Reconstruction of the 3D hydrodynamics in a baffled stirred tank using Proper Orthogonal Decomposition

C. Mayorga, J. Morchain, A. Liné

Reconstruction of the 3D hydrodynamics in a baffled stirred tank using Proper Orthogonal Decomposition

C. Mayorga, [1]J. Morchain,[1] A. Liné[1]
[1] Université Fédérale Toulouse Midi-Pyrénées, TBI, INSA 135 avenue de Rangueil - 31077 Toulouse cedex 4, France.

Email: mayorgae@insa-toulouse.fr, jerome.morchain@insa-toulouse.fr, line@insa-toulouse.fr

In this work, the unsteady turbulent flow in a four-baffled stirred tank equipped with a Rushton turbine is computed with ANSYS FLUENT® using the sliding mesh technique. The POD algorithm is applied first to each mesh zone (fixed and rotating) and, in a novel and original way, to the full 3D CFD data without distinguishing between the stationary and rotating zones. The resulting modes cannot be regarded as spatial modes revealing the underlying flow structure. The POD results showed that 99.9% of the system variance is recovered with the first three modes. The reconstruction of the velocity field from these modes reveals the trailing vortices plus the mean flow. The accuracy of the POD reconstructed flow fields is quantified. The present work is the first evidence of a POD reconstruction of the unsteady turbulent flow performed over the entire 3D domain in a baffled stirred tank.

Keywords:

Proper Orthogonal Decomposition; full 3D velocity reconstruction; Computational Fluid Dynamics; RANS turbulence model; baffled mixing tank; Rushton turbine.

Highlights:

- 3D POD global reconstruction of mean and organised components of a turbulent flow in a baffled stirred tank.
- 3D POD application on fixed and rotating reference frames.

- Implementation of 3D POD on a mesh with a stationary zone and a rotating zone without zone separation.
- Accuracy computation through instantaneous and localized error for 3D POD reconstruction.

1. Introduction

Stirred tank reactors are widely used in chemical and biochemical industrial applications. The hydrodynamics of such systems has been the subject of numerous experimental (Ng & Yianneskis, 2000; Bugay et al., 2002; Baldi et al., 2004; Escudié & Liné, 2004; Ducci & M, 2007) and Computational Fluid Dynamic (CFD) studies that revealed its complexity (Hartmann et al., 2004; Delafosse et al., 2008, 2009). Besides the hydrodynamic aspects, additional models describing the chemical or biochemical reactions have to be implemented in the CFD software in order to couple mixing and reaction and perform predictive simulations. In some applications, it is important to consider the time varying flow field but the computational cost for solving simultaneously momentum and chemical equations is often prohibitive. In such situations, it is advantageous to find a way to access the spatio temporal variation for the velocity field without solving the Reynolds-Averaged Navier-Stokes equations throughout the process duration. Since the flow in a stirred tank is periodic and organized, the reduced order model approaches, such as Proper Orthogonal Decomposition (POD), can be used to achieve this objective.

Due to its elegance and practicality, this mathematical procedure has been used to identify, study and model the dynamics of large-scale average spatial structures of the fluctuating velocity field for turbulent flow (Sirovich, 1987a, 1987c, 1987b; Berkooz et al., 1993; Borée, 2003; Joshi et al., 2009; Tirunagari et al., 2012; Du et al., 2013; El-Adawy et al., 2018). In terms of average kinetic energy, these large structures (named coherent structures) dominate in the fluctuating flow, it follows that their identification is crucial in the study of fluid dynamics. This fact evidences the relevance and convenience of the POD to identify such organized structures.

Specifically, POD allows the most energetic structures or modes (on average) of an ensemble of functions or experimental or numerical data to be extracted. Each of these modes will have a spatial part or field (topos, eigenvectors) and a temporal part (chronos). In this way, it is possible to identify and study the spatial organization and temporal variations of each POD mode. Such a decomposition is very convenient because it allows understanding the role of each component individually or in combination with other modes. Indeed, a linear superposition of selected POD modes can be used to represent all the members of the ensemble. The selection of modes to be used depends on the temporal and spatial scale of interest. It can also be shown that such a reduced order method is optimal in terms of minimizing the mean-square error between the input data and its modal decomposition. This fact represents an extremely attractive numerical advantage (Holmes et al., 1996).

The POD methodology has also been shown to be effective for studying instabilities (Hasal, 2000; Hasal et al., 2004) and the transition to the turbulent regime in fluids (Knight & Sirovich, 1990; Holmes et al., 1996). In Arányi et al., (2013) the technique was also used to study complex flows. This technique has also proved to be very useful for the analysis of turbulent flow fields in mixing tanks (Raju et al., 2005; Gabelle et al., 2013; de Lamotte et al., 2018b; Janiga, 2019; Fernandes del Pozo et al., 2020; Mikhaylov et al., 2021). For example Raju et al., (2005) implemented the POD algorithm to study the structures associated to the fluctuations about the mean flow induced by the impeller in a stirred tank. In this work, experimental PIV (particle image velocimetry) data were used as input to the method for different diameters of a Rushton-type turbine, and Reynolds numbers between 4,000 and 80,000. Moreover (Liné et al., 2013) used experimental PIV data from a mixing tank to study the presence of coherent structures and the turbulent phenomena associated. In addition, the kinetic energy and its viscous dissipation were successfully obtained from the POD methodology. Likewise (Gabelle et al., 2013, 2017; Fernandes del Pozo et al., 2020) used the decomposition method in 2D PIV data to reconstruct organized motion induced by impeller blades without performing angle-averaged sampling. The manner in which energy was transferred between the POD modes was also

studied. Also it is reported in the literature the application of the POD on CFD data for the study of turbulent swirling flow in an axisymmetric sudden pipe expansion (Howard et al., 2017). The order reduction technique used proved to be quite robust in reducing the transient database, showing its potential to provide valuable insights into the flow structure. By using a large array of pressure and velocity data, POD was able to pick out several key flow features, including the movement of vortices, and the structure and period of the precessing flow. This was followed by the reconstruction of the flow field de Lamotte et al., 2018a, 2018b assessed the dynamics for a stirred tank by applying POD to a 2D domain using CFD and PIV data. As reported in this work, the reduced order method was able to decompose the fluid flow into different structures: mean flow, coherent structures, and turbulent flow. Also (Janiga, 2019) applied the POD algorithm for the analysis of coherent structures and macro-stability present in a 3D (3-dimensional), 3C (3-component velocity) LES (Large Eddy Simulation) simulation for an unbaffled stirred tank. In this numerical experiment the sliding meshing technique was implemented, thus the simulation domain was divided into a fixed part and a rotating part. A similar work was recently published by (Mikhaylov et al., 2021b). The data was created through a Direct Numerical Simulation (DNS) with a "frozen rotor" approach and Reynolds numbers of 500, 600 and 700. In this research, POD was useful to reconstruct the temporal evolution of large-scale organized vortical structures behind the blades of a Rushton impeller in a non-baffled stirred tank. It was found that the first two modes dominate the energy spectrum by carrying 90% of the total mean kinetic energy. It also is relevant to mention that some higher modes (3^{rd}, 4^{th} and 5^{th}, 6^{th}) came in pairs in the energy ranking.

These two latest remarkable works show the feasibility of the numerical recipe for model order reduction in very complex flow fields. However, some important points were not addressed. For example in (Janiga, 2019), the POD decomposition was applied to each part of the mesh separately but a reconstruction of the 3D, 3C velocity field for the bulk was not performed. In general, it is neither simple nor straightforward to determine the relationship and similarity between the POD method for separate parts of the domain and the POD method corresponding to taking the domain

as a whole. This is due to the essentially statistical nature of the method. Additionally an analysis of the efficiency of the POD method in terms of calculation time, global and localized error was not provided. (Mikhaylov et al., 2021b) performed the reconstruction of coherent structures just in the zone around the impeller blades at low Reynolds number through the POD treatment of the velocity field expressed in a rotating frame. Finally, in both investigations the presence of baffles in the simulated stirred tanks was not considered. The incorporation of these elements is of great importance due to their contribution to the mean flow (circulation loops) and their decisive role in mixing phenomena. To sum up, running time-resolved CFD (Large Eddy Simulation or U-RANS) and applying Reduced Order Modeling such as POD (Proper Orthogonal Decomposition or Karhunen-Loève decomposition) is appealing because it generates a time-varying velocity fields at a moderate expense while preserving the spatial resolution (Du et al., 2013; de Lamotte et al., 2018b). This opens the route to both refined and cost-effective description of the unsteady velocity fields. From this brief review of the existing literature on the subject, it appears that the POD analysis of the flow field in a baffled stirred tank, in the fully turbulent regime (at a Reynolds number > 10^4), using CFD data computed with a sliding mesh approach has not been considered yet.

The main objective of the present work is to perform a numerical reconstruction of a three-dimensional velocity field in a four-baffled stirred tank, operated in a turbulent regime, making use of the Proper Orthogonal Decomposition. In that regime, DNS and LES are time consuming and POD post-processing relies on a huge set of instantaneous 3D velocity fields. In contrast, the unsteady RANS method addresses the calculation of the transient mean flow and, despite the filtering of turbulence scales; it was preferred here since our prior objective is to establish the methodology for the velocity field reconstruction in the entire domain. Whatever the approach used to produce CFD results, a major difficulty resides in the presence of a fixed and a rotating mesh zone. It is shown that the velocity field reconstruction using a per-zone treatment of the CFD data is complex and cumbersome whilst a global treatment of the entire volume offers significant advantages. Originally, it appears that the main flow features of the 3D velocity fields are correctly recovered with the global

treatment but the eigenvectors do not correspond to spatial modes, as is typically the case in POD treatment. In order to assess the relevancy of our reconstruction method in terms of efficiency and accuracy both the average error and the instantaneous and localized errors were calculated. Finally, an analysis of the numerical cost is presented. There is no record that such a scenario has been addressed before.

2. Theoretical background

As indicated by (Kosambi, 1943; Loève, 1945; Karhunen, 1946; Berkooz et al., 1993; Holmes et al., 1996) the proper orthogonal decomposition (POD) or Kosambi-Karhunen–Loève decomposition consists of a practical procedure that extracts the most energetic components (on average) from an ensemble of functions (in the continuous case) or experimental (physical or numerical) data (in the discrete case). The key idea of the technique is to provide a vector space basis with the special condition that the average square projection of the data ensemble in this vector space is maximum. In this scenario it is possible to reduce a large number of interdependent variables to a much smaller set of uncorrelated variables while maintaining the spatial and temporal resolution of the original ensemble of variables (Liang et al., 2002; Kerschen et al., 2005). When applied to fluid flow analysis, the data ensemble can be obtained from a series of experimental or simulated velocity fields.

Once the spatial structure of the ensemble is extracted and represented by a set of orthogonal functions named eigenvectors (or modes), the members of the original ensemble can be represented as linear combination of these modes. The most striking property of the POD is that the p-modes reconstruction (finite truncations in the modal expansion) is optimal as it minimizes the mean square of error in comparison with any other possible orthogonal basis having the same dimension p. Should the data ensemble be made of a collection of instantaneous three-dimensional velocity fields, the original field $\vec{U}(\boldsymbol{x}, t)$ is approximated as a linear combination of p modes, $\overrightarrow{\phi}_k(\boldsymbol{x})$ weighted by time varying coefficients, $a_k(t)$.

$$\vec{U}_{POD}(\boldsymbol{x},t,p)=\sum_{k=1}^{p}a_k(t)\vec{\phi}_k(\boldsymbol{x}) \qquad (1)$$

The Direct and the Snapshots methods are the most popular POD approaches. The choice of the method depends on the spatial and temporal resolution contained in the data ensemble. For example, when the spatial resolution is very high compared with the time resolution, the Snapshots Method is less demanding in terms of computational cost (Sirovich, 1987a, 1987b, 1987c; Smith et al., 2005).

