - ICMS 2010, Springer, Heidelberg, Germany, Lecture Notes in Computer Science, vol 6327, pp 28–31
13. Fleming WH, McEneaney WM (2000) A max-plus-based algorithm for a Hamilton-Jacobi-Bellman equation of nonlinear filtering. SIAM J Control Optim 38(3):683–710, DOI 10.1137/S0363012998332433
14. Gaubert S, McEneaney WM, Qu Z (2011) Curse of dimensionality reduction in max-plus based approximation methods: Theoretical estimates and improved pruning algorithms. In: CDC-ECC, IEEE, pp 1054–1061
15. Gil A, Segura J, Temme NM (2007) Numerical Methods for Special Functions, 1st edn. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA
16. Gruber PM (2007) Convex and discrete geometry. Springer, Berlin
17. Hales TC (1994) A proof of the kepler conjecture. Math Intelligencer 16:47–58
18. Hales TC (2005) A proof of the Kepler conjecture. Ann of Math (2) 162(3):1065–1185, DOI 10.4007/annals.2005.162.1065
19. Hansen E, Greenberg R (1983) An interval newton method. Applied Mathematics and Computation 12(2-3):89 – 98, DOI 10.1016/0096-3003(83)90001-2, URL `http://www.sciencedirect.com/science/article/pii/0096300383900012`
20. Hansen ER (2006) Sharpening interval computations. Reliable Computing 12(1):21–34
21. Kaltofen EL, Li B, Yang Z, Zhi L (2012) Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients. JSC 47(1):1–15, in memory of Wenda Wu (1929–2009)
22. Lakhoua A (2007) Max-plus finite element method for the numerical resolution of deterministic optimal control problems. PhD thesis, University of Paris 6
23. Lasserre J, Thanh T (2013) Convex underestimators of polynomials. Journal of Global Optimization 56(1):1–25, DOI 10.1007/s10898-012-9974-4
24. Lasserre JB (2001) Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization 11(3):796–817
25. Lasserre JB, Putinar M (2010) Positivity and optimization for semi-algebraic functions. SIAM Journal on Optimization 20(6):3364–3383
26. Magron V (2014) Nlcertify: A tool for formal nonlinear optimization. In: Hong H, Yap C (eds) Mathematical Software – ICMS 2014, Lecture Notes in Computer Science, vol 8592, Springer Berlin Heidelberg, pp 315–320, URL `http://dx.doi.org/10.1007/978-3-662-44199-2_49`
27. Maso G (1993) An Introduction to Gamma-Convergence. Birkhäuser
28. McEneaney WM (2006) Max-plus methods for nonlinear control and estimation. Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA
29. McEneaney WM (2007) A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs. SIAM J Control Optim 46(4):1239–1276, DOI 10.1137/040610830
30. McEneaney WM, Deshpande A, Gaubert S (2008) Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs. In: Proc. of the 2008 American Control Conference, Seattle, Washington, USA, pp 4684–4690, DOI 10.1109/ACC.2008.458723
31. Messine F (1999) Extensions of affine arithmetic: Application to unconstrained global optimization
32. Montanher TM (2009) Intsolver: An interval based toolbox for global optimization. Version 1.0, available from `www.mathworks.com`
33. Nagata J (1974) Modern general topology. Bibliotheca mathematica, North-Holland Pub. Co.
34. Parrilo PA, Sturmfels B (2003) Minimizing polynomial functions, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol 60, Amer. Math. Soc., Providence, RI, pp 83–99

35. Peyrl H, Parrilo PA (2008) Computing sum of squares decompositions with rational coefficients. Theor Comput Sci 409(2):269–281
36. Putinar M (1993) Positive polynomials on compact semi-algebraic sets. Indiana University Mathematics Journal 42(3):969–984
37. Rockafellar RT (1970) Convex Analysis. Princeton mathematical series, Princeton University Press
38. Sankaranarayanan S, Sipma HB, Manna Z (2005) Scalable analysis of linear systems using mathematical programming. In: Cousot R (ed) Proc. of Verification, Model Checking and Abstract Interpretation (VMCAI), Springer Verlag, Paris, France, LNCS, vol 3385, pp 21–47
39. Sridharan S, Gu M, James MR, McEneaney WM (2010) Reduced-complexity numerical method for optimal gate synthesis. Phys Rev A 82:042,319, DOI 10.1103/PhysRevA.82.042319
40. Waki H, Kim S, Kojima M, Muramatsu M (2006) Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. SIAM Journal on Optimization 17:218–242
41. Zumkeller R (2008) Rigorous global optimization. PhD thesis, École Polytechnique

A Appendix: global optimization problems issued from the literature

The following test examples are taken from Appendix B in [4]. Some of these examples involve functions that depend on numerical constants, the values of which can be found there.

