A reduced finite volume element formulation and numerical simulations based on POD for parabolic problems

Zhendong Luo\textsuperscript{a,*}, Zhenghui Xie\textsuperscript{b}, Yueqiang Shang\textsuperscript{c}, Jing Chen\textsuperscript{d,*}

\textsuperscript{a} School of Mathematics and Physics, North China Electric Power University, Beijing 102206, China
\textsuperscript{b} Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China
\textsuperscript{c} School of Mathematics and Computer Science, Guizhou Normal University, Guiyang 550001, China
\textsuperscript{d} College of Science, China Agricultural University, Beijing 100083, China

\begin{abstract}
A proper orthogonal decomposition (POD) method is applied to a usual finite volume element (FVE) formulation for parabolic equations such that it is reduced to a POD FVE formulation with lower dimensions and high enough accuracy. The error estimates between the reduced POD FVE solution and the usual FVE solution are analyzed. It is shown by numerical examples that the results of numerical computation are consistent with theoretical conclusions. Moreover, it is also shown that the reduced POD FVE formulation based on POD method is both feasible and highly efficient.
\end{abstract}

\section{1. Introduction}

Many physical phenomena of the natural environment, engineering equipments, and living organisms, such as the proliferation of gas, the infiltration of liquid, the conduction of heat, and the spread of impurities in semiconductor materials, are described with parabolic equations. It is not easy to find their exact solutions for practical engineering problems. On the contrary, it is an efficient approach for finding their numerical solutions. The finite volume element (FVE) method (see [1–3]), called box method (see [4]) earlier, discretize the integral form of conservation law of differential equation by choosing linear or bilinear finite element space as trial space and also called generalized difference method (see [5,6]) in China, can keep the conservation law of mass or energy. It has higher accuracy than finite difference method and keeps the same accuracy as finite element method but is simpler and more convenient than the finite element method. So it is regarded as one of the most effective numerical methods and its theory has been established very well and widely applied to finding numerical solutions of different types of partial differential equations, for example, second order elliptic equations and parabolic equations (see [1–5]). However, some usual FVE formulations for parabolic equations include too many degrees of freedom. Thus, an important problem is how to alleviate the computational load and save time for calculations and resource demands in the computational process in a way that guarantees a sufficiently accurate numerical solution.
A proper orthogonal decomposition (POD) is a technique offering adequate approximation for representing fluid flow with reduced number of degrees of freedom, i.e., with lower-dimensional models to alleviate the computational load and memory requirements (see [7]), which is also known as Karhunen–Loève expansions in signal analysis and pattern recognition (see [8]), or principal component analysis in statistics (see [9]), or the method of empirical orthogonal functions in geophysical fluid dynamics or meteorology (also see [9]). The POD method mainly provides a useful tool for efficiently approximating a large amount of data. The method essentially provides an orthogonal basis for representing the given data in a certain least squares optimal sense, that is, it provides a way to find optimal lower-dimensional approximations of the given data. Combined with the Galerkin projection procedure, POD provides a powerful method for generating lower-dimensional models of dynamical systems that have a very large or even infinite-dimensional phase space.

Though POD method has been widely applied in computations of statistics and fluid dynamics (see [7, 9–18]), it is mainly used to perform principal component analysis and search the main behavior of a dynamic system. More recently, some Galerkin POD methods for parabolic problems and a general equation in fluid dynamics are presented (see [19, 20]), the singular value decomposition approach combined with POD technique is used to treat the Burgers equation (see [21]) and the cavity flow problem (see [22]), some reduced order finite difference models and finite element (or mixed finite element) formulations and error estimates for the upper tropical pacific ocean model, parabolic problems, and the non-stationary Navier–Stokes equations based on POD are presented (see [23–27]). Moreover, there are related works available for POD applications in optimization, for instance, adaptive POD (see [28]), Trust-Region-POD (see [29]), OS-POD (see [30]), POD a posterior error estimates (see [31]).

However, to the best of our knowledge, there are no published results to address the POD approximate solutions of FVE formulation for parabolic problems. In this paper, the POD technique is used to reduce the FVE formulation for parabolic problems. The errors between the reduced POD FVE solution and the usual FVE solution are analyzed. It is shown by numerical examples that the results of numerical computation are consistent with theoretical conclusions. Moreover, it is also shown that the POD FVE formulation is feasible and efficient for solving parabolic problems.

The paper is organized as follows. Section 2 is to derive the usual FVE formulation for parabolic problems and to generate snapshots from transient solutions computed from the equation system derived by the usual FVE formulation. In Section 3, the optimal orthonormal bases are reconstructed from the elements of the snapshots with POD and a reduced FVE formulation with lower-dimensional number based on POD for parabolic problems is developed. In Section 4, the error estimates between usual FVE solutions and POD solutions of the reduced FVE formulation are derived. In Section 5, some numerical examples are presented illustrating that the errors between the reduced FVE approximate solutions and the usual FVE solutions are consistent with previously obtained theoretical results, thus validating the feasibility and efficiency of the POD formulation. Section 6 provides main conclusions and future tentative ideas.

2. Usual FVE formulation for parabolic problems

Let $\Omega \subset \mathbb{R}^2$ be a bounded, connected and polygonal domain with a Lipschitz continuous boundary. Consider the following parabolic problems.

**Problem (I).** Find $u$ such that

$$
\begin{align*}
&u_t - \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \quad t \in (0, T], \\
&u(x, y, 0) = u_0(x, y), \quad (x, y) \in \Omega, \\
&u(x, y, t) = 0, \quad (x, y) \in \partial \Omega, \quad t \in (0, T],
\end{align*}
$$

(2.1)

where $u$ represents the unknown function, $u_t = \frac{\partial u}{\partial t}$, $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, source term $f(x, y, t) \in L^2(\Omega)$ and initial condition $u_0(x, y) \in W^{1,p}(\Omega)$ ($p > 1$) are two given functions.

The Sobolev spaces used in this context are standard (see [32]). Let $U = H^1_0(\Omega)$. Then, the variational formulation for **Problem (I)** can be written as follows.

**Problem (II).** Find $u \in H^1((0, T); L^2(\Omega)) \cap L^2((0, T); U)$ such that

$$
\begin{align*}
&(u_t, v) + a(u, v) = (f, v), \quad \forall v \in U, \\
&u(x, y, 0) = u_0(x, y), \quad (x, y) \in \Omega,
\end{align*}
$$

(2.2)

where $a(u, v) = (\nabla u, \nabla v)$, $(\cdot, \cdot)$ denotes the inner product in $L^2(\Omega)$.

In order to get the numerical solution of FVE for **Problem (II)**, it is necessary to introduce a FVE approximation for the spatial variable of **Problem (II)** and to approximate the time derivative with difference quotient.