The POD equations

Let us consider N arbitrary 3D velocity fields, $\vec{U}_n$, sampled in L points at time t_n with $n \in \{1,\dots,N\}$. Using the POD terminology, these instantaneous 3D velocity fields are named snapshots. Then, the idea is to find the most "typical" or "characteristic" structures $\vec{\phi}_k(\boldsymbol{x})$ in the set of $\vec{U}_n$.

The data ensemble is first arranged in a matrix $\underline{M}_{3L\times N}$ of the form:

$$\underline{M}=\begin{pmatrix} u_1^{(1)} & \dots & u_1^{(N)} \\ \vdots & \dots & \vdots \\ u_L^{(1)} & \dots & u_L^{(N)} \\ v_1^{(1)} & \dots & v_1^{(N)} \\ \vdots & \dots & \vdots \\ v_L^{(1)} & \dots & v_L^{(N)} \\ w_1^{(1)} & \dots & w_1^{(N)} \\ \vdots & \dots & \vdots \\ w_L^{(1)} & \dots & w_L^{(N)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{U}^{(1)} & \dots & \boldsymbol{U}^{(N)} \end{pmatrix} \qquad (2)$$

Such a matrix is known as the *snapshots matrix* $\underline{M}$. From (2) the covariance or correlation tensor $\underline{R}$ can be constructed following the Direct or Snapshot method as follows:

$$\underline{R}=\begin{cases} \dfrac{1}{N}\,\underline{M}\left(\underline{M}\right)^T & \text{Direct method} \\ \dfrac{1}{N}\left(\underline{M}\right)^T\underline{M} & \text{Snapshot method} \end{cases} \qquad (3)$$

The correlation tensor contains the degree to which and the manner in which, the velocities at different points are correlated. Consider the velocity vector $\mathbf{U}_1$ at a point 1 and the velocity $\mathbf{U}_2$ at a point 2 of the fluid domain. If the velocities $\mathbf{U}_1$ and $\mathbf{U}_2$ are statistically dependent their correlation will be different from zero, otherwise such correlation will be zero. Using the Direct method, the covariance tensor is represented by a $3L \times 3L$ matrix, where L represents the number of data points. This array represents a time-averaged two-point spatial correlation. Such a process makes the computational cost immense when the number of points in the studied domain is large (the size of the tensor $\underline{R}$ is $(3L)^2$).

Alternatively, it is possible to apply the Snapshot method proposed by (Sirovich, 1987a; Smith et al., 2005) which is always valid when the ergodic hypothesis is fulfilled (all accessible microstates in the phase space have the same probability to occur over a long period). In this case, the covariance tensor is represented by a *N x N* matrix and a correlation is sought between the different realizations of the velocity. Such a process makes the computational cost less prohibitive compared with the Direct method (the size of the tensor $\underline{R}$ is only N^2). Due to this advantage, the Snapshot method is used in this work.

This Snapshot approach consists in solving an eigenvalue problem to find the eigenvalues λ_k and eigenvectors $\vec{\Upsilon}_k$ through a diagonalization of the correlation tensor $\underline{R}$ (snapshot case from (3)):

$$\underline{R}\vec{\Upsilon}_k = \lambda_k \vec{\Upsilon}_k \qquad (4)$$

The eigenvalues λ_k are the same for the Direct method and the Snapshot method because (3) is a Hermitian matrix. However, the eigenvectors are different for both approaches.

In order to recover the eigen-vectors in the spatial space, eigen-vectors $\vec{\Upsilon}_k$ are transformed into $\vec{\phi}_k$using equation (5):

$$\vec{\phi}_k(\boldsymbol{x}) = \underline{M}\vec{\Upsilon}_k \frac{1}{\lambda_k}, \qquad k = 1, \dots, N \qquad (5)$$

Since the eigenvectors constitute an orthonormal basis, they obey the following relationship.

$$\left(\vec{\phi}_i\right)^T \vec{\phi}_j = \delta_{ij}\ , i,j = 1, \ldots., N, \delta_{ij}\text{: delta de Kronecker} \qquad (6)$$

After the calculation of the eigenvector, $\vec{\phi}_k(\boldsymbol{x})$ the modal components $a_k(t)$ are obtained as:

$$a_k(t) = \underline{M}^T \vec{\phi}_k, \qquad k = 1, \ldots., N \qquad (7)$$

This completes all the elements necessary to reconstruct the velocity field according to (1).

3. Material and methods

3.1 Stirred tank and simulation set up.

Unsteady RANS CFD simulation in a 70 L baffled stirred tank is performed to produce our numerical data. The geometry of the flat-bottom stirred tank equipped with a Rushton turbine and four baffles is presented in Figure 1. Likewise, Table 1 provides the dimensions of the simulated tank.

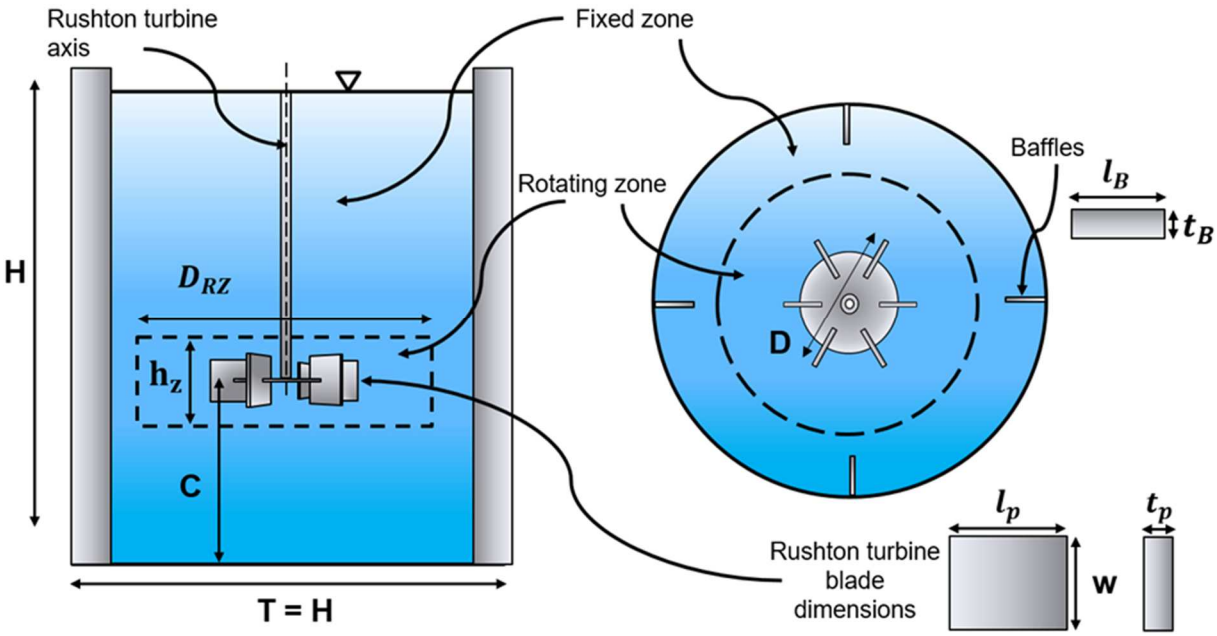

Figure 1 Configuration of the simulated stirred tank.

Table 1 Dimensions (in meters) of the simulated stirred tank

C	D	D_{RZ}	H	h_z	l_B	t_B	l_p	t_p	w
0.1500	0.1500	0.3000	0.4500	0.0600	0.0450	0.0050	0.0375	0.0020	0.0300

Water at room temperature and atmospheric pressure (density ρ = 998.2 $\mathrm{kg.m^{-3}}$, dynamic viscosity μ = 1.003 x $10^{-3}$$\mathrm{kg.m^{-1}.s^{-1}}$) is used as working medium. The stirring speed is 150 RPM (revolutions per minute) what corresponds to an angular velocity $\omega = 5\pi\ rad.s^{-1}$ or a frequency f = 2.5 Hz, according to (8) the Reynolds number is 56250.

$$Re = \frac{\rho\ f\ D^2}{\mu} \qquad (8)$$

The power number N_p for the Rushton turbine is deduced from the torque on the impeller ($\mathrm{P} = \mathrm{C}\ \omega = 6.5$ Watts) according to equation (9)

$$\boldsymbol{N}_\mathrm{p} = \frac{\mathrm{P}}{\rho f^3 D^5} \qquad (9)$$

The calculated value $\boldsymbol{N}_\mathrm{p} = 5.4$ allows determining the volume averaged viscous dissipation of kinetic energy $\bar{\varepsilon}$ computed from equation (10), we obtain a value of 0.090 $\mathrm{m^2.s^{-3}}$ or 90 $\mathrm{W.m^{-3}}$.

$$\bar{\varepsilon} = \frac{N_p\ f^3\ D^5}{V} = \frac{4}{27\ \pi}\ N_p\ f^3\ D^2 \qquad (10)$$

The mesh grid (1,129,140 cells and 1,184,282 nodes), models, settings and the numerical simulation procedure are identical to those used in a previous work (Delafosse et al., 2008). The standard $k-\varepsilon$ turbulence model is used with a standard wall function. A symmetry boundary condition is adopted on the free surface. The domain is divided into two zones: the fixed zone contains the walls, baffles, the major part of the shaft, and the volume outside the rotating zone, the latter is a cylindrical domain, which contains the impeller and a small portion of the shaft (see Figure 1). It must be noticed that a structured mesh made of hexahedrons was built in that zone (Figure 2). The starting point was a projection of all necessary edges on a horizontal plane followed by the creation of as many surfaces as needed to further build the impeller shape. Each face was meshed and the volume mesh obtained using the sweeping method. Owing to this strategy, the mesh is made of prismatic cells and it is almost invariant by rotation around the vertical axis.

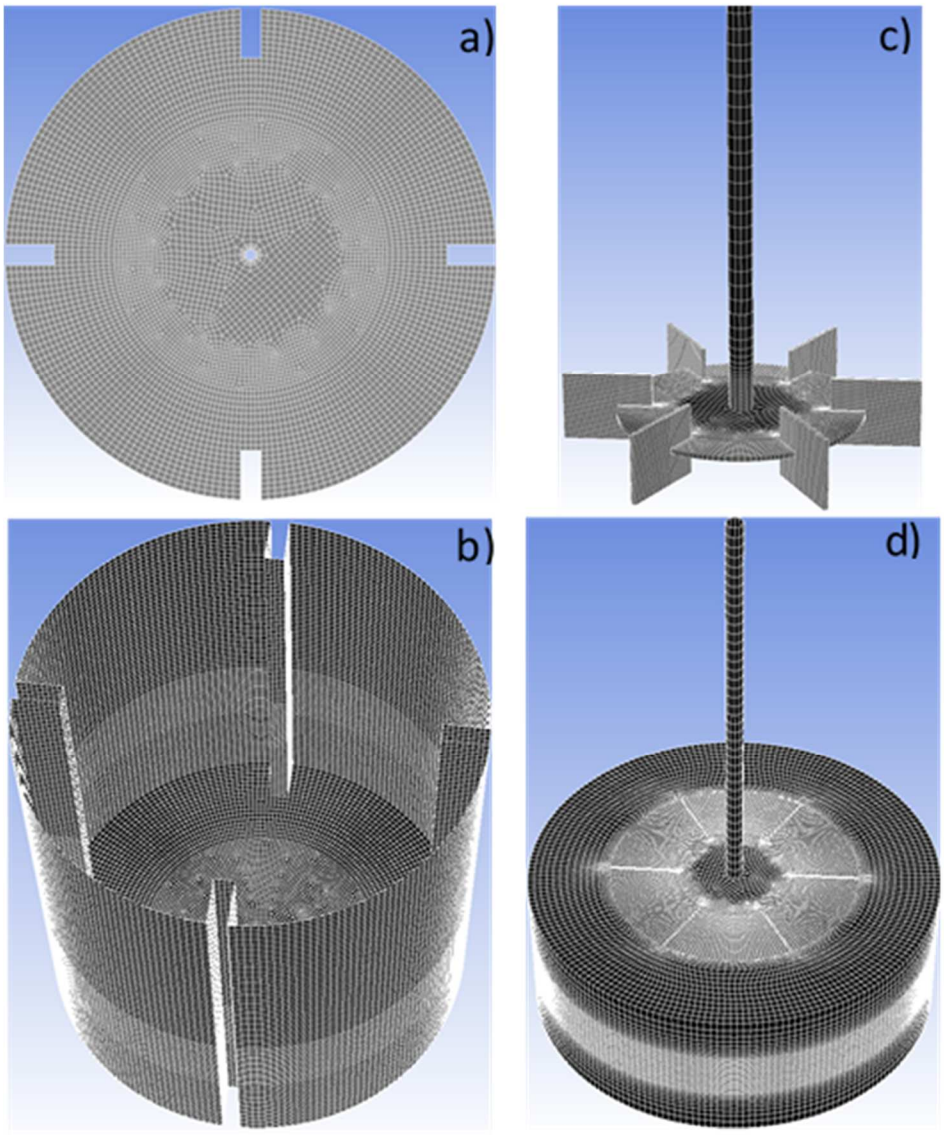

Figure 2.Mesh FLUENT views of the most important parts of the simulated tank: a) Top surface of the tank, b) External walls of the tank, c) Shaft and Rushton turbine, d) Shaft and rotating zone of the simulation domain.

