



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Numerical analysis of particulate flows with the finite element method

Guillermo Casas González

Supervisors:

Eugenio Oñate

Riccardo Rossi

Doctorate Program in Structural Analysis
Department of Civil and Environmental Engineering
BarcelonaTech (UPC)

This dissertation is submitted for the degree of
Doctor of Philosophy

March 2019

Dedico aquest treball, el més gran que he fet mai, als meus pares, Ana i Juan Ignacio, i a la Vera, el meu amor.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Guillermo Casas González
March 2019

Acknowledgements

Writing this doctoral dissertation often felt like writing a book, perhaps a long one. A somewhat lonely activity, a fight to keep a focus for a long time chasing a quality standard that keeps getting away. And as so often writers explain, this is best achieved in solitude: it is a lonely job.

On the other hand, one is not ready to write it the first day. I arrived at CIMNE knowing very little about programming, or physics, or about how to do research. I must thank Eugenio Oñate for, in spite of this, giving me the opportunity to enter the world of numerical methods and science in general. Eugenio has taken care of my career and has cultivated my confidence in a very generous way over these years. I am especially grateful for his welcoming me *a second time* after I came back from my failed entrepreneurial adventures.

A year into my doctorate, Riccardo Rossi kindly took me in as his candidate as co-director. He has been the one to listen to my worries and patiently help me every time I got stuck.

An advantage of working in CIMNE is that there is no shortage of smart researchers, always ready to discuss your problems. And I certainly took advantage of this during the course of this work: First of all, thank you Jordi Cotela. I owe most of my work on the extension of the finite element code to you (and also my understanding of subscale stabilization). You provided me with code, lessons and fruitful discussions. I am sorry I could not finish the work on turbulence we were collaborating on in time. . . I must also thank my other lab-mates: Julio Marti, Pavel Ryzhakov and Ricardo Reyes for their many discussions and advice.

Even before I began my doctorate, I had entered the DEMPack team. I owe a lot to my colleagues there too: Miguel Ángel Celigueta, with whom I have collaborated in many developments and held numerous stimulating discussions on programming and on the discrete element method; to Salvador Latorre, with whom I have shared many jobs; and Ferran Arrufat, who helped me with crucial developments at the end of this work. I am indebted to these people, as well as to Mercè López and Joaquin Irazábal for covering my back so many times when I was working on this manuscript. It has been a pleasure sharing all those good times during the lunch breaks with these friends.

I must thank several other researchers who have offered their help. I do not want to forget any of them: Roberto Flores, a friend and a such a talented scientist, who helped me sort out my ideas in several occasions; Enrique Ortega, who gave me valuable ideas on how to use polynomial smoothers; Ramón Codina and Joan Baiges, who helped me understand concepts on subscales stabilization and other questions; Prashanth Nadukandi, who helped me with matrix manipulations and with whom I enjoyed discussing possible research topics; and Pooyan Dadvant and Carlos Roig, who gave me programming solutions I would have taken an eternity to find.

I am convinced that the best research is done in collaboration. I must thank the very fruitful (and pleasant) collaborations I have had with Àlex Ferrer, who played a crucial role in the optimization process described in the third chapter and Ignasi Pouplana, with whom we worked on particle impact drilling.

On occasions, I encountered difficulties when reading crucial sections of the literature. Thankfully, most times the authors were kind to answer my questions. I want to thank Benoît Pouliot, Ksenia Guseva and Andrew Bragg for their patient and elaborate answers.

I spent two extremely productive months at UP Berkeley thanks to Tarek Zohdi, who kindly hosted me two times and collaborated with me in writing a paper. I thank him and Eugenio for making that exchange possible.

My research would have not been possible without financial support from Generalitat de Catalunya, under Doctorat Industrial program 2013 (DI 024). I also thank the support received from the project *PRECISE - Numerical methods for PREdicting the behaviour of CIVil StructurEs under water natural hazards (MINECO - BIA2017-83805-R - 01/01/2018 – 31/12/2020)*

Finally, I want to thank my family for their extreme patience and constant love and support.

Abstract

In this work we study the numerical simulation of particle-laden fluids, with an emphasis on Newtonian fluids and spherical, rigid particles.

Our general strategy consists in using the discrete element method (DEM) to model the particles and the finite element method (FEM) to discretize the continuous phase, such that the fluid is not resolved around the particles, but rather averaged over them. The effect of the particles on the fluid is taken into account by averaging (filtering) their individual volumes and particle-fluid interaction forces.

In the first part of the work we study the Maxey–Riley equation of motion for an isolated particle in a nonuniform flow; the equation used to calculate the trajectory of the DEM particles. In particular, we perform a detailed theoretical study of its range of applicability, reviewing the initial effects of breaking its fundamental hypotheses, such as small Reynolds number, sphericity of the particle, isolation etc. The output of this study is a set of tables containing order-of-magnitude inequalities to assess the validity of the method in practice.

The second part of the work deals with the numerical discretization of the MRE and, in particular, the study of different techniques for the treatment of the history-dependent term, which is difficult to calculate efficiently. We provide improvements on an existing method, proposed by van Hinsberg et al. (2011), and demonstrate its accuracy and efficiency in a sequence of tests of increasing complexity.

In the final part of the work we give three application examples representative of different regimes that may be encountered in the industry, demonstrating the versatility of our numerical tool. For that, we describe necessary generalizations to the MRE to cover problems outside its range of applicability. Furthermore, we give a detailed account of the stabilized FEM algorithm used to discretize the fluid phase and compare several derivative recovery tools necessary to calculate some of the interphase coupling terms. Finally, we generalize the algorithm to include the backward-coupling effects according to the theory of multicomponent continua, allowing the code to deal with arbitrarily dense flow regimes.

Table of contents

List of figures	xv
List of tables	xix
Nomenclature	xxiii
1 Introduction	1
1.1 Background and Motivation	1
1.1.1 Numerical simulations	4
1.1.2 Industrial Doctorate	9
1.2 Objectives and methodology	9
1.3 Outline of this document	11
2 The Maxey–Riley equation	15
2.1 Introduction	15
2.2 Range of validity	17
2.2.1 Inertial effects: first finite- Re_p effects	21
2.2.2 Finite radius effects	28
2.2.3 Nonsphericity effects	29
2.2.4 Effect of neighbouring particles	31
2.2.5 Small-size effects	49
2.2.6 Other effects	62
2.3 Scaling analysis	64
2.3.1 Analysis of a simplified MRE	65
2.3.2 Importance of the Faxén terms	72
2.4 Summary	74
3 The numerical solution of the Maxey–Riley equation	77
3.1 Introduction	77

3.2	Overview of approaches for the treatment of the Boussinesq–Basset term . . .	79
3.2.1	Fractional derivative approach	80
3.2.2	Hybrid polynomial interpolation/analytic approach	82
3.2.3	Comparing the accuracies of the different quadrature methods . . .	83
3.2.4	Addressing memory requirements: window methods	83
3.3	Improvements on the MAE	86
3.3.1	Introduction of quadrature substepping	89
3.3.2	How to choose the t_i parameters	96
3.4	Overall algorithm	104
3.5	The fractional calculus perspective	107
3.5.1	Exploring an idea: Richardson’s extrapolation	112
3.6	Performance of the methodology	114
3.6.1	First benchmark: an integral with analytical solution	117
3.6.2	Second benchmark: Candelier’s solution	121
3.6.3	Third benchmark: Sedimentation through synthetic vortices	128
3.7	Summary	131
4	Forward and backward-coupled particulate flows	137
4.1	Introduction	137
4.2	Beyond the MRE	141
4.2.1	Unperturbed fluid and added mass forces	143
4.2.2	Drag force	144
4.2.3	History force	144
4.2.4	Lift force	145
4.2.5	Torque	145
4.3	The continuous-phase problem	146
4.3.1	Variational form of the problem	146
4.3.2	VMS-stabilized finite element formulation	148
4.3.3	The overall algorithm	155
4.4	Derivative recovery	156
4.4.1	Overview of existing approaches	158
4.4.2	Comparison of the different recovery approaches	172
4.5	Forward coupling	179
4.5.1	Search	179
4.5.2	Interpolation	180
4.6	Application example: T-junction bubble trapping	181
4.6.1	Flow simulation	182

4.6.2	Trapping probability	184
4.6.3	Discussion	185
4.7	Application example: Particle impact drilling	188
4.7.1	Problem data	189
4.7.2	Drilling fluid model	192
4.7.3	Hydrodynamic interactions for the particles	194
4.7.4	Characteristic scales of the problem	197
4.7.5	Discussion of the coupling strategy	200
4.7.6	Final algorithm	202
4.7.7	Time and space resolution	202
4.7.8	Simulations	208
4.7.9	Discussion	223
4.8	Backward-coupled particle-laden flows	228
4.8.1	CFD-DEM model	229
4.8.2	The continuous-phase problem for the backward-coupled fluid	231
4.8.3	Finite Element Formulation	232
4.8.4	Backward coupling method	234
4.8.5	Final algorithm	237
4.9	Application example: fluidized bed	237
4.10	Summary	239
5	Conclusions and Future work	245
5.1	Concluding remarks	245
5.2	Future work	248
5.2.1	Range of applicability of the MRE and scaling analysis	248
5.2.2	More on the history force calculation	249
5.2.3	Derivative recovery	249
5.2.4	Future developments for backward coupling	249
5.2.5	Future developments for the applications	249
5.2.6	Backward Coupled CFD-DEM flows	250
5.2.7	Backward-coupling scheme	251
	References	253
	Appendix A DEM specifics	281
A.1	Basic Ingredients	281
A.1.1	Contact model	283

A.1.2	Time integration Scheme	284
A.2	Scales	284
A.3	Wear	286
A.4	Analytic tools	287
A.4.1	Analytic particles	287
A.4.2	Analytic surfaces	288
Appendix B	Quadrature substepping formulas	289
Appendix C	Alternative expression for the history force	295
Appendix D	Error bound for the kernel approximation	299
Appendix E	Quadratic character of the I_{2tH} problem	301
Appendix F	Optimal a_i and t_i values	303
Appendix G	FEM Discretization of the CFD-DEM equations	305
G.1	OSS FEM discretization	305
G.2	ASGS FEM discretization	313
Appendix H	Multicomponent theory fundamentals	315
H.1	The backward-coupled continuous problem	315
H.1.1	Single-component continuum theory: basic notions	315
H.1.2	Multi-component continuum theory: basic notions	318
H.1.3	An important particularization of the balance equations	321
H.1.4	Filtering of a particle-laden flow	322

List of figures

1.1	Numbers of particles at different scales	7
2.1	Different approximations to the drag force	23
2.2	Disturbance Stokes flow	32
2.3	Average nearest-neighbour distances	32
2.4	Drag correction for finite solid volume fraction	37
2.5	Fits for c_0 and c_1	45
2.6	Settling coefficient estimation	46
2.7	Small-size effects on the unsteady hydrodynamic force	54
2.8	Time scales of Brownian motion	57
2.9	Time evolution of the square root of the MSD	61
2.10	Ratio of the square root of the MSD over its asymptotic approximations	61
2.11	Importance of Brownian diffusion	63
2.12	Behaviour of the $(2\rho + 1)/ \rho - 1 $ factor	71
3.1	Convergence acceleration of the Daitche method	84
3.2	Time stepping process	92
3.3	Error bounds for the approximated kernel as a function of a_1 and t_1	100
3.4	Error bounds for the approximated kernel as a function of a_1 and a_2	101
3.5	Contour plot of the I_1 bound as a function of \tilde{t}_1 and \tilde{t}_2	102
3.6	Optimized I_1 values for different cost functions	103
3.7	Scaling of global error for different quadrature methods	114
3.8	Convergence acceleration of Bombardelli's method, $t = 2.5$	115
3.9	Convergence acceleration of Bombardelli's method, $t = 5$	115
3.10	Convergence acceleration of Bombardelli's method, $t = 10$	116
3.11	Convergence acceleration of Bombardelli's method, $t = 20$	116
3.12	Sinus convolution: error produced by the approximation $K \approx K_B$	118
3.13	Sinus convolution: minimal window time for an error of 1%	119

3.14	Sinus convolution: error vs. time step for Daitche's method and MAE	120
3.15	Candelier et al. benchmark: trajectories	121
3.16	Candelier et al. benchmark: average error in the trajectories (MAE vs. Daitche)	124
3.17	Candelier et al. benchmark: position error at the final time vs. memory requirements	124
3.18	Candelier et al. benchmark: error caused by substepping	126
3.19	Candelier et al. benchmark: error in radius at $t = 20$ with fixed number of historic points	127
3.20	Sedimentation through synthetic vortices: 10 000 particles' positions in time	132
3.21	Sedimentation through synthetic vortices: evolution of the difference in the x and y components in time (MAE vs. Daitche)	133
3.22	Sedimentation through synthetic vortices: Detail of positions at $t = 80$ s for different values of N_q	133
3.23	Sedimentation through synthetic vortices: Detail of positions at $t = 80$ s for different numbers of exponentials m	134
3.24	Sedimentation through synthetic vortices: wall-clock time per unit simulated second	134
4.1	Point-particle approach illustrated	138
4.2	Diagram of possible inter-phase interactions	140
4.3	Example of bad neighbours cloud	169
4.4	Derivative recovery: mesh types	173
4.5	Recovery errors for the product of sine functions	175
4.6	Recovery errors for the Ethier field	177
4.7	Geometry of circular cross-section of T-junction	182
4.8	T-junction: diagram of interactions	182
4.9	T-junction: mesh for trapping probability	186
4.10	T-junction: particles inlet mesh	186
4.11	T-junction: trapping probability	186
4.12	T-junction: pressure isosurfaces	187
4.13	T-junction: snapshot at $t/T = 26$	187
4.14	T-junction: effect of the history force	189
4.15	PID: sketch of wellbore	190
4.16	PID: diagram of interactions	190
4.17	PID: geometry	191
4.18	Rheology of power-law fluids	194
4.19	PID: Relaxation times for the steel particles	199

4.20	PID: meshes of mesh-convergence study	205
4.21	PID: point of measure (convergence study)	206
4.22	PID: mesh convergence	206
4.23	PID: internal flow mesh	207
4.24	PID: detail of mesh for nozzles	207
4.25	PID: total domain flow mesh	208
4.26	PID: evolution of the total number of particles	210
4.27	PID: snapshot at $t = 0.25$ s (internal flow)	211
4.28	PID: velocity cuts (internal flow)	212
4.29	PID: fluid streamlines (internal flow)	213
4.30	PID: particle trajectories (internal flow)	214
4.31	PID: sequence of snapshots (total domain flow)	215
4.32	PID: velocity cuts (total domain flow)	216
4.33	PID: streamlines (total domain flow)	217
4.34	PID: sequence of wear (total domain flow)	217
4.35	PID: effect of friction	219
4.36	PID: effect of friction on wear	219
4.37	PID: effect of COR	220
4.38	PID: effect of COR on wear	220
4.39	PID: effect of inter-particle contacts	222
4.40	PID: effect of inter-particle contacts on wear	223
4.41	PID: transverse velocity cuts	224
4.42	PID: standard vs. steady-state approaches (velocities)	224
4.43	PID: standard vs. steady-state approaches (snapshots)	225
4.44	PID: standard vs. steady-state approaches (detail snapshots)	225
4.45	PID: flux-measuring surfaces	226
4.46	PID: flux measurements	227
4.47	Fluidized bed: geometry and meshes	239
4.48	Fluidized bed: bubble (sequence of snapshots)	240
4.49	Fluidized bed: bubble (pressure evolution at a point)	241
4.50	Fluidized bed: diagram of interactions	241
B.1	Daitche inconsistent interpolation	290
B.2	Daitche interpolands support	290
B.3	Daitche corrected interpolation	292
B.4	Support of the second-order interpolands weighed by $f(t_i)$	292
B.5	Generic first-order interpolation	292

B.6	Support of the first-order interpolands	293
B.7	Second-order boundary interpolation	294

List of tables

2.1	Upper bounds for St	20
2.2	MRE: limits of applicability estimates	75
2.3	MRE: limits of applicability estimates with neighbours	76
2.4	MRE: relative importance of terms	76
2.5	MRE: importance of the Faxén corrections	76
3.1	Sedimentation through synthetic vortices: physical parameters	131
4.1	Derivatives required in the MRE	159
4.2	T-junction: material parameters	183
4.3	T-junction: simulation parameters	185
4.4	PID: geometric measurements	192
4.5	PID: physical parameters	192
4.6	PID: Characteristic scales	198
4.7	PID: time step choices	204
4.8	PID: standard input parameters	209
4.9	Fluidized bed: material parameters	238
F.1	Optimal a_i and t_i values (I_1 cost function)	303
F.2	Optimal a_i and t_i values (I_{2t} cost function)	304

List of Algorithms

1	Time integration algorithm for the MRE: single particle	108
2	Some functions for the time-stepping algorithm	109
3	Creation of clouds of neighbours	168
4	Creation of clouds of neighbours (modified)	169
5	PID: final algorithm	203
6	Solve fluid function	203
7	Fluidized bed: two-way coupled algorithm	238
8	Basic DEM algorithm	282
9	Solve DEM function	283

Nomenclature

Roman Symbols

a	particle's radius
A_p	particle cross-sectional area
\mathcal{B}	body (continuum mechanics)
\mathcal{B}_k	body belonging to component k (continuum mechanics)
$C_{D,\text{Shah}}$	drag coefficient in Shah's (2007) model
c_s	local sound speed
k	Stokes–Einstein–Smoluchowski diffusivity
\mathbb{E}^3	three-dimensional affine Euclidean space
f	friction factor (boundary layer)
\mathbf{F}_A	added mass force
F_A	characteristic scale of \mathbf{F}_A
\mathcal{F}_b	total body work
$\mathbf{F}_{\text{contact}}$	sum of contact forces
\mathbf{F}_D	steady drag force
F_D	characteristic scale of \mathbf{F}_D
\mathbf{F}_H	Boussinesq–Basset history force
F_H	characteristic scale of \mathbf{F}_H

F_{hydr}	hydrodynamic force
F_k	Momentum exchange between component k and the rest of components
$F_{L,\text{Herr}}$	Herron lift force
$F_{L,\text{Rub}}$	Rubinow and Keller lift force
$F_{L,\text{Rub}}$	characteristic value of $F_{L,\text{Rub}}$
$F_{L,\text{Saff}}$	Saffman lift force
F_r	Froude number
\mathcal{F}_s	total surface work
F_U	undisturbed-flow force
F_U	characteristic scale of F_U
F_W	static submerged weight
g	gravity acceleration
$H^1(\Omega)$	space of functions with square integrable derivatives in Ω
$H_0^1(\Omega)$	space of functions in $H^1(\Omega)$ vanishing on $\partial\Omega$
$\mathbf{H}_D^1(\Omega)$	space of vectors in $\mathbf{H}^1(\Omega)$ fulfilling the Dirichlet boundary conditions
I_p	spheric particle's moment of inertia
I_p	spheric particle's moment of inertia
K	consistency index
k	Boltzmann's constant
K_B	Boussinesq–Basset kernel function
Kn	Knudsen number
L	characteristic length scale of the smallest flow structures
$L^2(\Omega)$	space of square integrable functions in Ω
L_{filter}	characteristic filter length scale

\mathcal{M}	total momentum
m_f	displaced fluid's mass
m_p	particle mass
$m_{p,f}$	equivalent particle mass
n	behaviour index
\mathcal{N}	total moment of momentum
n_{dim}	number of space dimensions
n_p	number density
n_{nodes}	total number of mesh nodes
N_q	number of time steps between quadrature steps
n_{dim}	number of space dimensions
N_{Vor}	neighbourhood function
$\mathcal{P}(A)$	power set of set A
p	fluid pressure
Pe_g	Péclet number associated to the settling velocity
\mathbf{q}	heat flux
\mathcal{Q}	space of pressure scalars
\mathcal{Q}_h	discrete counterpart of \mathcal{Q}
\mathbf{r}	particle position
\mathbb{R}	field of the real numbers
\mathbb{R}^+	strictly positive real numbers
r	internal energy source/sink
Re	Reynolds number
Re_Ω	rotational Reynolds number

