

Dynamic optimization for robust path planning of horizontal oil wells

Zhaohua Gong^a, Ryan Loxton^b, Changjun Yu^b, Kok Lay Teo^b

^a*School of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai, China*

^b*Department of Mathematics and Statistics, Curtin University, Perth, Australia*

Abstract

This paper considers the three-dimensional path planning problem for horizontal oil wells. The decision variables in this problem are the curvature, tool-face angle and switching points for each turn segment in the path, and the optimization objective is to minimize the path length and target error. The optimal curvatures, tool-face angles and switching points can be readily determined using existing gradient-based dynamic optimization techniques. However, in a real drilling process, the actual curvatures and tool-face angles will inevitably deviate from the planned optimal values, thus causing an unexpected increase in the target error. This is a critical challenge that must be overcome for successful practical implementation. Accordingly, this paper introduces a sensitivity function that measures the rate of change in the target error with respect to the curvature and tool-face angle of each turn segment. Based on the sensitivity function, we propose a new optimization problem in which the switching points are adjusted to minimize target error sensitivity subject to continuous state inequality constraints arising from engineering specifications, and an additional constraint specifying the maximum allowable increase in the path length from the optimal value. Our main result shows that the sensitivity function can be evaluated by solving a set of auxiliary dynamic systems. By combining this result with the well-known time-scaling transformation, we obtain an equivalent transformed problem that can be solved using standard nonlinear programming algorithms. Finally, the paper concludes with a numerical example involving a practical path planning problem for a Ci-16-Cp146 well.

Keywords: Switched system; Parameter optimization; System sensitivity; Horizontal well; Time-scaling transformation

Email address: zhaohuagong@163.com (Zhaohua Gong)

1. Introduction

The path planning problem for oil wells involves designing an optimal drill path, subject to various engineering specifications, to connect given start and end points. There are many papers devoted to this problem (see, for example, references [2, 4, 6, 15]). The path planning problem can, in fact, be formulated as an optimal switching control problem and thereafter solved using standard gradient-based dynamic optimization techniques [9, 10]. In particular, the optimal control software packages MISER 3 [5] and Visual MISER [19] are both applicable.

In practical path planning problems, the path usually consists of a combination of turn/straight segments, where the curvature and tool-face angle (both decision variables) are constant in each segment. Although the segment arc lengths can be measured and controlled with high precision from the surface, this is not the case for the curvatures and tool-face angles. Indeed, the actual curvatures and tool-face angles will likely deviate from the optimal values during the real drilling process. Consequently, the well trajectory will deviate from the optimal trajectory, potentially leading to a large target error.

We are only aware of two papers (references [7] and [8]) that consider the issue of target error sensitivity with respect to inaccuracies in the curvatures and tool-face angles. In reference [7], the dispersion between the actual path and the optimal path is modelled by a stochastic perturbation in the dynamic system describing the path trajectory. In reference [8], the dispersion is instead modelled by a series of stochastic impulsive jumps applied at the turn segment end-points. Both references [7] and [8] apply the Hookes-Jeeves algorithm to solve a series of nonlinear optimization problems generated by realizations of the random variables modelling the dispersion. Thus, the methods proposed in these references require solving multiple optimization problems, not just one. Moreover, the cost and constraint gradients for these optimization problems are not provided in references [7] and [8], and hence fast gradient-based optimization algorithms such as sequential quadratic programming are not applicable.

In any dynamic system, small changes in the system parameters can cause large changes in the system cost. Therefore, when the system contains parameters whose values are uncertain (whether due to model inaccuracies or implementation errors), the sensitivity of the cost with respect to these parameters should be considered [14]. To this end, reference [16] proposes a novel optimal control problem in which the control is chosen to minimize the weighted sum of system cost and system sensitivity. This problem can be solved using the control parameterization technique [17], whereby the control function is discretized to yield an approximate optimal parameter selection problem. However, to generate the cost and constraint gradients in the approximate problem (as required to solve the problem using nonlinear programming methods), four dynamic systems must be solved—the state system, the costate system and two other dynamic systems. This contrasts with standard problems in which only the state and costate systems need to be solved. The method proposed in [16] is cumbersome to implement numerically and cannot be automated using the

MISER 3 software without significant changes to the internal source code.

Reference [13] introduces an alternative computational method for solving the same sensitivity optimal control problem considered in reference [16]. In this new method, the system sensitivity is computed by solving a set of auxiliary dynamic systems, and the non-standard sensitivity problem is converted into a standard optimal control problem that can be solved using standard techniques. This idea is extended to impulsive systems in [18].

The purpose of this paper is to apply the sensitivity penalization ideas introduced in references [13, 16, 18] to the path planning problem for drilling three-dimensional horizontal oil wells. Real drilling operations are susceptible to implementation inaccuracies in the segment curvatures and tool-face angles due to the difficulty in measuring these values accurately from the surface. Thus, our goal is to minimize the target error sensitivity with respect to the curvature and tool-face angle of each turn segment, subject to a constraint specifying the maximum path length compared with the theoretical minimum value. Note that the sensitivity optimal control problems in references [13, 16, 18] do not consider state constraints, but such constraints are present in the path planning problem due to practical requirements on the path. The computational algorithm we develop is capable of handling these constraints.

The remainder of this paper is organized as follows. In Section 2, both the traditional path planning problem and the new sensitivity problem are defined mathematically. In Section 3, a computational method is developed for evaluating the target error sensitivity. Then, in Section 4, this method is combined with a novel transformation procedure to convert the sensitivity problem into a tractable form. Gradient formulas are derived in Section 5. Simulation results for a real Ci-16-Cp146 well are presented in Section 6. Finally, some concluding remarks are given in Section 7.

2. Problem formulation

2.1. Path planning problem

Consider the three-dimensional path planning problem for horizontal oil wells described in reference [2]. Let s be an independent variable representing the distance along the path. Furthermore, let $\mathbf{x}(s) = (x_1(s), \dots, x_5(s))^T$ be the state vector whose components represent, respectively, the inclination, azimuth, and Cartesian coordinates of the point located at distance s along the path.

The model is based on the following hypotheses.

- (H₁). The path is a combination of n smooth turn segments.
- (H₂). The curvature and tool-face angle are constant in each turn segment.
- (H₃). The path contains no vertical segments.

Let τ_i , $i = 0, 1, \dots, n$, denote the switching points at which the path changes from one turn segment to another, where $\tau_0 = 0$ corresponds to the path's origin, and τ_n corresponds to the path's terminus. Under the above hypotheses, the dynamic system describing the path is

$$\dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\xi}^i), \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \quad (1)$$

with the initial and intermediate conditions:

$$\mathbf{x}(\tau_i) = \begin{cases} \mathbf{x}^0, & \text{if } i = 0, \\ \mathbf{x}(\tau_{i-}), & \text{if } i = 1, 2, \dots, n, \end{cases} \quad (2)$$

where $\boldsymbol{\xi}^i = (\xi_1^i, \xi_2^i)^\top$ is a vector containing the curvature and tool-face angle of the i th turn segment, \mathbf{x}^0 is a given start point, and

$$\mathbf{f}(\mathbf{x}(s), \boldsymbol{\xi}^i) = \begin{pmatrix} \xi_1^i \cos \xi_2^i \\ \xi_1^i \sin \xi_2^i / \sin x_1(s) \\ \sin x_1(s) \cos x_2(s) \\ \sin x_1(s) \sin x_2(s) \\ \cos x_1(s) \end{pmatrix}.$$

Due to physical limitations, the curvatures and tool-face angles are subject to the following constraints:

$$a_1 \leq \xi_1^i \leq b_1, \quad a_2 \leq \xi_2^i \leq b_2, \quad i = 1, 2, \dots, n, \quad (3)$$

where a_1, b_1, a_2, b_2 are given real numbers such that $0 \leq a_1 < b_1$ and $-\frac{1}{2}\pi < a_2 < b_2 < \frac{1}{2}\pi$. Let $\boldsymbol{\xi} = (\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^n)^\top \in R^{2n}$. Any $\boldsymbol{\xi} \in R^{2n}$ satisfying (3) is called an admissible parameter vector. Let Ξ denote the set of all admissible parameter vectors.

In addition to (3), we also have the following constraints on the turn segment lengths:

$$c_1 \leq \tau_i - \tau_{i-1} \leq d_1, \quad i = 1, 2, \dots, n, \quad (4)$$

where c_1 and d_1 are given real numbers such that $0 < c_1 < d_1$. Let $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)^\top$. Any $\boldsymbol{\tau} \in R^n$ satisfying (4) is called an admissible switching point vector. Let Γ denote the set of all admissible switching point vectors. Accordingly, any pair $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$ is called an admissible pair for system (1)-(2).

Let $\mathbf{x}(\cdot | \boldsymbol{\tau}, \boldsymbol{\xi})$ denote the solution of (1)-(2) corresponding to $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$. Since we only consider horizontal paths without vertical segments (as per hypothesis (H₃)), the solution $\mathbf{x}(\cdot | \boldsymbol{\tau}, \boldsymbol{\xi})$ must satisfy the following continuous state inequality constraints:

$$\omega_0 \leq x_1(s | \boldsymbol{\tau}, \boldsymbol{\xi}) \leq \frac{1}{2}\pi, \quad s \in [0, \tau_n], \quad (5a)$$

$$0 \leq x_2(s | \boldsymbol{\tau}, \boldsymbol{\xi}) \leq 2\pi, \quad s \in [0, \tau_n], \quad (5b)$$

where $\omega_0 > 0$ is a given constant.

