
Xin Zhao¹,², Jian Li², Wansuo Duan³ and Dongqian Xue¹,*

¹ College of Tourism and Environment, Shaanxi Normal University, Shaanxi 720062, China
² Weihe River Basin Resources and Environment and Ecological Civilization, Institute of Computational Mathematics and Its Applications, Baoji University of Arts and Sciences, Shaanxi 721013, China
³ LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

Received 7 October 2014; Accepted (in revised version) 13 November 2015

Abstract. In this paper, we propose and analyzes the mixed 4th-order Runge-Kutta scheme of conditional nonlinear perturbation (CNOP) approach for the El Niño-Southern Oscillation (ENSO) model. This method consists of solving the ENSO model by using a mixed 4th-order Runge-Kutta method. Convergence, the local and global truncation error of this mixed 4th-order Runge-Kutta method are proved. Furthermore, optimal control problem is developed and the gradient of the cost function is determined.

AMS subject classifications: 65M06, 65M12

Key words: The El Niño-Southern Oscillation (ENSO) model, 4th-order Runge-Kutta scheme, optimal control problem, conditional nonlinear perturbation.

1 Introduction

Optimal control theory is a mature mathematical discipline with numerous applications in both science and engineering. With the development of society and progress of science, the development of efficient numerical methods for the optimal control theory is

*Corresponding author.
Email: sungirl_zhx@163.com (X. Zhao), jiaaanli@gmail.com (J. Li), duanws@lasg.iap.ac.cn (W. S. Duan), xuedq@snnu.edu.cn (D. Q. Xue)
a fundamental component in the applied mathematics. Moreover, it plays an important role in many current scientific, engineering, and industrial applications.

In this paper, we mainly consider theoretical model for El Niño-Southern Oscillation (ENSO) including prognostic equations for sea surface temperature and for thermocline variation [1]. This model can reproduce essential aspects of ENSO evolution: cyclic, chaotic, and phase-locking with annual cycles. Moreover, two dimensionless equation is involved: one reveals the nonlinear evolution of the anomalous sea surface temperature in the equatorial western/estern Pacific, and the other show the variation of the anomalous thermocline depth. Here, a classical 4th-order Runge-Kutta method is applied and developed for the mixed coupled ocean-atmosphere model. Two 4th-order Runge Kutta scheme are respectively designed for two equations, which involve the nonlinear evolution, coupling with two physical objects. The convergence and global truncation model are analyzed and we here obtain the same results as the simple ODE for the ENSO model.

Many physical phenomena on the ocean-atmosphere can be viewed as perturbations on the basic flow for the scientific research. In mathematics, this can be viewed as development of initial perturbations’ evolution. Especially, it attracts more experts on the stability, sensitivity and predictability studies in geophysical fluid dynamics. In the past several decades, linear singular vector (LSV) [2], the fastest growing perturbation of the linearized model, is one of the dominant tools with the assumption that the initial perturbation is sufficiently small. Thus, its evolution can be governed approximately by the tangent linear model (TLM) of a nonlinear model. However, LSV neglects important issues on the nonlinearity and complexity of the physical phenomena. Then, conditional nonlinear optimal perturbation (CNOP) [3,4] is introduced with the initial perturbation of nonlinear evolution. Recently, CNOP has be extend to a comprehensive approach for the optimal combined mode of initial perturbations and model parameter perturbations [5]. This method has been succesfully applied in weather forecast and climatic prediction including El Niño-Southern Oscillation (EISO) [6,7].

Among these literatures related to CNOP, the conventional adjoint method is used to provide the gradient of cost function for the optimization process. Some iterative methods have been used and achieved the better performance. Here, we mainly choose the Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) for the presented problem, which only uses a limited memory variation of the BFGS update to approximate the inverse Hessian matrix and thus has an advantage over other corresponding methods. Moreover, CMA-ES method is a derivative free optimizer, the derivation argument is omitted and there are some other named parameters to control the behaviour of the algorithm. We will discuss this algorithm for the discontinuous cost function in the further study. However, to the best of our knowledge, [3–7] are the papers where numerical experiments have been provided to be efficient, and no rigorous error analysis of these methods has been done yet. In this paper, we mainly analyze and develop the mixed 4th-order Runger-Kutta scheme of CNOP approach for the ENSO model. The convergence and global truncation error are obtained. Finally, optimal control problem is developed and the gradient of the cost
function is determined.

