# Data-driven kinematics-consistent model order reduction of fluid-structure interaction problems: application to deformable microcapsules in a Stokes flow

Claire Dupont, Florian de Vuyst, Anne-Virginie Salsac

# Data-driven kinematics-consistent model order reduction of fluid-structure interaction problems: application to deformable microcapsules in a Stokes flow

**Claire Dupont[1], Florian De Vuyst[2] and Anne-Virginie Salsac[1] †**

[1]Biomechanics and Bioengineering Laboratory (UMR 7338), Université de Technologie de Compiègne - CNRS, 60203 Compiègne, France

[2]Laboratory of Applied Mathematics of Compiègne, Université de Technologie de Compiègne, CS 60319, 60203 Compiègne, France

(Received xx; revised xx; accepted xx)

In this paper, we present a generic approach of a dynamical data-driven model order reduction technique for three-dimensional fluid-structure interaction problems. A low-order continuous linear differential system is identified from snapshot solutions of a high-fidelity solver. The reduced order model (ROM) uses different ingredients like proper orthogonal decomposition (POD), dynamic mode decomposition (DMD) and Tikhonov-based robust identification techniques. An interpolation method is used to predict the capsule dynamics for any value of the governing non-dimensional parameters that are not in the training database. Then a dynamical system is built from the predicted solution. Numerical evidence shows the ability of the reduced model to predict the time-evolution of the capsule deformation from its initial state, whatever the parameter values. Accuracy and stability properties of the resulting low-order dynamical system are analysed numerically. The numerical experiments show a very good agreement, measured in terms of modified Hausdorff distance between capsule solutions of the full-order and low-order models both in the case of confined and unconfined flows. This work is a first milestone to move towards real time simulation of fluid-structure problems, which can be extended to non-linear low-order systems to account for strong material and flow non-linearities. It is a valuable innovation tool for rapid design and for the development of innovative devices.

**Key words:** Fluid-structure interaction, deformable capsule, dynamical system, reduced order model, non-intrusive, data-driven, dynamic mode decomposition

---

## 1. Introduction

Fluid-structure interaction (FSI) problems often occur in Engineering (aircraft and automotive industries, wind turbines) as well as in medical applications (cardiovascular systems,

† Email address for correspondence: anne-virginie.salsac@utc.fr

artificial organs, artificial valves, medical devices, etc.). Today the design of such systems usually requires advanced studies and high-fidelity (HF) numerical simulations become an essential tool of computed-aided analysis. However, computational FSI is known to be very time-consuming even when using high-performance computing facilities. Usually, engineering problems are parameterized and the search of suitable designs require numerous computer experiments leading to prohibitive computational times. For particular applications such as the tracking of drug carrier capsules flowing in blood vessels, it would be ideal to have real-time simulations for a better understanding of the behaviour of the dynamics and for efficiency assessment. Unfortunately, today high-fidelity real-time FSI simulations are far from being reached with current High Performance Computing (HPC) facilities.

A current trend is to use machine learning (ML) or artificial intelligence (AI) tools such as artificial neural networks (ANN). Such tools learn numerical simulations from HF solvers and try to map entry parameters with output criteria in an efficient way, with response times far less than HF ones, say 3 or 4 orders of magnitude smaller. In some sense, heavy HF computations and training stage are done in an offline stage, and learned ANNs can be used online for real time evaluations and analysis. However, ML and ANN today are not fully satisfactory for dynamical problems, and/or the training stage itself may be time consuming, thus requiring more Central Processing Unit (CPU) time. Another option is the use of model order reduction (MOR). Reduced-order modelling (ROM) can be seen as a 'grey-box' supervised ML methodology, taking advantage of the expected low-order dimensionality of the FSI mechanical problem. By 'grey-box' we mean that the low-dimensional encoding of the ML process is based on mechanical principles and a man-made preliminary dimensionality reduction study. This allows one for a better control of the ROM accuracy and behaviour. There are two families of MOR: intrusive and non-intrusive approaches. The intrusive approaches use physical equations. The low-order model is derived by setting the physical problem on a suitable low-dimensional space. The accuracy can be very good, but the price to pay is the generation of a new code which can be a tedious and long task. The non-intrusive approach does not require heavy code development. It is based on HF simulation results used as entry data. Although it is not based on high-fidelity physical equations, a non-intrusive approach can include a priori physical informations, like e.g. meaningful physical features, prototype of system of equations, pre-computed principal components, consistency with physical principles, etc.

In the recent literature, efficient intrusive ROMs for FSI have been proposed e.g. in (Quarteroni et al. 2016). But to our knowledge there are far less contributions in non-intrusive ROMs dedicated to FSI.

In this paper, we propose a data-driven model order reduction approach for FSI problem which is consistent with the equations of kinematics and is designed from a low-order meaningful system of equations. As case of study, we focus on the motion of a microcapsule, a droplet surrounded by a membrane, subjected to a confined and unconfined Stokes flow.

Artificial microcapsules can be used in various industrial applications such as in cosmetics (Miyazawa et al. 2000; Casanova & Santos 2016), food industry (Yun et al. 2021) and biotechnology, where drug targeting is a high potential application (Ma & Su 2013; Abuhamdan et al. 2021; Ghiman et al. 2022). Once in suspension in an external fluid, capsules are subjected to hydrodynamics forces, which may lead to large membrane deformation, wrinkle formation or damage. The numerical model must be able to capture the time-evolution of the non-linear 3D large deformations of the capsule membrane. Different numerical strategies are possible to solve the resulting large systems of equations (Lefebvre & Barthès-Biesel 2007; Hu et al. 2012; Ye et al. 2017; Tran et al. 2020). However, they all have long computational times.

Different approaches have been used over the past decade to accelerate the computations, such as HPC (e.g. Zhao et al. (2010)) and Graphics Processing units (e.g. Matsunaga et al. (2014)). More recently, reduced order models have been proposed to predict the motion of capsules suspended in an external fluid flow. In Quesada et al. (2021), the authors used the large amount of data generated by numerical simulations to show how relevant it is to recycle these data to produce lower-dimensional problem using physics-based reduced order models. However, their method can only predict the steady-state capsule deformed shape. Boubehziz et al. (2021) show for the first time the efficiency of data-driven model-order reduction technique to predict the dynamics of the capsule in a microchannel. However, the method is cumbersome as it requires two bases, one to predict the velocity field, the other to capture the shape evolution over time. And then they reconstruct the solution in the parameter space thanks to diffuse approximation (DA) strategy.

The proposed method serves different objectives. We have designed the method to be non-intrusive for practical uses of existing high-fidelity FSI solver (also referred to as the Full-Order Model, or FOM). That means that the ROM methodology should be data-driven. We also want the ROM to be consistent with the equations of kinematics. The model must thus return the displacement $\{u\}$ and velocity $\{v\}$ fields from a few snapshots provided by the FOM. It must otherwise be able to predict the solution for any parameter vector in predefined admissible domain. Finally, the kinematics-consistent data-driven reduced-order model of capsule dynamics must ideally open the way to real-time simulations. To do so, we use a coupling between methods that have been devised to analyse complex fluid problems:

- Proper Orthogonal Decomposition (POD) (Lumley 1967; Sirovich 1987)
- Dynamic Mode Decomposition (DMD) (Schmid 2010)

along with a Tikhonov regularization for robustness purposes. An interpolation method is implemented to predict the solution for any values of governing parameters that are not present in the training database.

As indicated above, we mainly consider the case of an initially spherical capsule flowing in a microfluidic channel with a square cross-section. The corresponding FOM was developed by Hu et al. (2012) and used to get a complete numerical database of the three-dimensional capsule dynamics as a function of the parameters of the problem: the capsule-to-tube confinement ratio, hereafter referred to as size ratio $a/\ell$ and the capillary number $Ca$, which measures the ratio between the viscous forces acting onto the capsule membrane and the membrane elastic forces. For clarity reasons, different ROMs are introduced with increasing levels of generality, as detailed in Table 1. First, we consider a fixed parameter vector, and get a space-time ROM in the form of a low-order dynamical system. Next, we generate such $N$ ROMs for the $N$ parameter samples that fill the admissible parameter domain, and then assess the uniform accuracy (space-time accuracy over the whole sample set). Finally, we propose a strategy to derive a general space-time-parameter ROM for any value of the parameter vector $(Ca, a/\ell)$ in the admissible space. To conclude the results section, we apply the ROM model to a capsule in a simple shear flow.

The paper is organized as follows. First, we present the physics of the problem and the FOM in Section 2. The strategy used to develop a non-intrusive space-time ROM is detailed in Section 3. We first present the results for an initially spherical flowing in a square channel. They first show the results for a given configuration in Section 4, generalize them in Section 5 on the entire database, formed by all the cases that have reached a stationary state and present in Section 6 the methodology of space-time-parameter ROM. In Section 7, we then apply the ROM model to a capsule in a simple shear flow before discussing the advantages and the limits of the method in Section 8.

| Nb of parameter samples for data | ROM output type | Verification (accuracy) | Related Section(s) in the paper |
|---|---|---|---|
| 1 | 1 space-time ROM | Space-time accuracy | Sections 3 and 4 |
| $N$ | $N$ space-time ROM | Uniform space-time accuracy on the sample set | Section 5 |
| $N$ | 1 space-time-parameter ROM (any parameter couple) | Uniform accuracy | Section 6 |

Table 1: Stepwise procedure for ROM construction of increasing level of generality.

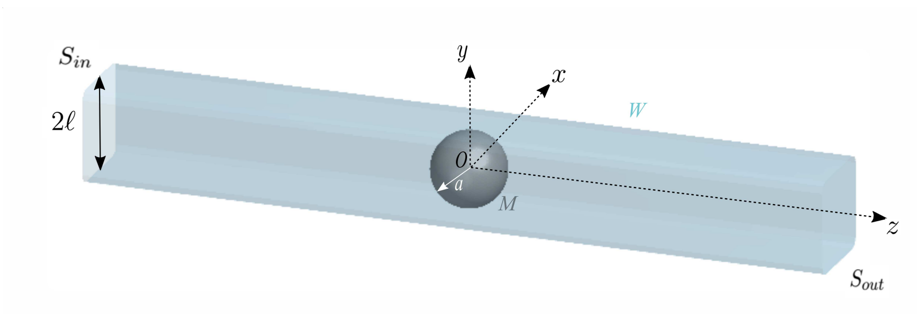


Figure 1: Sketch of the model geometry showing an initially spherical capsule of radius $a$ placed in a channel with a constant square section of side $2\ell$.

## 2. Full-order microcapsule model, parameters and quantities of interest

### 2.1. *Problem description for a spherical capsule in a channel flow*

An initially spherical capsule of radius $a$ flows within a long microfluidic channel having a constant square section of side $2\ell$ (Figure 1). The suspending fluid and capsule liquid core are incompressible Newtonian fluids with the same kinematic viscosity $\eta$.

The capsule liquid core is enclosed by a hyperelastic isotropic membrane. Its thickness is assumed to be negligible compared to the capsule dimension. The membrane is thus modelled as a surface devoid of bending stiffness with surface shear modulus $G_S$. The two non-dimensional governing parameters of the problem are the size ratio $a/\ell$ and the capillary number

$$Ca = \eta V / G_S \tag{2.1}$$

where $V$ is the mean axial velocity of the undisturbed external Poiseuille flow.

The flow Reynolds number is assumed to be very small. We solve the Stokes equations in the external ($\beta = 1$) and internal fluids ($\beta = 2$), together with the membrane equilibrium equation to determine the dynamics of the deformable capsule within the microchannel.

For the fluid problem, we denote $\boldsymbol{v}^{(\beta)}$, $\boldsymbol{\sigma}^{(\beta)}$ and $p^{(\beta)}$ the velocity, stress and pressure fields in the two fluids. These parameters are non-dimensionalized using $\ell$ as characteristic length, $\ell/V$ as characteristic time and $G_S\ell$ as characteristic force. The non-dimensional Stokes equations

$$\nabla p^{(\beta)} = Ca\nabla^2 \boldsymbol{v}^{(\beta)}, \quad \nabla\cdot\boldsymbol{v}^{(\beta)} = 0, \quad \beta = 1,2 \tag{2.2}$$

are solved in the domain bounded by the cross sections $S_{in}$ at the tube entrance and $S_{out}$ at the exit. These cross sections are assumed to be both located far from the capsule. The reference frame $(O, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is centred at each time step on the capsule center of mass $O$ in the high-fidelity code, but the displacement of the capsule center of mass along the tube axis $\boldsymbol{Oz}$ is computed.

The boundary conditions of the problem are the following ones:

• The velocity field is assumed to be the unperturbed flow field on $S_{in}$ and $S_{out}$, i.e. the flow disturbance vanish far from the capsule.

• The pressure is uniform on $S_{in}$ and $S_{out}$.

• A no-slip boundary condition is assumed at the channel wall $W$ and on the capsule membrane $M$:

$$\forall \boldsymbol{x} \in W, \boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{0}; \quad \forall \boldsymbol{x} \in M,\ \boldsymbol{v}(\boldsymbol{x}) = \frac{\partial \boldsymbol{u}}{\partial t}. \tag{2.3}$$

• The normal load $\boldsymbol{n}$ on the capsule membrane $M$ is continuous, i.e. the non-dimensionalized external load per unit area $\boldsymbol{q}$ exerted by both fluids is due to the viscous traction jump:

$$(\boldsymbol{\sigma}^{(1)} - \boldsymbol{\sigma}^{(2)}) \cdot \boldsymbol{n} = \boldsymbol{q} \tag{2.4}$$

where $\boldsymbol{n}$ is the unit normal vector pointing towards the suspending fluid.

To close the problem, the external load $\boldsymbol{q}$ on the membrane is deduced from the local equilibrium equation, which, in absence of inertia, can be written as

$$\boldsymbol{\nabla}_s \cdot \boldsymbol{\tau} + \boldsymbol{q} = \boldsymbol{0} \tag{2.5}$$

where $\boldsymbol{\tau}$ is the non-dimensionalized Cauchy tension tensor (forces per unit arclength in the deformed plane of the membrane) and $\boldsymbol{\nabla}_s\cdot$ is the surface divergence operator. We assume that the membrane deformation is governed by the strain-softening neo-Hookean law. The principal Cauchy tensions can then be expressed as

$$\tau_1 = \frac{G_S}{\lambda_1\lambda_2}\left[\lambda_1^2 - \frac{1}{(\lambda_1\lambda_2)^2}\right] \text{ (likewise for } \tau_2\text{)}, \tag{2.6}$$

where $\lambda_1$ and $\lambda_2$ are the principal extension ratios measuring the in-plane deformation.