The path should connect the start point \mathbf{x}^0 to a given target point \mathbf{x}^f . In addition, the path length should also be minimized considering the economic factor. This is done by adjusting the switching points, curvatures and tool-face angles to minimize the following weighted sum of path length and target error:

$$H(\boldsymbol{\tau}, \boldsymbol{\xi}) = \tau_n + \alpha \|\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}) - \mathbf{x}^f\|^2, \quad (6)$$

where α is a given weight coefficient. This problem can be formulated as the following dynamic optimization problem.

Problem (A). Given the dynamic system (1)-(2), choose an admissible pair $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$ such that the cost function (6) is minimized subject to the continuous state inequality constraints (5).

2.2. Sensitivity problem

Problem (A) can be solved using gradient-based dynamic optimization techniques such as those described in [9] and [10]. These techniques have been implemented as part of the optimal control software packages MISER 3 [5] and Visual MISER [19]. Let $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*) \in \Gamma \times \Xi$ be an optimal pair for Problem (A), where $\boldsymbol{\xi}^* = (\xi^{1*}, \dots, \xi^{n*})^\top \in R^{2n}$ and $\boldsymbol{\tau}^* = (\tau_1^*, \dots, \tau_n^*)^\top \in R^n$. In practice, due to limitations with the drilling equipment, it is impossible to implement the solution $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*)$ with perfect accuracy. In particular, the curvatures and tool-face angles cannot be measured accurately from the surface, and thus they will likely deviate from the optimal values during the drilling process. This will, in turn, cause the actual path to deviate from the optimal trajectory. Small inaccuracies in the curvatures and tool-face angles can propagate along the path and cause the target error to increase significantly. Therefore, the switching points must be chosen so that the target error is insensitive to small perturbations in $\boldsymbol{\xi}^*$. For this purpose, we introduce the following sensitivity function, which should be minimized with respect to $\boldsymbol{\tau}$:

$$J(\boldsymbol{\tau}) = \left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right) \left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right)^\top, \quad (7)$$

where

$$G(\boldsymbol{\tau}, \boldsymbol{\xi}^*) = \|\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*) - \mathbf{x}^f\|^2,$$

and

$$\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} = \left. \frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \right|_{\boldsymbol{\xi}=\boldsymbol{\xi}^*}.$$

With $\boldsymbol{\xi} = \boldsymbol{\xi}^*$, the continuous state inequality constraints (5) become

$$\omega_0 \leq x_1(s | \boldsymbol{\tau}, \boldsymbol{\xi}^*) \leq \frac{1}{2}\pi, \quad s \in [0, \tau_n], \quad (8a)$$

$$0 \leq x_2(s | \boldsymbol{\tau}, \boldsymbol{\xi}^*) \leq 2\pi, \quad s \in [0, \tau_n]. \quad (8b)$$

Moreover, we impose the following constraint to ensure that the new path length is sufficiently close to the optimal value:

$$\tau_n \leq (1 + \beta)\tau_n^*, \quad (9)$$

where $\beta > 0$ is a given constant specifying the maximum allowable increase in the path length from the optimal value.

Thus, the new dynamic optimization problem, which takes target error sensitivity into account, can be stated as follows.

Problem (B). Given an optimal solution $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*)$ for Problem (A) and the dynamic system (1)-(2) with $\boldsymbol{\xi} = \boldsymbol{\xi}^*$, choose an admissible switching point vector $\boldsymbol{\tau} \in \Gamma$ such that the target error sensitivity function (7) is minimized subject to the continuous state inequality constraints (8) and the additional inequality constraint (9).

For Problem (B), conventional dynamic optimization techniques are not directly applicable due to the non-standard sensitivity function (7). In the next section, we derive a tractable method for computing this sensitivity function.

3. Computing the sensitivity function

We first present conditions under which the state trajectory $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ exists and is unique. This is crucial for Problems (A) and (B) in Section 2 to be well-defined.

Theorem 1. Let $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$ be an admissible pair satisfying

$$0 < x_1^0 + \sum_{j=1}^i \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}) < \pi, \quad i = 0, 1, \dots, n. \quad (10)$$

Then there exists a unique state trajectory $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ for system (1)-(2) corresponding to $(\boldsymbol{\tau}, \boldsymbol{\xi})$.

Proof. The first state equation in system (1)-(2) can be integrated to yield

$$x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) = x_1^0 + \sum_{j=1}^{i-1} \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}) + \xi_1^i \cos \xi_2^i (s - \tau_{i-1}),$$

$$s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n.$$

Evaluating this equation at the switching points gives

$$x_1(\tau_i|\boldsymbol{\tau}, \boldsymbol{\xi}) = x_1^0 + \sum_{j=1}^i \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}), \quad i = 0, 1, \dots, n.$$

Condition (10) then implies that $0 < x_1(\tau_i|\boldsymbol{\tau}, \boldsymbol{\xi}) < \pi$ for each $i = 0, 1, \dots, n$. Hence, since $x_1(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ is piecewise-linear with break-points at $s = \tau_i$, $i = 0, 1, \dots, n$, we have $0 < x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) < \pi$ for all $s \in [0, \tau_n]$. This shows that, under condition (10), the right-hand side of (1) is well-defined.

Now, let $\rho_0 \in (0, \pi)$ be such that

$$\rho_0 \leq x_1^0 + \sum_{j=1}^i \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}) \leq \pi - \rho_0, \quad i = 0, 1, \dots, n.$$

Then clearly $\rho_0 \leq x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \leq \pi - \rho_0$ for all $s \in [0, \tau_n]$. Consider the last four state equations in (1)-(2):

$$\dot{\boldsymbol{x}}(s) = \hat{\boldsymbol{f}}(s, \hat{\boldsymbol{x}}(s), \boldsymbol{\xi}^i), \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \quad (11)$$

with the initial and intermediate conditions:

$$\hat{\boldsymbol{x}}(\tau_i) = \begin{cases} \hat{\boldsymbol{x}}^0, & \text{if } i = 0, \\ \hat{\boldsymbol{x}}(\tau_i^-), & \text{if } i = 1, 2, \dots, n, \end{cases} \quad (12)$$

where

$$\hat{\boldsymbol{x}}(s) = \begin{pmatrix} x_2(s) \\ x_3(s) \\ x_4(s) \\ x_5(s) \end{pmatrix}, \quad \hat{\boldsymbol{x}}^0 = \begin{pmatrix} x_2^0 \\ x_3^0 \\ x_4^0 \\ x_5^0 \end{pmatrix},$$

and

$$\hat{\boldsymbol{f}}(s, \hat{\boldsymbol{x}}(s), \boldsymbol{\xi}^i) = \begin{pmatrix} \xi_1^i \sin \xi_2^i / \sin x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \\ \sin x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \cos x_2(s) \\ \sin x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \sin x_2(s) \\ \cos x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \end{pmatrix}.$$

Clearly, $\hat{\boldsymbol{f}}$ is continuous with respect to s and $\hat{\boldsymbol{x}}$. Furthermore, for any $s \in [0, \tau_n]$ and $\hat{\boldsymbol{x}} = (x_2, x_3, x_4, x_5)^\top \in R^4$,

$$\begin{aligned} \|\hat{\boldsymbol{f}}(s, \hat{\boldsymbol{x}}, \boldsymbol{\xi}^i)\|^2 &= \frac{(\xi_1^i)^2 \sin^2 \xi_2^i}{\sin^2 x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi})} + \sin^2 x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \cos^2 x_2 \\ &\quad + \sin^2 x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \sin^2 x_2 + \cos^2 x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}) \\ &\leq \frac{b_1^2}{\sin^2 \rho_0} + 3. \end{aligned}$$

Thus, $\hat{\boldsymbol{f}}$ is uniformly bounded. Based on the results in [1], system (11)-(12) has a unique solution $\hat{\boldsymbol{x}}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$. Taking $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi}) = (x_1(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi}), \hat{\boldsymbol{x}}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})^\top)^\top$ completes the proof. \square

Let $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$ be a feasible pair for Problem (A) and consider fixed integers $k \in \{1, \dots, n\}$ and $l \in \{1, 2\}$. For notational simplicity, let $\boldsymbol{x}(\cdot) = \boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ and $\boldsymbol{x}^\epsilon(\cdot) = \boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi} + \epsilon \boldsymbol{e}_{2(k-1)+l})$, where $\boldsymbol{e}_{2(k-1)+l} \in R^{2n}$ denotes the unit basis vector whose $(2(k-1)+l)$ th component is one, and all other components are zero. Note that the i th subvector in $\boldsymbol{\xi} + \epsilon \boldsymbol{e}_{2(k-1)+l}$ (i.e., the subvector corresponding to $\boldsymbol{\xi}^i$) is

$$\begin{aligned} (\boldsymbol{\xi} + \epsilon \boldsymbol{e}_{2(k-1)+l})^i &= \begin{cases} \boldsymbol{\xi}^i, & \text{if } i \neq k, \\ (\xi_1^i + \epsilon, \xi_2^i)^\top, & \text{if } i = k \text{ and } l = 1, \\ (\xi_1^i, \xi_2^i + \epsilon)^\top, & \text{if } i = k \text{ and } l = 2, \end{cases} \\ &= (\xi_1^i + \epsilon \delta_{l1} \delta_{ik}, \xi_2^i + \epsilon \delta_{l2} \delta_{ik})^\top, \end{aligned}$$

where δ_{l1} , δ_{l2} , and δ_{ik} are the respective Kronecker delta functions.

