Trajectory Generation for the Unicycle Model Using Semidefinite Relaxations

Hannah Häring, Dennis Gramlich, Christian Ebenbauer and Carsten W. Scherer

Abstract— We develop a convex relaxation for the minimum energy control problem of the well-known unicycle model in the form of a semidefinite program. Through polynomialization techniques, the infinite-dimensional optimal control problem is first reformulated as a non-convex, infinite-dimensional quadratic program which can be viewed as a trajectory generation problem. This problem is then discretized to arrive at a finite-dimensional, albeit, non-convex quadratically constrained quadratic program. By applying the moment relaxation method to this quadratic program, we obtain a sequence of semidefinite relaxations. We construct an approximate solution for the infinite-dimensional trajectory generation problem by solving the first- or second-order moment relaxation. A comprehensive simulation study provided in this paper suggests that the second-order moment relaxation is lossless.

I. INTRODUCTION

The development of reliable and efficient trajectory planning techniques is of tremendous importance to meet the high safety standards imposed on autonomous vehicles. Considerable attention has been devoted to designing optimization-based trajectory generation methods relying on convex optimization due to their successful application to real-world systems, e.g., to space vehicles [1], [11], [14], [16], [24], [28]. This approach typically involves solving an infinite-dimensional optimal control problem which is usually neither readily amenable to computation nor can a closed-form solution be specified.

Thus, several approaches in the literature, e.g., [6], [7], [13], [17], [18], relax and reformulate general infinite-dimensional optimal control problems, such as the minimum energy control problem for the unicycle model, in some manner leading to finite-dimensional convex optimization problems which can be solved with a dedicated convex optimization solver.

More specifically, the authors in [17] and [18] rely on local optimization algorithms or, more precisely, sequential convex programming methods to produce an approximate solution to an optimal control problem. Using this approach, the non-convex components in the optimal control problem are successively discretized and linearized around a reference trajectory to arrive at a sequence of convex approximations of the original problem. In [7], possible non-convex input constraints are relaxed and the nonlinear dynamics are handled by a piecewise linear approximation. Combined with a discretization step, this leads to a mixed integer linear program to be solved.

In [6] and [13], sum-of-squares and moment relaxation techniques are invoked, respectively, to approximately solve optimal control problems. The result in [13] uses the weak formulation of optimal control problems and assumes that the cost function is a polynomial and the constraints form a semi-algebraic set. This weak formulation can be viewed as a special instance of the generalized moment problem. Accordingly, a hierarchy of semidefinite relaxations is constructed which approximately solves the optimal control problem. The authors in [6] build upon this result from the sum-of-squares perspective in a reproducing kernel Hilbert space, allowing for non-polynomial and smooth data in the optimal control problem. They obtain a semidefinite program (SDP) by subsampling the "kernel sum-of-squares relaxation", approximating the optimal value of the optimal control problem.

As a key contribution, we present a novel approach to construct global approximate solutions to the minimum energy control problem for the unicycle model through semidefinite relaxations. While this paper exclusively addresses this particular setup, the proposed approach should set the stage for constructing approximate solutions to general optimal control problems that can be "polynomialized", i.e., reformulated as an infinite-dimensional polynomial program through rational transformations. Nevertheless, the minimum energy control problem for the unicycle model is of high practical importance itself, e.g., in aerospace applications. A concrete practical example is landing a spacecraft using a parafoil as described in [8]. By extending the unicycle model to include the altitude component, the resulting model can be viewed as a kinematic model for a parafoil. Note that this extension does not alter the considered optimal control problem since this altitude component can be calculated independently from the optimization problem.

This paper is structured as follows. In Section II the considered optimal control problem setup is presented and cast into an equivalent formulation in the form of a trajectory optimization problem. Section III demonstrates that loosening a particular constraint in the trajectory generation problem enables a reformulation as a non-convex quadratically constrained quadratic program (QCQP). In Section IV we show how the first and second semidefinite relaxation

in the moment hierarchy can be exploited to construct an approximate solution for the trajectory generation problem. Numerical experiments are provided in Section V. A technical proof for the reformulation of the optimal control problem as well as proofs and properties of the solution space of the relaxed optimal control problem are found in the appendix.

Notation. The absolute value of $x \in \mathbb{R}$ is denoted as $|x|$. The closure of an open interval $\Omega \subset \mathbb{R}$ is denoted by $\bar{\Omega}$. For an interval $I = (a, b) \subset \mathbb{R}$ we denote by $C^k(I)$ the set of k-times continuously differentiable functions on I and by $C^k(\bar{I})$ the set of k-times continuously differentiable functions on I such that the derivatives up to order k can be continuously extended to $\bar{I}$. If we let $L_2(I)$ be the space of square integrable functions on I, then the Sobolev space $H^k(I)$ is the set of functions $f \in L_2(I)$ whose weak derivatives up to k-th order are in $L_2(I)$. The k-th (weak) derivative of a function x with respect to t is denoted as $\frac{\mathrm{d}^k}{\mathrm{d}t^k}x(t)$, provided that it exists. For first and second order (weak) derivatives, we use the notation $\dot{x}(t) := \frac{\mathrm{d}}{\mathrm{d}t}x(t)$ and $\ddot{x}(t) := \frac{\mathrm{d}^2}{\mathrm{d}t^2}x(t)$, respectively. We make use of the standard embeddings $H^1(I) \subset C(\bar{I})$ and $H^2(I) \subset C^1(\bar{I})$, see, e.g., Theorem 5.12 in [3].

II. PROBLEM STATEMENT

The minimum energy control problem for the unicycle model with fixed initial and final configuration is given by

$$\underset{u \in L_2(\Omega)}{\text{minimize}} \quad \int_\Omega |u(t)|^2 \, \mathrm{d}t \tag{1}$$

$$\begin{aligned}
\text{subject to} \quad & \dot{x}(t) = v_p \cos(\theta(t)), && t \in \bar{\Omega}, \\
& \dot{y}(t) = v_p \sin(\theta(t)), && t \in \bar{\Omega}, \\
& \dot{\theta}(t) = u(t), && t \in \Omega, \\
& x(0) = x_s \in \mathbb{R}, \ y(0) = y_s \in \mathbb{R}, \\
& x(t_f) = x_f \in \mathbb{R}, \ y(t_f) = y_f \in \mathbb{R}, \\
& \theta(0) = \theta_s \in \mathbb{R}, \ \theta(t_f) = \theta_f \in \mathbb{R},
\end{aligned}$$

where $\Omega := (0, t_f)$, $t_f \in \mathbb{R}$. Here, $x, y \in H^2(\Omega)$ denote the vehicle's position as functions of time, $\theta \in H^1(\Omega)$ its heading angle as function of time and $v_p \in \mathbb{R}$ its constant horizontal speed. The vehicle's angular velocity can be manipulated through the control input $u \in L_2(\Omega)$. Note that, given $u \in L_2(\Omega)$, the constraints in problem (1) uniquely define $\theta \in H^1(\Omega)$ and $x, y \in H^2(\Omega)$, by integration. The constraints in (1) are well-defined due to the embeddings $H^1(\Omega) \subset C(\bar{\Omega})$ and $H^2(\Omega) \subset C^1(\bar{\Omega})$. The question of existence of a solution of (1) is not addressed in this paper.