Firstly, let $\{\mathcal{T}_h\}$ be a uniformly regular family of triangulation of $\bar{\Omega}$ (see [33–36]), where $h$ is the maximum length of all the sides.

The following definition will be used throughout this paper.
**Fig. 1.** Barycenter dual decomposition.

**Fig. 2.** Circumcenter dual decomposition.

**Definition 1.** We use $PQ$ to denote the line segment with end-points $P$ and $Q$ on the plane, which may bear a direction from $P$ to $Q$ when e.g. it is a path of line integral. We also identify $PQ$ with the corresponding vector of $\mathbb{R}^2$ in the usual sense. Its length is denoted by $|PQ|$.

Next, we construct a dual decomposition $\mathcal{I}^*$ related to $\mathcal{I}_h$. Let $P_0$ be a node of a triangle, $P_i$ ($i = 1, 2, \ldots, 6$) the adjacent nodes of $P_0$, and $M_i$ the midpoint of $P_0P_i$ (cf. Fig. 1). Choose a point $Q_i$ in an element $\triangle P_0P_iP_{i+1}$ ($P_7 = P_1$) and connect successively $M_1, Q_1, \ldots, M_6, Q_6, M_1$ to form a polygonal region $K^*_P$, called a dual element. The modification of the definition is obvious when $P_0$ is on the boundary. All the dual elements constitute a new decomposition, called a dual decomposition (or a dual grid). $Q_i$ is called a node of the dual decomposition. The following two dual decompositions are most important for the triangulation $\mathcal{I}_h$:

1. **Barycenter dual decomposition.** Take the barycenter $Q_0$ of the triangle $\triangle P_0P_iP_{i+1}$ as the node of the dual decomposition, as shown in Fig. 1.

2. **Circumcenter dual decomposition.** Assume that the interior angles of any element of $\mathcal{I}_h$ are not greater than $90^\circ$. Then, take the circumcenter $Q_i$ of the element $\triangle P_0P_iP_{i+1}$ as the node of the dual decomposition. Now $Q_iQ_{i+1}$ is the perpendicular bisector of $P_0P_{i+1}$ (see Fig. 2).

In what follows we denote by $\Omega_h$ the set of the nodes of the decomposition $\mathcal{I}_h$, $\partial \Omega = \partial \Omega_h \setminus \partial \Omega$ the set of the interior nodes, and $\Omega^*_h$ the set of the nodes of the dual decomposition $\mathcal{I}^*_h$. For $Q_i \in \Omega^*_h$, $K_Q$ denotes the triangular element containing $Q_i$. Let $S_{Q_0}$ (or $S_Q$) and $S^*_{P_0}$ be the areas of the triangular element $K_Q$ and the dual element $K^*_P$, respectively. It is easy to check that if $\mathcal{I}_h$ and $\mathcal{I}^*_h$ are quasi-uniform (cf. [33–35]), then there exist constants $c_1, c_2$, and $c_3 > 0$ independent of $h$ such that

\begin{align}
  c_1 h^2 \leq S_Q \leq h^2, & \quad Q_i \in \Omega^*_h, \quad (2.3) \\
  c_2 h^2 \leq S^*_{P_0} \leq c_3 h^2, & \quad P_0 \in \Omega_h. \quad (2.4)
\end{align}

It can be easily shown that (2.3) is actually a necessary and sufficient condition for the triangulation $\mathcal{I}_h$ to be quasi-uniform. Besides, for barycenter and circumcenter dual decompositions, (2.4) can be deduced from (2.3). In what follows we always assume that the decomposition is quasi-uniform.
The trial function space $U_h$ chosen as the linear element space related to $\mathcal{S}_h$ is the set of all the functions $u_h$ satisfying the following conditions:

(i) $u_h \in C(\overline{\Omega})$, $u_h|_{\partial \Omega} = 0$;

(ii) $u_h|_K \in \mathcal{P}_1$, namely $u_h$ is a linear function of $x$ and $y$ on each triangular element $K \in \mathcal{S}_h$, determined only by its values on the three vertices. It is obvious that $U_h \subset U = H^1_0(\Omega)$. Let $K = \triangle P_i P_j P_k$ be any triangular element and $P(x, y)$ a point in the element (cf. Fig. 3).

Introduce the area coordinates $(\lambda_i, \lambda_j, \lambda_k)$ as follows:

\[
\lambda_i = \frac{S_i}{S} = \frac{1}{2S} \begin{vmatrix} 1 & x & y \\ 1 & x_i & y_i \\ 1 & x_k & y_k \end{vmatrix}, \quad \lambda_j = \frac{S_j}{S} = \frac{1}{2S} \begin{vmatrix} 1 & x & y \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix}, \quad \lambda_k = \frac{S_k}{S} = \frac{1}{2S} \begin{vmatrix} 1 & x & y \\ 1 & x_i & y_i \\ 1 & x_j & y_j \end{vmatrix},
\]

(2.5)

where $S_i, S_j, S_k$, and $S$ are the areas of $\triangle P_i P_k$, $\triangle P_i P_j$, $\triangle P_j P_k$, and $\triangle P_j P_k$, respectively. The mapping (2.5) maps $\triangle P_i P_j P_k$ onto a reference element $\hat{K}$ with vertices $\hat{P}_i(0, 0), \hat{P}_j(1, 0)$, and $\hat{P}_k(0, 1)$ on the $(\lambda_i, \lambda_k)$ plane (cf. Fig. 4).

The area coordinates and the orthogonal coordinates have the following relationship:

\[
\begin{align*}
x &= x_i \lambda_i + x_j \lambda_j + x_k \lambda_k, \\
y &= y_i \lambda_i + y_j \lambda_j + y_k \lambda_k, \\
1 &= \lambda_i + \lambda_j + \lambda_k.
\end{align*}
\]

(2.6)

It is easy to deduce that on the element $K$

\[
\begin{align*}
\frac{\partial u_h}{\partial x} &= \frac{1}{2S} \left[ \frac{\partial u_h}{\partial \lambda_i} (y_k - y_i) + \frac{\partial u_h}{\partial \lambda_j} (y_j - y_i) + \frac{\partial u_h}{\partial \lambda_k} (y_k - y_j) \right] = \frac{1}{2S} \left[ u_i (y_j - y_k) + u_j (y_k - y_i) + u_k (y_i - y_j) \right]; \\
\frac{\partial u_h}{\partial y} &= \frac{1}{2S} \left[ \frac{\partial u_h}{\partial \lambda_i} (x_k - x_i) + \frac{\partial u_h}{\partial \lambda_j} (x_j - x_i) + \frac{\partial u_h}{\partial \lambda_k} (x_k - x_j) \right] = \frac{1}{2S} \left[ u_i (x_j - x_k) + u_j (x_k - x_i) + u_k (x_i - x_j) \right],
\end{align*}
\]

(2.7)

where and in what follows, if there is no danger of confusion, we write in short $u_i = u_h(x_i, y_i)$, etc. For $u \in U = H^1_0(\Omega)$, let $\Pi_h u$ be the interpolation projection of $u$ onto the trial function space $U_h$. By the interpolation theory of Sobolev spaces (see [33–35]), we have, if $u \in H^2(\Omega)$, that

\[
|u - \Pi_h u|_m \leq Ch^2 |u|_2, \quad m = 0, 1.
\]

(2.8)
where C in this context indicates a positive constant which is possibly different at different occurrences, being independent of the spatial h and temporal mesh sizes.

