IEEE Transactions of Cybernetics, Volume 49, Issue 2, February 2019, pp. 467-480
DOI:10.1109/TCYB.2017.2778195

Trajectory Optimization of Space Maneuver Vehicle Using a Hybrid Optimal Control Solver

Runqi Chai, *Student Member, IEEE,* Al Savvaris, *Member, IEEE,* Antonios Tsourdos, *Member, IEEE,* Senchun Chai, and Yuanqing Xia, *Senior Member, IEEE,*

***Abstract*—In the paper, a constrained Space Maneuver Vehicles (SMV) trajectory optimization problem is formulated and solved using a new three-layer-hybrid optimal control solver. To decrease the sensitivity of the initial guess and enhance the stability of the algorithm, an initial guess generator based on a specific stochastic algorithm is applied. In addition, an improved gradient-based algorithm is used as the inner solver, which can offer the user more flexibility to control the optimization process. Furthermore, in order to analyze the quality of the solution, the optimality verification conditions are derived. Numerical simulations were carried out by using the proposed hybrid solver and the results indicate that the proposed strategy can have better performance in terms of convergence speed and convergence ability when compared with other typical optimal control solvers. A monto-carlo simulation was performed and the results show a robust performance of the proposed algorithm in dispersed conditions.**

***Index Terms*—Space Maneuver Vehicles, trajectory optimization, optimal control, initial guess, improved gradient-based algorithm, optimality verification.**

I. Introduction

SPACECRAFT trajectories planning is usually recognized as an optimal control problem [1]–[4]. Due to the high nonlinear nature and strict path constraints, direct transcription algorithms are commonly used to approximate the optimal solution [5], [6]. Direct transcription algorithms [7] can be divided into two major schemes: direct shooting scheme [8], in which only the control parameters are discretized at temporal nodes; and direct collocation scheme [9], where both the state and control are parameterized. Many both theoretical and experimental works have been carried out in this field. For example, Yakimenko et al. [10] applied an Inverse Dynamics in the Virtual Domain (IDVD) collocation method to generate a near-optimal aircraft trajectory. The work of Albin and Ritter [11] presents a multiple-shooting-based model predictive control scheme for generating the optimal control sequence of a two-stage turbocharged gasoline airpath problem. In [12], considering the model physical constraints, an energy-optimal control trajectory was generated for a wave energy converter based on a specific collocation method. In addition, Benson et al. [7] developed a Gauss pseudospectral (orthogonal collocation) method for transcribing general optimal control problems. In their latest works, a hp-adaptive strategy was embedded in the framework of the original solver [13].

Normally, after the discretization process, the resulting Nonlinear Programming problem (NLP) can be solved numerically via well-developed algorithms such as the Sequential Quadratic Programming (SQP) [14] or Interior Point method (IP) [15]. However, the main challenge for gradient-based algorithms is that an initial guess value must be provided to the NLP solver. For the user to provide an initial guess is usually problematic since the guess sequences might be far from the optimal solution or close to a local optimal solution. It is likely for the optimizer to converge to a local solution or a neighborhood of the initial guess. Using optimal control software such as GPOPS or ICLOCS, this process usually made by interpolating the boundary conditions provided by the user [16]. However, due to the lack of the physical knowledge, the initial guess value may not be physically meaningful and it can hardly satisfy the dynamic equations and path constraints. Therefore, the NLP solver may start at an infeasible point where most of the constraints cannot be satisfied.

In recent years, evolutionary-based algorithms have become popular and have been widely investigated for solving trajectory optimization problems [6]. The main advantages of using evolutionary algorithms are that it is simple to understand and easy to implement. Furthermore, it is more likely than other methods to locate the global optimum solution. Contributions made to apply this type of methods can be found in literatures. For example, Yokoyama et al. [17] implemented Genetic Algorithm (GA) to solve a constrained space plane reentry problem. In [18], a low-thrust interplanetary trajectory optimization problem was studied, and the main contribution is that a bi-level evolutionary-based optimal control solver was created. Kim et al. [19] proposed a Particle Swarm Optimization (PSO) to calculate the optimal trajectory for a manipulator motion planning problem. Similarly, a doubly fed induction generator optimal control problem was investigated in [2], wherein a modified PSO algorithm was applied to approximate the optimal control trajectory. Although all the aforementioned works have shown the feasibility of applying stochastic algorithms for trajectory optimization problems, the verification of solution optimality will become difficult. Moreover, the computational burden caused by the evolutionary optimization process is usually high [20].

In order to combine the advantages of the traditional and stochastic-based methods, this paper presents a hybrid structure for solving the spacecraft trajectory optimization

R. Chai, A. Savvaris and A. Tsourdos are with the School of Aerospace, Transport and Manufacturing, Cranfield University, UK, e-mail: (r.chai@cranfield.ac.uk), (a.savvaris@cranfield.ac.uk), and (a.tsourdos@cranfield.ac.uk).

S.Chai and Y. Xia are with the school of Automation, Beijing Institute of Technology, Beijing, China, e-mail: (chaisc97@163.com), (xia_yuanqing@bit.edu.cn).

©2017 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

problem. So far to the best of our knowledge, there are fewer integrated design results have been reported to generate the optimal spacecraft flight trajectory. Therefore, the present study is an attempt to address this concern. The proposed structure contains an outer layer that can generate a reference trajectory with respect to the state and control variables. The algorithm used in this layer is an adaptive differential evolution algorithm and the fundamental framework of this stochastic algorithm is based on [21]. Following that, the reference trajectory is provided to the inner layer, where a new improved gradient-based solver is applied. Compared with other typical optimal control methods, using this approach can improve both the stability and efficiency. Moreover, this integrated design can offer the designers more flexibility in order to control the optimization process and verify the optimality.

The main contributions of the work reported in this paper include the following three aspects:

1) We propose a new optimal trajectory generation method by constructing an integrated framework that contains a newly developed initial guess generator and an inner gradient solver.
2) An improved gradient-based solver is designed and embedded in the proposed algorithm framework. This method can better control the solution-finding process and achieve faster convergence speed for solving the spacecraft trajectory optimization problem.
3) Both optimality verification and simulation results are provided to illustrated the effectiveness of the proposed design.

Some preliminary results have appeared in [21]; however, the results obtained in this paper are more integrated, including detailed formulation, analysis and simulations.

The remainder of this paper is organized as follow: Section II presents the mathematical formulation of the Space Maneuver Vehicles (SMV) trajectory optimization problem. In Section III, the three-layer-hybrid optimal control solver is constructed. The optimality conditions of the SMV trajectory planning problem is derived and analyzed in Section IV. Section V presents the simulation results and verifies the optimality conditions numerically. The concluding remark is given in Section VI.

II. Space Maneuver Vehicle trajectory optimization problem

In this section, the mission scenario investigated in this paper is presented. The space vehicle re-enters the atmosphere at a predetermined altitude for observation and gathering of information of inaccessible areas. The space vehicle descends down to a minimum allowable altitude of around $50km$, once this altitude point is reached, the spacecraft fires its engine and starts the ascent phase, exiting the atmosphere and returning back to Low Earth Orbit (LEO). An engine model is embedded in the dynamics such that the vehicle can have enough kinetic energy to exit the atmosphere and return back into orbit.

A. Problem formulation

The overall formulation of the time-optimal SMV trajectory programming problem containing the flight dynamics is given as follows:

$$\begin{aligned}
\text{minimize} \quad & J = t_f \\
\text{subject to} \quad & \dot{r} = V\sin\gamma \\
& \dot{\theta} = \tfrac{V\cos\gamma\sin\psi}{r\cos\phi} \\
& \dot{\phi} = \tfrac{V\cos\gamma\cos\psi}{r} \\
& \dot{V} = \tfrac{T\cos\alpha - D}{m} - g\sin\gamma \\
& \dot{\gamma} = \tfrac{L\cos\sigma + T\sin\alpha}{mV} + (\tfrac{V^2 - gr}{rV})\cos\gamma \\
& \dot{\psi} = \tfrac{L\sin\sigma}{mV\cos\gamma} + \tfrac{V}{r}\cos\gamma\sin\psi\tan\phi \\
& \dot{m} = -\tfrac{T}{I_{sp}g} \\
& [r(0), \theta(0), \phi(0), V(0), \gamma(0), \psi(0), m(0)] \\
& = [r_0, \theta_0, \phi_0, V_0, \gamma_0, \psi_0, m_0]
\end{aligned} \tag{1}$$

where t_f is the terminal time. r, θ, ϕ, V, γ, ψ, m represent the radial distance, longitude, latitude, velocity, flight-path angle, heading angle and vehicle's mass, respectively. The control variables are the angle of attack α, bank angle σ and thrust T. In the paper, the dynamic model given by Eq.(1) is abbreviated as $\dot{x} = f(x,u), x(0) = x_0$, where $x \in \Re^7$ and $u \in \Re^3$ denote the state and control variables, respectively. Detailed information including the aerodynamic model and atmospheric model can be found in [22].

B. Constraints

The spacecraft reconnaissance mission should satisfy strict box and path constraints to protect the structure of the vehicle. These constraints usually depend on the mission requirements. Regarding the box constraints, during the mission, each state and control variable should vary in the tolerable region and this can be described as $x_{min} \le x \le x_{max}$ and $u_{min} \le u \le u_{max}$. Besides, the path constraints that the vehicle must satisfy during the entire mission are the aerodynamic heating, dynamic pressure and load factor, which can be summarised as:

$$\begin{aligned}
& \dot{Q}_d = K_Q \rho^{0.5} V^{3.07}(c_0 + c_1\alpha + c_2\alpha^2 + c_3\alpha^3) < \dot{Q}_{dmax} \\
& P_d = \tfrac{1}{2}\rho V^2 < P_{dmax} \\
& n_L = \tfrac{\sqrt{L^2 + D^2}}{mg} < n_{Lmax}
\end{aligned} \tag{2}$$

where Q_{dmax}, P_{dmax} and n_{Lmax} represent acceptable maximum heating rate, dynamic pressure and load factor, respectively.

