

An optimizing finite difference scheme based on proper orthogonal decomposition for CVD equations

Juan Du¹, Jiang Zhu², Zhendong Luo^{3,*},[†] and I. M. Navon⁴

¹*School of Science, Beijing Jiaotong University, Beijing 100044, People's Republic of China*

²*Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, People's Republic of China*

³*School of Mathematics and Physics, North China Electric Power University, Beijing 102206, People's Republic of China*

⁴*Department of Scientific Computing, Florida State University, Dirac Sci. Lib. Bldg., #483, Tallahassee, FL 32306-4120, U.S.A.*

SUMMARY

In this article, an optimizing reduced finite difference scheme (FDS) based on singular value decomposition (SVD) and proper orthogonal decomposition (POD) for the chemical vapor deposit (CVD) equations is presented. And the error estimates between the usual finite difference solution and the reduced POD solution of optimizing FDS are derived. At last, some examples of numerical simulation are given to demonstrate the consistency of the numerical and theoretical results. It is shown that the optimizing reduced FDS based on POD method is of great feasibility and efficiency. Copyright © 2009 John Wiley & Sons, Ltd.

Received 13 January 2009; Revised 22 April 2009; Accepted 26 May 2009

KEY WORDS: chemical vapor deposit equations; finite difference scheme; singular value decomposition; proper orthogonal decomposition; error estimate; numerical simulation

1. INTRODUCTION

The chemical vapor deposit (CVD) reaction processes, which are of extensive application, can be classified as a mathematical model by the following governing nonlinear partial differential equations containing velocity vector, temperature field, pressure field and gas mass field.

Problem (I) Find \mathbf{u} , p , T , C , and σ such that,

$$\nabla \cdot \mathbf{u} = 0 \quad (x, y, t) \in \Omega \times (0, T)$$

$$\frac{\partial \mathbf{u}}{\partial t} + c_1 \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nabla \sigma + T g \mathbf{j} \quad (x, y, t) \in \Omega \times (0, T)$$

$$\sigma = 2\mu_T \varepsilon(\mathbf{u}) - \text{tr} \sigma \mathbf{I} \quad (x, y, t) \in \Omega \times (0, T)$$

$$\frac{\partial T}{\partial t} + c_2 \mathbf{u} \cdot \nabla T = \lambda_T \nabla \cdot (\nabla T) \quad (x, y, t) \in \Omega \times (0, T)$$

*Correspondence to: Zhendong Luo, School of Mathematics and Physics, North China Electric Power University, Beijing 102206, People's Republic of China.

[†]E-mail: zhdluo@163.com

Contract/grant sponsor: The National Science Foundation of China; contract/grant numbers: 10871022, CCF-0635162

$$\begin{aligned} \frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C &= D_T \nabla \cdot (\nabla C) \quad (x, y, t) \in \Omega \times (0, T) \\ u(x, y, 0) = v(x, y, 0) &= T(x, y, 0) = C(x, y, 0) = 0 \quad (x, y) \in \Omega \\ u(x, y, t) = v(x, y, t) &= 0, T(x, y, t) = T_0, C(x, y, T) = C_0 \quad (x, y, t) \in \partial\Omega \times (0, T) \end{aligned} \quad (1)$$

where $\mathbf{u} = (u, v)$ represents the velocity vector, p the pressure, T the temperature, C the mass fraction of TMGa, \mathbf{g} the acceleration of gravity, μ_T the carrier gas viscosity, λ_T the thermal conductivity of the carrier gas, and D_T the diffusion coefficient of TMGa in the carrier gas. σ denotes the viscous stress tensor, $\sigma = (\sigma_{ij})_{2 \times 2}$. $\varepsilon(\mathbf{u}) = (\varepsilon_{ij})_{2 \times 2}$, $\varepsilon_{ij} = \frac{1}{2}(\partial u_i / \partial x_j + \partial u_j / \partial x_i)$, in which $u_i = u$, $u_j = v$, $x_i = x$, $x_j = y$. c_1 represents the density of the carrier gas. And $c_2 = c_p c_1$ in which c_p represents the heat capacity of the carrier gas.

The CVD equations (1) are a system of important physical equations with broad applied backgrounds. Although the finite difference scheme (FDS) is one of the most effective approaches to achieve numerical solution for many nonlinear partial equations [1–3], the usual FDS solutions usually involve many freedom degrees. So, how to simplify the computational load and save time becomes a key problem considering consuming calculations and resource demands in the actual computational process guarantees the sufficient accuracy of numerical solutions to some extent. Proper orthogonal decomposition (POD) is a technique that provides a useful tool for efficiently approximating a large amount of data and representing fluid flows with reduced number of degrees of freedom, i.e., with lower dimensional models to alleviate the computational load and memory requirements and has been successfully used in different fields including signal analysis and pattern recognition [4, 5], fluid dynamics and coherent structures [6–11], and optimal flow control problems [12–14]. More recently, some reduced-order finite difference models and MFE formulations and error estimates for the upper tropical pacific ocean model based on POD are presented [15–18], and a FDS based on POD for the non-stationary Navier–Stokes equations is established but its error analysis has not been derived [19]. Kunisch and Volkwein have presented some Galerkin POD methods for parabolic problems [20] and a general equation in fluid dynamics [21]. The singular value decomposition (SVD) approach combined with POD technique is used to treat the Burgers equation [22] and the cavity flow problem [23].

Though CVD equations are also dealt with POD in [13], to the best of our knowledge, there have been no published results addressing that POD is used to reduce the FDS for CVD equations and the error estimates between the usual FDS solutions and the reduced FDS solutions based on POD. In this paper, POD is used to reduce the FDS for CVD equations and the error estimates between the usual FDS solutions and the reduced FDS solutions are derived. At last, an example of numerical simulation is given to demonstrate that the errors between the reduced FDS solutions and the usual FDS solutions are consistent with theoretical results. Moreover, it is also shown that the reduced FDS based on POD is feasible and efficient in numerical simulations for CVD equations.

2. FDS FOR PROBLEM (1) AND SNAPSHOTS GENERATION

Let Δx and Δy be the spatial step increment in the x -direction and y -direction, respectively, and Δt be the time step increment, $u_{j+1/2, k}^n$ and $v_{j, k+1/2}^n$ denote function values of u and v at point $(x_{j+1/2}, y_k, t_n)$ and $(x_j, y_{k+1/2}, t_n)$ ($0 \leq j \leq J, 0 \leq k \leq K, 0 \leq n \leq N = T/\Delta t$), respectively.

In the following, we apply staggered net (see Figure 1) FDS to discretize Problem (1).

(1) Discretizing the continuous equation

$$\nabla \cdot \mathbf{u} = 0 \quad (2)$$

yields

$$\left[\frac{u_{j+1/2, k} - u_{j-1/2, k}}{\Delta x} + \frac{v_{j, k+1/2} - v_{j, k-1/2}}{\Delta y} \right]^n = 0 \quad (3)$$

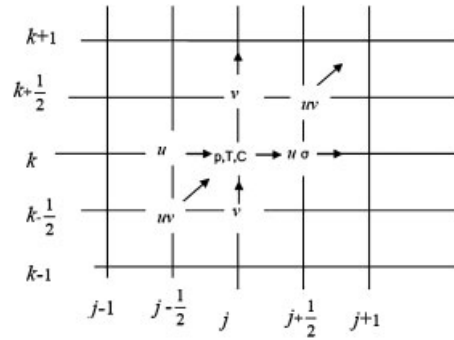


Figure 1. Staggered mesh graphics.

(2) The momentum equations

$$\frac{\partial \mathbf{u}}{\partial t} + c_1 \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nabla \sigma + T \mathbf{g} \mathbf{j} \quad (4)$$

$$\sigma + \text{tr} \sigma \mathbf{I} = 2\mu_T \varepsilon(\mathbf{u})$$

can be rewritten into the following equations

$$\frac{\partial u}{\partial t} + c_1 \left(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} \right) = -\frac{\partial p}{\partial x} + \mu_T \left(\frac{4}{3} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{1}{3} \frac{\partial^2 v}{\partial x \partial y} \right) \quad (5)$$

$$\frac{\partial v}{\partial t} + c_1 \left(v \frac{\partial v}{\partial y} + u \frac{\partial v}{\partial x} \right) = -\frac{\partial p}{\partial y} + \mu_T \left(\frac{4}{3} \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial x^2} + \frac{1}{3} \frac{\partial^2 u}{\partial x \partial y} \right) + Tg$$

Discretizing the momentum equation

$$\frac{\partial u}{\partial t} + c_1 \left(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} \right) = -\frac{\partial p}{\partial x} + \mu_T \left(\frac{4}{3} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{1}{3} \frac{\partial^2 v}{\partial x \partial y} \right) \quad (6)$$

on the x -direction at point $(x_{j+1/2}, y_k, t_n)$ yields

$$u_{j+1/2,k}^{n+1} = F_{j+1/2,k}^n - \frac{\Delta t}{\Delta x} (p_{j+1,k}^n - p_{j,k}^n) \quad (7)$$

where

$$F_{j+1/2,k}^n = u_{j+1/2,k}^n - c_1 \frac{\Delta t}{\Delta x} u_{j+1/2,k}^n (u_{j+1,k}^n - u_{j,k}^n)$$

$$- c_1 \frac{\Delta t}{\Delta y} v_{j+1/2,k}^n (u_{j+1/2,k+1/2}^n - u_{j+1/2,k-1/2}^n)$$

$$+ \Delta t \mu_T \left(\frac{u_{j+1/2,k-1} - 2u_{j+1/2,k} + u_{j+1/2,k+1}}{\Delta y^2} + \frac{4}{3} \frac{u_{j-1/2,k} - 2u_{j+1/2,k} + u_{j+3/2,k}}{\Delta x^2} \right.$$

$$\left. + \frac{1}{3} \frac{v_{j+1,k+1/2} - v_{j,k+1/2} - v_{j+1,k-1/2} + v_{j,k-1/2}}{\Delta x \Delta y} \right)^n \quad (8)$$

