Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Multi-Dimensional Singular Value Decomposition of Scale-Varying CFD Data: Analyzing Scale-Up Effects in Fermentation Processes

Pedro M. Pereira[a]*, Bruno S. Ferreira[b] and Fernando P. Bernardo[a]

[a] CERES, Department of Chemical Engineering, University of Coimbra, R. Sílvio Lima, 3030-790 Coimbra, Portugal
[b] Biotrend SA, Portugal.
* Corresponding Author: pedro.mesquita.pereira@tecnico.ulisboa.pt

ABSTRACT

The scale-up of processes with complex fluid flow presents significant challenges in process engineering, particularly in fermentation. Computational fluid dynamics (CFD) is a crucial tool for accurately modelling the hydrodynamic environment in bioreactors and understanding the effects of scale-up. This study utilizes Higher Order SVD (HOSVD), which is the multidimensional extension of Singular Value Decomposition (SVD), to identify the dominant structures (modes) of fluid flow in CFD data of fermentation process simulations. Similarly to Proper Orthogonal Decomposition (POD), also based on SVD, this method can be used to identify the dominant structures of fluid flow, and additionally explore the scale parameter space. As a first test case, we examined five scales of a reciprocally shaken flask bioreactor, from 125 mL to 10 L, specified using basic empirical scale-up rules. Results indicate a common set of spatial modes across all scales, thus confirming that the scale-up method assures some dynamic similarity. However, the relative importance of these common spatial modes changes sensibly across process scales. The main spatial mode that represents the large recirculating flow in the flask can account for 19.8% of the data variability for the 125 mL scale but only 11.7% for the 10 L scale case. In this larger scale, one needs a larger number of modes to capture flow dynamics, thus showing an increase in flow complexity. These findings illustrate the impact of scale on fluid dynamics in a particular bioreactor flow but also provide a proof of concept for this methodology.

Keywords: HOSVD, CFD, Scale-up, Fermentation

INTRODUCTION

Scaling up fermentation processes is a primary focus in bioprocess engineering. Like other areas in chemical and process engineering, moving fermentation operations from bench-scale to pilot-scale, and ultimately to commercial-scale, exacerbates mixing inefficiencies, which are then responsible for spatial gradients in species concentration and mechanical stimuli. The biomass experiences these gradients as oscillatory stimuli, which can affect cell phenotype. Such effects can alter metabolic activity, resulting in hard to predict outcomes and decreased productivity [1].

The effects of scale arise from the interaction of the biological system with the hydrodynamics of the reactor [1]. Therefore, Computational Fluid Dynamics (CFD) is one important tool to clarify how transfer phenomena change with scale. Due to the difficulty of analysing the 3D scalar and vector fields resulting from a multiphase CFD simulation, the need for a solid method to do fluid flow feature extraction arises.

One of the most well-established methods for flow feature extraction is Proper Orthogonal Decomposition (POD). Its primary applications in Fluid dynamics have been to study, analyse and model turbulent coherent structures both from experimental as well as simulated data [2]. POD allows for the identification of the set of spatial modes (basis vectors) that represent most of the kinetic energy of the system and their temporal coefficients. The linear combination of the main modes allows

for a reconstruction of the data. The analysis of the modes and temporal coefficients can clarify the main coherent structures in a fluid flow.

POD has been demonstrated to be equivalent to Principal Components Analysis (PCA), Karhunen–Loève decomposition and Singular Value Decomposition (SVD), with the nomenclature used depending mostly on the area of application. POD, as well as PCA and SVD, are in general terms methods to reduce a large number of correlated variables into a smaller number of uncorrelated variables, finding a few ordered basis vectors able to optimally express (mean squared error wise) the initial large data set [2]. These basis vectors are often denoted in the context of CFD data as “modes”.

As a tool for identifying coherent structures and their dynamics, POD can enhance the understanding of scale-up effects in aerated fermentation and thus improve process design. Notable works applied to CFD include studies on stirred tanks by Mayorga et al. for a single-phase case [4]. Multiphase cases, as the one treated in this work, present greater complexity due to the interphase interactions in a turbulent flow, and are rarer in the literature. One such example is our previous work on CFD data of an aerated stirred tank bioreactor [5].