Re_p	particle Reynolds number
$Re_{p,\text{Shah}}$	Reynolds number for power-law fluids
$Re_{p,\text{PL}}$	particle Reynolds number for power-law fluids (Metzner and Reed, 1955)
Re_s	Shear-related particle Reynolds number
\mathbb{R}^n	Cartesian product involving n copies of \mathbb{R}
\mathbf{S}	strain rate tensor
S	settling coefficient
S^*	Brinkman over Oseen length scales squared
Sc	Schmidt number
$S_{L_{\text{filter}}}$	sphere centred at zero with radius equal to L_{filter}
St	Stokes number
St_{bubble}	Stokes number for bubble rebound
St_{coll}	collisional Stokes number
St_{Ω}	rotational Stokes number
St_L	scale-dependent Stokes number
\mathbf{t}_N	surface traction on Γ_N
T	characteristic time scale of the smallest flow structures
τ_1	mass balance-related stabilization parameter
τ_1	momentum balance-related stabilization parameter
$\tau_{\Delta t, e}$	matrix of stabilization parameters
\mathcal{T}_b	total torque due to body actions
$\mathbf{T}_{\text{contact}}$	sum of contact torques
t_{DEM}	discrete time variable for the particles
T_{filter}	characteristic filter time scale

\mathbf{T}_{hydr}	hydrodynamic torque
\mathcal{T}_s	total torque due to surface actions
\mathbf{u}	fluid velocity
U	characteristic velocity scale of the smallest flow structures
u	internal energy
\mathcal{U}	total energy
\mathbf{v}	particle velocity
\mathcal{V}_D	space of velocity vectors
$\mathcal{V}_{D,h}$	discrete counterpart of \mathcal{V}_D
V_p	particle volume
\mathbf{w}	slip velocity
\mathcal{W}_b	power spent by the body forces
\mathcal{W}_s	power spent by the surface forces
\mathcal{X}	space of solutions
$\tilde{\mathcal{X}}$	space of subscales
$\tilde{\mathcal{X}}_0$	space of subscales that vanish at the boundary
\mathcal{X}_h	discrete counterpart of \mathcal{X}
y_{PL}^+	y^+ dimensionless distance for power-law fluids

Greek Symbols

$\alpha_{p,\rho}$	particle mass fraction
β_{slip}	Basset slip coefficient
χ	configuration (continuum mechanics)
$\bar{\chi}$	motion of a body (continuum mechanics)
$\bar{\chi}_k$	motion of a body of component k (continuum mechanics)

χ_t	configuration at time t (continuum mechanics)
$\chi_{k,t}$	configuration of component k at time t (continuum mechanics)
Δt	continuous phase time step
δt	disperse phase time step
ε_{ijk}	Levi–Civita symbol
ε_p	amplitude of surface irregularities
Γ	radial distribution function
γ	local strain rate
Γ_D	subset of $\partial\Omega$ where Dirichlet boundary conditions apply
Γ_N	subset of $\partial\Omega$ where Neumann boundary conditions apply
λ	mean free path of the fluid molecules
μ	fluid’s dynamic viscosity
μ_0	yield viscosity
ν	fluid’s kinematic viscosity
ν_m	molecular collision frequency
Ω	continuum phase domain
Ω_p	particle’s angular velocity
Ω_p	characteristic value of Ω_p
Ω_s	characteristic scale of the vorticity in pure-shear flow
ψ_k	moment exchange of component k with the rest of components
ρ	relative density of the disperse component
ρ_f	fluid’s density
σ	Cauchy’s stress tensor
σ_m	tangential accommodation coefficient

$\boldsymbol{\tau}$	deviator stress tensor
τ_{c_s}	sonic time scale
τ_D	molecular diffusion time scale
τ_ν	viscous diffusion time scale
$\tau_{p,\Omega}$	rotational particle relaxation time
τ_p	particle response time
τ_w	shear stress at the wall
Θ	dimensionless frequency normalized by ν_m
θ	absolute temperature
ξ	inverse of the particle mobility
ξ_U	activation coefficient for $\mathbf{F}_U + \mathbf{F}_A$ term

Other Symbols

$\langle \cdot \rangle_k$	filtering operator over the component k
$\frac{k}{\cdot}$	filtering operator per unit component volume over the component k
D/Dt	material derivative of the fluid continuum
Δ	Laplacian operator
∂A	boundary of set A
(\cdot, \cdot)	inner product
$\langle \cdot, \cdot \rangle$	integral of the product of two functions
$\overset{k}{\cdot}$	k -component variable over k -volume fraction

Acronyms / Abbreviations

ALE	arbitrary Lagrangian-Eulerian method
ALU	arithmetic logic unit

ASGS	algebraic sub-grid scales
BEM	boundary element method
CFD	computational fluid dynamics
CFL	Courant–Friedrichs–Lewy number
COR	coefficient of normal restitution
DEM	discrete element method
DNS	direct numerical simulation
DNS	direct numerical simulation
DOF	degree of freedom
EE	Euler–Euler (method)
EL	Euler–Lagrange (method)
FDE	fractional differential equations
FEM	finite element method
FFC	recovery method of Pouliot et al. (2012)
FFT	fast Fourier transform
FLOP	floating point operations
FPU	floating point unit
FSI	fluid-structure interaction
FVM	finite volume method
LBM	lattice Boltzmann method
LES	large eddy simulation
MAE	method of approximation by exponentials
MFP	mean free path
MPM	material point method

MRE	Maxey–Riley equation
MSD	mean square displacement
OSS	orthogonal subscales
PFEM	particle finite element method
PIC	particle-in-cell
PID	particle impact drilling
PPC	particles per cell
PPR	recovery method of Zhang and Naga (2005)
PSM	pseudo-spectral method
RDF	radial distribution function
RVE	representative elemental volume
SC	standard case
SM	streaming multiprocessors
VMS	Variational multiscale (method)

Chapter 1

Introduction

1.1 Background and Motivation

How is a moving particle affected by its surrounding fluid? This question has intrigued scientists since the Scientific Revolution and earlier, going back to Galileo [325], Leonardo Da Vinci [144], and even Aristotle [293]; perhaps because it refers to a simple system, ideal for thought experiments. The problem attracted attention for pragmatical reasons early on as well, as it is central to the study of ballistics (the science of predicting the trajectory of projectiles), which began to be systematized as a discipline by Tartaglia in the sixteenth century. Despite his ignorance of Newtonian dynamics, Tartaglia was able to provide useful range tables and instructions, although his understanding of the problem was still very rudimentary. For instance, he considered the trajectory of projectiles to be formed by the succession straight-circumference arch-straight, influenced by earlier tradition [335]. About a century later, Galileo provided the parabolic solution for a projectile in void; noting, in 1638, that he expected his solution to be valid only for slow-moving projectiles, while faster moving trajectories should make ‘the beginning of the parabola less tilted and curved than its end’. Thus, for a very long time, the intuitive notion that a resistance existed proved too difficult to translate into scientific knowledge.

In 1671 Newton wrote the following text in passing, as an analogy to describe the motion of his light corpuscles, when explaining the breakdown of a white ray into its different colors [259]¹:

¹Specifically, Newton is addressing the origins of the lateral force induced by simultaneous rotation and translation of a spherical particle, causing it to drift, curving its trajectory sideways. This phenomenon would later be rediscovered and empirically studied by Robins in the eighteenth century (although the phenomenon is commonly known as the ‘Magnus effect’, after Magnus, who studied it much later and who, in fact, cites Robins in his work [229]) [219, 325].

... I remembered that I had often seen a Tennis-ball struck with an oblique Racket describe such a curve line. for a circular as well as a progressive motion being communicated to it by that stroak, its parts on that side where the motions conspire must presse & beat the contiguous air more violently then on the other, & there excite a reluctancy & reaction of the air proportionally greater.

The quotation reflects Newton's very early insight into the fundamental mechanisms of fluid-particle interaction, made possible only thanks to the idea of force, the notion of continued action (as opposed to discrete) and its necessity to have the ball deviate from a straight line (law of inertia), and the action-reaction postulate; all of which were recent discoveries at the time. Moreover, it exemplifies the *reductionist approach*, the process of continued 'zooming in' that would ultimately give the scientific method its modern power and that particularly resonates with the common theme of numerical methods: 1) break down into simpler entities, 2) model the interactions, 3) aggregate the result of all these interactions.

By running experiments on pendulums with spherical bobs, Newton himself was able to establish that the air resistance was approximately proportional to the square of the velocity in his Principia [260], a result still standing today (for fast-moving particles, in the so-called *Newton-drag regime*; see e.g. , [218]). Together with his laws of motion, this finding was to set the science of ballistics on a much more solid ground².

Still, little progress could be made from the theoretical point of view for many years, due to limited understanding of fluid dynamics. Famously, in 1752 D'Alembert proved the paradoxical result that inviscid flows, for which Euler had correctly given their equations of motion, yielded zero resistance on submerged objects. Almost a century later, Saint-Venant (who had actually derived the Navier–Stokes equations before Stokes) took the first steps toward its resolution by suggesting that it was necessary to take the fluid viscosity into account, as it would surely introduce some tangential resistance [7].

But it was not until the work of Stokes [326] that the first (correct) quantitative solution to the problem could be derived from first principles, for the case of very small Reynolds number³. Stokes calculated the drag force on a small sphere moving at steady speed in an

²Although Newton was unable to improve upon Galileo's parabolic solution by introducing the quadratic air resistance into the equations. The first accurate ballistics tables were produced around mid-eighteenth century by Robin, who obtained them empirically, and by Euler who used Robin's data to fix the required empirical constants and a numerical method to integrate the nonlinear equation describing the trajectories based on Newton's squared-velocity drag law [325].

³This still did not resolve d'Alembert's paradox, since the fact that the Navier-Stokes equations coincide with Euler's equations in the limit of very large Reynolds number (large velocities over viscosity ratio), which yielded zero drag, seemed inconsistent with the known empirical observation that the drag did not actually tend to zero with viscosity at high velocities. A resolution of this problem did not come until 1904 with the work Prandtl, see [6], who introduced the concept of the boundary layer that provided a sound transition between both conflicting models.

infinite Newtonian fluid using a simplified version of his own discovery, the Navier–Stokes equations. The force turned out to depend linearly on the velocity in this regime, in contrast to the theory of Newton, whose pendulums operated at higher Reynolds numbers. Insightfully, Stokes observed that

Since in the case of minute globules falling with their terminal velocity the part of the resistance depending upon the square of the velocity, as determined by the common theory, is quite insignificant compared with the part which depends on the internal friction of the air, it follows that were the pressure equal in all directions in air in the state of motion, the quantity of water which would remain suspended in the state of cloud would be enormously diminished. The pendulum thus, in addition to its other uses, affords us some interesting information relating to the department of meteorology.

Where the reference to a pendulum is due to Stokes being in fact primarily interested in solving that problem, from which he derived his famous drag relation as the limiting case of an infinite oscillation period. As it turns out, the problem of cloud formation is still a subject of current studies, most often based on his equation [70, 150].

Indeed, Stokes' result became (and remains) very important to fundamental as well as applied science, having lead to at least three Nobel prizes, according to Dusenbery [111]. For instance, Einstein used it to calculate the Brownian diffusion rate (see [290]), a result that provided the first strong evidence of the existence of atoms (i.e., strictly speaking, water molecules) as well as a practical means to calculate their mass!

Most analytical successes in studying this problem have been achieved in the realm of the tiny to the very small; in a world where viscosity counteracts all external forces and where inertia is, to all effects, negligible (or at most a small correction). Neglecting inertia renders the Navier-Stokes equations linear, which removes the tremendous difficulties associated with nonlinear effects, including the analytical intractability of problems and the onset of turbulent chaos. For large particles, when the Reynolds number grows much past one, we have not been able to go too far past the empirical solution of Newton.

But even under low Reynolds number conditions, the analytical difficulties do appear and it took many years until a complete picture was available. Successive advances were made over several decades by many prominent scientists, such as Boussinesq (1885), who derived the unsteady terms that arise when the particle decelerates in otherwise the same conditions as Stokes; Oseen (1910, 1913), who correctly found the first effects of inertia on the steady drag (although not entirely satisfactory, see Section 2.2.1); Faxén, who accounted for the non-uniformity of the background flow and Saffman (1965), who calculated the lift

force in the presence of a linear shear flow (see the historical account by Michaelides [248]). The equation of motion that we will study in the following two chapters, the Maxey–Riley equation, that generalizes the works of Boussinesq and Faxén was not derived until the year 1983.

With this brief historical account we hope to motivate our interest in studying the motion of a single particle in a fluid. One of the lessons learned during the course of these years of research is to what extent this is a fundamental theoretical piece towards understanding particle-laden flows. Although initially unplanned, such endeavour has ended up taking a large part of our work (Chapters 2 and 3).

1.1.1 Numerical simulations

With the growing computational power available and the development of multiphase solvers, one may be tempted to drop the analytical route altogether and simply calculate the forces on individual particles by solving the Navier–Stokes equations around them. A particle-laden flow system is, after all, a set of solid bodies submerged in a fluid, and could in principle be simulated by using their surface to impose the corresponding boundary conditions on the fluid as in any fluid-structure interaction problems.

In fact, researchers are increasingly following this route, producing *fully-resolved* simulations to study systems of submerged particles (the number of particles in each study is indicated between parentheses): In Johnson and Tezduyar [186] the settling of a sphere in a container was studied with a body-fitted mesh (1000); in [273] the fluidization of a bed of spheres was studied numerically and experimentally (1204); in [138] the lattice-Boltzmann method was used to simulate spheres settling in a stationary fluid (6400); while in [349] an embedded-body approach was used to study particle clustering in turbulent flows ($\approx 20\,000$). These methods are particularly helpful in basic research, as they can be used to calibrate empirical models [384], to study qualitative aspects of the flow, including turbulent effects [344] and to study complex systems including non-Newtonian flow models and particles interaction in exquisite detail [107, 250].

But the numbers in the previous paragraph are still clearly way too small for the vast majority of particle-laden flows of interest: sand grains or pebbles in a river bed, soil slurries, avalanches and pyroclastic flows; sedimentation tanks and vessels used to clarify water or separate small particles by size; pneumatic conveying systems carrying sugar, flour, coal seeds, nuts and conglomerate pellets and in general solid matter in gas or liquid ducts; fluidized beds, used in the chemical industries to enhance chemical reactions and homogenize and dry particulates; bubbles in nuclear reactors, oil, rock and gas bubbles in oil wells; microscopic biological systems like blood cells and platelets in small vesicles or suspended particles in

the bronchioles, such as contaminants or vaporised sprays; plankton and organic matter in the ocean [231]; drops in a forming cloud, snow crystals and solid contaminant, like rubber particles, in the atmosphere; and even rocky particles in low-density atmospheres, relevant to the study of planet formation [358].

Thus, one soon realizes that this approach remains hopeless as a tool to study most systems as a whole; and it is expected to remain so for a long time, unless some unexpected, revolutionary discovery leads to a jump of several orders of magnitude in computational power.

On the other hand, note that if we are able to give up the resolution of the flow *around* the particles and instead *modelling* the hydrodynamic interactions using the coarse scale description of the flow we are able to immediately generate a radical cut-down in the computational cost.

A quick calculation illustrates this point very convincingly: First, take a suspension of N particles that we will consider to be of uniform size for simplicity. Let us assume that discretization size required to run a fully resolved simulation (where all the scales present in the fluid motion are considered) to be one tenth of the diameter of the particles. Note that it is unreasonable to think that a substantially coarser size should be sufficient to capture the dynamics of the flow around the particles in sufficient detail.

On the other hand, consider the same problem solved by an alternative, coarse-grained approach, where the hydrodynamic interactions are *modelled* as a distributed effect on the fluid, and where only the details of the flow much larger than the particles (say, ten times as big) are numerically resolved. Note that with such approach it would hardly make sense to consider finer resolutions at all because, at scales comparable to the particle size, the real flow would be locally distorted by the presence of the individual particles that are being averaged over. But without the information about the individual distortions the solution would necessarily be very poorly modelled at such fine scales. In other words, it would just not pay off to consider that level of detail.

Comparing both situations, one has that the number of computational cells required scales as $(10N)^3$ in the first case and $(0.1N)^3$ in the second case. That is a factor 1×10^6 in number of cells between the two approaches, which amounts to a significantly greater number in terms of computational resources⁴.

⁴In computational fluid dynamics, the computational cost is mostly determined by the resolution of the system of equations resulting from the spatial discretization, whose size scales as the number of unknowns. The computational cost of the resolution of the system can at the very best be linear for the latest-generation iterative solvers [146] (grow proportionally to the number of unknowns), although such scaling in practice is difficult to achieve. Furthermore, the smallest time step also needs to be modified, typically decreased proportionally to the element size, in order to preserve the accuracy increase desired; see the discussion around the Courant–Friedrichs–Lewy condition in Section 4.7.7

Note that such strategy corresponds to a mixed-scale method, where one phase (the particles) is described to a finer level of detail (individual particles) than the other (the fluid), which is described to a coarser level. This work is concerned with precisely this type of mixed methods.

Clearly, one can continue the process of coarsening of the description level further by homogenizing the particles phase as well, giving rise to a multicomponent continuum description of the particulate system. The method most consistent with this philosophy is the so-called Euler–Euler (EE) method [180], also called the Eulerian-Eulerian or two-fluid model; as opposed to the Euler–Lagrange (EL) method, which refers to the mixed method discussed above. In the EE method the motion of the particles is determined by a set of conservation differential equations, analogous to those of the fluid phase. Thus, both the particles and the fluid phases have their associated systems of conservation equations. The equations that correspond to each phase are coupled to each other through the averaged volume fractions (volumetric proportions), whose unknown values vary pointwise, as well as by momentum exchange terms (see Appendix H).

The equations involved can be derived following a purely formal procedure that relates the different averaged (large-scale) variables of interest. Several averaging methods exist, including time, volume and ensemble averages, but they all lead to similar equations [109]. As a consequence, and in a completely analogous situation to that encountered in turbulence modelling, the resulting equations include a number of terms that depend on finer-scale details, and that must be closed (expressed as a function of the averaged variables) to obtain a well-posed problem. This closure is normally based on the introduction of new physical models that most often include unknown parameters to be calibrated from experiments [182]. These methods have been applied to the description of particle-laden flows for many years; especially in the chemical [332, 139, 192] and the nuclear [196, 158] industries.

In fact, EE-type methods remain the only realistic approach in many industrial settings. The reason is found again in the large numbers involved, which penalize the explicit description of each and every particle in the domain. This is a limitation common to all particle-based methods, including the discrete element method (DEM), since most industrial problems involve a much larger number of particles than what is computationally feasible.

The latter point is illustrated in Fig. 1.1, where a bunch of industrial applications are identified in a domain size-number of particles graph. As a reference, simulations involving several million particles (green curve) are representative of what can typically be achieved in a personal computer and a few billion (red curve) particles that corresponding to what can ordinarily be achieved in powerful computer clusters, using massive parallelization techniques. The violet 2×10^{12} -curve is representative of the numbers of particles included

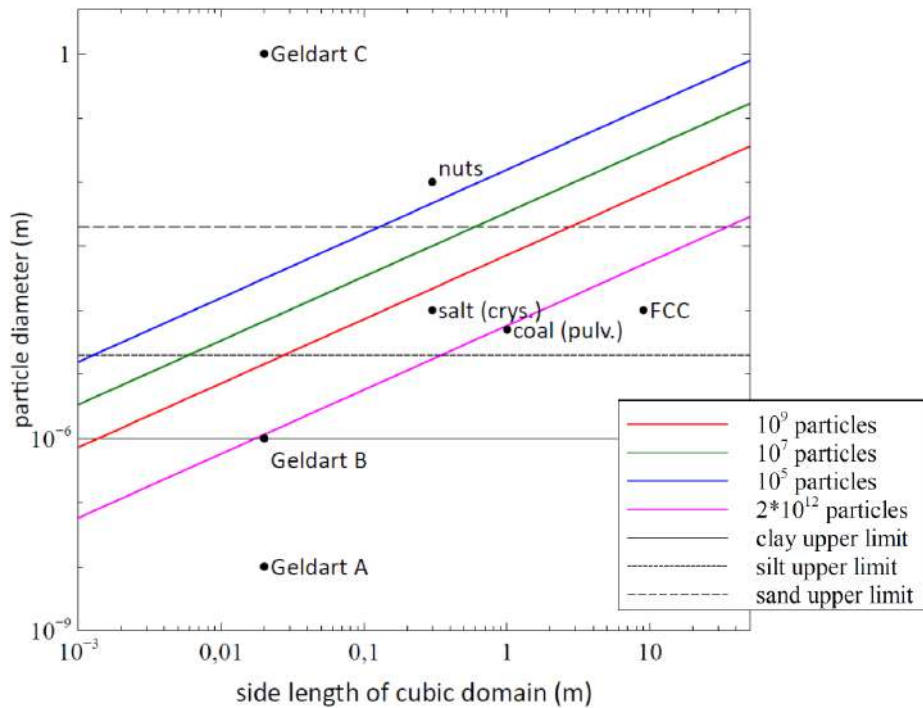


Fig. 1.1 Number of particles associated to cubic domains for a fixed proportion of volume occupied by the particles (0.1) as a function of the equivalent side length of the domain. The curves corresponding to different numbers of particles are included, as well as the upper limit sizes for clay, silt and sand, as representative granular materials. The scattered points correspond to different industrial application examples: coal, salt and nuts pneumatic conveying (volume per linear meter for typical pipe diameters); fluidized bed experiments of Geldart [143], showing the different Geldart categories; and catalytic cracking (FCC) for the production of gasoline.

in today's largest cosmological simulations [281] and can be interpreted as an extremely optimistic upper bound. By mentally locating one's preferred application in the plot of Fig. 1.1 one quickly arrives at an intuitive grasp of the limitations of particle-based methods as a numerical tool to simulate real systems with realistic numbers of particles.