The CFD calculations and data treatment were performed with ANSYS-FLUENT R20 on a parallel computer equipped with 40 processors Intel Xeon(R) E5-2660, 2.60 GHz. The time step used in the simulation is $\Delta t_{CFD} \approx 5.10^{-4}\, s$ corresponding to an angular rotation of 0.5 degrees per time step and a Courant number less than 1 in the whole domain to ensure numerical stability. Third order discretization scheme are used. Within each time step, 30 iterations of the non-linear solver were necessary to reach a plateau at 10^{-4} for the velocity and continuity residuals.

The convergence toward a stationary flow was assessed using two criteria: the equality between the torque on the turbine and the torque on the external walls of the tank, and the equality between the

volume integral of the turbulent dissipation rate and the power input. Regarding the first criterion, it is found that the torque on the turbine and the turbine shaft is practically identical to the torque generated on the external walls of the tank. Regarding the second criterion, the ratio of energy dissipation to the power input is 98%. The stationary flow field was established after 139 turbine revolutions (4 days of simulation) (see "Physical convergence of the numerical results" in Supplementary Information)

3.2 Sampling and data processing

The sampling of CFD data consists of a set of 3D velocity fields or snapshots collected in the entire simulated domain (1,129,140 cells). A total physical time of 11 s, representing 28 complete rounds of the turbine and requiring 26 days of computations, was spanned. During this time lapse, 386 snapshots were taken every 53 computational time steps. The time interval between two samples is $\Delta t_{st} = 0.03$ s ($\Delta t_{st} = 53\ \Delta t_{CFD}$). A sensitivity of the results to the number of snapshots is presented in the result section. Figure 3 provides an overview of the sampling data chosen in the present study.

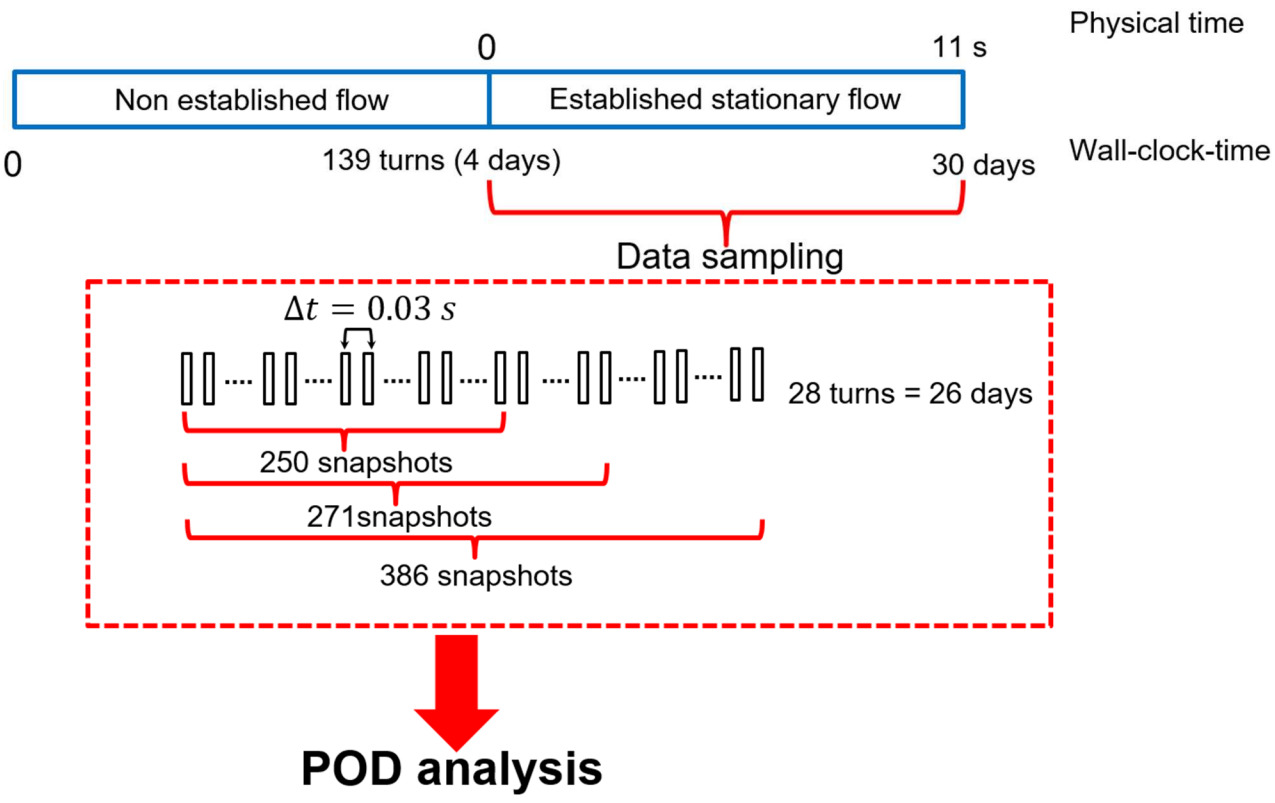

Figure 3. Sampling data sets collected for apply POD. Each rectangle represents a snapshot of registered velocity.

247 Finally, because the POD is a statistical method it is necessary to check that the number of snapshots
248 as well as the time window of data acquisition are sufficient to capture the structures carrying the
249 highest variance in the flow. The total simulation time must be long enough to record sufficient
250 information of the rotational motion while the sampling frequency must be high enough to prevent
251 the filtering of small-scale fluid motion. In addition, attention must be paid regarding the sampling
252 frequency to avoid the collection of snapshots "in phase" with the impeller rotation (see "Impeller
253 rotation and data sampling" in Supplementary Information).

254 A typical POD decomposition requires that velocity vectors and cell location be expressed in the
255 same frame. This is the case when POD is performed on experimental PIV data, since velocities are
256 generally measured in a fixed plane (Liné et al., 2013; Rodriguez et al., 2013). Some particular
257 attention must be paid when dealing with CFD data. In a previous work, Mikhaylov and co-workers
258 used CFD to compute the velocity field in an unbaffled stirred tank. The problem was solved in a
259 rotating reference frame using a single mesh domain (Mikhaylov et al., 2021a). Here also, velocities
260 are expressed in the rotating frame and calculated at some fixed position in that frame. In the
261 present work dealing with a stirred tank, the presence of baffles necessitates to split the fluid domain
262 into two mesh zones, a fixed one and a rotating one. In order to satisfy the above-mentioned
263 condition for a typical POD decomposition, velocities must be expressed in the inertial frame for the
264 fixed zone and in the rotating frame for the moving zone. The export of data, in FLUENT, extracts
265 velocities expressed in the inertial frame. Thus, prior to the POD treatment, the data collected in the
266 rotating zone have to be converted into velocities in the moving frame (see Figure 4). Then, the POD
267 treatment was performed separately for each set of velocities. This "per zone approach" corresponds
268 to the procedure adopted by Janiga when two mesh zones are present in the CFD model (Janiga,
269 2019).

270 Alternatively, a global POD treatment considering the two zones altogether (with the entire velocity
271 field expressed in the inertial frame) is attempted with a view of performing the 3D reconstruction of

the flow field in the entire fluid domain. This global approach has not been yet proposed in the literature for baffled stirred tanks and its feasibility remains to be demonstrated. We will show that the global treatment is feasible bearing in mind that its essential purpose is not to perform a hydrodynamic study of the fluid flow structure but a reconstruction of the velocity field by means of a linear combination of POD modes. Evidence of that will be provided in the section of 5.3. Assessing the accuracy of the flow field reconstruction.

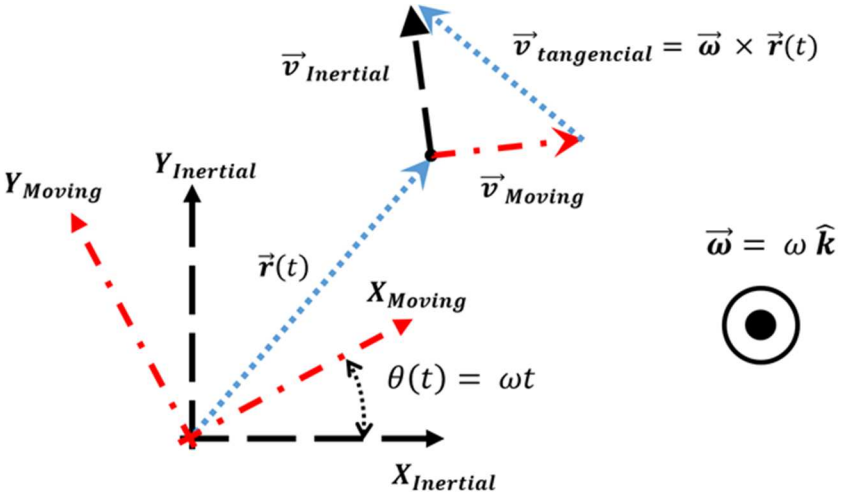

Figure 4. 2D schematic representation for the velocity vector expressed in the inertial and moving frames. The inertial frame is equivalent to the laboratory frame of reference. The moving and rotating frame share the same origin. The rotating frame rotates around the Z-axis with angular velocity ω. In that frame, the coordinates of the rotating cells are time independent.

Before going into the details, it is of utmost importance to observe that the data extracted from the CFD code are the components of the velocity vectors expressed in the inertial frame (laboratory reference frame). In addition, this information is referenced with respect to a cell index. Thus, each 3D snapshot results in an array containing the cell index in the first column followed by the three components of the velocity vector in the inertial reference frame ($\vec{U}_F(CI)$). The first step of the POD treatment consists in rewriting the three components of the velocity in a column vector corresponding to one snapshot. The rows of the matrix M represent the grid cells and the columns represent the time instants at which the data were registered. When moving along a row, the time changes but a given line always corresponds to the same cell index CI as illustrated in Figure 5.