The test space $V_h$ is chosen as the piecewise constant function space with respect to $\mathcal{T}^h$, spanned by the following basis functions: for any point $P_0 \in \hat{\Omega}_h$,

$$
\phi_{P_0}(P) = \begin{cases} 
1, & P \in K^*_P, \\
0, & \text{elsewhere}.
\end{cases}
$$

(2.9)

For any $v_h \in V_h$,

$$
v_h = \sum_{P_0 \in \hat{\Omega}_h} v_h(P_0) \phi_{P_0}.
$$

(2.10)

For $w \in U$, let $\Pi_h^* w$ be the interpolation projection of $w$ onto the test space $V_h$, i.e.,

$$
\Pi_h^* w = \sum_{P_0 \in \hat{\Omega}_h} w(P_0) \phi_{P_0}.
$$

(2.11)

By the interpolation theory we have

$$
\| w - \Pi_h^* w \|_0 \leq C h \| w \|_1.
$$

(2.12)

Though the trial function space $U_h$ of FVE methods satisfies $U_h \subset U$ like finite element methods, the test space $V_h \not\subset U_h$. As in the case of nonconforming finite element methods, this is due to the loss of continuity of the functions in $V_h$ on the boundary of two neighboring elements. So the bilinear form $a(u, v)$ must be revised accordingly. For nonconforming finite element methods, the idea is to write the integral on the whole region as a sum of the integrals on every element $K$, so $a(u, v)$ of (2.2) is rewritten as

$$
a(u, v) = \sum_k \int_K \left( \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \right) dx dy.
$$

(2.13)

Now $a(u, v)$ is well defined on $U_h \times V_h$. For the FVE methods, i.e., generalized difference methods, we place a dual grid and interpret (2.13) in the sense of generalized functions, i.e., $\delta$ functions on the boundary of neighboring dual elements. Or equivalently, we take $a(u, v)$ as the bilinear form resulting from the piecewise integrations by parts on the dual elements $K^*$:

$$
\int_\Omega \Delta u \cdot v dx dy = \sum_{K^*} \int_{K^*} \Delta u \cdot v dx dy.
$$

(2.14)

So we have

$$
a(u, v) = \sum_k \int_{K^*} \left( \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \right) dx dy - \sum_{K^*} \int_{\partial K^*} \left( \frac{\partial u}{\partial x} v dx - \frac{\partial u}{\partial y} v dy \right),
$$

(2.15)

where $\int_{\partial K^*}$ denotes the line integrals, with the counter-clockwise direction, on the boundary $\partial K^*$ of the dual element. Since $V_h$ is the piecewise constant function space with the characteristic functions of the dual elements $K^*$ as the basis functions, then Problem (I) becomes the integral interpolation method based on the integral conservation law (the balance equation)

$$
\int_{K^*} u_t dx dy - \int_{K^*} \Delta u dx dy = \int_{K^*} u_t dx dy - \int_{\partial K^*} \left( \frac{\partial u}{\partial x} dx - \frac{\partial u}{\partial y} dy \right) = \int_{K^*} f dx dy.
$$

(2.16)

Then the semi-discrete FVE (or generalized difference) scheme for (2.2) is rewritten as follows.

**Problem (II').** Find $u_h \in U_h$ such that, for $0 < t \leq T$,

$$
\begin{cases}
(u_{ht}, v_h) + a(u_h, v_h) = (f, v_h), & \forall v_h \in V_h, \\
u_h(x, y, 0) = \Pi_h^* u_0(x, y), & (x, y) \in \Omega. 
\end{cases}
$$

(2.17)

or equivalently

$$
\begin{cases}
(u_{ht}, \phi_{P_0}) + a(u_h, \phi_{P_0}) = (f, \phi_{P_0}), & \forall P_0 \in \hat{\Omega}_h, \\
u_h(x, y, 0) = \Pi_h^* u_0(x, y), & (x, y) \in \Omega,
\end{cases}
$$

(2.18)

where

$$
a(u_h, v_h) = \sum_{P_0 \in \hat{\Omega}_h} v_h(P_0)a(u_h, \phi_{P_0}),
$$

(2.19)

$$
a(u_h, \phi_{P_0}) = - \int_{\partial K^*_P} \frac{\partial u}{\partial x} dy + \int_{\partial K^*_P} \frac{\partial u}{\partial y} dx.
$$

(2.20)
So the FVE method, i.e., the generalized difference method is a significant generalization of the finite difference method. From [5] we have the following four lemmas.

**Lemma 1.** Set

\[
\|u_h\|_{0,h} \equiv \|\Pi_h^* u_h\|_0 = \left\{ \sum_{k \geq 0} u_h^2(P_0) |S_0| \right\}^{1/2} = \left\{ \frac{1}{3} \sum_{k \geq 0} [u_0^2(P_i) + u_0^2(P_j) + u_0^2(P_k)] |S_0| \right\}^{1/2},
\]

(2.21)

\[
|u_h|_{1,h} \equiv \left\{ \sum_{k \geq 0} \left[ \left( \frac{\partial u_0(Q)}{\partial x} \right)^2 + \left( \frac{\partial u_0(Q)}{\partial y} \right)^2 \right] |S_0| \right\}^{1/2},
\]

(2.22)

\[
|\tilde{u}_h|_{1,h} = \left[ |u_h|_{0,h}^2 + |u_h|_{1,h}^2 \right]^{1/2}.
\]

Then the pairs of norms \(| \cdot |_{1,h}\) and \(| \cdot |_{0,h}\) and \(| \cdot |_0\), and \(| \cdot |_{1,h}\) and \(| \cdot |_1\) are equivalent on \(U_h\), respectively.