### 2.2. *Numerical procedure*

The FSI problem is solved by coupling a finite element method that determines the capsule membrane mechanics with a boundary integral method that solves for the fluid flows (Walter et al. 2010; Hu et al. 2012). Thanks to the latter, only the boundaries of the flow domain, i.e the channel entrance $S_{in}$ and exit $S_{out}$, the channel wall and the capsule membrane have to be discretized to solve the problem. The mesh of the initially spherical capsule is generated by subdividing the faces of the icosahedron (regular polyhedron with 20 triangular faces) inscribed in the sphere until reaching the desired number of triangular elements. At the last step, nodes are added at the middle of all the element edges to obtain a capsule mesh with 1280 $P_2$ triangular elements and 2562 nodes, which correspond to a characteristic mesh size $\Delta h_C = 0.075\,a$. The channel mesh of the entrance surface $S_{in}$ and exit surface $S_{out}$ and of the channel wall is generated using `Modulef` (INRIA, France). The central portion of the channel, where the capsule is located, is refined. The channel mesh comprises 3768 $P_1$ triangular elements and 1905 nodes.

At time $t = 0$, a spherical capsule is positioned with its center of mass $O$ on the channel axis. At each time step, the in-plane stretch ratio $\lambda_1$ and $\lambda_2$ are computed from the nodes deformation. The elastic tension tensor $\boldsymbol{\tau}$ is then deduced from the values of $\lambda_1$ and $\lambda_2$. The finite element method is used to solve the weak form of the membrane equilibrium equation (2.5) and determine the external load $\boldsymbol{q}$.

Applying the boundary integral method, the velocity of the nodes on the capsule membrane reads (Pozrikidis 1992):

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{v}^{\infty}(\boldsymbol{x}) - \frac{1}{8\pi\mu_F}\left[\int_M \boldsymbol{J}(\boldsymbol{r})\cdot\boldsymbol{q}\,dS(\boldsymbol{y}) + \int_W \boldsymbol{J}(\boldsymbol{r})\cdot\boldsymbol{f}\,dS(\boldsymbol{y}) - \Delta P\int_{S_{out}} \boldsymbol{J}(\boldsymbol{r})\cdot\boldsymbol{n}\,dS(\boldsymbol{y})\right] \tag{2.7}$$

for any $\boldsymbol{x}$ in the spatial domain when the suspending and internal fluids have the same viscosity. The vector $\boldsymbol{f}$ is the disturbance wall friction due to the capsule, $\Delta P$ is the additional pressure drop and $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{x}$.

To update the position of the membrane nodes, the nodal displacement $\boldsymbol{u}$ is computed by integrating equation (2.3) in time. The procedure is repeated until the desired non-dimensional time $VT/\ell$.

For later development, it is more convenient to work on the condensed abstract form of the system. The full order semi-discrete FSI system to solve consists of the kinematics and the membrane equilibrium algebraic equations:

$$\{\dot{u}\} = \{v\}, \qquad t \in [0, T], \tag{2.8}$$

$$\{v\} = \varphi(\{u\}) \tag{2.9}$$

where $\varphi$ is a non-linear mapping from $\mathbb{R}^{3d}$ to $\mathbb{R}^{3d}$ and $d$ is the number of nodes on the membrane. Regarding time discretization, a Runge-Kutta Ralston scheme is used:

$$\begin{aligned}
\{\widehat{u}^{n+2/3}\} &= \{u^n\} + \frac{2}{3}\Delta t\,\{v^n\},\\
\{\widehat{v}^{n+2/3}\} &= \{\varphi\}(\{\widehat{u}^{n+2/3}\}),\\
\{u^{n+1}\} &= \{u^n\} + \Delta t\left(\frac{1}{4}\{v^n\} + \frac{3}{4}\{\widehat{v}^{n+2/3}\}\right),\\
\{v^{n+1}\} &= \{\varphi\}(\{u^{n+1}\}),\\
\{u^0\} &= \{0\},\ \{v^0\} = \{\varphi\}(\{0\})
\end{aligned}$$

where $\Delta t > 0$ is a constant time step and $\{u\}^n$ and $\{v\}^n$ respectively represent the discrete membrane displacement field and the discrete membrane velocity field at discrete time $t^n = n\Delta t$. The initial condition is simply $\{u\}^0 = \{0\}$.

The whole numerical scheme is subject to some Courant-Friedrichs-Lewy (CFL) type stability condition on the time step (Walter et al. 2010) because of its explicit nature. The numerical method is conditionally stable if the time step $\Delta t$ satisfies

$$\frac{V}{\ell}\Delta t < O\left(\frac{\Delta h_C}{\ell}Ca\right). \tag{2.10}$$

From the computational point of view, the resolution of (2.9) at each time step requires i) the computation of the disturbance wall friction $\boldsymbol{f}$ at all the wall nodes, ii) the additional pressure drop $\Delta P$, iii) the traction jump $\boldsymbol{q}$ at the membrane nodes and iv) the boundary integrals for each node. The resulting numerical FOM may thus be time-consuming, depending on the membrane discretization and the number of time steps. Figure 2 shows that the evolution of the computational cost when $a/\ell = 0.7$, considering the mesh discretization described above and a workstation equipped with 2 processors Intel® Xeon® Gold 6130 CPU (2.1 GHz).

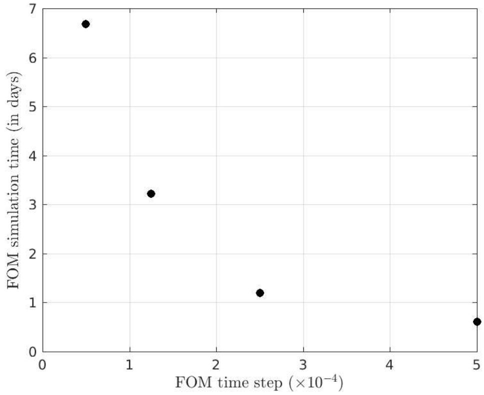


Figure 2: Simulation time of the dynamics of the capsule over a non-dimensional time $Vt/\ell = 10$ ($a/\ell = 0.7$) according to the time step. Simulations were performed on a workstation equipped with 2 processors Intel® Xeon® Gold 6130 CPU (2.1 GHz).

A week of computation is sometimes necessary to simulate the dynamics of an initially spherical capsule in a microchannel over the non-dimensional time $VT/\ell = 10$.

For that reason, a model-order reduction (MOR) strategy is studied in this paper, in order to reduce the computational time by several orders of magnitude. ROMs try to approximate solutions of the initial problem by strongly lowering the dimensionality of the numerical model, generally using a reduced basis (RB) of suitable functions, then derive a low-order system of equations.

In the case of differential algebraic equations (DAE) like (2.8)-(2.9), the reduced system of equations to find should also be of DAE nature. Remark that it is often possible to reformulate DAEs as a system of ordinary differential equations (ODEs) (Ascher & Petzold 1998). In the next section, we give details on the chosen ROM methodology for the particular case and context of FSI capsule problem.

## 3. Non-intrusive space-time model-order reduction strategy

In this section, the parameter couple $\boldsymbol{\theta} = (Ca, a/\ell)$ is fixed, thus we omit the dependency of the solutions with respect to $\boldsymbol{\theta}$ for the sake of simplicity. For the derivation of the ROM model, we consider the semi-discrete time-continuous version of the FOM, i.e. (2.8)-(2.9).

### 3.1. *Dimensionality reduction and reduced variables for displacements and velocities*

Assume first that, for any $t \in [0, T]$, the discrete velocity field can be accurately approximated according to the expansion

$$\{v\}(t) \approx \sum_{k=1}^{K} \beta_k(t)\,\{\phi_k\} \tag{3.1}$$

for some orthonormal modes $\{\phi_k\} \in \mathbb{R}^d$ and real coefficients $\beta_k(t)$. The truncation rank $K \leqslant d$ is of course expected to be far less than $d$ as expected in a general ROM methodology.

From the kinematics equations we have

$$\{u\}(t) = \int_0^t \{v\}(s)\, ds$$
$$\approx \int_0^t \beta_k(s)\, \{\phi_k\}\, ds$$

so that the displacement field can be accurately represented by

$$\{u\}(t) \approx \sum_{k=1}^{K} \alpha_k(t)\, \{\phi_k\} \tag{3.2}$$

where $\alpha_k(t) = \int_0^t \beta_k(s)\, ds$. The coefficients $(\alpha_k(t))_k$ and $(\beta_k(t))_k$ are called the reduced variables. For the sake of readability and mental correspondence between full-order unknowns and reduced ones, we will use the convenient notations

$$\boldsymbol{\alpha}(t) = (\alpha_1(t), ..., \alpha_K(t))^T, \quad \boldsymbol{\beta}(t) = (\beta_1(t), ..., \beta_K(t))^T$$

where the exponent $^T$ denotes the transpose of the matrix. The condensed matrix forms of (3.2) and (3.1) respectively are

$$\{u\}(t) \approx Q\, \boldsymbol{\alpha}(t), \quad \{v\}(t) \approx Q\, \boldsymbol{\beta}(t), \tag{3.3}$$

where $Q = [\{\phi_1\}, ..., \{\phi_K\}] \in \mathscr{M}_{dK}$. Since the modes $\{\phi_k\}$ are assumed to be orthonormal (for the standard Euclidean inner product), the matrix $Q$ is a semi-orthogonal matrix, i.e. $Q^T Q = I_K$. In particular, we have $\boldsymbol{\alpha}(t) \approx Q^T\, \{u\}(t)$ and $\boldsymbol{\beta}(t) = Q^T\, \{v\}(t)$.

Note that the modes $\{\phi_k\}$ and reduced variables $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ are determined for each parameter set $(Ca, a/\ell)$, but a common value of the truncation rank $K$ is chosen for all the sets. Its practical computation will be detailed in a next subsection, as well as that of the modes $\{\phi_k\}$.

### 3.2. *ROM prototype*

The expressions $\{\tilde{u}\}(t) = Q\, \boldsymbol{\alpha}(t)$ and $\{\tilde{v}\}(t) = Q\, \boldsymbol{\beta}(t)$ provide low-order representations of displacement and velocity fields respectively. We can now write equations for the reduced vectors $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\beta}(t)$ respectively. In this subsection, let us consider a projection Galerkin-type approach. Let us denote $\langle ., . \rangle$ the standard Euclidean scalar product in $\mathbb{R}^d$. Considering a test vector $\{w\}$ in $W = span(\{\varphi_1\}, ..., \{\varphi_K\})$, we look for an approximate displacement field $\{\tilde{u}\}(t)$ solution of the projected kinematics equations

$$\langle \frac{d}{dt}\{\tilde{u}\}(t), \{w\}\rangle = \langle \{\tilde{v}\}(t), \{w\}\rangle \quad \forall\, \{w\} \in W.$$

By considering each test vector $\{w\} = \{\varphi_k\}$, we get the consistent reduced kinematics equation

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{\beta}. \tag{3.4}$$

Consider now the projected field $\{\tilde{v}\}(t)$ which is solution of the system of algebraic equations (Galerkin approach):

$$\langle \{\tilde{v}\}(t), \{w\}\rangle = \langle \varphi(\{\tilde{u}\}(t)), \{w\}\rangle \quad \forall\, \{w\} \in W. \tag{3.5}$$

Again by taking the test vector $\{w\} = \{\phi_k\}$, we have

$$\{\phi_k\}^T Q\beta(t) = \{\phi_k\}^T \varphi(Q\boldsymbol{\alpha}(t)).$$

Considering all $k$ in $\{1, ..., K\}$, since $Q = [\{\phi_1\}, ..., \{\phi_K\}]$ and $Q^T Q = I_K$ we get

$$Q^T Q \boldsymbol{\beta}(t) = \boldsymbol{\beta}(t) = Q^T \varphi(Q\boldsymbol{\alpha}(t)).$$

It is in the form

$$\boldsymbol{\beta}(t) = \varphi_r(\boldsymbol{\alpha}(t)) \tag{3.6}$$

with the mapping $\varphi_r : \mathbb{R}^K \to \mathbb{R}^K$ defined by $\varphi_r(\boldsymbol{\alpha}) = Q^T \varphi(Q\boldsymbol{\alpha})$. We get a reduced-order algebraic equilibrium equation. Unfortunately, because of the non-linearities in $\varphi$, the computation of $\varphi_r(\boldsymbol{\alpha})$ requires high-dimensional $O(d)$ operations, making this approach irrelevant from the performance point of view. A possible solution to deal with the non-linear terms would be to use for example Empirical Interpolation Methods (EIM) (Barrault et al. 2004) but from the algorithm and implementation point of view, this would lead to an intrusive approach with specific code developments. We here rather adopt a linearization strategy in the following sense: by derivating (3.6) with respect to time, we get

$$\dot{\boldsymbol{\beta}}(t) = \frac{\partial \varphi_r}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}(t))\, \dot{\boldsymbol{\alpha}}(t).$$

Thanks to the reduced kinematics equation (3.4), we get

$$\dot{\boldsymbol{\beta}}(t) = \frac{\partial \varphi_r}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}(t))\, \boldsymbol{\beta}(t). \tag{3.7}$$

Since $\varphi_r$ is hard to evaluate, it is even harder to evaluate its differential. But the differential $\frac{\partial \varphi_r}{\partial \alpha}(\boldsymbol{\alpha}(t))$ can be approximated itself, or replaced by some matrix $A(t)$. Then we get a ROM structure (ROM prototype) in the form

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{\beta}(t), \tag{3.8}$$

$$\dot{\boldsymbol{\beta}}(t) = A(t)\, \boldsymbol{\beta}(t). \tag{3.9}$$

The differential system (3.8)-(3.9) is linear with variable coefficient matrix $A(t) \in \mathscr{M}_K(\mathbb{R})$. It can be written in matrix form

$$\frac{d}{dt}\begin{pmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\beta}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} [0] & I_K \\ [0] & A(t) \end{pmatrix}}_{=\mathbb{A}(t)} \begin{pmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\beta}(t) \end{pmatrix}. \tag{3.10}$$

The spectral properties of the differential system (3.10) are related to the spectral properties of matrix $A(t)$. In particular, if all the (complex) eigenvalues $\lambda_k(t)$ of $A(t)$ are such that $\mathrm{Re}(\lambda_k(t)) < 0$ for all $k$ (uniformly distributed in time), then the system is dissipative.