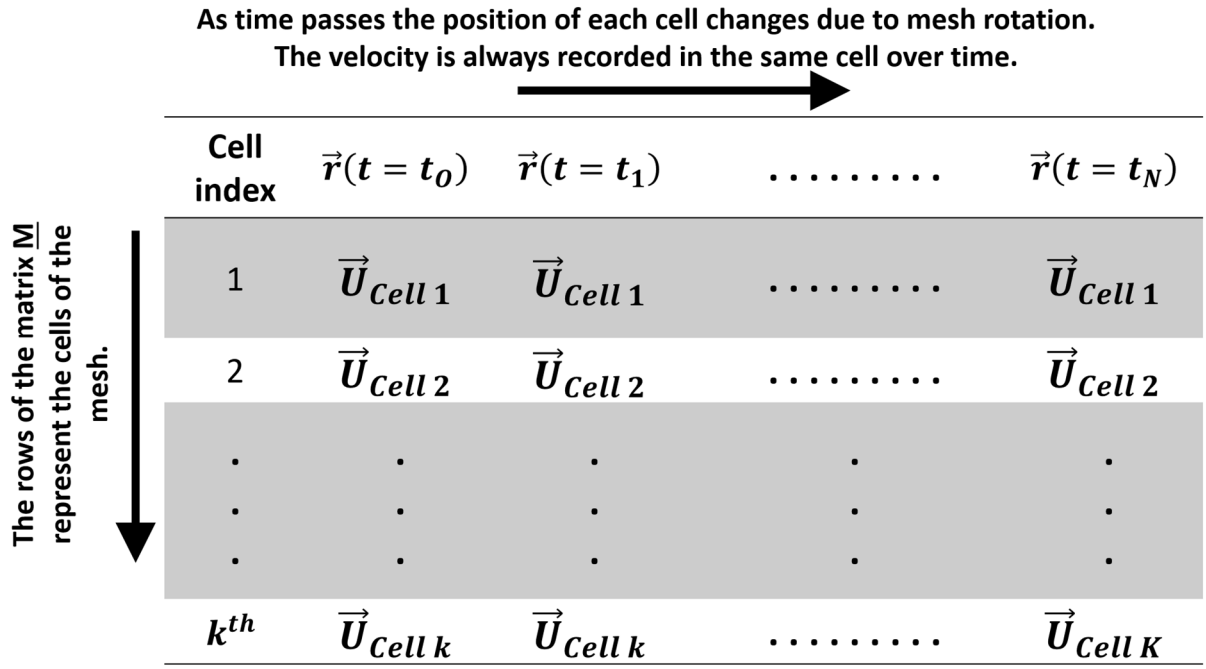

Figure 5. Schematic representation of the data recording used to build the snapshot matrix M for N snapshots and k cells. The velocity is always registered for a given cell although, for those cells belonging to the rotating zone, the spatial coordinates change over time due to the mesh rotation.

4. POD modal analysis

4.1 POD analysis in the fixed zone

First, the POD analysis is performed using 386 snapshots containing the velocity vector belonging to the fixed zone of the mesh, $\vec{U}_{IF} = \left(U_x, U_y, U_z\right)_{IF}$. Velocity components are here expressed in the inertial frame (IF). In that case, the relationship between the cell index and the cell location in the inertial reference frame is time independent. The mode spectrum is presented in Figure 6a, as well as the corresponding modal component and the vector field associated to the first mode. As expected, this first mode accounts for 90% of the total variance, its modal component is positive, almost time independent (Figure 6a,b) and thus this first mode reveals the structure of the mean flow. It is also remarkable the presence of two additional modes which carry about 10% of the total variance. These modes reveal a periodic fluid motion in the region outside the rotating zone of the grid, induced by the impeller rotation. As shown in Figure 6b, the corresponding modal components oscillate around zero with a period of 0.4 s corresponding to the impeller frequency of 2.5 Hz. The possibility of

identifying these oscillatory modes is an outstanding advantage that justifies the use of the POD methodology. In addition, Figure 6c and Figure 6d, provide different views of the spatial configuration contained in the first mode. Note that in Figure 6c an angular sector of 3.5 degrees is visualized in a vertical XZ-plane. Well-known organized structures are present: jet flow, recirculation loops, and vortices behind the baffles. Additionally, a predominantly axial fluid flow can be observed near the shaft and in the regions above and below the rotating zone.

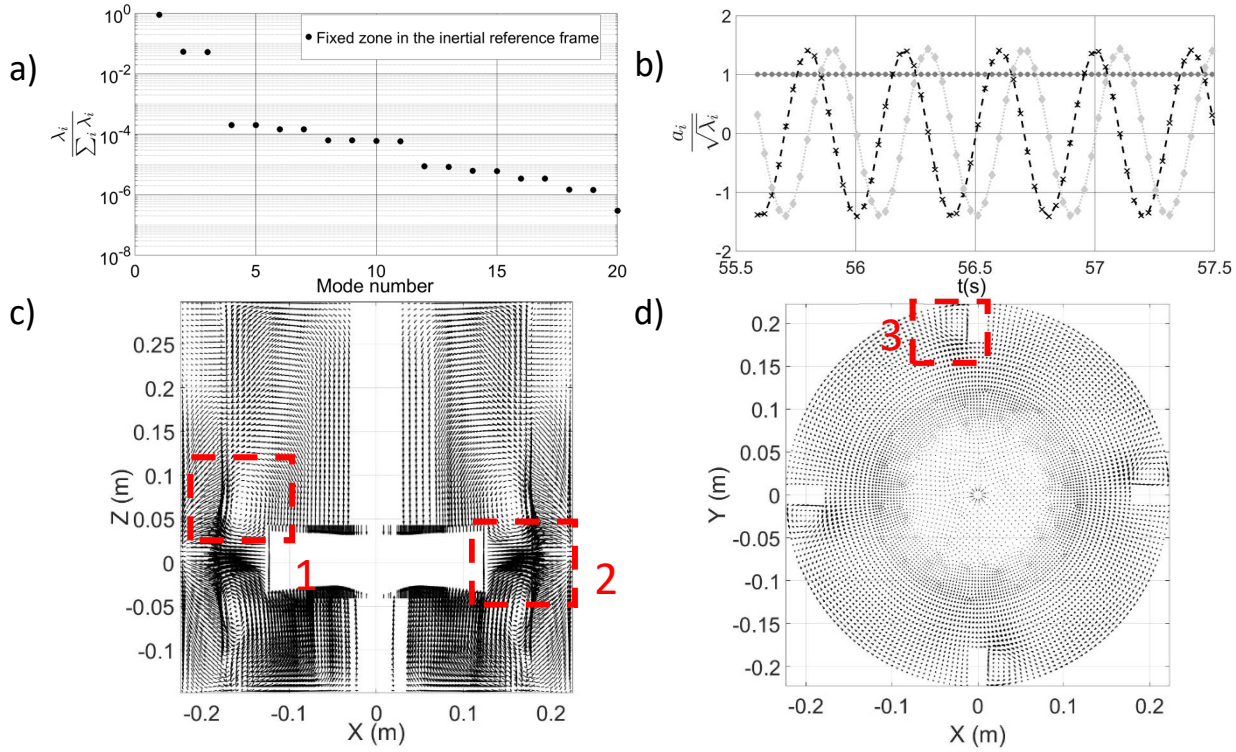

Figure 6. POD analysis of the flow in the fixed zone of the mesh: a) mode spectrum, b) mode component associated with the top three ranking modes, circles: first mode, x's second mode and rhomboids third mode (each data set is normalized by the variance corresponding to each mode.), c) First eigenvector visualized in a vertical XZ plane, d) Top view of the first eigen-vector. The main feature of the mean flow, i.e. radial jet flow, recirculation loops and coherent vortices behind the baffles are clearly visible (red boxes 1, 2 and 3).

The seemingly absence of velocity vectors in the central zone around the shaft (Figure 6c) is an artefact due to the small number of grid points used to describe the shaft surface itself. Since there is only one grid point every 30° around the shaft, the probability to find a grid node within the angular sector of only 3.5 °decreases as one approaches the shaft. Thus, the side view of the velocity field in the sector may not contain enough information close to the shaft resulting in an apparent absence of

data. However, it can be seen, from the top view (Figure 6d) that the flow field reconstruction at every grid point of the computational mesh is actually obtained close to the shaft.

4.2 POD in the rotating zone

Then, the POD analysis is performed using the velocity vectors belonging to the rotating zone of the mesh using again 386 snapshots. In this zone, the mesh rotates with respect to the inertial reference frame and the velocity fields $\vec{U}_{IF}$ are obtained in different locations for each time instant. Consequently, the velocity fields must be expressed in a reference frame in which the velocity measurement coordinates do not change in time. We recall here that the CFD code export procedure gives access to the Cartesian velocity components in the inertial reference frame. The velocity components in the moving (or rotating) frame, $\vec{U}_{MF}$, were deduced performing the next steps:

$$\text{i.} \quad \vec{U}'_{IF} = \vec{U}_{IF} + \vec{r}(t) \times \vec{\omega}$$
$$\text{ii.} \quad \vec{U}_{MF} = \underline{M_{Rotation}} \vec{U}'_{IF} \qquad (11)$$

$\vec{r}$ is the position vector of any cell expressed in the inertial reference frame. Importantly, the vector position of a cell differs from one snapshot to the other due to the rotation of the mesh. This rotation takes place around the vertical axis (Z-axis) aligned with the shaft and the position vector is known from the current cell coordinates $\vec{r}(t) = \vec{r}(\mathrm{x_{IF}(t)}, \mathrm{y_{IF}}(t), \mathrm{z_{IF}}(t))$. $\underline{M_{Rotation}}$ is the time dependent rotation matrix that is used to express the velocity $\vec{U}'_{IF}$ in a frame attached to the turbine (the position of the mesh in the first snapshot is used as a reference). Once the entire data set is treated, the snapshot database now contains a time series of the velocity components in the moving reference frame. Since the mesh is subjected to a solid rotation, the location of the cells, in the moving (rotating) frame, is now also time-independent. Once the transformations of (11) have been performed, the POD algorithm can be used. The POD results obtained are displayed in Figure 7.

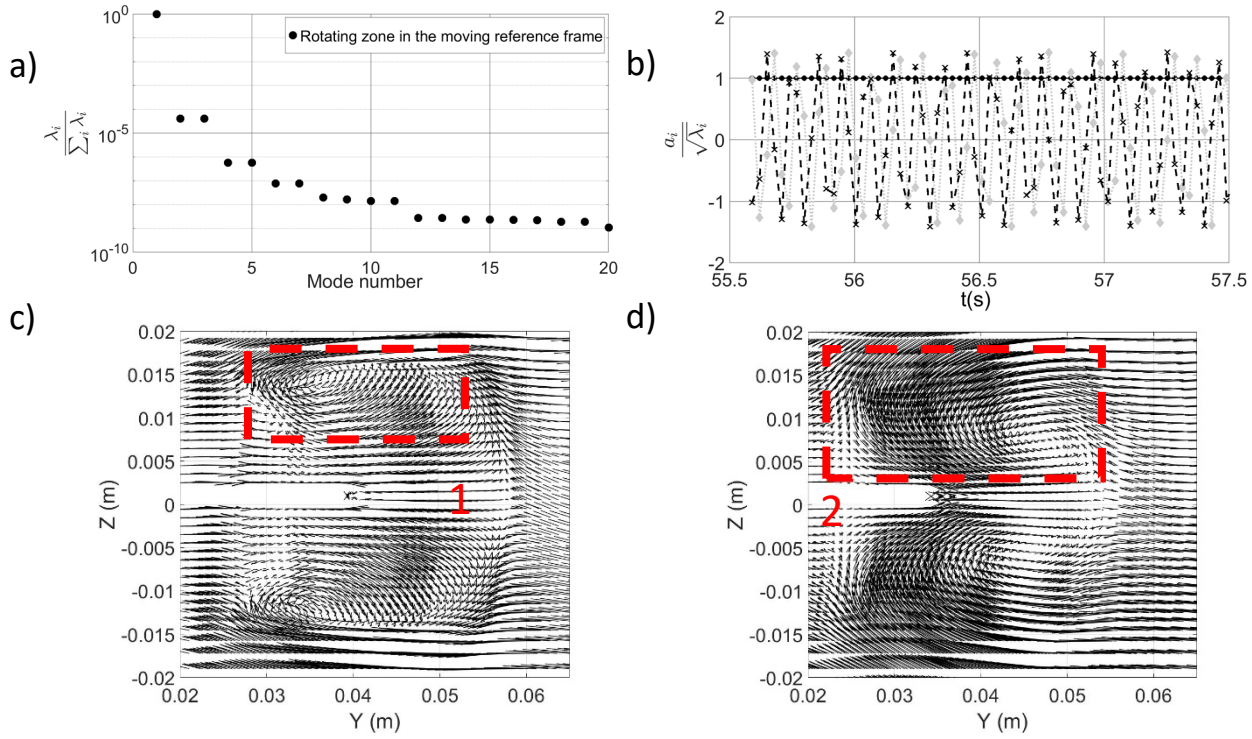

Figure 7. POD results for the velocity fields in rotating zone measured in the moving frame. a) Mode spectrum: one mode dominates the variance ranking; b) mode components associated to the three modes, circles: first mode, x's second mode and rhomboids third mode (each data set is normalized by the variance corresponding to each mode). The modal component of the first ranking mode is time independent and the second and third modes both oscillate with a frequency of 10 Hz. These oscillations are related to the passage of the baffles observed from the moving frame. c) The first eigenvector in the plane YZ three degrees behind one of the turbine blades reveal the presence of the trailing vortices (red box 1), d) The first eigenvector in the plane YZ six degrees behind one of the turbine blades reveals the radial displacement of the trailing vortices (red box 2).