Nonetheless, it must be mentioned at this point that there exists the possibility of using a coarse-grained description, where a smaller-than-real number of (larger) particles is used instead of the one-to-one description [125, 304, 271] in analogy to what is common practice in molecular dynamics [295, 176]. And, while the level of maturity of this approach is still low, with several problems still unresolved [200, 258], the technology seems to be yielding promising results with spectacular computational gains [225].

Furthermore, even if we ignore the possibility of coarse-graining, hybrid methods and, in particular, CFD-DEM methods retain an interest due to their many advantages over exclusively continuum-based methods:

- One avoids the complications associated with a continuum-based description of granular matter, with still several open theoretical questions [329, 297, 22]; despite recent promising advances, mainly in the area of dry granular matter [163, 110].
- The non-uniform granulometry of real materials can be reproduced naturally, while the same is extremely difficult to include in a model. Often, the only practical alternative is to simplify the granulometry into a finite number of fixed sizes with different proportions and include as many phases as different sizes are present, with the associated multiplicity of laws and computational resources required. This point is actually included in the point above but deserves special mention.
- A lower degree of empiricism is expected, thanks to the lower-scale description of the solid phase.
- The dynamics of the particles are not averaged-over, so that the actual mechanisms that dictate the behaviour of the particles can be understood directly, providing a more valuable qualitative picture that does not need to be reconstructed (which would be nontrivial or impossible) as a post-process. This advantage is particularly important close to the domain boundaries, where averaged methods tend to do most poorly, due to the sharp transitions.
- The consideration of additional effects such as new types of interaction laws (e.g. electrostatic interaction) is natural and does not require the overall modification of the models in place, but usually only a straightforward additive term.
- One can take advantage of existing programs (e.g. DEM programs) and modify them in a simple way to adapt them to the interaction effects with the surrounding fluid.

So, although the number of real-world applications where hybrid methods can be applied is in practice limited, it is important to recognize their value as validation tools that operate at an intermediate scale between fully resolved models and continuum-only models. Therefore, in addition to their value as direct simulation tools for the industry, they are useful to feed the empirical data needed to close EE-type equations and to understand the full-size objects of study through the analysis of simplified, small-size systems [132].

1.1.2 Industrial Doctorate

This dissertation has been completed with financial support from the *Doctorat Industrial* grant from the *Generalitat de Catalunya*, that supports research work performed as part of the work of a company in collaboration with a research center and with a focus on applied research. In our case, the research center role was played by the *International Center for Numerical Methods in Engineering* (CIMNE) and the role of the company was played by *Computational and Information Technologies S.A.* (CITECHSA).

The main objective of CITECHSA was to start the development of a computational fluid dynamics (CFD) tool able to simulate a wide range of particle-laden flows. Our realization at the time of beginning the work that a new DEM application tool was at that point at an early stage of development by the DEM group of Kratos (see next section) motivated our choice of the DEM for the simulation of the particles' phase. This choice also created a strong dependence between the DEM developments and those concerning the coupling, that explained the author's enrolment into the DEM group in CIMNE at around that time.

As a consequence, DEM-related tasks have occupied a substantial part of the research time. However most of the resulting algorithmic developments do not go beyond a state of the art, which is sufficiently documented. Furthermore, the application examples and analyses of interest are too detached from the main thread of this work to include them. Thus only a very brief account of some DEM-related elements are described in Appendix A. Nonetheless the reader should keep in mind the DEM background of the author, and the (perhaps) clear emphasis given to the particles' side of the research.

1.2 Objectives and methodology

The objectives of the research reported in this document are the following:

1. To develop an algorithm that combines the discrete element method and the finite element method to simulate particle-laden flows with the following list of requisites:
 - capability of dealing with a wide range of regimes, including the possibility to have regions with dense and dispersed suspensions simultaneously
 - use of the finite element method to discretize the fluid
 - use of the discrete element method to model the particles
2. To study the range of applicability of the Maxey–Riley equation as a model for the motion of the individual particles submerged in a fluid, improving the current knowledge

on the subject and generating, where possible, practical estimates of direct application to numerical modelling.

3. To study current alternatives for the numerical treatment of the history term in the equation of motion and compare them.
4. To improve on the method of quadrature (the history term involves a time integral, as will be explained later) proposed by van Hinsberg et al. [353] and provide a detailed study of its efficiency and accuracy, providing convincing evidence that it is not necessary to neglect this term to have an efficient numerical method.
5. To report an account of relevant application examples of the proposed strategy with interest to the industry, as well as of the different technologies developed for their particular requirements.
6. To generalize a stabilized finite element method and use it to discretize the backward-coupled flow equations.
7. To develop a suitable inter-phase coupling strategy.

Our research activities have consisted in a combination of bibliographical investigation, numerical analysis (designing and running numerical experiments using existing or new code, mainly on a personal computer and sometimes on a small cluster) and programming.

The bulk of the programming work has been carried out within the framework Kratos Multi-physics [88], or Kratos for short. Kratos is based on C++, with a Python-based external layer. It is organized according to a marked object-oriented philosophy that, at its coarsest level, can be modelled as a central core (Kratos Core) connected to a list of independent applications (FluidDynamicsApplication, StructuralApplication, DEMApplication, etc.).

The core contains the definition of the fundamental abstract classes common to a large part of the implementations. It provides a common language and a set generic protocols and algorithmic tools like linear solvers, search algorithms or input/output utilities.

On the other hand, the developments related to specific problems, such as solving the Navier–Stokes equations or performing structural analyses, are coded in the corresponding specialized applications (in this case perhaps FluidDynamicsApplication and StructuralApplication). Moreover, an application can relate any subset of other, existing applications to solve a more complex problem. In order to fulfil our particular goals, we have

created `SwimmingDEMApplication`, that couples `FluidDynamicApplication` [82, 311]⁵ to `DEMApplication` [64, 177].

`FluidDynamicsApplication` is an Eulerian-description, finite element method-based CFD simulator and it is in charge of solving the fluid equations of motion. `DEMApplication` is a general-purpose discrete element method simulation suite that is responsible of tracking the particles, computing contacts between them and evolving their motion in time.

`SwimmingDEMApplication` contains all the inter-phase coupling tools, the hydrodynamic calculation algorithms, the modified fluid finite elements necessary to solve the modified backward-coupled flow equations (see Chapter 4) and several utilities that were developed where no similar tool was found in the core, but that have not yet reached a stage of maturity to make them suitable for their implementation in the Kratos core (for instance, the derivative recovery tools, see Section 4.4).

The whole of Kratos, including all the applications mentioned above and many others, is freely available for download⁶. `FluidDynamicApplication` has an interface usable from the Pre-/Post-processor GiD [244], that, by default, is included in Kratos. Moreover, `DEMApplication` powers any of the four free packages grouped under the name `DEMPack` [65], also based on the GiD interface and also included in Kratos. Among these, `FDEMPack` gives access to most of the capabilities of `SwimmingDEMApplication`, developed during the course of the research reported in this work.

1.3 Outline of this document

The core of this work is contained in Chapters 2 to 4. Of these, Chapter 2 is the most theoretical in character. It is devoted to the analysis of the Maxey–Riley equation, as a fundamental model for the description of the motion of individual particles in a fluid, which at this point is assumed to be described by a known field. Section 2.2 systematically explores the range of validity of the model with respect to several criteria, first in terms of the nondimensional values that appear in the equation itself and later in terms of additional variables involving a selection of simplifications introduced a priori in the development of the theory. In Section 2.3 we apply a scaling analysis to the different terms in the model, providing estimates of their relative magnitude that may be applied in practice to simplify the basic model by neglecting the less important terms. Section 2.4 contains a summary of the

⁵Our application has been designed so as to facilitate changing the fluid-solving application. Incidentally, it has already been coupled to `PFEMApplication`, a Lagrangian description fluid solver based on the particle finite element method (PFEM) [268, 66].

⁶The code can be retrieved from <https://github.com/KratosMultiphysics/Kratos> (retrieved on June 15, 2018).

main results of the chapter, including Tables 2.2, 2.4 and 2.5, that list the most important numerical estimates.

Chapter 3 shifts from the analytic point of view of Chapter 2 to the study of different numerical techniques to solve the Maxey–Riley equation, providing a detailed description of the algorithms involved. We emphasize the treatment of the history force term, whose numerical solution remains challenging, and that is often ignored for this reason. In Section 3.2 we provide a state of the art concerning the numerical treatment of this term, comparing the accuracies of several methods of quadrature. Section 3.3 contains the description of our method of choice for the quadrature, proposed by van Hinsberg et al. [353], and of the extensions and improvements we have developed based on it. Section 3.4 contains the description of the overall algorithm for the time integration of the equation of motion including the history term. In Section 3.5 we make a stop to connect the general theory of fractional calculus to the Maxey–Riley equation, which had only been very superficially sketched before. This section appears at this point to take advantage of the terminology and concepts introduced in the previous developments. Section 3.6 is devoted to a systematic analysis of the accuracy and efficiency of the numerical method of solution, applying it to a sequence of benchmarks of increasing complexity. We close the chapter with a summary of the most important results and developments.

Chapter 4 moves even further toward practice, being eminently applied in nature. We consider a series of applications representative to different families of industrial problems, describing several developments involved in their solution as we go along. Section 4.2 discusses how the equation of motion considered in the previous chapters can be modified using empirical relations to extend its range of applicability beyond the limits studied in Chapter 2, for this will be necessary in the industrial applications that are considered later on. In Section 4.3 we introduce the fluid phase as a problem to be solved for the first time, for which we make use of a well-established stabilized finite element formulation. We describe it in sufficient detail to clarify the terminology and set the stage for the generalizations introduced in Section 4.3. In Section 4.4 we discuss the problem of derivative recovery, as a step necessary to obtain accurate estimates of the derivatives of the fluid field, once a solution is produced by the fluid solver. We give a brief state-of-the-art and compare several alternatives compatible to the finite element method. The description of the particles-fluid coupling is described very briefly in Section 4.5 where we basically bring together the different algorithmic parts involved. We quickly turn to the first application example, which is the subject of Section 4.6. It consists on the numerical simulation of the phenomenon of air bubble trapping in T-shaped pipe junctions, which was only recently first studied by Vigolo et al. [356]. This test exemplifies the use of the one-way coupled strategy, with

no inter-particle interactions, representative of low-Reynolds number internal flows with low-density suspensions of small particles. We continue on to the study of the next test application in Section 4.7, where we study particle impact drilling, a technology used in the oil and gas industries to produce high penetration rate drilling systems. Here we again make use of a one-way coupled strategy for the fluid-particles coupling, although this time we consider the inter-particle contact as well. This is by far the most detailed of the examples provided and corresponds to consultancy work made during the ‘Doctorat Industrial’. Before moving on to the final example, the theory related to the backward-coupled flow must be introduced first, based on the theory of multicomponent flows (its rudiments can be found in Appendix H) and the general finite element formulation introduced in Section 4.3, which we specialize for this problem. Following this theoretical interlude, we present the last application example, a fluidized bed of Geldard-D particles in Section 4.9. We close with a summary of the chapter in Section 4.10.

We end the work with Chapter 5, where we give a schematic summary of our developments. In Section 5.2 we discuss some of our current and future work. There is indeed a considerable number of research lines that have been opened with this work, but that we have not followed as far as we would have liked due to the breadth of our scope.

We provide a series of appendices that carry substantial content. Our intent in doing so has been more to lighten the main text than to expand it with additional content. Perhaps the idea of *modularity*, associated to the programming principle, has contaminated our writing style, but we hope it helps comprehension. For instance Appendix A contains most of the little text devoted to the discrete element method while Appendix H is devoted to the description of the continuum theory related to the backward-fluid flow. The other appendices include formulas and numerical data of interest mainly to the interested reader willing to program the associated algorithms.

Finally, the Nomenclature section provides a quick guide for the symbols used throughout. Only the symbols that are exclusively used in a single context, close to their definitions, have been left out of the list, to avoid overpopulating it.

Chapter 2

The Maxey–Riley equation

2.1 Introduction

The Maxey–Riley Equation (MRE) [238] describes the motion of a small, rigid, spherical particle immersed in an incompressible, Newtonian fluid. It is an expression of Newton's Second Law, relating the acceleration of the particle, taken as a point-mass, to the sum of the external forces (usually just its weight) plus the total force applied on it by its containing fluid. The smallness of the particle justifies the two fundamental assumptions made in its derivation: 1) that the relative flow can be described by the Stokes equations near the particle; and 2) that the flow is well represented by its second-order Taylor expansion about the particle's center (near the particle). The equation is thus applicable to the study of disperse particulate flows, for systems where the low volume fractions justify the consideration of each particle as an isolated body. Examples of such systems can be found in the study of warm-cloud rain initiation [122, 314, 363], turbulent dispersion of microscopic organisms [275, 365], the dynamics of suspended sediments [324, 155] and also in more fundamental areas, such as the study of turbulent dispersion [324, 27, 53].

The hypotheses involved in the derivation of the MRE are very often too stringent for the applications of interest. For such cases, there exist a huge variety of models applicable outside its strict range of applicability. These models, which usually involve a high degree of empiricism, are nonetheless very often constructed on the basis of the MRE (e.g. they are asymptotically convergent to it), showing similar mathematical properties and physical significance. Examples of such extensions apply to higher particle Reynolds numbers, nonspherical shape or finite-sized particles; see Section 4.2. The analysis and numerical treatment of the MRE is therefore of major importance from both the applied and the fundamental perspectives. This chapter is devoted to the study of several aspects related to

the MRE, more specifically to its range of applicability and scaling properties. Our goal is to provide a solid ground on which to base engineering decisions concerning its use.

Let us recall the exact formulation of the MRE. Let $\mathbf{u} : \mathcal{D} \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$ define the background fluid velocity field, where $\mathcal{D} \subseteq \mathbb{R}^n$ is some spatial open domain contained in the fluid and n is the ambient spatial dimension (either two or three). Let $\mathbf{r}(t) \in \mathcal{D}$ be the trajectory of a particle through this fluid, and \mathbf{v} its velocity. The MRE states:

$$\begin{aligned}
 m_p \frac{d\mathbf{v}}{dt} = & \underbrace{m_f \frac{D\mathbf{u}}{Dt}}_{\mathbf{F}_U} + \underbrace{\frac{1}{2}m_f \left(\frac{D}{Dt} \left(\mathbf{u} + \frac{1}{10}a^2\nabla^2\mathbf{u} \right) - \frac{d\mathbf{v}}{dt} \right)}_{\mathbf{F}_A} + \underbrace{6\pi a\mu \left(\mathbf{u} - \mathbf{v} + \frac{1}{6}a^2\nabla^2\mathbf{u} \right)}_{\mathbf{F}_D} \\
 & + \underbrace{6a^2\sqrt{\pi\rho_f\mu} \frac{d}{dt} \int_{t_0}^t \frac{1}{\sqrt{t-\tau}} \left(\mathbf{u} - \mathbf{v} + \frac{1}{6}a^2\nabla^2\mathbf{u} \right) d\tau}_{\mathbf{F}_H} + \underbrace{(m_f - m_p)\mathbf{g}}_{\mathbf{F}_W}
 \end{aligned} \tag{2.1}$$

where m_p is the mass of the particle, m_f is the mass of the displaced fluid volume, a is the particle radius, ρ_f is the density of the fluid, μ is the dynamic viscosity of the fluid ($\nu = \mu/\rho_f$ is the kinematic viscosity) and \mathbf{g} the acceleration due to gravity. The notation D/Dt refers to the material derivative of the fluid. Eq. (2.1), together with $\mathbf{v} = d\mathbf{v}/dt$ and the initial conditions $\mathbf{r}(t_0) = \mathbf{r}_0$ and $\mathbf{v}(t_0) = \mathbf{v}_0$ form an initial value problem that must be solved to obtain the trajectory of the particle.

The terms on the right hand side of Eq. (2.1) have distinct physical interpretations and can be identified as: the force applied to the volume displaced by the particle in the undisturbed flow (\mathbf{F}_U), the added mass or virtual mass term (\mathbf{F}_A), the Stokes drag term (\mathbf{F}_D), the Boussinesq–Basset history term (\mathbf{F}_H) and the term due to the weight of the particle minus its (hydrostatic) buoyancy (\mathbf{F}_W).

The terms bearing second order derivatives are the second order corrections, known as Faxén corrections, that take into account the finite size of the particle with respect to the second-order spatial variations of the flow field around it, see Section 2.2.2.

The exact formulation of the MRE as was originally given by Maxey [238] differs slightly from the version presented here, which adopts the Auton et al. [14] form of the added mass force. This formulation contains the material acceleration of the fluid continuum (i.e. the derivative following the fluid), evaluated at the current position of the particle. The original formulation presented instead the total derivative (i.e. the rate of change of the fluid velocity when following the particle). The former form, derived under the assumption of inviscid flow

in [14] is however valid to the same order as the other terms derived by Maxey and Riley within their assumptions, while remaining accurate at higher Reynolds numbers [361].

Another difference with respect to the original formulation involves the Boussinesq–Basset term, on which the time derivative appears *outside* the integral sign in Eq. (2.1), as opposed to inside, as in the original Maxey and Riley paper. As pointed out in [89] this is a more general formulation, which allows for a nonzero initial relative velocity (and is also equivalent to the generalization later given by Maxey [240], see [222]; see also Appendix C).

2.2 Range of validity

Let us rewrite the MRE in dimensionless form. This will serve to introduce some fundamental dimensionless numbers. For the sake of notational simplicity, the same symbols have been overloaded to designate the dependent and independent variables, although here they represent nondimensional quantities.

$$\begin{aligned} \frac{d\mathbf{v}}{dt} = & \frac{1}{\rho} \frac{D\mathbf{u}}{Dt} + \frac{1}{2\rho} \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} + \frac{9}{10} \frac{1}{2\rho+1} \frac{1}{StRe} \frac{d\nabla^2\mathbf{u}}{dt} \right) \\ & + \frac{2\rho+1}{2} St \left(\mathbf{u} - \mathbf{v} + \frac{3}{2} \frac{1}{2\rho+1} \frac{1}{StRe} \nabla^2\mathbf{u} \right) \\ & + \frac{3}{2\sqrt{\pi}} \sqrt{(2\rho+1)St} \int_{t_0}^t \frac{1}{\sqrt{t-\tau}} \left(\mathbf{u} - \mathbf{v} + \frac{3}{2} \frac{1}{2\rho+1} \frac{1}{StRe} \nabla^2\mathbf{u} \right) d\tau + (1-\rho) \frac{1}{Fr^2} \end{aligned} \quad (2.2)$$

where

$$\begin{aligned} \rho &= \frac{\rho_p}{\rho_f} \quad (\text{relative density}) \\ Re &= \frac{UL}{\nu} \quad (\text{Reynolds number of the ambient flow}) \\ St &= \frac{2\rho+1}{9} \frac{a^2}{\nu} \frac{1}{T} \quad (\text{Stokes number}) \\ Fr &= \frac{U}{\sqrt{gL}} \quad (\text{Froude number}) \end{aligned} \quad (2.3)$$

where $g := \|\mathbf{g}\|$ and U , L and T are the dimensional scalars by which \mathbf{u} (and \mathbf{v}), \mathbf{x} and t have been normalized. These quantities are conventionally defined such that L is a characteristic length scale of the flow and U is defined such that U/L is the characteristic magnitude of the gradient of the unperturbed velocity near the particle location, while T is defined such that $UT = L$. By the term ‘unperturbed’ we refer to the flow resulting from subtracting the

(Stokes) flow produced by the presence of the particle under the same far-field boundary conditions from the actual flow. The application of Buckingham’s Pi Theorem to this equation reveals that the set of dimensionless parameters that describe it is in fact minimal. By defining $\tau_p := TSt$, we obtain the relation

$$St = \frac{\tau_p}{T} \quad (2.4)$$

The quantity τ_p is commonly referred to as the *particle’s response time* [216]. It is equal to the time it takes for a static particle to accelerate to $1 - 1/e$ (where $e = 2.71 \dots$) parts of the surrounding fluid velocity, under the action of a constant, uniform flow (and neglecting F_H). The Stokes number can thus be interpreted as the ratio of the particle’s characteristic response time to the fluid’s characteristic time, and it measures the dynamical importance of the particle’s inertia. It is useful to additionally define

$$Re_p = \frac{aW}{\nu} \quad (2.5)$$

where W is the characteristic magnitude of the relative (or slip) velocity $\|\mathbf{v} - \mathbf{u}\|$. The quantity Re_p is known as the *particle’s Reynolds number* and it characterizes the importance of inertial effects versus viscous effects in the flow produced by the relative motion between the particle and the background flow.