In order to convexify (1) in the upcoming sections we rewrite this problem as

$$\underset{x, y \in H^2(\Omega)}{\text{minimize}} \quad J(x, y) := \int_\Omega |\ddot{x}(t)|^2 + |\ddot{y}(t)|^2 \, \mathrm{d}t \tag{2}$$

$$\text{subject to} \quad |\dot{x}(t)|^2 + |\dot{y}(t)|^2 = v_p^2, \quad t \in [0, t_f], \tag{2a}$$

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} x_s \\ y_s \end{pmatrix}, \begin{pmatrix} \dot{x}(t_f) \\ \dot{y}(t_f) \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_f) \\ \sin(\theta_f) \end{pmatrix},$$

$$\begin{pmatrix} x(t_f) \\ y(t_f) \end{pmatrix} = \begin{pmatrix} x_f \\ y_f \end{pmatrix}, \begin{pmatrix} \dot{x}(0) \\ \dot{y}(0) \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_s) \\ \sin(\theta_s) \end{pmatrix}.$$

Recall that the assumption $x, y \in H^2(\Omega)$ renders the constraint (2a) well-defined. In words, this reformulation means that a controller expending minimum energy corresponds to a trajectory with minimum bending energy, as stated in the next theorem, with a proof given in Appendix A.

Theorem 1: A solution of problem (1) can be obtained from a solution of problem (2) and vice versa.

In the sequel, we will relax problem (2) in order to arrive at a numerically tractable finite-dimensional formulation.

III. RELAXATION TO A FINITE-DIMENSIONAL PROBLEM

In a next step, we impose the constraint (2a) only at discrete points in time leading to the optimization problem

$$\underset{x, y \in H^2(\Omega)}{\text{minimize}} \quad \int_\Omega |\ddot{x}(t)|^2 + |\ddot{y}(t)|^2 \, \mathrm{d}t \tag{3}$$

$$\text{subject to} \quad |\dot{x}(t_i)|^2 + |\dot{y}(t_i)|^2 = v_p^2, \quad i = 0, \ldots, N+1,$$

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} x_s \\ y_s \end{pmatrix}, \begin{pmatrix} \dot{x}(t_f) \\ \dot{y}(t_f) \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_f) \\ \sin(\theta_f) \end{pmatrix},$$

$$\begin{pmatrix} x(t_f) \\ y(t_f) \end{pmatrix} = \begin{pmatrix} x_f \\ y_f \end{pmatrix}, \begin{pmatrix} \dot{x}(0) \\ \dot{y}(0) \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_s) \\ \sin(\theta_s) \end{pmatrix},$$

where $0 = t_0 < t_1 < \cdots < t_{N+1} = t_f$ is a partition of the interval $\bar{\Omega} = [0, t_f]$. As argued below, this relaxation of the constraint renders the corresponding solution space finite-dimensional. In fact, the solution space of (3) can be confined to a particular function space, which we call the space of variational generalized splines.

Definition 1: Let $a = t_0 < t_1 < \cdots < t_n = b$ be a partition of the interval $[a, b] \subset \mathbb{R}$ and let $I := (a, b)$. Then, the space of variational generalized spline functions is given by functions which solve the optimization problem

$$\underset{s \in H^2(I)}{\text{minimize}} \quad J(s) := \int_I |\ddot{s}(t)|^2 \, \mathrm{d}t \tag{4}$$

$$\text{subject to} \quad \dot{s}(t_i) = \beta_i, \quad i = 0, \ldots, n, \tag{4a}$$

$$s(a) = \alpha_0, \quad s(b) = \alpha_1, \tag{4b}$$

for some $\alpha_0, \alpha_1, \beta_i \in \mathbb{R}$, $i = 0, \ldots, n$. The space of variational generalized spline functions for a fixed partition of I in n subintervals is denoted as $S(\{t_i\}_{i=0}^n)$.

Remark 1: While Definition 1 may seem almost identical to definitions of other types of splines (see, e.g., [20], [25]), specifically cubic splines, our definition differs slightly. To be more precise, we do not require $s \in C^2(I)$ in Definition 1, only that $s \in H^2(I)$, i.e., the second derivative of s may be discontinuous when considering the entire interval I. In addition, while we require the function's first derivative to be interpolating for some given data points, we do not require this for the function itself.

In Appendix B we show that the space $S(\{t_i\}_{i=0}^n)$ is finite-dimensional, and in particular, that any $s \in S(\{t_i\}_{i=0}^n)$ satisfies

$$\frac{\mathrm{d}^4}{\mathrm{d}t^4} s(t) = 0, \quad \forall t \in I_i, \ i = 1, \ldots, n, \tag{5}$$

i.e., the function s is a cubic polynomial on the subintervals $I_i := (t_{i-1}, t_i)$, $i = 1, \ldots, n$. Most of the techniques used

to prove the results in Appendix B are reminiscent of proofs from spline theory. However, as mentioned in Remark 1, we did not find a type of spline in the literature which matches Definition 1 precisely. For this reason, we included Appendix B. The next theorem is a direct consequence of the results established thus far. Hence, its proof is omitted.

Theorem 2: Let $\Omega_i := (t_{i-1}, t_i)$ for $i = 1, \ldots, N + 1$. Further, let the pair (x, y) with $x, y \in H^2(\Omega)$ be a solution of (3). Then, $x, y \in S(\{t_i\}_{i=0}^{N+1})$.

By invoking Theorem 2, we derive a genuine finite-dimensional formulation of (3) as given in the next theorem.

Theorem 3: Define the set

$$\mathcal{T} := \{z \in \mathbb{R}^{2N} \mid h_i(z) := z_i^2 + z_{i+N}^2 - v_p^2 = 0,$$
$$i = 1, \ldots, N\}.$$

Then, there exists a matrix $Q \in \mathbb{R}^{2N+1 \times 2N+1}$ with $Q \succcurlyeq 0$ such that a solution to problem (3) can be obtained by solving the finite-dimensional QCQP

$$f^* := \inf_{z \in \mathcal{T}} f(z), \tag{6}$$

where $f(z) := \begin{pmatrix} 1 & z^\top \end{pmatrix} Q \begin{pmatrix} 1 & z^\top \end{pmatrix}^\top$.