POD is only defined for a rank 2 tensor, that is a matrix, and as such it is not possible to directly apply it to a problem with additional parameters. To study the effect of scale on the main flow features, process scale can be defined as a parameter of interest and POD applied to each set of data corresponding to a particular process scale. However, the data will not be decomposed into the same basis vectors, that is the same modes, and thus a quantitative analysis of the impact of scale on the relative importance of the different modes is not possible.

To apply a similar methodology to a parametric case it is necessary to use tensor decomposition. Multiple methods for tensor decomposition exist, the most common being canonical polyadic decomposition (CP) and variations of Tucker decomposition such as Higher Order Single Value Decomposition (HOSVD) [6].

Canonical polyadic decomposition decomposes a tensor into a sum of rank one tensor (vector) products. Each rank one tensor can be seen as a mode in each parametric space, which then results in an equal number of modes in the decomposition for each parameter. Due to the nature of CFD data, in almost all cases there are many more points in space (number of finite volumes in the CFD mesh) than in time. Also, due to the cost of CFD simulations, the number of simulations made varying a particular parameter, such as the process scale, will be much lower than the number of finite volumes. This imbalance in the data makes CP a worse option for CFD data.

Tucker decomposition follows a different approach: it decomposes the original data into one smaller core tensor of the same rank, and one factor matrix for each parameter in study. The core tensor encapsulates the level of interaction between the parameters while the factor matrices can be seen as analogous to the modes found by POD or CP. With some data manipulation, the resulting tensor and matrices can be easily interpreted. Due to its greater flexibility and better fitting, we choose Tucker decomposition to proceed with our work despite the greater computational cost to perform the decomposition. In particular, we used the Higher Order Single Value Decomposition (HOSVD) implementation of Tucker decomposition [7].

Tensor decomposition has been used in various machine learning applications such as signal or image processing for tensor completion, principal component analysis, regression and classification, but tensor decomposition is less commonly applied to CFD data [8]. Applications of HOSVD to CFD data include parametric wind speed models, parametric models for flow over an aerofoil and flow feature identification on channel flows [9][10][11]. While much of the relevant literature highlights the potential of HOSVD in parametric model development, some studies recognise its use in flow feature identification. Despite this, in our review of the literature no instances were found where parametric modelling and feature identification are combined. We propose one such application of HOSVD focusing on flow feature identification while conducting a parametric study of the effects of process scale. This introduces innovation both methodologically and in the application to scale-up of bioreactors.

The approach here proposed enables us to identify the main coherent structures (modes) that emerge at various scales of fermentation and quantify the contribution of each one of the modes to the total energy of the system. Our methodology includes CFD simulations of the fermentation process at multiple scales, with the operating parameters determined by traditional scale-up criteria. The main challenge we found was the construction of the data tensor which requires all data to be referring to equivalent points in space. Through interpolation, we aligned the different CFD meshes and constructed a snapshots tensor of the CFD results across all process scales, standardized on the same spatial reference grid. This systematic approach bridges a gap in the literature, combining parametric modelling with flow feature identification to provide new insights into scale-up effects in fermentation. By helping to uncover the phenomenological roots of the scale-up problem, our methodology could inform new protocols, making the process of fermentation scale-up more expedient and efficient.

METHODS

As a case study to illustrate tensor decomposition of CFD data modelling chemical and biochemical process scale-up we modelled a set of non-baffled Erlenmeyer flasks commonly used in small scale laboratory fermentation, at different process scales. We opted for this bioreactor due to its geometrical simplicity making it simpler to model. The geometry of the flasks was reproduced from commercially available models and is represented in Figure 1. We modelled flasks of scales 125 mL, 250 mL, 500 mL, 1 L and 10 L, maintaining geometrical similarity. The 10 L process scale is hypothetical but we expected it to provide enough scale contrast to illustrate the changes found with changing process scale. Mixing in these bioreactors is induced by a shaking motion [12], and in particular we have chosen reciprocal shaking, which is a periodic motion in one dimension. This isn't the most common shaking mode, but it was preferred due to its greater simplicity and suitability for 2D CFD simulations, with lower computational cost.