In order to describe the angular rate and range of the control variables, three first-order lag equations are appended to the equations of motion. Consequently, the control variables are divided into the actual controls α, σ, T and the demanded controls α_c, σ_c, T_c. In this way, the discontinuity in the control profiles can be removed.

$$\begin{cases} \dot{\alpha} = k_\alpha(\alpha_c - \alpha) \\ \dot{\sigma} = k_\sigma(\sigma_c - \sigma) \\ \dot{T} = k_T(T_c - T) \end{cases} \quad \begin{cases} \alpha_{cmin} \le \alpha_c \le \alpha_{cmax} \\ \sigma_{cmin} \le \sigma_c \le \sigma_{cmax} \\ T_{cmin} \le T_c \le T_{cmax} \end{cases} \tag{3}$$

III. A three-layer-hybrid optimal control solver

Traditional optimal control solvers tend to be sensitive with respect to the initial guess value provided by the users. However, according to the work carried-out by Conway [6], for solvers based on evolutionary algorithms or metaheuristics,

the problem needs to be discretized by a relatively small size of temporal set, and there is no guarantee of optimality. To address these problems, a new three-layer-hybrid optimal control solver was designed by combining the advantages of the traditional solvers and evolutionary-based solvers.

A. Initial guess generator

In the first layer of the proposed solver, an initial guess generator is designed to generate a reference trajectory, which can then be provided to the inner gradient-based solver as a start point of Newton iterations. The method used in this layer is an adaptive differential evolution algorithm similar to the framework proposed in [21]. This algorithm is then combined with a self-learning strategy proposed in [23]–[25]. Moreover, to handle different types of constraints entailing in the problem, a V-based constraint handling strategy is embedded in the algorithm framework, hence the name Violation Learning Differential Evolution (VLDE). For completeness, a brief description of this stochastic-based method is introduced in this section.

By using the discretization technique, the continuous-time optimal control problem given by Eq.(1) is then converted to a series of static Nonlinear Programming problems (NLPs). The discretization method used in this layer is the direct Runge-Kutta method [8] and the resulting static model can be written as:

$$\begin{array}{ll} \text{minimize} & J = t_f \\ \text{subject to} & x_{k+1} = x_k + h_k \sum_{i=1}^{s} b_i f(x_{ki}, u_{ki}), \\ & x_{ki} = x_k + h_k \sum_{j=1}^{s} a_{ij} f(x_{kj}, u_{kj}), \\ & g(x_{ki}, u_{ki}) \geq 0, \\ & x(0) = x_0 \\ & i = 1, ..., s \qquad k = 0, ..., N_{k-1} \end{array} \tag{4}$$

where N_k is the number of discretized time nodes, while g stands for the inequality constraints described in Eq.(2)-(3). It can be seen from Eq.(4) that only the control variables are parameterized at temporal nodes $[t_0, t_1, ..., t_f]$. Then, the equations of motion are integrated with a numerical integration method, e.g, forth-order Runge-Kutta method. In this way the advantage is that the control box constraints (described in Eq.(2)) and equations of motion (Eq.(1)) can be satisfied automatically by initializing all population members within the specified lower and upper bounds and by integrating the dynamic model forward via numerical integration. Specifically, if the initial population contains NP individuals, then all the decision variables can be generated randomly according to the limits of demanded angle of attack, bank angle and thrust given in Eq.(2) such that every decision variable can be in the feasible zone.

$$\begin{array}{l} \alpha_c = \alpha_c^{min} + rand(\cdot) \times (\alpha_c^{max} - \alpha_c^{min}) \\ \sigma_c = \sigma_c^{min} + rand(\cdot) \times (\sigma_c^{max} - \sigma_c^{min}) \\ T_c = T_c^{min} + rand(\cdot) \times (T_c^{max} - T_c^{min}) \end{array} \tag{5}$$

Remark 1. It is worth noting that by combining a shooting-scheme with the evolutionary solver, it offers the user more flexibility to choose the temporal set. (e.g. the designer can use a relatively large temporal set to parameterize the dynamics). This is because the resulting NLP problem contains less number of equality constraints compared with collocation discretization scheme. For collocation methods, both of the state and control variables will be discretized and consequently, the equations of motion are transcribed to a series of algebraic equations (equality constraints). This implies that the resulting number of optimization parameters and equality constraints tends to be larger and cannot be satisfied automatically. For instance, if an optimal control problem contains n_s state variables, n_c control variables and is parameterized using N_k collocation points, the number of decision variables of the NLP solver is $n_s \times (N_k + 1) + (n_c \times N_k) + 1$, which is large compared with shooting scheme since in a shooting method, only the control variable is parameterized. When a NLP problem contains too many optimization parameters and equality constraints, it tends to cost the stochastic-based solver significant number of iterations to capture the true behaviour of the model and fail to satisfy all of the constraints.

It should be noted that the constraint handling strategy used in this paper is based on the violation degree of constraints V_{ol}. Using this approach, each individual among the population can be associated with all the constraints; and the value of the violation function can directly reflect the magnitude of the solution infeasibility. For example, the violation degree for relation "$\leq$" (i.e. $g_j(u) \leq g_j^*$, where g_j^* is the required upper bound) can be defined as:

$$\mu_{g_j}(u) = \begin{cases} 0, & g_j(u) \leq g_j^*; \\ \frac{g_j(u) - g_j^*}{g_j max - g_j^*}, & g_j^* \leq g_j(u) \leq g_j^{max}; \\ 1, & g_j(u) \geq g_j^{max}. \end{cases} \tag{6}$$

where $\mu_{g_j}(u)$ is a continuous function and $g_j(u)$ is the value of jth constraint for each individual, while the tolerance range should be (g_j^*, g_j^{max}).

Consequently, the value of violation function V_{ol} can be obtained via $V_{ol} = \sum_{j=1}^{I} \mu_{g_j}(u) + \sum_{k=1}^{E} \mu_{h_k}(u)$, where h stands for the equality constraints. I and E are the number of inequality and equality constraints, respectively. As a result, the selection procedure is depended on the value of V_{ol}. The feasible individual ($V_{ol} = 0$) can always dominate the infeasible one ($V_{ol} \neq 0$), while the individual with small violation degree can always dominate the one with large violation degree and then be selected.

The main framework of the proposed VLDE method is constructed in Algorithm 1, from where it can be seen that VLDE applies the elitism strategy similar to that of traditional DE. The main new contributions in the VLDE algorithm lie in its offspring-creation component, where a simplex-based direct search operation is coupled with traditional generic operations such as crossover and mutation. In addition, a learning strategy is also designed according to the nature of the SMV optimal control problem. In this way, the diversity of the offspring generation can be further improved, thus helping avoid the premature convergence of VLDE. It was shown in [21] that by applying these modifications, the convergence speed of the VLDE method for solving spacecraft trajectory optimization problem is effectively improved. Hence this method is chosen as the outer initial guess generator in this work.

Since the aim of the initial trajectory generator is to

Algorithm 1 The main framework of the proposed VLDE

Input: the maximum number of generations G_{max}, number of temporal nodes N_k;
Output: the best candidate x_{best} among the final population $P_{G_{max}}$;
/*Initialization*/
Initialization: generate the initial population P_0 with N randomized individuals;
/*Main Loop*/
while $G < G_{max}$ **do**
    (a). choose the best number of the current population as P_G;
    (b). generate the offspring generation Q_G
        Q_G = VLDE offspring-creation (P_G);
    (c). L_G = learning strategy (P_G, Q_G);
    (d). set $P_G = P_G \cup Q_G \cup L_G$;
    (e). perform the elite selection based on the new dominant rule to get P_{G+1};
    (e). set $G = G + 1$;
end while

provide reference state and control sequences to the inner gradient solver, in order to keep a balance between the quality of solution and the computational burden, G_{max} is fixed as 100.

B. An improved gradient-based inner solver

After generating the reference state and control sequences, these results are provided to the inner gradient-based optimizer (layer 2). In the second layer, in order to achieve a higher accuracy, pseudospectral discretization scheme is applied to transform the dynamics. That is, both the control and state variables are parameterized at temporal nodes. It is discussed in [26], [27] that the approximation accuracy of pseudospectral methods can be directly controlled via the number of collocation nodes. For example, the approximate error order for Gauss pseudospectral method is $\mathcal{O}(N_k^{3-p})$ [27], whereas for Radau pseudospectral method, this order becomes $\mathcal{O}(N_k^{2-\eta})$ [26].

The optimization methodology used in the second layer is an improved gradient algorithm based on the Interior Point method (IP) and Sequential Quadratic Programming (SQP). The general idea of the proposed algorithm (referred to as the IPSQP algorithm) is to use the SQP method where the Quadratic Programming (QP) subproblem is solved using an IP method [28]. The general quadratic form of the transcribed trajectory optimization problem is given below:

$$\begin{aligned} \min \tfrac{1}{2} dx^T H(x_k, \lambda_k, u_k) dx + \nabla f(x_k)^T dx \\ h(x_k) + \nabla h(x_k) dx = 0 \\ g(x_k) + \nabla g(x_k) dx \leq 0 \end{aligned} \tag{7}$$

where $h(\cdot) = (h_1(\cdot), h_2(\cdot), ..., h_E(\cdot))^T$ and $g(\cdot) = (g_1(\cdot), g_2(\cdot), ..., g_I(\cdot))^T$. H stands for the Hessian matrix. At a fixed iteration time k, the IP strategy is used to solve the quadratic programming model given by Eq.(7). Then, the formulation of the improved gradient method can be formulated as follows:

$$\begin{aligned} \min \tfrac{1}{2} dx^T H(x_k, u_k) dx + \nabla f(x_k)^T dx - \mu_k \Sigma_{j=1}^m log(s_j^k + d_j^s) \\ st. \quad g(x_k) + \nabla g(x_k) dx + s_k + e^T ds = 0 \\ h(x_k) + \nabla h(x_k) dx = 0 \\ dx \in \Re^n, ds \in \Re^m \end{aligned} \tag{8}$$

where the primal and dual variables are dx and ds, respectively. The last two terms $s_k + e^T ds$ in the equality constraint are considered as the slack variables. The algorithm consists of two nested loops and correspondingly, it contains two iteration indices. To distinguish these two iterations, the internal iteration index is defined as l while the external iteration number is defined as k. Specifically, the primal, slack and dual variables in the inner iteration are $dx_{k,l}, ds_{k,l}, d\lambda_{k,l}$ and $du_{k,l}$, respectively. The main advantage of this algorithm is that the user can control the inner loop by setting the termination conditions at any time. Since the H_k is fixed at the internal circle, it is not required to solve the QP subproblem exactly, which means finding the time-consuming QP solution can be avoided.