### 3.3. *Nonintrusive approach, SVD decomposition and POD modes*

One of the requirements of this work is to achieve a non-intrusive reduced-order model, meaning that the effective implementation of the ROM does not involve tedious low-level code development into the FOM code. For that, a data-based approach is adopted: from the FOM code, it is possible to compute FOM solutions $(\{u\}^n, \{v\}^n)$ at discrete times $t^n$, $n = 0, ..., N$ ($t^N = N\Delta t = T$), then store some snapshot solutions (called snapshots) into a database for data analysis and later design of a ROM. Proper Orthogonal Decomposition (POD) (Berkooz et al. 1993) is today a well-known dimensionality reduction approach to determine the principal components from solutions of partial differential equations. The Sirovich's snapshot variant approach (Sirovich 1987) is based on snapshot solutions from a FOM to get *a posteriori* empirical POD modes $\{\varphi_k\}$. For the sake of simplicity, assume that the snapshot solutions are all the discrete FOM solution at simulation instants. Applying a singular value decomposition (SVD) to the displacement snapshot matrix

$$\mathbb{S}^u = \left[\boldsymbol{u}^1, \boldsymbol{u}^2, ..., \boldsymbol{u}^N\right],$$

of size $d \times N$, we get the SVD decomposition

$$\mathbb{S}^u = U \Sigma V^T \tag{3.11}$$

with orthogonal matrices $U \in \mathscr{M}_d(\mathbb{R})$, $V \in \mathscr{M}_N(\mathbb{R})$ and the singular value matrix $\Sigma = diag(\sigma_k) \in \mathscr{M}_{d\times N}(\mathbb{R})$, with $\sigma_k \geqslant 0$ for all $k$ organized in decreasing order: $\sigma_1 \geqslant \sigma_2 \geqslant ... \geqslant \sigma_{\min(d,N)} \geqslant 0$. From SVD theory, for a given accuracy threshold $\varepsilon > 0$, the truncation rank $K = K(\varepsilon)$ is computed as the smallest integer such that the inequality

$$\frac{\displaystyle\sum_{k=K+1}^{\min(d,N)} \sigma_k^2}{\displaystyle\sum_{k=1}^{\min(d,N)} \sigma_k^2} \leqslant \varepsilon \tag{3.12}$$

holds (Shawe-Taylor & Cristianini 2004). Proceeding like that, it is shown that the relative orthogonal projection error of the snapshots $\{v\}^n$ onto the linear subspace $W$ spanned by the $K$ first eigenvectors of $U$ is controlled by $\varepsilon$. Denoting $\pi^W$ the linear orthogonal projection operator over $W$, we have:

$$\sum_{n=1}^{N} \|\{v\}^n - \pi^W \{v\}^n\|^2 \leqslant \varepsilon \sum_{n=1}^{N} \|\{v\}^n\|^2.$$

The matrix $Q$ is obtained as the restriction of $U$ to its $K$ first columns.

### 3.4. *Data-driven identification of coefficient matrix*

The system (3.8)-(3.9) is still not closed since the coefficient matrices $A(t)$ are unknowns. From FOM data, one can try to identify the matrices by minimizing some residual function that measures the model discrepancy. The simplest linear model corresponds to the case where $A(t)$ is searched as a time-constant matrix $A$. In this case, equation (3.9) becomes $\dot{\boldsymbol{\beta}}(t) = A\,\boldsymbol{\beta}(t)$. This is the scope of this article. From the time continuous problem, one could determine the matrix $A$ by minimizing the least square functional

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2} \int_0^T \|\dot{\boldsymbol{\beta}}(t) - A\boldsymbol{\beta}(t)\|^2\, dt.$$

But practically, we only have velocity snapshot data at discrete times and we do not have access to the time derivatives of the velocity fields. So the following numerical procedure is adopted: from the velocity snapshot matrix $\mathbb{S}^v = [\{v\}^1, ..., \{v\}^N]$, we compute first the reduced snapshots variables:

$$\boldsymbol{\beta}^n = Q\,\{v\}^n, \quad n = 1, ..., N.$$

Next, we determine a matrix $A$ that minimizes the least square cost function:

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2} \sum_{n=1}^{N-1} \left\| \frac{\boldsymbol{\beta}^{n+1} - \boldsymbol{\beta}^n}{\Delta t} - A\boldsymbol{\beta}^n \right\|^2 \tag{3.13}$$

In (3.13), the finite difference $\dfrac{\boldsymbol{\beta}^{n+1} - \boldsymbol{\beta}^n}{\Delta t}$ is a first-order approximation (in $\Delta t$) of $\dot{\boldsymbol{\beta}}$ at time $t^n$. In appendix A, we provide a mathematical analysis of the effect of time discretization in (3.13) about the impact on the stability of the resulting identified differential system compared to the initial one.

The minimization problem (3.13) can be written in condensed matrix form

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\|\mathbb{Y} - A\mathbb{X}\|_F^2 \tag{3.14}$$

with the two data matrices

$$\mathbb{X} = \left[\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, ..., \boldsymbol{\beta}^{N-1}\right], \quad \mathbb{Y} = \left[\frac{\boldsymbol{\beta}^2 - \boldsymbol{\beta}^1}{\Delta t}, ..., \frac{\boldsymbol{\beta}^N - \boldsymbol{\beta}^{N-1}}{\Delta t}\right]. \tag{3.15}$$

Because $\mathbb{X}$ and $\mathbb{Y}$ store reduced variables (of size $K$), for a sufficient number of discrete snapshot times, these two matrices are horizontal ones. We will assume that the rank of $\mathbb{X}$ is always maximal, i.e. equal to $K$. The least-square solution $A$ of (3.14) is then given by

$$A = \mathbb{Y}\mathbb{X}^\dagger \tag{3.16}$$

where $\mathbb{X}^\dagger = \mathbb{X}^T(\mathbb{X}\mathbb{X}^T)^{-1}$ is the Moore-Penrose pseudo-inverse matrix of $\mathbb{X}$. This least square approach has close connections with SVD-based Dynamic Mode Decomposition (DMD) (Schmid 2010; Kutz et al. 2016).

### 3.5. *Tikhonov least-square regularized formulation*

From standard spectral theory arguments, it is expected that the POD coefficients rapidly decay when $k$ increases as soon as both displacement and velocity fields are smooth enough. A possible side effect is the bad condition number of the matrix $\mathbb{X}$, since the last rows of $\mathbb{X}$ have small coefficients (thus leading to row vectors close to zero 'at the scale' of the first row of $\mathbb{X}$). Even if the solution $A$ in (3.16) always exists, the solution may be sensitive to perturbations, noise or round-off errors. In order to get a robust identification approach, one can regularize the least-square problem (3.14) by adding a Tikhonov regularization term (see e.g. (Aster et al. 2019))

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\|\mathbb{Y} - A\mathbb{X}\|_F^2 + \frac{\mu}{2}\|\mathbb{X}\|_F^2\,\|A\|_F^2 \tag{3.17}$$

where the scalar $\mu > 0$ is the regularization coefficient. The factor $\|\mathbb{X}\|_F^2$ in the regularization term has been added for scaling purposes. The solution $A_\mu$ of (3.17) is given by

$$A_\mu = \mathbb{Y}\mathbb{X}^T\left(\mathbb{X}\mathbb{X}^T + \mu\|\mathbb{X}\|_F^2\, I_K\right)^{-1}. \tag{3.18}$$

*Choice of optimal regularization coefficient*

Of course, the solution matrix $A_\mu$ depends on the regularization coefficient $\mu$ and one can ask what is the optimal choice for $\mu$. There is a trade-off between the approximation quality measured by the residual $\|\mathbb{Y} - A_\mu\mathbb{X}\|_F$ and the norm solution $\|A_\mu\|_F$. The minimization of $\|A_\mu\|_F$ should ensure that unneeded features will not appear in the regularized solution. When plotted on the log-log scale, the curve of optimal values $\mu \mapsto \|A_\mu\|_F$ versus the residual $\mu \mapsto \|\mathbb{Y} - A_\mu\mathbb{X}\|_F$ often takes on a characteristic L shape (Aster et al. 2019). A design of experiment with the test of different values of $\mu$ (starting say from $10^{-12}$ to $10^{-3}$) generally allow to find quasi-optimal values of $\mu$ located at the corner of the L-curve, thus providing a good trade-off between the two criteria.

### 3.6. *Reduced-order continuous dynamical system*

Once the matrix $A_\mu$ has been determined, we get the reduced-order continuous dynamical system

$$\dot{\alpha} = \boldsymbol{\beta}, \tag{3.19}$$

$$\dot{\boldsymbol{\beta}} = A_\mu \boldsymbol{\beta} \tag{3.20}$$

with initial conditions $\boldsymbol{\alpha}(0) = \mathbf{0}$, $\boldsymbol{v}(0) = Q^T \varphi(\{0\})$. At any time $t$, one can go back to the high-dimensional physical space using the POD modes: $\{u\}(t) = Q\boldsymbol{\alpha}(t)$, $\{x\}(t) = \{X\} + \{u\}(t)$, $\{v\}(t) = Q\boldsymbol{\beta}(t)$. As already mentioned, the system can be written in condensed matrix form

$$\dot{\boldsymbol{w}} = \mathbb{A}_\mu \boldsymbol{w} \tag{3.21}$$

where $\boldsymbol{w}(t) = (\boldsymbol{\alpha}(t), \boldsymbol{\beta}(t))^T$ and $\mathbb{A}_\mu = \begin{pmatrix} [0]_K & I_K \\ [0]_K & A_\mu. \end{pmatrix}$.

The exact analytical solution of (3.21) is

$$\boldsymbol{w}(t) = \exp\big(\mathbb{A}_\mu t\big)\, \boldsymbol{w}(0). \tag{3.22}$$

The stability of the differential system depends on the spectral structure of $\mathbb{A}_\mu$, or equivalently on the spectrum of $A_\mu$. Because of the stability of the fluid-capsule coupled system and from accurate solutions of the FOM solver, one can hope that the solution $A_\mu$ of the least-square identification problem has the expected spectral properties. This will be studied and discussed in the numerical experimentation section. From the kinetic energy point of view, it is shown in appendix B that the stability of the kinetic energy is linked to the property of the (real) spectrum of the symmetric matrix $(A_\mu + A_\mu^T)/2$.

*Model consistency with steady states*

A steady state in our context is defined by a capsule that reaches a constant velocity $\{v\}_\infty$, so that the motion becomes a translation flow in time in the direction $\{v\}_\infty$. From (3.1), this shows that $\boldsymbol{\beta}(t)$ also reaches a constant vector $\boldsymbol{\beta}_\infty$, and $\dot{\boldsymbol{\beta}} = 0$ at steady state. As a consequence, from (3.20), we get $A_\mu \boldsymbol{\beta}_\infty = 0$, meaning that 0 is an eigenvalue of $A_\mu$ with $\boldsymbol{\beta}_\infty$ as eigenvector. As a conclusion, the matrix $A_\mu$ must have zero in its spectrum in order to be consistent with the existence of steady states.

### 3.7. *Reduced-order discrete dynamical system*

Of course, it is also possible to derive a discrete dynamical system from the continuous one by using a standard time advance scheme. For example the explicit forward Euler scheme with a constant time step $\Delta t$ gives

$$\boldsymbol{\alpha}^{n+1} = \boldsymbol{\alpha}^n + \Delta t\, \boldsymbol{\beta}^n, \tag{3.23}$$

$$\boldsymbol{\beta}^{n+1} = \boldsymbol{\beta}^n + \Delta t\, A_\mu \boldsymbol{\beta}^n. \tag{3.24}$$

By multiplying (3.23) by $Q$ we get the space-time approximate solution

$$\{u\}^{n+1} = \{u\}^n + \Delta t\, \{v\}^n,$$

so the ROM model is completely consistent with the kinematics equation. Stability properties of the discrete system are linked to the spectral properties of the matrix

$$A_\mu^\Delta = \begin{pmatrix} I_K & \Delta t\, I_K \\ [0]_K & (I_K + \Delta t\, A_\mu) \end{pmatrix}$$

For unconditional stability in time, it is required for the eigenvalues of $I_K + \Delta t A_\mu$ to stay in the unit disk of the complex plane.

More generally, it is possible to use any other time advance scheme, according to the expected order of accuracy or stability domain.

### 3.8. *Accuracy criteria and similarity distances between ROM and FOM solutions*

In order to quantify the error induced by approximations, we introduce 3 accuracy criteria. The first accuracy criterion is the relative information content (RIC), defined by

$$\mathrm{RIC}(K) = \frac{\sum_{k=K+1}^{\min(d,N)} \sigma_k^2}{\sum_{k=1}^{\min(d,N)} \sigma_k^2},$$

quantifies the relative amount of neglected information when truncating the number of modes at rank $K$. The truncation rank is determined such that the RIC is inferior to the accuracy threshold $\varepsilon$. The accuracy threshold $\varepsilon$ is fixed to $10^{-6}$.

The second accuracy criterion is the relative time residual $\mathcal{R}$. It quantifies the relative error induced by the minimization of the least square cost function (3.13) using $A_\mu$. It is given by

$$\mathcal{R}(j) = \frac{\|A_\mu \mathbb{X}_j - \mathbb{Y}_j\|^2}{\|\mathbb{Y}_j\|^2}$$

where $\mathbb{X}_j$ represents the $j^{th}$ column of $\mathbb{X}$ and $\mathbb{Y}_j$ the $j^{th}$ column of $\mathbb{Y}$. The index $j$ is thus linked to the snapshots ($j \in \{1, ..., N\}$). To better draw a parallel between the evolution of this parameter and the capsule dynamics, this parameter will be represented as a function of the non-dimensional time $Vt/\ell$ hereafter.

The third accuracy criteria $\varepsilon_{\text{Shape}}(Vt/\ell)$ measures the difference between the 3D reference capsule shape given by the FOM ($\mathcal{S}_{\text{FOM}}$) and the 3D shape predicted by the ROM ($\mathcal{S}_{\text{ROM}}$). It is defined at a given non-dimensional time $Vt/\ell$ as the ratio between the modified Hausdorff distance (MHD) computed between $\mathcal{S}_{FOM}$ and $\mathcal{S}_{ROM}$ and non-dimensionalized by $\ell$

$$\varepsilon_{\text{Shape}}(Vt/\ell) = \frac{\mathrm{MHD}(\mathcal{S}_{\text{FOM}}(Vt/\ell), \mathcal{S}_{\text{ROM}}(Vt/\ell))}{\ell}$$

The modified Hausdorff distance is the maximum value of the mean distance between $\mathcal{S}_{\text{FOM}}$ and $\mathcal{S}_{\text{ROM}}$ and the mean distance between $\mathcal{S}_{\text{ROM}}$ and $\mathcal{S}_{\text{FOM}}$ (Dubuisson & Jain 1994).

## 4. Numerical experimentation on a given configuration

The method is first applied to a given configuration, in order to set the model parameters and to study its stability and precision. We consider the dynamics of an initially spherical capsule flowing in a microchannel when $Ca = 0.13$ and $a/\ell = 0.8$. The time step between each snapshot $\Delta t$ equals to 0.04. The dynamics predicted by the FOM is illustrated in Fig. 3 up to a non-dimensional time $VT/\ell = 10$. As the capsule flows, its membrane is gradually deformed by the hydrodynamic forces inside the channel during a temporary time until a steady state is reached. We assume that the capsule has reached its steady-state shape, when the surface area of the capsule varies by less than $5 \times 10^{-4} \times (4\pi a^2)$ over a non-dimensional time $Vt/\ell = 1$. For ($Ca = 0.13$, $a/\ell = 0.8$), the steady state is reached at $VT_{SS}/\ell = 6.2$ and is characterized by a parachute capsule shape (Figure 3).