Eq. (2.1) is derived from the full Navier–Stokes problem as an asymptotic relation, valid as [238]

$$\frac{a^{\frac{a}{L}}U}{\nu}, Re_p, \frac{a}{L} \rightarrow 0 \quad (2.6)$$

In practice, these asymptotic relations are interpreted as requiring the left-hand sides to be *much smaller* than one ¹, that is

$$\frac{a^{\frac{a}{L}}U}{\nu} \ll 1 \quad (2.7)$$

$$Re_p \ll 1 \quad (2.8)$$

$$\frac{a}{L} \ll 1 \quad (2.9)$$

Note that the LHS of Eq. (2.7) can be understood as a sort of downscaled version of the ambient Re to the size of the particle, since the velocity variation seen by the particle scales as a/LU . This condition, along with Eq. (2.8) can be used to simplify the Navier–Stokes equations to the Stokes equation in the derivation of the MRE by resorting to scaling analysis

¹The imprecision in this requirement must be supplemented by experience or by some conventional rule adapted to the particular fields of application. A common criterion is to interpret $x \ll 1$ as meaning ‘at least two orders of magnitude smaller than one or $x < 0.01$ ’ [197].

of the full Navier–Stokes equations [238]. The approximation affects all the forces except F_U , which is treated independently and only relies on Eq. (2.9) (and a sufficient degree of smoothness of the flow).

The condition Eq. (2.9) is necessary to ensure that the disturbance flow caused by the presence of the sphere is well approximated at the surface by its second order Taylor expansion around its center. This approximation remains good as long as the mean flow around the particle is approximately quadratic, and the error is expected to drop very fast with decreasing a/L (as its third power). Note that, in practice, an accurate representation of the fluid velocity field will require a sufficient resolution of whichever numerical method is employed in its computation. The calculation of the spatial derivatives requires special attention. This issue will be addressed in Chapter 4.

Our main goal in this section is to gain a more complete understanding of the limits of applicability of the MRE than that directly offered by Eqs. (2.7)–(2.9). The original motivation for such an endeavour was to provide a bounded parametric space for the subsequent discussions about the relative importance of the different terms in the MRE, so as to make the analysis simpler and more precise. But we have come to recognize that such analysis should also be useful in itself, as a reference for the engineer or investigator concerned about the suitability of the equation for their particular problems. We are thus interested in answering questions such as

- What kind of effects are likely to appear first when moving away from its range of applicability?
- At what approximate pace are these effects expected to grow?

The survey we present is inevitably incomplete, though this kind of analysis is best realized through an iterative, cumulative process to which we hope to contribute a part. There have certainly been some notable efforts in this direction, the most important of which is the systematic analysis that E. Loth has undertaken over the last fifteen years [216, 219, 217, 218, 221]. However, his interest was mainly focussed outside the strict range of applicability of the MRE, exploring the ways in which it should be extended rather than its limits of applicability. Other relevant works are the best practice guideline [334] and the book [85].

Let us start by deriving some bounds to the dimensionless parameters in Eq. (2.3). First, the relative density ρ can be extremely large for gas flows, so that in practice it can be considered as any positive number. For liquid flows, however, it becomes bounded of order 10 for any commonly encountered materials (for a suspension of osmium particles in gasoline one has $\rho \approx 30$, most likely an upper value for ordinary applications). Note that for atmospheric air instead of water this value would be about a thousand times greater.

At the other extreme, the relative density of air bubbles in water can be practically taken to be zero with a minimal change in the value of St (typically, small bubbles and low Reynolds numbers lead to spherical bubbles; see [47] for an in-depth discussion on the circumstances under which bubbles can be modelled as rigid and spherical).

Moving on to the Stokes number, we find that it can also be bounded above by relating it to the particle Reynolds number as

$$\begin{aligned} St &= \frac{2\rho + 1}{9} \frac{a^2}{\nu} \frac{1}{T} \sim \frac{2\rho + 1}{9} \frac{a^2}{\nu} \frac{U}{L} < \max \left\{ 1, \frac{2\rho + 1}{90} \frac{a}{\nu} U \right\} \lesssim \\ &\lesssim \max \left\{ 1, \frac{2\rho + 1}{90} \frac{a}{\nu} W \right\} = \max \left\{ 1, \frac{2\rho + 1}{90} Re_p \right\} \end{aligned} \quad (2.10)$$

where we have interpreted Eq. (2.9) as $a < 0.1L$ and used the fact that for very large St the relative velocities will be of the order or larger than the absolute velocities of the smallest scales of the flow. We will come back to this issue in Section 2.3. Using Eq. (2.10) and the above argument for liquids allows us to bound St by the particle Reynolds number and in many cases (e.g. water suspended mineral particles) by an additional order of magnitude, so that if $Re_p < 0.1$, then $St < 0.01$, speaking in order-of-magnitude terms. In atmospheric air, a similar argument would lead to bound St below two to three orders of magnitude above Re_p . Using the more conservative $a < 0.01L$ would lead to lower upper bounds. We summarize the situation in Table 2.1, that includes some estimates for the upper bounds corresponding to different representative material combinations.

Table 2.1 Order of magnitude of the upper bounds to the admissible values (i.e. $Re_p < 0.1$) for St , given by Eq. (2.10) to half an order accuracy. The cases air-air and water-water are relevant to the study of neutrally buoyant particles; for example, marine plankton in sea water and some types of ice crystals [171]. Normal conditions have been assumed.

Continuous phase	Dispersed phase			
	air	water	sand	copper
air	1	10	50	100
water	1	1	1	5

Apart from Eqs. (2.7)–(2.9) and Eq. (2.10), there are additional assumptions involved in the derivation of Eq. (2.1) that we want to visit. To begin with, the problem is posed for a single sphere in an infinite domain. That is, the presence of nearby particles and of solid boundaries is not contemplated, though these elements are most often present in the

applications. Mere qualitative restrictions are of no practical use to the engineer, other than as a reminder of them being a source of error. Thus, there is a strong need to provide predictive quantitative measures too. Unfortunately, to the best of our knowledge, these are still open issues, and not much more than a few rough rules of thumb can be found in the literature.

It is also of interest to assess the range of validity of the different terms in the MRE, one at a time. The reason is that each term describes a distinct physical effect, with its own characteristic response to variations in the dimensionless parameters that characterize the flow. Indeed, in a given situation it might be of outermost importance to correctly calculate the steady drag force and not, say, the added mass force. Take for example the study of the deposition rate of microscopic particles in a container. Due to their tiny inertia the particles reach their terminal velocity very quickly. After the short transient the added mass force cancels, having a negligible impact on the much larger deposition time. In such cases, it is common practice to simply neglect the added mass force altogether, making to all effects the range of applicability of its particular formulation irrelevant. This sort of argument has allowed to justify the application of the MRE to a remarkably wide range of situations, that has been stretched even further by tweaking specific terms. A paradigmatic example of this practice is the use of empirical drag force laws combined with the usual formulation for the other terms (and usually neglecting the history force). We will revisit this type of approaches in Chapter 4, where we will need to move outside the range of applicability of the MRE.

It is important to realize that the use of this kind of empirical extensions almost invariably relies on the validity of the same additive structure of the different forces present in the MRE. Encouragingly, there has been an important body of research supporting such additive division both from a theoretical [223, 14] and an empirical [222] points of view in a variety of situations and well beyond the range of applicability of the MRE.

All these questions are the subject of the following sections. We will look at the different directions in which the hypotheses involved in the derivation of the MRE can fail, in a systematic way. Since we are interested in the first effects, we will assume that it is adequate to look at each term independently. Furthermore, the additive nature of the different forces suggests that each one can be treated independently. We will therefore do so when possible, making an effort to set specific numerical bounds to the applicability ranges.

2.2.1 Inertial effects: first finite- Re_p effects

The well known Stokes solution of the steady, low Reynolds number flow past a sphere (see Eq. (2.31)) is obtained by completely neglecting fluid inertia. By applying the no-slip boundary conditions on the particle surface and the far-field velocity conditions at infinity and expanding the stream function in powers of Re_p , the hydrodynamic force on the particle

can be calculated, to leading Re_p order, yielding the drag force of the MRE without the Faxén terms (this is the problem solved by Stokes in 1851). However, if one wishes to increase the order of this approximation following the same procedure, one soon realizes that there is no way to fulfil the far field boundary conditions in this case, since the higher order contributions to the perturbation caused by the particle do not vanish at infinity. This phenomenon is known as Whitehead’s paradox [248].

Its resolution was possible due to ideas from Oseen, who noted that the assumption by which inertial terms are disregarded (under Eq. (2.8)) is only valid near the particle, where viscous effects dominate. But, far from the particle (in particular at a distance such that $Re_p = \mathcal{O}(1)$, see [284]), the approximation breaks down. This implies that there is an inherent inconsistency in requiring that the Stokes solution be valid in an infinite domain, and that it is necessary to consider inertia far from the particle in order to calculate the higher order corrections to the drag force. The final word on the issue was nonetheless given by Proudman and Pearson [284], who, using the technique of matched asymptotic expansions, re-derived Oseen’s inertial, first-order correction to the steady drag force, corrected a flaw in Oseen’s original reasoning and added an additional term to the expansion:

$$\mathbf{F}_D = 6\pi a\mu \left(1 + \frac{3}{8}Re_p + \frac{9}{40}Re_p^2 \log Re_p + \mathcal{O}(Re_p^2) \right) (\mathbf{u} - \mathbf{v}) \quad (2.11)$$

This way, the contribution from the outer region becomes a correction of the steady drag force, whose leading term, given by $3/8Re_p$, could be used to provide the order of magnitude of this correction.

In order to illustrate how it could be used in practice, let us assume that we wish to establish the upper limit to the particle Reynolds number to ensure that the correction is smaller than 5% as a convention to fix the range of validity of the MRE. Then the leading term in Eq. (2.11) predicts this to happen for $Re_p \approx 0.14$, and by this point the error in this expression is only ≈ 0.03 , taking as a reference the empirical drag coefficient by Clift et al. (1978) (which has a root mean square error of ≈ 0.01 with respect to the empirical data gathered by Brown and Lawler [50] covering the range $0 < Re_p < 4000$). This means that its range of validity is large enough to estimate the first-effects of inertia for error tolerances lower than 5%. In Fig. 2.1 all these different approximations to the drag force are compared.

The above expression of the force is only valid for long-term steady motions. Lovalenti and Brady [223] studied the motion of a small sphere immerse in an arbitrary space and time varying flow, although considering the particle small enough, so that the undisturbed flow can be considered uniform over the diameter of the particle. The problem is thus the generalization of the MRE (excluding Faxén terms, as these authors assume flow uniformity;

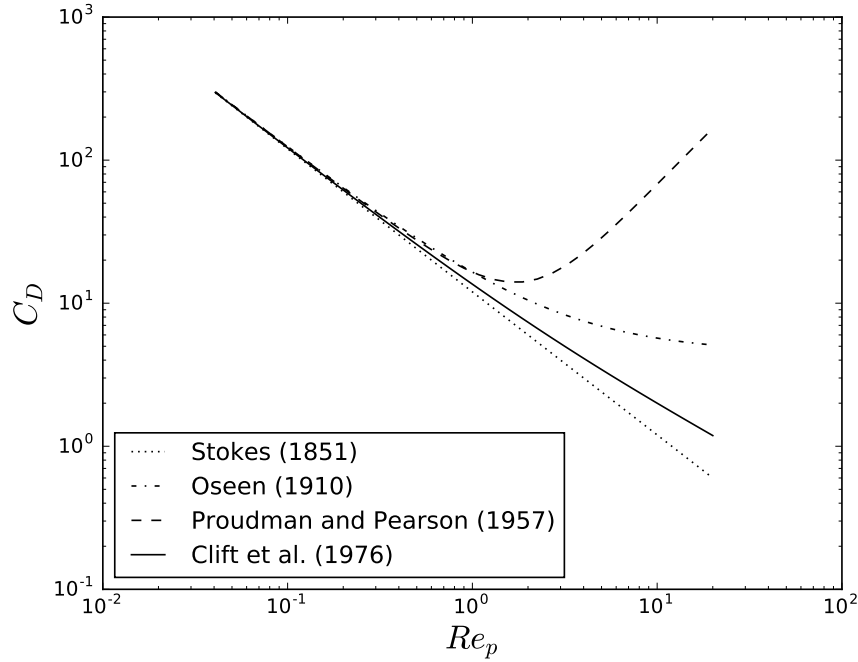


Fig. 2.1 Drag force predicted by different approximations, as measured by the drag coefficient $C_D = 2 \|\mathbf{F}_D\| / (\pi \rho_f \mu)$.

see Section 2.2.1) for small, though non-zero particle Reynolds number. Their analysis is based on a different approach compared to the asymptotic expansion matching used by Proudman and Pearson [284]). In fact, they find an expression for a uniformly valid flow over the whole domain to first order in Re_p . The hydrodynamic force on the particle is then computed to first order in Re_p by approximating the resulting integrals of the stress tensor. The result is an expression for the steady and unsteady forces at finite Reynolds number. The inertial effects include the steady Oseen correction to the drag, as well as corrections to the unsteady forces like the Basset Force and the added mass force. These corrections confirm the observations from fully-resolved numerical simulations [243] that higher rate of convergence of the particle velocity to terminal conditions (i.e. it varies but is t^{-2} for smoothly accelerating motions where the slip velocity tends to increase) than that predicted by the MRE equation, which is $t^{-1/2}$, as predicted by the Boussinesq–Basset kernel. According to Lovalenti and Brady [223]: ‘This fact may explain why experimentalists have measured a steady terminal velocity when the length of their apparatus would not have permitted this if the Basset force was correct.’.

This change in the decay regime of unsteady forces only applies at long times. For short-time-scale motions the unsteady force from the MRE is accurate. The boundary between both regimes is marked by the time it takes for vorticity to diffuse away from the particle up

to the so-called Oseen distance, where inertial and viscous contributions become of the same order. At that point, convection transport takes over as the dominant physical mechanism for the transport of vorticity, explaining the change in regime. From the application point of view it is important to realize whether the flow is dominated by high frequency flows, so that the motion is mainly explained by the MRE unsteady forces, or else the characteristic frequencies of the flow are low and thus the (substantially more complicated) expressions provided by Lovalenti and Brady apply. The key is to compare those characteristic frequencies of the particle's motion with the quantity $\tau_\nu = a^2/\nu$, the viscous diffusion time, which corresponds to the order of magnitude of the time it takes vorticity to be diffused away to the Oseen length. In other words, the MRE equation applies whenever [224]

$$\tau_p \ll \frac{\nu}{U^2} \quad (2.12)$$

After a little algebra, this condition can be rewritten as a function of the usual nondimensional parameters as

$$Re_p^2 \ll \frac{9}{2\rho + 1} \quad (2.13)$$

As the value of Re_p increases, the condition above may cease to be fulfilled and the form of the history term must be corrected using, for example, the formulation given by Lovalenti and Brady [223]; see also Lovalenti and Brady [224]. In turbulent dispersion, the importance of this correction will not be critical for most flows such that $Re_p \ll 1$, except perhaps for solid particles in gas. The same conclusion was reached by [211] based on the empirical formulation for the history force of Mei and Adrian [243].

The effect of vorticity

As stated, the derivations in Lovalenti and Brady [223] assume a uniform flow at the scale of the particle. If one allows for some non-uniformity to be present in the flow, additional effects appear. In particular, this allows for a break in the symmetry of the flow on the plane orthogonal to the direction of w , which gives rise to lateral forces on the particle that are known as *lift forces*. Famously, Saffman [299] (with the corrections in [300]) gave the correct expression, to first order in Re_p , for the lift force that an isolated, non-rotating particle experiments under constant shear (linear non-uniformity of the flow).

$$\mathbf{F}_{L,\text{Saff}} = 6.45a^2 \sqrt{\frac{\mu_f \rho_f}{\|\nabla \times \mathbf{u}\|}} (\nabla \times \mathbf{u}) \times \mathbf{w} \quad (2.14)$$

where $\mathbf{w} = \mathbf{v} - \mathbf{u}$ is the slip velocity. This result is valid under the restrictions $Re_p \ll 1$, $Re_s \ll 1$ and $Re_p \ll \sqrt{Re_s}$; where the shear Reynolds number is defined by

$$Re_s := Re_p \frac{a \|\nabla \times \mathbf{u}\|}{\|\mathbf{w}\|} \quad (2.15)$$

that is, the particle Reynolds number multiplied by a dimensionless shear rate gradient.

Unfortunately, the typical values for Re_s in turbulence are often smaller than those of Re_p [364], invalidating the application of the formulation. However, in this case the expression is in fact *overestimating* the lift force [364] and so, without loss of generality, we may apply the formulation as an upper bound for the remaining discussion.

But the fluid vorticity can arise due to the solid rotation of the fluid as well, also giving rise to a lift. Its low- Re_p analytical formula was provided by Herron et al. [164] and reads

$$\mathbf{F}_{L,Herr} = 8.6a^2 \sqrt{\frac{\mu_f \rho_f}{\|\nabla \times \mathbf{u}\|}} (\nabla \times \mathbf{u}) \times \mathbf{w} \quad (2.16)$$

which has a coefficient about 20% larger compared to $\mathbf{F}_{L,Saff}$. This expression is subject to the restriction $Re_p \ll 1$. Note that $\mathbf{F}_{L,Saff}$ and $\mathbf{F}_{L,Herr}$ are incompatible and are only strictly valid for their corresponding ideal fluid motions. In fact, Candelier and Angilella [56] have proved analytically that for a particle settling in a solid-body rotation fluid, the lift force can even take the opposite sign as that indicated by Eq. (2.14), which is valid only for pure shear, steady flows. In this case, furthermore, the relative motion is not stationary, due to the migration of the particle in the radial directions.

Indeed, as with rectilinear motion, the relative flow unsteadiness also generates history-dependent contributions. Those have been studied in [60, 55] for pure shear (generalizing Saffman's result) and solid-body rotation motions respectively.

Nonetheless, the order of magnitude of the Saffmann lift force, compared to the steady drag force gives an idea of the magnitude of this correction. An analogue estimate can be obtained if fluid motion is closer to that assumed in Eq. (2.16).

$$\frac{F_{L,Saff}}{F_D} \lesssim \frac{6.45a^2 \rho_f \nu^{1/2} \Omega_s^{1/2} W}{6\pi a \mu W} \sim \frac{1}{3} Re_s \quad (2.17)$$

We will thus consider the lift force to be very small within the range of applicability of the MRE. Nonetheless, one must keep in mind that the correction is orthogonal to the drag force and thus its effect can introduce important systematic changes in the particles' trajectories, especially in flows with a predominant direction. Since the largest values of the shear rate are

usually close to the boundaries of the domain, this is very often the case anyway, for instance in internal flows. When considering the flow of particles along ducts, it is therefore important to consider the possibility that the inertial lift force might play a role, except for extremely small particles. In any case, we leave the study of the effect of boundaries for future work, see Section 2.2.6, and thus we will not elaborate this point any further here.

The effect of particle rotation

The linearity of Stokes flow means that, for a spherical particle, rotation and translation are decoupled. Indeed, the combined motion is the result of adding the effects of one and the other, and it is clear that, by symmetry, the relative rotation of the particle alone cannot produce a force in any particular direction. It is however necessary to examine how fast this rotation can be until the first inertial effects appear, invalidating the MRE.

The first-effects of the force modification due to non-zero particle angular velocity were calculated by Rubinow and Keller [294], to zeroth-order accuracy in Re_p . This is enough for our arguments, since the MRE assumes this effect to be very small. This formulation predicts that due to the rotation of the particle, there arises a lift force given by

$$\mathbf{F}_{L,\text{Rub}} = \pi a^3 \rho_f \boldsymbol{\Omega}_p \times \mathbf{w} (1 + O(Re_p)) \quad (2.18)$$

where $\boldsymbol{\Omega}_p$ is the angular velocity of the particle. In other words: the Reynolds number defined by the slip velocity produced by rotation should be small enough. The lift force that arises due to the rotation of the particle relative to the fluid is generically called *Magnus effect* [219].