Since $(\boldsymbol{\tau}, \boldsymbol{\xi})$ must satisfy the continuous state inequality constraints (5a), we have

$$\omega_0 \leq x_1(\tau_i | \boldsymbol{\tau}, \boldsymbol{\xi}) = x_1^0 + \sum_{j=1}^i \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}) \leq \frac{1}{2} \pi, \quad i = 0, 1, \dots, n.$$

Thus, since $\xi_1^j + \epsilon \delta_{l1} \delta_{jk} \rightarrow \xi_1^j$ and $\cos(\xi_2^j + \epsilon \delta_{l2} \delta_{jk}) \rightarrow \cos \xi_2^j$ as $\epsilon \rightarrow 0$, there exists a constant $\rho_1 \in (0, \pi)$ and an interval $\Delta \subset [a_l - \xi_l^k, b_l - \xi_l^k]$ of positive Lebesgue measure such that for all $\epsilon \in \Delta$,

$$\rho_1 \leq x_1^0 + \sum_{j=1}^i (\xi_1^j + \epsilon \delta_{l1} \delta_{jk}) \cos(\xi_2^j + \epsilon \delta_{l2} \delta_{jk}) (\tau_j - \tau_{j-1}) \leq \pi - \rho_1, \quad i = 0, 1, \dots, n. \quad (13)$$

We may assume that $0 \in \Delta$. Moreover, since $\Delta \subset [a_l - \xi_l^k, b_l - \xi_l^k]$,

$$\epsilon \in \Delta \implies \boldsymbol{\xi} + \epsilon \mathbf{e}_{2(k-1)+l} \in \Xi.$$

The following lemma gives a uniform bound for the solution of system (1)-(2) corresponding to $(\boldsymbol{\tau}, \boldsymbol{\xi} + \epsilon \mathbf{e}_{2(k-1)+l})$.

Lemma 1. *There exists a constant $L_1 > 0$ such that for each $\epsilon \in \Delta$,*

$$\|\mathbf{x}^\epsilon(s)\| \leq L_1, \quad s \in [0, \tau_n].$$

Proof. In view of (13) and Theorem 1, $\mathbf{x}^\epsilon(\cdot)$ exists for each $\epsilon \in \Delta$. Moreover,

$$\mathbf{x}^\epsilon(s) = \mathbf{x}^0 + \sum_{j=1}^i \int_{\tau_{j-1}}^{\min\{\tau_j, s\}} \mathbf{f}(\mathbf{x}^\epsilon(\eta), (\boldsymbol{\xi} + \epsilon \mathbf{e}_{2(k-1)+l})^j) d\eta, \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n.$$

Taking the Euclidean norm of both sides yields

$$\begin{aligned} \|\mathbf{x}^\epsilon(s)\| &\leq \|\mathbf{x}^0\| + \sum_{j=1}^i \int_{\tau_{j-1}}^{\min\{\tau_j, s\}} \left\{ (\xi_1^j + \epsilon \delta_{l1} \delta_{jk})^2 \cos^2(\xi_2^j + \epsilon \delta_{l2} \delta_{jk}) \right. \\ &\quad + \frac{(\xi_1^j + \epsilon \delta_{l1} \delta_{jk})^2 \sin^2(\xi_2^j + \epsilon \delta_{l2} \delta_{jk})}{\sin^2 x_1^\epsilon(\eta)} + \sin^2 x_1^\epsilon(\eta) \cos^2 x_2^\epsilon(\eta) \\ &\quad \left. + \sin^2 x_1^\epsilon(\eta) \sin^2 x_2^\epsilon(\eta) + \cos^2 x_1^\epsilon(\eta) \right\}^{\frac{1}{2}} d\eta \\ &\leq \|\mathbf{x}^0\| + nd_1 \sqrt{b_1^2 + \frac{b_1^2}{\sin^2 \rho_1}} + 3, \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \end{aligned}$$

where ρ_1 is as defined in (13). Hence, $\mathbf{x}^\epsilon(s)$ is uniformly bounded with respect to ϵ , as required. \square

Now, for each $\epsilon \in \Delta$, define a function $\mathbf{v}^\epsilon(\cdot)$ as follows:

$$\mathbf{v}^\epsilon(s) = \mathbf{x}^\epsilon(s) - \mathbf{x}(s), \quad s \in [0, \tau_n]. \quad (14)$$

The next lemma shows that the partial derivatives of \mathbf{f} are bounded.

Lemma 2. *There exists a constant $L_2 > 0$ such that for each $\epsilon \in \Delta$ and $\eta \in [0, 1]$,*

$$\begin{aligned} \left\| \frac{\partial \mathbf{f}(\mathbf{x}(s) + \eta \mathbf{v}^\epsilon(s), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^i)}{\partial \mathbf{x}} \right\| &\leq L_2, \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \\ \left\| \frac{\partial \mathbf{f}(\mathbf{x}(s) + \eta \mathbf{v}^\epsilon(s), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^i)}{\partial \xi_l^i} \right\| &\leq L_2, \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n. \end{aligned}$$

Proof. It follows from Lemma 1 that for each $\epsilon \in \Delta$,

$$\mathbf{x}^\epsilon(s) \in B_5(L_1), \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n,$$

where $B_5(L_1)$ is a closed ball in R^5 of radius L_1 centered at the origin. Since $B_5(L_1)$ is convex, we conclude that for each $\epsilon \in \Delta$ and $\eta \in [0, 1]$,

$$\mathbf{x}(s) + \eta \mathbf{v}^\epsilon(s) \in B_5(L_1), \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n. \quad (15)$$

Moreover, for each $\epsilon \in \Delta$ and $\eta \in [0, 1]$,

$$(\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^i \in [a_1, b_1] \times [a_2, b_2], \quad i = 1, 2, \dots, n. \quad (16)$$

Since $\partial \mathbf{f} / \partial \mathbf{x}$ and $\partial \mathbf{f} / \partial \xi_l^i$ are continuous, the result follows from (15) and (16) and the compactness of $B_5(L_1)$ and $[a_1, b_1] \times [a_2, b_2]$. \square

We now turn our attention to the target error sensitivity function (7). Consider the following variational system:

$$\begin{aligned} \dot{\psi}_l^k(s) = \sum_{j=1}^i \delta_{jk} \frac{\partial \mathbf{f}(\mathbf{x}(s|\boldsymbol{\tau}, \boldsymbol{\xi}), \boldsymbol{\xi}^i)}{\partial \mathbf{x}} \psi_l^k(s) + \delta_{ik} \frac{\partial \mathbf{f}(\mathbf{x}(s|\boldsymbol{\tau}, \boldsymbol{\xi}), \boldsymbol{\xi}^i)}{\partial \xi_l^i}, \\ s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \end{aligned} \quad (17)$$

with the initial and intermediate conditions:

$$\psi_l^k(\tau_i) = \begin{cases} \mathbf{0}, & \text{if } i = 0, \\ \psi_l^k(\tau_{i-}), & \text{if } i = 1, 2, \dots, n. \end{cases} \quad (18)$$

Let $\psi_l^k(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ denote the solution of the variational system (17)-(18). The following theorem shows that the partial derivative of $\mathbf{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ with respect to ξ_l^k can be expressed in terms of the solution of the variational system (17)-(18).

Theorem 2. *Let $(\boldsymbol{\tau}, \boldsymbol{\xi}) \in \Gamma \times \Xi$ and consider fixed integers $k \in \{1, 2, \dots, n\}$ and $l \in \{1, 2\}$. Then*

$$\frac{\partial \mathbf{x}(s|\boldsymbol{\tau}, \boldsymbol{\xi})}{\partial \xi_l^k} = \psi_l^k(s|\boldsymbol{\tau}, \boldsymbol{\xi}), \quad s \in [0, \tau_n].$$

Proof. To prove the theorem, we need to show that

$$\lim_{\epsilon \rightarrow 0} \frac{\mathbf{x}^\epsilon(s) - \mathbf{x}(s)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \epsilon^{-1} \mathbf{v}^\epsilon(s) = \boldsymbol{\psi}_l^k(s | \boldsymbol{\tau}, \boldsymbol{\xi}), \quad s \in [0, \tau_n]. \quad (19)$$

We consider two cases: (i) $0 \leq s \leq \tau_{k-1}$; and (ii) $s > \tau_{k-1}$. It is clear that $\mathbf{v}^\epsilon(s) = \mathbf{0}$ and $\boldsymbol{\psi}_l^k(s | \boldsymbol{\tau}, \boldsymbol{\xi}) = \mathbf{0}$ for all $s \leq \tau_{k-1}$ and thus (19) holds for case (i). We will prove case (ii) in three steps.