The Karush-Kuhn-Tucker (KKT) system of Eq.(8) is described in Eq.(9), and it is solved iteratively.

$$\begin{pmatrix} H_k & 0 & \nabla h(x_k)^T & \nabla g(x_k)^T \\ 0 & Du_{k,l} & 0 & Ds_{k,l} \\ \nabla h(x_k) & 0 & 0 & 0 \\ \nabla g(x_k) & I & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta dx_{k,l} \\ \Delta ds_{k,l} \\ \Delta d\lambda_{k,l} \\ \Delta du_{k,l} \end{pmatrix} = \begin{pmatrix} -H_k dx_{k,l} - \nabla f(x_k) - \nabla h(x_k)^T d\lambda_{k,l} - \nabla g(x_k)^T du_{k,l} \\ -Ds_{k,l} du_{k,l} + \mu_{k,l} e \\ -h(x_k) - \nabla h(x_k) dx_{k,l} \\ -g(x_k) + \nabla g(x_k) dx_{k,l} - s_k + e^T ds_{k,l} \end{pmatrix} \tag{9}$$

where $\Delta d = [\Delta dx_{k,l}, \Delta ds_{k,l}, \Delta d\lambda_{k,l}, \Delta du_{k,l}]^T$. $Ds_{k,l}$ and $Du_{k,l}$ are positive diagonal matrices corresponding to the slack variables and multipliers; while λ and μ are Lagrangian multipliers and penalty factors related to equality constraints and inequality constraints, respectively.

Solving the KKT system, the new iteration can be calculated by:

$$\begin{aligned} dx_{k,l+1} &= dx_{k,l} + \alpha_{k,l} \Delta dx_{k,l} \\ du_{k,l+1} &= du_{k,l} + \alpha_{k,l} \Delta du_{k,l} \\ ds_{k,l+1} &= ds_{k,l} + \alpha_{k,l} \Delta ds_{k,l} \\ d\lambda_{k,l+1} &= d\lambda_{k,l} + \alpha_{k,l} \Delta d\lambda_{k,l} \end{aligned} \tag{10}$$

where the step length parameter $\alpha_{k,l} \in (0, 1]$ should be chosen to ensure that the merit function achieves sufficient decrease but the step is not too short. In this paper, the l_∞-merit function is implemented in order to measure the progress of each iteration and it can be written as:

$$\begin{aligned} M_{\mu,r}(x, s, \lambda, u, dx, ds, d\lambda, du) \\ = \tfrac{1}{2} dx^T H dx + \nabla f(x)^T dx - \mu_k \Sigma_{j=1}^m log(s_j + d_j^s) \\ + \lambda \Sigma_{i=1}^l (h_i(x) + \nabla h_i(x)) dx \\ + r \| g(x) + \nabla g(x) dx + s + e^T ds \|_\infty \end{aligned} \tag{11}$$

where r is the penalty factor and μ is a barrier parameter. The line search process shown in Eq.(10) should ensure that there is a progress on the merit function for each iteration and

therefore, the Goldstein conditions is applied.

$$\begin{aligned} & M_{\mu,r}(x_k, s_k, \lambda_k, u_k) + c_1\alpha_k \nabla M_{\mu,r}^T \Delta d_k \\ & \leq M_{\mu,r}(x_k + \alpha_k dx_k, s_k + \alpha_k ds_k, \\ & \qquad \lambda_k + \alpha_k d\lambda_k, u_k + \alpha_k du_k) \\ & \leq M_{\mu,r}(x_k, s_k, \lambda_k, u_k) + c_2\alpha_k \nabla M_{\mu,r}^T \Delta d_k \end{aligned} \tag{12}$$

with $0 < c_1 < c_2 < 1$. Δd_k denotes the directional derivative $(dx_k, ds_k, d\lambda_k, du_k)$. The second term of the inequality is the sufficient decrease condition while the first term of the inequality is to control the step length so that it will not be chosen too short.

In order to better show the structure of this two nested gradient optimization algorithm, the overall procedure is illustrated in the Pseudocode (see Algorithm 2).

Algorithm 2 Pseudocode for the IPSQP method

procedure (Two nested structure)
  Choose starting values $z_0 = (x_0, u_0, \lambda_0, s_0)$
  for $k := 0, 1, 2, ...$ **do**
    (a). Check stopping criteria for the outer loop
    (b). Choose $dx_{k,0}$, $du_{k,0}$, $d\lambda_{k,0}$ and $ds_{k,0}$
    for $l := 0, 1, 2, ..., l_{max}$ **do**
      i. Determine $Du_{k,l}$, $Ds_{k,l}$ and $\mu_{k,l}$.
      ii. Solve the KKT system described in Eq.(9).
      iii. Apply the line search algorithm shown in Eq.(10).
      iv. If the inner loop solution can satisfy the stopping condition of QP, break for-loop;
    end for
    (c). Find step length for the outer loop such that the merit function can have a proper improvement.
    (d). Update the current searching point and go back to line 3.
  end for
  Output the optimal solution
end procedure

Remark 2. It should be noted that commonly the first initial guess and active set provided by the user are far from the optimal solution. Therefore, the Lagrangian multipliers calculated by using SQP are inaccurate. If the quadratic model is solved using SQP and active set, it usually takes the solver a large number of iteration to converge. However, by applying the inner loop (controlled by the number of l_{max}) in the proposed algorithm and the reference solution generated from Layer 1, the identification of active set and the solution finding can be more accurate. Moreover, the convergence ability can also be improved and the computational burden can be decreased at the same time (less Newton iteration).

C. Mesh refinement

For direct transcription methods (e.g. direct multiple shooting and direct collocation), in general, the accuracy of the solution depends largely on the mesh refinement procedure [5], [29], [30]. The aim for carrying out mesh refinement is to determine whether the current mesh grid is proper and update the mesh grid so that the resulting mesh can contain a small number of temporal points and the mesh distribution is dense where the discontinuity order is high. Therefore, the third layer of the designed hybrid optimal control solver is coupled with the second layer (inner gradient-based optimizer) so as to refresh the mesh grid. The idea applied to do the mesh update is the hp-strategy. Detailed information including the analytical formulation of the hp-strategy can be found in [22], [30].

Let ε stand for an accuracy tolerance for the discrete algebraic function constraint. The errors of the dynamic equations and path constraints at the sth collocation point, kth time interval are $a_s^{(k)}$ and $b_s^{(k)}$, respectively. Therefore, the maximum error $e_{max}^{(k)}$ in the kth time interval can be approximated as $e_{max}^{(k)} = max[a_s^{(k)}, b_s^{(k)}]$. If the equation $e_{max}^{(k)} \leq \varepsilon$ can be satisfied, then the algorithm will stop the iteration because collocation points in the kth interval can reach the tolerance. Otherwise, it should be divided into subintervals or add more collocation points.

Suppose $k_{max}^{(k)}$ and $\overline{k}^{(k)}$ are the maximum curvature and average curvature of all the nodes, respectively. Furthermore, let $r^{(k)} = \frac{k_{max}^{(k)}}{\overline{k}^{(k)}}$ be the ratio of the maximum to the mean curvature. Setting the tolerance of curvature as r_{max} and if $r^{(k)} \geq r_{max}$, then the trajectory in this time interval tends to have oscillations and it should be divided into new subinterval. The number of the subinterval n_k is determined using $n_k = ceil(\log(e_{max}^{(k)}/\epsilon))$, where the function of $ceil(\cdot)$ is to round a number to the next larger integer. On the other hand, if the tolerance can satisfy $r^{(k)} < r_{max}$, the trajectory tends to be flat in this time interval and the accuracy can be improved by adding more collocation points. The number of points that should be added is determined by $N_k = N_k + ceil(\log(e_{max}^{(k)}/\epsilon))$.

Remark 3. In order to improve the efficiency of the proposed solver, in the simulation, each refinement iteration will use the previous mesh history as the start point to do the solution finding. The stopping criteria depends on the maximum number of iteration and the accuracy tolerance. The algorithm will carry on until one of the stopping criteria can be satisfied.

D. Algorithm framework

The framework of the proposed three-layer-hybrid optimal control solver is presented in Fig.1.