Although not fully representative we preferred 2D simulations, which should be good enough to illustrate our methods, and highlight the effect of process scale on the governing flow modes. All flasks were simulated with 20% of the capacity filled, consistent with common practice. The amplitude of the movement is known to have minor impact on mixing performance, so an arbitrary value of 1.5 cm was used for the smaller scale and then scaled accordingly for the larger scales [13]. The frequency of the movement was adjusted for each scale by keeping the power per unit volume of liquid (P/V) constant.

$$N_p = \frac{P}{\rho n^3 d^4 V^{1/3}} \approx 70Re^{-1} + 25Re^{-0.6} + 1.5Re^{-0.2} \quad (1)$$

The power number (N_p) was estimated using the correlation of Klöckner and Büchs as seen in eq. 1 [14]. The 10 L flask falls outside the scope of the experimental data fitted by this correlation, which is a common limitation in solving scale-up problems solely based on nondimensional numbers and other empirical rules (such as P/V constant). Table 1 summarizes the process scales under analysis and corresponding values of d_{max} (maximum flask diameter) and n (frequency of agitation).

Table 1: Volume, maximum diameter and agitation frequency of each flask.

Scale (mL)	125	250	500	1000	10000
d_{max} (mm)	66	83	101	128	271
n (s^{-1})	2.00	1.81	1.71	1.52	1.07

To model the fluid dynamics, we used the Volume of fluid (VOF) model implemented in the interFoam solver of the OpenFOAM v2206 finite volume opensource library. The interFOAM solver has been validated to model shake flask flows [15]. In eq. 2-4 we present the VOF RANS equations with k-epsilon closure equations not presented for brevity, the closure equation parameters used are the standard OpenFOAM values.

$$\nabla U = 0 \quad (2)$$

$$\frac{\partial}{\partial t}\rho U + \nabla(\rho U \cdot U) = -\nabla p + \nabla(\tau + \tau_t) + \rho g + F_\sigma \quad (3)$$

$$\frac{\partial \alpha}{\partial t} + (U \cdot \nabla \alpha) = 0 \quad (4)$$

To model the shaking motion, we used an oscillatingLinearMotion dynamic mesh. The CFD meshes are 2D axial cuts of the geometry and were constructed with the OpenFOAM meshing tools. The meshes are hexahedral-dominant with from 25k to 210k finite volumes. To construct a coherent data tensor the mesh of the 125 mL flask was chosen as a basis and velocity fields from all larger geometries were interpolated onto that basis using OpenFOAM's transformPoints (scaling) and mapFields (interpolation) functions, one additional advantage of this is the effective down sampling making the tensor smaller

