

**Lagrangian modeling of hydrodynamic–kinetic interactions in (bio)chemical reactors
Practical implementation and setup guidelines**

Haringa, Cees; Noorman, Henk J.; Mudde, Robert F.

DOI

[10.1016/j.ces.2016.07.031](https://doi.org/10.1016/j.ces.2016.07.031)

Publication date

2017

Document Version

Accepted author manuscript

Published in

Chemical Engineering Science

Citation (APA)

Haringa, C., Noorman, H. J., & Mudde, R. F. (2017). Lagrangian modeling of hydrodynamic–kinetic interactions in (bio)chemical reactors: Practical implementation and setup guidelines. *Chemical Engineering Science*, 157, 159-168. <https://doi.org/10.1016/j.ces.2016.07.031>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Lagrangian modeling of hydrodynamic-kinetic interactions in (bio)chemical reactors: practical implementation and setup guidelines

Cees Haringa^{a,*}, Henk J. Noorman^{b,c}, Robert F. Mudde^a

^a*Transport phenomena, Chemical engineering department, Delft University of Technology*

^b*DSM Biotechnology Center, Delft*

^c*Bioprocess engineering, Biotechnology department, Delft University of Technology*

Abstract

Large substrate concentration gradients can exist in chemical or biochemical reactions, resulting from a large circulation time compared to the turnover time of substrates. The influence of such gradients on the microbial metabolism can significantly compromise optimal bioreactor performance. Lapin et al. [1] proposed an Euler-Lagrange CFD method to study the impact of such gradients from the microbial point of view. The discrete representation of the biomass phase yields an advantageous perspective for studying the impact of extra-cellular variations on the metabolism, but at significant computational cost. In particular, the tracked number of particles, as well as the applied time resolution, have a large impact on both the accuracy of the simulation and the runtime of the simulation. In this work we study the influence of these parameters on both the simulation results and computation time, and provide guidelines for the accurate Euler-Lagrange bioreactor simulations at minimal computational cost.

Keywords: Euler-Lagrange, industrial scale, bioreactor, hydrodynamics, metabolic modelling
2015 MSC: 00-01, 99-00

1. Introduction

In many (bio)chemical reactors, reaction takes place inside a discrete phase such as micro-organisms or catalysts particles, with transport occurring in the bulk phase. If the timescale of bulk mixing is in the range of or longer than the reaction timescale, the competition between reaction and bulk mixing will result in spatial substrate heterogeneity. When the discrete phase is mobile, such as in a slurry reactor or fermentor, micro-organisms/particles will see continuous changes in their environment as they move around. The spatial substrate gradients inside the reactor, translate to temporal substrate variations from the organism or catalyst's reference frame.

Focusing now on a bioreactor, the biomass specific production rate q_p of the desired component is typically governed by a complex metabolic reaction network, with the reaction rates depending both on the availability of extracellular substrates (such as sugar and oxygen) and intracellular components (such as amino acids and ATP). The adaptation of organisms to their surroundings does not occur instantaneously [2], meaning that the intra-cellular and extra-cellular conditions will typically not be in equilibrium. Consequently, q_p may vary in time, being a function of the organism's trajectory through space. As a result, the observed production rate of the entire population may

differ considerably from an ideal mixing situation [3]. Second, there may be considerable heterogeneity *within* the population [4].

Substrate concentration gradients originate from the fact that the bulk circulation time equals or exceeds the turnover time of metabolites [5]. Since the circulation time is dependent on the reactor scale, but turnover times are not, the extracellular conditions in a (typical ideally mixed) laboratory scale fermentor will not reflect the non-ideal industrial scale operation. This indicates that microorganisms are tested and selected under conditions that do not represent their future working environment [6]. For a reliable process design, effects of substrate concentration gradients on q_p should be taken into consideration during both process and organism development.

Scale-down simulators offer the possibility to study microorganisms under industrially representative conditions by deliberately introducing spatial concentration gradients using multi-compartment systems, or temporal concentration gradients via a fluctuating feed. [7, 5]. A recurring problem in scale-down simulation is to select the proper setup to reasonably match the industrial environment. Detailed information about the industrial environment is often scarce, or unavailable. It has been proposed to use computational fluid dynamics (CFD) coupled with reaction dynamics (RD) to gain insight in the industrial environment [6, 8]. In this way it is possible to study the magnitude of gradients and their effect on microorganisms

*Corresponding author:

Email address: c.haringa@tudelft.nl (Cees Haringa)

[1, 9, 10], providing valuable input for the design of scale-down simulators. Although this communication focuses on bioreactor applications, the outlined methods are applicable to any system dealing with similar dynamics.

1.1. CFD-RD coupling

Due to computational constraints, early CFD-RD work related to bioreactors often remained limited to the application of unstructured kinetic models (i.e. [3]), simulating only the uptake of substrate, or linking growth and production rate directly to the substrate uptake rate via a Herbert-Pirt equation. As such models assume an instantaneous adaptation of the metabolism to the extra-cellular conditions, such an approach is unsuitable to assess the effect of substrate concentration gradients.

More recently, the adaptation of the metabolism to environmental fluctuations has been included via two approaches: the Population-balance (PB) approach and Euler-Lagrange (EL) approach. In the PB approach, microorganisms are modeled as a component of the liquid phase. The biomass specific growth rate μ is typically applied to describe population heterogeneity [10, 11, 12]. This method is suitable for situations where all relevant processes are coupled to μ . However, metabolic fluctuations may take place on shorter timescales than growth rate fluctuations, and may have complex mutual interactions. To capture such processes PB-approaches are unsuitable, as the biomass population heterogeneity is not described solely by μ .

In EL-approaches the biomass phase is represented by a large number of virtual particles carrying an internal parameter vector describing their state [1, 9]. These virtual particles are further referred to as parcels to distinguish between computational and physical biomass particles. A large number of intra-cellular components and their mutual interactions can be tracked for each particle via a structured metabolic model. This makes the EL method specifically suitable to study the effect of reactor heterogeneity on metabolic timescales, and to study how the variations in multiple metabolites lead to heterogeneity in (among others) μ and q_p [9]. In our opinion, coupling complex metabolic models in which the intra-cellular response is governed by more parameters than just μ , is more straightforward in the Euler-Lagrange approach.

Besides the easy coupling of complex metabolic models, a second advantage is the viewpoint that the EL method offers. The use of parcels makes it straightforward to study the observed extracellular conditions over time for each parcel, and to monitor the intracellular response to these conditions. From this viewpoint, it is straightforward to construct ‘parameter versus time’ series for each parcel, for example the q_s versus t series shown in figure 1. Statistical processing of these series, which we will discuss in future communications, provides direct insight in the response of the metabolism to extra-cellular variations and distributions of the magnitude and duration

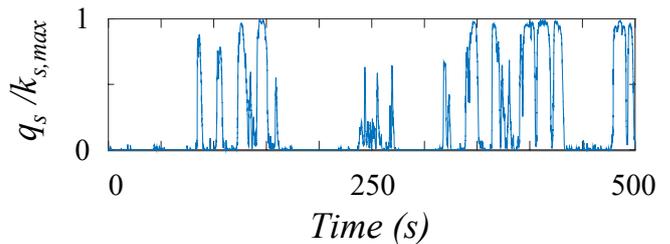


Figure 1: Example of a temporal uptake rate signal as experienced by a single parcel in an Euler-Lagrange simulation (scaled with maximum biomass specific substrate uptake rate $k_{s,max}$). This timeseries is a result of a spatial gradient as observed from the organism’s point of view.

of extra-cellular variations can easily be collected. Since experimental scale-down simulators should replicate such variations to test organisms under industrially representative conditions, these statistics are a valuable basis of scale-down design. In our opinion, these advantages make the EL-approach the preferred approach to study the impact of substrate gradients on the microbial metabolism, and to provide a basis of design for scale-down simulators.