The remainder of the paper is organized as follows. In the next section, an abstract setting of mathematical preliminaries for the El Niño-Southern Oscillation (ENSO) are recalled. In Section 3, a mixed 4th-order Runge-Kutta method for the ENSO model is studied and its convergence analysis and convergence order are analyzed. In Section 4, optimal control is developed and designed for the ENSO model.

2 Mathematical preliminaries

In this paper, a theoretical model for the El Niño-Southern Oscillation (ENSO) was used in [1]. This model is instrumental for understanding the nature of the couple model, the mechanisms of the irregular oscillation and the season dependence of ENSO evolution. The model is considered as follows:

\[
\begin{align*}
\frac{dT}{dt} &= a_1 T - a_3 h + \sqrt{\frac{3}{2}} T(T - a_3 h), \\
\frac{dh}{dt} &= b(2h - T), \\
T(0) &= T_0 \quad \text{in } [0, \tau], \\
h(0) &= h_0,
\end{align*}
\]  

(2.1a) (2.1b) (2.1c) (2.1d)

where the first ordinary differential equation describes the evolutions of the anomalous sea surface temperature \( T(t), t \in [0, \tau], \) and the second one represent the thermocline variation \( h(t), t \in [0, \tau], \) in the western/eastern Pacific. Moreover, the parameters can be expressed by

\[
\begin{align*}
a_1 &= \bar{T}_z + \bar{T}_x - \alpha_s, \\
a_2 &= (\mu + \delta_1) \bar{T}_x, \\
a_3 &= \mu + \delta_1, \\
b &= \frac{2 \alpha s}{p(1 - 3 \alpha^2)}, 
\end{align*}
\]

where the basic state parameters \( \bar{T}_x \) and \( \bar{T}_z \) respectively denote the difference between the eastern and western basins and between the surface and subsurface water; \( \alpha_s \) is Newtonian cooling coefficient for sea surface temperature anomaly; The parameter \( \delta_1 \) represents the contribution of the horizontal temperature advection by anomalous zonal currents to the variation of local sea surface temperature. The parameter

\[
p = \left(1 - \frac{H_1}{H}\right) \left(\frac{L_0}{L_s}\right)^2,
\]

where \( H \) and \( H_1 \) represent the mean depth of the thermocline and the depth of mixed layer, \( L_0 \) and \( L_s \) denote the oceanic Rossby radius of deformation and the Ekman spreading length scale. In this paper, the coefficients \( a_1 \) and \( a_2 \) vary with the state parameters \( \bar{T}_x \) and \( \bar{T}_z \) in each month. In other words, they are fixed on each month. For simplicity, we
assume that $a_1$ and $a_2$ are independent of time $t$ in theory. The other two parameters $a_3$ and $b$ are fixed for the given $\alpha$ and $\mu$.

For simplicity of presentation, it is assumed that the state vector function $U(t) = (T(t), h(t))$ is uniquely determined by the ODE system (2.1) for the given initial condition $U_0 = (T_0, h_0)$ and model parameter $P = (a_1, a_2)$. For convenience, we set

$$f(U) \equiv f(T, h) = a_1 T - a_2 h + \sqrt{\frac{3}{2}} T(T - a_3 h),$$

$$g(U) \equiv g(T, h) = b(2h - T).$$

Then, the state equation (2.1) can be rewritten by the following for $t \in [0, \tau]$

$$\frac{dT}{dt} = f(U),$$

$$\frac{dh}{dt} = g(U),$$

$$U(0) = U_0. \quad (2.2c)$$

Here, the initial vector-valued $U_0 \in \mathbb{R}^2$ and the functions $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$. The unknown is $U(t) : [0, \tau] \to \mathbb{R}^2$, which we interpret as the dynamical evolution of the state of some system. The picture illustrates the resulting evolution. The point is that the system may behave quite differently as we change the control parameters or initial value.