Step 1. The function $\mathbf{v}^\epsilon(\cdot)$ is of order ϵ

It follows from the definition of $\mathbf{v}^\epsilon(s)$ in (14) that for each $\epsilon \in \Delta$,

$$\mathbf{v}^\epsilon(s) = \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \{ \mathbf{f}(\mathbf{x}^\epsilon(\vartheta), (\boldsymbol{\xi} + \epsilon \mathbf{e}_{2(k-1)+l})^j) - \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j) \} d\vartheta, \\ s \in [\tau_{i-1}, \tau_i], \quad i = k, \dots, n.$$

By the mean value theorem, we obtain

$$\mathbf{v}^\epsilon(s) = \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \int_0^1 \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^j)}{\partial \mathbf{x}} \mathbf{v}^\epsilon(\vartheta) d\eta d\vartheta \\ + \int_{\tau_{k-1}}^{\min\{s, \tau_k\}} \int_0^1 \epsilon \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^k)}{\partial \boldsymbol{\xi}_l^k} d\eta d\vartheta, \\ s \in [\tau_{i-1}, \tau_i], \quad i = k, \dots, n. \quad (20)$$

Taking the norm of both sides of (20) and applying Lemma 2, we have

$$\|\mathbf{v}^\epsilon(s)\| \leq L_2 d_1 |\epsilon| + \int_{\tau_{k-1}}^s L_2 \|\mathbf{v}^\epsilon(\vartheta)\| d\vartheta, \quad s \in [\tau_{k-1}, \tau_n].$$

Hence, by the Gronwall-Bellman Lemma, we have

$$\|\mathbf{v}^\epsilon(s)\| \leq L_2 d_1 |\epsilon| \exp(L_2(s - \tau_{k-1})) \\ \leq L_2 d_1 |\epsilon| \exp(L_2 n d_1), \quad s \in [\tau_{k-1}, \tau_n]. \quad (21)$$

This shows that the function $\mathbf{v}^\epsilon(\cdot)$ is of order ϵ .

Step 2. Definition and limiting behavior of $\rho(\cdot)$

For each $\epsilon \in \Delta$, define functions $\boldsymbol{\gamma}^{1, \epsilon, j} : [\tau_{j-1}, \tau_j] \rightarrow R^5, j = 1, 2, \dots, n$, and $\boldsymbol{\gamma}^{2, \epsilon, k} : [\tau_{k-1}, \tau_k] \rightarrow R^5$ as follows:

$$\boldsymbol{\gamma}^{1, \epsilon, j}(\vartheta) := \int_0^1 \left\{ \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^j)}{\partial \mathbf{x}} - \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j)}{\partial \mathbf{x}} \right\} \mathbf{v}^\epsilon(\vartheta) d\eta$$

and

$$\gamma^{2,\epsilon,k}(\vartheta) := \int_0^1 \epsilon \left\{ \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^k)}{\partial \xi_l^k} - \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^k)}{\partial \xi_l^k} \right\} d\eta.$$

Furthermore, define another function $\rho : \Delta \setminus \{0\} \rightarrow R$ as follows:

$$\rho(\epsilon) := |\epsilon|^{-1} \sum_{j=1}^n \int_{\tau_{j-1}}^{\tau_j} \|\gamma^{1,\epsilon,j}(\vartheta)\| d\vartheta + |\epsilon|^{-1} \int_{\tau_{k-1}}^{\tau_k} \|\gamma^{2,\epsilon,k}(\vartheta)\| d\vartheta, \quad \epsilon \in \Delta \setminus \{0\}.$$

It follows from (21) that

$$\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta) \rightarrow \mathbf{x}(\vartheta) \text{ as } \epsilon \rightarrow 0, \quad (22)$$

uniformly with respect to $\vartheta \in [0, \tau_n]$ and $\eta \in [0, 1]$. In addition, it is obvious that

$$\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l} \rightarrow \boldsymbol{\xi} \text{ as } \epsilon \rightarrow 0, \quad (23)$$

uniformly with respect to $\eta \in [0, 1]$. Lemma 1 implies that the convergence in (22) takes place inside the ball $B_5(L_1)$, and the definition of Δ implies that the convergence in (23) takes place inside Ξ . Thus, since $\partial \mathbf{f} / \partial \mathbf{x}$ and $\partial \mathbf{f} / \partial \xi_l^k$ are uniformly continuous on the compact set $B_5(L_1) \times \Xi$, we have

$$\frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^j)}{\partial \mathbf{x}} \rightarrow \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j)}{\partial \mathbf{x}} \text{ as } \epsilon \rightarrow 0$$

and

$$\frac{\partial \mathbf{f}(\mathbf{x}(\vartheta) + \eta \mathbf{v}^\epsilon(\vartheta), (\boldsymbol{\xi} + \eta \epsilon \mathbf{e}_{2(k-1)+l})^k)}{\partial \xi_l^k} \rightarrow \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^k)}{\partial \xi_l^k} \text{ as } \epsilon \rightarrow 0$$

uniformly with respect to $\vartheta \in [0, \tau_n]$ and $\eta \in [0, 1]$. These results, together with inequality (21), imply that $\epsilon^{-1} \gamma^{1,\epsilon,j}(\vartheta) \rightarrow \mathbf{0}$ and $\epsilon^{-1} \gamma^{2,\epsilon,k}(\vartheta) \rightarrow \mathbf{0}$ uniformly with respect to $\vartheta \in [0, \tau_n]$ as $\epsilon \rightarrow 0$. Consequently,

$$\lim_{\epsilon \rightarrow 0} \rho(\epsilon) = 0. \quad (24)$$

Step 3. Comparing $\epsilon^{-1} \mathbf{v}^\epsilon(\cdot)$ with $\psi_l^k(\cdot | \tau, \boldsymbol{\xi})$

Using (20), we obtain

$$\begin{aligned} \mathbf{v}^\epsilon(s) &= \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \left\{ \gamma^{1,\epsilon,j}(\vartheta) + \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j)}{\partial \mathbf{x}} \mathbf{v}^\epsilon(\vartheta) \right\} d\vartheta \\ &\quad + \int_{\tau_{k-1}}^{\min\{s, \tau_k\}} \left\{ \gamma^{2,\epsilon,k}(\vartheta) + \epsilon \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^k)}{\partial \xi_l^k} \right\} d\vartheta, \\ &\quad s \in [\tau_{i-1}, \tau_i], i = k, \dots, n. \end{aligned} \quad (25)$$

Furthermore, since $\boldsymbol{\psi}_l^k(\tau_{k-1}) = \mathbf{0}$, integrating the auxiliary system (17)-(18) gives

$$\boldsymbol{\psi}_l^k(s) = \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j)}{\partial \mathbf{x}} \boldsymbol{\psi}_l^k(\vartheta) d\vartheta + \int_{\tau_{k-1}}^{\min\{s, \tau_k\}} \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^k)}{\partial \xi_l^k} d\vartheta, \\ s \in [\tau_{i-1}, \tau_i], \quad i = k, \dots, n. \quad (26)$$

Multiplying (25) by ϵ^{-1} and then subtracting (26) yields

$$\epsilon^{-1} \mathbf{v}^\epsilon(s) - \boldsymbol{\psi}_l^k(s) = \epsilon^{-1} \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \boldsymbol{\gamma}^{1, \epsilon, j}(\vartheta) d\vartheta + \epsilon^{-1} \int_{\tau_{k-1}}^{\min\{s, \tau_k\}} \boldsymbol{\gamma}^{2, \epsilon, k}(\vartheta) d\vartheta \\ + \sum_{j=k}^i \int_{\tau_{j-1}}^{\min\{s, \tau_j\}} \frac{\partial \mathbf{f}(\mathbf{x}(\vartheta), \boldsymbol{\xi}^j)}{\partial \mathbf{x}} (\epsilon^{-1} \mathbf{v}^\epsilon(\vartheta) - \boldsymbol{\psi}_l^k(\vartheta)) d\vartheta, \\ s \in [\tau_{i-1}, \tau_i], \quad i = k, \dots, n. \quad (27)$$

Hence, by taking the norm of both sides of (27), we have

$$\|\epsilon^{-1} \mathbf{v}^\epsilon(s) - \boldsymbol{\psi}_l^k(s)\| \leq \rho(\epsilon) + \int_{\tau_{k-1}}^s L_2 \|\epsilon^{-1} \mathbf{v}^\epsilon(\vartheta) - \boldsymbol{\psi}_l^k(\vartheta)\| d\vartheta, \quad s \in [\tau_{k-1}, \tau_n].$$

By the Gronwall-Bellman Lemma, we obtain

$$\|\epsilon^{-1} \mathbf{v}^\epsilon(s) - \boldsymbol{\psi}_l^k(s)\| \leq \rho(\epsilon) \exp(L_2 d_1 n), \quad s \in [\tau_{k-1}, \tau_n].$$

Taking the limit as $\epsilon \rightarrow 0$ in the above inequality and then applying (24), it follows that

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-1} \mathbf{v}^\epsilon(s) = \boldsymbol{\psi}_l^k(s).$$

This shows that equation (19) also holds for $s > \tau_{k-1}$. \square

Based on Theorem 2, we can now simplify the non-standard sensitivity term in Problem (B).