A major challenge of the EL method is the significant computational burden. First, the large range of timescales in the problem leads to the requirement to simulate a significant period of flow time at a high temporal resolution. Second, a large number of parcels may be required to obtain an accurate solution. In the EL-method the biomass is distributed over the parcels. With the metabolic model to the parcels, uptake of substrate from the liquid phase becomes an inter-phase exchange effect, only occurring at locations where a parcel is present. Hence, the number of parcels N_p influences the homogeneity of the biomass concentration in the reactor has a strong influence at the quality of the solution. Due to the point-like nature of parcels, the amount of biomass in a given volume element is directly proportional to the number of parcels in that volume element. An insufficient number of parcels leads to artificial spatial variations in the biomass concentration, C_X , and consequently to artificial spatial variations in q_s and C_s . As will be discussed in section 1.2, the computation time depends strongly on the number of tracked parcels N_p , and on the required time resolution Δt . In this work, we consider what the minimum number of parcels N_p and maximum allowed timestep size Δt is to provide a good trade-off between accuracy and computational costs. Based on an evaluation of timescales involved in the problem, we present how these values can be predicted based on simulation parameters. These guidelines provide a basis for the computationally efficient application of Euler-Lagrange CFD to study the effects of non-ideal mixing on the performance of industrial scale bioreactors.

1.2. Problem outline

With agitation speeds of $> 1/s$, the time resolution Δt required to resolve the particle motion is in the order $\Delta t \approx 5 - 50$ ms. A similar time resolution is necessary to model the fast reactions such as substrate assimilation and ATP consumption [2]. The mixing time is typically in the order of one or several minutes [13], and several (approx. 10) mixing times should be simulated to properly resolve the magnitude- and duration distributions for extracellular substrate variations. This means in the order of 30 min flow time must be resolved for a typical industrial case. Reactions related to (for example) enzyme production may act on the timescales of hours [14]. If the goal is to study a fermentation including the influence of enzyme dynamics on biomass specific growth, production and uptake rates, the resolved flowtime requirement may be in the order of hours to days. Clearly, such simulations will be computationally demanding, mainly originating from the number of timesteps required to resolve all relevant dynamics. As such, the computational burden per timestep must be kept to a minimum, and the maximum allowable Δt should be used.

Efficient modeling of stirred tank flows in a fully Eulerian fashion already received considerable interest in literature. Hence, we focus here on the Lagrangian aspects of the simulation. These aspects have received less interest as, in the conventional applications of Euler-Lagrange methods, the parcel computations are typically a minor contribution to the total computation time. In our particular application this is, however, not the case. Under the simplification that the flowfield is in steady state, it is not required to update the flowfield every timestep, leaving operations related to parcel tracking as the major contributor to the total computation time. It must be noted that if this steady-flowfield simplification cannot be made, the Eulerian updating is indeed the dominant factor in computation time. This assumption, and other aspects related to the Eulerian implementation, are presented in section 3.1.

Computational demands of parcel tracking. Typically, the Stokes number for a micro-organism $St < 0.01$ and massless parcel tracking may be used. This means parcels immediately adapt to the local flow velocity and no force balance has to be solved, reducing the computation time considerably. As such, the massless tracking methodology is relatively cheap. We will now briefly consider the computational cost; a more in-depth assessment including the background on the numbers presented here is found in Appendix A. Running parcel tracking only simulations on one core of a 2.4GHz XEON E5-2665 established that the computation time per parcel $t_{comp} \approx 5 - 10 \cdot 10^{-6}$ s without file output. Here, t_{comp} became roughly independent of N_p for $N_p > 10^5$, meaning the total calculation time scales linearly with N_p . In the case of file output, $t_{comp} \approx 20 - 30 \cdot 10^{-6}$ s when writing 5 tracking parameters to file for each parcel, each timestep. Clearly, file

output contributes significantly to the total computation time. In many cases, writing data on only a subset of the total parcel population suffices for the study of metabolic statistics, and it may not be necessary to write data with the same temporal resolution as the simulation timestepping. Taking this in consideration and writing only the necessary data can save considerable computation time. The specifics will be highly case-dependent however, and are therefore out of the scope of this paper.

Assuming an average computational demand of $t_{comp} = 10 \cdot 10^{-6}$ s/particle (including optimized file output) and a fairly typical time resolution of $\Delta t = 30$ ms/timestep, the total computation time is 67 h per hour flow-time for $2 \cdot 10^5$ parcels on a single processor; resolving one day of flowtime on a 16-core machine, assuming perfect scalability, would take around 4 days. As the runtime scales (approximately) linearly with N_p and inversely with Δt , the benefit of optimizing these two parameters is clear.

This leads to the questions: What is the minimum required number of parcels N_p and the maximum allowed timestep size Δt to achieve accurate results? In this case, accurate means the results are independent of N_p and Δt and, for a simple reaction model, comparable to results for an Eulerian simulation with a homogeneous biomass concentration C_X , with the same mesh. Preferably, we want to be able to determine the required N_p and Δt for a given simulation problem a priori. Therefore, we first turn to a theoretical approach for the prediction of N_p .

2. Predicting N_p

2.1. Artificial concentration variations

A typical CFD mesh of a reactor contains $N_c = O(10^5 - 10^6)$ grid cells where a factor 100 – 1000 difference in volume between cells is common due to local refinement. When $N_p \approx N_c$, large grid cells at all times contain tens of parcels while the smallest cells are rarely visited. This leads to artificial spatial biomass variations, as the amount of biomass in a grid cell is directly proportional to the amount of parcels in that cell. Taking $N_p \gg \gg N_c$ gives an approximately homogeneous biomass distribution, but the large N_p required will clearly lead to excessive computational demands in all but a few applications. Since the uptake of substrate(s) is directly coupled to the biomass availability, artificial biomass variations inherently give rise to artificial substrate concentration gradients in the vicinity of the parcel, as indicated in figure 2 a. This need not be a problem if mixing in the direct surrounding of the parcel is sufficiently fast compared to reaction, such that the magnitude of artificial substrate gradients is kept small. In order to estimate how many parcels are required to make the effect of artificial substrate gradients negligible, we first need to estimate the magnitude of the gradient around the particle as a function of process parameters.

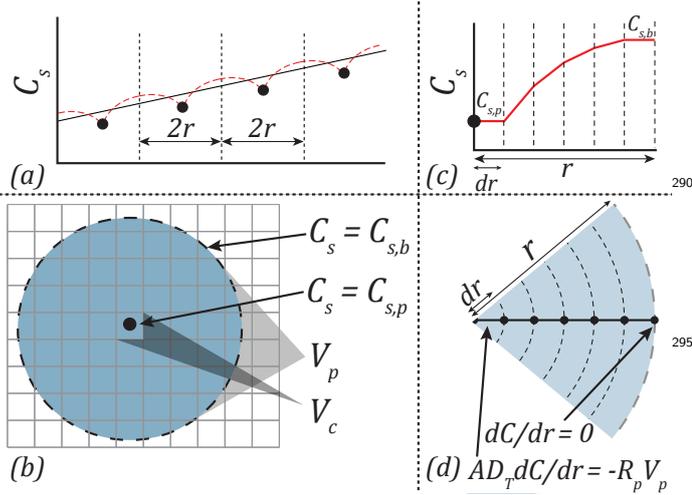


Figure 2: An outline of the computational problem: a) 1-D representation of computational substrate gradients. The black line represents the physical concentration profile. The dashed red line illustrates the concentration profile that might be found with parcel-coupled reactions. b) Hypothetical picture of a single parcel in a fine 2-D grid, outlining the defined volumes and concentrations. c) Enlarged view of the substrate concentration profile around a single spherical parcel assuming radial symmetry. The dashed lines indicate the grid cells. d) translation from the hypothetical image to the numerical implementation, including boundary conditions.

2.2. Predicting artificial gradients

Consider a 3-D domain with volume V_T containing N_p point-like biomass parcels with an equidistant particle spacing. Although the point parcels themselves have no true volume, each parcel has an associated volume $V_p = V_T/N_p$; the parcel represents the biomass that is distributed in this associated volume V_p . We study the concentration field in this volume V_p which, for simplicity, we assume as a spherical domain with radius r (figure 2 b). In reality spherical volumes cannot fill the entire domain of course, but we ignore this fact for the sake of the derivation. The background domain is discretized in grid cells with $V_c < V_p$. Because all biomass inside the volume V_p is concentrated in the central gridcell containing the parcel, substrate uptake only occurs inside this gridcell. The substrate concentration at the edge of V_p is assumed constant and homogeneous with a value $C_{s,b}$. Due to uptake, the concentration in the center $C_{s,p} < C_{s,b}$, causing a substrate flux from the domain edge to the center, eventually resulting in a steady concentration gradient within V_p . This is the artificial concentration gradient arising from the point-like biomass distribution; the situation is sketched in figure 2 c.