Moreover, we shall always assume that such a unique solution exists for this system presented in (2.1) or (2.2).

**Lemma 2.1.** If $f(U)$ and $g(U)$ are continuous and satisfy a Lipschitz condition

$$|f(U) - f(\tilde{U})| \leq L|U - \tilde{U}|, \quad (2.3a)$$

$$|g(U) - g(\tilde{U})| \leq L|U - \tilde{U}|, \quad (2.3b)$$

in $T$ in $[0, \tau]$, Then, there exists a unique, continuous, differentiable function $T(t)$ and $h(t)$. Here, $L$ is the Lipschitz constant which must exist for the condition to be satisfied.

We also assume that the system (2.1) satisfies some standard conditions needed for the existence of an optimal control below.

Based on theoretic results [5] with respect to CNOP, the solution to system (2.1) at time $\tau$ is given by

$$U(\tau) = M_\tau(P)(U_0), \quad (2.4)$$

where $M_\tau(P)$ propagates the initial state value $U_0$ to $U$ at time $\tau$. For a initial perturbation $u_0$ of time dependent state $U(t)$, there holds

$$U(\tau) + u(u_0, \tau) = M_\tau(P)(U_0 + u_0), \quad (2.5)$$
where $u(u_0, \tau)$ denotes the nonlinear evolution of this initial perturbation $u_0$. On the other hand, for a given parameter perturbation $p$, then it is valid that

$$U(\tau) + u(p, \tau) = M_\tau(P + p)(U_0),$$

where $u(p, \tau)$ describes the departure from the reference state $U(\tau)$ caused by $p$. If there exist both parameter perturbation and initial perturbation in this system (2.1), we can similarly obtain the following relationship

$$U(\tau) + u(u_0, p, \tau) = M_\tau(P + p)(U_0 + u_0),$$

where $u(u_0, p, \tau)$ is the nonlinear evolution of the initial perturbation and parameter perturbation. Then, the nonlinear optimization problems can be defined as follows

$$J(u_0, p) = \begin{cases} \max_{u_0 \in C_\delta} \|M_\tau(P)(U_0 + u_0) - M_\tau(P)(U_0)\|_0, & \text{CNOP-I}, \\ \max_{p \in C_p} \|M_\tau(P + p)(U_0 - M_\tau(P)(U_0)\|_0, & \text{CNOP-P}, \\ \max_{u_0 \in C_\delta, p \in C_p} \|M_\tau(P + p)(U_0 + u_0) - M_\tau(P)(U_0)\|_0, & \text{CNOP}, \end{cases}$$

for the equalities (2.5)-(2.7) in $L^2$-norm measurement $\| \cdot \|$, where the constraint condition is simply expressed in sets $C_\delta$ and $C_p$ with the specific measurement. We mark the first case, second case and third case, the CNOP-I, CNOP-P and CNOP optimal problems. In fact, it can be equivalent to the following cost function

$$\min J_1(u_0, p)$$

subject to (2.1) with the relation $J_1(u_0, p) = -J(u_0, p)$. By this conversion, the constrained optimization problem (2.8) becomes a corresponding minimization problem.

### 3 Numerical method for the ENSO model

For the ODE, we usually solve it by using the difference method. Through this section, we will define some denotations for convenience. Set

$$t_m = m\Delta t, \quad N\Delta t = \tau, \quad m = 0, 1, 2, \ldots, N,$$

$$U(t_m) = (T(t_m), h(t_m)),$$

$$U_m(T_m, h_m).$$

In particular, $T(t_m)$ and $h(t_m)$ are the analytic solutions of two functions $T(t)$ and $h(t)$ at time $t_m$. Moreover, $T_m$ and $h_m$ are the corresponding numerical solution of finite difference method at time $t_m$. There are many ways to evaluate the right-hand side of the ODE as $\frac{dy}{dt} = y(t, y(t))$, that all agree to first order, but that have different coefficients of higher-order error terms. Adding up the right combination of these, we can eliminate
the error terms order by order. By far the most popular scheme is the classical 4th-order Runge-Kutta formula. Here, we follow the idea to the ENSO model as follows [8]:

\[ U_{n+1} = U_n + \Pi(\Delta t, U_n) \Delta t, \] (3.1)

namely,

\[ T_{n+1} = T_n + \phi(\Delta t, U_n) \Delta t, \]
\[ h_{n+1} = h_n + \psi(\Delta t, U_n) \Delta t, \]

where

\[ \Pi(\Delta t, U_n) \equiv (\phi(\Delta t, U_n),\psi(\Delta t, U_n)) = (\phi(T_n, h_n),\psi(T_n, h_n)) \]

can be detailedly defined as follows

\[ \phi(T_n, h_n) = \sum_{j=1}^{r} c_{1j} K_{1j}, \quad \psi(T_n, h_n) = \sum_{j=1}^{r} c_{2j} K_{2j}. \]

Here,

\[ K_{1j} = f(T_n, h_n), \quad K_{21} = g(T_n, h_n), \]
\[ K_{1i} = f\left(T_n + \Delta t \sum_{j=1}^{i-1} \alpha_j K_{1j}, h_n + \Delta t \sum_{j=1}^{i-1} \beta_j K_{2j}\right), \]
\[ K_{2i} = g\left(T_n + \Delta t \sum_{j=1}^{i-1} \lambda_j K_{1j}, h_n + \Delta t \sum_{j=1}^{i-1} \chi_j K_{2j}\right), \quad i = 2, 3, \ldots, r, \]

with some constants \( \alpha_j, \beta_j, \lambda_j \) and \( \chi_j \). Now, we are in position to prove the consistency and convergence of the mixed Runge-Kutta method for the ENSO model.

**Proposition 3.1** (Consistency). The mixed Runge-Kutta method (3.1) is consistent with the initial value problem (2.1) if it holds

\[ f(U) = \phi(0, U), \quad g(U) = \psi(0, U), \] (3.2)

with \( \sum_{j=1}^{r} c_{ij} = 1, i = 1, 2. \)

**Remark 3.1.** In fact, that consistency presented in Proposition 3.1 is necessary and sufficient for convergence of Runge-Kutta methods.

**Theorem 3.1** (Convergence). Under the assumption of Proposition 3.1, let \( f(U(t)) \) and \( g(U(t)) \) satisfy the Lipschitz condition defined as above. Then, the local truncation error

\[ \lim_{\Delta t \to 0} e_{U_{n+1}} = 0. \] (3.3)
Using the Lipschitz condition for the function $f$

Especially, it follows from Proposition 3.1 that

$$\lim_{\Delta t \to 0} (e_{T_{n+1}} e_{h_{n+1}}) = 0,$$

where

$$e_{U_{n+1}} = |U_{n+1} - U(t_{n+1})|, \quad e_{T_{n+1}} = |T(t_{n+1}) - T_{n+1}|, \quad e_{h_{n+1}} = |h(t_{n+1}) - h_{n+1}|.$$ (3.4)

**Proof.** We only prove the first component of (3.1). Then, we can use the same approach to verify the second one. Firstly, applying the mean value theorem, we have

$$T(t_{n+1}) - T(t_n) = \Delta t \frac{d}{dt} T(t_n + \theta \Delta t) = f(U(t_n + \theta \Delta t)) \Delta t, \quad \theta \in (0, 1).$$ (3.5)

Using the definition of $e_{T_{n+1}}$ and $\phi(\Delta t, U_n)$, and (3.5), we obtain that

$$e_{T_{n+1}} = |T(t_{n+1}) - T_{n+1}|
= |T(t_n) + f(U(t_n + \theta \Delta t)) \Delta t - T_n - \phi(\Delta t, U_n) \Delta t|
\leq e_{T_n} + \Delta t |f(U(t_n + \theta \Delta t)) - \phi(\Delta t, U_n)|.$$ (3.6)