**Lemma 2.** The bilinear form \(a(u_h, \Pi_h^* \tilde{u}_h)\) can be expressed as

\[a(u_h, \Pi_h^* \tilde{u}_h) = a_h(u_h, \Pi_h^* \tilde{u}_h) + b(u_h, \Pi_h^* \tilde{u}_h),\]

(2.24)

where the leading term

\[a_h(u_h, \Pi_h^* \tilde{u}_h) = \sum_{k \geq 0} \left[ \frac{\partial u_0(Q)}{\partial x} \frac{\partial \tilde{u}_0(Q)}{\partial x} + \frac{\partial u_0(Q)}{\partial y} \frac{\partial \tilde{u}_0(Q)}{\partial y} \right] |S_0|\]

(2.25)

is symmetric, bounded, and coercive (or positive definite), i.e.,

\[a_h(u_h, \Pi_h^* \tilde{u}_h) = a_h(\tilde{u}_h, \Pi_h^* u_h),\]

(2.26)

and there are two positive constants \(C_1\) and \(C_2\) such that

\[C_1 \|u_h\|_1^2 \leq a_h(u_h, \Pi_h^* u_h) \leq C_2 \|u_h\|_1^2,\]

(2.27)

and the remainder \(b(u_h, \Pi_h^* \tilde{u}_h) = a(u_h, \Pi_h^* \tilde{u}_h) - a_h(u_h, \Pi_h^* \tilde{u}_h)\) satisfies

\[b(u_h, \Pi_h^* \tilde{u}_h) \leq Ch^2 \|u_h\|_1 \|\tilde{u}_h\|_1, \quad \forall u_h, \tilde{u}_h \in U_h.\]

(2.28)

If we put \(|\tilde{u}_h|_1 = \left[ a_h(u_h, \Pi_h^* u_h) \right]^{1/2}\), then \(|\tilde{u}_h|_1\) is equivalent to \(\|u_h\|_1\) on \(U_h\). Further,

\[|a(u_h, \Pi_h^* \tilde{u}_h) - a_h(\tilde{u}_h, \Pi_h^* u_h)| \leq Ch \|u_h\|_1 \|\tilde{u}_h\|_1, \quad \forall u_h, \tilde{u}_h \in U_h.\]

(2.29)

**Lemma 3.** There exist positive constants \(h_0\), \(\alpha\), and \(M\) such that when \(0 < h \leq h_0\)

\[a(u_h, \Pi_h^* u_h) \geq \alpha \|u_h\|_1^2, \quad \forall u_h \in U_h,\]

(2.30)

\[|a(u_h, \Pi_h^* \tilde{u}_h)| \leq M \|u_h\|_1 \|\tilde{u}_h\|_1, \quad \forall u_h, \tilde{u}_h \in U_h.\]

(2.31)

**Lemma 4.** There holds the following statement:

\[(u_h, \Pi_h^* \tilde{u}_h) = (\tilde{u}_h, \Pi_h^* u_h), \quad \forall u_h, \tilde{u}_h \in U_h.\]

(2.32)

Set \(|u_h|_0 = (u_h, \Pi_h^* u_h)^{1/2}\), then \(|\cdot|_0\) is equivalent to \(|\cdot|_0\) on \(U_h\), i.e., there exist two positive constants \(C_3\) and \(C_4\) such that

\[C_3 \|u_h\|_0 \leq |\cdot|_0 \leq C_4 \|u_h\|_0, \quad \forall u_h \in U_h.\]

(2.33)

Let \(\tau\) denote the time step size, and \(t_n = nr\) (\(n = 0, 1, \ldots, N = T/\tau\)), \(u_0^n = u_0(t_n)\). If the differential quotient \(u_{0h}^n\) in the semi-discrete scheme, i.e., Problem (II)’ is approximated with the backward difference quotient \(\tilde{u}_0^n = (u_0^n - u_0^{n-1})/\tau\) at time \(t = t_n\), then the fully discrete FVE approximation scheme of Problem (II)’ is read as follows.

**Problem (III).** Find \(u_0^n \in U_h (1 \leq n \leq N)\) such that

\[
\begin{cases}
\bar{\partial}_t u_0^n + a(u_0^n, \tilde{u}_0^n) = f(t_n, \nu_h), & \forall \nu_h \in V_h, \\
\tilde{u}_0^n = u_0(x, y, 0) = \Pi_h^* u_0(x, y), & (x, y) \in \Omega,
\end{cases}
\]

(2.34)

or, is equivalently read as follows.
**Problem (IV).** Find $u^n_h \in U_h$ ($1 \leq n \leq N$) such that

$$\begin{cases}
(u^n_h, v_h) + \tau a(u^n_h, v_h) = (u^{n-1}_h, v_h) + \tau (f(t_n), v_h), & \forall v_h \in V_h, \\
u^n_h = u_h(x, y, 0) = \Pi^h u_0(x, y), & (x, z) \in \Omega.
\end{cases} \tag{2.35}$$

**Problem (III) or (IV) is referred to as a backward Euler FVE (or generalized difference) scheme.** By Lemmas 3 and 4, we have

$$a(u^n_h, \Pi^h u^n_h) + \frac{1}{\tau} (u^n_h, \Pi^h u^n_h) \geq \alpha \|u^n_h\|^2_1, \quad \forall u^n_h \in U_h. \tag{2.36}$$

This guarantees the existence and uniqueness of the solution $u^n_h$ for **Problem (III)** or (IV) from the Lax–Milgram Theorem. And the following estimates hold (see Theorems 5.2.1 and 5.2.2 in [5]).

**Theorem 5.** Let $u_0 \in W^{3,p}(\Omega), u \in H^1((0, T); W^{3,p}(\Omega)) \cap H^2((0, T); L^2(\Omega)) (p > 1)$ be the solution to **Problem (II)**, and $u^n_h$ the solution to the backward Euler FVE scheme, i.e., **Problem (III)**, respectively. Then, for $n = 1, 2, \ldots, N,$

$$\|u(t_n) - u^n_h\|_0 \leq C \left[ h^2 \left( \|u_0\|_{3,p} + \int_0^{t_n} \|u_t\|_{3,p} \, dt \right) + \tau \int_0^{t_n} \|u_{tt}\| \, dt \right], \tag{2.37}$$

$$\|u(t_n) - u^n_h\|_1 \leq C \left[ h \|u_0\|_2 + h \int_0^{t_n} \|u_t\|_2 \, dt + h \left( \int_0^{t_n} \|u_{tt}\|_2 \, dt \right)^{1/2} + \tau \left( \int_0^{t_n} \|u_{tt}\|_2 \, dt \right)^{1/2} \right]. \tag{2.38}$$

If the source term $f(x, y, t)$, initial condition $u_0(x, y)$, the triangulation parameter $h$, the time step increment $\tau$, and trial function space $U_h$ are given, by solving **Problem (IV),** we can obtain a group of solutions ensemble $\{u^n_h\}_{n=1}^N$ for **Problem (III)** or (IV). And then we choose $L$ (in general, $L \ll N$, for example, $L = 20, N = 200$) instantaneous solutions $u^n_h(x, y)$ ($1 \leq n_1 < n_2 < \cdots < n_L \leq N$) (which are useful and of interest for us) from $N$ instantaneous solutions $\{u^n_h(x, y)\}_{n=1}^N$ for **Problem (III) or (IV),** which are referred to as snapshots and introduced by Sirovich in [18].