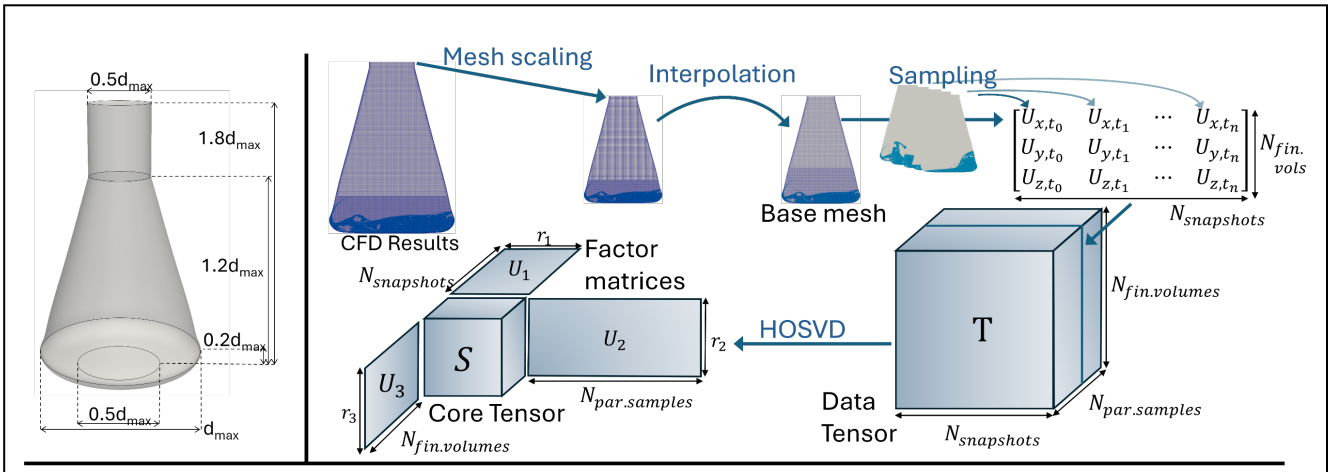

Figure 1. Schematic representation of the flask geometry and framework of the methods for data treatment.

and tensor decomposition computations lighter.

The snapshot matrix was constructed from velocity fields by arranging each velocity component per finite volume into columns, with each column representing a temporal snapshot. To ensure comparability across process scales, sampling was performed at 51 equally spaced times relative to the reciprocating movement's period, forming one full cycle. Snapshot matrices (shaped $N_{snapshots} \times 3N_{finite\_volumes}$) for five scales were concatenated into a data tensor ($N_{snapshots} \times 3N_{finite\_volumes} \times N_{scales}$).

The tensor was decomposed via truncated HOSVD using TensorLy a python library, with decomposition ranks of 20 (time), 5 (scale), and 50 (space). This decomposes the tensor into a core tensor (S), encoding parameter interactions (space, time, scale), and factor matrices (U_1, U_2 e U_3), analogous to POD modes (Eq. 5) in space, time, and scale, respectively.

$$T \approx S \times_1 U_1 \times_2 U_2 \times_3 U_3 \quad (5)$$

The method used here was proposed by Tucker and solves an eigenvalue problem to compute the leading left singular vectors of each matrix forming the data tensor. It then finds the core tensor by tensor multiplication of these vectors (transposed) with the data tensor. The number of left singular vectors corresponds to the rank of the decomposition in that parameter space. For the sake of brevity, we do not develop this methodology further, but further discussions on tensor decomposition and HOSVD can be found on the extensive review by Kolda and Bader [6]. The data tensor in this work includes ~2×10^7 elements, it takes 60 seconds to construct the tensor and only 4 seconds to perform the decomposition, using an in-house modest workstation.

RESULTS AND DISCUSSION

Figure 2 illustrates a set of snapshots of the volumetric fraction of water (α) for various scales taken at the same phase-matched times relative to the period of each movement (π). We present the α field as it is the most intuitive way to show the free-surface dynamics. Analysis of the snapshots reveals high dynamic similarity between the 125 mL and 1L scales. Particularly at the extremes of the flask's movement (at the relative times 2/5π and π) a similar wetted height of the flask wall is predicted and, the wave crests exhibit similar shapes. However, only at 2/5π are liquid ejections similarly predicted for both scales. At the 1/5π and 3/5π times, a similar wave pattern is predicted, but with greater amplitude for the smaller scale. For the 10 L scale less dynamic similarity is found. These differences illustrate the expected failure of the empirical scale-up correlation for the 10 L scale.

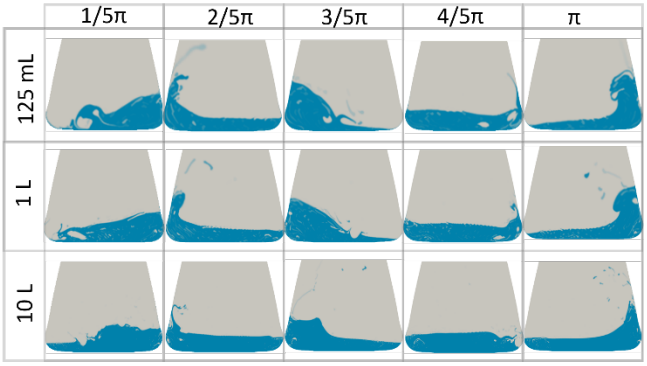

Figure 2. CFD results for different flask sizes, with the regions of water volumetric fraction greater than 0.95 painted in blue.