We relate concentrations at the edge and center of V_p as $C_{s,b} = (1 + \beta)C_{s,p}$, and use β to quantify the magnitude of the artificial gradient. Below, we will relate β to hydrodynamic and metabolic parameters. Subsequently, we consider that in reality the parcels are not stationary and homogeneously distributed. The random fluctuation of the number of parcels in a cell, $N_{p,c}$, is accounted for by

the standard deviation of the binomial distribution. Now, we turn to predicting the value of β .

Predicting β . The substrate gradient around a parcel depends on the balance between local mixing and reaction; we assume β is a function of the timescales describing these phenomena. Assuming mixing within V_p is governed by turbulent diffusion, D_t , penetration theory dictates $r_p = \sqrt{\pi D_t \tau_m}$; rewriting in terms of V_p and solving for τ_m gives eq. 1. The reaction timescale in gridcell c containing $N_{p,c}$ parcels, with parcel-bound reaction rate $R_{s,p}$ is given in eq. 2.

$$\tau_{m,p} = \left(\frac{3}{4\pi}\right)^{2/3} \frac{1}{\pi D_t} V_p^{2/3} \quad (1)$$

$$\tau_{r,c} = \frac{C_{s,c} V_c}{N_{p,c} R_{s,p} V_p} \quad (2)$$

Since we now study a single parcel, $N_{p,c} = 1$. We apply Monod kinetics with C_X being the biomass concentration: $R_{s,p} = k_{s,max} C_X \cdot [C_{s,p}/(K_s + C_{s,p})]$. These kinetics are often encountered in biological systems [15]. When $C_{s,p} \ll K_s$, these reduce to 1st order kinetics. Computations showed that β becomes independent of $C_{s,p}$ in this regime. For higher $C_{s,p}$, we find that β decreases as $C_{s,p}$ increases, converging to $\beta = 0$ for $C_{s,p} \rightarrow \infty$. Hence, we only have to consider the 1st order regime where β reaches a maximum ($\beta = \beta_m$), as a worst case-scenario to derive a criterion for N_p . Any N_p criterion that holds for $C_{s,p} \ll K_s$ also holds for higher values of $C_{s,p}$.

To study how β_m depends on the mentioned timescales, we solve the 1-D diffusion equation in MATLAB (version 8.2); A schematic view of the implementation is shown in figure 2d. Reaction takes place at the central node, leading to a boundary condition $AD_t \frac{dC_s}{dr}(t, 0) = -R_{p,c} V_p$. Two different boundary conditions have been applied at domain edge. Simulations were conducted with fixed $C_{s,b}$ - yielding a steady-state solution - or symmetry boundary condition $\frac{dC_s}{dr}(t, r_p) = 0$. The latter gives a transient batch process. As β_m is independent of C_s , both yielded similar results. Based on a large number of simulations with varying domain volume, hydrodynamic parameters and kinetic parameters, we find the relation of eq. 3 from the data shown in figure 3. The different symbols correspond with different numbers of nodes between simulations, their relative offset is attributed to the crude discretization.

$$\beta_m = \frac{4\pi^2}{3} \left(\frac{V_c}{V_p}\right)^{2/3} \cdot \frac{\tau_{m,p}}{\tau_{r,c}} \quad (3)$$

2.3. Distribution of parcels in a volume

Because of turbulent motion, the distribution of parcels inside the volume V_T is not equidistant, but more or less random and the probability of finding $N_{p,c}$ parcels in a volume V_c follows from the binomial distribution. For the

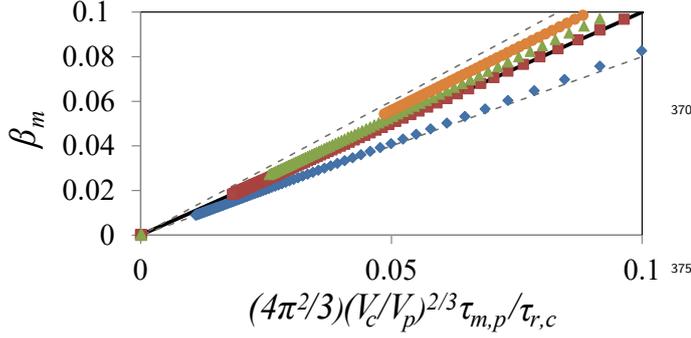


Figure 3: Calculated β_m versus predicting equation for a variety of settings. The different symbols indicate numerical resolution: $n_g = 3$ (blue diamonds), 5 (red squares), 9 (green triangles), 13 (orange circles). Black line: $X = Y$. Dashed lines: 20% deviation from the proposed relation

magnitude of artificial gradients the absolute number $N_{p,c}$ is not relevant; we have to consider the number of parcels inside a cell, compared to neighboring cells. This will fluctuate as the parcels are in constant motion. Therefore we use the standard deviation of the binomial distribution, eq. 4 to describe the effect of non-homogeneous distribution and parcel motion.

$$\sigma_{p,c} = \left(N_p \frac{V_c}{V_T} \left(1 - \frac{V_c}{V_T} \right) \right)^{1/2} \quad (4)$$

2.4. Quantifying artificial gradients in CFD simulations

In CFD software we cannot observe the artificial gradients surrounding a parcel directly (unless N_p is really small). Due to these gradients, the concentration C_s in any given gridcell c will be somewhat lower when it contains a parcel, then when it contains none. The artificial concentration gradients near parcels will hence result in temporal concentration fluctuations in each gridcell, which are easily quantified. Such variations are absent when the biomass phase is homogeneously distributed. To quantify these fluctuations we consider the concentration C_s in every gridcell c as a function time. The Coefficient-of-variation (COV_c) of concentration variations in cell c is defined as the ratio of the temporal standard deviation $\sigma_{s,c}$ over the temporal mean $\overline{C_{s,c}}$. Subsequently, COV_c is volume-averaged to yield a single number quantifying the fluctuations in the simulation, shown in equation 5.

$$\chi = \langle COV_c \rangle \equiv \left\langle \frac{\sigma_{s,c}}{\overline{C_{s,c}}} \right\rangle \quad (5)$$

With χ describing artificial gradients in simulations, we pose the proportionality $\chi \propto \langle \beta_{m,c} \sigma_{p,c} \rangle$ with a yet to be determined proportionality constant α . $\langle \beta_{m,c} \sigma_{p,c} \rangle$ is determined by solving eq. 3 and 4 for each gridcell, and volume-averaging their product over all cells. The constant α describes all effects that have not been accounted for in $\langle \beta_{m,c} \sigma_{p,c} \rangle$, including the fact that all above computations assumed the artificial gradient around the parcel

was completely developed, while in reality it will be very dynamic. Now combining $\chi = \alpha \langle \beta_{m,c} \sigma_{p,c} \rangle$ with equations 1 - 4, we can solve for N_p , yielding equation 6. For convenience, the numerical constant $4\pi^2/3$ is absorbed in α in this equation and the rest of this work; accordingly, β_m is scaled to $\beta_m^* = \beta_m / (4\pi^2/3)$. In eq. 6, the allowable artificial gradient magnitude is specified via χ (we use $\chi = 0.05$ as a default), and constant α has yet to be determined.

$$N_p = \alpha^2 \left[\frac{k_{s,max} X}{K_s \chi \sqrt{V_T}} \sum \left(\frac{V_c^{7/6}}{D_{t,c}} \left(1 - \left(\frac{V_c}{V_T} \right) \right)^{1/2} \right) \right]^2 \quad (6)$$

3. CFD implementation

Simulations were performed in the commercial software package ANSYS FLUENT 15.07, which includes Euler-Lagrange modeling as the DPM model. FLUENT is equipped with a parcel-reaction model, but this model is not compatible with the desired massless tracking. Hence, reaction routines were added as User-Defined Functions (UDFs). To allow direct comparison with Eulerian simulations for validation of the Euler-Lagrange implementation, the reaction rate used in this work is a function of a single external scalar only: $R_{s,p} = f(C_{s,c})$. Two different schemes for inter-phase reaction coupling were tested. In Approach I, the source term $S_{s,c}(t + \Delta t_c) = \sum N_{p,c} R_{s,p}(t) \Delta t_p$ with $\Sigma(\Delta t_p) = \Delta t_c$ and Δt_p a flexible timestep size determined by FLUENT. The value of the $S_{s,c}$ is determined at the beginning of each timestep, and is not updated in subsequent iterations of the Eulerian field. In essence, this makes approach I non-iterative. In Approach II, $S_{s,c}(t + \Delta t_c) = \sum N_{p,c} R_{s,p}(t + \Delta t_c) \Delta t_c$ where $S_{s,c}$ is recalculated per iteration, possibly increasing accuracy but also calculation time.