**Remark 1.** When one computes actual problems, one may obtain the ensemble of snapshots from physical system trajectories by drawing samples from experiments and interpolation (or data assimilation), then reconstruct the POD optimal basis for the ensemble of snapshots by using the following POD method, and finally the trial function space $U_h$ is substituted for the subspace generated with POD basis in order to derive a reduced order dynamical system with numbers of lower dimension. Thus, the future change of physical phenomenon can be quickly simulated, which is a result of much importance for real-life applications.

### 3. Generation of POD basis and reduced FVE formulation based on POD technique for **Problem (IV)**

For $u^n_h(x, y)$ ($1 \leq n_1 < n_2 < \cdots < n_L \leq N$) in Section 2, let $U_i(x, y) = u^n_h(x, y)$ ($1 \leq i \leq L$) and

$$\mathcal{V} = \text{span}\{U_1, U_2, \ldots, U_L\}, \tag{3.1}$$

and refer to $\mathcal{V}$ as the space generated by the snapshots $\{U_i\}_{i=1}^L$ at least one of which is assumed to be non-zero. Let $\{\psi_j\}_{j=1}^d$ denote an orthonormal basis of $\mathcal{V}$ with $l = \dim \mathcal{V}$. Then each member of the ensemble can be expressed as

$$U_i = \sum_{j=1}^d (U_i, \psi_j)_\Omega \psi_j, \quad i = 1, 2, \ldots, L, \tag{3.2}$$

where $(U_i, \psi_j)_\Omega = (\nabla u^n_h, \nabla \psi_j)$.

**Definition 2.** The method of POD consists in finding the orthonormal basis $\psi_i$ ($i = 1, 2, \ldots, L$) such that for every $d$ ($1 \leq d \leq l$) the mean square error between the elements $U_i$ ($1 \leq i \leq L$) and the corresponding $d$th partial sum of (3.2) is minimized on average

$$\min_{\{\psi_i\}_{i=1}^d} \frac{1}{L} \sum_{i=1}^L \left\| U_i - \sum_{j=1}^d (U_i, \psi_j)_\Omega \psi_j \right\|_\Omega^2 \tag{3.3}$$

subject to

$$(\psi_i, \psi_j)_\Omega = \delta_{ij}, \quad 1 \leq i, j \leq d, \quad 1 \leq i \leq L, \tag{3.4}$$

where $\|U_i\|_\Omega^2 = \|\nabla u^n_h\|_0^2$. A solution $\{\psi_j\}_{j=1}^d$ of (3.3) and (3.4) is known as a POD basis of rank $d$. 
By (3.2) and orthonormality of $\psi_j$, we can rewrite (3.3) as follows.

\[
\frac{1}{L} \sum_{i=1}^{l} \left\| U_i - \sum_{j=1}^{d} (U_i, \psi_j) U_i \psi_j \right\|_U^2 = \frac{1}{l} \sum_{i=1}^{l} \left\| \frac{1}{l} \sum_{j=d+1}^{l} (U_i, \psi_j) U_i \psi_j \right\|_U^2 \leq \left\| \sum_{j=d+1}^{l} (U_i, \psi_j) U_i \psi_j \right\|_U. \tag{3.5}
\]

Thus, in order to assure (3.5) minimum, it is equivalent to finding orthonormal basis $\psi_j$ ($j = 1, 2, \ldots, L$) such that

\[
\max_{\psi_j} \left\{ \left\| \frac{1}{l} \sum_{j=1}^{l} (U_i, \psi_j)^2 U_i \right\| \right\} \text{ subject to } (\psi_i, \psi_j)_U = \delta_{ij}, \quad 1 \leq i \leq d, \quad 1 \leq j \leq l. \tag{3.6}
\]

In other words, (3.3) and (3.4) are equivalent to looking for a function $\psi$, or the so-called POD basis element, such that most resembles $(U_l(x))_i^{d+1}$ in mean that it maximizes

\[
\frac{1}{L} \sum_{i=1}^{l} \left| (U_i, \psi)_U \right|^2 \text{ subject to } (\psi, \psi)_U = \| \nabla \psi \|_0^2 = 1. \tag{3.8}
\]

We cite the idea of snapshots introduced by Sirovich in [18] and choose a special class of trial functions for $\psi$ to be of the form:

\[
\psi = \sum_{i=1}^{l} a_i U_i, \tag{3.9}
\]

where the coefficients $a_i$ are to be determined so that $\psi$ given by expression (3.9) provides a maximum for (3.8). Thus, (3.8) is equivalent to the eigenvalue problem

\[
A \psi = \lambda \psi, \tag{3.10}
\]

where $A = (A_{ik})_{1 \times l}$ and

\[
A_{ik} = \frac{1}{L} \int L \nabla U_i(x, y) \cdot \nabla U_k(x, y) \, dx \, dy, \quad \psi = (a_1, a_2, \ldots, a_l)^T, \tag{3.11}
\]

and $\lambda$ depends on $\text{h}$ and $\tau$ due to $V$ depending on them. Since the matrix $A$ is a nonnegative Hermitian matrix which has rank $l$, it has a complete set of orthonormal eigenvectors

\[
\psi_1 = \frac{1}{\sqrt{\lambda_1}} \sum_{i=1}^{l} a_i^1 U_i, \tag{3.13}
\]

where $a_i^1$ are the elements of the eigenvector $\psi_1$ corresponding to the largest eigenvalue $\lambda_1$. The remaining POD basis elements $\psi_l$ ($l = 2, 3, \ldots, L$) are obtained by using the elements of other eigenvectors $\psi_i$ ($i = 2, 3, \ldots, L$), i.e.,

\[
\psi_i = \frac{1}{\sqrt{\lambda_i}} \sum_{k=1}^{l} a_i^k U_k. \tag{3.14}
\]

Moreover, the POD basis $\{\psi_1, \psi_2, \ldots, \psi_l\}$ forms an orthonormal set and holds the following results (see [18,21,22]).