Especially, it follows from Proposition 3.1 that

$$|f(U(t_n + \theta \Delta t)) - \phi(\Delta t, U_n)|
= |f(U(t_n + \theta \Delta t)) - \phi(0, U(t_n)) + \phi(0, U(t_n)) - \phi(\Delta t, U(t_n)) + \phi(\Delta t, U(t_n)) - \phi(\Delta t, U_n)|
\leq |f(U(t_n + \theta \Delta t)) - f(U(t_n))| + |\phi(0, U(t_n)) - \phi(\Delta t, U(t_n))| + |\phi(\Delta t, U(t_n)) - \phi(\Delta t, U_n)|
= I_1 + I_2 + I_3.$$

Using the Lipschitz condition for the function $f$ and the continuity of $U$ leads to

$$\lim_{\Delta t \to 0} I_1 = \lim_{\Delta t \to 0} |f(U(t_n + \theta \Delta t)) - f(U(t_n))|
\leq L \lim_{\Delta t \to 0} |U(t_n + \theta \Delta t) - U(t_n)|
= 0.$$

Using the Lipschitz condition for two functions $f$ and $g$ again gets the following

$$\lim_{\Delta t \to 0} I_2 = \lim_{\Delta t \to 0} |\phi(0, U(t_n)) - \phi(\Delta t, U(t_n))| = 0.$$

Since $f$ and $g$ are continuous and satisfies a Lipschitz condition, then $\phi(\Delta t, U(t_n))$ has the same property. Moreover, using the definition of $U$ leads to

$$I_3 = |\phi(\Delta t, U(t_n)) - \phi(\Delta t, U_n)|
\leq L |U(t_n) - U_n|
\leq L (e_{T_n} + e_{h_n}).$$
Thus,

\[ e_{T_{n+1}} \leq (1 + L \Delta t) e_{T_n} + L e_{h_n}. \]  

(3.7)

Similarly, we can get

\[ e_{h_{n+1}} \leq (1 + L \Delta t) e_{h_n} + L e_{T_n}. \]  

(3.8)

Adding (3.7) and (3.8) yields by recursion, we have

\[ e_{T_{n+1}} + e_{h_{n+1}} \leq (1 + L (\Delta t + 1)) (e_{T_n} + e_{h_n}) \leq (1 + L (\Delta t + 1))^{n+1} (e_{T_0} + e_{h_0}), \]  

(3.9)

which together with

\[ e_{T_0} = |T(t_0) - T_0| = 0, \quad e_{h_0} = |h(t_0) - h_0| = 0, \]

achieve the desired results.

Using Taylor expansion, we can similarly obtain the new mixed 4th-order Runge-Kutta formula for the ENSO model:

\[ T_{n+1} = T_n + \frac{\Delta t}{6} (K_{11} + 2K_{12} + 2K_{13} + K_{14}) + \mathcal{O}(h^r), \]  

(3.10a)

\[ h_{n+1} = h_n + \frac{\Delta t}{6} (K_{21} + 2K_{22} + 2K_{23} + K_{24}) + \mathcal{O}(h^r), \]  

(3.10b)

where

\[ K_{11} = f(T_n, h_n), \quad K_{21} = g(T_n, h_n), \]

\[ K_{12} = f\left(T_n + \frac{\Delta t}{2} K_{11}, h_n + \frac{\Delta t}{2} K_{21}\right), \]

\[ K_{22} = g\left(T_n + \frac{\Delta t}{2} K_{11}, h_n + \frac{\Delta t}{2} K_{21}\right), \]

\[ K_{13} = f\left(T_n + \frac{\Delta t}{2} K_{12}, h_n + \frac{\Delta t}{2} K_{22}\right), \]

\[ K_{23} = g\left(T_n + \frac{\Delta t}{2} K_{12}, h_n + \frac{\Delta t}{2} K_{22}\right), \]

\[ K_{14} = f\left(T_n + \Delta t K_{13}, h_n + \Delta t K_{23}\right), \]

\[ K_{24} = g\left(T_n + \Delta t K_{13}, h_n + \Delta t K_{23}\right), \]

and \( r \) is a positive integer. Note that the first index \( i \) of \( K_{ij} \) denotes the \( i \)-th equation in (2.1).