As shown in Figure 7 the variance ranking is dominated by a single mode. In addition, this mode component is strictly time constant, as shown in Figure 7b. The second and third modes both oscillate with a frequency of 10 Hz or a period of 0.1 s (see the table "Fitting for the second and third modes for the moving frame" in Supplementary information). The impeller period is 0.4 s; since there are four baffles, the period associated with each baffle is 0.1 s. The first eigenvector contains the typical trailing vortices. They appear here as coherent structures attached to the turbine blades. The presence of such structures is shown in Figure 7c and Figure 7d for different positions of one of the turbine blades. The above results agree with the description of the fluid flow in a reference frame moving with the turbine blades. However, transformation (11) modifies the energy content of the

initial data and there is no mathematical link between the kinetic energy in the rotating frame and the inertial frame.

4.3 Reconstruction of the velocity field from the POD results per zone.

Once the procedure of reconstruction of the velocity field in the fixed and rotating zones has been carried out, the velocity field was constructed (expressed in the inertial reference frame) in the whole domain of the stirred tank. Since the two zones mentioned above were processed independently, the reconstruction of the velocity field was done by placing the reconstructed vector fields in each region. Special attention must be paid to two important details. First, the entire velocity field was built as a sum of the velocities reconstructed in each zone because the POD components involved are synchronized (the same snapshots were used). Second, all the data (fixed and moving zone) must be expressed in the inertial reference frame.

In the case of the fixed zone, no additional operation is necessary because the POD technique was applied on a fixed grid. Therefore, for this area, the reconstruction of the velocity field is already expressed in the inertial reference frame. The foregoing case is not valid for the moving zone because the CFD data were processed through the transformations (11). Consequently, the reconstructed velocity field in this region must be manipulated to express it back in the inertial reference frame. Concretely, inverse transformations of (11) must be applied to the reconstructed vector fields showed in Figure 7.

To sum up, in order to produce a 3D velocity reconstruction in the entire fluid domain, the recorded 3D fields are split into two zones, the subset of data corresponding to the rotating or moving zone are transformed (through equation (11)), then typical POD is applied in each subset of data, velocity fields for both are reconstructed, velocities in the moving zone are transformed back to the inertial frame and finally, the two velocity fields are gathered in a single file. The procedure described is feasible but it is long, complex and cumbersome.

4.4 Global POD treatment

In the previous treatments, velocities are expressed either in the fixed or in the moving or rotating frame. Such velocity fields are also referenced by a cell index, which does not change in time, i.e. $\vec{U}_{IF}(CI,t)$, $CI = h(CI)$. In addition, as indicated in the mathematical presentation of the technique, POD is a statistical method in which a correlation is sought between the different realizations and the procedure in itself does not require that all data are recorded at the same location. In the present case, the cell index is time independent while the relationship between the cell index and the location of that cell at any time is readily accessible. Therefore, it was postulated that a global POD treatment considering the entire mesh, with all the velocities expressed in the inertial frame of reference could make sense. The collected snapshots were analyzed using the numerical procedure described in the previous sections, producing a decomposition of the time varying 3D flow field on an orthogonal basis of 3D vector fields called modes. It is crucial to observe that such a POD decomposition is made in the space of cell indices which means that the associated POD modes are not spatial modes and their physical interpretation (as flow structure descriptors) has to be questioned. From a vector field reconstruction point of view the above methodology means that the velocity is reconstructed for the whole domain in the geometrical conditions (angle of rotation of the moving zone) corresponding to a given snapshot. Once the reconstruction is carried out in such conditions, the velocity vectors belonging to the rotating zone have to be placed at the appropriate location, considering mesh rotation. Presumably, this will not affect the velocity reconstruction derived from the POD method; this assumption will be evaluated according to the error values calculated in the results section.

Figure 8 depicts the POD mode spectrum obtained from the global treatment. The first three modes carry 99.99% of the total system variance. The first and second modes share the same variance content while the third mode has a slightly smaller variance. In addition, from mode 4 to mode 11, pairs of modes of equal variance are observed. In general, a pair of modes with the same variance content is typical of coherent structures propagating in space (the phase difference between the

time coefficients and the shifted distribution in space of the corresponding modes lead to the motion of these structures). Thus, this decomposition reveals a strong spatial organization of the flow in the form of multiple coherent structures with decreasing variance content. Interestingly, the third mode is a single mode, which suggests that it will be associated to the reconstruction of the mean flow, the time-independent component of the 3D velocity field. Similar results were obtained using either 250, 286, and 386 snapshots, which suggests a statistical convergence of the POD methodology.

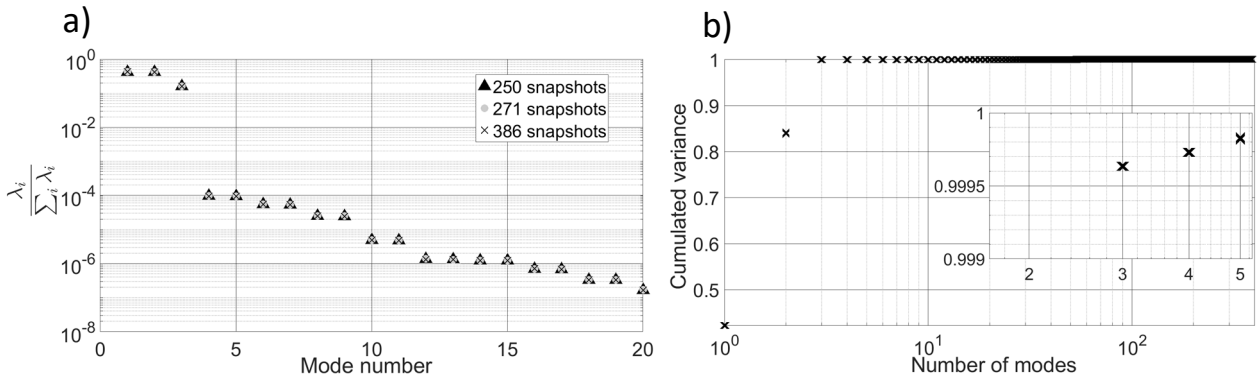

Figure 8. a) POD eigenvalue spectrum for the first twenty modes and three different time spans: the first three modes represents almost the 100% of the total system variance. It is also evident a correlation between pairs in the upper modes until the eleventh one. b) Cumulated variance POD spectrum for the time span of 386 snapshots. The cumulated variance until the third mode already represents the 99.9% of the total system variance.

Since the first three modes contain practically the totality of system variance, the subsequent discussion will first focus on them.

Figure 9a illustrates the time variation of the normalized amplitude factors associated to the first and second modes over 28 impeller turns (time span 386 snapshots). The temporal organization of these modes is clear; their behavior is oscillatory, with an identical period (0.4 s equal to the rotation period of the turbine) and amplitude. In addition, the phase shift between these modes is $\pi/2$ as can be seen from the circular configuration of the Figure 9b. Likewise, the amplitude factor of the third mode, presented in Figure 9c, shows a periodic time variation (again with a period of 0.4 s) but its mean value is different from zero. This fact together with its variance content suggests that this third mode will be associated to the reconstruction of the mean flow. The relatively small oscillations

presented in that POD component is related to the interaction between the periodic flow in the turbine region and the rest of the domain of the tank.

In order to verify statistically the periodic behavior of the coefficients a_1, a_2, and a_3 the pdf functions associated to each modal component are also provided (see "Examination of a_k statistical properties" in Supplementary information). It is evident that the functions are centered with respect to the origin (except a_3) , proving their periodical character, which is in agreement with what is reported by (Liné et al., 2013). With such an idea in mind, a fitting of the first eleven modes components, using as reference the function $\mathrm{A}\sin(\omega t + \varphi)$ was performed with the Matlab curve fitting toolbox®. The results provided in Table 2 show that the considered modes exhibit a very organized time pattern as all the squared correlation coefficient $\boldsymbol{R^2}$ are equal to 1. The first three modes oscillate with a period of 0.4 s (2.5 Hz) corresponding to the impeller rotation speed while the remaining modes turn out to be associated with harmonic frequencies of the main frequency.

Moreover, the presence of paired modes with the same frequency and amplitude is evidenced. This modal configuration implies a strong correlation between duets and is usually an indicator of coherent flow structure being present. Additionally, Figure 10 presents the relationship of the first mode with modes fourth to tenth. In agreement with the results of Table 2, each pair of modes is correlated with the first one in terms of Lissajous patterns. For each pair, represented by one member of each pair, the number of loops is related to the corresponding harmonic mode. Thus, for example, the fifth mode presents three loops since this is the third harmonic of the main frequency associated with the first mode. The same reasoning is applicable to the sixth, eighth and tenth mode.

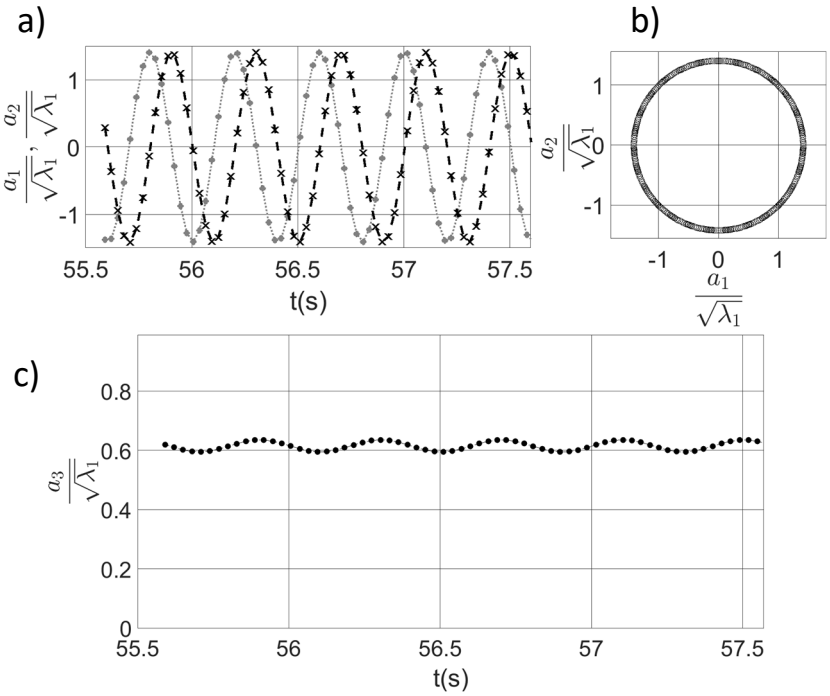

Figure 9. Graphical representation of the first three modes identified with the global treatment. For clarity, results are illustrated during 2 seconds only. a) Normalized time variation of the first and second mode. The circles and the crosses correspond to the modal component a_1 and a_2 respectively. The dotted and continuous lines correspond to their fitting by continuous sinusoidal functions. b) The circular configuration reveals a phase shift of π/2 between the first and second modes. c) Normalized time variation of the third mode. The solid circle corresponds to the modal component a_3. The dotted line corresponds to the sinusoidal fitting of the third mode.