3.1. Computational setup

To keep computational demands to a minimum, a Reynolds Average Navier Stokes (RANS) approach is favored over a more detailed, but more demanding Large Eddy Simulation (LES) approach. Although RANS models require several modeling assumptions and approximations, their engineering accuracy is sufficient for providing insight in reactor-scale substrate concentration gradients [3]. The standard $k - \epsilon$ model was used in all simulations. Although this model is fundamentally weak for strongly rotating flows [16], it has been well established that the $k - \epsilon$ model can reasonably the fluid flow [17], turbulence quantities [18] and mixing time [19]. Our own simulations yield comparable results; a brief flow validation section is added in Appendix B. While performance of the $k - \epsilon$ is sufficient for our goals, the guidelines presented in this paper are by no means limited to this turbulence model; they are compatible with any eddy viscosity model.

The steady state Multiple Reference Frame (MRF) model for impeller simulation is favored over the inherently transient Sliding Mesh (SM) approach. In many stirred tank

applications, the periodic velocity component arising from rotor-stator interaction is small and the MRF method can be applied [18]. As the MRF method calculates a steady state flowfield, there is no need to recalculate the flow every timestep saving an order of magnitude in computation time. Spatial discretization was 2^{nd} order upwind in space [18] and 2^{nd} order implicit in time. The SIMPLE algorithm was used for pressure-velocity coupling. Convergence was assumed if all residuals were below 10^{-5} and oscillations in the mean velocity magnitude were below 1%.

Reaction and parcel tracking. Species tracking was initially enabled including a volumetric (Eulerian) reaction term $R_{s,c} = k_{s,max} C_X \cdot (\frac{C_{s,c}}{K_s + C_{s,c}})$. The turbulent Schmidt number was kept at the default value of 0.7. After the concentration field converged to a steady state, the solver was switched to transient and massless parcels were added via the DPM model. The effect of turbulence was included using the *discrete random walk* model (DRW), where $\sqrt{u'^2} = \sqrt{2k/3}$ and isotropic turbulence is assumed. During parcel tracking, the baffles and impeller boundary conditions were switched from wall to interface. As k vanishes at these walls, parcels tend to get stuck when the wall boundary condition is retained. Parcels were first distributed for 30 impeller revolutions without particle-coupled reactions. After this time, the volumetric reaction model was switched off, and the particle-coupled reaction model was enabled.

Physical conditions. Water ($\rho = 1000 \text{ kg/m}^3$, $\mu_l = 1 \text{ mPa} \cdot \text{s}$) was used as a working fluid in all simulations. Glucose ($MW = 180 \text{ g/mol}$) was the substrate. It was assumed that the dissolution of glucose did not influence μ_l and ρ of the continuous phase. The molecular diffusion coefficient D_m was arbitrarily set to 10^{-9} ; the exact value is not of significant influence in turbulent simulations. Glucose uptake by the filamentous fungus *P. chrysogenum* was selected as a model reaction, described by Monod kinetics. De Jonge et al. determined the model parameters to be $k_{s,max} = 1600 \text{ } \mu\text{mol}/(\text{g}_{dw} \cdot \text{h})$ and $K_s = 7.8 \text{ } \mu\text{mol}/\text{L}$ [20]. In our simulations the biomass concentration is set to $C_X = 10 \text{ g/L}$. Since we assume a continuous fluid with the properties of water for model-development purposes here, mass transfer limitations towards the microorganism will play a negligible role [21] and are not included in the simulations.

As a hypothetical reactor, a flat bottomed tank with $H = T = 5 \text{ m}$ agitated by a 6-blade Rushton turbine (RT) with a clearance $\Delta C = T/4$, diameter $D = T/3$ and baffle width was $T/10$ was used. To reduce computational effort $1/6^{th}$ of the tank was modeled, containing a single blade and baffle. The additional baffling was shown to be of negligible influence [18]. Three hexahedral grids with respectively 51k, 166k and 341k gridcells, and one tetrahedral grid with 147k cells were used. All internals were modelled as thin surfaces. Substrate was fed to the

Case	RPS	No. Parcels	$\Delta t(\text{s})$	coupling
51kcI	1.026	5k - 2000k	0.001 - 5	I
51kcII	1.026	10k, 500k	0.01 - 3	II
51kcLP	0.750	10k - 100k	0.01	I
166kcI	1.026	10k, 500k	0.01	I
341kcI	1.026	10k, 500k	0.01	I
341kcII	1.026	10k, 500k	0.01	II
TET	1.026	10k, 500k	0.01	I
PBT	1.033	10k - 500k	0.01	I

Table 1: An overview of all conducted simulations. Several cases were conducted in duplo. Case names are based on mesh ($xxxk$) and parameter variation (I = coupling approach I, II = coupling approach II, LP = Low Power input, TET = tetrahedral, PBT = Pitch Blade Turbine).

domain via a source term in a region near the impeller top. In all simulations, the conditions were such that $\langle C_s \rangle \ll K_s$. Additional simulations were performed in a pitch blade tank (PBT), to test if the proportionality constant α is influenced by the geometry. The used dimensions are $\Delta C = T/2$, $D = 0.35T$ and $T = 5 \text{ m}$; the impeller contained 4 blades with an angle of 45° . For further geometric details, we refer to Bakker et al. [22]. The mesh was fully hexahedral with 99k grid cells and covered $1/4^{th}$ of the domain.

Monitoring. Each timestep, the mean concentration $\langle C_s \rangle$, total uptake rate $\Sigma R_{s,p} V_p$, concentration and parcels per gridcell ($C_{s,c}$, $N_{p,c}$) and concentration observed by each parcel ($C_{s,p}$) were monitored. To account for transients, the first 15 impeller revolutions with Lagrangian reaction were discarded - in all cases the concentration field was statistically steady after this time. All properties were monitored for 60 – 300 impeller rotations to derive the reported statistics.

Cases. Most simulations were conducted at an agitation speed of $N_s = 1.026 \text{ RPS}$; one series of simulations was executed with $N_s = 0.75 \text{ RPS}$ to test the influence of power input. The typical timestep was set to $\Delta t = 0.01 \text{ s}$, but other values have been applied to test the influence of Δt . An overview of cases is shown in table 1.

4. Results and discussion

We focus on four aspects of the simulations: 1) we compare χ against $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$ to study the value of α under different conditions, 2) we compare mean concentration $\langle C_s \rangle$ between Eulerian and Lagrangian simulations to validate the Lagrangian reaction approach, 3) we study the effect of Δt and 4) we check whether the interphase mass balance is closed.

4.1. Artificial gradients in FLUENT

Artificial gradients were quantified by monitoring $\sigma_{s,c}/\sqrt{C_{s,c}}$ in each grid cell, and averaging them to yield χ . Data was

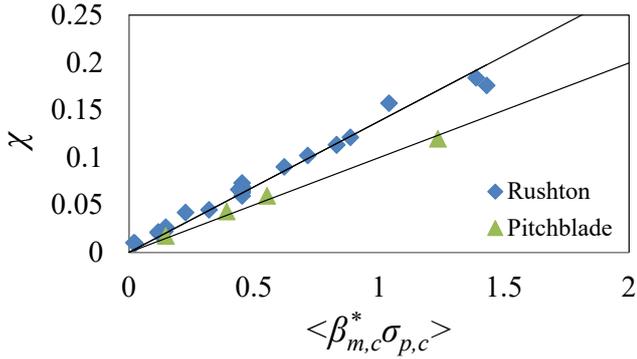


Figure 4: Observed fluctuations versus predicted value for the Rush-550 ton geometry and PBT geometry. The black lines show the obtained fits used for determining α .