In the above, we study the local truncation error in Theorem 3.1. Now, we will analyze the global truncation error in the following.
Proposition 3.2. A method is said to have order $r$ if $r$ is the largest integer for (3.10).

Theorem 3.2 (Global truncation error). Under the assumption of Propositions 3.1, 3.2, and Theorem 3.1, it holds that

$$|U(t_{n+1}) - U_n| = O(\Delta t^r),$$

where $U_{n+1}$ is a sequence of approximations of (2.1) by using the mixed 4th-order Runge-Kutta method (3.10).

Proof. Using a triangle inequality, the relationship $U_{n+1} = U_n - \Pi(\Delta t, U_n) \Delta t$, and the Lipschitz condition for $\Pi(\Delta t, \cdot)$, yields

$$E_{U_{n+1}} = |U(t_{n+1}) - U_n - \Pi(\Delta t, U_n)|$$

$$\leq |U(t_{n+1}) - U_n - \Pi(\Delta t, U_n)| + \Delta t |\Pi(\Delta t, U_n) - \Pi(\Delta t, U_n)| + |U_n - U_n|$$

$$\leq e_{U_{n+1}} + (1 + L \Delta t) E_{U_n}. \quad (3.11)$$

By recursion and sum formula of geometric progression,

$$E_{U_{n+1}} \leq e_{U_{n+1}} + (1 + L \Delta t) e_{U_n} + \cdots + (1 + L \Delta t)^n e_{U_1}$$

$$\leq \frac{(1 + L \Delta t)^n - 1}{L \Delta t} e_{U_{n+1}}$$

$$\leq \frac{(1 + L \Delta t)^n - 1}{L \Delta t}, \quad (3.12)$$

which together with $e_{U_{n+1}} = O(\Delta t^r)$, $E_{U_0} = (0,0)$ implies that the global truncation error is $\Delta t^4$.

4 Optimal control

To solve the optimal control problem (2.9), the corresponding minimization problem based on the Lagrange function $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t)) : [0, \tau] \to \mathbb{R}^3$ can be considered as follows

$$\min \{ J(u_0, p, \lambda) \}, \quad (4.1)$$

where

$$J(u_0, p, \lambda) = -J_1(u_0, p) + \left\langle \frac{dT}{dt} - f(U), \lambda_1(t) \right\rangle + \left\langle \frac{dh}{dt} - g(U), \lambda_2(t) \right\rangle + \left\langle \frac{dp}{dt}, \lambda_3(t) \right\rangle.$$

Here, $\lambda(t)$ denotes the associated co-state (“adjoint” or “dual” variable). $\langle \cdot, \cdot \rangle$ denotes the dual pairing $L^2(0, \tau)$, and is defined by

$$\langle u(t), v(t) \rangle = \int_0^\tau u(t)v(t)dt,$$
where given functions \( u(t), v(t) : [0, \tau] \to \mathbb{R} \). For the sake of readability, both forms of denoting the scalar or vector products will be subsequently used, depending on the situation. \( \frac{dT}{dt}, \frac{dh}{dt} \) and \( \frac{dp}{dt} - g(U) \) are continuously Fréchet-differentiable on variations \( u_0, p \) and \( \lambda(t) \). Furthermore, we refer [9, 10] for the following result for the Lagrangian multiplier.

The triplet \((u_0, p, \lambda(t))\) is determined by the saddle-point problem (so-called “optimality system” or “Karush-Kuhn-Tucker (KKT) system”).

**Lemma 4.1.** For any solution \((u_0, p)\) of the optimal control problem (4.1), there exists a unique function \( \lambda(t) \) such that \((u_0, p)\) satisfies the first order necessary optimality conditions of (4.1), i.e., the first order variation is equal to 0 with respect to the variables \( u_0, p \) and \( \lambda(t) \).

Besides, the tangent linear model are given for the ENSO model (2.1) as follows

\[
\begin{align*}
\frac{d\delta T}{dt} &- a_1 \delta T + a_2 \delta h - \sqrt{\frac{3}{2}} \left[ \delta T^2 - a_3 \delta (Th) \right] = 0, \quad (4.2a) \\
\frac{d\delta h}{dt} &- b(2\delta h - \delta T) = 0, \quad (4.2b) \\
\frac{d\delta p}{dt} & = 0, \quad (4.2c) \\
\delta T(0) & = \delta T_0, \quad (4.2d) \\
\delta h(0) & = \delta h_0, \quad (4.2e)
\end{align*}
\]

with \( \delta p \) is independent of time \( t \).