IV. Optimality verification

Since the SMV trajectory planning problem is formulated as an optimal control problem, to judge the quality of the solution generated by the proposed three-layer-hybrid optimal control solver, the first-order necessary conditions should be used. According to the problem formulation given by Eq.(1), the Hamiltonian function H_{am} is constructed as $H_{am} = \Phi(x, u, t; t_0, t_f) + \lambda^T(t) f(x, u, t; t_0, t_f)$. Taking into account the path constraints, the augmented Hamiltonian is

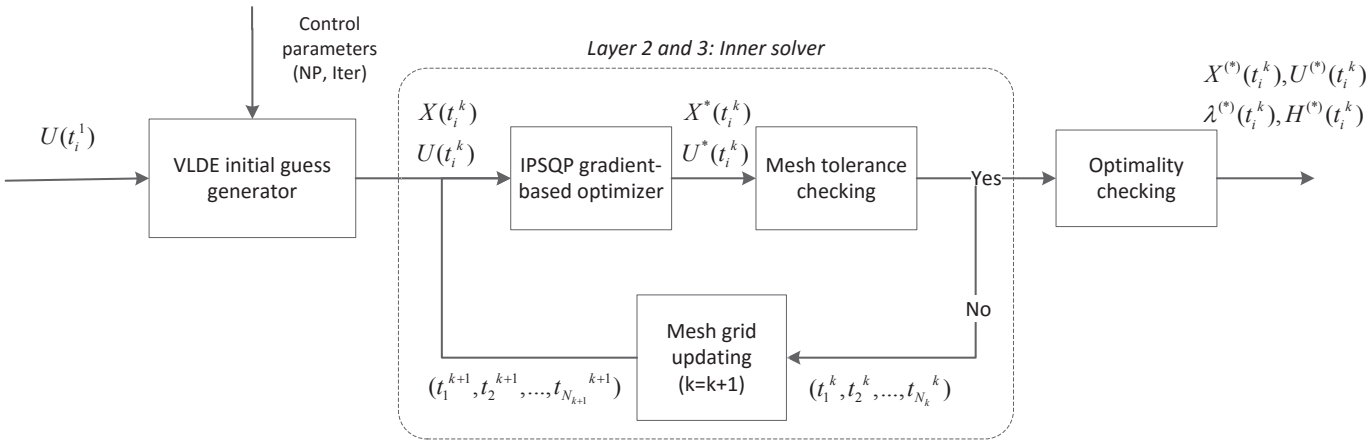

Fig. 1: Flow chart of the designed solver

defined as:

$$\begin{aligned} H_{am}(x,\lambda,u,\mu,t;t_0,t_f) = \quad & \Phi(x,u,t;t_0,t_f) \\ & +\lambda^T(t) f(x,u,t;t_0,t_f) \\ & -\mu^T(t) C(x,u,t;t_0,t_f) \end{aligned} \tag{13}$$

where Φ is the Lagrange form cost function, f is the right hand side of the equations of motion, $\lambda(t) = [\lambda_r, \lambda_\theta, \lambda_\phi, \lambda_V, \lambda_\gamma, \lambda_\psi, \lambda_m, \lambda_\alpha, \lambda_\sigma, \lambda_T]^T \in \Re^{10}$ is the costate corresponding to the dynamic equations and $\mu(t) \in \Re^3$ is the Lagrange multiplier associated with the path constraints.

A. First order necessary conditions

The continuous-time first-order necessary conditions in terms of costate and hamiltonian can be written as:

$$\begin{aligned} \lambda(t_0) &= -\frac{\partial J}{\partial x(t_0)} + \upsilon^T \frac{\partial \phi}{\partial x(t_0)} \\ \lambda(t_f) &= -\frac{\partial J}{\partial x(t_f)} + \upsilon^T \frac{\partial \phi}{\partial x(t_f)} \\ H_{am}(t_0) &= \frac{\partial J}{\partial t_0} - \upsilon^T \frac{\partial \phi}{\partial t_0} \\ H_{am}(t_f) &= \frac{\partial J}{\partial t_f} - \upsilon^T \frac{\partial \phi}{\partial t_f} \end{aligned} \tag{14}$$

where ϕ is the boundary condition and υ is the Lagrange multiplier associated with the boundary condition. Besides, the first-order Hamiltonian minimization condition is based on the minimum principle such that the optimal control $u^* = [\alpha_c^*, \sigma_c^*, T_c^*]$ must minimize the Hamiltonian with respect to control variables.

$$\frac{\partial H_{am}}{\partial u} = \frac{\partial g}{\partial u} + \left(\frac{\partial f}{\partial u}\right)^T \lambda - \left(\frac{\partial C}{\partial u}\right)^T \mu = 0 \tag{15}$$

It is obvious that the first term $\frac{\partial g}{\partial u} = 0$ and therefore, Eq.(15) can be rewritten as:

$$\begin{cases} \frac{\partial \lambda_\alpha [K_\alpha(\alpha_c - \alpha)]}{\partial \alpha_c} - q_{\alpha_c} = 0, & \text{(with respect to } \alpha_c) \text{ ;} \\ \frac{\partial \lambda_\sigma [K_\sigma(\sigma_c - \sigma)]}{\partial \sigma_c} - q_{\sigma_c} = 0, & \text{(with respect to } \sigma_c) \text{ ;} \\ \frac{\partial \lambda_T [K_T(T_c - T)]}{\partial T_c} - q_{T_c} = 0, & \text{(with respect to } T_c) \end{cases} \tag{16}$$

where $q_u = [q_{\alpha_c}, q_{\sigma_c}, q_{T_c}]^T$ is the Lagrange multiplier with respect to the control path constraints. Also, based on the KKT complementary condition, the control multipliers should satisfy the following equation:

$$q_u \begin{cases} \leq 0 & if, \quad u = u_{min}; \\ = 0 & if, \quad (u_{min} < u < u_{max}; \\ \geq 0 & if, \quad u = u_{max}. \end{cases} \tag{17}$$

Eq.(17) implies when the control constraints become active, the corresponding multipliers become nonzero.

B. Terminal transversality conditions

The following proposition illustrates the property of costate variables at the terminal time instant.

Proposition 1. Consider the time-optimal spacecraft trajectory optimization problem has a feasible optimal solution: The costate value $[\lambda_\theta, \lambda_\phi, \lambda_V, \lambda_\psi, \lambda_m, \lambda_\alpha, \lambda_\sigma, \lambda_T]$ must be zero when $t = t_f$.

Proof: Based on the Eq.(14), the transversality condition at the final time instant has the following form

$$\lambda(t_f) + \frac{\partial J}{\partial x(t_f)} - \upsilon^T \frac{\partial \phi}{\partial x(t_f)} = 0$$

For the time-optimal spacecraft trajectory planning problem, there are no constraints for $\theta, \phi, V, \psi, \alpha, \sigma, T$ at t_f. Therefore, the costate value for each state variable at the final time should hold Eq.(18).

$$\begin{cases} \lambda_\theta(t_f) = \frac{\partial J}{\partial \theta} \mid_{t=t_f} = 0 \\ \lambda_\phi(t_f) = \frac{\partial J}{\partial \phi} \mid_{t=t_f} = 0 \\ \lambda_V(t_f) = \frac{\partial J}{\partial V} \mid_{t=t_f} = 0 \\ \lambda_\psi(t_f) = \frac{\partial J}{\partial \psi} \mid_{t=t_f} = 0 \\ \lambda_m(t_f) = \frac{\partial J}{\partial m} \mid_{t=t_f} = 0 \\ \lambda_\alpha(t_f) = \frac{\partial J}{\partial \alpha} \mid_{t=t_f} = 0 \\ \lambda_\sigma(t_f) = \frac{\partial J}{\partial \sigma} \mid_{t=t_f} = 0 \\ \lambda_T(t_f) = \frac{\partial J}{\partial T} \mid_{t=t_f} = 0 \end{cases} \tag{18}$$

which completes the proof. ■

Applying the terminal transversality condition can give some indications of the final value of the dual variables λ that can be used later to verify the numerical results.

C. Hamiltonian function condition

The following proposition gives the nature of the Hamiltonian function with respect to time.

Proposition 2. Assume $x^*(t)$, $u^*(t)$ are the optimal state and control variables, the corresponding Hamiltonian function should satisfy $H_{am}(x^*(t), u^*(t)) = -1$ for $t \in [t_0, t_f]$.

Proof: Since the final time is a free variable, there is a stationary condition for the H_{am} function. This condition is obtained by applying the endpoint Lagrangian equation (e.g. $H_{am}(t_f) - \frac{\partial J}{\partial t_f} + \upsilon^T \frac{\partial \phi}{\partial t_f} = 0$). For the problem studied in this paper, this equation can be rewritten as:

$$\begin{aligned} H_{am}(X(t_f), t_f, \upsilon) &= -\frac{\partial J}{\partial t_f} + \upsilon_r \frac{\partial (r_f - r(t_f))}{\partial t_f} \\ &\quad + \upsilon_\gamma \frac{\partial (\gamma_f - \gamma(t_f))}{\partial t_f} \\ &= -1 \end{aligned} \tag{19}$$

Eq.(19) implies that the final value of the Hamiltonian function should be -1 for this problem. Then the Hamiltonian evolution equation is used to demonstrate the behaviour of the Hamiltonian with respect to time such that:

$$\frac{\partial H_{am}}{\partial t} = 0 \tag{20}$$

Eq.(20) means H_{am} is not explicitly a function of time and $H_{am} = c$, where c is a constant. Combining Eq.(19) and Eq.(20), it is clear that the following equation should be satisfied.

$$H_{am}(x^*(t), u^*(t)) = -1, \quad \forall t \in [t_0, t_f] \tag{21}$$

■

Therefore, it is obvious that the Hamiltonian function should be -1 during the entire time history. The aim of the analysis carried-out in this subsection is to verify the first-order quality of the numerical solution calculated using the designed optimal control solver.

D. Properties of the control variable

Similarly, the optimal control variable has the following property.

Proposition 3. If the time-optimal spacecraft trajectory optimization problem has the optimal control solution $u^*(t)$, then $u^*(t)$ can be expected to have a "bang-bang" behaviour for all $t \in [t_0, t_f]$.