- *Hartman 3 (H3)*: $\displaystyle\min_{\mathbf{x}\in[0,1]^3} f(\mathbf{x}) = -\sum_{i=1}^{4} c_i \exp\left[-\sum_{j=1}^{3} a_{ij}(x_j - p_{ij})^2\right]$.
- *Hartman 6 (H6)*: $\displaystyle\min_{\mathbf{x}\in[0,1]^6} f(\mathbf{x}) = -\sum_{i=1}^{4} c_i \exp\left[-\sum_{j=1}^{6} a_{ij}(x_j - p_{ij})^2\right]$.
- *Mc Cormick (MC)*, with $K = [-1.5, 4] \times [-3, 3]$:
  $\displaystyle\min_{\mathbf{x}\in K} f(\mathbf{x}) = \sin(x_1 + x_2) + (x_1 - x_2)^2 - 1.5x_1 + 2.5x_2 + 1$.
- *Modified Langerman (ML)*:
  $\displaystyle\min_{\mathbf{x}\in[0,10]^n} f(\mathbf{x}) = \sum_{j=1}^{5} c_j \cos(d_j/\pi) \exp(-\pi d_j),$ with $d_j = \sum_{i=1}^{n} (x_i - a_{ji})^2$.
- *Schwefel Problem (SWF)*: $\displaystyle\min_{\mathbf{x}\in[1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} x_i \sin(\sqrt{x_i})$.

B Appendix: proofs

B.1 Preliminary Results

For the sequel, we need to recall the following definition.

Definition 8 (Modulus of continuity) Let u be a real univariate function defined on an interval I. The modulus of continuity of u is defined as:

$$\omega(\delta) := \sup_{\substack{x_1,x_2\in I \\ |x_1-x_2|<\delta}} |u(x_1) - u(x_2)|$$

We shall also prove that `unary_approx` and `reduce_lift` return uniformly convergent approximations nets:

Proposition 4 *Suppose that Assumption 4 holds. For every function r of the dictionary $\mathcal{D}$, defined on a closed interval I, the procedure* `unary_approx` *returns two nets of univariate lower semialgebraic approximations $(r_p^-)_{p\in\mathcal{P}}$ and upper semialgebraic approximations $(r_p^+)_{p\in\mathcal{P}}$, that uniformly converge to r on I.*

For every semialgebraic function $f_{sa} \in \mathcal{A}$, defined on a compact semialgebraic set K, the procedure `reduce_lift` *returns two nets of lower semialgebraic approximations $(t_p^-)_{p\in\mathcal{P}}$ and upper semialgebraic approximations $(t_p^+)_{p\in\mathcal{P}}$, that uniformly converge to f_{sa} on K.*

Proof First, suppose that the precision p is the best uniform polynomial approximation degree. By Assumption 4, the procedure `unary_approx` returns the sequence of degree-d minimax polynomials, using the algorithm of Remez. This sequence uniformly converges to r on I, as a consequence of Jackson's Theorem [15, Chap. 3]. Alternatively, when considering maxplus approximations in which the precision is determined by certain sets of points, we can apply Theorem 2 that implies the uniform convergence of the maxplus approximations.

Next, for sufficiently large relaxation order, the `reduce_lift` procedure returns the best (for the L_1 norm) degree-d polynomial under-approximation of a given semialgebraic function, as a consequence of Theorem 3.   □

B.2 Proof of Lemma 3

Let us equip the vector space $\mathbb{R}_d[\mathbf{x}]$ of polynomials h of degree at most d with the norm $\|h\|_\infty := \sup_{|\boldsymbol{\alpha}| \leqslant d} \{|h_{\boldsymbol{\alpha}}|\}$.