Theorem 3. For each $\boldsymbol{\tau} \in \Gamma$,

$$\left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right) \left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right)^\top = 4 \sum_{k=1}^n \sum_{l=1}^2 [(\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*) - \mathbf{x}^f)^\top \boldsymbol{\psi}_l^k(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*)]^2.$$

Proof. By Theorem 2,

$$\frac{\partial \mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \xi_l^k} = \boldsymbol{\psi}_l^k(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*).$$

Thus,

$$\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \xi_l^k} = 2(\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*) - \mathbf{x}^f)^\top \boldsymbol{\psi}_l^k(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*).$$

It therefore follows that

$$\begin{aligned} \left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right) \left(\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \boldsymbol{\xi}} \right)^\top &= \sum_{k=1}^n \sum_{l=1}^2 \left[\frac{\partial G(\boldsymbol{\tau}, \boldsymbol{\xi}^*)}{\partial \xi_l^k} \right]^2 \\ &= 4 \sum_{k=1}^n \sum_{l=1}^2 [(\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*) - \mathbf{x}^f)^\top \boldsymbol{\psi}_l^k(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*)]^2, \end{aligned}$$

as required. \square

By virtue of Theorem 3, the target error sensitivity function (7) can be rewritten as

$$J(\boldsymbol{\tau}) = 4 \sum_{k=1}^n \sum_{l=1}^2 [(\mathbf{x}(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*) - \mathbf{x}^f)^\top \boldsymbol{\psi}_l^k(\tau_n | \boldsymbol{\tau}, \boldsymbol{\xi}^*)]^2. \quad (28)$$

It follows that Problem (B) is equivalent to the following optimization problem.

Problem (C). *Given an optimal solution $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*)$ for Problem (A) and the dynamic systems (1)-(2) and (17)-(18) with $\boldsymbol{\xi} = \boldsymbol{\xi}^*$, choose an admissible switching point vector $\boldsymbol{\tau} \in \Gamma$ such that the target error sensitivity function (28) is minimized subject to the continuous state inequality constraints (8) and the additional inequality constraint (9).*

4. Problem transformation

4.1. Continuous state inequality constraints

Recall from Section 3 that the well inclination trajectory is given by

$$\begin{aligned} x_1(s | \boldsymbol{\tau}, \boldsymbol{\xi}) &= x_1^0 + \sum_{j=1}^{i-1} \xi_1^j \cos \xi_2^j (\tau_j - \tau_{j-1}) + \xi_1^i \cos \xi_2^i (s - \tau_{i-1}), \\ & \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n. \end{aligned} \quad (29)$$

Note that (29) is a piecewise-linear function of s with break-points at $s = \tau_i$, $i = 0, 1, \dots, n$. Thus, constraint (8a) is equivalent to

$$\omega_0 \leq x_1(\tau_i | \boldsymbol{\tau}, \boldsymbol{\xi}^*) = x_1^0 + \sum_{j=1}^i \xi_1^{j*} \cos \xi_2^{j*} (\tau_j - \tau_{j-1}) \leq \frac{1}{2} \pi, \quad i = 0, 1, \dots, n, \quad (30)$$

where ξ_1^{j*} and ξ_2^{j*} denote, respectively, the optimal curvature and optimal tool-face angle for the j th turn section in Problem (A). Constraints (30) are conventional parameter constraints that can be handled directly using standard nonlinear programming algorithms.

Now, for the second continuous state inequality constraint (8b), we first consider the dynamic equation for x_2 :

$$\dot{x}_2(s) = \xi_1^i \sin \xi_2^i \csc x_1(s | \boldsymbol{\tau}, \boldsymbol{\xi}), \quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n, \quad (31)$$

where $x_1(\cdot|\boldsymbol{\tau}, \boldsymbol{\xi})$ is given by (29). Recall from Section 2 that $\xi_1^{i*} \geq 0$ for each $i = 1, 2, \dots, n$. Thus, if (30) holds (i.e., $\omega_0 \leq x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}^*) \leq \frac{1}{2}\pi$ for all $s \in [0, \tau_n]$), then the sign of $\dot{x}_2(s)$ on $[\tau_{i-1}, \tau_i]$ for $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ is the same as the sign of $\sin \xi_2^{i*}$. This immediately implies that $x_2(\cdot)$ for $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ is monotonic on $[\tau_{i-1}, \tau_i]$, and hence constraint (8b) can be replaced by

$$0 \leq x_2(\tau_i|\boldsymbol{\tau}, \boldsymbol{\xi}^*) \leq 2\pi, \quad i = 0, 1, \dots, n. \quad (32)$$

To simplify this constraint, we integrate (31) to obtain

$$\begin{aligned} x_2(s|\boldsymbol{\tau}, \boldsymbol{\xi}) &= x_2(\tau_{i-1}|\boldsymbol{\tau}, \boldsymbol{\xi}) + \int_{\tau_{i-1}}^s \xi_1^i \sin \xi_2^i \csc(x_1(\eta|\boldsymbol{\tau}, \boldsymbol{\xi})) d\eta \\ &= x_2(\tau_{i-1}|\boldsymbol{\tau}, \boldsymbol{\xi}) \\ &\quad + \tan \xi_2^i \ln \left[\frac{\csc(x_1(\tau_{i-1}|\boldsymbol{\tau}, \boldsymbol{\xi})) + \cot(x_1(\tau_{i-1}|\boldsymbol{\tau}, \boldsymbol{\xi}))}{\csc(x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi})) + \cot(x_1(s|\boldsymbol{\tau}, \boldsymbol{\xi}))} \right], \\ &\quad s \in [\tau_{i-1}, \tau_i], \quad i = 1, 2, \dots, n. \end{aligned}$$

Hence, by induction,

$$x_2(\tau_i|\boldsymbol{\tau}, \boldsymbol{\xi}) = x_2^0 + \sum_{j=1}^i \tan \xi_2^j \ln \left[\frac{\csc(x_1(\tau_{j-1}|\boldsymbol{\tau}, \boldsymbol{\xi})) + \cot(x_1(\tau_{j-1}|\boldsymbol{\tau}, \boldsymbol{\xi}))}{\csc(x_1(\tau_j|\boldsymbol{\tau}, \boldsymbol{\xi})) + \cot(x_1(\tau_j|\boldsymbol{\tau}, \boldsymbol{\xi}))} \right],$$

$i = 0, 1, \dots, n.$

Thus, constraint (32) becomes

$$0 \leq x_2^0 + \sum_{j=1}^i \tan \xi_2^{j*} \ln \left[\frac{\csc(x_1(\tau_{j-1}|\boldsymbol{\tau}, \boldsymbol{\xi}^*)) + \cot(x_1(\tau_{j-1}|\boldsymbol{\tau}, \boldsymbol{\xi}^*))}{\csc(x_1(\tau_j|\boldsymbol{\tau}, \boldsymbol{\xi}^*)) + \cot(x_1(\tau_j|\boldsymbol{\tau}, \boldsymbol{\xi}^*))} \right] \leq 2\pi,$$

$i = 0, 1, \dots, n. \quad (33)$

Replacing the continuous state inequality constraints (8) with constraints (30) and (33) yields the following transformed problem.

Problem (D). *Given an optimal solution $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*)$ for Problem (A) and the dynamic systems (1)-(2) and (17)-(18) with $\boldsymbol{\xi} = \boldsymbol{\xi}^*$, choose an admissible switching point vector $\boldsymbol{\tau} \in \Gamma$ such that the target error sensitivity function (28) is minimized subject to the inequality constraints (9), (30) and (33).*

4.2. Time-scaling transformation

In Problem (D), the decision variables are the switching points τ_i , $i = 1, 2, \dots, n$. It is well known that variable switching points pose major difficulties for gradient-based optimization techniques [11]. Thus, we now apply a time-scaling transformation [12] to replace the variable switching points with conventional decision parameters, thereby yielding a new problem in which the switching points are fixed.

First, we introduce a new independent variable $t \in [0, n]$. Next, we relate the new variable $t \in [0, n]$ to the original arc length variable $s \in [0, \tau_n]$ through the following differential equation:

$$\dot{s}(t) = \theta_i, \quad t \in [i-1, i), \quad i = 1, 2, \dots, n, \quad (34)$$

with initial condition

$$s(0) = 0, \quad (35)$$

where $\theta_i = \tau_i - \tau_{i-1}$ is the length of the i th turn segment in the original arc length horizon. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)^\top \in R^n$. From (4), we have the following constraints on $\boldsymbol{\theta}$:

$$c_1 \leq \theta_i \leq d_1, \quad i = 1, 2, \dots, n. \quad (36)$$

Any $\boldsymbol{\theta} \in R^n$ satisfying (36) is called an admissible segment length vector. Let Θ denote the set of all admissible segment length vectors. Clearly, from the dynamic equations (34) and (35),

$$s(i) = \theta_1 + \theta_2 + \dots + \theta_i = \tau_i, \quad i = 1, 2, \dots, n.$$

Hence, the i th switching point $s = \tau_i$ gets mapped to $t = i$ via equations (34) and (35). Let $\tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(s(t))$ and $\tilde{\boldsymbol{\psi}}_l^k(t) = \boldsymbol{\psi}_l^k(s(t))$. From (1)-(2), (17)-(18) and (34)-(35), we have

$$\dot{\tilde{\boldsymbol{x}}}(t) = \theta_i \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^i), \quad t \in [i-1, i), \quad i = 1, \dots, n, \quad (37a)$$

$$\dot{\tilde{\boldsymbol{\psi}}}_l^k(t) = \sum_{j=1}^i \delta_{jk} \theta_j \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^i)}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) + \delta_{ik} \theta_i \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^i)}{\partial \xi_l^i}, \quad (37b)$$

where $k = 1, 2, \dots, n$ and $l = 1, 2$, with initial and intermediate conditions

$$\tilde{\boldsymbol{x}}(i) = \begin{cases} \boldsymbol{x}^0, & \text{if } i = 0, \\ \tilde{\boldsymbol{x}}(i-), & \text{if } i = 1, 2, \dots, n, \end{cases} \quad (38a)$$

$$\tilde{\boldsymbol{\psi}}_l^k(i) = \begin{cases} \mathbf{0}, & \text{if } i = 0, \\ \tilde{\boldsymbol{\psi}}_l^k(i-), & \text{if } i = 1, 2, \dots, n. \end{cases} \quad (38b)$$

We denote the solutions of (37)-(38) by $\tilde{\boldsymbol{x}}(\cdot | \boldsymbol{\theta}, \boldsymbol{\xi})$ and $\tilde{\boldsymbol{\psi}}_l^k(\cdot | \boldsymbol{\theta}, \boldsymbol{\xi})$, $l = 1, 2$, $k = 1, \dots, n$.