**Proposition 6.** Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l > 0$ denote the positive eigenvalues of $A$ and $\psi_1, \psi_2, \ldots, \psi_l$ the associated orthonormal eigenvectors. Then a POD basis of rank $d \leq l$ is given by

\[
\psi_j = \frac{1}{\sqrt{\lambda_j}} \sum_{i=1}^{l} (\psi_j^i) U_i, \quad 1 \leq j \leq d \leq l, \tag{3.15}
\]

where $(\psi_j^i)$ denotes the $j$th component of the eigenvector $\psi_i$. Furthermore, the following error formula holds

\[
\frac{1}{L} \sum_{i=1}^{l} \left\| U_i - \sum_{j=1}^{d} (U_i, \psi_j) U_i \psi_j \right\|_U^2 = \sum_{j=d+1}^{l} \lambda_j. \tag{3.16}
\]

Let $U_d = \text{span} \{\psi_1, \psi_2, \ldots, \psi_d\}$. Define the Ritz projection $P^h: U \rightarrow U_h$ (if $P_h$ is restricted to Ritz projection from $U_h$ to $U_d$, it is written as $P^h$) such that $P^h|_{U_h} = P^h: U_h \rightarrow U^d$ and $P^h: U \setminus U_h \rightarrow U_h \setminus U^d$ denoted by

\[
(\nabla P^h u, \nabla w_h) = (\nabla u, \nabla w_h), \quad \forall w_h \in U_h. \tag{3.17}
\]
where $u \in U$. Due to (3.17) the linear operators $P^h$ are well defined and bounded
\[ \|\nabla (P^h u)\|_0 \leq \|\nabla u\|_0, \quad \forall u \in U. \] (3.18)

**Lemma 7.** For every $d$ ($1 \leq d \leq l$), the projection operators $P^d$ satisfy
\[ \frac{1}{L} \sum_{i=1}^{d} \|\nabla (u^h_i - P^d u^h_i)\|_0^2 \leq \sum_{j=d+1}^{l} \lambda_j, \] (3.19)
\[ \frac{1}{L} \sum_{i=1}^{d} \|u^h_i - P^d u^h_i\|_0^2 \leq Ch^2 \sum_{j=d+1}^{l} \lambda_j, \] (3.20)
where $u^h_i \in V$ is the solution of **Problem (IV)**.

**Proof.** The proof of inequality (3.19) has been given in [19,20]. It is only necessary to prove inequality (3.20). We consider the following variational problem
\[ \langle \nabla w, \nabla \varphi \rangle = (u - P^h u, \varphi), \quad \forall \varphi \in U. \] (3.21)
Since $u - P^h u \in U$, Eq. (3.21) has a unique solution $w \in H^1_0(\Omega) \cap H^2(\Omega)$ such that $\|w\|_2 \leq C\|u - P^h u\|_0$. Taking $\varphi = u - P^h u$ in (3.21) and using (3.18), we get that
\[ \|u - P^h u\|_0^2 = \langle \nabla w, (u - P^h u) \rangle = \langle (\nabla w - \nabla u), (u - P^h u) \rangle \leq \|\nabla (w - u)\|_0 \|\nabla (u - P^d u)\|_0, \quad \forall u^h \in U^h. \] (3.22)
Taking $w^h = \Pi^1 h w$ as interpolation function of $w$ in $U^h$ and using (2.8) or interpolation theory (see [33-35]) and (3.22), we obtain that
\[ \|u - P^h u\|_0^2 \leq Ch\|w\|_2 \|\nabla (u - P^h u)\|_0 \leq Ch\|u - P^h u\|_0 \|\nabla (u - P^h u)\|_0. \] (3.23)
Therefore, we get
\[ \|u - P^h u\|_0 \leq Ch\|\nabla (u - P^h u)\|_0. \] (3.24)
Thus, if $u = u^h_i$, $P^h$ is restricted to Ritz projection from $U^h$ to $U^d$ such that $P^h|_{U^h} = P^d : U^h \to U^d$, i.e., $P^h u^h_i = P^d u^h_i \in U^d$, then from (3.24) and (3.19) we derive (3.20), which completes the proof of Lemma 7.  

Thus, by using $U^d$, we can obtain the reduced formulation based on POD for **Problem (IV)** as follows.

**Problem (V).** Find $u^d_n \in U^d$ ($1 \leq n \leq N$) such that
\[ \begin{cases} (u^d_n, \Pi^1 h w_d) + \tau a(u^d_n, \Pi^1 h w_d) = (u^{d-1}_n, \Pi^1 h w_d) + \tau (\phi(t_n), \Pi^1 h w_d), \quad \forall w_d \in U^d, \\
\end{cases} \] (3.25)
where $\Pi^1 h$ is defined by (2.11).

**Remark 2.** If $\mathcal{T}_h$ is a uniformly regular triangulation and $U^h$ is the space of piecewise linear function, the total degrees of freedom for **Problem (IV)**, i.e., the number of unknown quantities is $N^h$ (where $N^h$ is the number of vertices of triangles in $\mathcal{T}_h$; see [33,34]), while the number of total degrees of freedom for **Problem (V)** is $d \ll l \ll N$. For scientific engineering problems, the number of vertices of triangles in $\mathcal{T}_h$ is more than tens of thousands or even more than a hundred million, while $d$ is only the number of few maximal eigenvalues which is chosen $L$ snapshots from the $N$ snapshots so that it is very very small (for example, in Section 5, $d = 6$, while $N^h = 100 \times 100 = 10000$). Therefore, **Problem (V)** is a reduced FVE formulation based on POD method for **Problem (IV)**. Moreover, since the development and change of a large number of future nature phenomena are closely related to previous results, for example, weather change, biology anagenesis, and so on, one may truly capture laws of change of natural phenomena by using existing results as snapshots to structure POD basis and solving corresponding PDEs. Therefore, the POD methods provide useful and important applications.

4. Error estimates of solution for **Problem (V)**

In this section, we refer to the usual FVE method to derive the error estimates for **Problem (V)**.

We have the following main result for **Problem (V)**.

**Theorem 8.** Under hypotheses of **Theorem 5**, **Problem (V)** has a unique group of solutions $u^d_n \in X^d$ such that
\[ \|u^d_n\|_0^2 + \tau C_1 \sum_{i=1}^{n} \|u^d_i\|_1^2 \leq C\|u_0\|_1^2 + \tau C_1^{-1} \sum_{i=1}^{n} \|f(t_i)\|_1^2, \quad 1 \leq n \leq N. \] (4.1)
And if \( \tau = O(h) \), \( L^{3/2} = O(N) \), and snapshots are taken at uniform intervals, then the following error estimates hold

\[
\| u^n_h - u^n_0 \|_0 + \tau \sum_{i=1}^n \| \nabla (u^n_h - u^n_0) \|_0 \leq C \tau + C \left( \sum_{i=d+1}^l \lambda_j \right)^{1/2}, \quad 1 \leq n \leq N. \tag{4.2}
\]

**Proof.** Using the same approach as proof of existence and uniqueness of solution of Problem (IV), based on (2.36), we can prove that Problem (V) has a unique group of solutions \( u^n_0 \in U^d \).