Proof: Due to Theorem 2, we can restrict the solution space of (3) to the space of variational generalized spline functions. Let us now introduce basis functions $\varphi_j \in S(\{t_i\}_{i=0}^{N+1})$, $j = -1, \ldots, N+2$ for our variational generalized spline space, which can be determined by solving an interpolation problem. Here, a representation in the Lagrange-basis is adopted, where the functions φ_j are defined such that for $i = 0, \ldots, N + 1$ we have

$$\varphi_j(t_0) = \begin{cases} 1 & \text{if } -1 = j \\ 0 & \text{if } -1 \neq j \end{cases}, \quad \varphi_j(t_{N+1}) = \begin{cases} 1 & \text{if } N + 2 = j \\ 0 & \text{if } N + 2 \neq j \end{cases},$$
$$\dot{\varphi}_j(t_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \text{for } j = -1, \ldots, N + 2. \tag{7}$$

Then, the functions x and y can be parameterized as

$$x(t) = \sum_{j=-1}^{N+2} x_j \varphi_j(t), \quad y(t) = \sum_{j=-1}^{N+2} y_j \varphi_j(t), \tag{8}$$

where $x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{R}$ are unknown coefficients and represent the new decision variables. The coefficients $x_{-1}, x_0, x_{N+1}, x_{N+2}, y_{-1}, y_0, y_{N+1}, y_{N+2} \in \mathbb{R}$ are obtained by using the constructed basis functions (7) and evaluating the boundary values of x and y in (3), yielding

$$\begin{pmatrix} x_{-1} \\ y_{-1} \end{pmatrix} = \begin{pmatrix} x_s \\ y_s \end{pmatrix}, \qquad \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_s) \\ \sin(\theta_s) \end{pmatrix},$$
$$\begin{pmatrix} x_{N+2} \\ y_{N+2} \end{pmatrix} = \begin{pmatrix} x_f \\ y_f \end{pmatrix}, \qquad \begin{pmatrix} x_{N+1} \\ y_{N+1} \end{pmatrix} = v_p \begin{pmatrix} \cos(\theta_f) \\ \sin(\theta_f) \end{pmatrix}.$$

Next, we represent our discretized constraint in problem (3) in the basis (7), resulting in

$$|\dot{x}(t_i)|^2 + |\dot{y}(t_i)|^2 = \left| \sum_{j=-1}^{N+2} x_j \dot{\varphi}_j(t_i) \right|^2 + \left| \sum_{j=-1}^{N+2} y_j \dot{\varphi}_j(t_i) \right|^2$$
$$= x_i^2 + y_i^2 = v_p^2, \quad i = 1, \ldots, N. \tag{9}$$

In order to transform the cost functional, denote

$$x_b(t) := \sum_{k=-1}^{0} x_k \varphi_k(t) + \sum_{k=N+1}^{N+2} x_k \varphi_k(t),$$
$$y_b(t) := \sum_{k=-1}^{0} y_k \varphi_k(t) + \sum_{k=N+1}^{N+2} y_k \varphi_k(t).$$

Then, defining $Q \in \mathbb{R}^{2N+1 \times 2N+1}$ as

$$Q := \begin{pmatrix} r & b_1^\top & b_2^\top \\ b_1 & H & 0_{N \times N} \\ b_2 & 0_{N \times N} & H \end{pmatrix} \tag{10}$$

with $H = H^T \in \mathbb{R}^{N \times N}, b_1 \in \mathbb{R}^N, b_2 \in \mathbb{R}^N, r \in \mathbb{R}$, where

$$H_{i,j} := \int_\Omega \ddot{\varphi}_i(t) \ddot{\varphi}_j(t) \, dt, \quad r := \int_\Omega |\ddot{x}_b(t)|^2 + |\ddot{y}_b(t)|^2 \, dt$$

as well as

$$b_1 := \begin{pmatrix} \int_\Omega \ddot{x}_b(t) \ddot{\varphi}_1(t) \, dt \\ \vdots \\ \int_\Omega \ddot{x}_b(t) \ddot{\varphi}_N(t) \, dt \end{pmatrix}, \quad b_2 := \begin{pmatrix} \int_\Omega \ddot{y}_b(t) \ddot{\varphi}_1(t) \, dt \\ \vdots \\ \int_\Omega \ddot{y}_b(t) \ddot{\varphi}_N(t) \, dt \end{pmatrix}$$

yields $Q \succcurlyeq 0$ and

$$J(x, y) = \int_\Omega \left| \sum_{j=-1}^{N+2} x_j \ddot{\varphi}_j(t) \right|^2 + \left| \sum_{j=-1}^{N+2} y_j \ddot{\varphi}_j(t) \right|^2 \, dt$$
$$= \begin{pmatrix} 1 & x^\top & y^\top \end{pmatrix} Q \begin{pmatrix} 1 & x^\top & y^\top \end{pmatrix}^\top. \tag{11}$$

Let $z := \begin{pmatrix} x^\top & y^\top \end{pmatrix}^\top \in \mathbb{R}^{2N}$. Then, (11) corresponds to the cost function f and (9) to the constraint set $\mathcal{T}$ of (6). Since the constraint set $\mathcal{T}$ is diffeomorphic to a torus, a compact set in $\mathbb{R}^{2N}$, and since the function f is continuous, problem (6) admits a minimizer. Hence, problem (6) can be solved to obtain a solution $z = \begin{pmatrix} x_1 & \ldots & x_N & y_1 & \ldots & y_N \end{pmatrix}^\top$. By inserting the coefficients $x_1, \ldots, x_N, y_1, \ldots, y_N$ into (8) a solution for (3) can be recovered. ∎

Even though we were able to render problem (2) finite-dimensional by loosening the constraint (2a), the resulting problem (6) remains non-convex due to the non-convexity of the constraint set $\mathcal{T}$. Subsequently, we concentrate on solving (6) and discuss how to generate a (inexact) solution for (6) using convex relaxations based on the theory of moments. This allows us to retrieve an approximate solution for the original problem (2) through the parameterization (8).

IV. SEMIDEFINITE RELAXATIONS OF THE QCQP

As we are interested in global solutions of (6), as opposed to local solutions, we turn to the technique of constructing hierarchies of semidefinite relaxations based on the theory of moments and the dual theory of sum-of-squares. These types of semidefinite relaxations mainly originate from the works of Lasserre [12], Nesterov [21], Parrilo [22], [23] and Shor [27]. In our setup we require the availability of (inexact) solutions for (6) in order to generate suboptimal trajectories for (2). Below, we consider two procedures to construct (inexact) solutions for (6). Both of these methods involve solving an SDP from the moment perspective. Thus, we

follow Lasserre [12] and consider the sequence of moment relaxations of (6) given by

$$f_d^{mom} := \inf_y \quad L_y(f) \tag{12}$$
$$\text{subject to} \quad M_d(y) \succcurlyeq 0,$$
$$y_0 = 1,$$
$$M_{d-1}(h_i y) = 0, \quad i = 1, \dots, N,$$

with relaxation order $d > 0$, where L_y denotes the Riesz linear functional, $M_d(y)$ the moment matrix and $M_{d-1}(h_i y)$ the localizing matrix. Each successive value f_d^{mom} improves upon the optimum value f^* and can be obtained by solving an SDP. Since the considered constraint set $\mathcal{T}$ satisfies the so-called Archimedean property, we can establish asymptotic convergence, i.e., $\lim_{d \to \infty} f_d = f^*$ due to Theorem 2.9 in [12]. However, note that the computational burden grows significantly as the relaxation order d is increased.