The condition of having a negligible contribution to the force could then be expressed as a function of the particle Reynolds number associated to the rotational motion, given by

$$Re_\Omega := \frac{a^2 \Omega_p}{\nu} \quad (2.19)$$

where $\Omega_p = \|\boldsymbol{\Omega}_p\|$. Note that $Re_\Omega = a/L Re_p$, and so this is assumed to be very small within the theory of the MRE.

Maxey [238] argues that in turbulent flows the particle will tend to acquire a rotational velocity of the same order as the local shear rate. Taking into account that the first-order rotation-induced effect of inertia is a lift force with a magnitude of order Re_Ω [294], the effect results of order $a^2 U / (L\nu)$; more precisely, we have

$$\frac{F_{L,\text{Rub}}}{F_D} \sim \frac{\pi a^3 \rho_f \Omega_p W}{6\pi a \mu W} \sim \frac{\pi a^3 \rho_f U / LW}{6\pi a \mu W} = \frac{1}{6} Re_\Omega \quad (2.20)$$

where U/L give the estimation of local shear rate; U and L being the characteristic velocity and distance of the flow, which is certainly a small quantity by Eq. (2.7).

There are circumstances in which the particle angular velocity might be substantially higher. For instance, a collision against a wall or another particle might transform a good part of the translational energy into rotational energies, leading to high angular velocities, especially for the smaller particles ². The fact that this force is perpendicular to the direction of translation makes this effect even more important. Nevertheless, collisions are not expected to be too frequent in this regime (see Section 2.2.4), substantially alleviating this effect.

In this chapter we ignore the rotational degrees of freedom of the particle, except for the present discussion, with the argument that these effects are, to first order of approximation, decoupled. Nonetheless, it is worth to consider momentarily what the rotational dynamics of a particle within the range of applicability of the Maxey–Riley regime looks like. Feuillebois and Lasek [128] derived the expression of the instationary motion of a small, rigid sphere spinning in a viscous, Newtonian fluid, providing history-dependent terms that decay much faster than the translational analogues. In the steady-state limit, the equation reads

$$I_p \frac{d\Omega_p}{dx} = -8\pi a^3 \mu \Omega_p \quad (2.21)$$

This equation allows us to calculate the rotational relaxation time as

$$\tau_{p,\Omega} = \frac{I_p}{8\pi a^3 \mu} = \frac{\rho a^2}{15\nu} \quad (2.22)$$

from which a rotational Stokes number can be defined (assuming the steady-state forces are dominant)

$$St_\Omega = \frac{\tau_{p,\Omega}}{T} = \frac{1}{15} \rho \frac{a^2}{\nu} \frac{1}{T} \sim St \quad (2.23)$$

which shows that this Stokes number is of the same order as the translational Stokes number, under the assumption that the fluid time scales are similar. Note that for large Stokes numbers the assumption above that the particles rotation will be of the order of the fluid vorticity most of the time (in fact equal to the local angular velocity or one half of the vorticity in the limit of zero inertia) can be violated, as the particles may not have time to accommodate to the fluid vorticity. In such cases the analysis becomes more involved and it is likely that finite Reynolds number effects must be taken into account for the angular dynamics. Nonetheless,

²Indeed, a fixed percentage of energy conversion from translation to rotation leads to higher angular velocities for smaller particles. Specifically, the rotational kinetic energy is $\frac{1}{2}\Omega_p^2 I \propto \Omega_p^2 a^5$ and the translational kinetic energy is $\frac{1}{2}\|\mathbf{v}\|^2 m_p \propto \|\mathbf{v}\|^2 a^3$. Thus, after a collision we have $\Omega_p \propto \|\mathbf{v}\|/a \propto U/a$, for light and/or small particles. Furthermore, this gives $Re_\Omega \sim aU/L \sim Re_p U/W$; and the ratio U/W can become quite large.

for this order-of-magnitude analysis the present formulation will suffice, especially since the rate at which the steady-state formulation becomes inaccurate is quite slow, and it still yields reasonable results at Re_Ω well past unity, see [219].

2.2.2 Finite radius effects

As the particle radius grows with respect to the characteristic length scale of the unperturbed flow the Faxén corrections become increasingly important. Eventually, even these second-order corrections become insufficient to accurately characterize the surrounding flow and the MRE breaks down as an appropriate model. The rate at which this happens can be analysed by taking one more term in the series expansion that leads to the appearance of the Faxén corrections [222] (see also Eq. 8.175 in [91]). It is useful to consider the expression of the drag force as a surface integral to measure the error. The Faxén corrected drag force becomes, after taking one more term in the same expansion considered in [222]

$$\mathbf{F}_D = \frac{3\mu}{2a} \int_{S_p} \mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x}_p) \, dA \approx 6\pi a\mu \left(\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x}_p) + \frac{a^2}{6} \nabla^2 \mathbf{u} + \frac{a^4}{120} \nabla^4 \mathbf{u} \right) \quad (2.24)$$

Where the ∇^{2k} is the k -th successive composition of the vector Laplacian.

Now, for low Re_p , the fluid is expected to present only long-wave length variations across the particle diameter. Furthermore, since our interest lies on the integrated value of such variations, it appears reasonable to assume that the characteristic length scale of the new term is also the particle radius. This allows us to compare its order of magnitude to that of the usual Faxén correction. Let us therefore assume that the effect of the new term is of the same order as that produced by quadratic variation of $\nabla^2 \mathbf{u}$ across the particle diameter. Then, an adequate order-of-magnitude estimate of its value is given by (see, e.g. [245] for a discussion on the estimation of order of magnitude scales)

$$\left\| \frac{a^4}{120} \nabla^4 \mathbf{u} \right\| \sim \frac{a^4}{120} \frac{\|\nabla^2 \mathbf{u}\|}{2a^2} = \frac{1}{40} \left\| \frac{a^2}{6} \nabla^2 \mathbf{u} \right\| \quad (2.25)$$

So that the next order correction will be significantly smaller than the classical Faxén correction up to the point when Re_p is not small any more (in the limit of quasi-steady Stokes flow it holds $\nabla^4 \mathbf{u} = 0$ exactly). This means that in practice the small Re requirement will guarantee that the Faxén corrections will make a good enough job and thus the restriction Eq. (2.9), that is, $a/L \ll 1$ by itself does not appear to be restrictive, but rather indirectly through the violation of $Re_p \ll 1$. A similar argument can be made with the other forces bearing their corresponding Faxén correction. A significant study supporting this conclusion

in the context of isotropic turbulence was provided by Homann and Bec [171], who tested the performance of the Faxén terms for a neutrally buoyant particle in direct numerical simulation (DNS) of isotropic turbulence. Their conclusion was that the first effects of the finite size of the particles were well captured by these terms up to $a \sim 4L$ (where L is the Kolmogorov microscale). From then on, inertial (finite Re_p) effects quickly kick in.

2.2.3 Nonsphericity effects

No particle is perfectly spherical in reality and thus it is necessary to acknowledge this fact in assessing the reliability of the MRE for predicting the trajectories of real particles. Departure from the ideal spherical shape greatly complicates rigorous analysis and, most often, precludes it completely. The situation is somewhat similar to the consideration of nonsphericity in granular flows using DEM. It is known that its effects are important although only very rough approximations to them can be captured when using only spheres. And while the possibility of using nonspherical shapes also exists, it is greatly limited in practice due to increased numerical costs and difficulties in achieving a good characterization of realistic shapes.

Nevertheless, there has been substantial theoretical progress in characterizing the hydrodynamic interaction of particles of nonspherical shape at low Re_p ; see [219] and references therein, [114, 59, 58]. Gavze [142] derived a generalized version of Eq. (2.1) for a body of arbitrary shape written in terms of its viscous resistance tensors [48]. Within this model translation and rotation become coupled.

Of particular relevance to our arguments is the study of Zhang and Stone [389], who derived the first order asymptotic correction to the unsteady equation of motion of a sphere in quiescent fluid in an amplitude parameter, ε , of the deviation from an ideal spherical shape using the reciprocal theorem (see, e.g. [152]). Their equations particularize the work in [142] to the limit of weak nonsphericity, in which the coupling between rotation and translation comes in as a higher-order correction and can thus be neglected to first-order approximation.

Specifically, let the particle's surface S_p be defined in polar coordinates around its center as the set of points with coordinates r , θ and ϕ such that

$$\frac{r}{a} = 1 + \varepsilon_p f(\theta, \phi) \quad (2.26)$$

where a is the radius of the equivalent-volume sphere.

Their equation for the hydrodynamic force on the translating particle reads:

$$\mathbf{F} = (\mathbf{I} - \varepsilon_p \mathbf{H}) \mathbf{F}_D + (\mathbf{I} - 2\varepsilon_p \mathbf{H}) \mathbf{F}_H + (\mathbf{I} - \varepsilon_p \mathbf{H}) \mathbf{F}_A \quad (2.27)$$

where

$$\mathbf{H} = \frac{3}{8\pi} \int_{S_1} f \mathbf{n} \mathbf{n} \, dS \quad (2.28)$$

and the integral is over the unit sphere S_1 , after applying a uniform scaling by its radius over the whole space. Note that \mathbf{n} represent the exterior unit normals to the unit sphere. The surface of the particle is by construction (although this is not emphasized in [389]) assumed to be star-shaped with respect to its center of mass³. We can then use these results as a conservative estimate of the first perturbations due to nonsphericity to the MRE.

Eq. (2.27) allows us to estimate an upper bound to the magnitude of the corrections to the MRE that are needed to take into account generic (weak) nonsphericity. Let us consider the induced matrix norm on matrices denoted by $\|\cdot\|$, based on the usual Euclidean norm for vectors, denoted with the same symbol. Then, for example, the corrected drag force $\mathbf{F}_{D,\varepsilon_p} = (\mathbf{I} - \varepsilon_p \mathbf{H}) \mathbf{F}_D$ fulfils

$$\frac{\|\mathbf{F}_{D,\varepsilon_p} - \mathbf{F}_D\|}{\|\mathbf{F}_D\|} \leq \|\varepsilon_p \mathbf{H}\| = \frac{3\varepsilon_p}{8\pi} \left\| \int_{S_p} f \mathbf{n} \mathbf{n} \, dS \right\| \leq \frac{3\varepsilon_p}{8\pi} \max_{S_p} f \frac{4}{3}\pi = \frac{\varepsilon_p}{2} \max_{S_p} f \leq \varepsilon_p \quad (2.29)$$

Similar estimates can be worked out for the other hydrodynamic effects, yielding similar relative orders in their corrections. This bound is related, although a bit less sharp than the one provided by the application of a theorem by Hill and Power [166], which establishes that the magnitude of the steady drag (it only refers to the steady drag force) on the particle lies between that obtained on an inscribed and a circumscribed spherical particles. That is, the center of the circumscribed and inscribed spheres do not have to be the center of gravity. In this case therefore, for steady-state motions, the relative error made by taking the Stokes expression with the average diameter is smaller than

$$\mathcal{O}\left(\frac{a_{\text{circ}} - a_{\text{insc}}}{2a_{\text{insc}}}\right) \quad (2.30)$$

In the limit of small corrections to the spherical shape, however, all inscribed and circumscribed spheres tend to collapse to the ones around the center of mass and the result then coincides, for the drag force, with the one provided by the asymptotic technique by Zhang and Stone [389].

³In geometry, a body is said to be *star-shaped* if there exists a point within it such that, for any other point in the body, the straight segment that joins both points is itself fully contained in the body. This appears to be a reasonably weak assumption for approximately spherical particles. In particular, it is much less restrictive than convexity. See [156] for an analytic correction to the added mass force for a (non-star shaped) rough sphere.

2.2.4 Effect of neighbouring particles

The MRE is based on the assumption that the particle is isolated, far away from any neighbouring particles or any other boundary. But, for many applications, such assumption can become too restrictive. Even in the absence of interactions such as electrical potentials or van der Waals forces between the particles, there are still two types of inter-particle influences:

- Mid/long-range hydrodynamic interactions, due to the flow disturbances generated by the presence of all the other particles
- Short-range forces of various types that appear when the inter-particle gaps become very small, such as lubrication forces and inter-particle contact forces

Starting with the long-range hydrodynamic interactions, it is clear that they depend on the configuration of neighbours and their velocities, requiring a statistical approach for a general analysis. Under the low-Reynolds number hypothesis, such influences are linear on the particles' separations (and velocities, including angular velocities), meaning that their spatial decay is slow. Every particle in a suspension produces a disturbance flow around it which, under the hypothesis of a low particle Reynolds number, is accurately represented by the creeping flow equations solution, at least up to a distance comparable to the Oseen distance (a/Re_p)[248] from the particle's center, i.e.

$$\mathbf{u} = \|\mathbf{w}\| \begin{pmatrix} \cos \theta \left(1 - \frac{3}{2} \frac{1}{r} + \frac{1}{2} \left(\frac{1}{r} \right)^3 \right) \\ \sin \theta \left(1 - \frac{3}{4} \frac{1}{r} - \frac{1}{4} \left(\frac{1}{r} \right)^3 \right) \\ \sin \theta \left(1 - \frac{3}{4} \frac{1}{r} - \frac{1}{4} \left(\frac{1}{r} \right)^3 \right) \end{pmatrix} \quad (2.31)$$

where the velocity components are Cartesian (with the first component being aligned with the relative flow) but have been parametrized in spherical coordinates, with r the radial coordinate (normalized by the sphere's radius) and θ the azimuthal angle. Note that, by symmetry, the field must be independent of the polar angle, as it is. Eq. (2.31) confirms that the decay of the disturbance intensity is dominated by a linear behaviour, which means that its influence can still be appreciable at significant distances as shown in Fig. 2.2. In Fig. 2.3 the average distance to the nearest 2^k -th neighbour, as a function of α_p , is shown for reference; for $k = 0 \dots 5$ in a random array of spheres. The calculation of this distance is done according to the formulation in [340], that takes into account the space taken up by the finite-size particles and is robust for all values of α_p up to close packing.

A description of the flow around a given particle, under the influence of its neighbours, would require the solution of the Stokes equations around all of them simultaneously. And

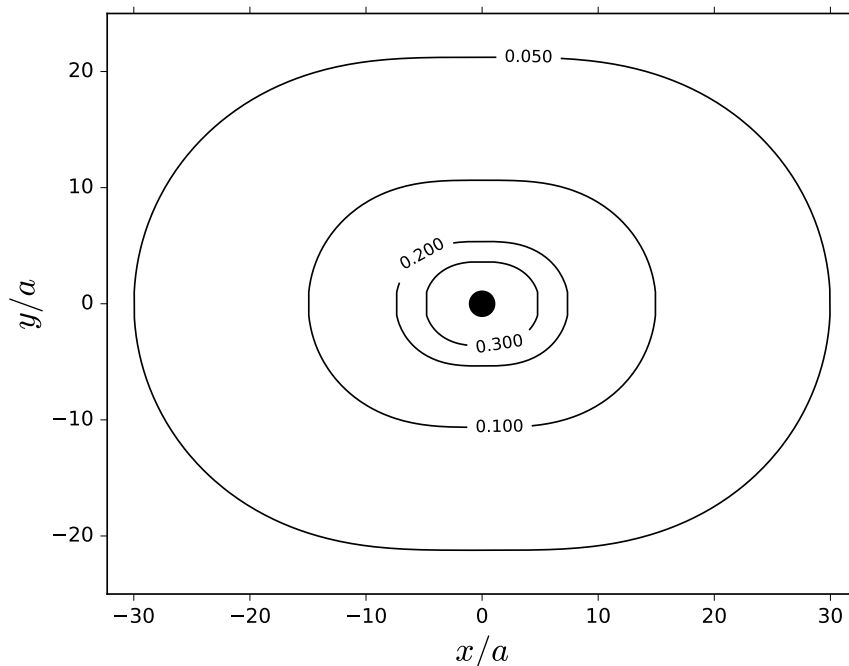


Fig. 2.2 Modulus of the flow disturbance around a particle (black circle), calculated according to Eq. (2.31) and normalized by the modulus of the slip velocity, $\|\mathbf{w}\|$. All distances have been normalized by the radius of the particle, a .

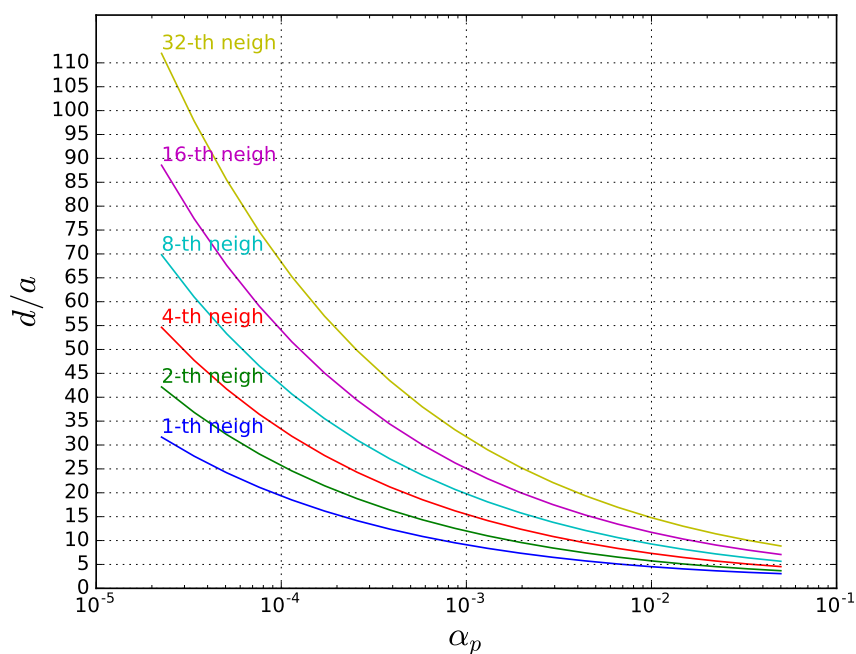


Fig. 2.3 Average distance (normalized by the radius of the sphere) to k -th closest neighbour in a random array of monodisperse spheres.

since the solution depends on the velocities of the particles, which are in turn instantaneously determined by the fluid velocity around them, all interactions are completely coupled, making the problem fully multi-body in general. In the dilute limit that we are interested in (we are studying the first-effects of the presence of neighbours), it is, however, possible to give the interactions a *pair-wise* treatment. This means that they can be calculated as the sum of the interaction between the individual pairs of neighbours with only a small error [23]. Still, the main difficulty lies in describing a representative configuration of relative positions and velocities for a representative ensemble of particles, valid for the general case.

Switching to the short-range interactions, these include two types of forces: the shortest-range hydrodynamic interactions, termed *lubrication effects*, and the direct contact interactions, that appear once the fluid has been completely squeezed out of the gap between pairs of particles coming into contact. Here too, a statistical analysis seems unavoidable. Fortunately, these effects are more adaptable to the general DEM approach, as these are by definition pairwise, short-range forces that can be naturally included in a contact model. Nonetheless, in some situations it may still be interesting to neglect them completely as:

1. The complicated modelling issue associated with the two types of short-range interactions is avoided.
2. The computational cost is greatly reduced, as the most expensive parts of the DEM algorithm (force calculation and search) are avoided.
3. The time step of the simulation, which is otherwise upper bounded by a fraction of the contact duration, can then be much increased, further reducing the cost.
4. It becomes easier to give a completely statistical interpretation to the particles, making it possible to alter the real concentration of particles to speed up the simulations, avoiding for instance the need to consider a realistic number of particles when that is too costly.

We wish to characterize the conditions for which these effects start to become important, causing the MRE to start losing accuracy in modelling particle-laden flows. Such characterization has been attempted before [116, 118] in the context of general particle-laden flows. Specifically, Elghobashi [116] establishes limits on the global solid fraction (the proportion of volume occupied by the particles in the whole domain; see Section 4.8 and Appendix H) for which three-way interactions become important is

$$\alpha_p \lesssim 10^{-3} \quad (2.32)$$

where α_p is the local solid fraction. This corresponds to an average nearest-neighbour separation of about 10 particle radii. This is an oft-cited limit[232, 159, 206]

We will attempt to enrich the general picture here, surveying the most relevant results and the different aspects to take into account with the ultimate goal of producing a useful guide for practical applications. Once again, we focus on turbulent flows to provide the order-of-magnitude estimates.