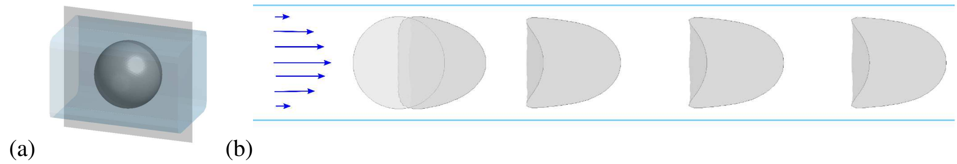


Figure 3: Dynamics of a microcapsule flowing in a microchannel with a square cross-section predicted by FOM in the vertical cutting plane represented in grey in (a). The in-plane capsule profiles are shown for $Ca = 0.13$ and $a/\ell = 0.8$ at the non-dimensional times $Vt/\ell = 0, 0.4, 2, 4, 6$ in (b). The horizontal lines on (b) represent the channel borders. The capsule will always be represented flowing from left to right.

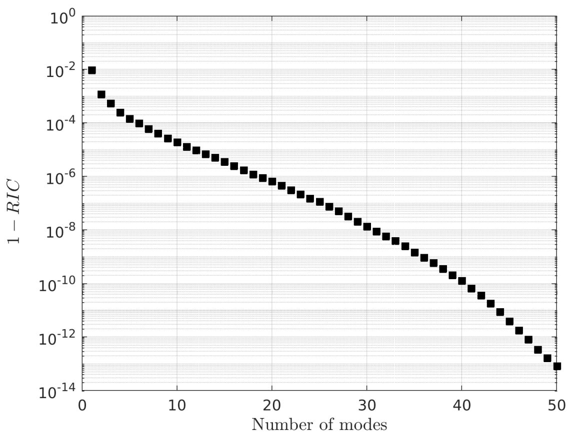


Figure 4: Evolution of the relative amount of neglected information 1-RIC, as a function of the number of modes ($Ca = 0.13, a/\ell = 0.8$).

### 4.1. *Proper orthogonal decomposition, truncation and modes*

The singular value decomposition is first applied to the displacement snapshot matrix. To determine the truncation rank, the evolution of 1- RIC is illustrated in Figure 4 as a function of number of modes considered. The RIC is about 1% only with one mode. The more modes is kept, the less information is neglected. In the following, we fix the number of modes to 20. The accuracy threshold $\varepsilon$ is thus equal to $10^{-6}$.

The modes are determined from the displacement snapshot matrix. They are added to the sphere of radius 1 and amplified by a factor 2 to be visualized. The first six modes are represented in Figure 5 for ($Ca = 0.13, a/\ell = 0.8$).

The first six modes are mostly dedicated to change the shape of the rear of the capsule. The following modes appear to become noisy (not shown). However, these modes are not negligible, if one wants to get an accuracy of $10^{-6}$.

### 4.2. *Dynamic Mode Decomposition: empirical regularization*

Before determining the matrix $A$, we check the condition number of the matrices $\mathbb{X}$ and $\mathbb{X}\mathbb{X}^{\mathbb{T}}$. They are respectively equal to $6.5 \times 10^4$ and $4.3 \times 10^9$. The condition numbers of these matrices are very high and the matrix $A$, determined by solving (3.16), may be sensitive to perturbations or noise. To improve the robustness, a Tikhonov regularization is applied to solve the least-square problem (3.13) and the matrix $A_\mu$ is computed using (3.18), which depends on the regularization coefficient $\mu$. To determine the optimal value of $\mu$, the relative least square error $\|A_\mu \mathbb{X} - \mathbb{Y}\|_F / \|\mathbb{Y}\|_F$ is represented according to the norm solution $\|A_\mu\|_F$

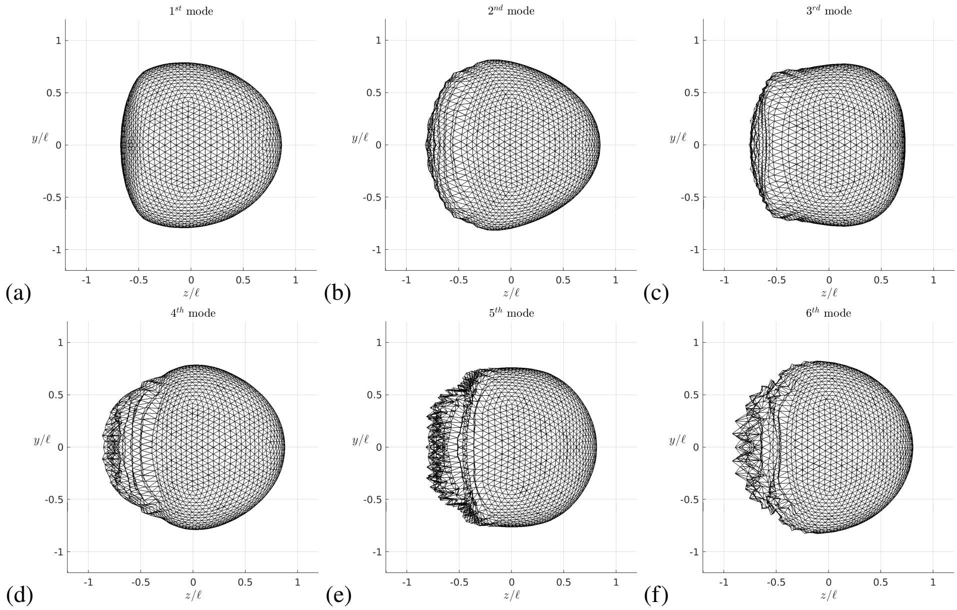


Figure 5: Representation of the first six modes of the capsule dynamics when $a/\ell = 0.80$ and $Ca = 0.13$. To be visualized the modes of displacement were added to the sphere of radius 1 and amplified by a factor 2.

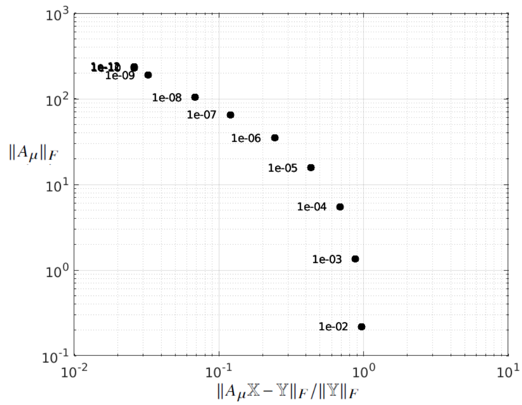


Figure 6: Evolution of the norm solution $\|A_\mu\|_F$ as a function of the least square error $\|A_\mu \mathbb{X} - \mathbb{Y}\|_F / \|\mathbb{Y}\|_F$ when the number of modes is fixed to 20 and ($Ca = 0.13$, $a/\ell = 0.8$).

when 20 modes are considered and when $\mu$ is varied between $10^{-12}$ and $10^{-3}$ (Figure 6). The least square error $\|A_\mu \mathbb{X} - \mathbb{Y}\|_F$ and the norm solution $\|A_\mu\|_F$ are minimal when $\mu = 10^{-9}$. In the following, $\mu$ is thus fixed to $\mu = 10^{-9}$ and the number of modes to 20.

### 4.3. *Validity check of the ROM: spectral study of the resulting matrix*

In order to detect anomalies, a spectral analysis of the reduced-order model learned by the DMD method is carried out. The spectrum of the matrix $A_\mu$ is represented in Figure 7. All the eigenvalues $\lambda_k$ of the matrix $A_\mu$ have non-positive real parts. The resulting linear ROM is thus stable.

The temporal evolution of the residual $\mathcal{R}$ (Figure 8) shows that the error is less than

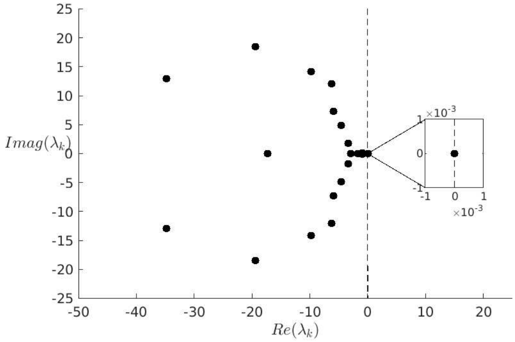


Figure 7: Eigenvalues $\lambda_k$ of $A_\mu$ ($Ca = 0.13$, $a/\ell = 0.8$, 20 modes, $\mu = 10^{-9}$). Note that the maximum real part of the eigenvalues is exactly equal to zero.

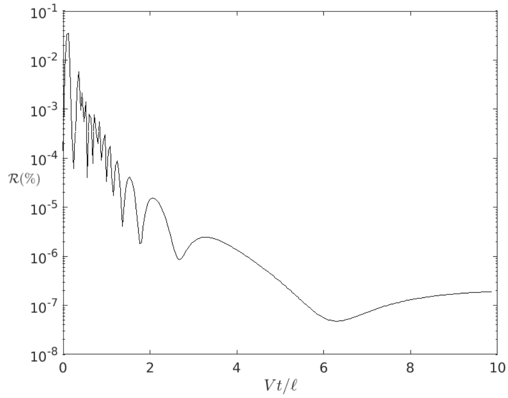


Figure 8: Temporal evolution of the normalized time residual with $Ca = 0.13$, $a/\ell = 0.8$, 20 modes and $\mu = 10^{-9}$.

0.7‰. The maximal value is reached at the beginning of the simulation ($Vt/\ell < 0.3$) and $\mathcal{R}$ decreases afterwards. When $Vt/\ell \lesssim 6$, i.e. before the capsule has reached its steady state, high frequency oscillations are observed. This probably means that a high frequency mode is neglected, even if 20 modes are considered. For $Vt/\ell > 6$, $\varepsilon_{ROM}$ is of order $10^{-9}$. The stationary state is thus well predicted by the model and the error during the transient stage is more than acceptable.

### 4.4. *ROM online stage and accuracy assessment*

The displacement of all the nodes of the capsule mesh estimated by the ROM is then added to the corresponding node of the sphere of radius 1 to visualize the temporal evolution of the capsule shape in three dimensions. Figure 9 shows the capsule dynamics for the reference case ($Ca = 0.13, a/\ell = 0.8$). The ROM allows us to reproduce the capsule deformation from the initial state up to the parachute-shaped steady state. For the FOM and the ROM, the capsule profile is then determined in the cutting plane passing through the middle of the microchannel. Figure 10 shows that the two capsule profiles perfectly overlap at $Vt/\ell = 0, 0.4, 2, 4, 6$. The temporal evolution of $\varepsilon_{\text{Shape}}$ is shown in Figure 11a. The maximum value of the error

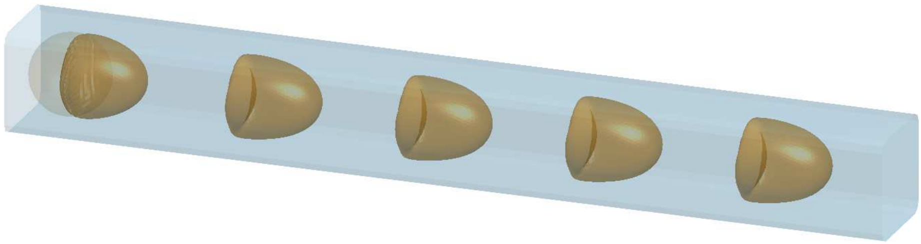

Figure 9: Dynamics of a microcapsule flowing in microchannel with a square cross-section predicted by ROM at the non-dimensional time $Vt/\ell = 0.4, 2.8, 5.2, 7.6, 10$ with $Ca = 0.13$, $a/\ell = 0.8$, 20 modes and $\mu = 10^{-9}$. The initial spherical capsule is shown on the left by transparency.

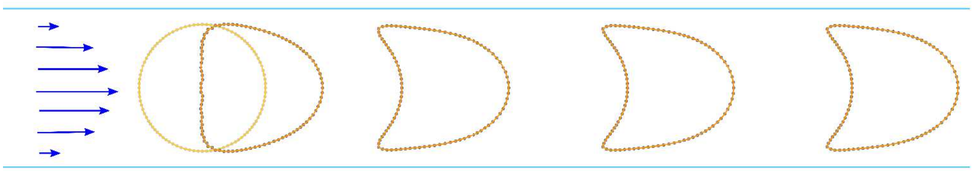

Figure 10: Comparison of the capsule contours given by the FOM (dotted line) and estimated by the ROM (orange line). The capsule is shown for ($Ca = 0.13$, $a/\ell = 0.8$) at the non-dimensional time $Vt/\ell = 0, 0.4, 2, 4, 6$. The horizontal lines represent the channel borders. The number of modes is fixed to 20 and $\mu = 10^{-9}$.

committed on the 3D shape $\varepsilon_{\text{Shape}}$ equals to 0.004%. The error on the capsule shape $\varepsilon_{\text{Shape}}$ is thus negligible. The deformation of the capsule from its initially spherical shape to its steady state over an non-dimensional time $Vt/\ell = 10$ can thus be estimated very precisely with the developed reduced-order model.

The DMD method predicts the capsule displacement at time $t^{n+1}$ from that at time $t^n$. The model has been constructed until now by considering the dynamics of the capsule over a non-dimensional time $Vt/\ell$ of 10.

In order to study the sensitivity of the ROM on the learning time $VT_L/\ell$, i.e. the non-dimensional time over which the model is trained, we modify it with values between 2 and 8, knowing that the time to reach the steady state is in this case $VT_{SS}/\ell = 6.2$. We estimate the capsule dynamics using the ROM model up to a non-dimensional time $Vt/\ell$ of 10. The number of modes is always equal to 20 and $\mu = 10^{-9}$. The comparison of the estimated shape at $Vt/\ell = 10$ with the one simulated with the FOM (Figure 11a) shows that $VT_L/\ell \geqslant 4$ is sufficient to predict very well the capsule dynamics. The Figure 11b confirms that the error on the capsule shape is negligible when $VT_L/\ell \geqslant 4$. It is interesting that the ROM model could predict the steady state even when $T_L < T_{SS}$. This could be due to the fact that the maximum real part of the eigenvalues is exactly equal to zero from $VT_L/\ell \geqslant 4$. The maximum real part of the eigenvalues is negative and close to zero for $VT_L/\ell = 2$. Figure 11c shows that, in the worst case ($VT_L/\ell = 2$), the error on the capsule shape increases with time but reaches a plateau from a time $Vt/\ell$ of 8. The system is thus stable without exponential drift as proven by the negative value of the maximum real eigenvalue. In the zoom insert, the error also increases for the other learning times but remains very small (below 0.2%).