Now, in terms of $\boldsymbol{\theta}$, constraint (9) can be written as

$$\sum_{j=1}^n \theta_j \leq (1 + \beta) \tau_n^*. \quad (39)$$

Furthermore, constraints (30) and (33) become

$$\omega_0 \leq x_1^0 + \sum_{j=1}^i \theta_j \xi_1^{j*} \cos \xi_2^{j*} \leq \frac{1}{2} \pi, \quad i = 0, 1, \dots, n, \quad (40)$$

and

$$0 \leq x_2^0 + \sum_{j=1}^i \tan \xi_2^{j*} \ln \left[\frac{\csc(\tilde{x}_1(j-1|\boldsymbol{\theta}, \boldsymbol{\xi}^*)) + \cot(\tilde{x}_1(j-1|\boldsymbol{\theta}, \boldsymbol{\xi}^*))}{\csc(\tilde{x}_1(j|\boldsymbol{\theta}, \boldsymbol{\xi}^*)) + \cot(\tilde{x}_1(j|\boldsymbol{\theta}, \boldsymbol{\xi}^*))} \right] \leq 2\pi, \quad i = 0, 1, \dots, n. \quad (41)$$

Similarly, the target error sensitivity function (28) becomes

$$\tilde{J}(\boldsymbol{\theta}) = 4 \sum_{k=1}^n \sum_{l=1}^2 [(\tilde{\mathbf{x}}(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*) - \mathbf{x}^f)^\top \tilde{\boldsymbol{\psi}}_l^k(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*)]^2. \quad (42)$$

Based on the time-scaling transformation defined by equations (34) and (35), Problem (D) can be transformed into the following equivalent problem with fixed switching points.

Problem (E). *Given an optimal solution $(\boldsymbol{\tau}^*, \boldsymbol{\xi}^*)$ for Problem (A) and the dynamic system (37)-(38) with $\boldsymbol{\xi} = \boldsymbol{\xi}^*$, choose an admissible segment length vector $\boldsymbol{\theta} \in \Theta$ such that the target error sensitivity function (42) is minimized subject to the inequality constraints (39), (40) and (41).*

Problem (E) is a standard problem that can be solved using existing gradient-based optimization techniques. However, to do this, the gradients of the cost and constraints are required. These gradients are derived in the next section.

5. Gradient computation

Define

$$g_1(\boldsymbol{\theta}) = \sum_{j=1}^n \theta_j, \quad (43)$$

$$g_{i+2}(\boldsymbol{\theta}) = x_1^0 + \sum_{j=1}^i \theta_j \xi_1^{j*} \cos \xi_2^{j*}, \quad i = 0, 1, \dots, n, \quad (44)$$

$$g_{i+n+3}(\boldsymbol{\theta}) = x_2^0 + \sum_{j=1}^i \tan \xi_2^{j*} \ln \left[\frac{\csc(\tilde{x}_1(j-1|\boldsymbol{\theta}, \boldsymbol{\xi}^*)) + \cot(\tilde{x}_1(j-1|\boldsymbol{\theta}, \boldsymbol{\xi}^*))}{\csc(\tilde{x}_1(j|\boldsymbol{\theta}, \boldsymbol{\xi}^*)) + \cot(\tilde{x}_1(j|\boldsymbol{\theta}, \boldsymbol{\xi}^*))} \right], \quad i = 0, 1, \dots, n. \quad (45)$$

Then the constraints in Problem (E) can be written as follows:

$$\begin{aligned} g_1(\boldsymbol{\theta}) &\leq (1 + \beta)\tau_n^*, \\ \omega_0 &\leq g_{i+2}(\boldsymbol{\theta}) \leq \frac{1}{2}\pi, \quad i = 0, 1, \dots, n, \\ 0 &\leq g_{i+n+3}(\boldsymbol{\theta}) \leq 2\pi, \quad i = 0, 1, \dots, n. \end{aligned}$$

The gradients of constraint functions $g_i(\cdot)$, $i = 1, 2, \dots, n+2$, are easily obtained:

$$\begin{aligned} \frac{\partial g_1(\boldsymbol{\theta})}{\partial \theta_q} &= 1, \\ \frac{\partial g_{i+2}(\boldsymbol{\theta})}{\partial \theta_q} &= \begin{cases} \xi_1^{q*} \cos \xi_2^{q*}, & \text{if } q \leq i, \\ 0, & \text{otherwise,} \end{cases} \quad i = 0, 1, \dots, n. \end{aligned}$$

The gradients of $g_{i+n+3}(\cdot)$, $i = 0, 1, \dots, n$, are more challenging because these constraints depend on $\tilde{x}_1(\cdot | \boldsymbol{\theta}, \boldsymbol{\xi}^*)$. From (37a) and (38a),

$$\begin{aligned} \tilde{x}_1(t | \boldsymbol{\theta}, \boldsymbol{\xi}^*) &= x_1^0 + \sum_{j=1}^{i-1} \theta_j \xi_1^{j*} \cos \xi_2^{j*} + \theta_i \xi_1^{i*} \cos \xi_2^{i*} (t - i + 1), \\ & \quad t \in [i - 1, i], \quad i = 1, 2, \dots, n. \end{aligned}$$

Therefore, the gradients of $g_{i+n+3}(\cdot)$, $i = 0, 1, \dots, n$, are given by

$$\frac{\partial g_{i+n+3}(\boldsymbol{\theta})}{\partial \theta_q} = \sum_{j=1}^i \tan \xi_2^{j*} \left\{ \csc(\tilde{x}_1(j)) \frac{\partial \tilde{x}_1(j)}{\partial \theta_q} - \csc(\tilde{x}_1(j-1)) \frac{\partial \tilde{x}_1(j-1)}{\partial \theta_q} \right\},$$

where

$$\frac{\partial \tilde{x}_1(j)}{\partial \theta_q} = \begin{cases} \xi_1^{q*} \cos \xi_2^{q*}, & \text{if } q \leq j, \\ 0, & \text{if } q > j, \end{cases} \quad j = 0, 1, \dots, n,$$

and

$$\frac{\partial \tilde{x}_1(j-1)}{\partial \theta_q} = \begin{cases} \xi_1^{q*} \cos \xi_2^{q*}, & \text{if } q \leq j-1, \\ 0, & \text{if } q > j-1, \end{cases} \quad j = 1, 2, \dots, n.$$

Now, for the gradients of the sensitivity function (42), consider the following costate system:

$$\begin{aligned} \dot{\boldsymbol{\lambda}}(t) &= - \sum_{k=1}^n \sum_{l=1}^2 \sum_{j=1}^i \delta_{jk} \theta_i \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \mathbf{x}} \tilde{\boldsymbol{\psi}}_i^k(t) \right) \right)^\top \boldsymbol{\mu}^{k,l}(t) \\ &\quad - \sum_{k=1}^n \sum_{l=1}^2 \delta_{ik} \theta_i \left(\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \xi_l^i} \right) \right)^\top \boldsymbol{\mu}^{k,l}(t) - \theta_i \left(\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \mathbf{x}} \right)^\top \boldsymbol{\lambda}(t), \\ & \quad t \in [i - 1, i], \quad i = 1, 2, \dots, n, \end{aligned} \quad (46a)$$

$$\begin{aligned} \dot{\boldsymbol{\mu}}^{k,l}(t) &= - \sum_{j=1}^i \delta_{jk} \theta_i \left(\frac{\partial \mathbf{f}(\tilde{\mathbf{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \mathbf{x}} \right)^\top \boldsymbol{\mu}^{k,l}(t), \\ & \quad t \in [i - 1, i], \quad i = 1, 2, \dots, n, \end{aligned} \quad (46b)$$

where $k = 1, 2, \dots, n$, and $l = 1, 2$, with the terminal conditions:

$$\boldsymbol{\lambda}(n) = 8 \sum_{k=1}^n \sum_{l=1}^2 \tilde{\boldsymbol{\psi}}_l^k(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*) \tilde{\boldsymbol{\psi}}_l^k(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*)^\top (\tilde{\boldsymbol{x}}(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*) - \boldsymbol{x}^f), \quad (47a)$$

$$\boldsymbol{\mu}^{k,l}(n) = 8(\tilde{\boldsymbol{x}}(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*) - \boldsymbol{x}^f) \tilde{\boldsymbol{\psi}}_l^k(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*)^\top (\tilde{\boldsymbol{x}}(n|\boldsymbol{\theta}, \boldsymbol{\xi}^*) - \boldsymbol{x}^f). \quad (47b)$$

Let $\boldsymbol{\lambda}(\cdot|\boldsymbol{\theta})$ and $\boldsymbol{\mu}^{k,l}(\cdot|\boldsymbol{\theta})$, $k = 1, \dots, n$, $l = 1, 2$, denote the solutions of the costate system (46)-(47). The following theorem gives the gradient of the sensitivity function (42) with respect to $\boldsymbol{\theta}$.