Discretizing the momentum equation

$$\frac{\partial v}{\partial t} + c_1 \left(v \frac{\partial v}{\partial y} + u \frac{\partial v}{\partial x} \right) = -\frac{\partial p}{\partial y} + \mu_T \left(\frac{4}{3} \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial x^2} + \frac{1}{3} \frac{\partial^2 u}{\partial x \partial y} \right) + Tg \quad (9)$$

on the y-direction at point $(x_j, y_{k+1/2}, t_n)$ yields

$$v_{j,k+1/2}^{n+1} = G_{j,k+1/2}^n - \frac{\Delta t}{\Delta y} (p_{j,k+1}^n - p_{j,k}^n) \tag{10}$$

where

$$\begin{aligned} G_{j,k+1/2}^n &= v_{j,k+1/2}^n - c_1 \frac{\Delta t}{\Delta x} u_{j,k+1/2}^n (v_{j+1/2,k+1/2}^n - v_{j-1/2,k+1/2}^n) \\ &\quad - c_1 \frac{\Delta t}{\Delta y} v_{j,k+1/2}^n (v_{j,k+1}^n - v_{j,k}^n) \\ &\quad + \Delta t \mu_T \left(\frac{v_{j-1,k+1/2} - 2v_{j,k+1/2} + v_{j+1,k+1/2}}{\Delta x^2} + \frac{4}{3} \frac{v_{j,k-1/2} - 2v_{j,k+1/2} + v_{j,k+3/2}}{\Delta y^2} \right. \\ &\quad \left. + \frac{1}{3} \frac{u_{j+1/2,k+1} - u_{j-1/2,k+1} - u_{j+1/2,k} + u_{j-1/2,k}}{\Delta x \Delta y} \right)^n + \Delta t g T_{j,k}^n \end{aligned} \tag{11}$$

The stress tensor equation

$$\sigma = 2\mu_T \varepsilon(\mathbf{u}) - \text{tr} \sigma \mathbf{I} \tag{12}$$

can be rewritten as

$$\begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_3 \end{pmatrix} = 2\mu_T \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) & \frac{\partial v}{\partial y} \end{pmatrix} - \begin{pmatrix} \sigma_1 + \sigma_3 & 0 \\ 0 & \sigma_1 + \sigma_3 \end{pmatrix} \tag{13}$$

such that

$$\begin{aligned} \sigma_1 &= \frac{1}{3} \left(4\mu_T \frac{\partial u}{\partial x} - 2\mu_T \frac{\partial v}{\partial y} \right) \\ \sigma_2 &= \mu_T \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \sigma_3 &= \frac{1}{3} \left(4\mu_T \frac{\partial v}{\partial y} - 2\mu_T \frac{\partial u}{\partial x} \right) \end{aligned} \tag{14}$$

at point $(x_{j+1/2}, y_k, t_n)$ yield

$$\begin{aligned} \sigma_{1,j,k}^n &= \left[\frac{4}{3} \mu_T \frac{u_{j+1/2,k} - u_{j-1/2,k}}{\Delta x} - \frac{2}{3} \mu_T \frac{v_{j,k+1/2} - v_{j,k-1/2}}{\Delta y} \right]^n \\ \sigma_{2,j,k}^n &= \left[\mu_T \frac{u_{j,k+1/2} - u_{j,k-1/2}}{\Delta y} + \mu_T \frac{v_{j+1/2,k} - v_{j-1/2,k}}{\Delta x} \right]^n \\ \sigma_{3,j,k}^n &= \left[\frac{4}{3} \mu_T \frac{v_{j,k+1/2} - v_{j,k-1/2}}{\Delta y} - \frac{2}{3} \mu_T \frac{u_{j+1/2,k} - u_{j-1/2,k}}{\Delta x} \right]^n \end{aligned} \tag{15}$$

(3) Expanding the energy equation

$$\frac{\partial T}{\partial t} + c_2 \mathbf{u} \cdot \nabla T = \lambda_T \nabla \cdot (\nabla T) \tag{16}$$

at point (x_j, y_k, t_n) yields

$$\begin{aligned} T_{j,k}^{n+1} &= T_{j,k}^n - \frac{c_2 \Delta t}{\Delta x} u_{j,k}^n (T_{j+1/2,k} - T_{j-1/2,k})^n - \frac{c_2 \Delta t}{\Delta y} v_{j,k}^n (T_{j,k+1/2} - T_{j,k-1/2})^n \\ &\quad + \lambda_T \Delta t \left(\frac{T_{j-1,k} - 2T_{j,k} + T_{j+1,k}}{\Delta x^2} + \frac{T_{j,k-1} - 2T_{j,k} + T_{j,k+1}}{\Delta y^2} \right)^n \end{aligned} \tag{17}$$

(4) Expanding the mass equation

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C = D_T \nabla \cdot (\nabla C) \quad (18)$$

at point (x_j, y_k, t_n) yields

$$\begin{aligned} C_{j,k}^{n+1} = & C_{j,k}^n - \frac{\Delta t}{\Delta x} u_{j,k}^n (C_{j+1/2,k} - C_{j-1/2,k})^n - \frac{\Delta t}{\Delta y} v_{j,k}^n (C_{j,k+1/2} - C_{j,k-1/2})^n \\ & + D_T \Delta t \left(\frac{C_{j-1,k} - 2C_{j,k} + C_{j+1,k}}{\Delta x^2} + \frac{C_{j,k-1} - 2C_{j,k} + C_{j,k+1}}{\Delta y^2} \right)^n \end{aligned} \quad (19)$$

Inserting (7) and (10) into (3) could obtain the Poisson equation for p as follows

$$\left[\frac{p_{j-1,k} - 2p_{j,k} + p_{j+1,k}}{\Delta x^2} + \frac{p_{j,k-1} - 2p_{j,k} + p_{j,k+1}}{\Delta y^2} \right]^n = \text{RHS} \quad (20)$$

where $\text{RHS} = \frac{1}{\Delta t \Delta x} (F_{j+1/2,k} - F_{j-1/2,k})^n + \frac{1}{\Delta t \Delta y} (G_{j,k+1/2} - G_{j,k-1/2})^n$.

Using the same approaches as the proof of the convergence and stability of finite difference equations of the non-stationary Navier–Stokes equation in [1] or [2], it is not difficult to prove the convergence and stability of usual FDS (3), (7), (10), (15), (17) and (19) for CVD equations, if $0.25(|u|^2 + |v|^2)\Delta t \leq \mu_T$, $\mu_T \Delta t \leq 0.25\Delta x^2$, and $\Delta t \mu_T \leq 0.25\Delta y^2$. And we can conclude the following result using the Taylor expansion.

Theorem 1

Usual FDS (3), (7), (10), (15), (17) and (19) for CVD equations has the following error estimates

$$\begin{aligned} & |E_n(u_{j+1/2,k}^n, v_{j,k+1/2}^n, p_{j,k}^n, T_{j,k}^n, C_{j,k}^n)| \\ & = \|(u(x_{j+1/2}, y_k, t_n), v(x_j, y_{k+1/2}, t_n), p(x_j, y_k, t_n), T(x_j, y_k, t_n), C(x_j, y_k, t_n)) \\ & \quad - (u_{j+1/2,k}^n, v_{j,k+1/2}^n, p_{j,k}^n, T_{j,k}^n, C_{j,k}^n)\| = O(\Delta t, \Delta x^2, \Delta y^2), \quad 1 \leq n \leq N \end{aligned} \quad (21)$$

where $\|\cdot\|$ denotes the usual norm of vector.

$$|E_n(\sigma_{j+1/2,k}^n)| = \|(\sigma(x_{j+1/2}, y_k, t_n) - \sigma_{j+1/2,k}^n)\| = O(\Delta x^2, \Delta y^2), \quad 1 \leq n \leq N \quad (22)$$

where $\|\cdot\|$ denotes the usual norm of matrix.

The procedure to achieve FDS numerical solutions for problem (I) is first to solve (20), then (7) and (10), and then (15), (17) and (19), from which we obtain a set of discrete numerical solutions.

Thus, if the carrier gas viscosity μ_T , the thermal conductivity of the carrier gas λ_T , the diffusion coefficient D_T , time step increment Δt , and spacial step increment Δx and Δy in the x -direction and y -direction are given, by solving (3), (7), (10), (15), (17) and (19) one could obtain $u_{j+1/2,k}^n$, $v_{j,k+1/2}^n$, $p_{j,k}^n$, $T_{j,k}^n$, $C_{j,k}^n$, and $\sigma_{j+1/2,k}^n$ ($0 \leq j \leq J$, $0 \leq k \leq K$, $1 \leq n \leq N$).

Write $u_i^n = u_{j+1/2,k}^n$, $v_i^n = v_{j,k+1/2}^n$, $p_i^n = p_{j,k}^n$, $T_i^n = T_{j,k}^n$, and $C_i^n = C_{j,k}^n$ ($i = k(J+1) + j + 1$, $m = JK$, $1 \leq i \leq m$, $0 \leq j \leq J$, $0 \leq k \leq K$, $1 \leq n \leq N$). $L \times m$ group of values consisting of the ensemble $\{u_i^{n_l}, v_i^{n_l}, p_i^{n_l}, T_i^{n_l}, C_i^{n_l}\}_{l=1}^L$ ($1 \leq i \leq m$) (usually $L \ll N$), known as ‘snapshots’, are chosen from $N \times m$ group of $\{u_i^n, v_i^n, p_i^n, T_i^n, C_i^n\}_{n=1}^N$ ($1 \leq i \leq m$).

Remark 1

When one computes actual problems, one may obtain the ensemble of snapshots from physical system trajectories via drawing samples of experiments and interpolation (or data assimilation). For example of weather forecast, one can use the previous weather results to structure the ensemble of snapshots, then restructure the optimal basis for the ensemble of snapshots by the following SVD and POD, and finally combine with POD projection to derive a reduced optimizing dynamical system. Thus, the situation of future weather change can be quickly simulated and the future weather change can be forecasted, which is of major importance for actual real-life applications.

3. POD REDUCED MODEL FOR CVD EQUATIONS

In the following, we derive POD basis from the snapshots generated above with SVD, and then use the POD basis to develop a reduced optimizing FDS for CVD equations.