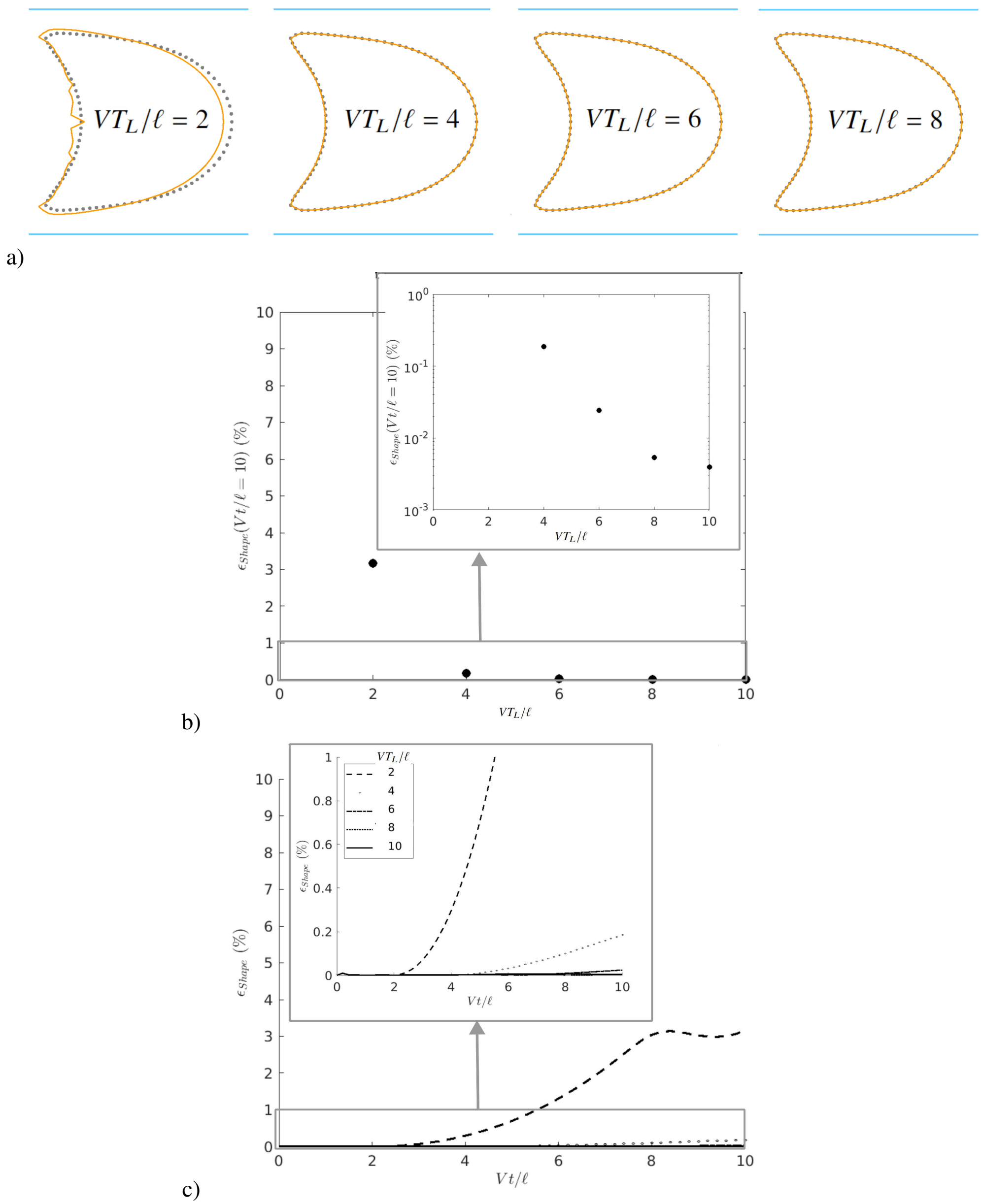


Figure 11: a) Comparison of the capsule contours given by the FOM (dotted line) and estimated by the ROM (orange line) for the different learning times $VT_L/\ell$. b) Evolution of $\varepsilon_{\text{Shape}}$ measured at $Vt/\ell = 10$ as a function of the learning time $VT_L/\ell$. c) Influence of the learning time $VT_L/\ell$ on the temporal evolution of the error on the capsule shape $\varepsilon_{\text{Shape}}$. The error during the learning time is shown in solid line. For this case, the parameters are 20 modes, $\mu = 10^{-9}$, $Ca = 0.13$ and $a/\ell = 0.8$.

## 5. Space-time ROM accuracy assessment over the full parameter sample set

The capillary number $Ca$ and the size ratio $a/\ell$ are now considered as variable parameters. A database of 119 simulations of the deformation of an initially spherical capsule in a microchannel has been generated using the FOM with the same time step and mesh size as in

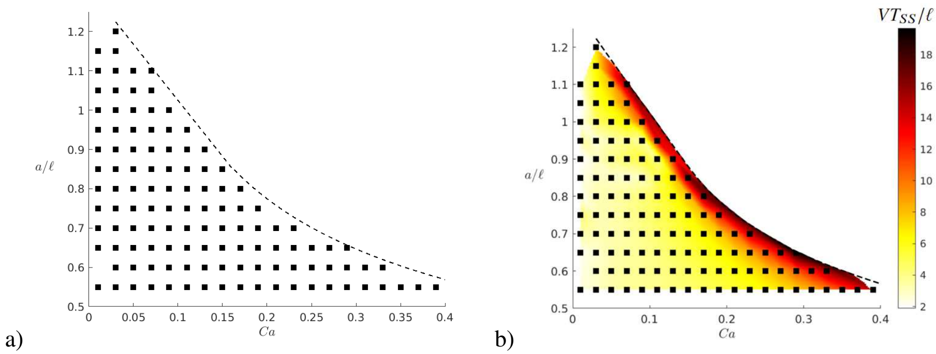


Figure 12: a) Values of $Ca$ and $a/\ell$ included in the training database. b) Evolution of the time $VT_{SS}/\ell$ needed to reach the steady state, on the training database. The dotted line delimits the domain where a steady-state capsule deformation exists for capsules following the neo-Hookean law.

section 4. Figure 12a shows the different values of $Ca$ and $a/\ell$ for which the simulations have been computed to create the training database. When the capsule initial radius is close to or larger than the microchannel cross-dimension ($a/\ell \geqslant 0.90$), the capsule is pre-deformed into a prolate spheroid to fit in the channel. For a given $a/\ell$, a limit value of $Ca$ exists beyond which a capsule does not reach a steady-state (Figure 12). This is due to the softening behavior of the neo-Hookean law.

For the following, we have considered a learning time $VT_L/\ell = 10$. The evolution of the time $VT_{SS}/\ell$ needed to reach the steady state is illustrated in Figure 12b on the whole training database. The steady state is reached on average at a time $VT_{SS}/\ell$ of 6.2. However, we notice that for the cases close to the steady state limit, $VT_{SS}/\ell$ increases and exceeds the considered learning time.

For all the couples $(Ca, a/\ell)$ of the training database, the capsule shape is reconstructed from the ROM results at given non-dimensional times and compared to the shape predicted by the FOM at the same non-dimensional time. The evolution of the error committed on the capsule shape $\varepsilon_{Shape}$ on the full database is illustrated in Figure 13 at $Vt/\ell = 0, 0.4, 1, 2, 5, 10$. $\varepsilon_{Shape}$ is null at $Vt/\ell = 0$. The ROM is therefore able to predict the initial capsule shape correctly, whether it is spherical or slightly ellipsoidal. Until $Vt/\ell \leqslant 2$, $\varepsilon_{Shape}$ essentially remains zero on the majority of the database. Otherwise, it is equal to 0.15% at maximum. At $Vt/\ell = 5$ and 10, the error $\varepsilon_{Shape}$ slightly increases for most of the couples $(Ca, a/\ell)$ of the database. It remains fully acceptable since it is equal to 0.35% at maximum. When considering 20 modes and $\mu = 10^{-9}$, the developed ROM allows us to estimate with great precision the dynamics of an initially spherical capsule in a microchannel with a square cross-section.

To respect the stability condition (see Equation 2.10), the time step imposed to simulate the capsule dynamics with the FOM decreases, when the $Ca$ decreases. The lower the $Ca$, the longer the simulation lasts (Figure 2). The time needed to calculate the capsule shape and write the results was estimated on the same workstation used to simulate and generate the result files with the FOM (2-CPU Intel® Xeon® Gold 6130, 2.1 GHz). The speedup is the ratio between the FOM runtime and the ROM runtime. Its evolution according to the FOM time step is illustrated in Figure 14. It was estimated from the ROM and FOM simulation time obtained when $a/\ell = 0.7$. The speedup varies between 52106 for a FOM time step of $10^{-4}$ (i.e for the lowest value of $Ca$ tested) and 4200 for $5 \times 10^{-4}$ (i.e $Ca \geqslant 0.05$). It is thus

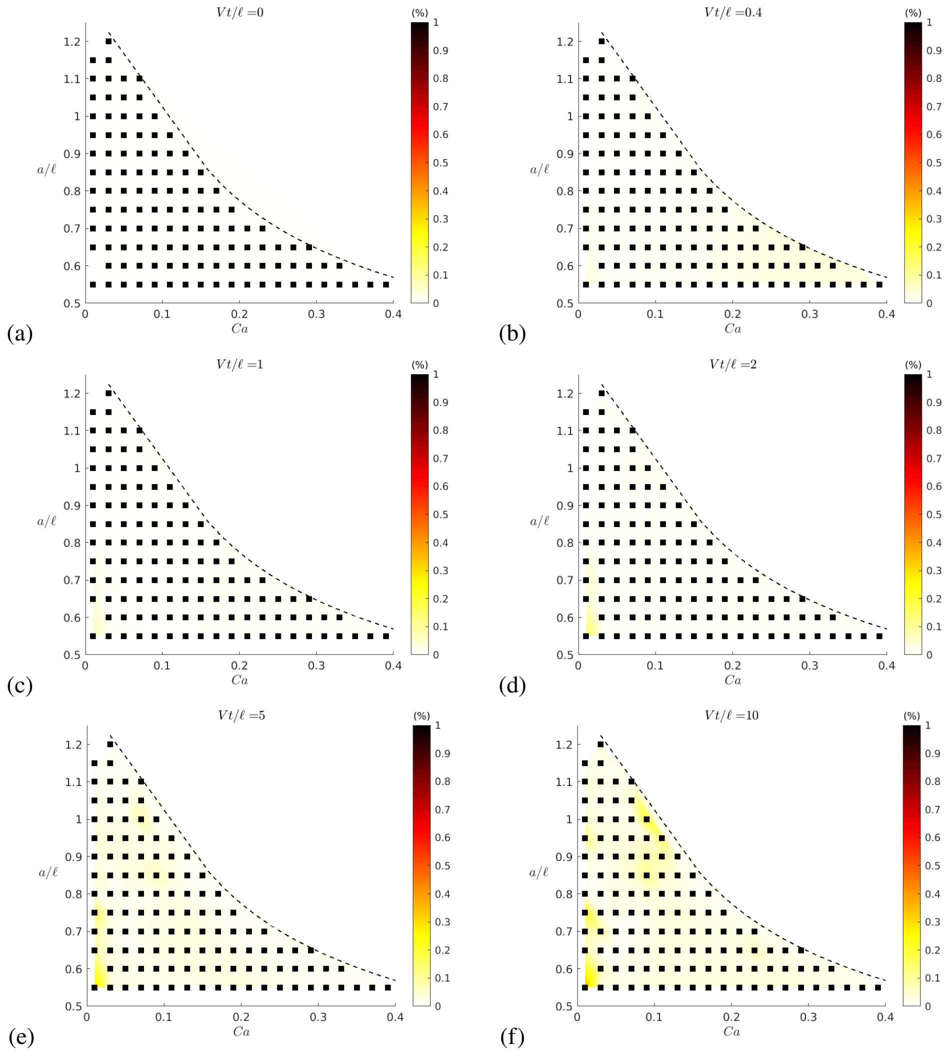


Figure 13: Heat maps of $\epsilon_{Shape}$ on the training database as a function of $Ca$ and $a/\ell$ at (a) $\dot{\gamma}t$=0, (b) 0.4, (c) 1, (d) 2, (e) 5, (f) 10 (obtained with 20 modes and $\mu = 10^{-9}$). The dotted line delimits the domain where a steady-state capsule deformation exists.

possible to estimate the capsule dynamics very precisely with the developed ROM, while considerably reducing the computational time.

Another significant advantage is the gain in storage of the simulation results. By storing only the reduced variables $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, the modes $\{\phi_k\}$ and the initial position of the nodes of each couple $\theta = (Ca, a/\ell)$, the training database is reduced from 1.9 GB, when computed with the FOM, to 0.15 GB with the ROM. It can therefore be more easily shared.

## 6. Full space-time-parameter ROM (for any admissible parameter value)

### 6.1. *General methodology*

It is here again assumed that a training database of $N$ precomputed FOM results is available. Now we would like to derive a ROM for any parameter couple $\boldsymbol{\theta} = (Ca, a/\ell)$ in the admissible

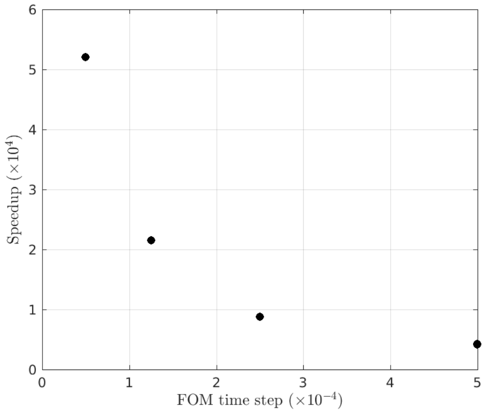


Figure 14: Evolution of the speedup as function to the time step imposed to simulate the capsule dynamics with the FOM ($a/\ell = 0.7$).

parameter domain. The proposed space-time-parameter ROM is made of two steps. The first step consists in predicting the space-time solution $\{\boldsymbol{u}\}(t;\boldsymbol{\theta})$ by means of a robust interpolation procedure. The second step consists in deriving a ROM in the form of a low-order dynamical system by using the predicted solutions of the first step as training data. Then we apply the former procedure detailed in Section 3. Below we give a detailed explanation of the two steps.