471 *Table 2. Sinusoidal fitting of the mode components, $a_k(t)$ associated to the first eleven POD modes.*

Modal coefficient	$\boldsymbol{A_k}$: *Amplitude (m/s)*	$\boldsymbol{\omega_k}$: *Angular frequency (rad/s)*	$\boldsymbol{f_k}$: *Frequency (Hz)*	$\boldsymbol{\varphi_k}$ *Phase (rad)*	$\boldsymbol{R^2}$	$\dfrac{\boldsymbol{f}}{(\boldsymbol{2.50\ Hz})}$
a_1	454.50	15.71	2.50	17.21	1.00	1.00
a_2	454.60	15.71	2.50	15.65	1.00	1.00
a_3	6.51	15.71	2.50	15.64	1.00	1.00
a_4	6.84	47.12	7.50	7.84	1.00	3.00
a_5	6.84	47.12	7.50	12.55	1.00	3.00
a_6	5.15	62.83	10.00	21.91	1.00	4.00
a_7	5.15	62.83	10.00	26.62	1.00	4.00
a_8	3.47	78.54	12.50	5.97	1.00	5.00
a_9	3.47	78.54	12.50	4.39	1.00	5.00
a_{10}	1.53	94.25	15.00	21.61	1.00	6.00
a_{11}	1.53	94.25	15.00	23.17	1.00	6.00

472

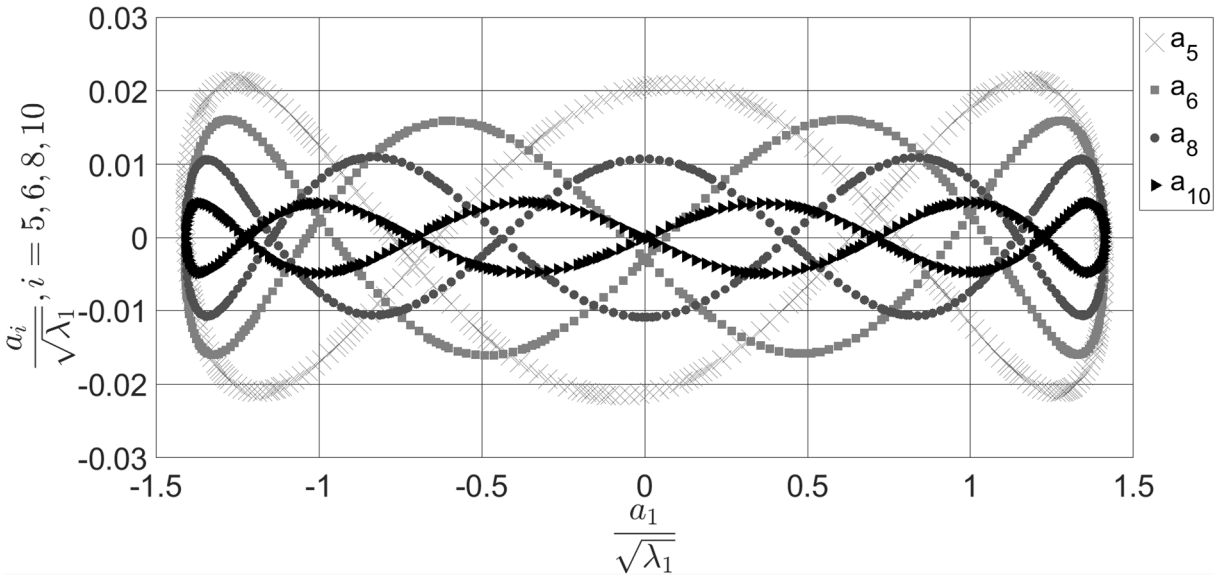

Figure 10.Lissajous patterns obtained from the temporal variation of the first to tenth modes.

The information provided above concerning the mode components a_k, clearly shows that the global POD approach produces an unexpected pattern regarding the temporal behavior of the corresponding modes. Indeed, the strong correlation between all frequencies and the impeller rotation speed suggests that the mode components actually reflect some information related to the mesh rotation rather than to the flow structure.

Now that the amplitude factors associated with each mode are expressed as continuous sinusoidal functions, it becomes possible to perform a continuous reconstruction of the 3D velocity field at any instant and not only at the instants corresponding to the snapshots. However, it is crucial to note that the mode produced by the global POD treatment are *cell index modes* and not spatial modes as is typically the case in POD analysis:

$$\vec{U}_{POD}(\boldsymbol{CI}, t) = \Sigma_k A_k \sin(\omega_k t + \varphi_k)\, \vec{\phi}_k(\boldsymbol{CI}) \qquad (12)$$

The exact location of any cell at any instant can be easily recovered from its initial position knowing the rotation speed of the moving zone. For those cells belonging to the fixed zone, the relationship between the cell index and the spatial location is time independent. For those cells belonging to the rotating zone, the relationship between the cell index and the location in the inertial reference frame

490 is a rotation around the z-axis. The angle of rotation is $\theta(t) = \omega\,(t - t_0)$ where t_0 is the instant
491 corresponding to the first snapshot. As a result, the global procedure to perform the velocity field
492 reconstruction contains only three steps:

493 i. Perform the POD on the entire domain
494 ii. Reconstruct the velocity field in the entire domain using *cell index POD modes*
495 iii. Relocate the velocity vector of the rotating zone at their actual location.

496 Figure 11a provides a close up of such a reconstructed flow field in the region of the blade impeller.
497 Reconstruction is here performed at the instant when the blade is located in the middle of the
498 angular sector, the first, second and third POD modes are included. The three major features of a
499 flow field generated by a Rushton turbine are clearly visible. Firstly, there is an axial flow moving
500 towards the blade as consequence of the suction produced by the passage of the blade turbine (red
501 boxes n° 1). Secondly, a pair of symmetrical vortices develop behind the blade (red box n° 2). These
502 very organized structures represent a manifestation of the well-known trailing vortices, which play
503 an extremely important role in mixing and transport processes. Noteworthy, these organized flow
504 structures would not be present if a time averaging of instantaneous velocity field (snapshots) was
505 performed. This fact represents an important advantage in the use of POD for velocity field
506 reconstruction in comparison to a mean flow field description based on the time averaging of
507 velocity fields. Thirdly, an axial jet flow is produced at the periphery of the turbine (red boxes n° 3).
508 Although wiggling, the jet flow is, on average, slightly deflected upwards in concordance with
509 (Delafosse, 2008; Escudié & Liné, 2004). This flow feature is present in the third mode and can
510 therefore be related to the mean flow. Figure 11b shows the velocity field obtained from the CFD
511 simulation at the same instant of the POD reconstruction (Figure 11a). A simple inspection of both
512 figures shows a very close resemblance between the reconstructed field and the field generated by
513 the simulator.

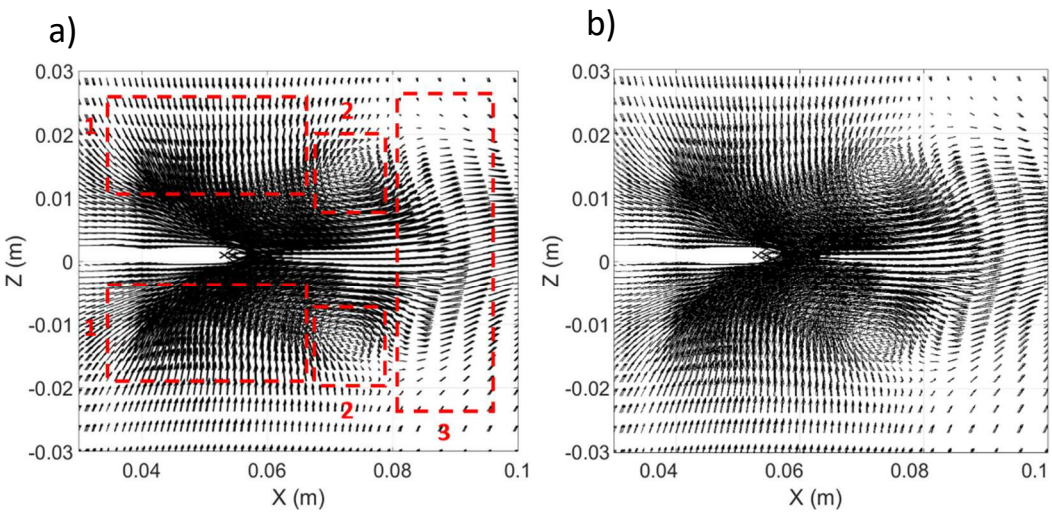

Figure 11. a) Reconstruction of the velocity using the first three modes. From the left to the right side the most important features of the field are marked by numbered dot rectangles: Upward and downward axial pumping flow (red boxes #1), vortices generated by the passage of the blade (red boxes #2), and radial flow out of the impeller-sweeping zone (red box #3), b) Velocity field from the CFD data in the same snapshot of the POD reconstruction.

5 Discussion

5.1. On the Global POD treatment eigenvectors

Having shown that the proposed treatment allows an accurate reconstruction of the fluid flow, we now comment on the mode decomposition. As already indicated, the global POD approach yields results that are not as intuitive as those usually provided by the typical POD approach are. In particular, the eigenvector obtained through the global treatment do not reveal the structure of the flow. Nevertheless, this “non-standard” treatment offers some striking features worth noting.

Firstly, the resulting modes preserve the variance content and allow a direct reconstruction without applying the per-zone method. This fact is an advantage in terms of computational resource savings by avoiding the use of equation (11).

Secondly, the rotational frequency of the mesh appears in the temporal behavior of all modes. In contrast with the typical treatment, this does not reveal a periodic fluid motion but rather reflects a correlation between velocity data when the mesh occupies the same location. In that sense, this observation is related to the statistical nature of POD. Although, the snapshot method is not aimed

534 at identifying temporal correlation, this type of correlation emerges here in the mode component
535 that are all scaled by the mesh rotation frequency: the velocity field in every cell is strongly
536 correlated with that obtained one turn later. Higher frequencies can be related to a strong angular
537 periodicity of the mesh.

538 Thirdly, the inspection of the local features of the cell index modes helps understanding the
539 outcomes of the global POD treatment. The first and second modes are significant in the rotating
540 region and almost negligible outside this zone (the same apply to high order modes also). It was
541 observed that both of them contain only radial and angular components and practically no axial
542 component. Since such modes do not come out from a typical POD procedure, and refer to a cell
543 index, it makes no sense to try to visualize them as such. However, as already emphasized, their
544 contributions to the reconstructed velocity field can be identified. It appears that these first two
545 modes are involved in the description of trailing vortices. The third mode is present throughout the
546 tank, including the rotating zone, and the three components (axial, radial and angular) are present in
547 this mode. However, the three components are not present everywhere. In the area swept by the
548 rotating mesh, the third mode contains essentially an axial component above and below the turbine
549 and a strong radial component away from the blades (see Figure 11a, Figure 3a from Supplementary
550 Information). Considering that the related mode component is almost constant, the third mode
551 seems to contain the axial flow corresponding to the pumping flowrate of the turbine and a radial
552 flow rate associated to the jet flow created by the radial impeller. Note that both of them are, on
553 average, independent of the actual mesh position. Accordingly, the reconstruction of the actual flow
554 field in the rotating zone requires using the first three modes altogether. In the fixed zone, the third
555 mode of the global treatment is very similar to the first mode of the POD treatment of the fixed zone
556 only (first mode of the per-zone treatment). In that zone, the contribution of the third eigenvector
557 can be physically interpreted, since the cell position is time independent. This result is sounded
558 because in a fixed domain, the per-zone and global approaches are equivalent. Indeed, the
559 configuration of this mode is very similar to the spatial mode presented in Figure 6d, Figure 6c. In the

upper part of the fixed zone, within a layer between 0.05 m and 0.1 m, this third eigenvector reveals the presence of clockwise vortices behind each baffle as is marked with the red boxes n°1 and n° 2 (see “Flow structure analysis using cell index modes” in Supplementary information).