As we known, it is usually required to compute the gradient of the cost function \( J_1(u_0, p) \) for solving the consider state constrained control system (2.9). In the following, we analyze the gradient of the objective function \( J_1(u_0, p) \), which is determined by the adjoint equation of the corresponding state equation (2.4).

**Theorem 4.1.** Under the assumption of Lemma 4.1, the gradient of \( J(u_0, p, \lambda) \) can be determined by the following equation:

\[
\begin{align*}
\frac{d\lambda_1(t)}{dt} + \left[ a_1 + 2 \sqrt{\frac{2}{3}} T - ah \right] \lambda_1(t) &- 2b\lambda_2(t) = 0, \quad (4.3a) \\
\frac{d\lambda_2(t)}{dt} - \left[ a_2 + a_3 T \right] \lambda_1(t) &+ 2b\lambda_2(t) = 0, \quad (4.3b) \\
\frac{d\lambda_3(t)}{dt} & = \lambda_3(t), \quad (4.3c) \\
\lambda_1(\tau) & = 0, \quad (4.3d) \\
\lambda_2(\tau) & = 0. \quad (4.3e)
\end{align*}
\]

**Proof.** Based on Lemma 4.1, we are now in position to analyze the 1th-order variational of \( J \) as follows:

\[
\delta J(u_0, p, \lambda) = 0, \quad (4.4)
\]
Then, (2.4) and (2.7) gives
\[
\delta J(u_0, p, \lambda) = - \delta J_1(u_0, p) + \left\langle \frac{d\delta T}{dt} - a_1 \delta T + a_2 \delta h - \sqrt{\frac{3}{2}} [\delta T^2 - a_3 \delta (Th)], \lambda_1(t) \right\rangle \\
+ \left\langle \frac{d\delta h}{dt} - b(2\delta h - \delta T), \lambda_2(t) \right\rangle + \left\langle \delta \frac{dp}{dt}, \lambda_3(t) \right\rangle.
\]
(4.5)

Then, (2.4) and (2.7) gives
\[
\delta J_1(u_0, p) = - \frac{1}{2} \delta [f(u_0, p)]^2 - \left\langle u(u_0, p, \tau), \delta u(u_0, p, \tau) \right\rangle \\
= \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial T_0}, \delta T_0 \right\rangle + \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial h_0}, \delta h_0 \right\rangle + \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial p}, \delta p \right\rangle.
\]
(4.6)

For the second, third and fourth terms in (4.5), it follows from an integration by parts that
\[
\left\langle \frac{d\delta T}{dt}, \lambda_1(t) \right\rangle = \frac{d}{dt} \left\langle \lambda_1(t), \delta T \right\rangle - \left\langle \frac{d\lambda_1(t)}{dt}, \delta T \right\rangle \\
= \left\langle \lambda_1(\tau), \delta T(\tau) \right\rangle - \left\langle \lambda_1(0), \delta T_0 \right\rangle - \left\langle \frac{d\lambda_1(t)}{dt}, \delta T \right\rangle,
\]
\[
\left\langle \frac{d\delta h}{dt}, \lambda_2(t) \right\rangle = \frac{d}{dt} \left\langle \lambda_2(t), \delta h \right\rangle - \left\langle \frac{d\lambda_2(t)}{dt}, \delta h \right\rangle \\
= \left\langle \lambda_2(\tau), \delta h(\tau) \right\rangle - \left\langle \lambda_2(0), \delta h_0 \right\rangle - \left\langle \frac{d\lambda_2(t)}{dt}, \delta h \right\rangle,
\]
\[
\left\langle \frac{d\delta p}{dt}, \lambda_3(t) \right\rangle = \left\langle \lambda_3(\tau), \delta p \right\rangle - \left\langle \lambda_3(0), \delta p \right\rangle - \left\langle \frac{d\lambda_3(t)}{dt}, \delta p \right\rangle.
\]