**Step 1: predictor step.** Considering a parameter couple $\boldsymbol{\theta}$, we first search the three nearest neighbour parameters in the sample set that form a non-degenerate triangle in the plane $(Ca, a/\ell)$. Let us denote them by $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$. We will define a linear operator in the triangle $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$. For that, let us introduce the barycentric coordinates $(\lambda_1, \lambda_2, \lambda_3)$, $\lambda \in [0, 1]$, $i = 1, 2, 3$ such that

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \tag{6.1}$$

$$\boldsymbol{\theta}_1\lambda_1 + \boldsymbol{\theta}_2\lambda_2 + \boldsymbol{\theta}_3\lambda_3 = \boldsymbol{\theta}. \tag{6.2}$$

The $3 \times 3$ linear system (6.1),(6.2) is invertible as soon as the triangle $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$ is non-degenerate. Notice that the $\lambda_i$ ($i = 1, 2, 3$) are actually functions of $\boldsymbol{\theta}$. Let us now denote by $\{u_1\}$, $\{u_2\}$ and $\{u_3\}$ the displacement fields for the parameter vectors $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ respectively. Then we can consider the predicted velocity field $\hat{\boldsymbol{u}}(t;\boldsymbol{\theta})$ defined by

$$\{\hat{u}\}(t, \boldsymbol{\theta}) = \lambda_1\{u_1\}(t) + \lambda_2\{u_2\}(t) + \lambda_3\{u_3\}(t). \tag{6.3}$$

**Step 2: low-order dynamical system ROM.** Expression (6.3) can be evaluated at some discrete instants in order to generate new training data. Then the SVD-DMD ROM methodology presented in Section 3 can be applied to these data to get a reduced dynamical system in the form

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \boldsymbol{\beta}(\boldsymbol{\theta}),$$
$$\dot{\boldsymbol{\beta}}(\boldsymbol{\theta}) = A_\mu(\boldsymbol{\theta})\,\boldsymbol{\beta}(\boldsymbol{\theta}).$$

We also have a matrix $Q(\boldsymbol{\theta})$ of orthogonal POD modes and we can go back to the high-dimensional physical space by the standard operations

$$\{\hat{u}\}(t, \boldsymbol{\theta}) \approx Q(\boldsymbol{\theta})\,\boldsymbol{\alpha}(t, \boldsymbol{\theta}), \quad \{\hat{v}\}(t, \boldsymbol{\theta}) \approx Q(\boldsymbol{\theta})\,\boldsymbol{\beta}(t, \boldsymbol{\theta}). \tag{6.4}$$

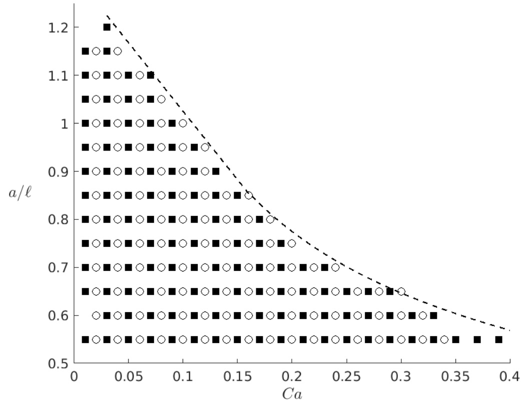


Figure 15: Values of $Ca$ and $a/\ell$ included in the testing database (open circle). The filled squares represent the cases in the training database. The dotted line delimits the domain where a steady-state capsule deformation exists for capsules following the neo-Hookean law.

Notice that the capsule position field $\{\boldsymbol{x}\}(t, \boldsymbol{\theta})$ is given by

$$\{x\}(t; \boldsymbol{\theta}) = \{X\}(\boldsymbol{\theta}) + \{\hat{u}\}(t, \boldsymbol{\theta})$$

with an initial capsule position $\{X\}(\boldsymbol{\theta})$ that may depend on $\boldsymbol{\theta}$ because of the pre-deformation preprocessing if $a/\ell \geqslant 0.95$.

A testing database is created using the FOM as in Section 5 and considering $(Ca, a/\ell)$-couples which are not in the training database. A set of 110 $(Ca, a/\ell)$-couples are included in this database (Figure 15). For all the $(Ca, a/\ell)$-couples of the testing database, the capsule dynamics is interpolated from the dynamics of the 3 closest neighbours at a given non-dimensional time. Capsule shapes obtained by the ROM are compared to the ones predicted by the FOM at the same non-dimensional time. Figure 16 represents the evolution of the error committed on the capsule shape $\varepsilon_{Shape}$ on the training database at $Vt/\ell = 0,\ 0.4,\ 1,\ 2,\ 5,\ 10$. At initial time, $\varepsilon_{\text{Shape}}$ is zero. The interpolation method is therefore able to capture the initial capsule shape. When the time increases, $\varepsilon_{\text{Shape}}$ increases and greater than if we apply directly the POD-DMD method on the FOM results and reconstruct the dynamics. However, $\varepsilon_{\text{Shape}}$ remains less than 0.3% on the majority of the testing database. It remains fully acceptable. $\varepsilon_{\text{Shape}}$ is more important near the steady-state limit and when we approach the lowest values of $Ca$ because we are close to the limits of the training base.

## 7. Application of the ROM to a capsule in simple shear flow

To prove the generality of the proposed approach, we additionally apply the ROM to a capsule in simple shear flow. This classical case was extensively studied over the past years (Ramanujan & Pozrikidis 1998; Lac & Barthès-Biesel 2005; Li & Sarkar 2008; Walter et al. 2010; Foessel et al. 2011; Barthès-Biesel et al. 2010; Dupont et al. 2015). The FOM results of an initially spherical capsule subjected to a shear rate $\dot{\gamma}$ are simulated using the unconfined version of the boundary integral - finite element method presented in section 2 (see Walter et al. (2010) for a detailed description of the method). The time step $\Delta t$ between each snapshot is equal to 0.04.

We first build a ROM model that predicts the capsule dynamics until $\dot{\gamma}t = 10$ with 15 modes, a learning time of $\dot{\gamma}T_L = 10$ and $\mu = 10^{-6}$. We retrieve that the initial spherical

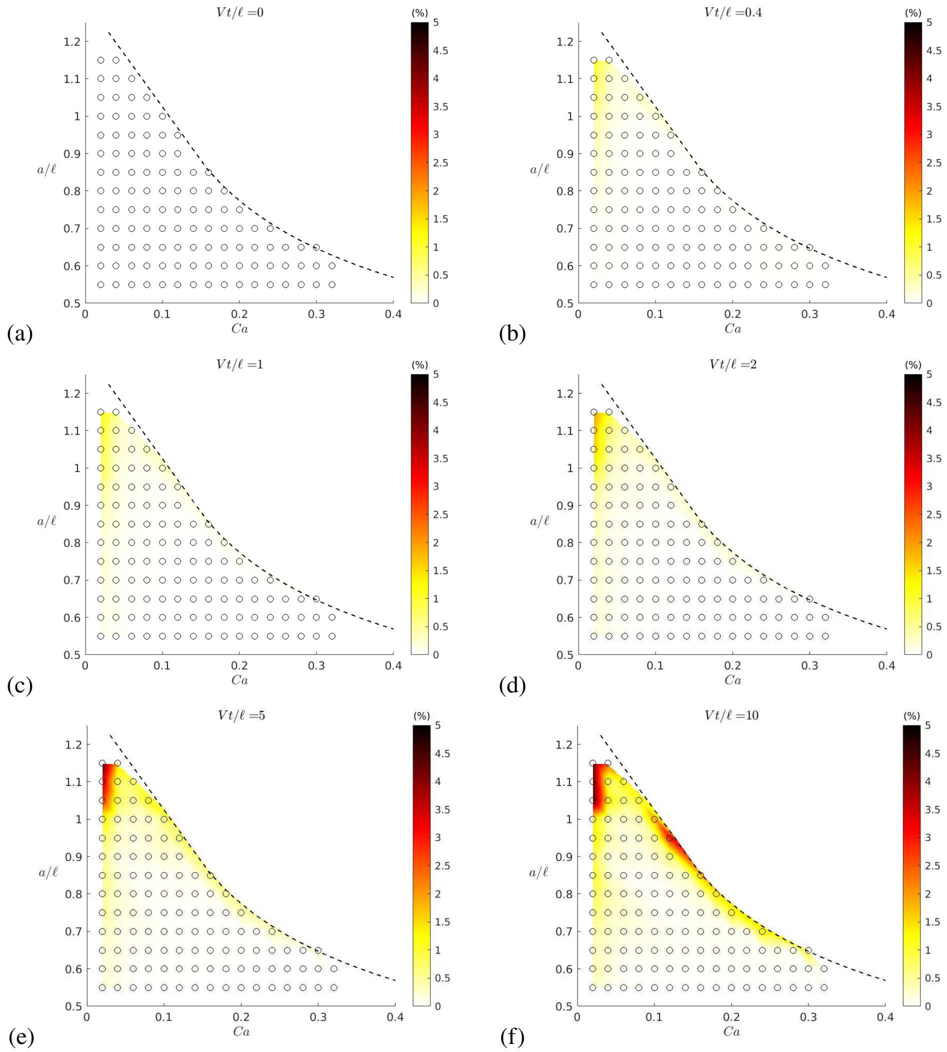


Figure 16: Heat maps of $\varepsilon_{\text{Shape}}$ on the testing database as a function of $Ca$ and $a/\ell$ at (a) $\dot{\gamma}t$=0, (b) 0.4, (c) 1, (d) 2, (e) 5, (f) 10. The dotted line delimits the domain for which a steady-state capsule deformation exists.

capsule elongates under the effect of the external flow in the straining direction and that the membrane rotates around the deformed shape due to the flow vorticity (Figure 17). The ROM is thus able to recover the tank-treading motion. A very good agreement between the ROM and FOM is seen in Figure 18 for the capsule profiles in the shear and perpendicular planes. Figure 19 shows the evolution of the maximum error on the capsule shape for different values of $Ca$. At $Ca$ = 0.1, the ROM model predicts well the time evolution of the global capsule shape but not precisely the wrinkle formation, leading to 2% error on average. But from $Ca \geqslant 0.3$, the error is reduced by an order of magnitude and is below 0.2%.

We then perform some tests to be sure that the model is able to predict the tank-treading motion correctly after the learning time. Since the period is equal to 17.6 for $Ca$ = 0.3, the learning time $T_L$ = 10 appears to be too short to capture the periodical motion. We consider a (safe) learning time $T_L$ = 20 and increase the number of modes to 60 to capture the

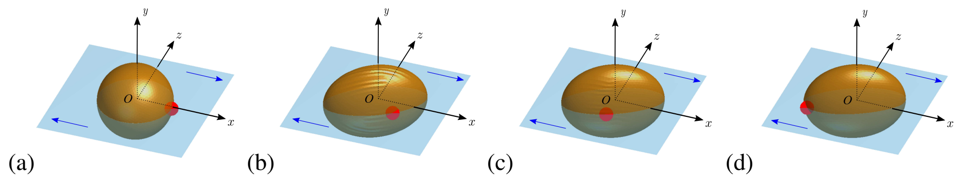

Figure 17: Snapshots of a capsule subjected to a simple shear flow estimated by the ROM ($Ca = 0.3$, 15 modes and $\mu = 10^{-6}$): $\dot{\gamma}t = 0$ (a), 1.6 (b), 4.8 (c), 6.4 (d). A red point is placed on the membrane to visualize the tank-treading motion.

Lagrangian motion of the mesh along the capsule (Eulerian) steady shape. This convection-dominated motion of the capsule is known to be an unfavourable condition for dimensionality reduction and this is the reason why it is adequate to increase the number of modes. We have obtained the best tradeoff between accuracy, numerical conditioning and complexity using 60 modes.

The error on the 3D shape $\varepsilon_{\text{Shape}}$, represented in Figure 20a, does not exceed 2%. Indeed, after a quasi-monotonic increase, it reaches a value of 1.7 at the end of the learning time ($\dot{\gamma}t \leqslant 20$) and remains almost constant during the extended prediction time ($20 < \dot{\gamma}t \leqslant 30$). This is very comforting for long-time stability and accuracy of the simulation. Furthermore, we study the spectral structure of the matrix $A_\mu$ and plot its eigenvalues in the complex plane in Figure 20b. All the eigenvalues have non-positive real parts showing the asymptotic stability property of the dynamical system.

One may still wonder whether the DMD-ROM model is accurate only for capsule flows that converge towards a steady state. To answer the question, we have investigated the feasibility of applying the method to an initially ellipsoidal capsule in simple shear flow. Depending on the parameters, such a capsule exhibits a variety of dynamical regimes, which are periodical in many cases (Walter et al. 2011; Dupont et al. 2013, 2016). We apply the ROM model to the full dataset of FOM simulations for the same initial capsule. It is thus a case where the ergodicity hypothesis cannot be applied to improve the filling of the state space (see Tu et al. (2014)).

At large $Ca$, when the capsule exhibits a fluid-like behaviour, a large number of modes is required to capture the membrane rotation around the deformed shape. When the capsule behaves like a solid particle at low $Ca$ and exhibits a tumbling motion, it is preferable to place the capsule within its own reference frame before applying the ROM method. The error is typically of a few percent and the capsule motion is well reproduced. An example of the tumbling dynamics predicted by the ROM is compared to the one simulated by the FOM in Figure 21. The ROM is able to reproduce more complex capsule dynamics (e.g. with periodical motion) and to capture deformation features including wrinkles, all this with a speedup of about 35 000.

## 8. Discussion and conclusion

As a summary, in this paper we have considered a $\boldsymbol{\theta}$-parametrized reduced-order model of microcapsule dynamics in the form

$$\dot{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\beta}(\boldsymbol{\theta}),$$
$$\dot{\boldsymbol{\beta}}(\boldsymbol{\theta}) = A_\mu(\boldsymbol{\theta})\,\boldsymbol{\beta}(\boldsymbol{\theta}).$$

The vector $\boldsymbol{\theta} = (Ca, a/\ell)$ contains the governing parameters, the coefficients $\alpha_k(t,\boldsymbol{\theta})$ and $\beta_k(t,\boldsymbol{\theta})$ are spectral coefficients of POD decomposition for the displacement and velocity

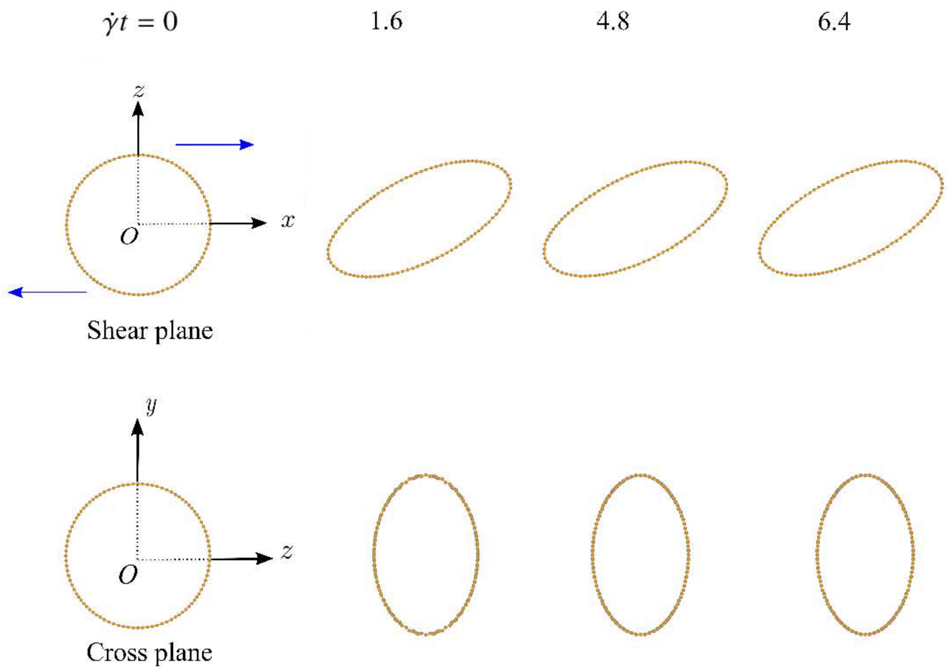


Figure 18: Capsule subjected to a simple shear flow for $Ca = 0.3$: Comparison of the contours in the shear and cross planes given by the FOM (dotted line) and estimated by the ROM (orange line, obtained with 15 modes and $\mu = 10^{-6}$).