3.1. SVD and POD basis

The ensemble of snapshots $\{u_i^{n1}, v_i^{n1}, p_i^{n1}, T_i^{n1}, C_i^{n1}\}_{i=1}^L$ ($1 \leq i \leq m$) (usually $L \ll N$) can be expressed as the following $m \times L$ matrices $\mathbf{A}_u, \mathbf{A}_v, \mathbf{A}_p, \mathbf{A}_T$, and \mathbf{A}_C .

$$\begin{aligned} \mathbf{A}_u &= \begin{pmatrix} u_1^{n1} & u_1^{n2} & \cdots & u_1^{nL} \\ u_2^{n1} & u_2^{n2} & \cdots & u_2^{nL} \\ \vdots & \vdots & \vdots & \vdots \\ u_m^{n1} & u_m^{n2} & \cdots & u_m^{nL} \end{pmatrix}, & \mathbf{A}_v &= \begin{pmatrix} v_1^{n1} & v_1^{n2} & \cdots & v_1^{nL} \\ v_2^{n1} & v_2^{n2} & \cdots & v_2^{nL} \\ \vdots & \vdots & \vdots & \vdots \\ v_m^{n1} & v_m^{n2} & \cdots & v_m^{nL} \end{pmatrix} \\ \mathbf{A}_p &= \begin{pmatrix} p_1^{n1} & p_1^{n2} & \cdots & p_1^{nL} \\ p_2^{n1} & p_2^{n2} & \cdots & p_2^{nL} \\ \vdots & \vdots & \vdots & \vdots \\ p_m^{n1} & p_m^{n2} & \cdots & p_m^{nL} \end{pmatrix}, & \mathbf{A}_T &= \begin{pmatrix} T_1^{n1} & T_1^{n2} & \cdots & T_1^{nL} \\ T_2^{n1} & T_2^{n2} & \cdots & T_2^{nL} \\ \vdots & \vdots & \vdots & \vdots \\ T_m^{n1} & T_m^{n2} & \cdots & T_m^{nL} \end{pmatrix} \\ \mathbf{A}_C &= \begin{pmatrix} C_1^{n1} & C_1^{n2} & \cdots & C_1^{nL} \\ C_2^{n1} & C_2^{n2} & \cdots & C_2^{nL} \\ \vdots & \vdots & \vdots & \vdots \\ C_m^{n1} & C_m^{n2} & \cdots & C_m^{nL} \end{pmatrix} \end{aligned} \quad (23)$$

In order to obtain optimal representation for \mathbf{A}_u ($\mathbf{A}_v, \mathbf{A}_p, \mathbf{A}_T$, and \mathbf{A}_C are similar), we employ SVD to research FDS for CVD equations, which is an important tool to construct optimal basis of optimization approximation. For matrix $\mathbf{A}_u \in R^{m \times L}$, there exists the SVD

$$\mathbf{A}_u = \mathbf{U}_u \begin{pmatrix} \mathbf{S}_u & 0 \\ 0 & 0 \end{pmatrix} \mathbf{V}_u^T \quad (24)$$

where $\mathbf{U}_u \in R^{m \times m}$ and $\mathbf{V}_u \in R^{L \times L}$ are all orthogonal matrices, $\mathbf{S}_u = \text{diag}\{\sigma_{u1}, \sigma_{u2}, \dots, \sigma_{u\ell}\} \in R^{\ell \times \ell}$ is the diagonal matrix correspondent to \mathbf{A}_u , and σ_{ui} ($i = 1, 2, \dots, \ell$) are the positive singular values. The matrices $\mathbf{U}_u = (\boldsymbol{\phi}_{u1}, \boldsymbol{\phi}_{u2}, \dots, \boldsymbol{\phi}_{um}) \in R^{m \times m}$ and $\mathbf{V}_u = (\boldsymbol{\phi}_{u1}, \boldsymbol{\phi}_{u2}, \dots, \boldsymbol{\phi}_{uL}) \in R^{L \times L}$ contain the orthogonal eigenvectors to the $\mathbf{A}_u \mathbf{A}_u^T$ and $\mathbf{A}_u^T \mathbf{A}_u$, respectively. The columns of these eigenvector matrices are organized such that corresponding to the singular values σ_{ui} are comprised in \mathbf{S}_u in a non-increasing order. And the singular values of the decomposition are connected to the eigenvalues of the matrices $\mathbf{A}_u \mathbf{A}_u^T$ and $\mathbf{A}_u^T \mathbf{A}_u$ in the manner such that $\lambda_{ui} = \sigma_{ui}^2$ ($i = 1, 2, \dots, \ell$). Since the number of mesh points is far larger than that of transient moment points, i.e. $m \gg L$, that is also that the order m for matrix $\mathbf{A}_u \mathbf{A}_u^T$ is far larger than the order L for matrix $\mathbf{A}_u^T \mathbf{A}_u$, however, their null eigenvalues are identical, therefore, we may first solve the eigen equation corresponding to matrix $\mathbf{A}_u^T \mathbf{A}_u$ to find the eigenvectors $\boldsymbol{\phi}_{uj}$ ($j = 1, 2, \dots, L$), and then by

$$\boldsymbol{\phi}_{uj} = \frac{1}{\sigma_{uj}} \mathbf{A}_u \boldsymbol{\phi}_{uj}, \quad j = 1, 2, \dots, \ell \quad (25)$$

we may obtain ℓ ($\ell \leq L$) eigenvectors corresponding to the non-null eigenvalues for matrix $\mathbf{A}_u \mathbf{A}_u^T$.

Define matrix norm $\|\cdot\|_{\alpha, \beta}$ as $\|\mathbf{A}_u\|_{\alpha, \beta} = \sup_{\mathbf{x} \neq 0} \|\mathbf{A}_u \mathbf{x}\|_{\alpha} / \|\mathbf{x}\|_{\beta}$ (where $\|\cdot\|_{\alpha}$ and $\|\cdot\|_{\beta}$ are the norm of vector). Let $\mathbf{A}_{M_u} = \sum_{i=1}^{M_u} \sigma_{ui} \boldsymbol{\phi}_{ui} \boldsymbol{\phi}_{ui}^T$, $\boldsymbol{\phi}_{ui}$ ($i = 1, 2, \dots, M_u$) and $\boldsymbol{\phi}_{uj}$ ($j = 1, 2, \dots, M_u$) are M_u

first column vectors of matrices \mathbf{U}_u and \mathbf{V}_u , respectively. Then, by the relationship properties between spectral radius and $\|\cdot\|_{2,2}$ for matrix, if $M_u < r = \text{rank} \mathbf{A}_u$ ($r \leq \ell \leq L$), then there is the following equation

$$\min_{\text{rank}(\mathbf{B}) \leq M_u} \|\mathbf{A}_u - \mathbf{B}\|_{2,2} = \|\mathbf{A}_u - \mathbf{A}_{M_u}\|_{2,2} = \sigma_{u(M_u+1)} \quad (26)$$

which shows that \mathbf{A}_{M_u} is an optimal representation of \mathbf{A}_u , i.e., \mathbf{A}_{M_u} is an optimal approximation of \mathbf{A}_u and the error is $\sigma_{u(M_u+1)} = \sqrt{\lambda_{u(M_u+1)}}$.

Denote the L column vectors of matrix \mathbf{A}_u by $\mathbf{a}_u^l = (u_1^{n_l}, u_2^{n_l}, \dots, u_m^{n_l})^T$ ($l = 1, 2, \dots, L$), and ε_l ($l = 1, 2, \dots, L$) by unit column vectors except that l th component is 1, while the other components are 0. Then by the compatibility of the norm for matrixes and vectors, we obtain that

$$\|\mathbf{a}_u^l - P_{M_u}(\mathbf{a}_u^l)\|_2 = \|(A_u - A_{M_u})\varepsilon_l\|_2 \leq \|A_u - A_{M_u}\|_{2,2} \|\varepsilon_l\|_2 = \sqrt{\lambda_{u(M_u+1)}} \quad (27)$$

where $P_{M_u}(\mathbf{a}_u^l) = \sum_{j=1}^{M_u} (\Phi_{uj}, \mathbf{a}_u^l) \Phi_{uj}$, $(\Phi_{uj}, \mathbf{a}_u^l)$ are the canonical inner products for vector Φ_{uj} and vector \mathbf{a}_u^l . Inequality (27) shows that $P_{M_u}(\mathbf{a}_u^l)$ are the optimal approximations to \mathbf{a}_u^l , whose errors are all $\sqrt{\lambda_{u(M_u+1)}}$. Thus, a group of optimal basis is found in the construction of A_{M_u} . By the property of eigenvectors, it is well known that $\Phi_u = (\Phi_{u1}, \Phi_{u2}, \dots, \Phi_{uM_u})$ ($M_u \ll L$) is an orthonormal matrix and $\{\Phi_{uj}\}_{j=1}^{M_u}$ is a group of POD optimal bases.

By the same approach as the above (27), if $\mathbf{a}_v^l = (v_1^{n_l}, v_2^{n_l}, \dots, v_m^{n_l})^T$, $\mathbf{a}_p^l = (p_1^{n_l}, p_2^{n_l}, \dots, p_m^{n_l})^T$, $\mathbf{a}_T^l = (T_1^{n_l}, T_2^{n_l}, \dots, T_m^{n_l})^T$, and $\mathbf{a}_C^l = (C_1^{n_l}, C_2^{n_l}, \dots, C_m^{n_l})^T$ ($l = 1, 2, \dots, L$) are the L column vectors of matrices \mathbf{A}_v , \mathbf{A}_p , \mathbf{A}_T , and \mathbf{A}_C , then $P_{M_v}(\mathbf{a}_v^l) = \sum_{j=1}^{M_v} (\Phi_{vj}, \mathbf{a}_v^l) \Phi_{vj}$, $P_{M_p}(\mathbf{a}_p^l) = \sum_{j=1}^{M_p} (\Phi_{pj}, \mathbf{a}_p^l) \Phi_{pj}$, $P_{M_T}(\mathbf{a}_T^l) = \sum_{j=1}^{M_T} (\Phi_{Tj}, \mathbf{a}_T^l) \Phi_{Tj}$, and $P_{M_C}(\mathbf{a}_C^l) = \sum_{j=1}^{M_C} (\Phi_{Cj}, \mathbf{a}_C^l) \Phi_{Cj}$ are the optimal approximations to \mathbf{a}_v^l , \mathbf{a}_p^l , \mathbf{a}_T^l , and \mathbf{a}_C^l , whose errors are

$$\|\mathbf{a}_v^l - P_{M_v}(\mathbf{a}_v^l)\|_2 \leq \sqrt{\lambda_{v(M_v+1)}} \quad (28)$$

$$\|\mathbf{a}_p^l - P_{M_p}(\mathbf{a}_p^l)\|_2 \leq \sqrt{\lambda_{p(M_p+1)}} \quad (29)$$

$$\|\mathbf{a}_T^l - P_{M_T}(\mathbf{a}_T^l)\|_2 \leq \sqrt{\lambda_{T(M_T+1)}} \quad (30)$$