Thus, the global treatment of the entire volume leads to identical results in those regions where the cell location is not changing in time. In the particular case studied here, it was observed that the constant mode in the moving zone contains a contribution to the flow, which is independent of the mesh rotation. Possibly, the structured mesh configuration played a role here since the Z location of the cell in the rotating zone is truly time independent while the mesh is invariant by rotation outside the zone swept by the impeller.

5.2. Dynamical representation of the reconstructed flow using the three first POD modes

The sinusoidal fit of the amplitude factor (Table 2) associated to each mode allow reconstructing the flow field at any instant. Therefore, a reconstruction was performed for a complete turn of the turbine, using a temporal resolution of the CFD (Δt_{CFD} = 0.5 ms), which corresponds to 715 snapshots. These snapshots were then superposed to generate a dynamic visualization or video of the plane reconstruction shown in Figure 11a. The video is provided in the complementary material (1+2+3 POD_Reconst stp 0.03s_XZ_5.5degrees_FITTING_6.avi). Such dynamic representation makes visible the dynamic or temporal evolution of the structures presented in Figure 11a. As can be seen in the video, both the axial flow (box n° 1 in Figure 11a) and the jet flow (box n° 3 in Figure 11a) are essentially constant. Finally, it is possible to identify the appearance of traveling vortices. These appear behind the blades every time they pass. Once such vertical structures are generated, they move radially away from their point of origin.

It is worth mentioning that all this dynamic information is contained in arrays whose size is much smaller than the CFD analogous. More explicitly, each POD vector has a dimension of 3L x 1 and its respective modal coefficient has a size of N x 1. In the present case study, L = 1.129.140 elements, N

= 386 snapshots, and only the first three POD eigenvectors with their respective modal components are necessary to recover 99% of the total variance of the system. This represents a storage requirement of (3L x 3) = 9L for the POD vectors plus the temporal components (Nx3) = 3N, the total number of elements being 1 x 10^7 or 0.08 Gb in total. The CFD data size 3xLxN = 1.3 10^9 elements which is equivalent to a data storage of 9.7 Gb. Thus, 121 times more storage capacity is required when using CFD data.

5.3. Assessing the accuracy of the flow field reconstruction using global POD approach

The accuracy of the reconstruction obtained from a global treatment is now examined. The equation (13) estimates the quality of the approximation as a function of the number of modes p used for the flow field reconstruction from global approach. The expression (13) estimates a maximum average relative error E_{MNE}. First, the average of the maximum error between the reconstructed velocity field and the CFD velocity field is calculated for every snapshot. Then by dividing this quantity by the tip blade velocity $\left(U_{tip} = 1.18\, m.s^{-1}\right)$, a normalized averaged error is obtained. This quantity is relevant because the POD technique is designed to minimize the average error between the reconstructed data and the experimental data (CFD data in this case).

$$E_{NAE} = \frac{\frac{\sum_{i=1}^{N} max\left(\left\|\vec{U}_{POD}^{i}(\boldsymbol{x},p) - \vec{U}_{CFD}^{i}(\boldsymbol{x})\right\|\right)}{N}}{U_{tip}} \qquad (13)$$

The relation (13) was calculated using 2, 3, 5 and 10 POD modes (time span of 386 snapshots). The results are provided in the Table 3 and presented graphically in the supplementary information ("Visual representation of the normalized error"):

Table 3 : Normalized Averaged Error as a function of the number of modes used for the reconstruction.

Number of modes	2	3	5	10
Normalized error (%)	62	6	5	3.5

The use of the first and second modes produces an average error of 62% with respect to the tip velocity of the turbine. Such components are independent of the third mode, which is associated with the reconstruction of the average flow. In this way the POD reconstruction associated with the two first modes do not contain the mean flow information, implying a considerably high average error. When the third mode is included in the reconstruction, the error decreases noticeably, which is fully consistent with the previous analysis. As can be seen in the Table 3, the maximum error becomes approximately 6.0 %. Adding more modes to the reconstruction only slightly reduces the error down to 3.5 %. The above information shows that the global treatment allows the reconstruction of the velocity field with a low average error.

Beside the normalized error, it seems interesting to quantify the error considering the number of cells in which the error is significant, i.e. the error as a function of the cumulative percentage of cells. Equation (14) estimates a local error between the global POD reconstruction of the velocity field and the CFD velocity field for each cell in the domain. From this, the error distribution is built and represented in Figure 12. Unlike expression (13), (14) provides the error due to the reconstruction in a localized and instantaneous manner because the respective calculation is performed for each cell. The definition of (14) allows a more adequate and accurate evaluation of the POD reconstruction for the unsteady case studied in the present work.

$$\mathrm{E}(\boldsymbol{x}, \mathrm{t}, p) = \frac{\left\| \vec{U}_{POD}(\boldsymbol{x}, t, p) - \vec{U}_{CFD}(\boldsymbol{x}, t) \right\|}{\left\| \vec{U}_{CFD}(\boldsymbol{x}, t) \right\|} \qquad (14)$$

Error distribution obtained through equation (14) for three different snapshots and using $p = 5$ and $p = 10$ POD modes for the reconstruction is presented in Figure 12. For the case of 5 POD modes, it was found that around 90.5% of the total number of cells exhibit an error smaller than 3%. The relative error of the rest of the cells is distributed according to the following categories: 4.5% of the cells have an error greater than 3% and less than 5%. The final 5% has an error greater than 5% and less than 27%.

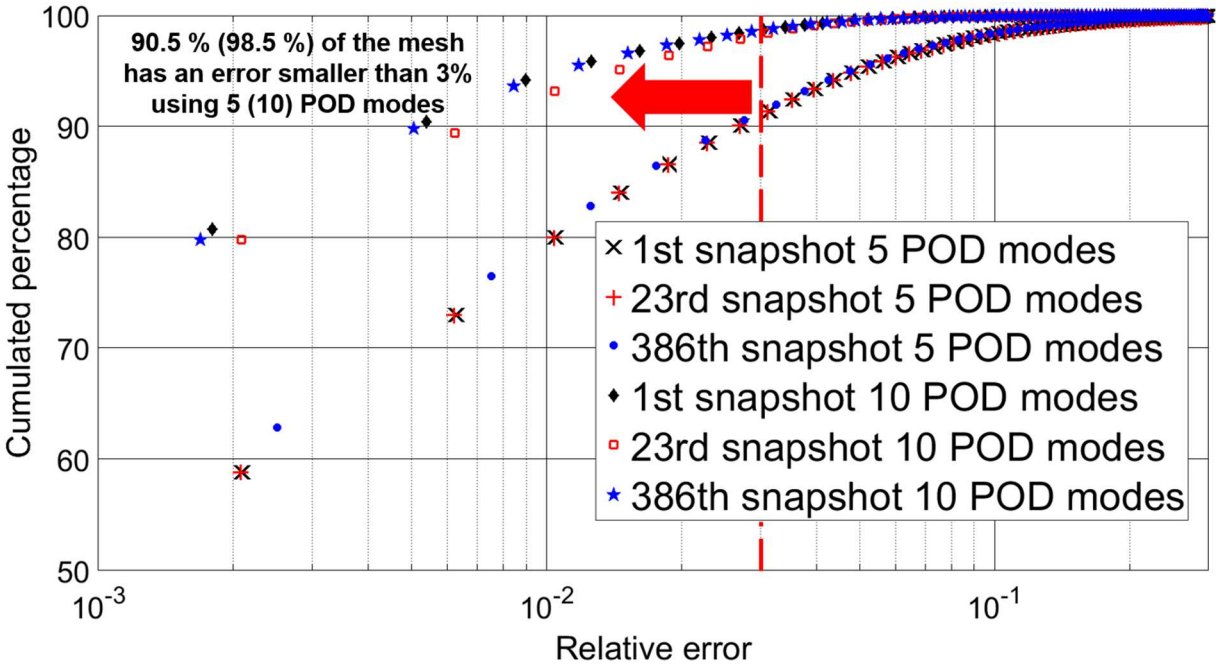

Figure 12. Cumulative percentage of the number of cells as a function of the relative error calculated by equation (14),using either 5 and 10 POD modes An increase in the number of POD modes leads to an upward shift of the curve. The vertical red line indicates 3% relative error.

The previous results are improved when 10 POD modes are used to perform the flow reconstruction: 98.5% of the total number of cells have an error less than 3%, 1% of the cells have an error between 3% and 5%. The final 0.5% has an error greater than 5% and less than 27%. Table 4 proposes the same results in terms of percentage of cells with an error comprised in a given range.

Table 4 : Percentage of cells with an error $E(\boldsymbol{x},t,p)$ comprised in a given range for p = 5 and 10 modes.

Snapshot #	E($\boldsymbol{x}$,t,p) < 3%		3% < E($\boldsymbol{x}$,t,p) < 5%		5% < E($\boldsymbol{x}$,t,p)	
	p = 5	p = 10	p = 5	p = 10	p = 5	p = 10
1	90.5	98.5	4.5	1	5	0.5
23	90.4	98.0	4.5	1.4	5.1	0.6
386	90.5	98.5	4.5	1	5	0.5

It was observed that the cells in which the relative error is high are mainly located on the interface between the fixed and rotating mesh grids (see 'locations of cell with significant error' in Supplementary Information).

5.4. Numerical cost of a spatio-temporal reconstruction using the global POD technique

The use of order reduction methodologies is not justified solely in terms of flow structure analysis. Another relevant aspect is the efficiency in the implementation of the method compared with the standard use of CFD tools. At this point of the present work, it is possible to propose a continuous spatio-temporal reconstruction of the 3D unsteady flow field, i.e. at a higher temporal resolution than the initial CFD data. Starting from the previous POD results and after identifying the continuous function associated with the mode components, the 3D field reconstruction was performed 714 times over 0.4 s, the duration of one turn. The time necessary to perform these reconstructions is compared with the time necessary to perform the CFD calculation with the same temporal resolution. Table 5 presents a summary of the physical time required for the CFD simulation of a full turn on the complete tank and the corresponding time necessary to perform the POD reconstruction via a global treatment using 5 and 10 modes. The first two lines provide the results for five and ten modes respectively. For both cases, the set-up time (time to read data and perform POD decomposition) is equal because the input data and numerical operations are identical. The reconstruction of a single 3D flow field is longer when ten modes are considered, 133.6 seconds for 10 modes against 64.5 seconds for 5 modes reconstruction. The corresponding time for the CFD simulation is 2.7 x 10^4 seconds, which corresponds to 58 and 50 times the duration of the POD reconstruction using 5 and 10 POD modes respectively.

Table 5. Comparative summary of the time duration for the CFD and POD reconstruction procedure.

Procedure	**Number of POD modes**	**Set up time (s)**	**Reconstruction time (s)**	**Total time (s)**	$\frac{t_{Recons\_POD}}{t_{CFD}}$
POD Reconstruction	5	400	64.5	464.5	1.7 x 10^{-2}
	10		133.6	533.6	2.0 x 10^{-2}
CFD				2.7 x 10^4	

This saving of calculation time is important especially considering that the reconstruction process is carried out only once. As indicated in the last column of the table it is also advantageous to use POD reconstruction with respect to CFD data. The difference in processing time between the two numerical methodologies evidences the advantage of using the POD technique for flow reconstruction and its subsequent use as input to other physical and biochemical models.

In particular, a continuous 3D reconstruction of the velocity field at a minimal computational cost can be particularly interesting to perform Lagrangian Particle Tracking and/or a compartmentalization of the fluid domain.

Conclusions

By using a per-zone and global POD approaches, it is possible to reconstruct the 3D mean and organized unsteady flow of a four-baffled agitated tank equipped with a Rushton turbine.