In the sequel, we consider the first and second relaxation in the hierarchy (12). Using the first-order moment relaxation, we can always retrieve inexact solutions to (6) through a randomized projection algorithm inspired by the approximation scheme for the maximum cut (MAXCUT) problem [9]. On the other hand, in all cases considered numerically, see Section V-C, we obtain an optimality certificate through the second-order moment relaxation, meaning that this relaxation is an exact convex formulation of (6). For this reason, higher relaxation orders are not considered in this work.

A. First-order moment relaxation

The first moment relaxation in hierarchy (12) is also referred to as standard SDP relaxation and is equivalent to the dual of the well-known Shor relaxation for QCQPs. For (6) this relaxation can be written as

$$\inf_{Z \in \mathbb{R}^{2N+1 \times 2N+1}} \quad \text{trace}(QZ) \tag{13}$$
$$\text{subject to} \quad Z \succcurlyeq 0,$$
$$Z_{1,1} = 1,$$
$$Z_{i,i} + Z_{i+N,i+N} = v_p^2, \quad i = 2, \dots, N+1.$$

If a solution Z^* of (13) has rank one, then this relaxation is lossless and a solution to (6) can be extracted from the first column of Z^*.

Whenever $\text{rank}(Z^*) > 1$, the above relaxation is inexact and we do not directly recover an optimal or even feasible point for (6). For this purpose, we propose a randomized approximation algorithm, based on the MAXCUT algorithm [9], to produce an inexact solution for (6). Like the MAXCUT algorithm, our algorithm exploits that any positive semidefinite matrix is admissible for a sequence of second order moments of a measure. As a consequence, if the matrix the matrix Z^* is partitioned as

$$Z^* = \begin{pmatrix} 1 & \hat{z}^\top \\ \hat{z} & \hat{Z} \end{pmatrix}, \tag{14}$$

then, the matrix $\hat{Z} - \hat{z}\hat{z}^\top$ can be interpreted as the covariance matrix for a multivariate normal distribution. This can be

used to generate a candidate point ζ from this normal distribution, i.e., $\zeta \sim \mathcal{N}(\hat{z}, \hat{Z} - \hat{z}\hat{z}^\top)$. By projecting this candidate onto the set $\mathcal{T}$ in some way, a feasible point for (6) can be produced. In contrast to the MAXCUT algorithm, our algorithm, given in Algorithm 1, performs a different projection step. More specifically, we exploit that $\mathcal{T}$ is diffeomorphic to a torus and project onto this set. The reason for this modification is that for the considered setup, we have observed through numerical experiments that Algorithm 1 yields feasible points associated with a lower objective value when compared to the MAXCUT algorithm. However, we could not establish a performance guarantee for (6) with either algorithm.

Algorithm 1 Approximation algorithm for (6)

Require: Z^* as in (14), obtained by solving (13)
  for $k = 1 : 2000$ **do**   ▷ Heuristically determined sample size
    Draw a candidate $\zeta \sim \mathcal{N}(\hat{z}, \hat{Z} - \hat{z}\hat{z}^\top)$
    for $i = 1 : N$ **do**   ▷ Project ζ onto $\mathcal{T}$
      $z_i^a = v_p \frac{\zeta_i}{\sqrt{\zeta_i^2 + \zeta_{i+N}^2}}, \quad z_{i+N}^a = v_p \frac{\zeta_{i+N}}{\sqrt{\zeta_i^2 + \zeta_{i+N}^2}}$
    end for
  end for
  Choose the vector from $((z^a)_{k=1}^{2000})$ which is associated with the lowest objective value.

B. Second-order moment relaxation

Let us now consider the next relaxation in hierarchy (12), i.e., the second-order moment relaxation. As supported by numerical results, see Section V-C, we conjecture the following statement and emphasize that this conjecture does not rely on the concrete structure of the cost function. In other words, this conjecture includes problem (6) as a special case.

Conjecture 1: Let p be a quadratic polynomial with real coefficients in the variable z. Then, the second-order moment relaxation of

$$p^* := \inf_{z \in \mathcal{T}} \ p(z) \tag{15}$$

is lossless

Remark 2: One could also attempt to prove Conjecture 1 from the sum-of-squares perspective, in which case it needs to be shown that for real-valued quadratic polynomials $(\psi_i(z))_{i=1}^N$ we have that $p(z) - p^* - \sum_{i=1}^N \psi_i(z)h_i(z)$ has a sum-of-squares decomposition whose degree is at most 4. Another approach which might help to prove Conjecture 1 is to replace the ordinary polynomials in the sum-of-squares decomposition by trigonometric polynomials, e.g., as in [19].

By implementing the linear algebra routine devised in [10], we can numerically determine whether the second-order moment relaxation of (15) is tight and also extract some global minimizer for (15). This optimality certificate relies on checking a certain rank condition of the moment matrix. We present a convincing numerical indicator in Section V-C to support Conjecture 1, especially concerning the specific QCQP (6). We note that numerically examining this conjecture involves considering all natural numbers for the parameter N as well as gridding the space of quadratic polynomials and the parameter v_p. For the particular case (6),

this translates into varying the number of samples N and all model parameters consisting of the velocity v_p as well as the initial and final values $x_s, y_s, \theta_s, x_f, y_f, \theta_f, t_f$.

V. NUMERICAL RESULTS

This section presents some numerical results for the trajectory planning problem (2) as well as a numerical indicator for the validity of Conjecture 1. The involved SDPs were solved using the parser Yalmip [15] and the solver Mosek [2].

We generate an approximate solution for (2) by constructing a (inexact) solution for (3), which translates into problem (6) by using the parameterization (8). In case the first-order moment relaxation (13) is lossless, we obtain an optimal trajectory to (3) for a given N. Otherwise, either Algorithm 1 is used to generate a suboptimal trajectory for (3) or, the second-order moment relaxation is solved, which always yielded an optimal trajectory for (3) in our simulation studies through the algorithm presented in [10].

To numerically construct an approximate solution for (2) using our approach, the discretization parameter N as well as the model parameters $v_p, x_s, y_s, \theta_s, x_f, y_f, \theta_f, t_f$ need to be fixed. We note that for a given set of model parameters, the exactness of the moment relaxations (12) seems to be independent of N when $N > 5$. In particular, if t_f, x_s, y_s, x_f, y_f are held constant, we have observed that the exactness of the first-order moment relaxation depends on the parameter values v_p, θ_s, θ_f. This aspect is further elaborated on throughout this section. In the following simulation results, the constraint (2a) is sampled equidistantly. Moreover, we fix the final time to $t_f = 1$ as well as the initial and final position to $(x_s, y_s) = (1, -1)$ and $(x_f, y_f) = (0, 0)$, respectively. The values of the horizontal speed v_p, initial angle θ_s, final angle θ_f and the number of samples N are varied depending on the specific simulation goal. More specifically, we vary the values of v_p, θ_s, θ_f to examine the exactness of the first- and second-order moment relaxation and N to investigate the empirical convergence behavior as discussed next.