$$\|\mathbf{a}_C^l - P_{M_C}(\mathbf{a}_C^l)\|_2 \leq \sqrt{\lambda_{C(M_C+1)}} \quad (31)$$

where $\lambda_{v(M_v+1)}$, $\lambda_{p(M_p+1)}$, $\lambda_{T(M_T+1)}$, and $\lambda_{C(M_C+1)}$ are (M_v+1) th eigenvalue for $\mathbf{A}_v \mathbf{A}_v^T$, (M_p+1) th eigenvalue for $\mathbf{A}_p \mathbf{A}_p^T$, (M_T+1) th eigenvalue for $\mathbf{A}_T \mathbf{A}_T^T$, and (M_C+1) th eigenvalue for $\mathbf{A}_C \mathbf{A}_C^T$, and $\Phi_v = (\Phi_{v1}, \Phi_{v2}, \dots, \Phi_{vM_v})$, $\Phi_p = (\Phi_{p1}, \Phi_{p2}, \dots, \Phi_{pM_p})$, $\Phi_T = (\Phi_{T1}, \Phi_{T2}, \dots, \Phi_{TM_T})$, and $\Phi_C = (\Phi_{C1}, \Phi_{C2}, \dots, \Phi_{CM_C})$ are orthonormal matrixes and $\{\Phi_{vj}\}_{j=1}^{M_v}$, $\{\Phi_{pj}\}_{j=1}^{M_p}$, $\{\Phi_{Tj}\}_{j=1}^{M_T}$, and $\{\Phi_{Cj}\}_{j=1}^{M_C}$ are three groups of optimal basis corresponding to \mathbf{A}_v , \mathbf{A}_p , \mathbf{A}_T , and \mathbf{A}_C , respectively.

3.2. Reduced optimizing FDS based on POD for CVD equations

In the following, we use POD basis to develop a reduced optimizing FDS for CVD equations.

Write

$$\begin{aligned} \mathbf{u}_m(t) &= (u_1(t), u_2(t), \dots, u_m(t))^T \\ \mathbf{v}_m(t) &= (v_1(t), v_2(t), \dots, v_m(t))^T \\ \mathbf{p}_m(t) &= (p_1(t), p_2(t), \dots, p_m(t))^T \\ \mathbf{T}_m(t) &= (T_1(t), T_2(t), \dots, T_m(t))^T \\ \mathbf{C}_m(t) &= (C_1(t), C_2(t), \dots, C_m(t))^T \end{aligned} \quad (32)$$

where $u_i = u_{j+1/2,k}$, $v_i = v_{j,k+1/2}$, $p_i = p_{j,k}$, $T_i = T_{j,k}$, and $C_i = C_{j,k}$ ($1 \leq i \leq m$, $i = k(J+1) + j + 1$, $m = KJ$, $0 \leq j \leq J-1$, $0 \leq k \leq K-1$). Thus, (3), (7), (10), (13), and (15) are rewritten as the following vector formulation.

$$(\mathbf{u}_m^{n+1}, \mathbf{v}_m^{n+1}, \mathbf{p}_m^{n+1}, \mathbf{T}_m^{n+1}, \mathbf{C}_m^{n+1})^T = (\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n)^T + \Delta t \tilde{\mathbf{F}}(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n), \quad 0 \leq n \leq N \quad (33)$$

where $\tilde{\mathbf{F}}(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n) = (\tilde{F}_1(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n), \dots, \tilde{F}_5(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n))^T$ is the vector function obtained from (3), (7), (10), (13), and (15). Put

$$(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n)^T = (\Phi_u \alpha_{M_u}^{u,n}, \Phi_v \alpha_{M_v}^{v,n}, \Phi_p \alpha_{M_p}^{p,n}, \Phi_T \alpha_{M_T}^{T,n}, \Phi_C \alpha_{M_C}^{C,n})^T \quad (34)$$

where $\mathbf{u}_m^n = (u_1^n, u_2^n, \dots, u_m^n)^T$, $\mathbf{v}_m^n = (v_1^n, v_2^n, \dots, v_m^n)^T$ and $\mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n$ likewise. Inserting (34) into (33) and noting that $\Phi_u, \Phi_v, \Phi_p, \Phi_T, \Phi_C$ are orthogonal matrices, we may obtain a reduced optimizing model that has $M_u + M_v + M_p + M_T + M_C$ ($M_u, M_v, M_p, M_T, M_C \ll L \ll m$) unknown values:

$$\begin{pmatrix} \alpha_{M_u}^{u,n+1} \\ \alpha_{M_v}^{v,n+1} \\ \alpha_{M_p}^{p,n+1} \\ \alpha_{M_T}^{T,n+1} \\ \alpha_{M_C}^{C,n+1} \end{pmatrix} = \begin{pmatrix} \alpha_{M_u}^{u,n} \\ \alpha_{M_v}^{v,n} \\ \alpha_{M_p}^{p,n} \\ \alpha_{M_T}^{T,n} \\ \alpha_{M_C}^{C,n} \end{pmatrix} + \Delta t \begin{pmatrix} \Phi_u^T \tilde{F}_1(\Phi_u \alpha_{M_u}^{u,n}, \dots, \Phi_C \alpha_{M_C}^{C,n}) \\ \Phi_v^T \tilde{F}_2(\Phi_v \alpha_{M_v}^{v,n}, \dots, \Phi_C \alpha_{M_C}^{C,n}) \\ \Phi_p^T \tilde{F}_3(\Phi_p \alpha_{M_p}^{p,n}, \dots, \Phi_C \alpha_{M_C}^{C,n}) \\ \Phi_T^T \tilde{F}_4(\Phi_T \alpha_{M_T}^{T,n}, \dots, \Phi_C \alpha_{M_C}^{C,n}) \\ \Phi_C^T \tilde{F}_5(\Phi_C \alpha_{M_C}^{C,n}, \dots, \Phi_C \alpha_{M_C}^{C,n}) \end{pmatrix} \quad (35)$$

where $n=0, 1, 2, \dots, N$, initial values are $\alpha_{M_u}^{u,0} = \Phi_u^T \mathbf{u}_m^0$, $\alpha_{M_v}^{v,0} = \Phi_v^T \mathbf{v}_m^0$ and $\alpha_{M_p}^{p,0}$, $\alpha_{M_T}^{T,0}$, $\alpha_{M_C}^{C,0}$ respectively.

After one has obtained $\alpha_{M_u}^{u,n}$, $\alpha_{M_v}^{v,n}$, $\alpha_{M_p}^{p,n}$, $\alpha_{M_T}^{T,n}$, and $\alpha_{M_C}^{C,n}$ from (35), one obtains the POD optimal solutions, which are written as $\mathbf{u}_i^{*n} = \Phi_u \alpha_{M_u}^{u,n}$, $\mathbf{v}_i^{*n} = \Phi_v \alpha_{M_v}^{v,n}$, $\mathbf{p}_i^{*n} = \Phi_p \alpha_{M_p}^{p,n}$, $\mathbf{T}_i^{*n} = \Phi_T \alpha_{M_T}^{T,n}$, and $\mathbf{C}_i^{*n} = \Phi_C \alpha_{M_C}^{C,n}$ for Problem (I) by (34). Thus, we get the optimal numerical solutions, which are written as $(u_{j+1/2,k}^{*n}, v_{j,k+1/2}^{*n}, p_{j,k}^{*n}, T_{j,k}^{*n}, C_{j,k}^{*n})$ ($0 \leq j \leq J-1, 0 \leq k \leq K-1, 0 \leq n \leq N$) for Problem (I), where $u_{j+1/2,k}^{*n} = u_i^{*n}$, $v_{j,k+1/2}^{*n} = v_i^{*n}$, $p_{j,k}^{*n} = p_i^{*n}$, $T_{j,k}^{*n} = T_i^{*n}$, and $C_{j,k}^{*n} = C_i^{*n}$ ($j = i - 1 - k(J+1) \geq 0, 1 \leq i \leq m = KJ, 0 \leq k \leq K-1, 0 \leq n \leq N$).

Remark 2

Formula (35) with (34) and (50) is the reduced optimizing FDS based on SVD and POD for Problem (I), because it only involves $(M_u + M_v + M_p + M_T + M_C) \times N$ ($M_u, M_v, M_p, M_T, M_C \ll L \ll m$) freedom degrees while usual FDS (3), (7), (10), (13), and (15) involves $5m \times N$. When it comes to actual problems, the POD basis can be structured by the known ensemble of snapshots, and then combine it with POD projection to derive a reduced optimizing FDS, i.e. one needs only to solve the above formula (35) with (34) and (50) which has much fewer degrees of freedom instead of the usual FDS (3), (7), (10), (13), and (15). Thus, the computational load and memory requirements can be saved greatly.

4. ERROR ANALYSIS OF POD REDUCED FDS

In this part, the error estimates of reduced optimizing FDS (34), (35), and (50) for Problem (I) are discussed.

Let

$$\begin{aligned} \mathcal{X}_u &= \text{span}\{\phi_{u1}, \phi_{u2}, \dots, \phi_{uM_u}\} \\ \mathcal{X}_v &= \text{span}\{\phi_{v1}, \phi_{v2}, \dots, \phi_{vM_v}\} \\ \mathcal{X}_p &= \text{span}\{\phi_{p1}, \phi_{p2}, \dots, \phi_{pM_p}\} \\ \mathcal{X}_T &= \text{span}\{\phi_{T1}, \phi_{T2}, \dots, \phi_{TM_T}\} \\ \mathcal{X}_C &= \text{span}\{\phi_{C1}, \phi_{C2}, \dots, \phi_{CM_C}\} \end{aligned} \tag{36}$$

Then, for column vectors \mathbf{a}_u^l ($1 \leq l \leq L$) of \mathbf{A}_u , by (27) we have $\mathbf{a}_u^l = \mathbf{u}_m^{n_l}$, and there is a $P_{M_u}(\mathbf{u}_m^{n_l}) = P_{M_u}(\mathbf{a}_u^l) = \sum_{j=1}^{M_u} (\phi_{uj}, \mathbf{a}_u^l) \phi_{uj} = \sum_{j=1}^{M_u} (\phi_{uj}, \mathbf{u}_m^{n_l}) \phi_{uj} \in \mathcal{X}_u$ such that

$$\|\mathbf{u}_m^{n_l} - P_{M_u}(\mathbf{u}_m^{n_l})\|_2 \leq \sqrt{\lambda_u(M_u+1)}, \quad 1 \leq l \leq L \tag{37}$$

while $n \in \{n_1, n_2, \dots, n_L\}$, $\mathbf{u}_m^{*n} = P_{M_u}(\mathbf{u}_m^n) = \sum_{j=1}^{M_u} (\phi_{uj}, \mathbf{u}_m^n) \phi_{uj}$ is obtained by (34) and (35), therefore, we obtain that

$$\|\mathbf{u}_m^n - \mathbf{u}_m^{*n}\|_2 \leq \sqrt{\lambda_u(M_u+1)}, \quad n \in \{n_1, n_2, \dots, n_L\} \tag{38}$$