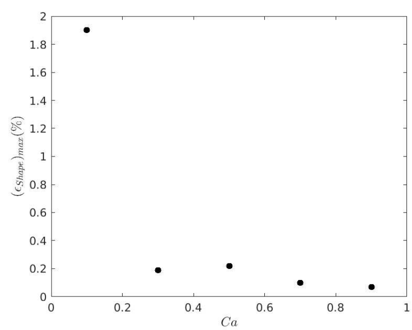


Figure 19: Evolution of the maximum error committed on the shape of a capsule subjected to a simple shear flow as a function of the capillary number $Ca$ (obtained with 15 modes and $\mu = 10^{-6}$). The capsule dynamics was simulated up to a non-dimensional time $\dot{\gamma} = 10$.

fields respectively, and the matrix $A_\mu(\boldsymbol{\theta})$ is identified from data using a dynamic mode decomposition least-square procedure. We have numerically proven for a broad range of capillary numbers $Ca$ and size ratios $a/\ell$ that it is able to capture the dynamics up to the steady state of a capsule flowing in a channel and its large deformations. As a first approach, we have presently chosen to use a DMD method that is linear in time to build the ROM model. Still the ROM model captures spatial non-linearity by means of the POD modes. The resulting reduced-order model is of great fidelity, weak discrepancies being only observed in the early transient stage. We have also shown that the learning time need to be larger than

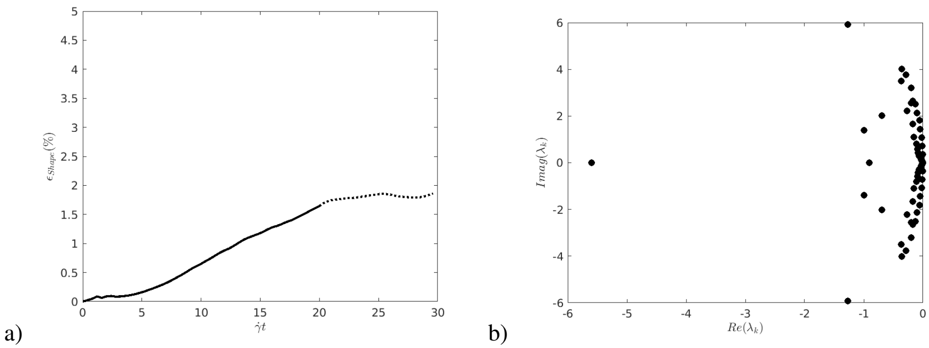


Figure 20: a) Evolution of the error committed on the shape of a capsule subjected to a simple shear flow during the learning time (full line) and the extended prediction time (dotted line). b) Representation of the eigenvalues of $A_\mu$ when 60 modes, $\mu = 10^{-6}$ and $Ca = 0.3$ are considered.

the transient stage duration and that we can go beyond the FOM time window used for the training of the ROM model.

For generalisation, we have computed the capsule dynamics for any parameter set. The generalization algorithm is based on interpolation: we first pre-calculate the ROM dynamic model at a finite number of points in the parameter space domain and determine the $\alpha$, $\beta$ and $\phi_k$ (and thus the capsule displacement) at these points. For any other value of the parameters, we first predict the time-evolution of the capsule node displacements using a linear interpolation procedure in the parameter space and then build a dynamical system based the DMD methodology. The error is mostly below 0.3% over the entire domain, which proves the precision and utility of the ROM approach.

Like any other data-driven model, the model requires a certain number of high-fidelity simulations to provide accurate predictions. By discretizing the parameter space in a regular and homogeneous way (Figure 12), we have not presently tried to optimize the number of FOM simulations. But sampling strategies like the Latin Hypercube Sampling (LHS) exist and result in a net reduction in FOM simulation number. The empirical law, conventional among the data-driven model community, is that one needs between $10 \times D$ and $50 \times D$ points, where $D$ is the dimension of the problem ($D = 2$ in our case). This law shows that the number of high-fidelity simulations does not explode with the problem dimension, owing to the linear dependence of the law.

The linear differential model is stable as soon as the eigenvalues of $A_\mu$ have non-positive real parts, and is consistent with steady states as soon as zero is an eigenvalue. Numerical experiments show that identified matrices $A_\mu$ from data have eigenvalues with negative real parts and one of the eigenvalues is very close to zero.

As it is often the case with spectral-like methods, there is a trade-off between accuracy and ill-conditioning effects: when a large number of POD modes are used ($K > 20$), the data matrix $\mathbb{X}$ of snapshot POD coefficients is ill-conditioned. For the determination of $A_\mu$, we have used a Tikhonov regularization in the least square cost function (see (3.17)) in order to have a better conditioned problem and a $L$-curve procedure to determine the best regularization coefficient $\mu$. Unfortunately we observe some limitations in the accuracy. A perspective would be to use a proximal approach: within an iterative procedure, at

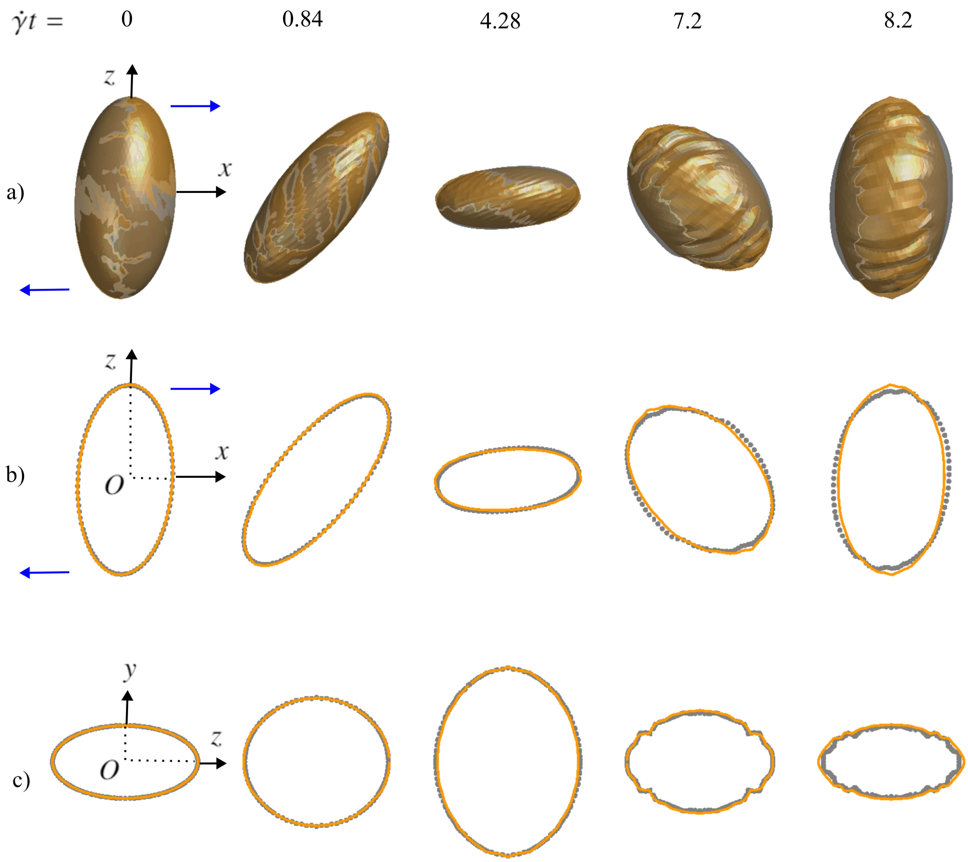


Figure 21: Tumbling motion of a prolate capsule (aspect ratio = 2) subjected to a simple shear flow ($Ca = 0.1$): a) Comparison of the 3D shape given by the FOM (in gray) and estimated by the ROM (in orange, obtained with 50 modes and $\mu = 10^{-6}$). Comparison of the 2D profil b) in the shear plane and c) in the crossed plane. The time step $\Delta t$ between each snapshot is equal to 0.04.

iteration $(p+1)$, compute the matrix $A_\mu^{(p+1)}$ solution of

$$A_\mu^{(p+1)} = \arg \min_{A \in \mathscr{M}_K(\mathbb{R})} \frac{1}{2} \|\mathbb{Y} - A\mathbb{X}\|_F^2 + \frac{\mu}{2} \|\mathbb{X}\|_F^2 \, \|A - A_\mu^{(p)}\|_F^2$$

using $A_\mu^{(0)} = 0$. At convergence, one can observe that the regularization term vanishes, so that one can expect better accuracy with this approach. This will be investigated in a future work.

We have proposed a successful and very efficient ROM for FSI problems. It is an alternative to the use of HPC. It must be seen as a complimentary (and non-competing) approach to full-order models, and has many advantages. Among them, one can mention the easiness in implementation. It leads to a very handy set of ODEs, that are easy to determine from an algorithmic point of view. Furthermore, the system can be run on any computer. The size of the matrices is, indeed, reduced from ($3 \times 2562$ nodes $\times 250$ snapshots) to about ($3 \times 2562$ nodes $\times (K+1)$), where the number of modes is $K = 20$. The computation required time is a few milliseconds for one parameter set. The current speedups are between 5 000 and 52 000, which out-performs any full-order model approach. We believe that this work is an encouraging milestone to move toward real time simulation of general coupled problems and to deal with high-level parametric studies, sensitivity analysis, optimization and uncertainty quantification.

The next milestone following this work would be to go toward non-linear differential dynamical systems as reduced-order models. There is three natural ways for that. The first one is to use Kernel Dynamic Model Decomposition (KDMD) rather than DMD. But we have recently shown in De Vuyst et al. (2022) that a non-linear low-order dynamical model does not provide significant improvement. The second one is to use Extended Dynamic Model Decomposition (EDMD) (Williams et al. 2015). The EDMD method adds some suitable non-linear observables (or features) in the data, so that a linear 'augmented' dynamical system is searched for. A third option is would be to directly use artificial neural networks (ANN), in particular recurrent neural networks (RNN) (Trischler & D'Euleuterio 2016). The RNN training would replace the DMD procedure, and would be trained with the same POD coefficient matrices $\mathbb{X}$ and $\mathbb{Y}$. As shown in the recent study by Lin et al. (2021), artificial intelligent may prove to be efficient and precise to predict capsule deformation.

**Acknowledgements.** The authors warmly thank Prof. Pierre Villon for fruitful discussions on model-order reduction and related topics.

**Funding.** This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. ERC-2017-COG - MultiphysMicroCaps).

**Declaration of interests.** The authors report no conflict of interest.

**Author ORCID.** C. Dupont, https://orcid.org/0000-0002-7727-3846; F. De Vuyst, https://orcid.org/0000-0003-0854-4670; A.-V. Salsac, https://orcid.org/0000-0001-8652-5411

**Author contributions.** A.V.S and F.D.V. created the research plan and formulated the numerical problem. C.D. implemented the numerical method and performed the tests. All authors contributed to analysing data and reaching conclusions, and in writing the paper.

## Appendix A. Effects of time derivative discretization on matrix estimation

In section 3.4, we explain how to identify the coefficient matrix $A$ from a least square problem that tries to minimize the squared residual $\int_0^T \|\dot{\boldsymbol{\beta}}(t) - A\,\boldsymbol{\beta}(t)\|^2\,dt$. For practical reasons and because of a finite number of data, we have to discretize the functional and in particular the time derivatives by means of finite differences. This section is dedicated to the analysis of the effect of discretization on the estimation on $A$, and in particular on the effect on the spectrum of $A$ and the impact on the stability of the identified model.

The notations here are specific to this section. Suppose we have a reference linear dynamical system whose equations and initial data are respectively

$$\dot{\boldsymbol{v}} = A^{ref}\,\boldsymbol{v},\ t \in [0,T], \quad \boldsymbol{v}(0) = \boldsymbol{v}^0 \in \mathbb{R}^K,$$

where $A^{ref} \in \mathscr{M}_K(\mathbb{R})$. The solution of the differential problem problem is given by $\boldsymbol{v}(t) = \exp(A^{ref}\,t)\boldsymbol{v}^0$, $t \in [0,T]$. Suppose that we don't know $A^{ref}$ but we have access to the exact solutions $\boldsymbol{v}^n = \boldsymbol{v}(t^n)$ at discrete times $t^n = n\Delta t$, $n \in \{0, ..., N\}$ where with $\Delta t = T/N$. The $(\boldsymbol{v}^n)_n$ will be used as data for the identification (estimation) of the matrix $A^{ref}$. Consider the least square minimization problem

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\sum_{i=0}^{N-1}\left\| \frac{\boldsymbol{v}^{n+1}-\boldsymbol{v}^n}{\Delta t} - A\boldsymbol{v}^n \right\|^2. \tag{A 1}$$

Since $\boldsymbol{v}^n = \exp\big(A^{ref} n\Delta t\big)\boldsymbol{v}^0$ for all $n$, we have also $\boldsymbol{v}^{n+1}-\boldsymbol{v}^n = \exp\big(A^{ref}\Delta t\big)\boldsymbol{v}^n$. So (A 1) is equivalent to

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\sum_{i=0}^{N-1}\left\| \left(\frac{\exp\big(A^{ref}\Delta t\big)-I}{\Delta t} - A\right)\boldsymbol{v}^n \right\|^2 = \min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\left\| \left(\frac{\exp\big(A^{ref}\Delta t\big)-I}{\Delta t} - A\right)\mathbb{X} \right\|_F^2$$

with $\mathbb{X} = [\boldsymbol{v}^0, \boldsymbol{v}^1, \ldots, \boldsymbol{v}^{N-1}] \in \mathscr{M}_{KN}(\mathbb{R})$. The first-order optimality conditions are

$$A\,\mathbb{X}\mathbb{X}^T = \left(\frac{\exp\big(A^{ref}\Delta t\big)-I}{\Delta t}\right)\mathbb{X}\mathbb{X}^T.$$

As soon as $\mathbb{X}$ is a full-rank matrix (meaning that $N \geqslant K$ and we reasonably have $K$ linearly independent measurements of $\boldsymbol{v}^n$), the matrix $\mathbb{X}\mathbb{X}^T$ is invertible and we get the estimate

$$A = \frac{\exp\big(A^{ref}\Delta t\big)-I}{\Delta t}. \tag{A 2}$$

Let us denote by $\lambda_k^{ref}$ (resp. $\lambda_k$) the (complex) eigenvalues of $A^{ref}$ (resp. $A$). We have $\lambda_k = \dfrac{e^{\lambda_k^{ref}\Delta t}-1}{\Delta t}$. Suppose now that we use a small time step $\Delta t$. From a Taylor expansion, we observe that

$$\lambda_k = \lambda_k^{ref} + \frac{\Delta t}{2}(\lambda_k^{ref})^2 + o(\Delta t).$$

We would like to study what is the effect of the first-order error term $\frac{\Delta t}{2}(\lambda_k^{ref})^2$ on the stability of the reconstructed dynamical system $\dot{\boldsymbol{v}} = A\boldsymbol{v}$. Suppose that the complex number $\lambda_k^{ref}$ has real and imaginary parts $a$ and $b$ respectively. Then

$$\lambda_k = \left(a + \frac{\Delta t}{2}(a^2-b^2)\right) + ib(1+a\Delta t) + o(\Delta t).$$

If $a = \mathrm{Re}(\lambda_k^{ref}) \leqslant 0$, what are the conditions to keep $\mathrm{Re}(\lambda_k) \leqslant 0$ ? We consider two cases:

•If $a = 0$ (with $b \neq 0$), $\lambda_k^{ref}$ is pure imaginary, meaning that the $k$-th field is a center for the reference dynamical system. In this case $\lambda_k = -\frac{\Delta t}{2}b^2 + o(\Delta t) < 0$ for a small enough $\Delta t$.