Proof: The control variables should be chosen such that the augmented Hamiltonian function can be minimized. Since the rate constraint of the control variable is achieved using the first order lag equations (shown in Eq.(3)), the path constraints do not involve the control variables explicitly, which means the optimal solution may contain corners. Moreover, according to Eq.(3), the demanded control appears linearly in the differential equations. Therefore to minimize the Hamiltonian function H_{am} with respect to the demanded control, taking into account the control variable constraints, the demanded control variables should move from one point from the boundary of the feasible control region to another point on the boundary, which can be expressed as

$$u^*(t) = \begin{cases} u_{min} & if, \quad \rho < 0; \\ u_{max} & if, \quad \rho > 0. \end{cases} \tag{22}$$

In Eq.(22), ρ is the switching function with the expression $\rho = -\lambda^{'} K - q_u$, where $\lambda^{'} = [\lambda_\alpha, \lambda_\sigma, \lambda_T]$ and $K = [k_\alpha, k_\sigma, k_T]$, respectively. The sign of the switching function dominates the magnitude of the controls according to Eq.(22). ■

Consequently, as for the SMV model considered in this paper, it can be expected to experience a switching structure in terms of the demanded angle of attack, bank angle and thrust profiles.

E. Bellman's principle

Another way to verify the optimality is Bellman's principle. The main idea of this principle is that the optimal result will not change if several points on the original optimal trajectory are selected as the initial condition to a new problem. This principle is also the main theory of Dynamic Programming (DP) in terms of optimality. If several time points on the original trajectory are selected as initial conditions and the results show that there is no better or different solutions, then the Bellman's optimality principle can be satisfied.

V. Simulation results

A. Parameter setting

All the mission-dependent and vehicle-dependent parameters including the initial, terminal boundary conditions, box constraints and aerodynamic coefficients of the skip problem can be found in [22]. Regarding the path constraints for mission 1, the maximum allowable heating, dynamic pressure and load factor are set as: $Q_{max} = 200BTU$; $Pd_{max} = 13406.4583Pa$; $nl_{max} = 2.5$, respectively.

B. Optimal solution

According to the dynamic model, objective function and path constraints that were given in Section II, the numerical solutions obtained using the VLDE-based initial guess generator and the hybrid solver are shown in Fig.2. The characteristic arcs of the obtained trajectories are analyzed in this subsection. The trajectory is split into two subintervals: descent and climb. In the descent phase, to reach the minimum predetermined altitude (around 50km above sea level) and minimize the objective function (final time), Fig.2 shows that the SMV descends directly at the start of the mission. The angle of attack increases to slow down the vehicle so that the heating and dynamic pressure do not increase significantly, hence to avoid the path constraints from becoming active. While rapid descent is necessary, it should be noted that there is a slight dip in terms of the curvature of the altitude (a small hop) before reaching the minimum altitude point. This is because if the spacecraft descends without increasing the flight path angle, then the dynamic pressure and load factor constraints may become active. To preserve kinetic energy and satisfy the path constraints, the curvature of the altitude is decreased such that the SMV can have sufficient time to use the drag force

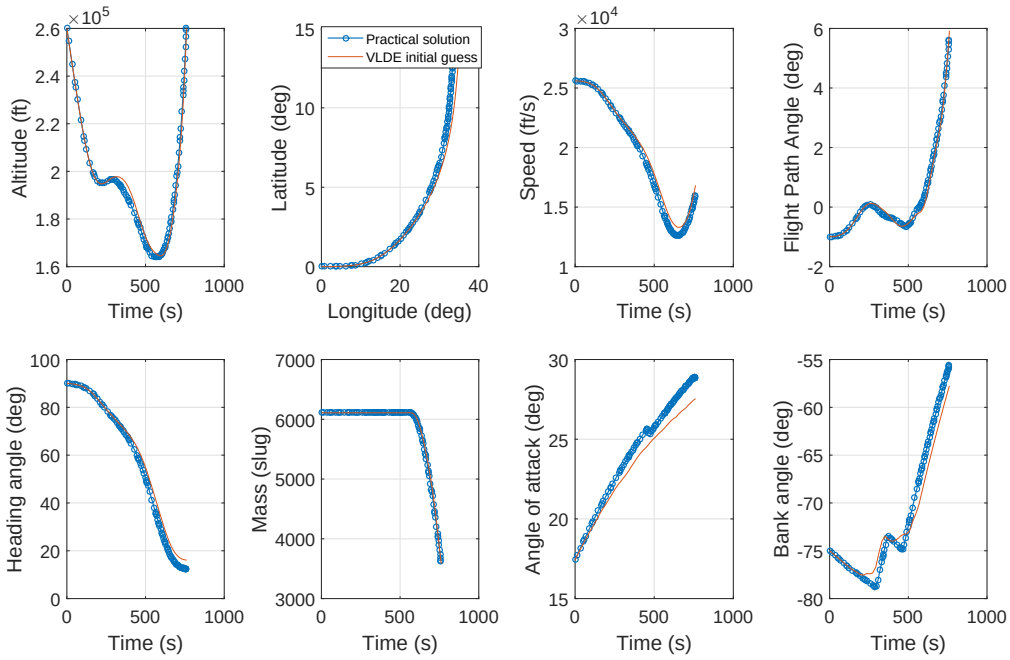

Fig. 2: Time history for states and controls (Mission 1: $Q_{max} = 200BTU$; $Pd_{max} = 13406.4583Pa$; $nl_{max} = 2.5$)

(i.e. slow its speed down). Therefore, the structure integrity of the vehicle can be guaranteed at the expense of the objective function.

In the climb phase, the SMV fires it engine so that the vehicle can accelerate and have enough kinetic energy to return back to orbit. The trend for the angle of attack can be seen in Fig.2, where the angle of attack is increased during the whole climbing phase. The reason is that in the climbing phase, without violating the path constraints it can have positive influences in terms of acceleration.

To further test and analyze the performance of the proposed hybrid strategy, another mission scenario, which contains more strict limits on the path constraints was carried out using the designed hybrid optimal solver. In this case, the maximum allowable heating, dynamic pressure and load factor for mission 2 are restricted to: $Q_{max} = 150BTU$; $Pd_{max} = 11970.05Pa$; $nl_{max} = 2.0$, respectively. The time history of the optimization variables and path constraints are shown in Fig.3 and Fig.4.

As can be seen from Fig.3 and Fig.4, the proposed hybrid strategy can still generate high quality solutions without violating path constraints and box constraints. Compared with the results shown in Fig.2, the solution obtained in Fig.3 and Fig.4 shows a slight difference. The reason of the additional hop in the altitude profile after around 400s is to prevent the load factor constraint from becoming active. Therefore, based on there observations and results, it can be concluded that the initial guess generator embedded in the algorithm framework can have positive influences for the inner gradient-based solver in terms of increasing convergence ability and getting rid of infeasibility.

Regarding to the mesh refinement process, the way that this strategy adds points in a specific time interval or divides an interval into subintervals was described in Section III.C. Using the hp strategy, the time history for the states and controls can be much smoother. This can be seen in Fig.2 to 4, where the distribution of grid points tends to be dense at the areas having a high value of curvature (the trajectory at those areas tends to have more oscillations), while the distribution of temporal points tends to be sparse at those flat areas. It is important to combine the mesh refinement with a specific optimal control solver since for most real-world problems, the optimal state and control profiles may contain discontinuous points or non-smooth segments. If a fixed temporal set is used, it is usually hard to detect all the discontinuity or non-smoothness, and the obtained solution may fail to describe the true behaviours of the state and control variables.

C. *Verification of optimality*

After generating all the state and control profiles using the proposed solver, the next step is to verify the optimality of the obtained solutions. The conditions used to analyze the optimality were presented in Section IV. For the simulation results shown in Fig.2, the corresponding actual and demanded control profiles including angle of attack, bank angle, and thrust are plotted in Fig.5. As discussed in Section IV, since the demanded control variables that appear on the right hand side of the equations of motion are linear and do not appear in the path constraints, a "bang-bang" behaviour with respect to the demanded control profile can be expected. This behaviour can be used partly to show the optimality of the calculated trajectories. Moreover, by introducing the three lag equations (Eq.(3)), the actual controls become much smoother (no discontinuous point), which can offer more flexibilities in terms of the design of online guidance law.

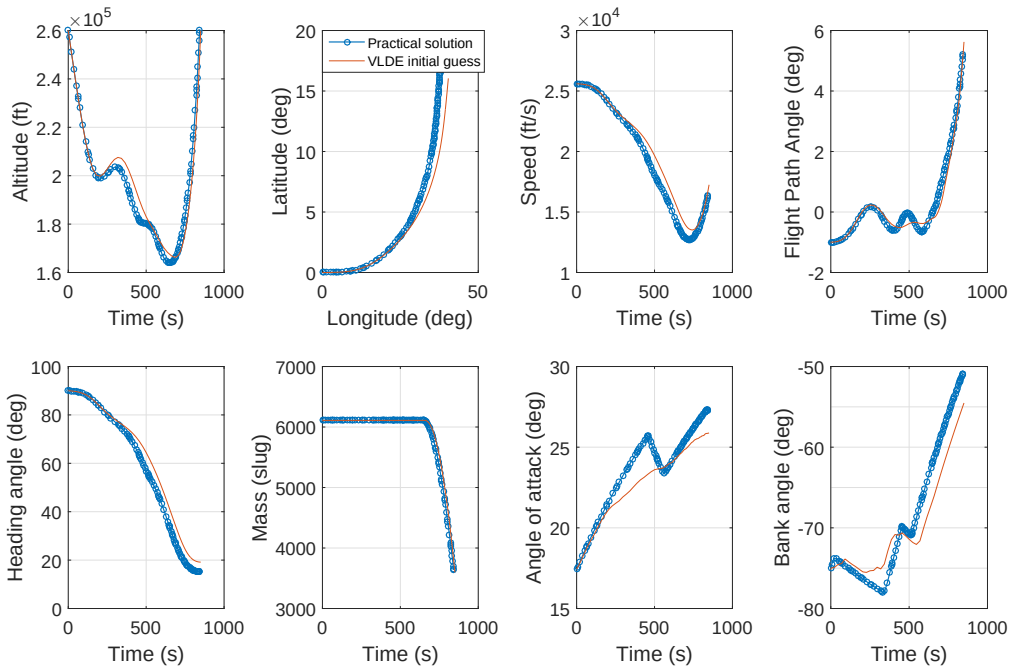

Fig. 3: Time history for states and controls (Mission 2: $Q_{max} = 150BTU$; $Pd_{max} = 11970.05Pa$; $nl_{max} = 2.0$)