Using the same approach as (38), we could obtain that

$$\|\mathbf{v}_m^n - \mathbf{v}_m^{*n}\|_2 \leq \sqrt{\lambda_v(M_v+1)}, \quad n \in \{n_1, n_2, \dots, n_L\} \tag{39}$$

$$\|\mathbf{p}_m^n - \mathbf{p}_m^{*n}\|_2 \leq \sqrt{\lambda_p(M_p+1)}, \quad n \in \{n_1, n_2, \dots, n_L\} \tag{40}$$

$$\|\mathbf{T}_m^n - \mathbf{T}_m^{*n}\|_2 \leq \sqrt{\lambda_T(M_T+1)}, \quad n \in \{n_1, n_2, \dots, n_L\} \tag{41}$$

$$\|\mathbf{C}_m^n - \mathbf{C}_m^{*n}\|_2 \leq \sqrt{\lambda_C(M_C+1)}, \quad n \in \{n_1, n_2, \dots, n_L\} \tag{42}$$

When $n \notin \{n_1, n_2, \dots, n_L\}$, we may as well let $t_n \in (t_{n_l}, t_{n_{l+1}})$ and t_n be the nearest point to t_{n_l} . Comparing (34)–(35) with (33), (34)–(35) can be written in similar forms as (3), (7), (10), (13), and (15) as follows.

$$u_{j+1/2,k}^{*n+1} = F_{j+1/2,k}^{*n} - \frac{\Delta t}{\Delta x} (p_{j+1,k}^{*n} - p_{j,k}^{*n}) \tag{43}$$

where

$$\begin{aligned} F_{j+1/2,k}^{*n} &= u_{j+1/2,k}^{*n} - c_1 \frac{\Delta t}{\Delta x} u_{j+1/2,k}^{*n} (u_{j+1,k}^{*n} - u_{j,k}^{*n}) - c_1 \frac{\Delta t}{\Delta y} v_{j+1/2,k}^{*n} (u_{j+1/2,k+1/2}^{*n} - u_{j+1/2,k-1/2}^{*n}) \\ &\quad + \Delta t \mu_T \left(\frac{u_{j+1/2,k-1} + 2u_{j+1/2,k} + u_{j+1/2,k+1}}{\Delta y^2} + \frac{4}{3} \frac{u_{j-1/2,k} - 2u_{j+1/2,k} + u_{j+3/2,k}}{\Delta x^2} \right. \\ &\quad \left. + \frac{1}{3} \frac{v_{j+1,k+1/2} - v_{j,k+1/2} - v_{j+1,k-1/2} + v_{j,k-1/2}}{\Delta x \Delta y} \right)^{*n} \end{aligned} \tag{44}$$

$$v_{j,k+1/2}^{*n+1} = G_{j,k+1/2}^{*n} - \frac{\Delta t}{\Delta y} (p_{j,k+1}^{*n} - p_{j,k}^{*n}) \tag{45}$$

where

$$\begin{aligned} G_{j,k+1/2}^{*n} &= v_{j,k+1/2}^{*n} - c_1 \frac{\Delta t}{\Delta x} u_{j,k+1/2}^{*n} (v_{j+1/2,k+1/2}^{*n} - v_{j-1/2,k+1/2}^{*n}) - c_1 \frac{\Delta t}{\Delta y} v_{j,k+1/2}^{*n} (v_{j,k+1}^{*n} - v_{j,k}^{*n}) \\ &\quad + \Delta t \mu_T \left(\frac{v_{j-1,k+1/2} - 2v_{j,k+1/2} + v_{j+1,k+1/2}}{\Delta x^2} + \frac{4}{3} \frac{v_{j,k-1/2} - 2v_{j,k+1/2} + v_{j,k+3/2}}{\Delta y^2} \right) \end{aligned}$$

$$+\frac{1}{3}\frac{u_{j+1/2,k+1}-u_{j-1/2,k+1}-u_{j+1/2,k}+u_{j-1/2,k}}{\Delta x\Delta y}\Big)^{*n}+\Delta t g T_{j,k}^{*n} \tag{46}$$

$$T_{j,k}^{*n+1}=T_{j,k}^{*n}-\frac{c_2\Delta t}{\Delta x}u_{j,k}^{*n}(T_{j+1/2,k}-T_{j-1/2,k})^{*n}-\frac{c_2\Delta t}{\Delta y}v_{j,k}^{*n}(T_{j,k+1/2}-T_{j,k-1/2})^{*n} \\ +\lambda_T\Delta t\left(\frac{T_{j-1,k}-2T_{j,k}+T_{j+1,k}}{\Delta x^2}+\frac{T_{j,k-1}-2T_{j,k}+T_{j,k+1}}{\Delta y^2}\right)^{*n} \tag{47}$$

$$C_{j,k}^{*n+1}=C_{j,k}^{*n}-\frac{\Delta t}{\Delta x}u_{j,k}^{*n}(C_{j+1/2,k}-C_{j-1/2,k})^{*n}-\frac{\Delta t}{\Delta y}v_{j,k}^{*n}(C_{j,k+1/2}-C_{j,k-1/2})^{*n} \\ +D_T\Delta t\left(\frac{C_{j-1,k}-2C_{j,k}+C_{j+1,k}}{\Delta x^2}+\frac{C_{j,k-1}-2C_{j,k}+C_{j,k+1}}{\Delta y^2}\right)^{*n} \tag{48}$$

$$\left[\frac{p_{j-1,k}-2p_{j,k}+p_{j+1,k}}{\Delta x^2}+\frac{p_{j,k-1}-2p_{j,k}+p_{j,k+1}}{\Delta y^2}\right]^{*n}=\text{RHS} \tag{49}$$

where $\text{RHS}=(F_{j+1/2,k}-F_{j-1/2,k})^{*n}/(\Delta t\Delta x)+(G_{j,k+1/2}-G_{j,k-1/2})^{*n}/(\Delta t\Delta y)$, and

$$\sigma_{1,j,k}^{*n}=\left[\frac{4}{3}\mu_T\frac{u_{j+1/2,k}-u_{j-1/2,k}}{\Delta x}-\frac{2}{3}\mu_T\frac{v_{j,k+1/2}-v_{j,k-1/2}}{\Delta y}\right]^{*n} \\ \sigma_{2,j,k}^{*n}=\left[\mu_T\frac{u_{j,k+1/2}-u_{j,k-1/2}}{\Delta y}+\mu_T\frac{v_{j+1/2,k}-v_{j-1/2,k}}{\Delta x}\right]^{*n} \tag{50} \\ \sigma_{3,j,k}^{*n}=\left[\frac{4}{3}\mu_T\frac{v_{j,k+1/2}-v_{j,k-1/2}}{\Delta y}-\frac{2}{3}\mu_T\frac{u_{j+1/2,k}-u_{j-1/2,k}}{\Delta x}\right]^{*n}$$

If $|u_{j+1/2,k}^n|, |v_{j+1/2,k}^n|, |u_{j,k+1/2}^n|, |v_{j,k+1/2}^n|, |u_{j+1/2,k}^{*n}|, |v_{j+1/2,k}^{*n}|, |u_{j,k+1/2}^{*n}|$, and $|v_{j,k+1/2}^{*n}|$ are all bounded, by subtracting (43)–(50) from (3), (7), (10), (15), (17), and (19), respectively, and using (3) and

$$\left[\frac{u_{j+1/2,k}-u_{j-1/2,k}}{\Delta x}+\frac{v_{j,k+1/2}-v_{j,k-1/2}}{\Delta y}\right]^{*n}=0 \tag{51}$$

we can obtain that

$$\|\mathbf{u}_m^{n+1}-\mathbf{u}_m^{*n+1}\|_2+\|\mathbf{v}_m^{n+1}-\mathbf{v}_m^{*n+1}\|_2+\|\mathbf{T}_m^{n+1}-\mathbf{T}_m^{*n+1}\|_2+\|\mathbf{C}_m^{n+1}-\mathbf{C}_m^{*n+1}\|_2 \\ \leq M(\|\mathbf{u}_m^n-\mathbf{u}_m^{*n}\|_2+\|\mathbf{v}_m^n-\mathbf{v}_m^{*n}\|_2+\|\mathbf{T}_m^n-\mathbf{T}_m^{*n}\|_2+\|\mathbf{C}_m^n-\mathbf{C}_m^{*n}\|_2) \tag{52}$$

$$\|\mathbf{p}_m^n-\mathbf{p}_m^{*n}\|_2\leq M_0(\|\mathbf{u}_m^n-\mathbf{u}_m^{*n}\|_2+\|\mathbf{v}_m^n-\mathbf{v}_m^{*n}\|_2+\|\mathbf{T}_m^n-\mathbf{T}_m^{*n}\|_2) \tag{53}$$

and

$$\|\sigma_{im}^n-\sigma_{im}^{*n}\|_2\leq M_0(\|\mathbf{u}_m^n-\mathbf{u}_m^{*n}\|_2+\|\mathbf{v}_m^n-\mathbf{v}_m^{*n}\|_2), \quad i=1,2,3 \tag{54}$$

where $M=1+M_0$, $M_0=C\Delta t/\min(\Delta x, \Delta y, \mu_T^{-1}\Delta x^{-2}, \mu_T^{-1}\Delta y^{-2}, \mu_T^{-1}\Delta x^{-1}\Delta y^{-1})$, C is a constant independent of $\Delta t, \Delta x^2$, and Δy^2 . Summing (52) from $n_l, n_l+1, \dots, n-1$ can yield that

$$\|\mathbf{u}_m^n-\mathbf{u}_m^{*n}\|_2+\|\mathbf{v}_m^n-\mathbf{v}_m^{*n}\|_2+\|\mathbf{T}_m^n-\mathbf{T}_m^{*n}\|_2+\|\mathbf{C}_m^n-\mathbf{C}_m^{*n}\|_2 \\ \leq\|\mathbf{u}_m^{n_l}-\mathbf{u}_m^{*n_l}\|_2+\|\mathbf{v}_m^{n_l}-\mathbf{v}_m^{*n_l}\|_2+\|\mathbf{T}_m^{n_l}-\mathbf{T}_m^{*n_l}\|_2+\|\mathbf{C}_m^{n_l}-\mathbf{C}_m^{*n_l}\|_2 \\ +M_0\sum_{j=n_l}^{n-1}(\|\mathbf{u}_m^j-\mathbf{u}_m^{*j}\|_2+\|\mathbf{v}_m^j-\mathbf{v}_m^{*j}\|_2+\|\mathbf{T}_m^j-\mathbf{T}_m^{*j}\|_2+\|\mathbf{C}_m^j-\mathbf{C}_m^{*j}\|_2) \tag{55}$$