A. On the empirical convergence of solutions to problem (3)

We emphasize that the only difference between (2) and (3) lies in the discretization of constraint (2a). Hence, by reducing the mesh size $\frac{t_f}{N-1}$, i.e., by increasing N, we expect a better approximation of an optimal trajectory for (3) to an optimal trajectory for (2). This means that for increasing N, the constraint (2a) should be violated less in-between the sampling instances and the trajectory's shape should appear smoother. Indeed, if (13) is lossless, we have observed that values already between $N = 10$ and $N = 30$ yield satisfactory results. This is exemplified in Fig. 1 which depicts trajectories resulting from an exact first-order moment relaxation for increasing N in (a), using the parameter values $v_p = 4, \theta_s = 0, \theta_f = \frac{3\pi}{2}$. The corresponding velocity functions $v(t)$ are shown in (c), illustrating the satisfaction of constraint (2a). Magnified portions of (a) and (c) are shown in (b) and (d), respectively. Starting with $N = 30$, the trajectories resulting from larger N can barely be distinguished from each other. In addition, the constraint (2a) is violated only by a negligible discretization error.

(a) All trajectories.

(b) Last 4 trajectories magnified.

(c) Associated velocity functions.

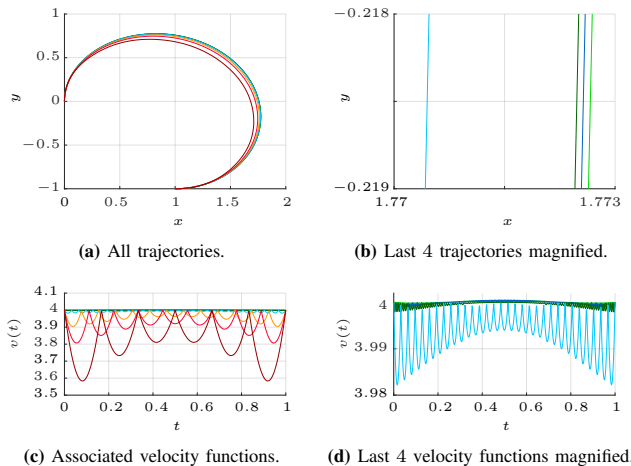**(d)** Last 4 velocity functions magnified.

Fig. 1: Optimal trajectories for (3) with their associated velocity $v(t)$ using the parameter values $v_p = 4, \theta_s = 0, \theta_f = \frac{3\pi}{2}$ and (▬): $N = 5$, (▬): $N = 8$, (▬): $N = 12$, (▬): $N = 30$, (▬): $N = 100$, (▬): $N = 150$, (▬): $N = 250$.

If the standard SDP relaxation is not tight and Algorithm 1 is employed, this discretization error is larger. Depending on the particular parameter configuration, the number of samples N can be increased to reduce this error. Still, in selected cases a significant error persists which is also apparent in the corresponding non-smooth appearing trajectory. To mitigate this problem, the second-order moment relaxation can be solved instead, albeit at the expense of an increased computational cost. Nevertheless, comparatively low values of N are sufficient to improve the trajectory's quality. Unfortunately, due to this high computational burden, the convergence behavior resulting from the second-order moment relaxation could not be studied thoroughly. However, for small N, it could be observed that trajectories resulting from the second-order moment relaxation display similar properties compared to trajectories stemming from an exact first-order moment relaxation. On the other hand, a slightly larger error in the satisfaction of constraint (2a) in-between the sampling instances could be noticed. Hence, we expect the convergence rate for these trajectories to be comparable but slightly slower.

Remark 3: The problem size of (12) for $d = 2$ can be reduced by eliminating variables through the constraints $h_i(z) = 0, i = 1, \ldots, N$. For instance, Yalmip translates the second-order moment relaxation for $N = 10$ into a problem with $10\,625$ constraints, 2310 scalar variables and one 231×231 matrix variable, if the equality constraints are not exploited. In contrast, by eliminating decision variables through the given constraints, the resulting problem reduces to 8360 constraints and one 231×231 matrix variable. While certain decision variables in the moment matrix can be eliminated in this way, we could not reduce the size of the moment matrix without sacrificing exactness of the second-order moment relaxation for some parameter configurations.

B. Trajectory generation for problem (2)

For the simulation result depicted in Fig. 2, we set $v_p = 4$, $\theta_s = 0$ and $N = 12$, while θ_f is varied. We note that a symmetric behavior is observed if θ_f is held constant

while θ_s is varied, i.e., it does not matter whether θ_f is held fixed while varying θ_s or vice versa. This is coherent with our problem setup since any trajectory generated for (1) is undirected. Further, we note that the value of the parameter v_p can have an impact on the exactness of the first-order moment relaxation. More precisely, for certain fixed parameters θ_s, θ_f, N, it is possible to induce an inexact first-order moment relaxation to become exact by increasing v_p.

Each (a) and (b) in Fig. 2 display a trajectory produced by Algorithm 1, and an optimal trajectory for (3) with $\theta_f = 0$ and $\theta_f = \frac{\pi}{4}$, respectively. This optimal trajectory is obtained through the second-order moment relaxation combined with the algorithm from [10]. The associated velocity functions are depicted in (c) and (d). In the configuration $\theta_f = \theta_s = 0$ there exists an additional optimal trajectory which is not plotted in (a) for the purpose of clarity. As can be inferred from this figure, the shape of the suboptimal trajectory with $\theta_f = \theta_s$ deviates more from the optimal trajectory compared to the case where $\theta_f \neq \theta_s$. Indeed, the cost value associated with the suboptimal trajectory drifts further apart from the optimal value of (6) as the configuration $\theta_f = \theta_s$ is approached. We note that throughout our simulations, we have never identified a case for which the first-order moment relaxation is exact if the values of θ_s and θ_f coincide, even if v_p is increased. In fact, it seems as if the rank of a solution Z^* to (13) becomes more "unstable" as θ_s and θ_f approach each other, i.e., as $\operatorname{rank}(Z^*) = 1$ transitions to $\operatorname{rank}(Z^*) = 3$.

(a) Trajectories using $\theta_f = 0$.

(b) Trajectories using $\theta_f = \frac{\pi}{4}$.

(c) Constraint satisfaction (2a) of (a).
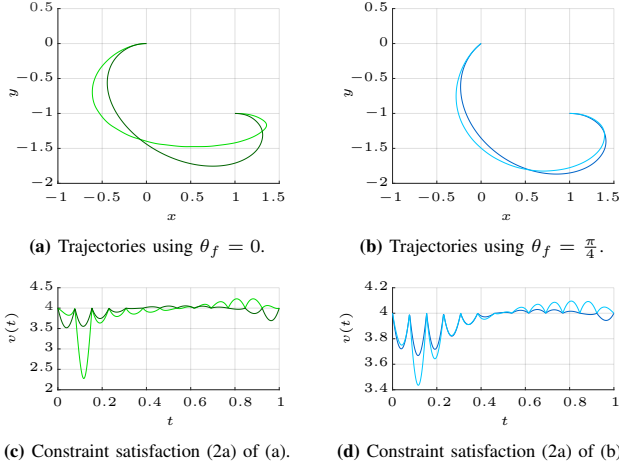
(d) Constraint satisfaction (2a) of (b).