510 stored every timestep and post-processing was performed in MATLAB. In figure 4 χ is plotted against $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$. 555 We observe a linear relation between $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$ and χ , in line with our expectations. The proportionality constant α is found to be 0.27 for the RT geometry and 0.19 for the PBT. This indicates α is dependent on the fermentor geometry. This is attributed to the difference in 560 global flow pattern and thereby global parcel motion. In the RT case, runs with different meshes (hexa- and tetrahedral), different power input and different Δt all collapse on the same line. Also, we see no influence of the coupling 520 mechanism on χ and thereby α . 565

Additional runs were conducted in a tank stirred by 2 Rushton turbines, using the geometry of [23]. Here, α is also found to be 0.27, which is not surprising as the geometry can be seen as 2 single-RT tanks stacked together. 525 Finally, sliding mesh simulations too yield $\alpha = 0.27$. For brevity, a detailed discussion of the sliding mesh results is omitted. 570

4.2. Mean concentration 575

530 A second prerequisite for the successful application of parcel-based reactions, is that the time-average substrate concentration field agrees with the steady state substrate concentration field found in a fully Eulerian simulation, for the simple kinetics applied here. To establish this, we com-580 pare the volumetric mean substrate concentration for the Lagrangian and Eulerian reaction implementation. First we consider how $\overline{\langle C_{s,c} \rangle}$ (the Eulerian mean in EL simulations) develops with N_p , by comparing $\overline{\langle C_{s,c} \rangle}$ with $\langle \beta_{m,c}^* \sigma_p \rangle$ (figure 5). As can be expected, $\overline{\langle C_{s,c} \rangle}$ con-585 verges to a steady value for $\langle \beta_{m,c}^* \sigma_p \rangle \rightarrow 0$, with $\overline{\langle C_{s,c} \rangle}$ nearly independent of N_p for $\langle \beta_{m,c}^* \sigma_p \rangle$ below 0.5. 590

In table 2 we compare $\langle C_{s,c} \rangle$ for an Eulerian simulation (first column), with the Lagrangian $\overline{\langle C_{s,p} \rangle}$ and $\overline{\langle C_{s,c} \rangle}$ for the same mesh. The agreement between $\langle C_{s,c} \rangle$ and $\overline{\langle C_{s,p} \rangle}$ is expected; with the essentially 1st order-595 kinetics, there is a unique $\langle C_{s,c} \rangle$ at which substrate

Case	$\langle C_{s,c} \rangle$	$\overline{\langle C_{s,c} \rangle}$	$\overline{\langle C_{s,p} \rangle}$
51kcE-100kp	1.31	0.522	1.27
166kcE-500kp	1.23	1.14	1.23
341kcE-500kp	1.23	1.31	1.23

Table 2: Mean concentration in the volumetric reaction case versus Lagrangian reaction case for selected cases. For the Lagrangian case, both the field mean, and mean observed by parcels is shown. For all cases, $\Delta t_c = 0.01$.

uptake and feed balance. This is the average concentration in an Eulerian simulation, and must also be the average concentration observed by the parcels in a Lagrangian simulation, regardless of N_p . The small difference between the meshes shows the Eulerian solution is not yet mesh independent for the 51k grid - also observed in Appendix B. More striking is the significant mesh dependency of $\overline{\langle C_{s,c} \rangle}$. We attribute this effect to the DRW model for turbulent parcel motion.

As noted before, the turbulent kinetic energy k vanishes near walls, and the convective velocity is low consequently parcels tend to have an unrealistically long residence time in these zones [9]. The effect is most pronounced in the crudest mesh, which has a relatively large ratio of boundary gridcells over bulk gridcells. The result is a somewhat elevated average biomass concentration in some wall regions, most notably the vessel top - near the substrate feed. This leads to a comparatively large uptake in this region and thereby lower $\langle C_s \rangle$ in the bulk of the vessel. Higher mesh densities result in fewer wall cells compared to bulk gridcells and the offset decreases with increasing mesh resolution. Improvements over the DRW have been suggested, such as anisotropic-DRW or continuous random walk (CRW)[24] and probabilistic methods such as explored by Lapin [9]. We consider the implementation of such methods outside the scope of this work. The main message is that the agreement between $\overline{\langle C_{s,c} \rangle}$ and $\langle C_{s,c} \rangle$ should be checked, as an offset is not fixed by simply increasing N_p . If an offset is detected, it should be checked whether parcels accumulate in certain regions, and measures (either mesh refinement or alternative tracking models) should be implemented.

Turning back to figure 5, increase in $\overline{\langle C_{s,c} \rangle}$ at high $\langle \beta_{m,c}^* \sigma_p \rangle$, can be explained by the N_p -independent value of $\overline{\langle C_{s,p} \rangle}$. Looking back at figure 2 b and c, the concentration in the parcel-containing grid cell is lower than in the surrounding cells as a result of uptake, an effect that becomes more pronounced with decreasing N_p ; hence upon decreasing N_p , $\overline{\langle C_{s,c} \rangle}$ will increase. This notion still holds for the 51k grid.

4.3. Interphase mass balance

Having established the value of α , dependence of $\overline{\langle C_{s,c} \rangle}$ on N_p we have established the minimum required N_p for an accurate Euler-Lagrange bioreactor simulation. A final important aspect is the closure of the inter-phase mass balance. The substrate uptake from the field $\Sigma S_{s,c} V_c$, must

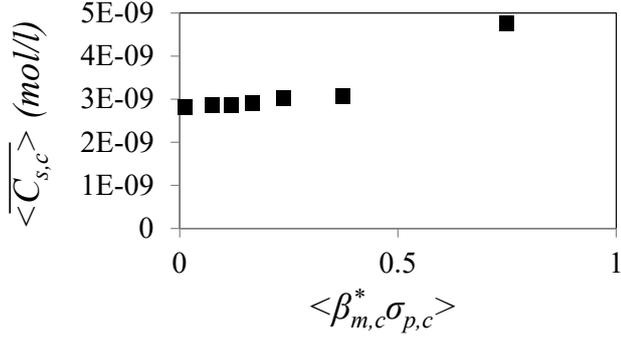


Figure 5: Mean concentration as a function of $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$ showing the convergence of the $\langle C_s \rangle$ as a function of N_p , for the $51kcE$ cases. The concentration is nearly independent of N_p for $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$ below 0.5. For all cases, $\Delta t_c = 0.01$.

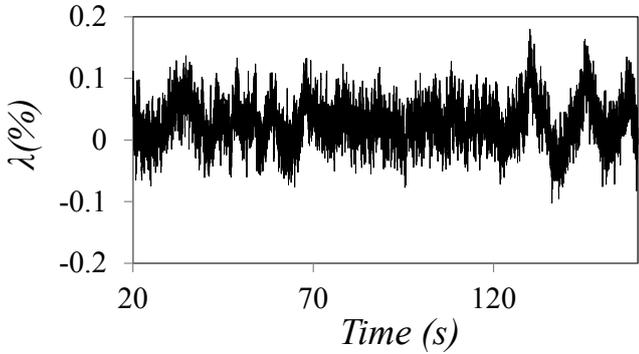


Figure 6: Signed inter-phase mass imbalance λ versus time for case $166kcI - 500kp$. The variations between timesteps are mainly due to numerical errors in the iterative solution procedure. Overall, λ is slightly positive with mean offset is $2.8 \cdot 10^{-2}\%$. This offset occurs when in some gridcells, substrate uptake exceeds availability and the uptake must be clipped. The low offset magnitude indicates this effect is of negligible.

be balanced by an equal accumulation term ($\Sigma R_p V_p$) in the Lagrangian phase in order to properly study intracellular reaction dynamics when more complex models are used. In practice, some imbalance will occur, due to numerical errors in the iterative solution of the scalar field, and due to uptake clipping. The latter occurs when the uptake from a gridcell, $S_{s,c} V_c \Delta t_c$, exceeds the availability in that cell, $C_{s,c} V_c$.