•Consider now the case $a \neq 0$. There are two sub-cases. If $a^2 \leqslant b^2$, then $\mathrm{Re}(\lambda_k) \leqslant 0$ for a small enough $\Delta t$. If $a^2 < b^2$, the condition $\mathrm{Re}(\lambda_k) \leqslant 0$ gives

$$\Delta t + o(\Delta t) = -\frac{2a}{a^2-b^2}.$$

So there is again a time step $\Delta t^\star > 0$ for which, for any $\Delta t < \Delta t^\star$ we have $\mathrm{Re}(\lambda_k) \leqslant 0$.

As a conclusion, starting from a stable linear dynamical system (in the sense that $\mathrm{Re}(\lambda_k^{ref}) \leqslant 0$ for all $k$), using a small enough time step $\Delta t$ and the forward Euler time discretization, the identification method leads to an estimated dynamical system which is also stable.

Let us underline that this could not be the case using another time discretization as e.g. for the backward Euler time discretization and the associated least square problem

$$\min_{A\in\mathscr{M}_K(\mathbb{R})} \frac{1}{2}\sum_{i=0}^{N-1}\left\|\frac{\boldsymbol{v}^{n+1}-\boldsymbol{v}^{n}}{\Delta t} - A\boldsymbol{v}^{n+1}\right\|^2. \tag{A 3}$$

Using identical developments, we would get in this case $A = \frac{I-\exp\left(-A^{ref}\Delta t\right)}{\Delta t}$ and

$$\lambda_k = \left(a - \frac{\Delta t}{2}(a^2-b^2)\right) + ib(1-a\Delta t) + o(\Delta t).$$

We observe that for a center with a pure imaginary eigenvalue $\lambda_k^{ref} = i\,b,\ b \neq 0$, one gets $\lambda_k = \frac{\Delta t}{2}b^2 + o(\Delta t)$ therefore $\lambda_k > 0$ for a small enough $\Delta t$. This is a counter-intuitive result: for numerical simulations, it is known that the backward Euler scheme provide more stability than the forward one. For system identification with time discretization of the residual term, it is safer to use the forward Euler scheme for stability of the estimated dynamical model.

## Appendix B. Kinetic energy dissipation

Another quantity of interest is the capsule kinetic energy $\|\{v\}\|^2$. Since the capsules are expected to reach a steady state after a transient stage in the Stokes pipe flow, the kinetic energy should also reach a constant value. From the differential equations, semi-orthogonality of $Q$ and symmetry property of the scalar product, we successively have

$$\begin{aligned}
\frac{d}{dt}\left(\frac{1}{2}\|\{v\}\|^2\right) &= \frac{d}{dt}\left(\frac{1}{2}\langle Q\boldsymbol{\beta}, Q\boldsymbol{\beta}\rangle\right)\\
&= \frac{d}{dt}\left(\frac{1}{2}\|\boldsymbol{\beta}\|^2\right)\\
&= \langle\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}\rangle\\
&= \langle\boldsymbol{\beta}, A_\mu\boldsymbol{\beta}\rangle\\
&= \frac{1}{2}\langle\boldsymbol{\beta}, A_\mu\boldsymbol{\beta}\rangle + \frac{1}{2}\langle\boldsymbol{\beta}, A_\mu^T\boldsymbol{\beta}\rangle\\
&= \langle\boldsymbol{\beta}, \frac{A_\mu + A_\mu^T}{2}\boldsymbol{\beta}\rangle.
\end{aligned}$$

So stability properties on the kinetic energy are related to the spectral nature of the (symmetric) matrix $A_\mu^S = \frac{A_\mu + A_\mu^T}{2}$. Dissipation property is linked to the non-positiveness of the (real) eigenvalues of $A_\mu^S$.

## Appendix C. Practical computation of the pseudo-inverse matrix

The Moore-Penrose pseudo-inverse $\mathbb{X}^\dagger$ of a matrix $\mathbb{X}$ of size $d\times K$, $d \geqslant K$, with $\mathrm{rank}(\mathbb{X}) = K$ is defined by

$$\mathbb{X}^\dagger = \mathbb{X}^T(\mathbb{X}\mathbb{X}^T)^{-1}. \tag{C 1}$$

For an ill-conditioned matrix $\mathbb{X}$, the direct computation of $\mathbb{X}^\dagger$ by formula (C 1) is unsuitable because the condition number of $\mathbb{X}\mathbb{X}^T$ is the square of that of $\mathbb{X}$. A more robust procedure can be derived by help of the QR factorization. There exists a semi-orthogonal matrix $\hat{Q}$ of size $d\times K$ and an upper triangular square matrix $R$ of size $K\times K$ such that $\mathbb{X}^T = \hat{Q}R$.

Moreover, $R$ is invertible because $\mathbb{X}$ is assumed to be a full-rank matrix. Since $\mathbb{X}^\dagger$ is the solution of the matrix system

$$\mathbb{X}^\dagger \, (\mathbb{X}\mathbb{X}^T) = \mathbb{X}^T,$$

we get

$$\mathbb{X}^\dagger \, R^T \hat{Q}^T \hat{Q} R = \hat{Q} R.$$

By multiplying by $R^{-1}$ to the right, since $\hat{Q}^T \hat{Q} = I_K$ we get

$$\mathbb{X}^\dagger = \hat{Q} \, (R^T)^{-1}.$$

## REFERENCES


Abuhamdan, R. M., Al-Anati, B. H., Al Thaher, Y., Shraideh, Z. A., Alkawareek, M. Y. & Abulateefeh, S. R. 2021 Aqueous core microcapsules as potential long-acting release systems for hydrophilic drugs. *International Journal of Pharmaceutics* **606**, 120926.

Ascher, U. M. & Petzold, L. R. 1998 Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations, 1st edn. USA: Society for Industrial and Applied Mathematics.

Aster, R. C., Borchers, B. & Thurber, C. H. 2019 Parameter Estimation and Inverse Problems, third edition edn. Elsevier.

Barrault, M., Maday, Y., Nguyen, N.C. & Patera, A.T. 2004 An 'impirical interpolation' method: application to efficient reduced basis discretization of partial diffeerential equations. *C.R. MAthematique* **339**, 667 – 672.

Barthès-Biesel, D., Walter, J. & Salsac, A.-V. 2010 Computational hydrodynamics of capsules and biological cells, chap. Flow-induced deformation of artificial capsules, pp. 35–70. Taylor & Francis.

Berkooz, G., Holmes, Ph. & Lumley, J. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Ann. Rev. Fluid Mech.* **25**, 539 – 575.

Boubehziz, T., Quesada-Granja, C., Dupont, C., Villon, P., Vuyst, F. De & Salsac, A.-V. 2021 A data-driven space-time-parameter reduced-order model with manifold learning for coupled problems: Application to deformable capsules flowing in microchannels. *Entropy* **23** (9).

Casanova, F. & Santos, L. 2016 Encapsulation of cosmetic active ingredients for topical application – a review. *Journal of Microencapsulation* **33** (1), 1–17.

De Vuyst, F., Dupont, C. & Salsac, A.-V. 2022 Space-time-parameter pca for data-driven modeling with application to bioengineering. In Advances in Principal Component Analysis (ed. Prof. Fausto Pedro García Márquez), chap. 13. Rijeka: IntechOpen.

Dubuisson, M. . & Jain, A. K. 1994 A modified hausdorff distance for object matching. In Proceedings of 12th International Conference on Pattern Recognition, , vol. 1, pp. 566–568 vol.1.

Dupont, C., Delahaye, F., Barthès-Biesel, D. & Salsac, A.-V. 2016 Stable equilibrium configurations of an oblate capsule in simple shear flow. *J. Fluid Mech.* **791**, 738–757.

Dupont, C., Salsac, A.-V. & Barthès-Biesel, D. 2013 Off–plane motion of a prolate capsule in shear flow. *J. Fluid Mech.* **721**, 180 – 198.

Dupont, C., Salsac, A.-V., Barthès-Biesel, D., Vidrascu, M. & Tallec, P. Le 2015 Influence of bending resistance on the dynamics of a spherical capsule in shear flow. *Physics of Fluids* **27** (5), 051902.

Foessel, E., Walter, J., Salsac, A.-V. & Barthès-Biesel, D. 2011 Influence of internal viscosity on the large deformation and buckling of a spherical capsule in a simple shear flow. *J. Fluid Mech.* **672**, 477 – 486.

Ghiman, R., Pop, R., Rugina, D. & Focsan, M. 2022 Recent progress in preparation of microcapsules with tailored structures for bio-medical applications. *Journal of Molecular Structure* **1248**, 131366.

Hu, X.-Q., Salsac, A.-V. & Barthès-Biesel, D. 2012 Flow of a spherical capsule in a pore with circular or square cross–section. *J. Fluid Mech.* **705**, 176 – 194.

Kutz, J., Brunton, S., Brunton, B. & Proctor, J. 2016 Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems. SIAM.

Lac, E. & Barthès-Biesel, D. 2005 Deformation of a capsule in simple shear flow: effect of membrane prestress. *Phys. Fluids* **17**, 0721051 – 0721058.

Lefebvre, Y. & Barthès-Biesel, D. 2007 Motion of a capsule in a cylindrical tube: effect of membrane pre–stress. *J. Fluid Mech.* **589**, 157 – 181.

Li, X. & Sarkar, K. 2008 Front tracking simulation of deformation and buckling instability of a liquid capsule enclosed by an elastic membrane. *J. Comput. Phys.* **227**, 4998 – 5018.

Lin, T., Wang, Z., Wang, W. & Sui, Y. 2021 A neural network-based algorithm for high-throughput characterisation of viscoelastic properties of flowing microcapsules. *Soft Matter* **17**, 4027–4039.

Lumley, J. L. 1967 The structure of inhomogeneous turbulent flows. *Atmospheric turbulence and radio wave propagation* pp. 166–178.

Ma, G. & Su, Z.-G 2013 Microspheres and Microcapsules in Biotechnology: Design, Preparation and Applications. Pan Stanford Publishing.

Matsunaga, D., Imai, Y., Omori, T., Ishikawa, T. & Yamaguchi, T. 2014 A full gpu implementation of a numerical method for simulating capsule suspensions. *Journal of Biomechanical Science and Engineering* **9** (3), 1–16.

Miyazawa, K., Yajima, I., Kaneda, I. & Yanaki, T. 2000 Preparation of a new soft capsule for cosmetics. *J. Cosmet. Sci.* **51**, 239 – 252.

Pozrikidis, C. 1992 Boundary Integral and Singularity Methods for Linearized Viscous Flow. Cambridge Univiversity Press.

Quarteroni, A., Manzoni, A. & Negri, F. 2016 Reduced basis methods for partial differential equations - An introduction. Springer International Publishing.

Quesada, C., Villon, P. & Salsac, A.-V. 2021 Real-time prediction of the deformation of microcapsules using proper orthogonal decomposition. *J. Fluids Struct.* **101**, 103193.

Ramanujan, S. & Pozrikidis, C. 1998 Deformation of liquid capsules enclosed by elastic membranes in simple shear flow: Large deformations and the effect of capsule viscosity. *J. Fluid Mech.* **361**, 117 – 143.

Schmid, P. J. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28.

Shawe-Taylor, J. & Cristianini, N. 2004 Kernel methods for pattern analysis. Cambridge University Press.

Sirovich, L. 1987 Turbulence and the dynamics of coherent structures, parts i, ii and iii. *Quart. Appl. Math.* pp. 561–590.

Tran, S. B. Q., Le, Q. T., Leong, F. Y. & Le, D. V. 2020 Modeling deformable capsules in viscous flow using immersed boundary method. *Phys. Fluids* **32** (9), 093602.

Trischler, A.P. & D'Euleuterio, G.M.T. 2016 Synthesis of recurrent neural networks for dynamical system simulation. *Neural Netw.* **80**, 67 – 78.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. & Kutz, J. Nathan 2014 On dynamic mode decomposition: Theory and applications. *J. Comp. Dyn.* **1** (2), 391–421.

Walter, J., Salsac, A.-V. & Barthès-Biesel, D. 2011 Ellipsoidal capsules in simple shear flow: prolate versus oblate initial shapes. *J. Fluid Mech.* **676**, 318 – 347.

Walter, J., Salsac, A.-V., Barthès-Biesel, D. & Le Tallec, P. 2010 Coupling of finite element and boundary integral methods for a capsule in a Stokes flow. *Int. J. Num. Meth. Engng* **83**, 829 – 850.

Williams, M.O., Kevrekidis, I.G. & Rowley, C.W. 2015 A driven-approximation of the koopman operator: Extending dynamic mode decomposition. *J.Nonlinear Sci.* **25**, 1307 – 1346.

Ye, T., Shi, H., Peng, L. & Li, Y. 2017 Numerical studies of a red blood cell in rectangular microchannels. *J. Appl. Phys.* **122** (8), 084701.

Yun, P., Devahastin, S. & Chiewchan, N. 2021 Microstructures of encapsulates and their relations with encapsulation efficiency and controlled release of bioactive constituents: A review. *Comprehensive Reviews in Food Science and Food Safety* **20** (2), 1768–1799.

Zhao, H., Isfahani, A. H.G., Olson, L. N. & Freund, J. B. 2010 A spectral boundary integral method for flowing blood cells. *Journal of Computational Physics* **229** (10), 3726–3744.