If $\Delta t = O(\Delta x^2, \Delta y^2)$, $(\mu_T)^2 \leq \Delta t$, by discrete Gronwall Lemma (see [24]), we get that

$$\begin{aligned} & \| \mathbf{u}_m^n - \mathbf{u}_m^{*n} \|_2 + \| \mathbf{v}_m^n - \mathbf{v}_m^{*n} \|_2 + \| \mathbf{T}_m^n - \mathbf{T}_m^{*n} \|_2 + \| \mathbf{C}_m^n - \mathbf{C}_m^{*n} \|_2 \\ & \leq (\| \mathbf{u}_m^{n_l} - \mathbf{u}_m^{*n_l} \|_2 + \| \mathbf{v}_m^{n_l} - \mathbf{v}_m^{*n_l} \|_2 + \| \mathbf{T}_m^{n_l} - \mathbf{T}_m^{*n_l} \|_2 + \| \mathbf{C}_m^{n_l} - \mathbf{C}_m^{*n_l} \|_2) \exp[C \Delta t^{1/2} (n - n_l)] \end{aligned} \quad (56)$$

If $t_l (1 \leq l \leq L)$ are uniformly chosen from $t_n (1 \leq n \leq N)$, then $(n - n_l) \leq N / (2L)$. If $L^2 = O(N)$, we obtain from (56) and (38)–(42) that

$$\begin{aligned} & \| \mathbf{u}_m^n - \mathbf{u}_m^{*n} \|_2 + \| \mathbf{v}_m^n - \mathbf{v}_m^{*n} \|_2 + \| \mathbf{T}_m^n - \mathbf{T}_m^{*n} \|_2 + \| \mathbf{C}_m^n - \mathbf{C}_m^{*n} \|_2 \\ & \leq C_0 (\sqrt{\lambda_u(M_u+1)} + \sqrt{\lambda_v(M_v+1)} + \sqrt{\lambda_T(M_T+1)} + \sqrt{\lambda_C(M_C+1)}) \end{aligned} \quad (57)$$

in which C_0 represents a constant.

Then the error estimates including u , v , p , T , and C can be put as follows.

Theorem 2

Let $(\mathbf{u}_m^n, \mathbf{v}_m^n, \mathbf{p}_m^n, \mathbf{T}_m^n, \mathbf{C}_m^n)$ ($n = 1, 2, \dots, N$) be vectors constituted with solutions of usual FDS (3), (7), (10), (15), (17), and (19), $(\mathbf{u}_m^{*n}, \mathbf{v}_m^{*n}, \mathbf{p}_m^{*n}, \mathbf{T}_m^{*n}, \mathbf{C}_m^{*n})$ be the vectors of the reduced optimizing FDS (34), (35), and (50), if $n \in \{n_1, n_2, \dots, n_L\}$, then the following error estimates exist

$$\begin{aligned} & \| \mathbf{u}_m^n - \mathbf{u}_m^{*n} \|_2 \leq \sqrt{\lambda_u(M_u+1)}, \quad \| \mathbf{v}_m^n - \mathbf{v}_m^{*n} \|_2 \leq \sqrt{\lambda_v(M_v+1)}, \quad \| \mathbf{p}_m^n - \mathbf{p}_m^{*n} \|_2 \leq \sqrt{\lambda_p(M_p+1)} \\ & \| \mathbf{T}_m^n - \mathbf{T}_m^{*n} \|_2 \leq \sqrt{\lambda_T(M_T+1)}, \quad \| \mathbf{C}_m^n - \mathbf{C}_m^{*n} \|_2 \leq \sqrt{\lambda_C(M_C+1)} \\ & \| \boldsymbol{\sigma}_{im}^n - \boldsymbol{\sigma}_{im}^{*n} \|_2 \leq C (\sqrt{\lambda_u(M_u+1)} + \sqrt{\lambda_v(M_v+1)}), \quad i = 1, 2, 3 \end{aligned} \quad (58)$$

Moreover, if $n \notin \{n_1, n_2, \dots, n_L\}$, $\Delta t = O(\Delta x^2, \Delta y^2)$, $(\mu_T)^2 \leq \Delta t$, $|u_{j+1/2,k}^n|, |v_{j,k+1/2}^n|, |p_{j,k}^n|, |T_{j,k}^n|, |C_{j,k}^n|, |u_{j+1/2,k}^{*n}|, |v_{j,k+1/2}^{*n}|, |p_{j,k}^{*n}|, |T_{j,k}^{*n}|, |C_{j,k}^{*n}|$ are all bounded, snapshots $\{u_{j+1/2,k}^{n_l}, v_{j+1/2,k}^{n_l}, p_{j,k}^{n_l}, T_{j,k}^{n_l}, C_{j,k}^{n_l}\}_{l=1}^L$ are uniformly chosen from $\{u_{j+1/2,k}^n, v_{j+1/2,k}^n, p_{j,k}^n, T_{j,k}^n, C_{j,k}^n\}_{n=1}^N$, $L^2 = O(N)$, then the following error estimates hold:

$$\begin{aligned} & \| \mathbf{u}_m^n - \mathbf{u}_m^{*n} \|_2 + \| \mathbf{v}_m^n - \mathbf{v}_m^{*n} \|_2 + \| \mathbf{p}_m^n - \mathbf{p}_m^{*n} \|_2 + \| \mathbf{T}_m^n - \mathbf{T}_m^{*n} \|_2 + \| \mathbf{C}_m^n - \mathbf{C}_m^{*n} \|_2 \\ & \leq C (\sqrt{\lambda_u(M_u+1)} + \sqrt{\lambda_v(M_v+1)} + \sqrt{\lambda_p(M_p+1)} + \sqrt{\lambda_T(M_T+1)} + \sqrt{\lambda_C(M_C+1)}) \\ & \| \boldsymbol{\sigma}_{im}^n - \boldsymbol{\sigma}_{im}^{*n} \|_2 \leq C (\sqrt{\lambda_u(M_u+1)} + \sqrt{\lambda_v(M_v+1)} + \sqrt{\lambda_p(M_p+1)} + \sqrt{\lambda_T(M_T+1)} \\ & \quad + \sqrt{\lambda_C(M_C+1)}), \quad i = 1, 2, 3 \end{aligned} \quad (59)$$

$$\quad (60)$$

where $N = T / \Delta t$ and L is the number of snapshots.

Given that the absolute value of each component of a vector is not more than its any norm, the following result could be achieved by combining Theorems 1 and 2.

Theorem 3

Under the assumptions of Theorem 2, the following error estimate holds

$$\begin{aligned} & |u(x_{j+1/2}, y_k, t_n) - u_{j+1/2,k}^{*n}| + |v(x_j, y_{k+1/2}, t_n) - v_{j,k+1/2}^{*n}| + |p(x_j, y_k, t_n) - p_{j,k}^{*n}| \\ & + |T(x_j, y_k, t_n) - T_{j,k}^{*n}| + |C(x_j, y_k, t_n) - C_{j,k}^{*n}| + |\sigma_1(x_{j+1/2}, y_k, t_n) - \sigma_{1j+1/2,k}^{*n}| \\ & + |\sigma_2(x_{j+1/2}, y_k, t_n) - \sigma_{2j+1/2,k}^{*n}| + |\sigma_3(x_{j+1/2}, y_k, t_n) - \sigma_{3j+1/2,k}^{*n}| \\ & \leq O(\sqrt{\lambda_u(M_u+1)} + \sqrt{\lambda_v(M_v+1)} + \sqrt{\lambda_p(M_p+1)} + \sqrt{\lambda_T(M_T+1)} + \sqrt{\lambda_C(M_C+1)}, \Delta t, \Delta x^2, \Delta y^2) \\ & 1 \leq n \leq N \end{aligned} \quad (61)$$

Remark 3

The condition $L^2 = O(N)$ in Theorem 2 shows the relation between the number L of snapshots and the number N of all time instances in fact. Therefore, it is unnecessary to take total transient solutions at all time instances t_n as snapshots (see [20, 21]). Theorems 2 and 3 have presented the error estimate between the solution of the reduced optimizing FDS (34), (35), and (50) and the solution of usual FDS (3), (7), (10), (15), (17), and (19), and the error estimate between the solution of the reduced FDS and Problem (I), respectively. Because our method employs the computed FDS solutions $(u_{j+1/2,k}^n, v_{j,k+1/2}^n, p_{j,k}^n, T_{j,k}^n, C_{j,k}^n)$ ($n=1, 2, \dots, N$) for Problem (I) as initial data for the reduced model, the error estimates in Theorem 3 are correlated to the gridding scale Δx and Δy , and time step size Δt . However, when it comes to actual problems, the ensemble of snapshots could be obtained from physical system trajectories by drawing samples from experiments and interpolation (or data assimilation). Thus, the assistant data $(u_{j+1/2,k}^n, v_{j,k+1/2}^n, p_{j,k}^n, T_{j,k}^n, C_{j,k}^n)$ ($n=1, 2, \dots, N$) could be replaced by the interpolation functions of experimental or previous results.

5. NUMERICAL SIMULATIONS

In this section, some numerical examples of the physical model of the cavity flows with the POD reduced FDS (34)–(35), and (50) are presented to demonstrate the feasibility and efficiency of the POD method.

Let the side length of the cavity be 1 (see Figure 2). We take spatial step increments as $\Delta x = \Delta y = \frac{1}{32}$ and time step increment as $\Delta t = 0.0001$. Except that

$$T = 4y(1 - y) \quad \text{and} \quad \left. \frac{\partial p}{\partial y} \right|_{\partial\Omega} = \mu_T \left(\frac{4}{3} \frac{\partial^2 v}{\partial y^2} \right) \Big|_{\partial\Omega}$$

on the right boundary, other initial and boundary values are all taken as 0. Take the constants $c_1 = c_2 = 0.05$, the viscosity $\mu_T = 0.01$, the thermal conductivity $\lambda_T = 1$, and the diffusion coefficient $D_T = 0.5$.