Combining all these equalities, leads to
\[
\delta J(u_0, p, \lambda) = - \left\langle \frac{d\lambda_1(t)}{dt}, \lambda_1(t) \right\rangle + \lambda_1(t) \left[ a_1 + 2 \sqrt{\frac{3}{2}} T - a_3 h \right] - b \lambda_2(t), \delta T \right\rangle \\
- \left\langle \frac{d\lambda_2(t)}{dt}, \lambda_1(t) \right\rangle \left[ a_2 + a_3 T \right] + 2 b \lambda_2(t), \delta h \right\rangle - \left\langle \frac{d\lambda_3(t)}{dt}, \delta p \right\rangle \\
- \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial T_0}, \delta T_0 \right\rangle - \left\langle \lambda_1(0), \delta T_0 \right\rangle \\
- \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial h_0}, \delta h_0 \right\rangle - \left\langle \lambda_2(0), \delta h_0 \right\rangle \\
- \left\langle \frac{\partial J_1(u_0, p, \tau)}{\partial p}, \delta p \right\rangle - \left\langle \lambda_3(0), \delta p \right\rangle \\
+ \left\langle \lambda_1(\tau), \delta T(\tau) \right\rangle + \left\langle \lambda_2(\tau), \delta h(\tau) \right\rangle + \left\langle \lambda_3(\tau), \delta p \right\rangle.
\]
(4.7)

Then,
\[
\frac{\partial J_1(u_0, p, \tau)}{\partial T} = -\lambda_1(0), \quad \frac{\partial J_1(u_0, p, \tau)}{\partial h} = -\lambda_2(0), \quad \frac{\partial J_1(u_0, p, \tau)}{\partial p} = -\lambda_3(0),
\]
(4.8)
can be solved by the following adjoint equation of the Eq. (4.3).
**Remark 4.1.** The adjoint state \( \lambda(t) \) is determined by the Eq. (4.3), which is called the adjoint equation of the control problem (4.1). Set

\[
A_{11} = a_1 + 2\sqrt{\frac{2}{3}} T - a_3 h \quad \text{and} \quad A_{21} = a_2 + a_3 T,
\]

then the adjoint state can be clear presented.

The optimization problem is discretized by the mixed 4th-order Runge-Kutta method. Its discretization of adjoint equation can be written as follows: For the discrete solutions \( (\lambda_{n+1,3}, \lambda_{n+1,2}, \lambda_{n+1,1}) \) such that

\[
\begin{align*}
\frac{d\lambda_{n+1,1}(t)}{dt} + A_{11}\lambda_{n+1,1}(t) - b\lambda_{n+1,2}(t) &= 0, \\ \frac{d\lambda_{n+1,2}(t)}{dt} - A_{21}\lambda_{n+1,1}(t) + 2b\lambda_{n+1,2}(t) &= 0, \\ \frac{d\lambda_3(t)}{dt} &= \lambda_3(\tau), \\ \lambda_{n+1,1}(\tau) &= 0, \\ \lambda_{n+1,2}(\tau) &= 0.
\end{align*}
\]

Then, the discretization of gradient of \( J_1 \) can be determined by the vector

\[
(-\lambda_{n+1,1}(0), -\lambda_{n+1,2}(0), -\lambda_{n+1,3}(0)).
\]

**Remark 4.2.** Usually, the constrained optimization problem is solved by the gradient-based optimization algorithms. The gradient of cost function is required, in other words, the cost function is at least continuous, and first derivative with respect to the corresponding variables. However, the gradient-based optimization algorithm is still convergent faster than others. If the cost function is sufficiently smooth, the gradient-based algorithm is the best choice.

**Acknowledgements**

The authors are supported in part by NSF of China (No. 11371031), Technology Infrastructure Work (No. 2014FY210100), Baoji Science and Technology Plan Projects (No. 14SFGG-2-7), and the Key Project of Baoji University of Arts and Sciences (No. ZK15033).

**References**