We quantify the instantaneous uptake imbalance as $\lambda(\%) = 100 \cdot \frac{R_{p,c} \Delta V_p \Delta t - S_{s,c} \Delta V_c}{R_{p,c} \Delta V_p \Delta t}$. A typical λ versus time plot is shown in figure 6. Both positive and negative imbalances are observed, the average $\bar{\lambda}$ is slightly positive. This is attributed to the two types of error mentioned above; the iteration error is distributed around 0, clipping errors lead to a positive λ . In all cases where $\chi < 0.05$, the absolute time average mass imbalance $|\bar{\lambda}|$ was well below 1%, indicating the inter-phase mass balance is sufficiently closed.

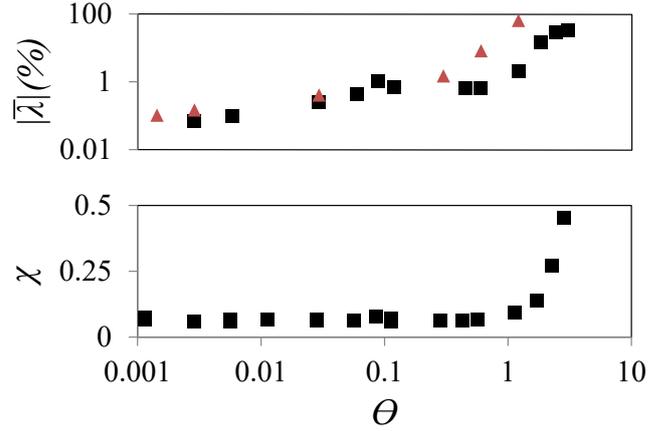


Figure 7: Uptake mass imbalance E (top) and artificial gradients χ (bottom) as a function of time. Both parameters are shown to be independent of Δt_c for $\theta < 1$. All shown datapoints are for the $51kc$ mesh. Black squares: Coupling approach I. Red triangles: Coupling approach II.

4.4. Time step size

As noted in the introduction, the time resolution Δt_c will significantly affect the computation time, given a certain flowtime; a higher Δt_c simply means less timesteps are required for the same simulation. The influence of timestep size Δt_c has been studied in the $51kc$ Rush-ton grid. First, the effect of Δt_c on the mass imbalance $|\bar{\lambda}|$ is briefly considered. We scale Δt with the reaction timescale, $\theta = \frac{\Delta t c k_{s,max} X}{K_s}$; for $\theta > 1$ significant clipping is expected since the uptake per timestep is bigger than the substrate availability. The effect of clipping is clearly visible in figure 7 (top). The value of $|\bar{\lambda}|$ is sufficiently small (consistently $< 1\%$) for $\theta < 0.1$. It is noteworthy that coupling approach I (non-iterative) and II (iterative) yield no significant difference in error, meaning the additional computational burden of recalculating uptake each field iteration is not justified. These approaches are outlined in section 3.

To conclude, the effect of timestep size on the χ is checked. The data in figure 7 (top) clearly shows that for $\theta < 1$ there is no effect of Δt on χ . This is in line with our expectations; the parcel timestep Δt_p (controlling uptake and motion) is determined independent of the field timestep Δt_c ; as long as there is no significant change in the external scalar field during Δt_c , no effect on χ is expected. Still, for a too large Δt_c parcel trajectories could not be computed successfully - this occurred when $\Delta t_c \approx N_s^{-1}$. Hence, we advise a maximum timestep size $\Delta t_c < \frac{1}{10N_s}$ and $\frac{\Delta t_c k_{s,max} X}{K_s} < 0.1$ to ensure successful trajectory calculation, time-independent χ and an acceptably low imbalance in the inter-phase mass balance.

5. Conclusion

Due to the large number of timesteps required when running an Euler-Lagrange simulation of a fermentation

645 process Lagrangian reaction coupling, computational efficiency is essential. Inherently, the clustering of biomass in Lagrangian parcels leads to artificial substrate concentration fluctuations in the substrate concentration field. A sufficient number of parcels N_p must be used to minimize⁷⁰⁰ these effects, with higher N_p inherently leading to a longer computation time. In this work, we set out to predict the minimum value of N_p , as well as the maximum allowable timestep size Δt_c , necessary to acquire accurate simulation results at minimal computational cost. ⁷⁰⁵

655 Artificial substrate concentration fluctuations, caused by non-homogeneous distribution of the biomass phase in Euler-Lagrange(EL) simulations, scale linearly with $\langle \beta_{m,c}^* \sigma_{p,c} \rangle$. Here, $\beta_{m,c}^*$ is derived from the substrate gradient around a single, stationary parcel and $\sigma_{p,c}$ is the standard deviation of the binomial distribution, describing the random distribution of parcels. The proportionality constant α is found to depend solely on the geometry of the system; no effect of mesh, power input or timestep size was observed, provided the timestep size was below $K_s/k_{s,max}X$. Knowing α , the number of parcels N_p required to keep the artificial substrate fluctuations below a certain threshold can be calculated from equation 6, where N_p typically is of $O(10^5)$ for a threshold $\chi = 0.05$. Because of the iterative nature of the method, there is a mass-imbalance $|\bar{\lambda}|$ between the Eulerian and Lagrangian uptake. This imbalance was found to be $< 1\%$ when $\alpha < \beta_{m,c}^* \sigma_{p,c} < 0.05$ and $\Delta t < 0.1K_s/k_{s,max}X$, for all cases. Both explicit and implicit uptake coupling were tested. No consistent improvement in performance was found regarding error or fluctuations for the implicit method, while a significant increase in calculation time was observed. Hence, we provide the following primary guidelines for setting up Euler-Lagrange fermentation simulations:

1. Calculate the minimum number of required parcels using equation 6.
2. A timestep size $\Delta t < \min\left(\frac{0.1K_s}{k_m X}, \frac{1}{10N_s}\right)$ is recommended.
3. Calculation inter-phase exchange terms only during the first iteration of a timestep is recommended to reduce computation time.

The size of microorganisms allows massless parcel tracking to be used, with reaction models supplied via user-defined functions, resulting in significant savings in computation time. Furthermore, computation time can be reduced by using the Multiple Reference Frame method for impeller simulation; the impeller and baffles have to be switched from 'wall' to 'interface' boundary condition to prevent parcels from getting stuck. Use of the discrete random walk model for turbulence is not advised. Although its drawbacks can be diminished by increasing mesh density, a more systematic solution is preferred.

6. Acknowledgments

We want to thank our colleagues at ECUST Shanghai, DSM Sinochem pharmaceuticals and the DSM biotechnology center for our fruitful collaboration. Special thanks to Dr. Amit T. Deshmukh for hosting our discussions, and drafting this work. Prof. Joseph J. Heijnen and Dr. Saša Kenjereš are acknowledged for their discussions on the biological aspects and parcel turbulence, respectively. Thanks to Prof. Matthias Reuss for our discussions on fluid-reaction coupling. This work has been conducted within a multi-party research project, between DSM-Sinochem Pharmaceuticals, TU Delft, East China University of Science and Technology and Guojia, subsidized by NWO and MoST. All sponsors are gratefully acknowledged.

Roman	Units	Description
T	m	Tank diameter
A	m ²	Area
C_s	mol/m ³	Substrate concentration
C_X	kg/m ³	Biomass concentration
D_t	m ² /s	Turbulent diffusion coeff.
D_m	m ² /s	Molecular diffusion coeff.
ΔC	m	Off-bottom clearance
D	m	Impeller Diameter
H	m	Tank height
K_s	mol/m ³	Affinity constant for s
k	m ² /s ²	Turbulent kinetic energy
$k_{s,max}$	–	Max. reaction rate of s
M	N · m	Impeller moment
N_c	–	Total no. grid cells
N_p	–	Total no. parcels
$N_{p,c}$	–	No parcels in cell c
N_s	1/s	Impeller revolutions
n	–	Radial divisions
q_s	mol/g/h	biomass specific glucose uptake rate
q_p	mol/g/h	biomass specific production rate
$R_{s,c}$	mol/m ³ /s	Vol. reaction of s
$R_{s,p}$	mol/m ³ /s	Parcel-coupled reaction of s
r	m	Radius
$S_{s,c}$	mol/m ³	Source of s in cell c
t	s	Time
V	m ³	Volume
Re	–	Reynolds number
St	–	Stokes number
Sc_t	–	Turbulent Schmidt no.
Po	–	Power number