We obtain 20 discrete values (i.e. snapshots) at time $t = 10, 20, 30, \dots, 200$ by solving usual FDS (3), (7), (10), (15), (17), and (19). And it is shown by computing the eigenvalues that $\sqrt{\lambda_{u6}} + \sqrt{\lambda_{v6}} + \sqrt{\lambda_{p8}} + \sqrt{\lambda_{T8}} + \sqrt{\lambda_{C8}} = O(10^{-3})$.

When $t = 200$, we obtain the solutions of the POD reduced FDS (34)–(35), and (50) depicted in Figure 3 and 4 on the right-hand side used $M_u = M_v = 5$, $M_p = M_T = M_C = 7$

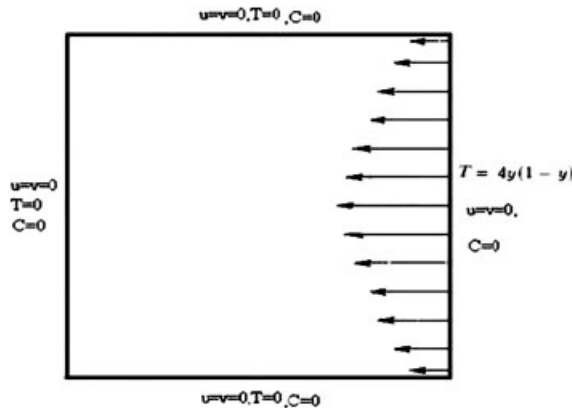


Figure 2. Physics model of the transport gas: $t=0$, i.e. $n=0$ initial values on boundary.

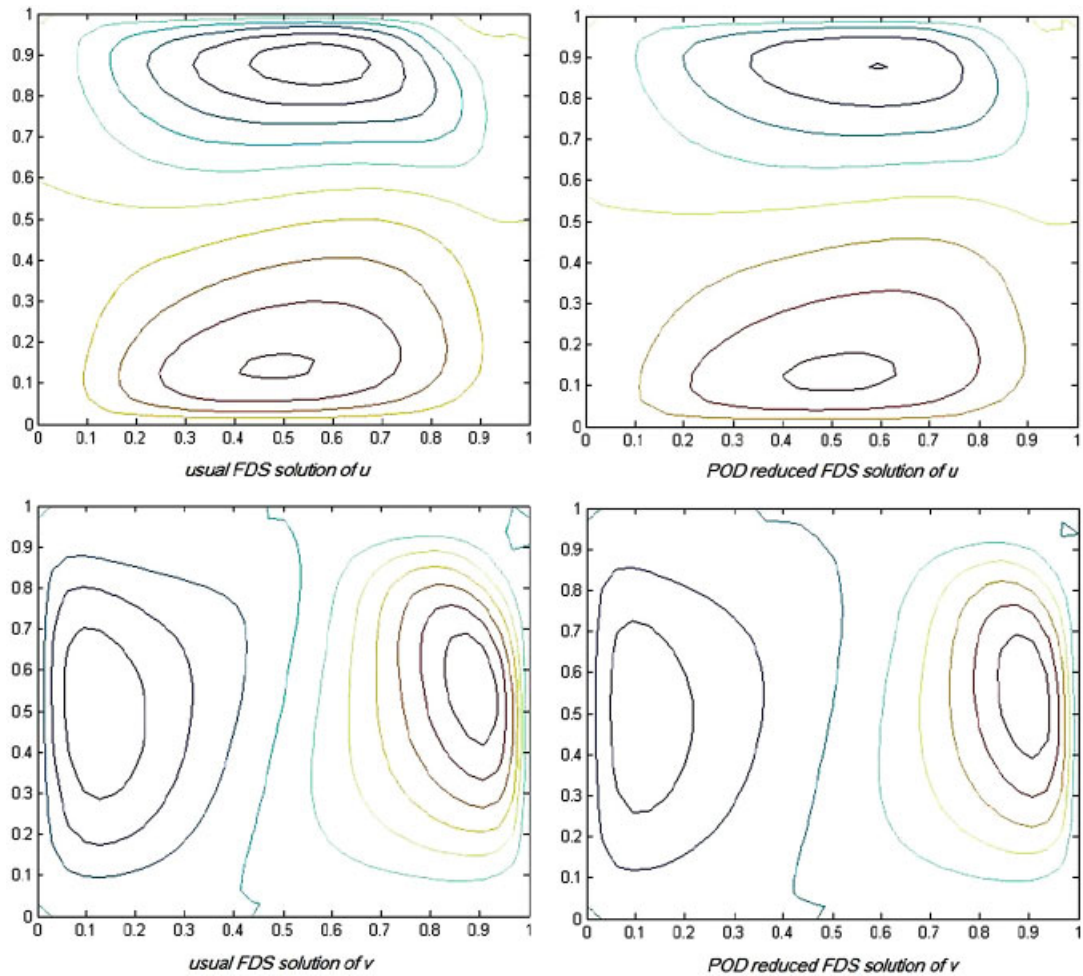


Figure 3. Velocity (u, v) stream line figure for usual FDS solution (on left-hand side figure) and when $M_u = M_v = 5$, solutions of the reduced FDS based on POD (on right-hand side figure).

optimal POD basis, respectively, and the solutions of usual FDS, i.e. (3), (7), (10), (15), (17), and (19) are depicted in Figures 3 and 4 on the left-hand side (because these figures are equal to solutions obtained with 20 bases, they are also known as the figures of full basis solution).

Figure 5 shows the errors between solutions obtained with different numbers of optimizing POD bases and solutions with full basis. Comparing the usual FDS with the POD reduced FDS by implementing 3000 times numerical simulation computations, we find that time-consuming calculations with usual FDS are 5 min, while those with the reduced optimizing FDS time-consuming expend only 3 s, i.e. the running time with usual FDS, i.e. (3), (7), (10), (15), (17), and (19) is almost as 100 multiples as that with the POD reduced FDS (34), (35), and (50) with POD optimal basis and the errors between the solutions are not more than 8×10^{-3} . Although what we have done here is a kind of recomputing what we have already computed by usual FDS, when it comes to actual problems, we may structure the snapshots and POD basis with interpolation or data assimilation by drawing samples from experiments, then solve directly the reduced optimizing FDS (35) with (34) and (50). Thus, the calculating time and memory demands in the computational process will be greatly saved. It is also shown that the POD approach in solving the CVD equations is very effective and the consistency of the numerical and theoretical results is demonstrated (Figure 6).

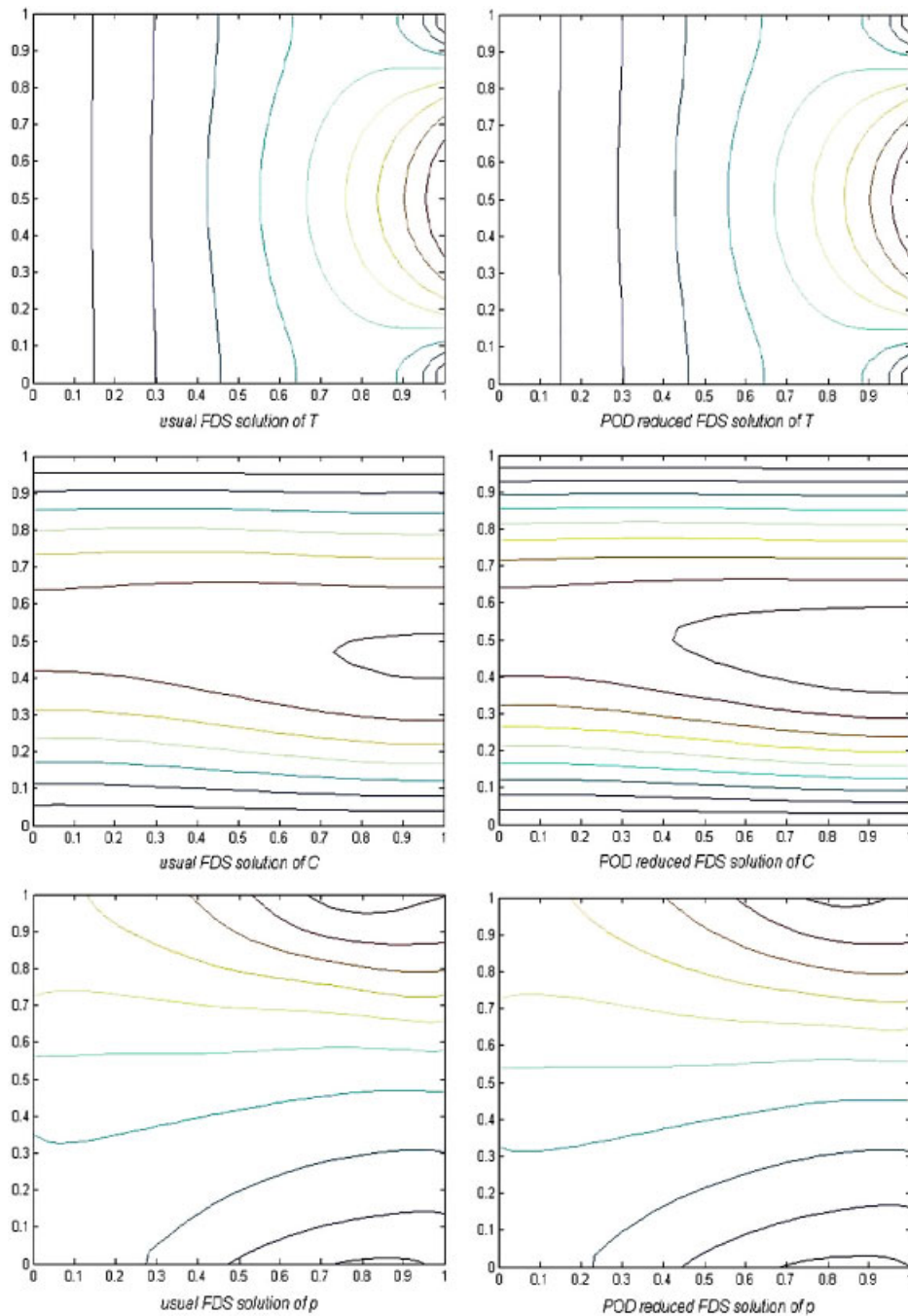


Figure 4. Temperature figure, mass figure, and pressure figure for usual FDS solution (on left-hand side figure) and when $M_p = M_T = M_C = 7$ solutions of the reduced FDS based on POD (on right-hand side figure).