Following the preliminary qualitative analysis of the alpha fields, we propose tensor decomposition as a method to quantify the dynamic similarity of the fluid flow across process scales. The columns of the spatial factor matrix can be seen as spatial modes analogous to POD modes. Individual modes or combinations of modes are well-established to represent one or more correlated flow features. By quantifying the importance of each mode for each process scale, we can compare the fluid flow patterns at different scales. Applying tensor multiplication to the spatial factor matrix and the original data tensor provides a projection of the CFD data into the space spanned by the spatial modes. By taking the norm of the results of this operation over the temporal domain, we obtain the energy of each spatial mode for each process scale. The energy in this context means the amount of variability in the original data that is explained by each mode. Figure 3 presents these results, normalized by the sum of the energy of all modes for each process scale.

In Figure 3 we observe that the main two modes are the same across all process scales. But for less important modes the order of importance can change between process scales. Further, the combined importance of the two main modes decreases as the scale increases, from a combined normalized energy of 28.5% for 125 mL to 17.8% at 10 L scale. This loss of importance of the two top modes could be accompanied by an increase in the number of modes (ie. more relevant flow features), but we do not observe this. Depending on scale 43 or 44 modes are required to achieve 95% of the total energy represented in the decomposition. The top 30 modes represent a similar cumulative relative energy across all scales and the top 10 modes for each process scale represent 54%, 57%, 50%, 52% and 48% in ascending

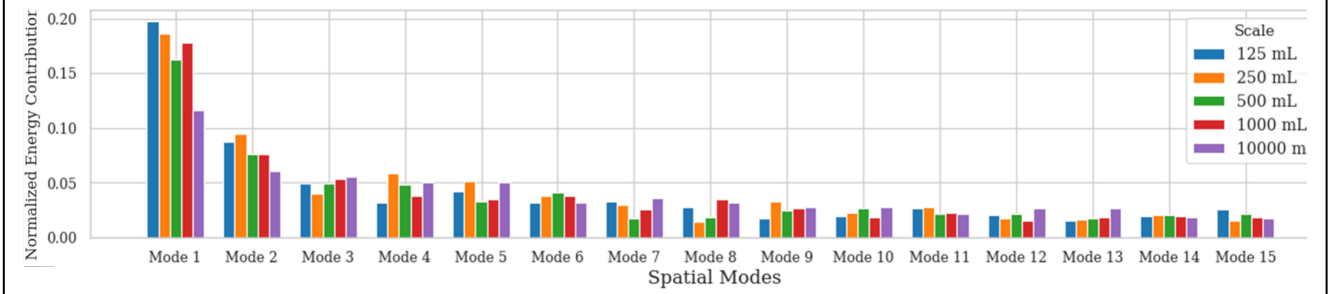

Figure 3: Relative normalized energy contribution (explained variability) of the 15 main spatial modes

process scale order. From these results we conclude that the relevant spatial modes are generally conserved across process scales, but their energy becomes more distributed and less concentrated on the top modes as scale increases. This redistribution of energy suggests a complexification of the velocity field with the increase of scale, likely due to enhanced turbulent effects for the larger scales.

To explore the connection between fluid flow features and spatial modes, we graphically analysed the spatial distribution of the modes and their temporal distribution. In Figure 4 the first five modes are presented and their normalized temporal coefficients for each process scale. The main modes exhibit a clear periodic oscillation in the normalized temporal coefficient curve, while the remaining modes appear to have a periodic nature with increased noise. Analysis of the Lissajous figures for each combination of the temporal coefficients (not presented for brevity) reveals that the two main modes have the same frequency and are π/2 out of phase for all process scales. Based on their spatial and temporal distribution mode 1 represents the rise of the fluid as the flask slows before reverting direction and mode 2 is the waving liquid transport during the movement of the flask. Mode 3 is π/4 out of phase and has twice the frequency of mode 1. It is then likely connected to the formation of wave crests at the extremes of the flask's movement.