Let H be the admissible set of Problem (P^{sa}). Observe that H is closed in the topology of the latter norm. Moreover, the objective function of Problem (P^{sa}) can be written as $\phi : h \in H \mapsto \|f_{\mathrm{sa}} - h\|_{L_1(K)}$, where $\|\cdot\|_{L_1(K)}$ is the norm of the space $L^1(K, \lambda_n)$. The function ϕ is continuous in the topology of $\|\cdot\|_\infty$ (for polynomials of bounded degree, the convergence of the coefficients implies the uniform convergence on every bounded set for the associated polynomial functions, and a fortiori the convergence of these polynomial functions in $L^1(K, \lambda_n)$). Note also that $\int_{[0,1]^n} h \ d\lambda_n = \int_{[0,1]^n} h(\mathbf{x}) \ d\lambda_n(\mathbf{x}) = \int_{[0,1]^{n+p}} h(\mathbf{x}, \mathbf{z}) \ d\lambda_{n+p}(\mathbf{x}, \mathbf{z})$. We claim that for every $t \in \mathbb{R}$, the sub-level set $S_t := \{h \in H \mid \phi(h) \leqslant t\}$ is bounded. Indeed, when $\phi(h) \leqslant t$, we have:
$$\|h\|_{L_1(K)} \leqslant \|f_{\mathrm{sa}} - h\|_{L_1(K)} + \|f_{\mathrm{sa}}\|_{L_1(K)} \leqslant t + \|f_{\mathrm{sa}}\|_{L_1(K)} \ .$$

Since on a finite dimensional vector space, all the norms are equivalent, there exists a constant $C > 0$ such that $\|h\|_\infty \leqslant C\|h\|_{L_1(K)}$ for all $h \in H$, so we deduce that $\|h\|_\infty \leqslant C(t + \|f_{\mathrm{sa}}\|_{L_1(K)})$ for all $h \in S_t$, which shows the claim. Since ϕ is continuous, it follows that every sublevel set of ϕ, which is a closed bounded subset of a finite dimensional vector space, is compact. Hence, the minimum of Problem (P^{sa}) is attained.   □

B.3 Proof of Proposition 3

The proof is by induction on the structure of t.

- When t represents a semialgebraic function of $\mathcal{A}$, the under-approximation (resp. over-approximation) net $(t_p^-)_p$ (resp. $(t_p^+)_p$) converges uniformly to t by Proposition 4.
- The second case occurs when the root of t is an univariate function $r \in \mathcal{D}$ with the single child c. Suppose that r is increasing without loss of generality. We consider the net of under-approximations $(c_p^-)_p$ (resp. over-approximations $(c_p^+)_p$) as well as lower and upper bounds m_{c_p} and M_{c_p} which are obtained recursively. Since K is a compact semialgebraic set, one can always find an interval I_0 enclosing the values of r_p^+ (*i.e.* such that $[m_{c_p}, M_{c_p}] \subset I_0$), for all p.

The induction hypothesis is the uniform convergence of $(c_p^-)_p$ (resp. $(c_p^+)_p$) to c on K. Now, we prove the uniform convergence of $(t_p^+)_p$ to t on K. One has:

$$\|t - t_p^+\|_\infty \leqslant \|r \circ c - r_p^+ \circ c\|_\infty + \|r_p^+ \circ c - t_p^+\|_\infty. \tag{B.1}$$

Let note ω the modulus of continuity of r_p^+ on I_0. Thus, the following holds:

$$\|r_p^+ \circ c - r_p^+ \circ c_p^+\|_\infty \leqslant \omega(\|c - c_p^+\|_\infty). \tag{B.2}$$

Let $\epsilon > 0$ be given. The univariate function r_p^+ is uniformly continuous on I_0, thus there exists $\delta > 0$ such that $\omega(\delta) \leqslant \epsilon/2$. Let choose such a δ. By induction hypothesis, there exists a precision p_0 such that for all $p \geqslant p_0$, $\|c - c_p^+\|_\infty \leqslant \delta$. Hence, using (B.2), the following holds:
$$\|r_p^+ \circ c - r_p^+ \circ c_p^+\|_\infty \leqslant \epsilon/2. \tag{B.3}$$

Moreover, from the uniform convergence of $(r_p^+)_{p \in \mathbb{N}}$ to r on K (by Proposition 4), there exists a precision p_1 such that for all $p \geqslant p_1$:

$$\|r \circ c - r_p^+ \circ c\|_\infty \leqslant \epsilon/2. \tag{B.4}$$

Using (B.1) together with (B.3) and (B.4) yield the desired result. The proof of the uniform convergence of the under-approximations is analogous.