Before proceeding, we must make one further remark concerning the scope of the discussion. In some particle-laden systems, the particles may interact forming aggregates, coalescing or breaking. We do not consider these cases here, but rather the case in which the total number of particles and their shape is preserved throughout the simulation.

Hindered settling

In a random arrangement of particles, the added effect of all particles around any one of them can be very substantial, due to the long-range character of the interactions that we discussed above. In fact, the naive summation of the moduli of the disturbances would not work at all for our order-of-magnitude purposes, because of the slow decay of the interactions with distance. Indeed, take the sum of the norms of all the pairwise interactions between the particles inside a ball centred at the target particle and the particle itself, which bounds the contribution of the particles considered above. By considering a sequence of balls of growing radii, one obtains a sequence of sums that, due to the linear decay of the forces with distance, does not lead to an absolutely convergent series, because the number of terms grows with the third power of the radius in a statistically homogeneous distribution. Thus, such straightforward strategy fails to produce any bound at all.

This difficulty was overcome by Batchelor [23], who was able to rigorously calculate the (ensemble) average, to first-order in the solid fraction, of the effect of a stable, random uniformly distributed suspension of identical rigid spheres on the settling velocity of any given particle. This author used an ingenious method in which known averaged quantities were used to cancel out the slow varying contributions exactly, reducing the remaining terms to rapidly varying quantities that could be calculated, to first order in the concentration, based on the interaction of the two closest particles only. The result of his analysis can be written as

$$\frac{V_{\text{settling}} - V_0}{V_0} = S\alpha_p \quad (2.33)$$

where U_0 is the terminal velocity of the target particle when isolated, V_{settling} the actual settling velocity of the particle in the suspension, $\alpha_p := 1 - \alpha_f$ is the solid volume fraction and S is the *settling coefficient*, which under the particular assumptions of the theorem is $S = -6.55$;

and α_p . This result is based on the further assumption that the probability distribution of the distance between pairs of particles is uniform. A similar expression has been derived for polydisperse suspensions [24, 93], yielding smaller values for S . For instance, $S \approx -2.5$ for both small spheres surrounded by a suspension of larger spheres or the opposite situation (See [100]). Therefore, one may use Eq. (2.33) as a conservative order of magnitude of the first effect of the presence of neighbouring particles in sedimentation of statistically homogeneous suspensions. Note that in this case (S is negative) the presence of other particles tends to *slow down* the target particle as it sediments. This is true specifically *because* of the assumption of having a uniform suspension that can be treated as unbounded, since non-uniformity often leads to a reduced resistance and thus an increased settling velocity with respect to an isolated particle, as will be shown below.

Drag modification

Another relevant theoretical result can be taken as a reference to study the effects of the presence of a random array of particles around a target particle for the calculation of the drag force, which is enough to assess first-order effects due to the surrounding suspension. The theory is based in the asymptotic analysis of the hydrodynamic forces in the low Reynolds number and solid fraction limit, which is relevant to the present discussion. It was put forward by Kaneda [189] and experimentally verified up to solid fractions of around 0.04 for small Reynolds numbers based on Brinkman's screening length, $\frac{\sqrt{3}}{2}\phi^{-1/2}a$, as the simulations for higher values of Re were found to be 'prohibitively expensive', in [167].

The derivations by Kaneda were performed 'for simplicity' under the hypothesis $\alpha_p \sim Re_p^2$, although the resulting expression is shown to have the correct asymptotic behaviour at the opposite limits $(\alpha_p, Re_p) \rightarrow (0, Re_p)$ and $(\alpha_p, Re_p) \rightarrow (\alpha_p, 0)$. In this work it is argued that the first effects of inertia (see Section 2.2.1) and those of the nonvanishing solid fraction cannot be treated independently as both are inextricably linked for all but essentially zero Reynolds numbers. The expression for the modified drag is

$$\mathbf{F}_D^{\text{Kaneda}} = C_K \mathbf{F}_D \quad (2.34)$$

with

$$C_K := (1 + Re_p K^{\text{Kaneda}}(S^*)) \quad (2.35)$$

where S^* ⁴ is a nondimensional constant such that $S^* = 9/2\alpha_p/Re_p^2 + o(\alpha_p)$ (when $\alpha_p \rightarrow 0$) and where \mathbf{F}_D is the usual Stokes drag; and

$$K^{\text{Kaneda}}(S^*) = \frac{3}{8} \left((2S^* + 1)\sqrt{(4S^* + 1)} - 4S^{*2} \log \frac{\sqrt{(4S^* + 1)} + 1}{\sqrt{(4S^* + 1)} - 1} \right) \quad (2.36)$$

Note that

$$\begin{aligned} \mathbf{F}_D^{\text{Kaneda}} &\sim \left(1 + \frac{3}{8} Re_p + \frac{3}{2} Re_p S^* \right) \mathbf{F}_D \\ &\sim \left(1 + \frac{3}{8} Re_p + \frac{27}{4} \alpha_p Re_p^{-1} \right) \mathbf{F}_D \quad \text{for } S^* \rightarrow 0 \\ \mathbf{F}_D^{\text{Kaneda}} &\sim \left(1 + Re_p \sqrt{S^*} + \frac{3}{40} \frac{Re_p}{\sqrt{S^*}} \right) \mathbf{F}_D \\ &\sim \left(1 + \frac{3}{\sqrt{2}} \alpha_p^{1/2} + \frac{\sqrt{2}}{40} Re_p^2 \alpha_p^{-1/2} \right) \mathbf{F}_D \quad \text{for } S^* \rightarrow +\infty \end{aligned} \quad (2.37)$$

Which, as stated, recover the first order correction due to Oseen for the drag force at $S^* \rightarrow 0$ (and thus $\alpha_p \rightarrow 0$), see Eq. (2.11), and also the classic expression by Brinkman [49], at $Re_p \rightarrow 0$.

Fig. 2.4 shows the magnitude of the correction coefficient C_K for a wide range of solid volume fractions, for different values of Re_p . Note that the correction to the drag is significant, even for very small Reynolds numbers if the solid volume fraction is greater than 1×10^{-3} , in accordance with the rule of thumb of Eq. (2.32), and it still appreciable past the 1×10^{-4} mark.

Although Kaneda's result applies to arrays of *fixed spheres* under a uniform background flow, it clearly reflects the inextricable relationship between the low-Reynolds number and low-concentration hypotheses through the dimensionless number S^* . This number can be understood as the quotient between two length scales:

$$S^* = \frac{9}{2} \left(\frac{L_{\text{Oss}}}{L_{\text{Brink}}} \right)^2 \quad (2.38)$$

where $L_{\text{Oss}} = a/Re_p$ is the Oseen length (distance from the particle at which inertia becomes as important as viscosity) and $L_{\text{Brink}} = a/\sqrt{\alpha_p}$ (distance from the particle at which the Stokesian disturbance caused by it switches from a linear to a cubic-order decay due to the presence of neighbours). The physical interpretation of the two regimes in Eq.(2.36)

⁴The asterisk has been added to avoid a symbolic clash with the settling coefficient, S .

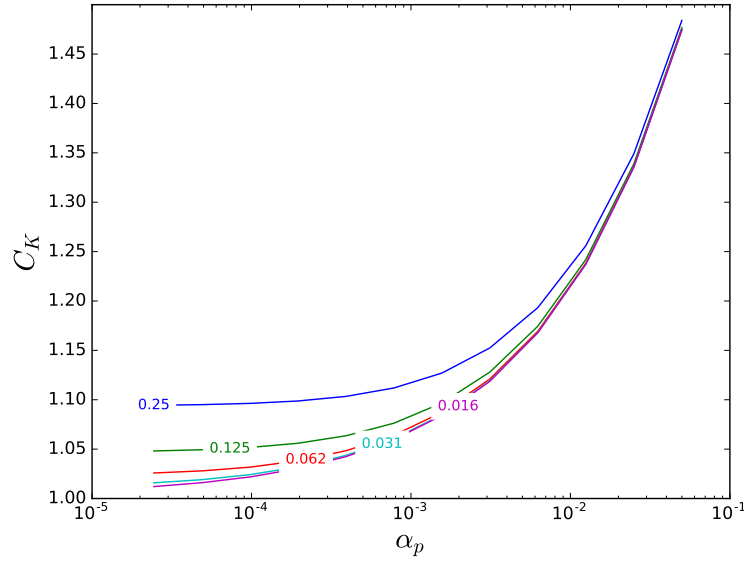


Fig. 2.4 Correction factor of the steady drag force as a function of the average solid fraction for a randomly distributed array of spheres. The different curves are labelled according to their corresponding Re_p .

goes as follows: when S^* is very small ($L_{\text{Brink}} \gg L_{\text{Oss}}$), inertial effects enter much before the Stokesian disturbance has been screened by other particles and thus their influence is dictated by finite- Re_p effects. On the contrary, if $L_{\text{Brink}} \ll L_{\text{Oss}}$, the influence of distant particles is screened before it can reach the target particle, and only the effect of Stokesian flow disturbances is felt.

Finally, note that the concentration dependence of the drag force is proportional to the square root of the solid fraction, to first order, under the assumption of having a small Reynolds number based on the system size. This contrasts with the result by Batchelor discussed in Section 2.2.4, under which the resistance depends linearly on the solid fraction under the same hypothesis. The reason is the hypothesis of a fixed bed used by Kaneda, which constraints the possibilities that the particles have to accommodate to the resistance, causing it to be more important. The hypothesis of a free bed adopted by Batchelor is more relevant to the behaviour of a cloud of particles suspended in a fluid [92]. Nonetheless, the description of the two regimes studied by Kaneda does remind us that the assumption of having a small Reynolds number used by Batchelor is actually tied to the condition of having a large enough solid fraction to make sure that $S^* \ll 1$, a condition that was not explicitly stated by this author.

Recently, Pignatelli et al. [278] have experimentally studied the transition between regimes of a falling (finite) cloud of particles. In this case, the Oseen length scale is compared to the cloud's initial diameter. Furthermore, a Reynolds number associated with the cloud's length

scale is also considered. Several regimes are identified, but what is most interesting to the present discussion is their identification of the Stokes regime (Batchelor’s hypothesis) for $L_{\text{cloud}}/L_{\text{Oss}} < 0.1$, $Re_{\text{cloud}} < 0.1$. We will come back to considering sedimenting clouds in the proceeding paragraphs, which address inhomogeneous suspensions.

Clustering

In reality, the assumption of homogeneity in the spatial distribution of a suspension is rarely fulfilled. Instead, it is often the case that the particles approach each other forming more or less defined clumps, a phenomenon known as *clustering*. This is observed very prominently in turbulence, where inertial particles tend to concentrate in filament-like structures causing the solid fraction to fluctuate strongly within the domain [127, 305, 350, 262].

A number of theories have been proposed to explain these inhomogeneities, and it seems that, rather than a single mechanism, there are several [309, 44, 157], and their relative importance may depend on several factors, including the particle Reynolds number and, especially, the Stokes number. The most well-known of these mechanisms consists in the progressive expulsion of heavy particles from high-vorticity regions into straining-flow regions [112] as a result of their inertia. This is known as the ‘centrifuge’ mechanism and it has been well documented both from simulations [239] and experiments [305]. In other words, the inhomogeneities arise because the particles spend more time in high-strain regions than in high vorticity regions, so they tend to concentrate there. Such oversampling of regions of the flow with specific properties is generally known as *preferential concentration* [323, 145]⁵ in the literature.

While the preferential concentration effect is the most predominant at small Stokes numbers [45], other mechanisms become important as St grows. For instance, the *sweep-stick* mechanism [149] consists in the tendency of particles to move away from regions of high accelerations and *stick* to the low-acceleration zones. The sweep-stick mechanism becomes important at around $St = 1$ and higher (inertial range).

Another effect that becomes important at finite Stokes numbers is the so-called *sling effect* [122] which develops in intense turbulence. This effect is caused by rare events where very large gradients develop, producing jets of particles that are ejected from the trajectories defined by the streamlines of the vortices from which they come, just as a sling does with a stone. The ejected particles can enter regions where the local suspended particles have significantly different velocities, rendering the *particles velocity field* (the hypothetical field formed by the trajectories of an infinitude of particles taking up all space) *multivalued*,

⁵For particles lighter than the surrounding fluid, such as air bubbles in water, the inverse type of oversampling is observed [5], as expected.

a phenomenon known as caustics [367]. These caustics are responsible for an important increase of the collision rate of particles, and is currently believed this plays a crucial role in explaining the fast rain initiation times in turbulent warm clouds [122, 123].

Other mechanisms for the appearance of inhomogeneities are turbophoresis [257] and clustering related to boundary layers [234], although we will stick to homogeneous, isotropic turbulence in our discussion for the sake of simplicity. An important and perhaps surprising fact about the clustering effects is that they are known to continue below the smallest scales of the flow, the Kolmogorov microscales in turbulence. Indeed, inertial particles tend to cluster forming multifractal structures [26] at sub-Kolmogorov scales, in a process that is still not fully understood. Furthermore, this phenomenon is not exclusive of turbulent flows, and can even be produced by random, uncorrelated fields that contain a smallest length-scale, as shown by Bec [25].

Important efforts are being made to characterize the phenomenon and some statistical models, and the development of physical theories to gain insight into clustering phenomena are currently being developed, along with predictive models that are yielding correct qualitative as well as quantitative predictions [382, 381, 73, 131, 157]; see [157] for a review. These models contain a number of assumptions and simplifications that limit their range of validity. Recently, Bragg and Collins [44] reviewed some of these formulations, recommending the one by Zaichik and Alipchenkov [381] as the most comprehensive and robust over a wide range of St values (see below).

Nonetheless, there remain a number of open questions and inconsistencies in the literature that need to be addressed. For instance, there is a significant consensus that the strength of clustering peaks at $St \sim 1$ [253, 178], However Sumbekova et al. [328] concludes that the level of clustering, measured using a Voronoi tessellation of space, is most strongly related to the Taylor Reynolds number, less so to the average particle volume fraction and negligibly on the Stokes number. Moreover, Uhlmann and Chouippe [349] has found that clustering scaling seems to depend on the mode in which it is measured.

Understanding the physical mechanisms that govern the appearance of clustering is of extraordinary importance, not only due to its relevance to the assessment of the validity of the single-particle theory (which includes the MRE), but also in understanding the statistical distribution of particles in space, statistics on the flow properties being sampled (of relevance to the simulation of chemical reactions, for example) and the prediction of the collision rates (of relevance to the study of initiation of rain and snow, for example). Of direct relevance to the present discussion is the study by Aliseda et al. [4], where the settling velocity of spheres was measured experimentally in a turbulent channel flow. A notable increase in the settling velocity was observed with respect to quiescent fluid conditions, especially for Stokes

numbers on the order of unity. Furthermore, this velocity increased monotonically with the overall volume fraction, indicating the effect of three-way coupling. A phenomenological model based on the idea of particle clusters locally altering the average velocity seen by its constituent particles turned out to explain the observations very well. It is remarkable that the whole study was performed under volume fractions well below the limit given in Eq. (2.32), which clearly demonstrates the limitations of the single-particle paradigm.

The estimates provided in Section 2.2.4 were derived under the assumption of an unbounded array of particles characterized by an average solid fraction under the hypothesis of having a homogeneous suspension. Therefore, the existence of significant inhomogeneities as those caused by preferential concentration in turbulent flows, invalidate the theory and the derived estimates. Aliseda et al. [4] proposed a phenomenological model to explain the effect of preferential concentration in a turbulent suspension of settling particles. They observed that the average settling velocity of the particles belonging to a cluster was reasonably well predicted by Eq. (2.33), with

$$S = K_T \alpha_p L_{\text{cluster}}^2 \quad (2.39)$$

where K_T is the volumetric shape factor [285], an order-one constant that depends on the shape of the cluster and is equal to one for spherical clusters and L_{cluster} is the characteristic size of the cluster. This formula was derived under the hypothesis of quasi-steady Stokesian flow *around the cluster*. That is, this work is concerned with the low-Reynolds regime

Interestingly, this model was theoretically ratified by Jabin and Otto [181] under similar conditions, though no mention is made in their paper of the previous investigations by Aliseda et al. This remarkable work seems to have remained relatively unknown, perhaps because of this omission. This work refines the picture sketched by Aliseda et al. by

1. Rigorously proving the following relation

$$V_{\text{settling}} \sim \frac{(\rho - 1)}{\nu} C N_{\text{cluster}}^{2/3} \frac{a^3}{\Delta d_{\text{cluster}}} g \quad (2.40)$$

where C is an unknown constant and where $\Delta d_{\text{cluster}}$ is the average separation between the particles in the cluster. This expression is equivalent to Eq. (2.39) in the regime ($V_0 \ll V_{\text{settling}}$), as it can be immediately shown using $\alpha_p \sim a^3 / \Delta d_{\text{cluster}}^3$ and $L_{\text{cluster}}^2 \sim N_{\text{cluster}}^{2/3} \Delta d_{\text{cluster}}^2$.

2. Establishing the existence of a critical, minimum number of particles ($N_{\text{crit}} = C \left(\frac{d}{R}\right)^{3/2}$) that must be contained in the cluster, so that it behaves as a macro-particle and where Eq. (2.39) applies. The macro-particle argument was precisely the one used in [4].

3. Establishing a lower bound to the time scale in which the particles will not come close enough to invalidate the model, which is the time scale of the time taken by the cluster to settle a distance equivalent to L_{cluster} . This point is fundamental in proving that the macro-particle model is stable for a sufficiently long time to apply the theory effectively.

Further support to this model can be found by using an argument due to Zaichik and Alipchenkov [381] (ZA henceforth). Note that this is the same work describing the model advocated by Bragg and Collins [44] for the description of clustering in turbulence, mentioned above. ZA used the theory of Batchelor [24], a generalization of that discussed in Section 2.2.4 allowing for non-uniform spatial distribution of particles, to provide an expression for settling coefficient S in the much more general setting of isotropic turbulence. The general expression reads:

$$\begin{aligned}
 S &= S_1 + S_2 - 5 \\
 S_1 &= \int_2^\infty \left(A_{11} + A_{12} + 2B_{11} + 2B_{12} - 3 - \frac{3}{r_a} \right) \Gamma r_a^2 dr_a \\
 S_2 &= 3 \int_2^\infty (\Gamma - 1) r_a dr_a
 \end{aligned} \tag{2.41}$$

where r_a is the radial distance normalized by the particle radius a ; A_{11} , A_{12} , B_{11} and B_{12} are the mobility functions of the particle pair [24] and Γ is the *radial distribution function* (RDF), defined as the ratio of the probability density of particle pair relative to the same quantity in a uniform suspension, see, e.g., [380].

In simple words, the RDF gives, assuming a particle is found at a certain location, the average number of neighbours found at any given distance from the former, expressed as a proportion to the average number found at that distance in a homogeneous distribution. Note that this is an infinitesimal quantity. The latter can be approximated as

$$N(r_a) = \underbrace{\frac{\alpha_p}{V_p}}_{n_p} \underbrace{4\pi r_a^2 dr_a a^3}_{|B(r_a+dr_a)-B(r_a)|} \tag{2.42}$$

where the number density n_p is the number of particles per unit volume and $B(r_a)$ is the ball of (nondimensional) radius r_a centred at the particle.

The RDF is often employed in the fundamental study of turbulence dispersion [310, 26]. This function would be exactly equal to one for a statistically homogeneous distribution. When a certain degree of inhomogeneity is present, one observes peaks in its value at distances of the order of the characteristic distance of the inhomogeneities (the diameter of the clusters, say). Typically, the RDF asymptotes to one at infinity, as homogeneity is

often attained at large enough distances. This is roughly the case for many physical systems, including particle suspensions in turbulence [381] and the molecules in an ideal gas in thermal equilibrium [161] and even the distribution of matter in the universe [347].

The different terms in Eq.(2.41) can be identified as the effect of the particle-pair interactions (S_1), the effect of preferential concentration (S_2) and the effect of backflow (the resistance caused by the displaced fluid raising due to the settling of the particles cloud).

Therefore, Eq. (2.41) provides a means to calculate the effect of generic inhomogeneities on the settling velocity of the individual particles. Note that it can also be interpreted as the *effect on the drag force* (since in a time-averaged sense both notions are one-to-one related) of the presence of neighbours in remarkably general conditions, and in a time-averaged sense. This is of interest to produce a first set of estimates for the effect of neighbouring particles in a wide range of situations involving a predominant drag force and arbitrary distribution of particles. We believe it can provide an order-of-magnitude guide for engineers that wonder about the magnitude of this effect. Nonetheless, the whole theory is subject to the assumptions underlying Batchelor’s theory of sedimentation, and thus the restrictions mentioned at the end of Section 2.2.4, defining the Stokes regime for the whole cluster, apply here too.