Modes 4 and 5 exhibit interesting interactions with each other and with the two main modes. For the 125 mL and 250 mL scales, modes 2 and 4 are π/2 out of phase, as are modes 2 and 5. Modes 1 and 4 are 3π/4 out phase, while modes 1 and 5 are π/4 out of phase. This indicates that for these scales modes 1, 2, 4 and 5 represent coherent structures that peak sequentially at π/4 phase intervals. Modes 4 and 5 likely represent vortical structures related to the liquid jump at the extremes of the flask's movement. For the 500 mL scale a similar arrangement is found but with mode 6 (not pictured) showing similar dynamics to mode 4 as before, and mode 4 becoming an harmonic of mode 1. The results for 1 L and 10 L scales align with the results for 500 mL scale. These results indicate that the fluid flow during the travelling parts of the flasks movement represented by modes 1 and 2 is conserved with scale, at least for the adopted scale-up protocol. However, the phenomena at the extremes of the flask's movement show important changes with scale. Unexpectedly, an apparent discontinuity appears between the 250 mL and 500 mL scales, which seems to indicate that there is a particular loss of effectiveness in this scale-up protocol between these two scales.

Our findings demonstrate that as process scale increases, the main modes connected with dominant free-surface phenomena (e.g., wave formation, surface renewal) lose prominence relative to turbulence-driven dissipation within the bulk fluid phases (air and water). This redistribution of power dissipation highlights a decoupling of power dissipation and interfacial phenomena. The interfacial phenomena being the driving factor for oxygen mass transfer. From our analysis of this case we reach a central conclusion: scale-up strategies that neglect the evolving balance between surface and bulk dissipation mechanisms in free-surface aeration may fail to preserve mass transfer efficiency, despite achieving target power consumption metrics.

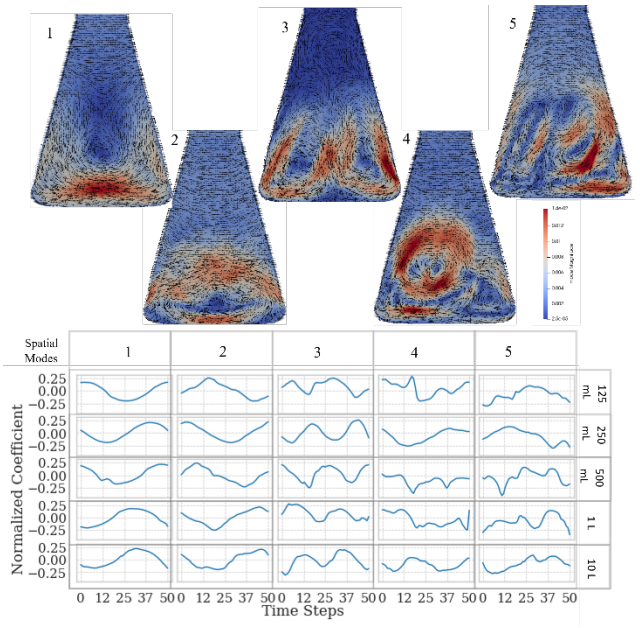

Figure 4: Spatial modes for the 5 main modes and normalized time coefficient curve for each mode/scale.

CONCLUSION

In this work we introduced a new methodology to study the effects of scale-up of fermentation processes, demonstrated with the scale-up of a shake flask bioreactor fermentation. Our findings reveal the impact of scale on fluid dynamics in a particular bioreactor flow and provide a proof of concept for this methodology, which can give insights for a more rational approach to bioreactor scale-up and optimization. We believe there is wide future application for bioreactor types, where geometric similarity is maintained. While new methods are required for future applications where geometrical similarity is broken. The application of this method can be extended to other scale-up rules and beyond scale-up, to the construction of parametric reduced models of multiple processes including reaction and mass-transfer models.