- If the root of t is a binary operation whose arguments are two children c_1 and c_2, then by induction hypothesis, we obtain semialgebraic approximations $c_{1,p}^-, c_{2,p}^-, c_{1,p}^+, c_{2,p}^+$ that verify:

$$\lim_{p \to \infty} \|c_1 - c_{1,p}^-\|_\infty = 0, \qquad \lim_{p \to \infty} \|c_1 - c_{1,p}^+\|_\infty = 0, \tag{B.5}$$

$$\lim_{p \to \infty} \|c_2 - c_{2,p}^-\|_\infty = 0, \qquad \lim_{p \to \infty} \|c_2 - c_{2,p}^+\|_\infty = 0. \tag{B.6}$$

If $\mathtt{bop} = +$, by using the triangle inequality:

$$\|c_1 + c_2 - c_{1,p}^- - c_{2,p}^-\|_\infty \leqslant \|c_1 - c_{1,p}^-\|_\infty + \|c_2 - c_{2,p}^-\|_\infty,$$

$$\|c_1 + c_2 - c_{1,p}^+ - c_{2,p}^+\|_\infty \leqslant \|c_1 - c_{1,p}^+\|_\infty + \|c_2 - c_{2,p}^+\|_\infty.$$

Then, the uniform convergence comes from (B.5) and (B.6). The proof for the other cases is analogous. $\qquad \square$

B.4 Convergence of the `template_optim` Algorithm

B.4.1 Preliminaries: Γ and Uniform Convergence

To study the convergence of the minimizers of t_p^-, we first introduce some background on the Γ-convergence (we refer the reader to [27] for more details) and the lower semicontinuous envelope. The topology of Γ-Convergence is known to be metrizable hence, we shall consider the Γ-Convergence of sequences (rather than nets).

Definition 9 (Γ-Convergence) The sequence $(t_p)_{p \in \mathbb{N}}$ Γ-converges to t if the following two conditions hold:
1. (Asymptotic common lower bound) For all $\mathbf{x} \in K$ and all $(\mathbf{x}_p)_{p \in \mathbb{N}}$ such that $\lim_{p \to \infty} \mathbf{x}_p = \mathbf{x}$, one has $t(\mathbf{x}) \leqslant \liminf_{p \to \infty} t_p(\mathbf{x}_p)$.
2. (Existence of recovery sequences) For all $\mathbf{x} \in K$, there exists some $(\mathbf{x}_p)_{p \in \mathbb{N}}$ such that $\lim_{p \to \infty} \mathbf{x}_p = \mathbf{x}$ and $\limsup_{p \to \infty} t_p(\mathbf{x}_p) \geqslant t(\mathbf{x})$.

Define $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ to be the extended real number line.

Definition 10 (Lower Semicontinuous Envelope) Given $t : K \mapsto \overline{\mathbb{R}}$, the lower semicontinuous envelope of t is defined by:

$$t^{\mathrm{lsc}}(\mathbf{x}) := \sup\{g(\mathbf{x}) \mid g : K \mapsto \overline{\mathbb{R}} \text{ is lower semicontinuous and } g \leqslant f \text{ on } K\}.$$

If t is continuous, then $t^{\mathrm{lsc}} := t$.

Theorem 5 (Fundamental Theorem of Γ-Convergence [27]) *Suppose that the sequence $(t_p)_{p \in \mathbb{N}}$ Γ-converges to t and $\mathbf{x}_p$ minimizes t_p. Then every limit point of the sequence $(\mathbf{x}_p)_{p \in \mathbb{N}}$ is a global minimizer of t.*

Theorem 6 (Γ and Uniform Convergence [27]) *If $(t_p)_{p \in \mathbb{N}}$ uniformly converges to t, then $(t_p)_{p \in \mathbb{N}}$ Γ-converges to t^{lsc}.*

Theorem 6 also holds for nets, since the topology of Γ-Convergence is metrizable.

B.4.2 Proof of Corollary 2

From Proposition 3, the under-approximations net $(t_p^-)_{p \in \mathbb{N}}$ uniformly converge to t on K. Then, by using Theorem 6, the net $(t_p^-)_{p \in \mathbb{N}}$ Γ-converges to $t^{\mathrm{lsc}} := t$ (by continuity of t). It follows from the fundamental Theorem of Γ-Convergence 5 that every limit point of the net of minimizers $(\mathbf{x}_p^*)_{p \in \mathbb{N}}$ is a global minimizer of t over K. $\qquad \square$