Fig. 2: Trajectories generated for (3) with their associated velocity $v(t)$ using the parameter values $N = 12$, $v_p = 4$ and $\theta_s = 0$. (▬), (▬) correspond to optimal trajectories for (3) and (▬), (▬) to suboptimal trajectories produced by Algorithm 1.

C. Exactness of the second-order moment relaxation

This subsection presents the results of two simulation studies dedicated to showing that the second-order moment relaxation of (6) or (15) is sufficient to exactly solve these two problems. To render the simulation time tolerable, while retaining an adequate dimension of the sample space $\mathbb{R}^{2N}$, the number of samples of constraint (2a) is set to $N = 5$ for the larger study and to $N = 7$ for the smaller study. While larger values of N were not considered, we expect

that the exactness of the second-order moment relaxation of (6) and (15) is maintained.

The larger simulation study focuses on our specific problem setup (6). For this simulation we proceeded as follows.

We equidistantly sample the parameter space $(\theta_s, \theta_f, v_p) \in \mathbb{R}^3$, where $\theta_s, \theta_f \in [0, 2\pi]$ and $v_p \in [3, 6]$. We iterate through each parameter configuration and check whether a solution Z^* of (13) is of rank one. If $\operatorname{rank}(Z^*) > 1$, the second-order moment relaxation, i.e., (12) for $d = 2$ is solved. In total, this procedure was conducted 10^6 times. In about 74.91% of the time, (13) was lossless whereas in $24.85\% + 0.24\%$ of the time we established that $\operatorname{rank}(Z^*) = 3$ such that the second-order moment relaxation, was solved. More specifically, the 24.85% signify that the second-order moment relaxation yielded a global minimizer for (6) while in the remaining 0.24% numerical problems occurred in solving this SDP. We note that for increasing values of v_p, numerical issues in solving (12) for $d = 2$ emerged more frequently. This could not be observed for $d = 1$.

In an effort to prove the exactness of (12) with $d = 2$ for (6), we noticed that the particular structure of the cost function, as determined by Q in (10), may not be relevant for this phenomenon. This is supported by the smaller simulation study which we discuss in the following. First, we express each quadratic polynomial p from Conjecture 1 as $p(z) = \begin{pmatrix} 1 & z^\top \end{pmatrix} A \begin{pmatrix} 1 & z^\top \end{pmatrix}^\top$ and randomly draw the entries of the coefficient matrix $A = A^\top$ from the uniform distribution in the interval $[-T, T]$, with different values for $T \in \mathbb{R}$. Then, for varying v_p and T, the second-order moment relaxation is solved in case the first-order moment relaxation did not yield an optimal point for (15). This procedure was performed 10^5 times in total with different parameter choices of T and v_p, which are specified in Table I. In this study, we decided to conduct more trials for lower values v_p since we have observed that higher values of v_p typically tighten the first-order moment relaxation for some fixed A. Therefore, if lower values of v_p induce an exact second-order moment relaxation, we also expect this relaxation to be exact for higher values of v_p. Otherwise, the chosen parameter values and number of trails have not been chosen according to any system. For a more convincing result in this setup, covering a wider range of parameter configurations, more trials would need to be performed. In almost all cases, specified in Table I, we obtained an optimality certificate by recovering a global minimizer for (15). Only negligible selected cases could be identified for which a minimizer with the algorithm from [10] could not be constructed. In these cases, however, the first- and second-order moment relaxation had the same optimal value, indicating that the global optimum of (15) is attained.

v_p \ T	1	10	100	5000	70 000	500 000
$\sqrt{7}$	10 000	16 000	11 000	3000	1500	6200
4	3000	21 000	7000	2000	5000	4000
$\sqrt{70}$	800	400	1600	1400	300	500
10	500	2000	500	1300	700	300

TABLE I: Number of times the first- or second-order moment relaxation of (15) is solved, partitioned according to different parameter combinations.

VI. CONCLUSIONS

We have proposed a novel approach for convexifying the minimum energy problem for the unicycle model through semidefinite moment relaxations. As has been supported by numerical results, the constructed approximate solutions are highly satisfactory. Beyond that, we conjecture that the second-order moment relaxation of minimizing a quadratic polynomial over an n-dimensional torus represents a tight and tractable convex formulation.

Possible future work includes incorporating constraints in the problem setup, e.g., hard input constraints. Furthermore, this approach could be applied to other optimal control problems, e.g., the Brockett integrator. In this context it should be considered that with the proposed approach, an infinite-dimensional optimization problem with more decision variables directly leads to a finite-dimensional optimization problem with more decision variables and thus, we are faced with a higher computational burden. In this case, it may be possible to exploit symmetries or further sparsities in the corresponding semidefinite relaxation, see, e.g., [26] or [29].

REFERENCES

[1] B. Açıkmeşe and S. R. Ploen. Convex Programming Approach to Powered Descent Guidance for Mars Landing. *Journal of Guidance, Control, and Dynamics*, 30(5):1353–1366, 2007.

[2] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.3.*, 2022.

[3] W. Arendt and K. Urban. *Partial Differential Equations*. Graduate Texts in Mathematics. Springer Cham, 1 edition, 2023.

[4] K. Atkinson and W. Han. *Theoretical Numerical Analysis: A Functional Analysis Framework*. Texts in Applied Mathematics. Springer New York, 2009.

[5] H. Bauer. *Probability Theory*. De Gruyter studies in mathematics. Walter de Gruyter, 1996.

[6] E. Berthier, J. Carpentier, A. Rudi, and F. Bach. Infinite-Dimensional Sums-of-Squares for Optimal Control. arXiv:2110.07396, 2021.

[7] L. Blackmore, B. Açıkmeşe, and J. M. Carson III. Lossless Convexification of Control Constraints for a Class of Nonlinear Optimal Control Problems. In *2012 American Control Conference (ACC)*, pages 5519–5525. IEEE, 2012.

[8] A. Figueroa-González, F. Cacciatore, and R. Haya-Ramos. Landing Guidance Strategy of Space Rider. *Journal of Spacecraft and Rockets*, 58(4):1220–1231, 2021.

[9] M. X. Goemans and D. P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[10] D. Henrion and J. B. Lasserre. Detecting Global Optimality and Extracting Solutions in GloptiPoly. In D. Henrion and A. Garulli, editors, *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Sciences*, pages 293–310. Springer, Berlin, Heidelberg, 2005.

[11] J. M. Carson III, M. M. Munk, R. R. Sostaric, J. N. Estes, F. Amzajerdian, J. B. Blair, D. K. Rutishauser, C. I. Restrepo, A. M. Dwyer-Cianciolo, G. T. Chen, and T. Tse. The SPLICE Project: Continuing NASA Development of GN&C Technologies for Safe and Precise Landing. In *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, 2019.

[12] J. B. Lasserre. Global Optimization with Polynomials and The Problem of Moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[13] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear Optimal Control via Occupation Measures and LMI-Relaxations. *SIAM Journal on Control and Optimization*, 47(4):1643–1666, 2008.