Taking \( w_d = u^n_d \) in Problem (V) from Lemmas 2–4 yields

\[
\| u^n_0 \|_0^2 + \tau C_1 \| u^n_0 \|_1^2 \leq \| u^n_{d-1} \|_0 \| u^n_0 \|_0 + \tau \| f(t_0) \|_1 \| u^n_0 \|_1 \\
\leq \frac{1}{2} \left[ \| u^n_{d-1} \|_0^2 + \| u^n_0 \|_0^2 + \tau C_1 \| f(t_0) \|_1 + \tau C_1 \| u^n_0 \|_1^2 \right]. \tag{4.3}
\]

i.e.,

\[
\| u^n_0 \|_0^2 + \tau C_1 \| u^n_0 \|_1^2 \leq \| u^n_{d-1} \|_0^2 + \tau C_1 \| f(t_0) \|_1. \tag{4.4}
\]

By summing (4.4) from 1 to \( n \), we obtain that

\[
\| u^n_0 \|_0^2 + \tau C_1 \sum_{i=1}^n \| u^n_i \|_1^2 \leq \| \Pi_h u^n_0 \|_0^2 + \tau C_1 \sum_{i=1}^n \| f(t_i) \|_1^2 \leq C \| u_0 \|_1^2 + \tau C_1 \sum_{i=1}^n \| f(t_i) \|_1^2. \tag{4.5}
\]

Since \( U^d \subset U_h \), subtracting Problem (V) from Problem (IV) taking \( w_h = \Pi_h w_d \in V_h \) can yield that

\[
(u^n_0 - u^n_0, \Pi_h w_d) + \tau a(u^n_0 - u^n_0, \Pi_h w_d) = (u^n_{d-1} - u^n_{d-1}, \Pi_h w_d), \quad \forall w_d \in U^d. \tag{4.6}
\]

Note that we have \( \| P^d u^n_0 - u^n_0 \|_0 \leq Ch \| \nabla (P^d u^n_0 - u^n_0) \|_0 \) from (3.24) and \( \| P^d u^n_0 - u^n_0 \|_0 \leq C \| \nabla (P^d u^n_0 - u^n_0) \|_0 \). Thus, if \( \tau = O(h) \), we obtain from (4.6) and by Lemmas 2–4 that

\[
\| P^n u^n_0 - u^n_0 \|_0 + \tau \| P^n u^n_0 - u^n_0 \|_0 \leq \| P^n u^n_0 - u^n_0 \|_0 + \tau \| P^n u^n_0 - u^n_0 \|_0 \leq C \| u_0 \|_1 + \tau C \| u_0 \|_1 \\
\leq C (\| u_0 \|_1 + \tau C_1 \| f(t_0) \|_1). \tag{4.7}
\]

Furthermore, we get that

\[
\| P^n u^n_0 - u^n_0 \|_1 + \tau \| P^n u^n_0 - u^n_0 \|_1 \leq C (\| P^n u^n_0 - u^n_0 \|_1 + \tau C_1 \| f(t_0) \|_1). \tag{4.8}
\]

Summing (4.8) for 1, 2, \ldots, \( n \) and using Lemmas 2–4 yield

\[
\| P^n u^n_0 - u^n_0 \|_1 + \tau \sum_{i=1}^n \| P^n u^n_0 - u^n_0 \|_1 \leq C \sum_{i=1}^n \| \nabla (u^n_0 - P^n u^n_0) \|_0. \tag{4.9}
\]

For \( n \) satisfying \( 1 \leq n \leq N \), we might well let \( n \geq n_{i+1} \leq N (i = 1, 2, \ldots, L - 1) \) and \( n_i \leq n \leq (n_i + n_{i+1})/2 \). Expanding \( u^n_0 \) into Taylor series with respect to \( t_0 \) yields that

\[
u^n_0 = u^n_0 - \varepsilon_i u_{ht}(t_i), \quad t_i \leq \xi_i \leq t_n, \quad i = 1, 2, \ldots, L. \tag{4.10}
\]

where \( \varepsilon_i \) is the step number from \( t_0 \) to \( t_n \) \( (i = 1, 2, \ldots, L) \). If snapshots are taken at uniform intervals, then \( \varepsilon_i \leq N/(2L) \). If \( u_{ht} \) is bound, we obtain from (4.12)
\[ \| P^d u^n_h - u^n_0 \|_0^2 + \tau \sum_{i=1}^{n} \| \nabla (P^d u^n_h - u^n_0) \|_0^2 \leq C \tau^2 h (N/L)^3 + Ch^2 \sum_{j=1}^{n} \| \nabla (P^d u^n_h - u^n_0) \|_0^2, \quad 1 \leq n \leq N. \] (4.11)

Thus, if \( L^{3/2} = O(N) \) and \( k = O(h) \), by using Lemma 7, we obtain from (4.14) that

\[ \| P^d u^n_h - u^n_0 \|_0^2 + \tau \sum_{i=1}^{n} \| \nabla (P^d u^n_h - u^n_0) \|_0^2 \leq C \tau^2 + C \sum_{j=d+1}^{l} \lambda_j. \] (4.12)

Using triangular inequality and noting that \( k = O(h) \) and \( L^{3/2} = O(N) = O(h^{-1}) \), we can obtain that

\[ \| u^n_h - u^n_0 \|_0 + \tau \sum_{i=1}^{n} \| \nabla (u^n_h - u^n_0) \|_0 \leq C \tau + C \left( \sum_{j=d+1}^{l} \lambda_j \right)^{1/2}, \quad 1 \leq n \leq N, \] (4.13)

which completes the proof of Theorem 8. \( \square \)

Combining Theorems 5 and 8 yields the following result.

**Theorem 9.** Under hypotheses of Theorem 8, the error estimates between the solutions for Problem (II) and the solutions for the reduced Problem (V) are

\[ \| u(t_n) - u^n_0 \|_0 + \tau \sum_{i=1}^{n} \| \nabla (u(t_i) - u^n_0) \|_0 \leq C h^2 + C \tau + C \left( \sum_{j=d+1}^{l} \lambda_j \right)^{1/2}, \quad 1 \leq n \leq N, \] (4.14)

where \( \lambda_j \)'s rely on \( h \) and \( \tau \) since \( V \) depends on them.

**Remark 3.** Inequality (4.1) has shown that the solution of Problem (V) is stable and continuously dependent on source term \( f(x, y, t) \) and initial condition \( u_0(x, y) \). The condition \( L^{3/2} = O(N) \) in Theorem 8 shows the relation between the number \( L \) of snapshots and the number \( N \) of all time instances. Therefore, it is unnecessary to take total transient solutions at all time instances \( t_n \) as snapshots (see [19,20]). Theorems 8 and 9 have presented the error estimates between the solution of the reduced FVE formulation Problem (V) and the solution of the usual FVE formulation Problem (IV) and Problem (II), respectively. Our method here employs some FVE solutions \( u^n_h(n = 1, 2, \ldots, N) \) for Problem (IV) as assistant analysis. However, when one computes actual problems, one may obtain the ensemble of snapshots from physical system trajectories by drawing samples from experiments and interpolation (or data assimilation) or previous results. Therefore, the assistant \( u^n_h(n = 1, 2, \ldots, N) \) could be replaced with the interpolation functions of experimental and previous results, thus rendering it unnecessary to solve Problem (IV), and requiring only to solve directly Problem (V) such that Theorem 8 is satisfied. And then, time instances are continuously extrapolated forward and POD basis is ceaselessly renewed, the rules of future development and change of natural phenomenon would be very well simulated.