The key is thus to provide an expression for Γ to calculate the settling coefficient. ZA distinguish between three different regimes, valid for heavy particles ($\rho \gg 1$):

Vanishing inertia ($St < 0.01$) **with no turbulence** In this case there is no clustering and thus $\Gamma \equiv 1$, which leads to the same expression derived by Batchelor in [23], where a random, uniform distribution was assumed. In such case, the settling coefficient is given by Eq. (2.33).

Vanishing inertia ($St < 0.01$) In general, in a turbulent flow field the particles interactions produce a drift velocity of the non-inertial particles towards one another [382, 73]. Taking this drift into account modifies the RDF (see [381]), leading to $S = -4$. That is, there is a small increase of the settling velocity that cannot however compensate the hindering effect completely, making the particles settle slower than isolated, by an amount that is about 40 % less than that corresponding to a laminar settling situation. In terms of order-of-magnitude arguments this is in any case a negligible correction.

Finite but small inertia ($0.01 < St < 0.1$) In this case, the RDF is derived from the equations presented in [381] for low inertia particles, based on a power-law model for the RDF:

$$\Gamma(r_L) = \begin{cases} c_0 r_L^{-c_1} & \text{if } r_L < L_{\text{cluster}} \\ 1 & \text{if } r_L \geq L_{\text{cluster}} \end{cases} \quad (2.43)$$

where L_{cluster} is defined so as to guarantee continuity and $c_0 > 0$ and $c_1 > 0$ depend only on the Stokes number. L_{cluster} can be interpreted as a characteristic size of a cluster, since beyond this length the RDF becomes uniform [381]. While this formulation is only strictly valid (within the ZA theory) at distances smaller than the Kolmogorov micro-length scale L , they use it anyway, warning of its exclusively qualitative value beyond that point. But since we are interested in distances of only up to the order of a few times this distance (the effect at larger distances would be more appropriately dealt with through a two-way coupled approach), and taking into account the excellent agreement of ZA's model with the results reported in [178], fitted from DNS results in the range $r_L \in [0.75, 2.75]$, we will consider it reasonable for order-of-magnitude arguments.

Moreover, ZA point out that for heavy particles, with $\rho \geq 1000$ (although, judging from their [381, Figure 14], this argument could be extended to $\rho \geq 100$), one has $S_2 \gg S_1$ and so using Eq. (2.43) we can write

$$S \approx S_2 = \frac{3}{(a/L)^2} \frac{c_1 c_0^{1/c_1}}{2 - c_1} \quad (2.44)$$

which can be rewritten as

$$S \approx b \left(\frac{L_{\text{cluster}}}{a} \right)^2 \quad (2.45)$$

where (c_1 corresponds to $-\chi$ in [381])

$$L_{\text{cluster}} = c_0^{1/c_1}; \quad b = \frac{3c_1}{4 - 2c_1} \quad (2.46)$$

The values of c_0 and c_1 are reasonably approximated in $St \in [0.01, 0.1]$ by [381]

$$\begin{aligned} c_0 &= 1 + 12St^2 \\ c_1 &= 6St^2 - 9.8St^3 + 6.2St^4 \end{aligned} \quad (2.47)$$

Large inertia ($St > 0.1$) In this case the S_2 outweighs S_1 by a large margin, since the small disturbances caused by the surrounding neighbours have a relatively small effect on the sluggish, heavy particles. Therefore it is possible to take ⁶

$$S \approx S_2 = \frac{3}{(a/L)^2} \int_{2a/L}^{\infty} (\Gamma - 1) r_L dr_L \quad (2.48)$$

⁶Note that there appears in Equation 79 in [381] the lower limit of the integral is 0. Note however that this is inconsequential, since there is no contribution of the integral between 0 and $2a$ because no particle centres can be found closer than $2a$.

where now the radial coordinate (r_L) is normalized by the Kolmogorov micro-scale. ZA provide a set of stochastic, probability conservation equations that can be numerically solved to determine Γ .

The formulation by ZA is reviewed and compared to the alternative formulation by Bragg and Collins [44], who, while not addressing the issue of settling, compare it to the alternative formulation of Chun et al. [73] and give a synthetic review of the most important physical mechanisms that explain the development of inhomogeneities in particulate suspensions in turbulence. Furthermore, these authors propose to introduce a change in the original formulation by ZA to include the non-local diffusion effect, as had already been applied by [73]. This mechanism is due to the over-sampling of strain vs. rotation regions by inertial particles. Based on an analogy between the two formulations, they were able to apply the same modification to the model by ZA, yielding more accurate results when compared to experimental results for all Stokes numbers, but especially for $St < 0.1$. For this range of values of St the modification is reduced to modifying the value of c_1 in the model of Chun et al. by dividing it by the value $B_{NL} \approx 0.42$. And since in this regime the values for c_1 for both models are extremely close, we can apply the same multiplicative factor to the value obtained from Eq. (2.47), obtaining a corrected approximation for small Stokes numbers. Doing so results in a much more perfect coincidence between the ZA theory and numerical results for c_1 at small values of St (see Fig. 2.5 to appreciate the corrected level of matching).

Still, for larger values of St the full ZA formulation must be solved numerically. This is cumbersome and does not allow us to obtain explicit analytical estimates as we would like. Furthermore, Ireland et al. [178] have shown the estimation of ZA to be accurate only up to $St \approx 0.1$, even when the non-local diffusion correction is applied due, they argue, to its poor prediction of the relative velocity statistics at $St \gtrsim 0.1$.

We have thus taken a different path at $St > 0.1$. We use the fitted expression for Γ given by the former authors in this regime. It is again based on the power-law model of Eq. (2.43), where the c_0 and c_1 are empirically derived. Evidence supporting the power-law scaling of the RDF has been presented for $St \lesssim 1$ [288, 178].

Once again, the coefficients c_0 and c_1 are functions of the St only. This hypothesis can be seen to be reasonable up to $St \sim 1$ and is independent (or only weakly dependent) of the Reynolds number, at least for the range of Reynolds numbers investigated so far [178]. A set of values of c_0 and c_1 , as a function of St where obtained from DNS results and plotted in [178]. We provide a simple fit to this data points from the digitalized images of these plots, see Fig. 2.5. We found that a very convenient model for our fit was given by

$$f(a, b, c; x) = \frac{a}{x} e^{-\frac{\log x - b}{2c^2}} \quad (2.49)$$

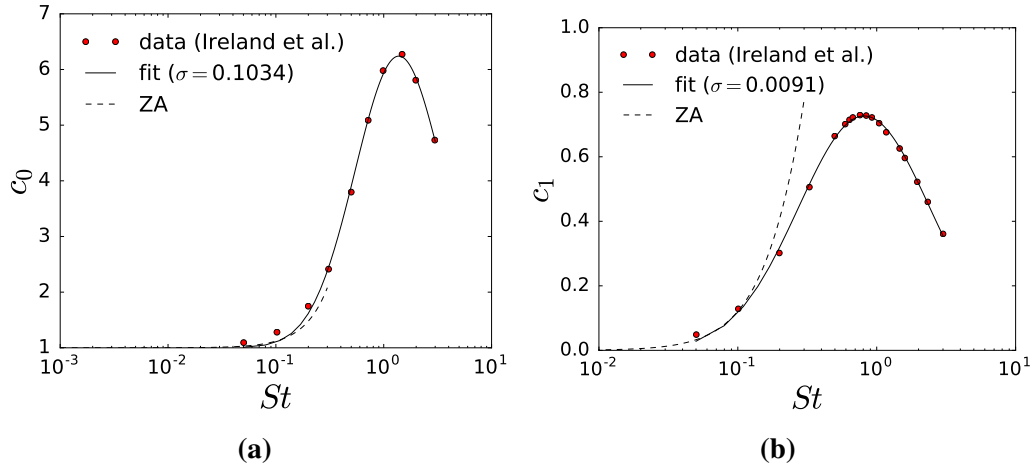


Fig. 2.5 Fits for the c_0 and c_1 as a function of St . The data were extracted from the digitalized Figure 22 in [178]. The continuous curves correspond to the fits in Eq. (2.49). The values of σ in parenthesis show the root mean square error of the approximation.

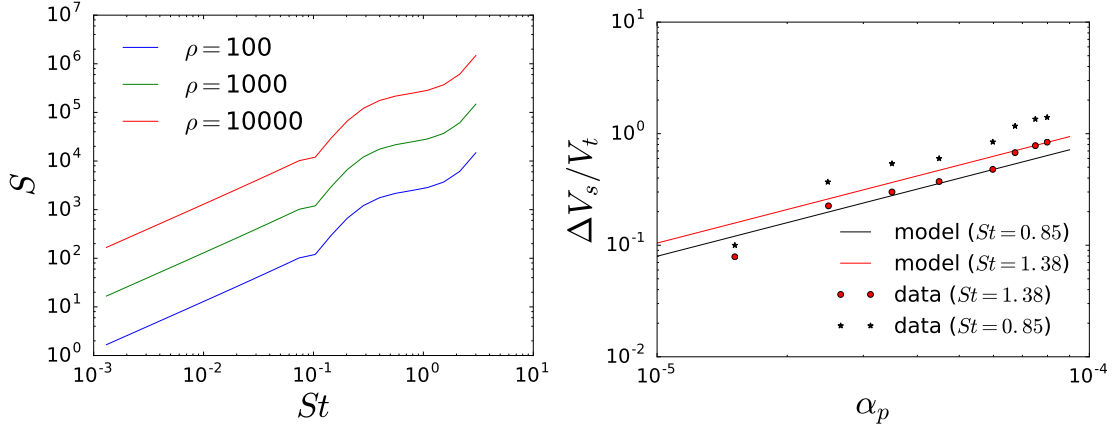
which is proportional to a lognormal distribution probability density function. The optimal parameters c_0 and c_1 were found to be

$$\begin{aligned} c_0(St) &= 1 + f(1.0713, 0.985322, 1.09278; St) \\ c_1(St) &= f(11.1892, 1.19499, 0.93432; St) \end{aligned} \quad (2.50)$$

Now, using Eq. (2.43) and stitching together the corrected version of ZA with the fits above for c_0 and c_1 we are able to produce an explicit analytic formula for the settling coefficient S that can be used to establish the first-effects of the influence of the neighbouring particles in a situation where non-uniform suspensions are to be expected, due to turbulence. This formulation is furthermore limited to heavy particles, but can be used as a conservative estimation for lighter particles while a more complete theory is still unavailable. The condition for neglecting the collective effects of neighbours then becomes, based on Eq.(2.33), $\alpha_p \ll S^{-1}$ or, assuming a maximum value for the acceptable influence at 1% change in the settling velocity, the maximum solid fraction can be expressed as

$$\alpha_p < \frac{0.01}{S(St)} \quad (2.51)$$

where S has the expression of Eq.(2.45) and where the expression for c_0 and c_1 can be taken from Eq. (2.47) (dividing c_1 by 0.42) for small values of St and using Eqs.(2.49) and (2.50)



(a) Settling coefficient as a function of the Stokes number using Eq.(2.45) with Eq. (2.47) for $St < 0.1$ (with the non-local diffusion correction) and Eq. (2.50) for larger values.

(b) Estimation of the settling velocity using the same criterion with the data from [4]; the dots correspond to empirical data from [4], digitalized from Figure 17 in [381].

Fig. 2.6 Effect of the Stokes number on the settling velocity based on the theory of Zaichik and Alipchenkov [381], with corrections from Bragg and Collins [44], Ireland et al. [178].

for larger values. The behaviour of S as a function of St have been plotted in Fig. 2.6a for different values of the relative density.

Note that the curves in Fig. 2.6a show a change in concavity at around $St = 1$ toward a positive one. This is unexpected, as it has been empirically observed that S tends to peak at one, and then monotonically decrease for larger Stokes numbers. This signals a clear flaw of the present model. Nonetheless, we believe the predictions are much more likely to be correct up to $St = 1$, as it is at around this point that the first signs of Reynolds number dependence in the c_0 and c_1 values are observed [44]. Therefore, Fig. 2.6a should be treated with caution and probably only for values of the Stokes number below one. Beyond this point, the value attained at one could be used as a conservative upper bound, as we know the maximum clustering is expected at around this point.

A final remark about this subject is the following. The power-law model for the RDF is supposed to be valid all the way down, below sub-Kolmogorov scales, in accordance with DNS results for monodisperse suspensions. However, such model might suffer from the effects of a certain level of over-idealization. Chun et al. [73] studied this issue theoretically for small-inertia particles and concluded that arbitrarily small differences between the diameters of the particle pairs lead to the appearance of a cut-off distance below which the RDF approaches a constant. This observation goes in the direction of moderating the amount of clustering expected at very small separations in real-world situations and should be kept

in mind when interpreting our results, since our analysis has been limited to monodisperse dispersions.

Effect of collisions

In this work we are mainly interested in CFD-DEM methods, where the MRE is enriched by the addition of contact forces to its RHS. In such a case the effect of collisions is naturally accounted for and the need to estimate the effects of its neglect diminishes. Nonetheless, in practice there are many situations in which it becomes advantageous to turn off inter-particle interactions:

- Removing interactions eliminates the smallest scales in the system (as long as the particle-wall interactions are also eliminated), which leads to the possibility of using much larger time steps
- It can be difficult to accurately model inter-particle interactions, due to uncertainties in the rheology involved
- The number of particles is sometimes made smaller in the simulation compared to the real system to reduce the overall cost, which alters the collision frequency artificially, rendering the effort pointless
- The overall computations become simpler and, thus, cheaper

Therefore, it is interesting to study under what conditions inter-particle interactions can be eliminated without incurring in serious error.

The evaluation of the first effects of collisions must be based on a statistical analysis because any useful characterization of a *typical collision* has to consider both the magnitude and angle of the approaching particles' velocities, which are stochastic variables by virtue of the chaotic nature of turbulent flows. Such characterization is challenging, as it must be derived from local and history-dependent knowledge about the flow and the particles, including (but not limited to) the modelling of clustering that we discussed in the previous section.

The problem of estimating the rate of collision is central to the study of rain formation in clouds and, for a long time, it has been (and continuous to be) a topic of constant developments motivated by this problem and others [288, 123, 70].

For our purposes, though, we can use relatively crude estimates to bound the importance of these effects that can be useful for engineering considerations. In this regard, Loth [216] has established that inter-particle collisions have a negligible effect to the overall trajectories

of the particles if the average inter-collision time is substantially larger than the particle relaxation time. This criterion naturally gives rise to the definition of a *collisional Stokes number* as

$$St_{\text{coll}} := \frac{\tau_p}{\tau_{\text{coll}}} \ll 1 \quad (2.52)$$

where τ_{coll} is the reciprocal of the average collision rate, f_{coll} . The above criterion thus requires an estimation of the collision rate, valid in as wide a range of regimes as possible.

In order to do so, one may use the following formula, that relates the collision rate of two generic neighbours, labelled i, j , with their expected relative velocity:

$$f_{\text{coll}} = A_{\text{coll}} \mathbf{v}_{ij} n_j \quad (2.53)$$

where

$$A_{\text{coll}} = \pi(a_i + a_j)^2 \quad (2.54)$$

where A_{coll} is the swept area formed by the projected silhouette of the two particles onto the plane orthogonal to \mathbf{v}_{ij} , a_i and a_j are the respective radii and n_j is the number density of the species of the particle j . With this formula, the problem reduces to the determination of the expected number density (which can be obtained from the RDF) and the expected relative velocities, where the value of A_{coll} can be determined as a function of these values for spheres.

Loth [216] distinguishes between two sources of collisions, corresponding to different causes for the existence of the relative velocity: those due to turbulent fluctuations of the flow and those due to the different terminal velocities of particles of different terminal velocities under the influence of gravity.

For the case of gravity-driven settling, one can estimate the collision frequency experienced by a particle by calculating the swept volume per unit time of its neighbours as seen from a frame of reference moving with the average speed of the particles ensemble. Assuming monodispersity, the average number of collisions δN in a given time interval δt can be calculated as $\delta N = 4n\pi a^2 v \delta t$, corresponding to a cylindrical volume of base $4\pi a^2$ and height vt , where v is an average relative velocity of a particle with respect to its surrounding ensemble. For small differences in the settling velocity, due to weak polydispersity this yields

$$St_{\text{coll,sett}} = \frac{2}{27} \frac{g}{\nu} \alpha_p (2\rho + 1) (\rho - 1) \varepsilon a^3 \quad (2.55)$$

and the equation is valid to first order in ε , where ε is the range in diameters normalized by the average diameter.

The turbulence-driven collision frequency has been intensely studied [199, 383, 358]. Here we follow Loth [220] who bases the analysis on two limiting expressions are valid for two extreme regimes: the very low and very high Stokes numbers. For very low St Saffman and Turner [301] derived the following expression, which has been validated in numerical simulations in the limit of very small Stokes numbers [331].

$$St_{\text{coll}} \approx 1.24\alpha_p St \quad (2.56)$$

where St represents the usual Stokes number based on the Kolmogorov microscales.

At the opposite extreme, in the so-called *uncorrelated regime*, the following expression can be derived [220]

$$St_{\text{coll, turb}} = \frac{13}{2}\alpha_p \frac{St_{\Lambda\mathcal{L}}}{\sqrt{1 + St_{\Lambda\mathcal{L}}}} \frac{\Lambda_{\perp}}{a} \quad (2.57)$$

where $St_{\Lambda\mathcal{L}}$ is the Stokes number based on the eddy turnover time and where Λ_{\perp} is the lateral integral scale of the turbulent flows [280].

A simple approximation was proposed by Loth [220] to bridge both regimes in between (intermediate regime). This approximation is rough but sufficient to provide a first estimation of the importance of tracking collisions in dilute particle-laden simulations. The result is

$$St_{\text{coll, turb}} = \frac{\alpha_p}{1 + St_{\Lambda\mathcal{L}}} \left(1.2St + \frac{13}{2} \frac{St_{\Lambda\mathcal{L}}^2}{\sqrt{1 + St_{\Lambda\mathcal{L}}}} \frac{\Lambda_{\perp}}{a} \right) \quad (2.58)$$

For more accurate estimates see [331, 383, 71], where more advanced models for the collision frequency in the intermediate regime are discussed.

In order to unify most criteria, we can take, as a first approximation, the minimum of the two definitions Eqs.(2.55) and (2.58), so that the condition to be able to neglect collisions in turbulence becomes

$$St_{\text{coll}} = \min(St_{\text{coll, sett}}, St_{\text{coll, turb}}) \ll 1 \quad (2.59)$$

2.2.5 Small-size effects

In theory, the validity of the MRE should tend to be fully attained in the limit of a vanishingly small particle radius. Indeed, Eqs. (2.7)–(2.9) can all be fulfilled by making a small enough. Small particles are thus the realm of the MRE. However, the implicit assumption of the applicability of the Navier–Stokes equations to the continuous phase in this limit is in fact flawed because the continuum hypothesis fails at sufficiently small sizes. In this section we look to determine the order of magnitude of the lower acceptable limits for the particle size for which the MRE still applies. Since the behaviour of a fluid near the referred limit is

different for gases and liquids, we treat each one separately, keeping in mind mainly water and air as the archetypical examples.

Gas flow

At large scales, the evolution of gases is well described by the Navier–Stokes equations, which are based on the hypotheses of continuum and of pointwise thermodynamic equilibrium (or at least quasi-equilibrium) that lead to the linear stress-strain rate relation [135]. The progressive deterioration of the validity of this assumption as the length scales shrink manifests itself as the so-called *rarefaction effects*. These effects arise due to the growing mean free path (the average distance that the molecules travel between successive impacts [313]) relative to the particle dimensions and the corresponding lower frequency rate of collisions ⁷. The Knudsen number, Kn , specifically quantifies this phenomenon and is defined as

$$Kn = \frac{\lambda}{a} \quad (2.60)$$

with λ the mean free path (MFP) and a the representative physical dimension of the flow, which is usually taken as the particle radius in this context. In reducing the particle radius, the first rarefaction effects to be expected to appear are related to the breakdown of the no-slip boundary condition [193], which is used to derive the MRE. Let us thus investigate the issue a bit further.

Upon impacting on a solid wall, the gas molecules exchange momentum in the tangential direction in other than idealized conditions. If all the reflections were merely specular (symmetric with respect to the surface average normal), there would be no average exchange of tangential momentum and one would have slip boundary conditions. For most materials, however, this is not the case [21], resulting in a net tangential momentum exchange. The main mechanism for such exchange is mediated by the molecular-scale roughness on the wall, that randomizes the reflection angle of the particles, absorbing a part of the statistical bias in the tangential component of their momentum. The exchange is statistically systematic as long as the average relative velocity between gas and wall is nonzero.