[14] X. Liu, Z. Shen, and P. Lu. Entry Trajectory Optimization by Second-Order Cone Programming. *Journal of Guidance, Control, and Dynamics*, 39(2):227–241, 2015.

[15] J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *IEEE International Conference on Robotics and Automation*, pages 284–289, 2004.

[16] P. Lu and X. Liu. Autonomous Trajectory Planning for Rendezvous and Proximity Operations by Conic Optimization. *Journal of Guidance, Control, and Dynamics*, 36(2):375–389, 2013.

[17] Y. Mao, M. Szmuk, and B. Açıkmeşe. Successive Convexification of Non-Convex Optimal Control Problems and Its Convergence Properties. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3636–3641. IEEE, 2016.

[18] Y. Mao, M. Szmuk, X. Xu, and B. Açıkmeşe. Successive Convexification: A Superlinearly Convergent Algorithm for Non-convex Optimal Control Problems. arXiv:1804.06539, 2018.

[19] A. Megretski. Positivity of Trigonometric Polynomials. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, volume 4, pages 3814–3817, 2003.

[20] G. Micula. A variational approach to spline functions theory. *General Mathematics*, 10:21–50, 2002.

[21] Y. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, Applied Optimization (APOP, volume 33), pages 405–440. Springer Boston, 2000.

[22] P. A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness*. PhD thesis, California Institute of Technology, 2000.

[23] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96:293–320, 2003.

[24] R. M. Pinson and P. Lu. Trajectory Design Employing Convex Optimization for Landing on Irregularly Shaped Asteroids. *Journal of Guidance, Control, and Dynamics*, 41(6):1243–1256, 2018.

[25] P. M. Prenter. *Splines and Variational Methods*. John Wiley & Sons, 1989.

[26] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre. Exploiting Symmetries in SDP-Relaxations for Polynomial Optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013.

[27] N. Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23:695–700, 1987.

[28] M. Szmuk, T. P. Reynolds, and B. Açıkmeşe. Successive Convexification for Real-Time Six-Degree-of-Freedom Powered Descent Guidance with State-Triggered Constraints. *Journal of Guidance, Control, and Dynamics*, 43(8):1399–1413, 2020.

[29] J. Wang, V. Magron, J. B. Lasserre, and N. H. A. Mai. CS-TSSOS: Correlative and Term Sparsity for Large-Scale Polynomial Optimization. *ACM Trans. Math. Softw.*, 48(4), 2022.

APPENDIX

A. Proof of Theorem 1

To infer that a solution to (2) can be obtained from a solution of (1), we show that a solution of (2) satisfies the constraints in problem (2) and that the cost functionals are equal up to a constant. By inserting a solution of (1) into the constraints of problem (2), we readily conclude that the constraints are satisfied. Due to the relation

$$\ddot{x}(t)^2 + \ddot{y}(t)^2 = \left(\frac{\mathrm{d}}{\mathrm{d}t}\cos(\theta(t))v_p\right)^2 + \left(\frac{\mathrm{d}}{\mathrm{d}t}\sin(\theta(t))v_p\right)^2$$
$$= v_p^2\dot{\theta}(t)^2 = v_p^2 u(t)^2, \qquad t \in \Omega, \qquad (16)$$

we establish that the cost functional in (2) is equivalent to the one in (1), where v_p can be omitted since it is a constant.

It remains to show that, given a solution $x, y \in H^2(\Omega) \subset C^1(\bar{\Omega})$ of (2), there exists some $\theta \in H^1(\Omega)$ satisfying the constraints in problem (1). In the following, we assume w.l.o.g. that the constraint (2a) is scaled such that $\dot{x}(t)^2 + \dot{y}(t)^2 = 1$ for $t \in \bar{\Omega}$. Let $u, v : \mathbb{R} \to \mathbb{R}$ be given by $u(t) = \dot{x}(t)$, $v(t) = \dot{y}(t)$ for $t \in \bar{\Omega} = [0, t_f]$ and

$$u(t) = \dot{x}(0) \text{ for } t < 0, \ u(t) = \dot{x}(t_f) \text{ for } t > t_f,$$
$$v(t) = \dot{y}(0) \text{ for } t < 0, \ v(t) = \dot{y}(t_f) \text{ for } t > t_f.$$

Then, the mapping $\mathbb{R} \to \mathbb{R}^2$ given by $t \mapsto (u(t), v(t))$ into the unit circle is continuous. By Theorem 26.2 from [5], there exists some $\theta \in C(\mathbb{R})$ with $u(t) = \cos(\theta(t))$ and $v(t) = \sin(\theta(t))$ for $t \in \mathbb{R}$. It remains to show that $\theta \in H^1(\Omega)$.

First note that for some $a < 0$ and $b > t_f$ with $[0, t_f] \subset (a, b)$ we have that $u, v \in C([a, b])$, $u|_{(a,0)}, v|_{(a,0)} \in H^1(a, 0)$, $u|_{(0,t_f)}, v|_{(0,t_f)} \in H^1(0, t_f)$, $u|_{(t_f,b)}, v|_{(t_f,b)} \in H^1(t_f, b)$. By Theorem 5.14 in [3] we infer $u, v \in H^1(a, b)$.

Next, we show that for any $t_0 \in [0, t_f]$ there exists some $\epsilon > 0$ with $I_{t_0} := (t_0 - \epsilon, t_0 + \epsilon) \subset (a, b)$ and $\theta \in H^1(I_{t_0})$. Assume $\cos(\theta(t_0)) \neq 0$ and let $f \in C^1(\mathbb{R})$ be given by $\theta \mapsto \sin(\theta)$. Then, the derivative of f does not vanish at $\theta(t_0)$. By the inverse function theorem, there exists some $\delta > 0$ such that $f : U \to V$ for $U := (t_0 - \delta_0, t_0 + \delta_0)$ and $V := f(U)$ is a diffeomorphism with inverse $g : V \to U$. If $\theta(t) \in U$, then $f(\theta(t)) = v(t)$ and thus $\theta(t) = g(v(t))$. Since $\theta \in C(\mathbb{R})$, we can pick some $\epsilon > 0$ with $\theta(I_{t_0}) \subset U$ such that $\theta(t) = g(v(t))$ for $t \in I_{t_0}$. Due to our assumption, we can ensure that the derivative of g is bounded. Then, by applying the chain rule from Proposition 7.1.12 in [4] we conclude that $\theta \in H^1(I_{t_0})$ and $\dot{\theta}(t) = g'(v(t))\dot{v}(t)$ for almost all $t \in I_{t_0}$. Whenever $\cos(\theta(t_0)) = 0$, then $\sin(\theta(t_0)) \neq 0$ and we pick an inverse of the map $\theta \mapsto \cos(\theta)$ instead.