Theorem 4. For each $q = 1, 2, \dots, n$,

$$\begin{aligned} \frac{\partial \tilde{J}(\boldsymbol{\theta})}{\partial \theta_q} &= \sum_{k=1}^q \sum_{l=1}^2 \int_{q-1}^q \boldsymbol{\mu}^{k,l}(t|\boldsymbol{\theta})^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) dt \\ &+ \sum_{l=1}^2 \int_{q-1}^q \boldsymbol{\mu}^{q,l}(t|\boldsymbol{\theta})^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*})}{\partial \xi_l^q} dt + \int_{q-1}^q \boldsymbol{\lambda}(t|\boldsymbol{\theta})^\top \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*}) dt. \end{aligned}$$

Proof. Let $\boldsymbol{\varpi} : [0, n] \rightarrow R^5$ and $\boldsymbol{u}^{k,l} : [0, n] \rightarrow R^5$, $k = 1, \dots, n$, $l = 1, 2$, be arbitrary functions that are continuous and differentiable almost everywhere. Then, the sensitivity function (42) can be expressed as follows:

$$\begin{aligned} \tilde{J}(\boldsymbol{\theta}) &= 4 \sum_{k=1}^n \sum_{l=1}^2 \left[(\tilde{\boldsymbol{x}}(n) - \boldsymbol{x}^f)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \right]^2 \\ &+ \sum_{i=1}^n \int_{i-1}^i \boldsymbol{\varpi}(t)^\top [\theta_i \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*}) - \dot{\tilde{\boldsymbol{x}}}(t)] dt \\ &+ \sum_{k=1}^n \sum_{l=1}^2 \sum_{i=1}^n \int_{i-1}^i \boldsymbol{u}^{k,l}(t)^\top \\ &\times \left[\sum_{j=1}^i \delta_{jk} \theta_j \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) + \delta_{ik} \theta_i \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \xi_l^i} - \dot{\tilde{\boldsymbol{\psi}}}_l^k(t) \right] dt, \end{aligned}$$

where we have omitted the arguments $\boldsymbol{\theta}$ and $\boldsymbol{\xi}^*$ in $\tilde{\boldsymbol{x}}(\cdot|\boldsymbol{\theta}, \boldsymbol{\xi}^*)$ and $\tilde{\boldsymbol{\psi}}_l^k(\cdot|\boldsymbol{\theta}, \boldsymbol{\xi}^*)$ for brevity. Applying integration by parts gives

$$\begin{aligned} \tilde{J}(\boldsymbol{\theta}) &= 4 \sum_{k=1}^n \sum_{l=1}^2 \left[(\tilde{\boldsymbol{x}}(n) - \boldsymbol{x}^f)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \right]^2 - \boldsymbol{\varpi}(n)^\top \tilde{\boldsymbol{x}}(n) + \boldsymbol{\varpi}(0)^\top \boldsymbol{x}^0 \\ &+ \sum_{i=1}^n \int_{i-1}^i \left[\theta_i \boldsymbol{\varpi}(t)^\top \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*}) + \dot{\boldsymbol{\varpi}}(t)^\top \tilde{\boldsymbol{x}}(t) \right] dt - \sum_{k=1}^n \sum_{l=1}^2 \boldsymbol{u}^{k,l}(n)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \\ &+ \sum_{k=1}^n \sum_{l=1}^2 \sum_{i=1}^n \int_{i-1}^i \left[\sum_{j=1}^i \delta_{jk} \theta_j \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) \right. \\ &\left. + \delta_{ik} \theta_i \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \xi_l^i} + \dot{\boldsymbol{u}}^{k,l}(t)^\top \tilde{\boldsymbol{\psi}}_l^k(t) \right] dt. \end{aligned}$$

Hence, by applying the definition of the Kronecker delta function, we obtain

$$\begin{aligned}
\tilde{J}(\boldsymbol{\theta}) &= 4 \sum_{k=1}^n \sum_{l=1}^2 \left[(\tilde{\boldsymbol{x}}(n) - \boldsymbol{x}^f)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \right]^2 - \boldsymbol{\varpi}(n)^\top \tilde{\boldsymbol{x}}(n) + \boldsymbol{\varpi}(0)^\top \boldsymbol{x}^0 \\
&+ \sum_{i=1}^n \int_{i-1}^i \left[\theta_i \boldsymbol{\varpi}(t)^\top \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*}) + \dot{\boldsymbol{\varpi}}(t)^\top \tilde{\boldsymbol{x}}(t) \right] dt - \sum_{k=1}^n \sum_{l=1}^2 \boldsymbol{u}^{k,l}(n)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \sum_{i=k}^n \int_{i-1}^i \theta_i \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) dt \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \int_{k-1}^k \theta_k \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{k*})}{\partial \xi_l^k} dt + \sum_{k=1}^n \sum_{l=1}^2 \int_0^n \dot{\boldsymbol{u}}^{k,l}(t)^\top \tilde{\boldsymbol{\psi}}_l^k(t) dt.
\end{aligned}$$

Differentiating this equation with respect to θ_q yields

$$\begin{aligned}
\frac{\partial \tilde{J}(\boldsymbol{\theta})}{\partial \theta_q} &= 8 \sum_{k=1}^n \sum_{l=1}^2 (\tilde{\boldsymbol{x}}(n) - \boldsymbol{x}^f)^\top \tilde{\boldsymbol{\psi}}_l^k(n) \left(\tilde{\boldsymbol{\psi}}_l^k(n)^\top \frac{\partial \tilde{\boldsymbol{x}}(n)}{\partial \theta_q} + (\tilde{\boldsymbol{x}}(n) - \boldsymbol{x}^f)^\top \frac{\partial \tilde{\boldsymbol{\psi}}_l^k(n)}{\partial \theta_q} \right) \\
&- \boldsymbol{\varpi}(n)^\top \frac{\partial \tilde{\boldsymbol{x}}(n)}{\partial \theta_q} - \sum_{k=1}^n \sum_{l=1}^2 \boldsymbol{u}^{k,l}(n)^\top \frac{\partial \tilde{\boldsymbol{\psi}}_l^k(n)}{\partial \theta_q} + \int_{q-1}^q \boldsymbol{\varpi}(t)^\top \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*}) dt \\
&+ \sum_{k=1}^q \sum_{l=1}^2 \int_{q-1}^q \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) dt \\
&+ \sum_{l=1}^2 \int_{q-1}^q \boldsymbol{u}^{q,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{q*})}{\partial \xi_l^q} dt \\
&+ \sum_{i=1}^n \int_{i-1}^i \left[\theta_i \boldsymbol{\varpi}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\varpi}}(t)^\top \right] \frac{\partial \tilde{\boldsymbol{x}}(t)}{\partial \theta_q} dt \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \int_0^n \dot{\boldsymbol{u}}^{k,l}(t)^\top \frac{\partial \tilde{\boldsymbol{\psi}}_l^k(t)}{\partial \theta_q} dt \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \sum_{i=k}^n \int_{i-1}^i \theta_i \boldsymbol{u}^{k,l}(t)^\top \frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} \frac{\partial \tilde{\boldsymbol{\psi}}_l^k(t)}{\partial \theta_q} dt \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \int_{k-1}^k \theta_k \boldsymbol{u}^{k,l}(t)^\top \frac{\partial}{\partial \boldsymbol{x}} \left(\frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{k*})}{\partial \xi_l^k} \right) \frac{\partial \tilde{\boldsymbol{x}}(t)}{\partial \theta_q} dt \\
&+ \sum_{k=1}^n \sum_{l=1}^2 \sum_{i=k}^n \int_{i-1}^i \theta_i \boldsymbol{u}^{k,l}(t)^\top \frac{\partial}{\partial \boldsymbol{x}} \left(\frac{\partial \boldsymbol{f}(\tilde{\boldsymbol{x}}(t), \boldsymbol{\xi}^{i*})}{\partial \boldsymbol{x}} \tilde{\boldsymbol{\psi}}_l^k(t) \right) \frac{\partial \tilde{\boldsymbol{x}}(t)}{\partial \theta_q} dt.
\end{aligned}$$

Choosing $\boldsymbol{\varpi}(\cdot) = \boldsymbol{\lambda}(\cdot | \boldsymbol{\theta})$ and $\boldsymbol{u}^{k,l}(\cdot) = \boldsymbol{\mu}^{k,l}(\cdot | \boldsymbol{\theta})$, $k = 1, \dots, n$, $l = 1, 2$, and then substituting (46)-(47) into the above equation completes the proof. \square

6. Numerical example

In this section, we consider the example problem in [3], which relates to a Ci-16-Cp146 well in the Liaohe oil field in China. The well consists of three turn segments, each of which is between 10 and 100 meters long. The tool-face angle in each turn segment should be between -50° and 50° , and the radius of curvature should be between 40 and 60 meters.