Greek	Units	Description
α	–	Proportionality parameter
β	–	Theoretical error
$\beta_{m,c}$	–	Max. β in cell c
$\beta_{m,c}^*$	–	$\beta_{m,c}/(4\pi^2/3)$
χ	–	Mean fluctuations ($\langle \overline{COV_x} \rangle$)
ϵ	m^3/s^2	Turbulent energy dissipation
$\tau_{r,c}$	–	Reaction time in cell c
$\tau_{m,p}$	–	Parcel-mixing time
σ_x	–	St. dev. of parameter x
ρ	kg/m^3	Density
μ_l	$\text{Pa} \cdot \text{s}$	Molecular viscosity
μ	$1/\text{s}$	Biomass specific growth rate
λ	%	Interphase substrate imbalance
θ	–	Mean reaction time $\frac{\Delta t_c k_s \max X}{K_s}$
Subsc.	Units	Description
t	–	Turbulent
T	–	Total
p	–	Particle
c	–	Gridcell
s	–	Substrate
Other	Units	Description
\bar{x}	–	Time-average of x
$\langle x \rangle$	–	Volume-average of x
COV_x	–	Coeff. of variation of x , $\frac{\sigma_x}{\langle x \rangle}$

- [1] A. Lapin, D. Müller, M. Reuss, Dynamic Behavior of Microbial Populations in Stirred Bioreactors Simulated with Euler-Lagrange Methods: Traveling along the Lifelines of Single Cells, *Industrial & Engineering Chemistry Research* 43 (16) (2004) 4647–4656. doi:10.1021/ie030786k. URL <http://pubs.acs.org/doi/abs/10.1021/ie030786k>
- [2] J. J. Heijnen, Impact of Thermodynamic Principles in Systems Biology doi:10.1007/10.
- [3] G. Larsson, M. Trnkvist, E. S. Wernersson, C. Trgrdh, H. Noorman, S. O. Enfors, Substrate gradients in bioreactors: origin and consequences, *Bioprocess Engineering* 14 (6) (1996) 281–289. doi:10.1007/BF00369471. URL <http://link.springer.com/10.1007/BF00369471>
- [4] F. Delvigne, P. Goffin, Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations., *Biotechnology journal* 9 (1) (2014) 61–72. doi:10.1002/biot.201300119. URL <http://www.ncbi.nlm.nih.gov/pubmed/24408611>
- [5] G. Wang, W. Tang, J. Xia, J. Chu, H. Noorman, W. M. van Gulik, Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses, *Engineering in Life Sciences* 15 (1) (2015) 20–29. doi:10.1002/elsc.201400172. URL <http://doi.wiley.com/10.1002/elsc.201400172>
- [6] H. Noorman, An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies., *Biotechnology journal* 6 (8) (2011) 934–43. doi:10.1002/biot.201000406. URL <http://www.ncbi.nlm.nih.gov/pubmed/21695785>
- [7] P. Neubauer, S. Junne, Scale-down simulators for metabolic analysis of large-scale bioprocesses., *Current opinion in biotechnology* 21 (1) (2010) 114–21. doi:10.1016/j.copbio.2010.02.001. URL <http://www.sciencedirect.com/science/article/pii/S0958166910000157> <http://www.ncbi.nlm.nih.gov/pubmed/20185293>
- [8] G. Wang, J. Chu, H. Noorman, J. Xia, W. Tang, Y. Zhuang, S. Zhang, Prelude to rational scale-up of penicillin production: a scale-down study., *Applied microbiology and biotechnology* 98 (6) (2014) 2359–69. doi:10.1007/s00253-013-5497-2. URL <http://www.ncbi.nlm.nih.gov/pubmed/24413978>
- [9] A. Lapin, J. Schmid, M. Reuss, Modeling the dynamics of E. coli populations in the three-dimensional turbulent field of a stirred-tank bioreactor A structured-segregated approach, *Chemical Engineering Science* 61 (14) (2006) 4783–4797. doi:10.1016/j.ces.2006.03.003. URL <http://www.sciencedirect.com/science/article/pii/S0009250906001606> <http://linkinghub.elsevier.com/retrieve/pii/S0009250906001606>
- [10] J. Morchain, J.-C. Gabelle, A. Cockx, Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors, *AIChE Journal* 59 (2) (2013) 369–379. doi:10.1002/aic.13820. URL <http://dx.doi.org/10.1002/aic.13820> <http://doi.wiley.com/10.1002/aic.13820>
- [11] J. Morchain, J.-C. Gabelle, A. Cockx, A coupled population balance model and CFD approach for the simulation of mixing issues in lab-scale and industrial bioreactors, *AIChE Journal* 60 (1) (2014) 27–40. doi:10.1002/aic.14238. URL <http://doi.wiley.com/10.1002/aic.14238>
- [12] M. Pigou, J. Morchain, Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models, *Chemical Engineering Science* doi:10.1016/j.ces.2014.11.035. URL <http://www.sciencedirect.com/science/article/pii/S0009250914006897>
- [13] P. Vrabel, R. G. van der Lans, K. C. Luyben, L. Boon, A. W. Nienow, Mixing in large-scale vessels stirred with multiple radial or radial and axial up-pumping impellers: modelling and measurements, *Chemical Engineering Science* 55 (23) (2000) 5881–5896. doi:10.1016/S0009-2509(00)00175-5. URL <http://www.sciencedirect.com/science/article/pii/S0009250900001755>
- [14] R. D. Douma, A. T. Deshmukh, L. P. de Jonge, B. W. de Jong, R. M. Seifar, J. J. Heijnen, W. M. van Gulik, Novel insights in transport mechanisms and kinetics of phenylacetic acid and penicillin-G in *Penicillium chrysogenum*., *Biotechnology progress* 28 (2) (2011) 337–48. doi:10.1002/btpr.1503. URL <http://www.ncbi.nlm.nih.gov/pubmed/22223600>
- [15] J. Monod, Bacteria growth kinetics, *Ann. Rev. Microbiol* 3 (1949) 371–377. URL <https://scholar.google.nl/scholar?start=0{&}q=monod+kinetics+1949{&}hl=nl{&}as{&}sdt=0,5{&}as{&}ylo=1948{&}as{&}yhi=1950{#}1>
- [16] H. K. Versteeg, W. Malalasekera, An introduction to Computational Fluid Dynamics - the finite volume method, Longman Scientific and Technical, Essex, England, 1995.
- [17] A. Kukuková, M. Moštěk, M. Jahoda, V. Machon, CFD Prediction of Flow and Homogenization in a Stirred Vessel: Part I Vessel with One and Two Impellers, *Chemical Engineering & Technology* 28 (10) (2005) 1125–1133. doi:10.1002/ceat.200500094. URL <http://doi.wiley.com/10.1002/ceat.200500094>
- [18] O. Gunyol, R. F. Mudde, Computational Study of Hydrodynamics of a Standard Stirred Tank Reactor and a Large-Scale Multi-Impeller Fermenter, *International Journal for Multiscale Computational Engineering* 7 (6) (2009) 559–576.
- [19] M. Coroneo, G. Montante, A. Paglianti, F. Magelli, CFD prediction of fluid flow and mixing in stirred tanks: Numerical issues about the RANS simulations, *Computers & Chemical Engineering* 35 (10) (2011) 1959–1968. doi:10.1016/j.compchemeng.2010.12.007. URL <http://www.sciencedirect.com/science/article/pii/S0098135410003686>
- [20] L. P. de Jonge, N. A. A. Buijs, A. ten Pierick, A. Deshmukh, Z. Zhao, J. A. K. W. Kiel, J. J. Heijnen, W. M. van Gulik, Scale-down of penicillin production in *Penicillium chrysogenum*., *Biotechnology journal* 6 (8) (2011) 944–58. doi:10.1002/biot.201000409.