For every $t \in \bar{\Omega}$, we can hence construct an interval $I_t := (t - \epsilon, t + \epsilon)$ with some $\epsilon > 0$ as above. Let $I_{\tilde{t}_1} \cup \cdots \cup I_{\tilde{t}_n}$ be a sub-covering of the open-covering $\cup_{t \in \bar{\Omega}} I_t$ of $\bar{\Omega}$. Then, we construct a partition $0 = t_0 < t_1 < \cdots < t_m = t_f$ of $\bar{\Omega}$ such that for any $j = 0, \ldots, m-1$, there exists some $k = 1, \ldots, n$ with $(t_j, t_{j+1}) \subset I_{\tilde{t}_k}$. We then infer $\theta \in H^1(t_j, t_{j+1})$, $j = 0, \ldots, m - 1$. Again by Theorem 5.14 in [3], we conclude that $\theta \in H^1(\Omega)$ as was to be shown. ∎

B. Some properties of the space $S(\{t_i\}_{i=0}^n)$

In the following, let the interval $I = (a, b)$ and its partition be as in Definition 1 and denote $I_i := (t_{i-1}, t_i)$ for $i = 1, \ldots, n$. Further, let $\alpha_0, \alpha_1, \beta_i \in \mathbb{R}$, $i = 0, \ldots, n$ in (4) be fixed. We introduce some additional notation. The space of infinitely differentiable functions with compact support on I is denoted as $C_0^\infty(I)$. We denote the closure of $C_0^\infty(I)$ by $H_0^{2,2}(I)$. Lastly, we define $|s|_{L_2(I)}^2 := \int_I |s(t)|^2 \, dt$ and $|s|_{H^2(I)}^2 := \int_I |s(t)|^2 + |\dot{s}(t)|^2 + |\ddot{s}(t)|^2 \, dt$.

Before we show that any $s \in S(\{t_i\}_{i=0}^n)$ has to satisfy (5), we need to establish the following lemma.

Lemma 1: To any feasible $s \in H^2(I)$ of (4), we can pair a feasible $s_1 \in H^2(I)$ satisfying (5) such that $J(s_1) \leq J(s)$.

Proof: We begin the proof by choosing s_1 such that

$$s_1(t_i) = s(t_i), \quad i = 1, \ldots, n-1. \tag{17}$$

Note that

$$0 \leq \int_I |\ddot{s}(t) - \ddot{s}_1(t)|^2 \, dt$$
$$= \int_I |\ddot{s}(t)|^2 - 2\ddot{s}(t)\ddot{s}_1(t) + 2|\ddot{s}_1(t)|^2 - |\ddot{s}_1(t)|^2 \, dt$$
$$= J(s) - J(s_1) - 2 \sum_{i=1}^n \int_{I_i} (\ddot{s}(t) - \ddot{s}_1(t)) \, \ddot{s}_1(t) \, dt. \tag{18}$$

Now, let I_i be fixed but arbitrary. Integrating by parts twice combined with (4a), (4b), (5) and (17) yields

$$\int_{I_i} (\ddot{s}(t) - \ddot{s}_1(t)) \, \ddot{s}_1(t) \, dt$$
$$= \lim_{\epsilon \to 0} \left([(\dot{s}(t) - \dot{s}_1(t)) \ddot{s}_1|_{I_i}(t)]_{t_{i-1}+\epsilon}^{t_i - \epsilon} \right)$$
$$\quad - \int_{I_i} (\dot{s}(t) - \dot{s}_1(t)) \frac{d^3}{dt^3}(s_1|_{I_i}(t)) \, dt$$
$$= \int_{I_i} (s(t) - s_1(t)) \frac{d^4}{dt^4}(s_1|_{I_i}(t)) \, dt$$
$$\quad - \lim_{\epsilon \to 0} \left(\left[(s(t) - s_1(t)) \frac{d^3}{dt^3}(s_1|_{I_i}(t)) \right]_{t_{i-1}+\epsilon}^{t_i - \epsilon} \right) = 0,$$

where $\ddot{s}|_{I_i}(t) := \frac{d^2}{dt^2} s|_{I_i}(t)$ and we take the limit $\epsilon \to 0$ since the second derivative of s and s_1 might be discontinuous. As I_i was chosen arbitrarily, (18) shows $J(s_1) \leq J(s)$. ∎

Equipped with this lemma, we can show that any $s \in S(\{t_i\}_{i=0}^n)$ is a cubic polynomial on I_i.

Lemma 2: The space $S(\{t_i\}_{i=0}^n)$ is finite-dimensional and linear. Specifically, any $s \in S(\{t_i\}_{i=0}^n)$ satisfies (5).

Proof: The strong form of the cost functional in problem (4) is given by (5), which can be derived through the Euler-Lagrange equations. We argue now that any solution to (4) needs to satisfy (5). To this end, let I_i be fixed but arbitrary. Further, let $s_1 \in H^2(I)$ be some solution for (4). Due to Lemma 1 we can construct another solution $s_2 \in H^2(I)$ to (4) satisfying (5) such that

$$s_1(t_{i-1}) = s_2(t_{i-1}), \quad s_1(t_i) = s_2(t_i). \tag{19}$$

Pick arbitrary but fixed admissible variations $\eta_{s_1}, \eta_{s_2} \in H_0^2(I_i)$. Then, with $J^i(s) := \int_{I_i} |\ddot{s}(t)|^2 \, dt$, we have

$$\frac{d}{d\epsilon} J^i(s_1 + \epsilon\eta_{s_1}) \Big|_{\epsilon=0} = 0 \tag{20}$$

and $\frac{d}{d\epsilon} J^i(s_2 + \epsilon\eta_{s_2}) \big|_{\epsilon=0} = 0$. Thus,

$$\int_{I_i} \ddot{s}_1(t)\ddot{\eta}_{s_1}(t) \, dt = \int_{I_i} \ddot{s}_2(t)\ddot{\eta}_{s_2}(t) \, dt = 0. \tag{21}$$

Now, consider the particular choices $\eta_{s_1} = s_1 - s_2 \in H_0^2(I_i)$ and $\eta_{s_2} = s_2 - s_1 \in H_0^2(I_i)$. Then, (21) implies

$$\int_{I_i} \ddot{s}_1(t)\ddot{\eta}_{s_1}(t) + \ddot{s}_2(t)\ddot{\eta}_{s_2}(t) \, dt = |\ddot{s}_1 - \ddot{s}_2|_{L_2(I_i)}^2 = 0.$$

Since $|\ddot{s}_1|_{L_2(I_i)}$ and $|s_1|_{H^2(I_i)}$ are equivalent in $H_0^{2,2}(I_i)$, this implies $|s_1 - s_2|_{H^2(I_i)} = 0$. Hence, a weak solution to (20) is unique in $H^2(I_i)$. A classical solution to (20) is given by a function satisfying (5). Due to the condition (19) and since a classical solution is also a weak solution, and therefore unique, every solution of (4) has to satisfy (5).

Consequently, any minimizing function for (4) is a cubic polynomial on I_i satisfying (4a) and (4b). Since the space of polynomials of fixed order is a finite-dimensional linear space, only finitely many unknowns need to be determined in order to characterize the space $S(\{t_i\}_{i=0}^n)$. ∎