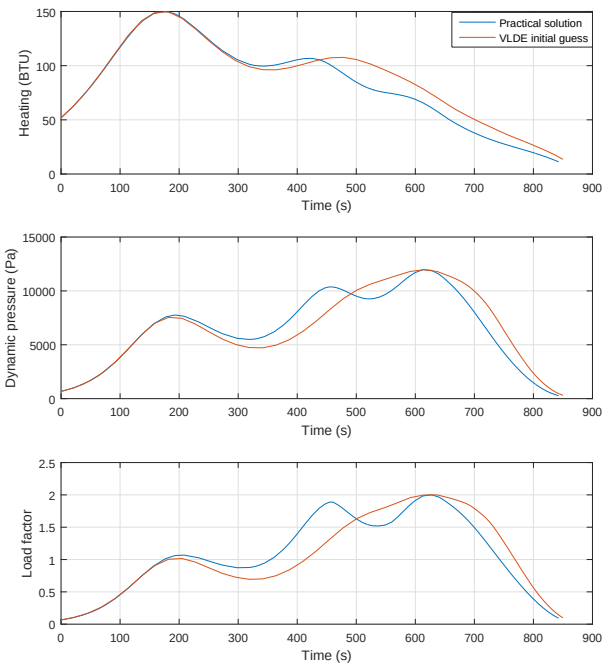

Fig. 4: Time history for three path constraints (Mission 2)

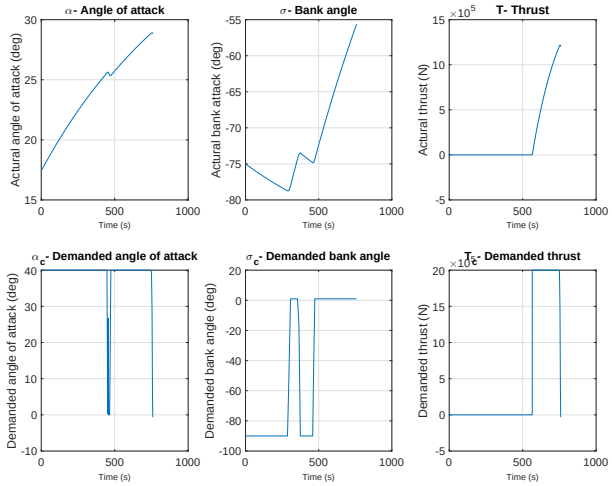

Fig. 5: Actual and demanded control profiles (Mission 1)

Moreover, to check the Terminal transversality conditions, the final values of the costate (λ) corresponding to the state obtained from the optimization process are calculated, such that $\lambda_\theta(t_f) = -5.33 \times 10^{-10}$, $\lambda_\phi(t_f) = -2.28 \times 10^{-9}$, $\lambda_V(t_f) = -4.93 \times 10^{-2}$, $\lambda_\psi(t_f) = 7.52 \times 10^{-10}$, $\lambda_\alpha(t_f) = 6.09 \times 10^{-21}$, $\lambda_\sigma(t_f) = 2.17 \times 10^{-17}$ and $\lambda_T(t_f) = -3.14 \times 10^{-23}$.

The Hamiltonian value condition shows that the Hamiltonian function should be -1 at the final time (i.e. $H_{am}(t_f) = -1$), see Section IV.C. From the Hamiltonian evolution equation, it can be shown that the Hamiltonian is constant during the whole time history. Therefore, combining these two conditions, the Hamiltonian should keep -1, as illustrated in Fig.6.

Fig.6 and the final costate values illustrate the optimality of the obtained solutions. Specifically, the value of the Hamil-

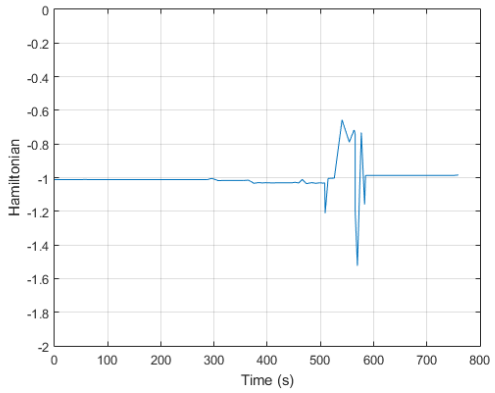

Fig. 6: Hamiltonian function for Mission Scenario 1

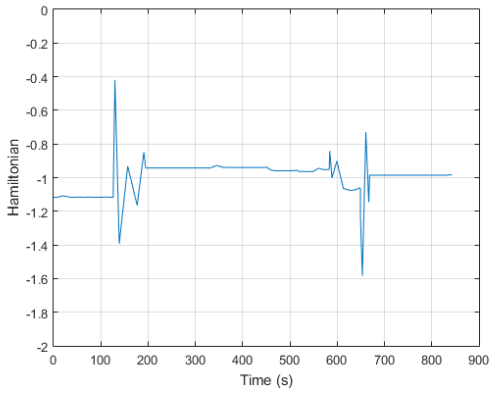

Fig. 8: Hamiltonian function for Mission Scenario 2

tonian function generated using the proposed approach tends to remain flat around -1 with small variance, while the final costate values for states which have free final conditions are equal to 0 or approximately 0. That means the performance of the designed strategy is closer to the theoretical behaviour and it further confirms that the newly designed solver is feasible and effective for handling SMV trajectory optimization problem.

For the mission scenario containing strict path constraints, the control profiles are plotted in Fig.7. It is clear that the numerical solution can still satisfy the control profile conditions discussed in Section IV. In terms of the Hamiltonian profile shown in Fig.8, again, the theoretical behaviours discussed in Section IV can be satisfied.

In order to verify Bellman's principle described in [31], and presented in Section IV of this paper, 15 time points on the obtained trajectory were selected randomly as initial conditions for the simulation. The results show that there is no better or different solution for both mission cases.

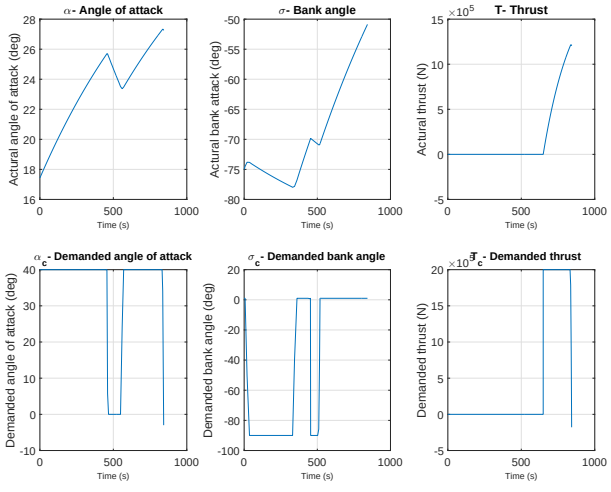

Fig. 7: Actual and demanded control profiles (Mission 2)

Remark 4. It is clear from Fig.6 and Fig.8 that the Hamiltonian profiles do not stay flat during the entire time period and contain some oscillations. For example, in Fig.6, there are some fluctuations around 500s, while in Fig.8, the trajectory tend to oscillate in the $[100s, 200s]$ and $[550s, 650s]$ regions. This is because during these time periods, the path constraints (see Fig.4) become active. That is, in the Hamiltonian equation, the multipliers $\mu^T(t)$ associated with C (path constraints) become non-zero. More precisely, at these time periods, the numerical solutions will lose some optimality.

Remark 5. One main advantage of the proposed optimal control solver is that it is possible for the user to analyze the optimality of the obtained solution. Since the inner solver is based on gradient-based optimization methods, it is possible to calculate the information about the Lagrange multipliers, which implies that the adjoint variable associated with each state can then be calculated via the mapping principle discussed in [7]. Therefore, the first-order necessary conditions for the general optimal control problems can be verified numerically.

D. Comparison with existing evolutionary solvers

In this subsection, comparative studies were performed to analyze the optimal trajectories achieved by applying the VLDE evolutionary solver with other typical heuristic strategies. For example, the PSO method reported in [19], and a DE design studied in [32]. These strategies use stochastic optimization processes and were shown as promising methods to calculate the optimal state and control trajectories. For the purpose of comparison, all the methods are applied to solve the first mission scenario. Fig.9 illustrates the time history with respect to the state variable calculated using different evolutionary algorithms, whereas Fig.10 gives the control and path constraint trajectories.

As can be seen from Fig.9 and Fig.10, all the solutions calculated by applying these three methods can be accepted as feasible solutions. The VLDE approach can perform better than its counterparts in terms of achieving better fitness values for this mission case. Moreover, it is obtained that the final boundary error values calculated using different global optimization methods are $e_{r_VLDE} = 0.08$ft, $e_{r_PSO} = 14.16$ft and $e_{r_DE} = 203.88$ft, respectively. This further confirms that the VLDE method investigated in this work can better lead the boundary error to a small value without violating various path constraints for the SMV trajectory planning problem.