URL <http://www.ncbi.nlm.nih.gov/pubmed/21751388>

825 [21] M. Linkès, P. Fedé, J. Morchain, P. Schmitz, Numerical investigation of subgrid mixing effects on the calculation of biological reaction rates, *Chemical Engineering Science* 116 (2014) 473–485. doi:10.1016/j.ces.2014.05.005.
URL <http://www.sciencedirect.com/science/article/pii/S0009250914002152>

830 [22] A. Bakker, K. J. Myers, R. W. Ward, C. K. Lee, The laminar and turbulent flow pattern of a pitched blade turbine (1996).
URL <http://www.bakker.org/cfm/publications/BakkerMyersWardLee.pdf>

835 [23] M. Jahoda, M. Moštk, A. Kukuková, V. Macho, CFD Modelling of Liquid Homogenization in Stirred Tanks with One and Two Impellers Using Large Eddy Simulation, *Chemical Engineering Research and Design* 85 (5) (2007) 616–625. doi:10.1205/cherd06183.
URL <http://www.sciencedirect.com/science/article/pii/S0263876207730915>

840 [24] A. Dehbi, A CFD model for particle dispersion in turbulent boundary layer flows, *Nuclear Engineering and Design* 238 (3) (2008) 707–715. doi:10.1016/j.nucengdes.2007.02.055.
URL <http://www.sciencedirect.com/science/article/pii/S0029549307003561>

845 [25] H. Wu, G. Patterson, Laser-Doppler measurements of turbulent-flow parameters in a stirred mixer, *Chemical Engineering Science* 44 (10) (1989) 2207–2221. doi:10.1016/0009-2509(89)85155-3.
URL <http://www.sciencedirect.com/science/article/pii/S0009250989851553>

Case	N_c	N_p	Δt	CPU_s	t_{comp}
A	51k	10k	0.01	1	23.85
A	51k	100k	0.01	1	8.79
A	51k	500k	0.01	1	7.54
A	51k	5000k	0.01	1	6.44
A	51k	500k	0.01	4	2.76
A	51k	500k	0.02	1	9.24
A	51k	500k	0.005	1	7.02
B	51k	10k	0.01	1	23.10
B	51k	100k	0.01	1	8.42
B	51k	500k	0.01	1	7.10
C	51k	10k	0.01	1	37.95
C	51k	100k	0.01	1	21.94
C	51k	500k	0.01	1	20.91
D	51k	10k	0.01	1	33.67
D	51k	100k	0.01	1	20.43
D	51k	500k	0.01	1	21.16
E	51k	500k	0.01	1	26.30
E	51k	500k	0.01	4	17.08
A	341k	10k	0.01	1	121.2
A	341k	100k	0.01	1	17.75
A	341k	500k	0.01	1	8.92
C	341k	100k	0.01	1	33.48
C	341k	500k	0.01	1	21.21

Table 3: Computation time t_{comp} in μs /parcel/ timestep for a variety of cases

7. Appendix A: Computation times

7.1. Computation time

855 The computation time will be a strong function of used computing platform (specifically to the CPU, and the speed of the mass storage system in cases where data is frequently written to file). The computation times presented⁸⁸⁰ here should therefore be used for relative comparison purposes only.

860 All presented simulations were executed on a desktop 8-core Intel Xeon E5-2665 system, clocked at 2.4GHz with hyperthreading disabled. 32Gb DDR3 RAM, optionally⁸⁸⁵ data was written to a Crucial BX100 SSD.

865 The following tracking cases were considered:

- A: Massless tracking (movement) only
- B: Massless tracking + reaction (1-eq) ⁸⁹⁰
- C: Massless tracking + reaction (1-eq) + data output (5-param)
- 870 • D: Massless tracking + reaction (10-eq) + data output (5-param) ⁸⁹⁵
- E: Massless tracking + reaction (1-eq) + data output (10-param)

This led to the computation times presented in figure

875 3: ⁹⁰⁰

The following observations are made:

- The computation time per parcel, t_{comp} , asymptotically approaches a minimum value upon increasing N_p . The minimum t_{comp} per parcel is estimated to be $5\mu s$ for tracking only for this system. Every timestep, an approximately fixed amount of time is required to update the Eulerian fields, and the FLU-ENT solver itself. The contribution of this step becomes smaller as N_p increases, hence the asymptotic behaviour.
- For the 51k mesh, the influence of the Eulerian updating is hardly noticeable for $N_p > 100k$. With increasing mesh size, the effect of the Eulerian step becomes more pronounced. Crudely speaking, the results indicate the Eulerian update has the most significant effect when $N_p < N_c$, and vice versa. This clearly indicates that a good mesh dependency study, such that N_c too can be kept to a minimum, is advisable. For the 341k grid, increasing N_p from 10^5 to $5 \cdot 10^5$ still leads to a $2.5x$ longer computation - showing that also in this case optimizing N_p may be very beneficial.
- The effect of the metabolic model equations on total computation time is negligible, both for a model including 1 and 10 hyperbolic reaction equations. Of course, this observation will not hold for much larger models, and will depend on the complexity of the model equations.

- File output has a very distinct influence on the computation time. Minimizing the number of output occasions, depending on the goal of the simulation and the statistical requirements is highly advised. As these settings are very case-dependent, we do not treat them here.
- The writing time increases somewhat when increasing the number of output parameters per parcel, but not proportionally; this signifies that file lookup and handling contribute significantly to the writing time.
- Increasing the number of CPUs with a factor 4 leads to a factor 2.76 reduction in computation time; although this is based on only 1 run, scalability does not seem to be perfect - but using multiple cores can bring down computation time significantly.
- When writing files, the scalability becomes poorer, as the communication with the SSD takes up a significant portion of the computation time.

8. Appendix B: Flow validation

We focused the flow validation on k and ϵ as they are often underpredicted by the $k - \epsilon$ model. Results are compared with LDA data by Wu and Patterson [25]. Although the employed scale in our simulation was much larger ($D = 5\text{m}$ versus $D = 1\text{m}$), the employed non-dimensionalization allows direct comparison with the experimental value [18]. Figure 8 shows good agreement for all hexahedral meshes. Similar agreement was found for the power numbers calculated from the impeller moment ($Po = \frac{2\pi M_y N_s}{\rho N^3 D^5}$), shown in table 4. The Power number

based on energy dissipation, $Po_\epsilon = \frac{\int \rho \epsilon dV}{\rho N^3 D^5}$, shows a greater mesh dependency. These results are in agreement with the observations of [18]. For the tetrahedral mesh, we do observe a significant under-prediction of turbulence parameters. This can lead to longer mixing times which influence substrate concentration gradients, and must be taken into account when studying real bioreactors. For our current purposes, however, all Lagrangian results are compared only with Eulerian results of the same mesh. Hence, our conclusions should not be influenced by such differences between meshes.

For the PBT, no k and ϵ data was available in the work of Bakker et al. [22]; only the observed power number was compared, and we find decent agreement (table 4). We do observe a stronger axial and weaker radial flow than Bakker et al. Since our interest is in testing reaction methods, this difference is not of current concern, and was not further addressed.

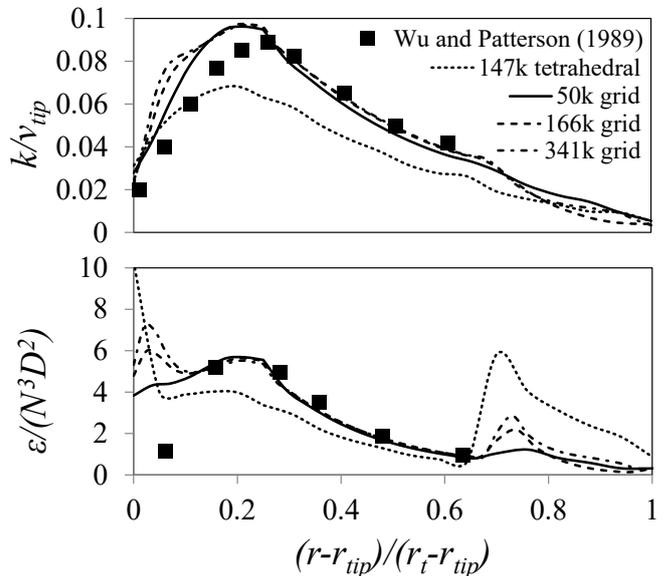


Figure 8: Dimensionless turbulence profiles in the impeller discharge stream of a Rushton profile; results at different mesh densities compared with LDA-data by Wu and Patterson.

Mesh	Po	Po_ϵ	literature
51kc HEX	5.32	4.46	5.2
166kc HEX	5.47	4.92	5.2
341kc HEX	5.53	5.11	5.2
147kc TET	4.80	3.45	5.2
99kc PBT	1.16	0.98	1.27

Table 4: Power Numbers for various meshes, based on torque and epsilon, compared with experimental value.