**5. Some numerical experiments**

In this section, we present some numerical examples of the two-dimensional parabolic problems to show the advantage of the reduced POD FVE formulation, i.e., Problem (V).

Without losing generality, we might as well take source term \( f(x, y, t) = 0 \), then the two-dimensional parabolic problems with the initial condition \( u(x, y, 0) = \sin \pi x \sin \pi y \) can be written as follows.

\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \quad (x, y) \in \Omega, \quad 0 < t \leq T, \] (5.1)

\[ u(x, y, 0) = \sin \pi x \sin \pi y, \quad (x, y) \in \Omega, \] (5.2)

\[ u(x, y, t) = 0, \quad (x, y) \in \partial \Omega, \quad 0 < t \leq T, \] (5.3)

where \( \Omega = \{(x, y); \ 0 \leq x \leq 1, 0 \leq y \leq 1\}, \partial \Omega \) denotes the boundary of \( \Omega \).

We first divide the field \( \Omega \) into 100 \times 100 small squares with side length \( \Delta x = \Delta y = 0.01 \), and then link the diagonal of the square to divide each square into two triangles in the same direction which consists of triangulation \( \mathcal{T}_h \). Thus \( h = \sqrt{2} \times 0.01 \). In order to satisfy \( \tau = O(h) \), we take time step as \( \tau = 0.01 \), \( T = 200r \). The dual decomposition \( \mathcal{T}_h^* \) is taken as barycenter dual decomposition, i.e., the barycenter of the right triangle \( K \in \mathcal{T}_h \) is taken as the node of the dual decomposition.

Next, we find a group of numerical solutions \( u^n_h \) of the usual FVE method (i.e., Problem (IV)) when \( n = 1, 2, \ldots, 200 \), i.e., at time \( t = 1r, 2r, \ldots, 200r \), constructing 200 numerical solutions. And then, we choose 20 values from 200 values every 10 values to consist of a set of snapshots. Finally, using Matlab software, we find 20 eigenvalues which are arranged in a non-decreasing order, and 20 eigenvectors corresponding to the twenty eigenvalues and using (3.13) and (3.14) we
construct a group of POD bases. Take the first 6 POD bases from 20 POD bases to expand into subspace $U^d$ and compute the errors of the POD FVE numerical solution to Problem (V) and the usual FVE solutions to Eqs. (5.1)–(5.3) at $t = 200\tau$ which are depicted graphically in Figs. 5 and 6, respectively. Fig. 7 is the error between the POD FVE solution and the usual FVE solution at $t = 200\tau$.

When we take 6 POD bases and $\tau = 0.01$, by computing we obtain that $\sum_{j=6+1}^{20} \lambda_j \leq 0.03$. Fig. 8 theoretically shows the errors (log10) between the solutions $u^n_d$ of Problem (V) with 20 different numbers of POD bases and a solution $u^n_h$ of the usual FVE formulation Problem (IV) at $t = 100\tau$ (i.e., $n = 100$) and $t = 200\tau$ (i.e., $n = 200$), respectively. Comparing the usual FVE formulation Problem (IV) with the reduced FVE formulation Problem (V) containing 6 POD bases implementing the numerical simulation computations when total time $t = 200\tau$, we find that for usual FVE formulation Problem (IV) with piecewise linear polynomials for $u^n_h$, which has $100 \times 100 = 10,000$ degrees of freedom, the required computing time is 18 min, while for the reduced FVE formulation Problem (V) with 6 POD bases, which has only 6 degrees of freedom, the corresponding time is only six seconds, i.e., the required computing time to solve the usual FVE formulation Problem (IV) is 180 times as that to do the reduced FVE formulation Problem (V) with 6 POD bases, while the errors between their respective solutions do not exceed $3 \times 10^{-2}$. Though our examples are in a sense recomputing what we have already computed by the usual FVE formulation, when we compute 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 FVE formulation, while it is unnecessary to solve the usual FVE formulation. Thus, the time-consuming calculations and resource demands in the computational process will be greatly saved. It is also shown that finding the approximate solutions for two-dimensional solute transport problems with the reduced FVE formulation Problem (V) is computationally very effective. And the results for numerical examples are consistent with those obtained for the theoretical case.
Fig. 7. Error between the POD FVE solution and the usual FVE solution at $t = 200\tau$.

Fig. 8. When $t = 200\tau$, the errors (log10) between solutions of Problem (V) with different number of POD bases for a group of 20 snapshots and the usual FVE formulation Problem (IV) with piecewise first degree polynomials.

6. Conclusions and perspective

In this paper, we have employed the POD basis to derive a reduced FVE formulation for two-dimensional parabolic problems, analyzed the errors between the solution of their usual FVE formulation and solution of the POD reduced FVE formulation, and discussed theoretically the relation between the number of snapshots and the number of solutions at all time instances, which have shown that our present method has improved and innovated the existing methods. We have validated the correctness of our theoretical results with numerical examples. Though snapshots and POD basis of our numerical examples are structured with the solution of the usual FVE formulation, when one computes actual problems, this process can be omitted in actual applications and one may structure the snapshots and POD basis with interpolation or data assimilation by drawing samples from experiments, then solve Problem (V), while it is unnecessary to solve Problem (IV). Thus, the time-consuming calculations and resource demands in the computational process are greatly saved and the computational efficiency is vastly improved. Therefore, the method in this paper holds a good prospect of extensive applications.

In this paper, we use only the forward FVE formulation based POD to deal with parabolic equations. However, to solve an inverse problem of parabolic equations, for example, to find the initial conditions, boundary conditions, source term, coefficients (if needed), discussing the POD basic sensitivity of initial condition, and so on, by using existing data with the POD technique is very interesting work and an important applied topic of POD, which is our future research work. In future another research work in this area will aim at extending the reduced FVE formulation as well as applying it to a realistic atmospheric operational forecast system and to a set of more complicated PDEs such as the atmosphere quality forecast system, the ocean fluid forecast system, and so on. Moreover, there are still many interesting works for POD applications, for instance, FVE methods based on adaptive POD, FVE methods based on Trust-Region-POD, FVE methods based on POD a posterior error estimates, and so on, which are also worth studying.
References