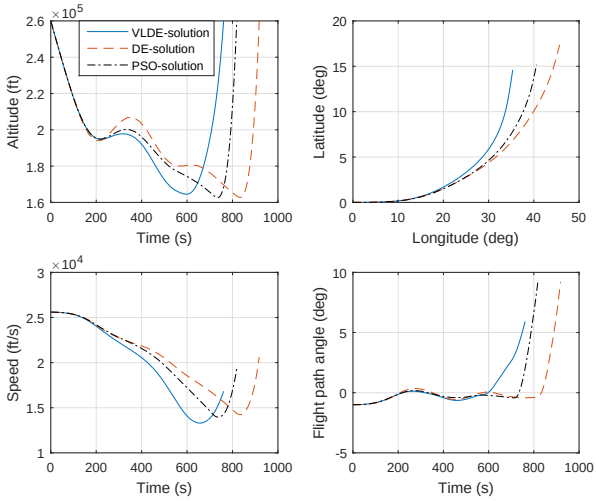

Fig. 9: Results obtained using different heuristic solvers

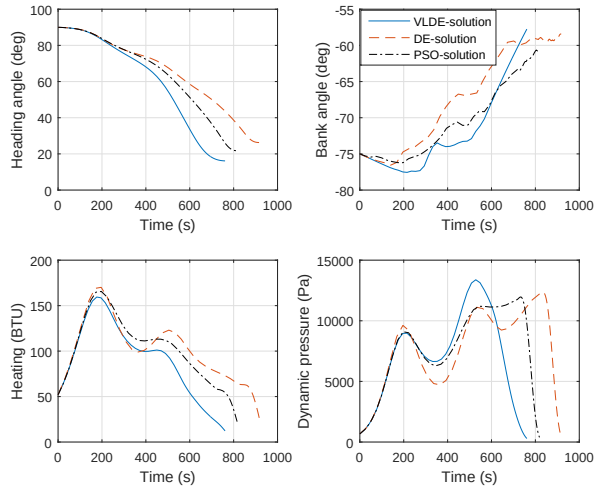

Fig. 10: Results obtained using different heuristic solvers

E. Dispersion models

The aim of dispersion simulations is to illustrate the stability of the proposed three-layer-hybrid optimal control solver in the presence of significant deviations in trajectory initial state variables (x_0) and vehicle uncertainties [33]. All the random initialization data used in the dispersion model are tabulated in Table.I. Furthermore, the vehicle mass was perturbed uniformly up to 5% with the nominal mass, which gives a range of value of $95\%m_0$ to $105\%m_0$.

The dispersion simulation was carried out for the Mission scenario 1 and 1000 Monte-Carlo simulations were performed. Simulation results show that most of the cases can successfully converge to the optimal solution and it is not sensitive with respect to the random initialization. The time histories of 100 dispersed trajectories in terms of the state variables obtained using the proposed hybrid optimal control algorithm are plotted in Fig.10. Correspondingly, Fig.11 shows the Hamiltonian profile in order to illustrate the optimality of the calculated

TABLE I: Dispersions in the entry initial conditions

State/parameter	Distribution	3-σ range
Altitude,ft	Zero-mean Gaussian	500
Longitude,deg	Zero-mean Gaussian	0.0749
Latitude,deg	Zero-mean Gaussian	0.3202
Velocity,ft/s	Zero-mean Gaussian	100
Flight-path angle,deg	Zero-mean Gaussian	0.1084
Heading angle,deg	Zero-mean Gaussian	0.0973
Vehicle's mass,slug	Uniform	$\pm 5\%$

solutions. As can be seen from Fig.10, all the trajectories can be accepted as feasible solutions. Moreover, based on the time history of the Hamiltonian function demonstrated in Fig.11, the optimality of the dispersion simulation can be guaranteed.

A comparative study was made in order to compare the convergence ability and stability of the proposed hybrid solver with other typical optimal control solvers such as GPOPS (e.g. orthogonal collocation method) and ICLOCS (direct collocation or multiple shooting method). By setting the maximum number of Newton iterations as 3000, the convergence results are tabulated in Table.II. The table summarises the times of optimal solution found, infeasible point detected and maximum iterations exceeded for different solvers. It is worth noting that the term "infeasible point detected" in Table.II means that the algorithm is stuck at a locally infeasible point.

As can be seen from Table.II based on the same dispersion model, the proposed hybrid solver can perform a higher probability in terms of finding optimal solutions compared with other typical strategies. A comparison between the algorithm with and without the layer 1 was also performed. The results show that the use of the layer 1 can effectively improve the convergence ability of the proposed hybrid algorithm. Moreover, the number of Newton iterations required for the proposed method is also smaller than its counterparts. The results not only confirm that the design philosophy of the proposed solver can have positive influences in terms of improving the convergence ability and convergence speed, but also indicate that the method designed in this study can have a better performance over other typical optimal control methods.

Remark 6. It is worth noting that the stochastic optimization approach used in this paper can be a feasible way to generate skip entry trajectories. This can be seen from the reference trajectories shown in Figs.2-4. Clearly, all the obtained solutions can be accepted as feasible solutions. If there is no accuracy requirement, then the initial guess can be accept as a near-optimal solution. In addition, when the nonlinearity of the cost functions or path constraints become higher, which means it is difficult to calculate the gradient information for gradient techniques, the stochastic method can also be an efficient way to generate a high quality reference trajectory.

Remark 7. It should be mentioned that the hybrid solver developed in this study can be easily extended and implemented in other control optimization problems. For instance, recent

TABLE II: Convergence results for different optimal control solvers

Different methods	Optimal solution found	Infeasible point detected	Maximum iterations exceeded	Successful rate (%)
Proposed hybrid strategy (With layer 1)	907	24	69	90.7
Proposed hybrid strategy (Without layer 1)	811	33	56	81.1
GPOPS (Orthogonal collocation)	798	53	149	79.8
ICLOCS (Direct collocation)	663	69	268	66.3
ICLOCS (Multiple shooting)	691	77	232	69.1

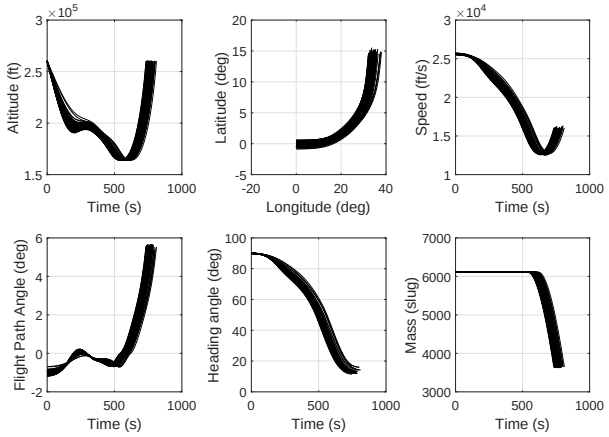

Fig. 11: 100 dispersed trajectories for the state variables

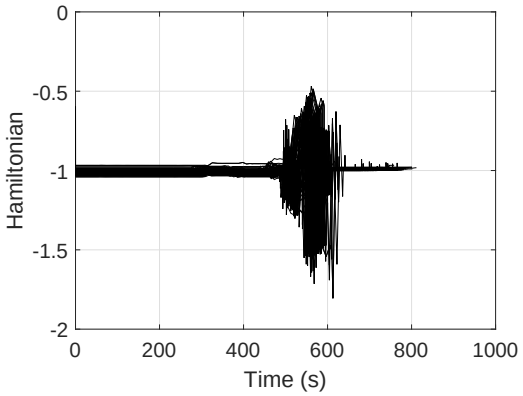

Fig. 12: 100 dispersed trajectories for the Hamiltonian function

research [34] has demonstrated the capability to utilize optimal control-based solver in dealing with agent/robot obstacle-avoidance path planning problems [35]. If the proposed solver is applied to solve such a problem, some adjustments should be made in the optimization process. Firstly, different obstacles should be reformulated as a series of path constraints entailed in the optimization model. Besides, since the nonlinearity of the dynamics and constraints in these problems is usually high, the problem tends to be more sensitive with respect to the optimization parameters. Therefore, the proposed solver should start with a relatively small temporal set.

VI. Conclusion

In this work, a three-layer-hybrid optimal control solver was constructed and applied to solve the Space Maneuver Vehicle trajectory optimization problem. In order to effectively evaluate a reference trajectory, an initial guess generator using V-based adaptive differential evolution algorithm was applied. Since it is hard to use a large temporal set for stochastic-based algorithms, a specific discretization scheme is implemented to tackle this problem. In addition, a new gradient-based algorithm is used as the inner solver, thereby allowing the designer more flexibility to control the optimization process. Comparative simulations show that the proposed method can have better performance in terms of convergence ability and stability than other typical optimal control solvers. Moreover, it was verified that the obtained solution can satisfy all the optimality conditions, which illustrates the effectiveness of the proposed solver.

Our follow-up research will focus on extending the proposed design such that it can be applied in handling stochastic trajectory planning problems. In addition, it would be also worthwhile to use the proposed method for other path planning applications, especially in the multi-agent path planning with collision avoidance.

References

[1] C. Lian, X. Xu, H. Chen, and H. He, "Near-optimal tracking control of mobile robots via receding-horizon dual heuristic programming," *IEEE Transactions on Cybernetics*, vol. 46, no. 11, pp. 2484–2496, 2016.

[2] R. Ruiz-Cruz, E. N. Sanchez, F. Ornelas-Tellez, A. G. Loukianov, and R. G. Harley, "Particle swarm optimization for discrete-time inverse optimal control of a doubly fed induction generator," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1698–1709, 2013.

[3] M. F. Corapsiz and K. Erenturk, "Trajectory tracking control and contouring performance of three-dimensional cnc," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2212–2220, 2016.

[4] Q. Wei, D. Liu, and H. Lin, "Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems," *IEEE Transactions on Cybernetics*, vol. 46, no. 3, pp. 840–853, 2016.

[5] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Cambridge University Press, 2009.

[6] B. A. Conway, "A survey of methods available for the numerical optimization of continuous dynamic systems," *Journal of Optimization Theory and Applications*, vol. 152, no. 2, pp. 271–306, 2012.

[7] D. A. Benson, G. T. Huntington, T. P. Thorvaldsen, and A. V. Rao, "Direct trajectory optimization and costate estimation via an orthogonal collocation method," *Journal of Guidance, Control, and Dynamics*, vol. 29, no. 6, pp. 1435–1440, 2006.

[8] W. W. Hager, "Runge-kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, vol. 87, no. 2, pp. 247–282, 2000.