6. CONCLUSIONS

In this paper, the SVD and POD methods are applied to derive a reduced optimizing FDS for CVD equations. First, ensembles of data are compiled from transient solutions obtained with usual FDS, and this process can be omitted in actual applications where the ensemble of snapshots can be obtained from physical system trajectories by drawing samples from experiments and

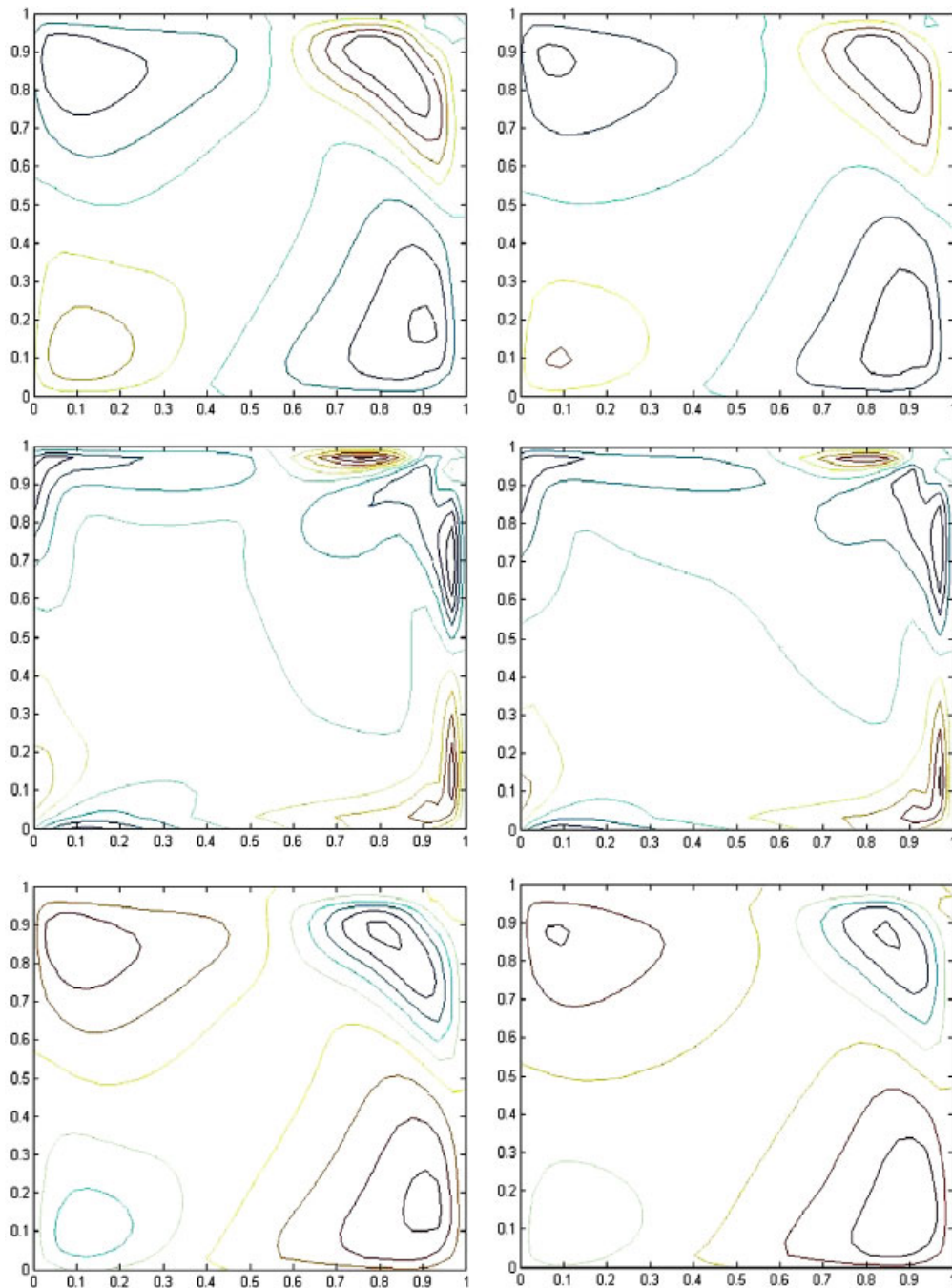


Figure 5. Stress tensor $\sigma = (\sigma_{ij})_{2 \times 2}$ for usual FDS solutions (on left-hand side figure) and when $M_u = M_v = 5$, POD reduced FDS solutions (on right-hand side figure). From top to bottom, figure of σ_{11} , figure of $\sigma_{12} (= \sigma_{21})$ and figure of σ_{22} .

interpolation (or data assimilation). Then we employ SVD to derive POD basis from the ensembles of data and substitute the unknowns of usual FDS with the linear combinations of POD basis to develop the reduced optimizing FDS for CVD equations, in which the much fewer basis make the POD reduced FDS optimal in a sense. Finally, we have analyzed the errors between the POD reduced FDS approximate solutions and the usual finite difference solutions, which are consistent with the theoretical results of error estimate and validating the feasibility and efficiency of our

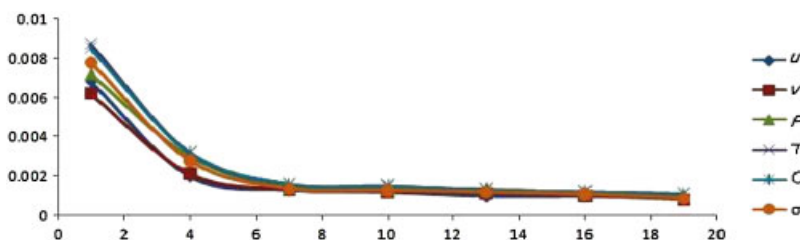


Figure 6. Error estimate.

reduced optimizing FDS at the same time. The theoretical and numerical results in this article also demonstrate that the POD method has extensive applications in complicated nonlinear PDEs, which makes it more feasible to be applied in some other complicated PDEs and some actual problems such as the atmosphere quality forecast system, the ocean fluid forecast system and so on.

ACKNOWLEDGEMENTS

Prof. I. M. Navon acknowledges support of NSF grant CCF-0635162.

REFERENCES

1. Chung T. *Computational Fluid Dynamics*. Cambridge University Press: Cambridge, U.K., 2002.
2. Xin X, Liu R, Jiang B. *Computational Fluid Dynamics*. National Defense Science Technology Press: Changsha, 1989 (in Chinese).
3. Quarteroni A, Valli A. *Numerical Approximation of Partial Differential Equations*. Springer: Berlin, Heidelberg, New York, 1997.
4. Fukunaga K. *Introduction to Statistical Recognition*. Academic Press: New York, 1990.
5. Jolliffe IT. *Principal Component Analysis*. Springer: Berlin, 2002.
6. Holmes P, Lumley JL, Berkooz G. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press: Cambridge, U.K., 1996.
7. Lumley JL. Coherent structures in turbulence. In *Transition and Turbulence*, Meyer RE (ed.). Academic Press: New York, 1981.
8. Aubry N, Holmes P, Lumley JL, Stone E. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Dynamics* 1988; **192**:115–173.
9. Sirovich L. Turbulence and the dynamics of coherent structures: part I–III. *Quarterly of Applied Mathematics* 1987; **45**(3):561–590.
10. Moin P, Moser RD. Characteristic-eddy decomposition of turbulence in channel. *Journal of Fluid Mechanics* 1989; **200**:417–509.
11. Rajaei M, Karlsson SKF, Sirovich L. Low dimensional description of free shear flow coherent structures and their dynamical behavior. *Journal of Fluid Mechanics* 1994; **258**:1401–1402.
12. Joslin RD, Gunzburger MD, Nicolaides RA, Erlebacher G, Hussaini MY. A self-contained automated methodology for optimal flow control validated for transition delay. *AIAA Journal* 1997; **35**:816–824.
13. Ly HV, Tran HT. Proper orthogonal decomposition for flow calculations and optimal control in a horizontal CVD reactor. *Quarterly of Applied Mathematics* 2002; **60**:631–656.
14. Rediniotis OK, Ko J, Yue X, Kurdila AJ. Synthetic jets, their reduced order modeling and applications to flow control. *AIAA Paper Number 99-1000, 37 Aerospace Sciences Meeting and Exhibit*, Reno, 1999.
15. Cao Y, Zhu J, Luo Z, Navon IM. Reduced order modeling of the upper tropical Pacific ocean model using proper orthogonal decomposition. *Computers and Mathematics with Applications* 2006; **52**:1373–1386.
16. Cao YH, Zhu J, Navon IM, Luo Z. A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. *International Journal for Numerical Methods in Fluids* 2007; **53**: 1571–1583.
17. Luo Z, Zhu J, Wang R, Navon IM. Proper orthogonal decomposition approach and error estimation of mixed finite element methods for the tropical Pacific Ocean reduced gravity model. *Computer Methods in Applied Mechanics and Engineering* 2007; **196**(41–44):4184–4195.
18. Luo Z, Chen J, Zhu J, Wang R, Navon IM. An optimizing reduced order FDS for the tropical Pacific Ocean reduced gravity model. *International Journal for Numerical Methods in Fluids* 2007; **55**(2):143–161.
19. Luo Z, Wang R, Zhu J. Finite difference scheme based on proper orthogonal decomposition for the non-stationary Navier–Stokes equations. *Science in China Series A: Mathematics* 2007; **50**(8):1186–1196.

20. Kunisch K, Volkwein S. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik* 2001; **90**:117–148.
21. Kunisch K, Volkwein S. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis* 2002; **40**:492–515.
22. Kunisch K, Volkwein S. Control of Burgers' equation by a reduced order approach using proper orthogonal decomposition. *Journal of Optimization Theory and Applications* 1999; **102**:345–371.
23. Ahlman D, Södelund F, Jackson J, Kurdila A, Shyy W. Proper orthogonal decomposition for time-dependent lid-driven cavity flows. *Numerical Heat Transfer Part B—Fundamentals* 2002; **42**:285–306.
24. Luo Z. *Mixed Finite Element Methods and Applications*. Chinese Science Press: Beijing, 2006.