The per-zone method turns out to be an option to reconstruct the velocity field but such a methodology involves a larger number of steps compared with the global treatment of the entire volume.

For the global POD approach, the fact that in an inertial reference frame, the spatial coordinates of the cells belonging to the rotating zone are changing does not actually matter for the POD reconstruction. Such findings show the feasibility of using the POD technique for velocity field reconstruction using a CFD simulation on a mesh with a stationary zone and a rotating zone without zone separation.

The global POD decomposition results provide a series of components with different variances and frequencies. The top ranking modes present temporally periodic, these ones are associated with the reconstruction of coherent structures. The frequency associated to the first three modes (carrying 99.9% of the total system variance) is identical to the rotation frequency of the turbine. The first and

second modes are localized in the rotating zone and their axial component is essentially zero. The third mode carries the information related to the mean flow so its contribution is present in the whole domain. The largest horizontal (circulation loops) and vertical vortices are contained in it.

The coherent structures found by means of global POD reconstruction do not depend on the number of snapshots implying a statistical convergence of such organized movements. There exist modes of very low variance associated with the reconstruction of very organized structures. Such modes turned out to be harmonic frequencies of the main frequency associated the rotation of the turbine.

When three modes are considered for the reconstruction of the mean and organized velocity field, the maximum average error is of the order of six percent with respect to the CFD data. The addition of higher modes reveals an improvement in the maximum average error but it is not large unless several modes are considered.

The POD technique proves to be accurate even when evaluating a localized and instantaneous error. As expected, such accuracy is greatly improved when a larger number of POD modes are included.

Funding: This work was supported by the Instituto Tecnológico de Costa Rica; Institute Français d'Amérique Centrale IFAC, Campus France; and the Institute National des Sciences Appliquées Toulouse.

References

Arányi, P., Janiga, G., Zähringer, K., & Thévenin, D. (2013). Analysis of different POD methods for PIV-measurements in complex unsteady flows. *International Journal of Heat and Fluid Flow*, *43*, 204–211. https://doi.org/10.1016/j.ijheatfluidflow.2013.07.001

Baldi, S., Ducci, A., & Yianneskis, M. (2004). Determination of Dissipation Rate in Stirred Vessels Through Direct Measurement of Fluctuating Velocity Gradients. *Chemical Engineering & Technology*, *27*(3), 275–281. https://doi.org/10.1002/ceat.200401979

Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows. *Annual Review of Fluid Mechanics*, *25*(1), 539–575. https://doi.org/10.1146/annurev.fl.25.010193.002543

Borée, J. (2003). Extended proper orthogonal decomposition: A tool to analyse correlated events in turbulent flows. *Experiments in Fluids*, *35*(2), 188–192. https://doi.org/10.1007/s00348-003-0656-3

Bugay, S., Escudié, R., & Liné, A. (2002). Experimental analysis of hydrodynamics in axially agitated tank. *AIChE Journal*, *48*(3), 463–475.

de Lamotte, A., Delafosse, A., Calvo, S., & Toye, D. (2018a). Analysis of PIV measurements using modal decomposition techniques, POD and DMD, to study flow structures and their dynamics within a stirred-tank reactor. *Chemical Engineering Science*, *178*, 348–366. https://doi.org/10.1016/j.ces.2017.12.047

de Lamotte, A., Delafosse, A., Calvo, S., & Toye, D. (2018b). Identifying dominant spatial and time characteristics of flow dynamics within free-surface baffled stirred-tanks from CFD simulations. *Chemical Engineering Science*, *192*, 128–142. https://doi.org/10.1016/j.ces.2018.07.024

Delafosse, A. (2008). *Analyse et étude numérique des effets de mélange dans un bioréacteur* [Université Fédérale Toulouse Midi-Pyrénées]. http://www.theses.fr/2008ISAT0029/document

Delafosse, A., Liné, A., Morchain, J., & Guiraud, P. (2008). LES and URANS simulations of hydrodynamics in mixing tank: Comparison to PIV experiments. *Chemical Engineering Research and Design*, *86*(12), 1322–1330.

Delafosse, A., Morchain, J., Guiraud, P., & Liné, A. (2009). Trailing vortices generated by a Rushton turbine: Assessment of URANS and large Eddy simulations. *Chemical Engineering Research and Design*, *87*(4), 401-411.

Du, J., Fang, F., Pain, C. C., Navon, I. M., Zhu, J., & Ham, D. A. (2013). POD reduced-order unstructured mesh modeling applied to 2D and 3D fluid flow. *Computers & Mathematics with Applications*, *65*(3), 362–379. https://doi.org/10.1016/j.camwa.2012.06.009

Ducci, A., & M, Y. (2007). Vortex tracking and mixing enhancement in stirred processes. *AIChE Journal*, *53*(2), 305–315. https://doi.org/10.1002/aic.11076

El-Adawy, M., Heikal, M., A. Aziz, A., Adam, I., Ismael, M., Babiker, M., Baharom, M., Firmansyah, & Abidin, E. (2018). On the Application of Proper Orthogonal Decomposition (POD) for In-Cylinder Flow Analysis. *Energies*, *11*(9), 2261. https://doi.org/10.3390/en11092261

Escudié, R., & Liné, A. (2004). Experimental analysis of hydrodynamics in a radially agitated tank. *AIChE Journal*, *49*(3), 585–603. https://doi.org/10.1002/aic.690490306

Fernandes del Pozo, D., Liné, A., Van Geem, K. M., Le Men, C., & Nopens, I. (2020). Hydrodynamic analysis of an axial impeller in a non-Newtonian fluid through particle image velocimetry. *AIChE Journal*, *66*(6): e16939. https://doi.org/10.1002/aic.16939

Gabelle, J. C., Morchain, J., Anne-Archard, D., Augier, F., & Liné, A. (2013). Experimental determination of the shear rate in a stirred tank with a non-Newtonian fluid: Carbopol. *AIChE Journal*, *59*(6), 2251–2266. https://doi.org/10.1002/aic.13973

Gabelle, J. C., Morchain, J., & Liné, A. (2017). Kinetic Energy Transfer between First Proper Orthogonal Decomposition Modes in a Mixing Tank. *Chemical Engineering & Technology*, *40*(5), 927–937. https://doi.org/10.1002/ceat.201600674

Hartmann, H., Derksen, J. J., Montavon, C., Pearson, J., Hamill, I. S., & van den Akker, H. E. A. (2004). Assessment of large eddy and RANS stirred tank simulations by means of LDA. *Chemical Engineering Science*, *59*(12), 2419–2432.

Hasal, P. (2000). Macro-instabilities of velocity field in stirred vessel: Detection and analysis. *Chemical Engineering Science*, 11.

Hasal, P., Fort, I., & Kratena, J. (2004). Force Effects of the Macro-Instability of Flow Pattern on Radial Baffles in a Stirred Vessel With Pitched-Blade and Rushton Turbine Impellers. *Chemical Engineering Research and Design*, *82*(9), 1268–1281. https://doi.org/10.1205/cerd.82.9.1268.44169

Holmes, P., Lumley, J. L., & Berkooz, G. (1996). *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge: Cambridge University Press.

Howard, C., Gupta, S., Abbas, A., Langrish, T. A. G., & Fletcher, D. F. (2017). Proper Orthogonal Decomposition (POD) analysis of CFD data for flow in an axisymmetric sudden expansion. *Chemical Engineering Research and Design*, *123*, 333–346. https://doi.org/10.1016/j.cherd.2017.05.017

Janiga, G. (2019). Large-eddy simulation and 3D proper orthogonal decomposition of the hydrodynamics in a stirred tank. *Chemical Engineering Science*, *201*, 132–144. https://doi.org/10.1016/j.ces.2019.01.058

Joshi, J. B., Tabib, M. V., Deshpande, S. S., & Mathpati, C. S. (2009). Dynamics of Flow Structures and Transport Phenomena, 1. Experimental and Numerical Techniques for Identification and Energy Content of Flow Structures. *Industrial & Engineering Chemistry Research*, *48*(17), 8244–8284. https://doi.org/10.1021/ie8012506

Karhunen, K. (1946). Zur spektral Theorie stochastischer Prozesse. *Ann. Acad. Sci. Fennicae Ser.*, *34*.

Kerschen, G., Golinval, J., Vakakis, A. F., & Bergman, L. A. (2005). The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview. *Nonlinear Dynamics*, *41*(1–3), 147–169. https://doi.org/10.1007/s11071-005-2803-2

Knight, B., & Sirovich, L. (1990). Kolmogorov inertial range for inhomogeneous turbulent flows. *Physical Review Letters*, *65*(11), 1356–1359. https://doi.org/10.1103/PhysRevLett.65.1356

Kosambi, D. D. (1943). Statistics in function space. *Journal of the Indian Mathematical Society*, *7*, 76–88.

Liang, Y. C., Lee, H. P., Lim, S. P., Lin, W. Z., Lee, K. H., & Wu, C. G. (2002). PROPER ORTHOGONAL DECOMPOSITION AND ITS APPLICATIONS—PART I: THEORY. *Journal of Sound and Vibration*, *252*(3), 527–544. https://doi.org/10.1006/jsvi.2001.4041

Liné, A. (2016). Eigenvalue spectrum versus energy density spectrum in a mixing tank. *Chemical Engineering Research and Design*, *108*, 13–22. https://doi.org/10.1016/j.cherd.2015.10.023

Liné, A., Gabelle, J.-C., Morchain, J., Anne-Archard, D., & Augier, F. (2013). On POD analysis of PIV measurements applied to mixing in a stirred vessel with a shear thinning fluid. *Chemical Engineering Research and Design*, *91*(11), 2073–2083. https://doi.org/10.1016/j.cherd.2013.05.002

Loève, M. (1945). Functions aleatoire de second ordre. *C.R. Acad. Sci. Paris*, *220*.

Mikhaylov, K., Rigopoulos, S., & Papadakis, G. (2021). Reconstruction of large-scale flow structures in a stirred tank from limited sensor data. *AIChE Journal*. https://doi.org/10.1002/aic.17348

Ng, K., & Yianneskis, M. (2000). Observations on the Distribution of Energy Dissipation in Stirred Vessels. *Chemical Engineering Research and Design*, *78*(3), 334–341.

Raju, R., Balachandar, S., Hill, D. F., & Adrian, R. J. (2005). Reynolds number scaling of flow in a stirred tank with Rushton turbine. Part II — Eigen decomposition of fluctuation. *Chemical Engineering Science*, *60*(12), 3185–3198. https://doi.org/10.1016/j.ces.2004.12.040

Rodriguez, G., Weheliye, W., Anderlei, T., Micheletti, M., Yianneskis, M., & Ducci, A. (2013). Mixing time and kinetic energy measurements in a shaken cylindrical bioreactor. *Mixing*, *91*(11), 2084–2097. https://doi.org/10.1016/j.cherd.2013.03.005

Sirovich, L. (1987a). Turbulence and the dynamics of coherent structures. I. Coherent Structures. *Quarterly of Applied Mathematics*, *45*(3), 573–582. https://doi.org/10.1090/qam/910463

Sirovich, L. (1987b). Turbulence and the dynamics of coherent structures. II. Symmetries and transformations. *Quarterly of Applied Mathematics*, *45*(3), 573–582. https://doi.org/10.1090/qam/910463

Sirovich, L. (1987c). Turbulence and the dynamics of coherent structures. III. Dynamics and Scalling. *Quarterly of Applied Mathematics*, *45*(3), 573–582. https://doi.org/10.1090/qam/910463

Smith, T. R., Moehlis, J., & Holmes, P. (2005). Low-Dimensional Modelling of Turbulence Using the Proper Orthogonal Decomposition: A Tutorial. *Nonlinear Dynamics Springer*, *41*, 275–307.

Tirunagari, S., Vuorinen, V., Kaario, O., & Larmi, M. (2012). Analysis of Proper Orthogonal Decomposition and Dynamic Mode Decomposition on LES of Subsonic Jets. *CSI Journal of Computing, 1(3), 46-50*.