The start point for the horizontal well is defined by the vector

$$(x_1^0, x_2^0, x_3^0, x_4^0, x_5^0) = (10.4^\circ, 228.18^\circ, 102.69\text{m}, -156.39\text{m}, 1673.15\text{m}),$$

where x_1^0 is the initial inclination, x_2^0 is the initial azimuth, and (x_3^0, x_4^0, x_5^0) is the Cartesian coordinate vector of the start point. Furthermore, the target point is defined by the vector

$$(x_1^f, x_2^f, x_3^f, x_4^f, x_5^f) = (89.5^\circ, 205.5^\circ, 62.5\text{m}, -192.9\text{m}, 1718.0\text{m}),$$

where x_1^f is the final inclination, x_2^f is the final azimuth, and (x_3^f, x_4^f, x_5^f) is the Cartesian coordinate vector of the target point.

We first considered Problem (A) with $\alpha = 10$. To solve this problem, we applied the time-scaling transformation [12] to yield an equivalent problem with fixed switching instants. The equivalent problem was then solved using the software MISER 3 [5]. The optimal parameters obtained are given in Table 1, and the corresponding path length is 74.4483m with a target error of 0.0481.

We next used the solution of Problem (A) reported in Table 1 to form Problem (E) with $\beta = 0.2$. This problem can also be solved using MISER 3 [5]. The optimal segment lengths for Problem (E) are listed in Table 2, where the corresponding switching instants are denoted by ς_i^* , $i = 1, 2, 3$. The optimal well trajectory for Problem (E) has a total well length of 74.7222m and a target error of 0.0564.

To compare the solutions of Problems (A) and (E), we randomly generated 200 perturbations of the nominal parameter vector ξ^* . For each $i = 1, 2, \dots, 200$, the i th perturbed vector belongs to the set of all $\xi \in R^6$ such that

$$0.0001(i-1)\|\xi^*\| \leq \|\xi - \xi^*\| \leq 0.0001i\|\xi^*\|.$$

The coefficient of $\|\xi^*\|$ in the upper bound of this range is called the disturbance percentage of the i th perturbed vector (note that the disturbance percentage increases monotonically from 0.01% for $i = 1$ to 2% for $i = 200$). The target errors of the optimal switching points from Problems (A) and (E) under the 200 randomly-selected perturbed vectors are shown in Figure 1. Note that the red dashed curve (which corresponds to the solution from Problem (A)) predominantly lies above the blue solid curve (which corresponds to the solution from Problem (E)). This indicates that the solution of Problem (E) is more robust than the solution of Problem (A).

As another demonstration, for each $i = 1, 2, \dots, 5$, we generated 100 random points from the disturbance region defined by the inequality

$$\|\xi - \xi^*\| \leq 0.01i\|\xi^*\|.$$

Here, $0.01i$ is called the disturbance radius percentage. Both the average target error for each problem's optimal solution and the relative frequency that each problem's solution gives the best result are recorded in Table 3. Again, as with Figure 1, the results in Table 3 demonstrate that Problem (E) gives a more robust solution than Problem (A).

7. Conclusion

This paper has considered an optimal path planning problem arising in the design of horizontal oil wells. The aim of the problem is to strike a balance between minimizing target error and maximizing solution robustness with respect to unwanted parameter disturbances. The problem is formulated as a dynamic optimization problem governed by a switched system, where the switching points must be chosen to minimize target error sensitivity subject to a bound on the maximum path length. In closing, we note that the properties of the soil may exert significant influence on the direction of the well path. Thus, constructing appropriate constraints to penalize certain regions of the soil is an interesting area to pursue for future research.

Acknowledgments

This work was supported by grants from the Natural Science Foundation for the Youth of China (Grant Numbers 11201267 and 11001153), Natural Science Foundation of China (Grant Number 71471103), Natural Science Foundation of Shandong Province, China (Grant Number ZR2015AL010), and the Australian Research Council (Grant Number LP130100451).

- [1] N.U. Ahmed, *Dynamic Systems and Control with Applications*, World Scientific, Singapore, 2006.
- [2] Z.H. Gong, C.Y. Liu, E.M. Feng, Optimal control and properties of nonlinear multistage dynamical system for planning horizontal well paths, *Appl. Math. Model.* 33 (2009) 2992–3001.
- [3] Z.H. Gong, K.L. Teo, C.Y. Liu, E.M. Feng, Horizontal wells path planning: An optimal switching control approach, *Appl. Math. Model.* 39 (2015) 4022–4032.
- [4] M.W. Helmy, F. Khalf, T.A. Darwish, Well design using a computer model, *SPE Drill. Complet.* 13 (1998) 42–46.
- [5] L.S. Jennings, M.E. Fisher, K.L. Teo, C.J. Goh, *MISER 3 optimal control software: Theory and user manual (version 3)*, University of Western Australia, Australia, 2004.
- [6] S.Z. Jiang, E.M. Feng, Model and algorithm for designing 3D trajectory in sidetracking horizontal wells, *J. Dalian Univ. Technol.* 42 (2004) 261–264.

- [7] A. Li, E.M. Feng, X.L. Sun, Stochastic optimal control and algorithm of the trajectory of horizontal wells, *J. Comput. Appl. Math.* 212 (2008) 419–430.
- [8] A. Li, E.M. Feng, L. Wang, Impulsive optimal control model for the trajectory of horizontal wells, *J. Comput. Appl. Math.* 223 (2009) 893–900.
- [9] Q. Lin, R. Loxton, K.L. Teo, Y.H. Wu, A new computational method for optimizing nonlinear impulsive systems, *Dyn. Contin. Discret. Ser. B* 18 (2011) 59–76.
- [10] Q. Lin, R. Loxton, K.L. Teo, Optimal control of nonlinear switched systems: Computational methods and applications, *J. Oper. Res. Soc. China.* 1 (2013) 275–311.
- [11] Q. Lin, R. Loxton, K.L. Teo, The control parameterization method for nonlinear optimal control: A survey, *J. Ind. Manag. Optim.* 10 (2014) 275–309.
- [12] R. Loxton, Q. Lin, K.L. Teo, Switching time optimization for nonlinear switched systems: Direct optimization and the time-scaling transformation, *Pac. J. Optim.* 10 (2014) 537–560.
- [13] R. Loxton, K.L. Teo, V. Rehbock, Robust suboptimal control of nonlinear systems, *Appl. Math. Comput.* 217 (2011) 6566–6576.
- [14] K. Malanowski, Sensitivity analysis for parametric optimal control of semi-linear parabolic equations, *J. Convex Anal.* 9 (2002) 543–561.
- [15] R.C. McCann, P.V.R. Suryanarayana, Horizontal well path planning and correction using optimization techniques, *J. Energy Resour.* 123 (2001) 187–193.
- [16] V. Rehbock, K.L. Teo, L.S. Jennings, A computational procedure for suboptimal robust controls, *Dyn. Control* 2 (1992) 331–348.
- [17] K.L. Teo, C.J. Goh, K.H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Longman Scientific and Technical, Essex, 1991.
- [18] W. Wei, K.L. Teo, Z. Zhan, A numerical method for an optimal control problem with minimum sensitivity on coefficient variation, *Appl. Math. Comput.* 218 (2011) 1180–1190.
- [19] F. Yang, K.L. Teo, R. Loxton, V. Rehbock, B. Li, C.J. Yu, L. Jennings, Visual MISER: An efficient user-friendly visual program for solving optimal control problems, *J. Ind. Manag. Optim.* 12 (2016) 781–810.

Table 1: Optimal parameter values for Problem (A) in the numerical example.

Parameter	Notation	Turn Segment		
		$i = 1$	$i = 2$	$i = 3$
Curvature radius (m)	$1/\xi_1^{i*}$	40.0000	59.9880	59.9840
Tool-face angle (degrees)	ξ_2^{i*}	-6.2682	49.9963	-43.7018
Segment length (m)	$\tau_i^* - \tau_{i-1}^*$	24.5479	15.9807	33.9198

Table 2: Optimal segment lengths for Problem (E) in the numerical example.

		Turn Segment		
		$i = 1$	$i = 2$	$i = 3$
Segment length (m)	$\varsigma_i^* - \varsigma_{i-1}^*$	24.5122	16.2908	33.9192

Table 3: Target error for the solutions of Problems (A) and (E) under parameter disturbances.

Disturbance Radius	Average Target Error		Best Result Percentage	
	Problem (A)	Problem (E)	Problem (A)	Problem (E)
0%	0.0481	0.0564	—	—
1%	0.0988	0.0910	25%	75%
2%	0.2352	0.2010	14%	86%
3%	0.5323	0.4902	6%	94%
4%	0.7525	0.7010	4%	96%
5%	1.0028	0.9406	3%	97%

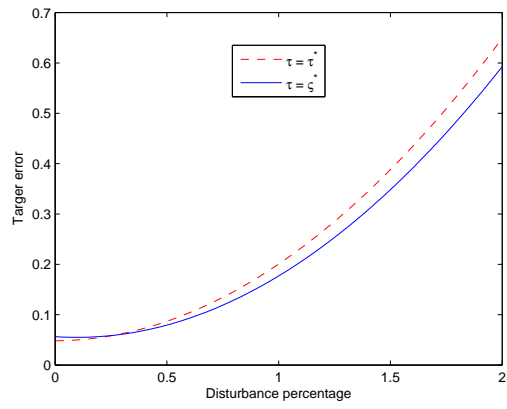


Figure 1: Target error variation for the solutions of Problems (A) and (E) under parameter disturbances.