[9] P. Williams, "Jacobi pseudospectral method for solving optimal control problems," *Journal of Guidance, Control, and Dynamics*, vol. 27, no. 2, pp. 293–297, 2004.

[10] O. A. Yakimenko, "Direct method for rapid prototyping of near-optimal aircraft trajectories," *Journal of Guidance, Control, and Dynamics*, vol. 23, no. 5, pp. 865–875, 2000.

[11] T. Albin, D. Ritter, N. Liberda, R. Quirynen, and M. Diehl, "In-vehicle realization of nonlinear mpc for gasoline two-stage turbocharging airpath control," *IEEE Transactions on Control Systems Technology*, vol. PP, no. 99, pp. 1–13, 2017.

[12] R. Genest and J. V. Ringwood, "Receding horizon pseudospectral control for energy maximization with application to wave energy devices," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 29–38, 2017.

[13] D. Christopher, H. William, and R. Anil, *An Improved Adaptive hp Algorithm using Pseudospectral Methods for Optimal Control*, ser. Guidance, Navigation, and Control and Co-located Conferences. American Institute of Aeronautics and Astronautics, 2010.

[14] M. Heinkenschloss and D. Ridzal, "A matrix-free trust-region sqp method for equality constrained optimization," *SIAM Journal on Optimization*, vol. 24, no. 3, pp. 1507–1541, 2014.

[15] J. Laurent-Varin, F. Bonnans, N. Berend, M. Haddou, and C. Talbot, "Interior-point approach to trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 5, pp. 1228–1238, 2007.

[16] B. Tian, W. Fan, R. Su, and Q. Zong, "Real-time trajectory and attitude coordination control for reusable launch vehicle in reentry phase," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1639–1650, 2015.

[17] N. Yokoyama and S. Suzuki, "Modified genetic algorithm for constrained trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 28, no. 1, pp. 139–144, 2005.

[18] J. A. Englander and B. A. Conway, "Automated solution of the low-thrust interplanetary trajectory problem," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 1, pp. 15–27, 2017.

[19] J. J. Kim and J. J. Lee, "Trajectory optimization with particle swarm optimization for manipulator motion planning," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 620–631, 2015.

[20] P. C. Roy, M. M. Islam, K. Murase, and X. Yao, "Evolutionary path control strategy for solving many-objective optimization problem," *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 702–715, 2015.

[21] R. Chai, A. Savvaris, and A. Tsourdos, "Violation learning differential evolution-based hp-adaptive pseudospectral method for trajectory optimization of space maneuver vehicle," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 4, pp. 2031–2044, 2017.

[22] ——, "Fuzzy physical programming for space manoeuvre vehicles trajectory optimization based on hp-adaptive pseudospectral method," *Acta Astronautica*, vol. 123, pp. 62–70, 2016.

[23] S. M. Elsayed, R. A. Sarker, and D. L. Essam, "An improved self-adaptive differential evolution algorithm for optimization problems," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 89–99, 2013.

[24] D. Liu, X. Yang, D. Wang, and Q. Wei, "Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints," *IEEE Transactions on Cybernetics*, vol. 45, no. 7, pp. 1372–1385, 2015.

[25] Y. J. Gong, J. J. Li, Y. Zhou, Y. Li, H. S. H. Chung, Y. H. Shi, and J. Zhang, "Genetic learning particle swarm optimization," *IEEE Transactions on Cybernetics*, vol. 46, no. 10, pp. 2277–2290, 2016.

[26] W. W. Hager, H. Hou, and A. V. Rao, "Convergence rate for a radau collocation method applied to unconstrained optimal control," *Journal of Optimization Theory and Applications*, pp. 1–24, 2015.

[27] ——, "Convergence rate for a gauss collocation method applied to unconstrained optimal control," *Journal of Optimization Theory and Applications*, pp. 1–24, 2016.

[28] R. Chai, A. Savvaris, A. Tsourdos, S. Chai, and Y. Xia, "Improved gradient-based algorithm for solving aeroassisted vehicle trajectory optimization problems," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 8, pp. 2093–2101, 2017.

[29] J. T. Betts and W. P. Huffman, "Mesh refinement in direct transcription methods for optimal control," *Optimal Control Applications and Methods*, vol. 19, no. 1, pp. 1–21, 1998.

[30] F. Liu, W. W. Hager, and A. V. Rao, "Adaptive mesh refinement method for optimal control using nonsmoothness detection and mesh size reduction," *Journal of the Franklin Institute*, vol. 352, no. 10, pp. 4081–4106, 2015.

[31] B. Kevin, R. Michael, and D. David, *Optimal Nonlinear Feedback Guidance for Reentry Vehicles*, ser. Guidance, Navigation, and Control and Co-located Conferences. American Institute of Aeronautics and Astronautics, 2006, doi:10.2514/6.2006-6074.

[32] M. Ergezer and D. Simon, "Mathematical and experimental analyses of oppositional algorithms," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2178–2189, 2014.

[33] Y. Xia, F. Pu, S. Li, and Y. Gao, "Lateral path tracking control of autonomous land vehicle based on adrc and differential flatness," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3091–3099, 2016.

[34] A. J. Hausler, A. Saccon, A. P. Aguiar, J. Hauser, and A. M. Pascoal, "Energy-optimal motion planning for multiple robotic vehicles with collision avoidance," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 867–883, 2016.

[35] Z. Chen and H. T. Zhang, "A minimal control multiagent for collision avoidance and velocity alignment," *IEEE Transactions on Cybernetics*, vol. 47, no. 8, pp. 2185–2192, 2017.

Runqi Chai (SM'15) was born in Beijing, China, in 1993. He is currently pursuing the Ph.D. degree in Aerospace Engineering at Cranfield University. His research interests include trajectory optimization, guidance and control.

Al Savvaris (M'08) received the M.Eng. in aerospace systems engineering from the University of Hertfordshire, in 1998, and a Ph.D. on radiowave propagation and system design from University of South Wales, in 2004. He is a Reader in the Centre for Cyber-Physical Systems at Cranfield University. He established the Autonomous Vehicle Dynamics and Control MSc course and the COMAC training programme at Cranfield. His research interests and activities include systems integration, hybrid energy management, communication systems, embedded systems, guidance and control. Currently he is working on InnovateUK funded AirStart and USMOOTH projects. In the past he worked on the FLAVIIR and ASTRAEA UAS projects, developing new technologies for unmanned systems; working on hardware and system integration. He participated in FP6, in the scope of the FLYSAFE project, working on next generation integrated safety systems. He was a member of the Autonomous Systems National Technical Committee; EPSRC College review member and reviewer on several international publications including IMechE and IEEE. He has published over 100 peer-reviewed journal and conference papers.

Antonios Tsourdos (M'99) obtained an M.Eng. on electronic, control and systems engineering from the University of Sheffield, in 1995, an M.Sc. on systems engineering from Cardiff University, in 1996, and a Ph.D. on nonlinear robust missile autopilot design and analysis from Cranfield University, in 1999. He is currently a professor of control systems at Cranfield University. Appointed Head of the Autonomous Systems Group in 2007. Professor Tsourdos was member of the Team Stellar, the winning team for the UK MoD Grand Challenge (2008) and the IET Innovation Award (Category Team, 2009). Antonios is an editorial board member on several international publications including IMechE and IEEE. He is member if the IFAC Technical Committee on Intelligent Autonomous Vehicles, the IET Executive Team on Robotics and Automation and the ATI Autonomous Systems National Technical Committee. He was also involved in the SEAS DTC on Autonomous Systems Verifications. Professor Tsourdos has also been engaged in research on guidance & control for single and multiple vehicles as well as verifiable autonomy of autonomous systems and lately dealing with the newly-important subjects of integrated system health management and cyber-physical systems. He has published over 100 peer-reviewed journal and conference papers.

Senchun Chai was born in Beijing, China. He received his Ph.D. degree and then held a Postdoc fellowship with School of Electronics, University of Glamorgan, UK, in 2007 and 2009, respectively. He joined the School of Automation in September 2010, where he is currently an associate Professor in control. He was a researcher at Cranfield University, UK, from 2009 to 2010, and was a visiting scholar at University of Illinois at Urbana-Champaign Urbana, USA, from January 2010 to May 2010. His current research interests include design of unmanned aerial vehicles (UAV), wireless sensor network control, networked control systems and multi-agent control systems.

Yuanqing Xia (M'15-SM'16) was born in Anhui Province, China, in 1971. He received the B.S. degree from the Department of Mathematics, Chuzhou University, Chuzhou, China, in 1991, the M.S. degree in fundamental mathematics from Anhui University, Wuhu, China, in 1998, and the Ph.D. degree in control theory and control engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2001. His current research interests are in the fields of networked control systems, robust control and signal processing, active disturbance rejection control and flight control. He has published 8 monographs with Springer and Wiley, and more than 200 papers in journals. He has obtained Second Award of the Beijing Municipal Science and Technology (No. 1) in 2010, Second National Award for Science and Technology (No. 2) in 2011, and Second Natural Science Award of The Ministry of Education (No. 1) in 2012. He is a Deputy Editor of the Journal of the Beijing Institute of Technology, Associate Editor of Acta Automatica Sinica, Control Theory and Applications, the International Journal of Innovative Computing, Information and Control, and the International Journal of Automation and Computing.

https://dspace.org/

Trajectory optimization of space maneuver vehicle using a hybrid optimal control solver

Chai, Runqi; Savvaris, Al; Tsourdos, Antonios; Chai, Senchun; Xia, Yuanqing

2017-12-12
IEEE
http://dx.doi.org/10.1109/TCYB.2017.2778195
Chai R, Savvaris A, Tsourdos A, Chai S, Xia Y, Trajectory optimization of space maneuver vehicle using a hybrid optimal control solver, IEEE Transactions on Cybernetics, Volume 49, Issue 2, February 2019, pp. 467-480

Downloaded from a DSpace repository. https://dspace.org/