ACKNOWLEDGEMENTS

The authors thank financial support (PhD scholarship) from Fundação para a Ciência e Tecnologia, an organization within the Ministry of Science, Technology and Higher Education in Portugal, and also from Project "CiNTech - Technological Hub for Innovation, Translation and Industrialization of Complex Injectable Drugs" (reference: 7131), in the scope of PRR - Recovery and Resilience Plan and by the Next Generation EU European Funds, following NOTICE No. 02/C05-i01/2022, Component 5 - Capitalization and Business Innovation - Mobilizing Agendas for Business Innovation.

REFERENCES

1. Nadal-Rey G, McClure DD, Kavanagh JM, Cornelissen S, Fletcher DF, Gernaey KV. Understanding gradients in industrial bioreactors. Biotechnol Adv. 46:107660. (2021) https://doi.org/10.1016/j.biotechadv.2020.107660
2. P. Holmes, Turbulence, Coherent Structures, Dynamical Systems and Symmetry. Cambridge University Press (2012)
3. Liang YC, Lee HP, Lim SP, Lin WZ, Lee KH, Wu CG. Proper orthogonal decomposition and its applications—Part I: Theory. J Sound Vib. 252(3):527-544 (2002) https://doi.org/10.1006/jsvi.2001.4041
4. Mayorga C, Morchain J, Liné A. Reconstruction of the 3D hydrodynamics in a baffled stirred tank using Proper Orthogonal Decomposition. Chem Eng Sci. 248:117220 (2022) https://doi.org/10.1016/j.ces.2021.117220
5. Pereira PM, Martins RC, Ferreira BS, Bernardo FP. Reduction of an aerated fermenter CFD model using proper orthogonal decomposition. In: Computer Aided Chemical Engineering. Vol. 44. Elsevier (2024). Ed: Gernaey KV, Huusom JK, Gani R, editors. https://doi.org/10.1016/B978-0-443-28824-1.50144-7
6. Kolda TG, Bader BW. Tensor Decompositions and Applications. SIAM Rev. 51(3):455-500. (2009) https://doi.org/10.1137/07070111X
7. De Lathauwer L, De Moor B, Vandewalle J. A multilinear singular value decomposition. SIAM J Matrix Anal Appl. 21(4):1253-1278. (2000) https://doi.org/10.1137/S0895479896305696
8. Bi Y, Lu Y, Zhang L, Zhu C, Liu Y. Tensor decompositions: computations, applications, and challenges. In: Tensors for Data Processing. Ed: Liu Y. Elsevier (2022).
9. Zhang Y, Li X, Chen L, Yang Y. Three-dimensional wind velocity reconstruction based on tensor decomposition and CFD data with experimental verification. Energy Convers Manag. 316:115322. (2022) https://doi.org/10.1016/j.enconman.2022.115322
10. Lorente LS, Vega JM, Velazquez A. Generation of aerodynamic databases using high-order singular value decomposition. *J Aircr*. 45(5):1779 (2008) https://doi.org/10.2514/1.35258
11. von Larcher T, Klein R. On identification of self-similar characteristics using the Tensor Train decomposition method with application to channel turbulence flow. *Theor Comput Fluid Dyn*. 33(2):141-159. (2019) https://doi.org/10.1007/s00162-019-00485-z
12. Takahashi M, Aoyagi H. Practices of shake-flask culture and advances in monitoring CO_2 and O_2. *Appl Microbiol Biotechnol*. 102(10):4279-4289 (2018) https://doi.org/10.1007/s00253-018-8922-8
13. Kato Y, Hiraoka S, Tada Y, Hirose K, Büchs J. Mixing performance of a reciprocally shaking vessel. *J Chem Eng Jpn*. 36(6):663-667 (2003) https://doi.org/10.1252/jcej.36.663
14. Klöckner W, Büchs J. Shake-flask bioreactors. In: Comprehensive Biotechnology Ed: Moo-Young M, Elsevier (2011)
15. Dinter C, Gumprecht A, Menze MA, Azizan A, Niehoff P-J, Hansen S, et al. Validation of computational fluid dynamics of shake flask experiments at moderate viscosity by liquid distributions and volumetric power inputs. *Sci Rep*. 14(1):3658 (2024) https://10.1038/s41598-024-53980-7