As a result of the solid movement, near to the wall, the gas becomes statistically out of equilibrium, because there is a large number of molecules biased toward the wall velocity. This spatial shift in the mean velocity is rapidly homogenized over the neighbouring particles, moving away from the wall. This process is nonlinear (unlike the transmission of the macroscopic momentum, which, at close distances from the wall is linear) and spans a

⁷Such phenomena can arise equally due to a decreasing particle radius, or, as well, due to a growing distance between molecules due to a drop in the average pressure.

few mean free paths. Such out-of-equilibrium region is known as the Knudsen layer [193]. We can conclude that the assumption of no-slip boundary condition at the solid wall must introduce an error of the order of λ , that is, proportional to Kn .

While such an error is the manifestation of the onset of non-continuum effects, the continuum model still has application for a range of Knudsen number values after these effects are first noticeable. This can be achieved through a generalization of the boundary conditions to partial slip (see [390] for a recent review) if Kn is small enough. The set of low Kn conditions for which this strategy is appropriate is called the *slip regime*. If one pretends to extend the applicability of the MRE to this regime, it is evident that one would need to modify the equation to adapt to the replacement of the no-slip condition. Let us look into this.

At the very small particles limit that we are examining, the particle Reynolds number is expected to be extremely low, and so is the Stokes number (unless one is interested in high frequency oscillations, see [79]). This means that the prevailing force is the Stokes drag in this region (see Section 2.3). So to estimate the order of the error introduced by the use of the MRE in this range, we will first look at the drag force corrections for the low-Knudsen number limit (weak non-continuum effects), neglecting for now the other forces.

As previously mentioned, the Navier-Stokes equations must be considered along slip boundary conditions. The simplest and most commonly used model is given by the Navier–Maxwell–Basset formula [390], that assumes the tangential slip velocity at any surface point on the sphere to be proportional to the local tangential stress. That is:

$$\frac{a}{\mu\beta_{\text{slip}}}\tau \quad (2.61)$$

where τ is the local tangential stress and β_{slip} is known as the Basset slip coefficient, in general ranging from zero (perfect slip) to infinity (no-slip conditions). The corresponding drag force on a sphere moving in uniform motion through an otherwise quiescent infinite fluid is (see, e.g. [283])

$$\frac{|\mathbf{F}_{D,1\text{st order}}|}{|\mathbf{F}_{D,\text{Stokes}}|} = \frac{1 + 2\beta_{\text{slip}}^{-1}}{1 + 3\beta_{\text{slip}}^{-1}} \Rightarrow \frac{|\mathbf{F}_{D,\text{Stokes}} - \mathbf{F}_{D,1\text{st order}}|}{|\mathbf{F}_{D,\text{Stokes}}|} = \frac{2\beta_{\text{slip}}^{-1}}{1 + 3\beta_{\text{slip}}^{-1}} \quad (2.62)$$

since β_{slip} is a positive number.

In 1879, Maxwell derived such boundary condition from the kinetic theory of gases under isothermal conditions (we will assume isothermal conditions are a good approximation for the applications we are interested in). Those are given by the following expression, which relates the coefficient β_{slip} with the mean free path from the kinetic theory of monatomic

gases in isothermal conditions.

$$\beta_{\text{slip}}^{-1} = \frac{2 - \sigma_m}{\sigma_m} Kn \quad (2.63)$$

where σ_m is the *tangential momentum accommodation coefficient*⁸. This parameter is defined as the proportion of molecule collisions that result in a diffusive reflection (as opposed to specular) over the total. A molecule is said to undergo diffusive collision when its incidence angle (with respect to the wall's normal) is uncorrelated with the reflected angle, which is randomly distributed. On the contrary, specular reflection means that both incident and reflected angle coincide. The average of the tangential momentum exchange over a number of diffusive collisions equals the average incident momentum, since the average over the reflected momenta has null mean (no-slip). In contrast, the averaged momentum exchange for a number of specular collision has zero mean, and the momentum transfer is null (full slip). The average in reality is somewhere in between these extremes. The value of σ_m depends on the gas and the surface material and finish, but experimental evidence shows σ_m to be between 0.2 and 0.8 (or $\beta_{\text{slip}}^{-1} \in (0.11Kn, 0.67Kn)$), with the former corresponding to specially treated, ultra-smooth surfaces and the latter to most practical surfaces [133]. Maxwell model is derived, by taking the first order (in Kn) effects into account, and thus its validity is bounded below by $Kn < 0.1$; the so-called *slip regime* [193]. Higher order approximations have been derived, but this is unnecessary for our purposes.

Substituting these numbers into Eq.(2.62), we come up with the approximate requirement that $Kn \ll 1$, which imposing our accustomed one percent error bound reads

$$Kn < 1 \times 10^{-2} \quad (2.64)$$

for typical (most) surfaces. Taking the mean free path of atmospheric air as a reference, this means the particles radius should be larger than about 3 μm . Note that this requirement can become more strict for highly polished surfaces, dropping to $Kn \lesssim 1 \times 10^{-3}$ for $\sigma_m = 0.2$, putting the minimum radius at a ten times larger value.

While small-frequency conditions are the most relevant for a majority of applications with low Kn , it is still possible that higher frequencies arise, either due to the flow's own turbulent fluctuations or due to external forces, such as forced vibrations due to the workings of certain machinery etc. While an analysis analogous to that of Coimbra and Rangel is not available yet for small Kn , a recent paper [376] addresses the case of imposed harmonic motion on

⁸Here the Knudsen number is defined based on the mean free path given by $\lambda = \frac{\mu}{p} \sqrt{\frac{\pi RT}{2}}$, with p the gas pressure, R the gas constant and T the temperature; in accordance with the kinetic theory of gases [198]. Note that it is common to encounter variations of Eq. (2.63) based on a different definition of the Knudsen number. Other, frequently used, definitions include $Kn_1 := \sqrt{2/\pi} Kn$ and $Kn_2 := 8/5 Kn$; see [327].

the particle submerged in quiescent gas (rather than the perhaps more representative case of forced motion on the fluid and a free particle treated by the former authors) for small Kn . We consider this case to be representative for unsteady motion in low Kn conditions, at least for order-of-magnitude arguments. In the referred paper, expressions are provided for the asymptotic limits in which we are interested, that is, the limit of $Kn \ll 1$, $\Theta \ll 1$, where Θ is a nondimensional frequency defined as (see [376])

$$\Theta = \frac{\omega}{\nu_m} = Kn^2 \beta \quad (2.65)$$

where ν_m is the molecular collision frequency, ω is the characteristic frequency of the relative motion and β is a Stokes number ⁹, defined as

$$\beta = \frac{a^2 \omega}{2\nu} \quad (2.66)$$

The first of these nondimensional numbers measures how the characteristic length scales of the system compare to the mean free path of the gas, which is related to the statistical convergence of the spatial homogenization, while the second number measures the ratio of the characteristic frequencies to the molecular collision frequency, which is related to the temporal homogenization.

Yap and Sader [376] provide the expression for the force in the slip regime, along with the force in the fully no-slip limit (which can be found, for example, as a solved exercise in [202]). The ratio of the amplitudes of both forces (with the no-slip case in the denominator), which we refer to as α_{Kn} , provides an adequate measure of the distance from the MRE range of applicability, to first order in the Kn . For conciseness, we do not reproduce its complicated expression here, but instead directly plot it in Fig. 3.1. In it, we highlight the one percent and ten percent error curves. Note how the force decreases both as Θ and Kn grow, as the momentum transfer at the particle's surface weakens.

Note that the intersection of the graph with the plane $\beta = 0$ recovers the steady-state asymptotic behaviour discussed above. For the case shown, $\beta_{\text{slip}}^{-1} = 1.016191$, consistent with [376]. Such value comes as a result of refined gas-kinetic simulations, which correct expression Eq. (2.63) and taking $\sigma_m = 1$, see [21]. While the final value is about twice as big as the typical values expected following the above arguments, its order of magnitude is well within the mentioned limits and thus consistent for this matter.

In Fig. 2.7 we have limited the range of Kn and β so that Θ is bounded above by 0.1 since, as mentioned, the validity of this formulation is confined to small values of Kn and

⁹This is a more general notion of the Stokes number, where the motion is not assumed to be determined by the balance between inertia and viscous forces, as we have assumed elsewhere.

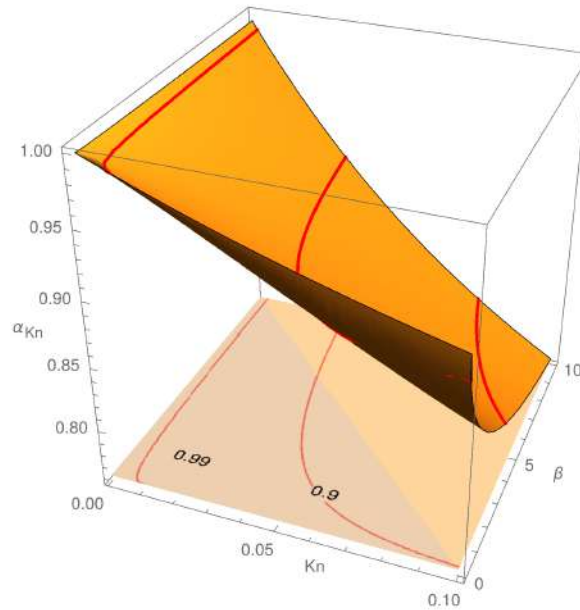


Fig. 2.7 Ratio of force calculated as in [376], valid for small Kn and β , to the magnitude of the force calculated with the MRE, valid under the continuum, no-slip conditions on the surface of the sphere.

$\Theta = Kn^2\beta$. Nonetheless, within these restrictions, we have chosen the range of values for β in Fig. 2.7 to cover all relative motions that could be driven by turbulence within the range of applicability of the MRE, but leaving out extremely large frequencies that would have to be excited by some external force. Indeed, the modified Stokes number β can be related to Eq. (2.4) by using the estimate $\omega \sim 2\pi/T$, with T the characteristic time of the relative motion (which in turbulence-driven flows will be governed by the fluid's small scales motion). This gives

$$St \sim \frac{2\rho + 1}{9\pi}\beta \quad (2.67)$$

The order of magnitude of the coefficient of β can reach values of over 1×10^3 for some materials (e.g. heavy metals suspended in air). Thus, roughly speaking, the graph covers $St \in (0, 10^4)$, well beyond our range of interest (see Table 2.1).

In summary, at the small-size limit that we are analysing it is unlikely to encounter large Stokes numbers due to the tiny inertia of the particles. Nonetheless, one should keep in mind that under unsteady circumstances, the low Kn limitation is intensified, so particular attention must be paid to these situations if the MRE is to be applied.

Liquid flow

The molecules in a liquid are typically much closer than in a gas under the same conditions. Specifically, while the typical mean free path in atmospheric air is in the order of ten molecular lengths, the same distance is only in the order of a single molecule in water [134]. This explains the low compressibility of fluids with respect to gases. The result is that the concept of mean free path, and thus Kn , is less useful for the theoretical characterization of liquids than for that of gases. Furthermore, the hypothesis of low density gas, where the molecular collisions can be regarded as binary and instantaneous, which is at the heart of the classical statistical treatment of gases does not apply. The result is that the molecular theory for liquids is much less developed than that of gases and, for example, a reliable analysis of the breakdown of the no-slip boundary condition is not available yet [134]. Another consequence of the higher density of gases is the change in the order of appearance of small-size effects (see [249]; compare Figures 1 and 4). Indeed, here non-continuum effects come before thermodynamic non-equilibrium effects. That is, statistical fluctuations become significant before the no-slip condition breaks down. These statistical fluctuations are due to a finite number of molecular collisions, leading to an erratic movement of the particles known as Brownian motion. Indeed, the assumption of a valid no-slip boundary condition was employed by Einstein in his famous analysis of Brownian motion. Thus, the first small-size effects in liquid suspensions are discussed in the following subsection, which deals with Brownian motion. The theory is however relevant to both gases and liquids.

Statistical fluctuations: Brownian motion

In order to study the onset of effect of Brownian fluctuations on the movement of the suspended particles, it is necessary to specify the length and time scales of interest. The situation is reminiscent of the study of turbulent dispersion of suspensions, where the averaged dispersion can be studied without having to resolve the smallest scales of the flow. This however means that their effect must be accounted for in an averaged sense. Similarly, Brownian motion presents a wide spectrum of time and length scales. Here too it is possible to avoid resolving all these scales and instead focus on their averaged effect on the particles motion, if the length scales of interest are significantly larger. Moreover, and in contrast with the turbulent dispersion problem, the statistical theory of Brownian motion has been well understood (at least for large times) for over a century, since the works of Einstein, Sutherland and Smoluchowski (see [39], where the relevant citations can be found, as well as a brief historical account of the subject). Let us thus begin by introducing the relevant time scales associated to Brownian motion.

Following Bian et al. [39] ¹⁰

$$\begin{aligned}
 \tau_{c_s} &= \frac{a}{c_s} && \text{(sonic time)} \\
 \tau_p &= \frac{m_{p,f}}{\xi} && \text{(particle's response time)} \\
 \tau_\nu &= \frac{a^2}{\nu} && \text{(viscous diffusion time)} \\
 \tau_D &= \frac{a^2}{D} && \text{(molecular diffusion time)}
 \end{aligned} \tag{2.68}$$

where $m_{p,f} = m_p + \frac{1}{2}m_f$ is the equivalent mass (particle mass plus added mass), c_s is the local sound speed in the fluid, ξ is the inverse of the mobility (that is, the constant quantity in the expression of F_D drag coefficient ¹¹); and D is the Stokes–Einstein–Smoluchowski diffusivity, given by the formula

$$D = \frac{k\theta}{\xi} \tag{2.69}$$

where k is Boltzmann's constant and θ is the absolute temperature.

The meaning of the different time scales is as follows: τ_{c_s} is the time it takes a sound wave to travel through the fluid a distance equivalent to the particle's radius, a ; τ_p is the time it takes for a particle to lose a $1 - 1/e$ fraction of its initial slip velocity, due to the action of the steady drag force and added mass forces alone; τ_ν is the time it takes vorticity to diffuse over a length a from the particle's surface, while the time scale τ_D is the long-term average time it takes for a Brownian particle to diffuse, again, to a distance a (see the Langevin equation Eq. (2.79)). Note that the definitions of τ_p (see beginning of this Section 2.2) and of τ_ν (see Section 2.2.1) had already been introduced. We repeat them here for the reader's convenience.

Fig. 2.8 shows the orders of magnitude of the different time scales for the same material combinations of Table 2.1 as a function of the particle radius. The individual combinations have not been identified for simplicity, but they can be inferred. Note that the values roughly reflect the succession in Eq. (2.68). That is, τ_{c_s} is much smaller than the rest, τ_ν is typically smaller than τ_p for gases, while both are comparable for liquids and τ_D is much larger than the rest for all but the tiniest length-scales, where the continuum hydrodynamic theory has started to break down anyway. Manifestly, such ordering is robust, at least for particles larger than 100 nm.

¹⁰We have slightly altered their notation, in that we have replaced their Brownian relaxation time (τ_B) with τ_p , which includes the added mass effect; and also in calling τ_ν the viscous diffusion time (instead of viscous time) and τ_D the molecular diffusion time (instead of diffusive time).

¹¹Perhaps appropriately modified due to the slip boundary conditions, according to Eq.(2.62)

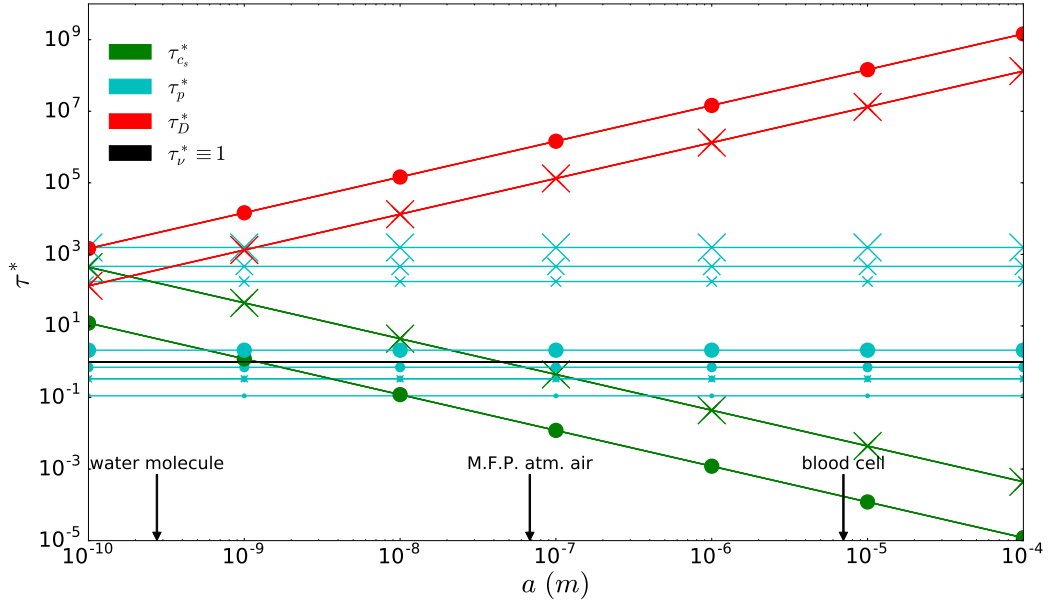


Fig. 2.8 Time scales of Brownian motion normalized by the viscous diffusion time for the same combinations considered in Table 2.1. Dots indicate water as the suspending fluid, while crosses imply air. The size of the markers grows with the density of the particles.

Sticking to the common theme of this section, we will now attempt to link this theory to the MRE. Our goal is to mark the onset of Brownian motion, providing a simple rule to assess whether Brownian effects can be safely neglected. We proceed as before, comparing these effects to that of other forces as the radius of the particle shrinks, since we know Brownian motion is mostly relevant at small sizes. The focus is placed on hydrodynamics, so we should compare Brownian motion to the motion caused by hydrodynamic forces. We would also like to take the buoyancy force into account.

Comparing different hydrodynamic forces is meaningful only when their characteristic times of application are the same, which is equivalent to comparing impulses. In turbulence, the natural reference is the microscopic Kolmogorov time-scale, since it determines the time scale of the relative motion between fluid and particles (when turbulence is driving the relative motion). From hydrodynamic theory, we know that the Reynolds number represents the ratio of inertial versus diffusive momentum transport rates, that is $Re \sim \frac{L^2/\nu}{L/U}$. At the Kolmogorov scales the Reynolds number is by definition of order one, so

$$Re \sim \frac{\eta^2/\nu}{\tau_\eta} \sim 1 \Rightarrow a^2/\nu = \tau_\nu \ll \tau_\eta \quad (2.70)$$

since we have by assumption that $a \ll \eta$. In fact, by the same argument, we have the even stronger

$$\sqrt{\tau_\nu} \ll \sqrt{\tau_\eta} \quad (2.71)$$

This will have immediate consequence in our analysis, as we will see in the following lines.

On the other hand, Brownian motion presents a full spectrum of time and length scales for all time intervals below τ_D , due to its pseudo-fractal nature, down to the ballistic regime¹², where the velocity of the particle becomes well defined as its trajectory starts to smooth out when observed at such small scales. The ballistic regime starts roughly below τ_p [39]. This kind of motion is more appropriately described in terms of statistical averages. Here we consider the mean squared displacement (MSD) or, rather, its square root as a measure of the displacement caused by Brownian motion:

$$\delta(t) := \sqrt{\text{MSD}(t)} \quad (2.72)$$

where

$$\text{MSD}(t) = \langle \|\mathbf{r}(t) - \mathbf{r}(0)\|^2 \rangle \quad (2.73)$$

That is, the ensemble average of the Euclidean norm of the displacement after a given time lapse t . In this way, $\delta(t)$ can be interpreted as the amplitude of the Brownian motion at the time scale t . So that, instead of comparing forces, we can compare the characteristic displacements $\delta(\tau_\eta)$ and $\delta_f(\tau_\eta)$ to each other, where the latter is the characteristic displacement due to the hydrodynamic forces, still to be determined.

We expect the very small particles for which Brownian motion might be relevant to follow the background flow closely, due to their small inertia¹³. So their velocity is expected to be overwhelmingly dominated by this contribution. However, it is the component of the motion *relative* to the fluid motion that determines many interesting phenomena in particulate flows, such as preferential concentration [112] or particle collisions [30], and so we will consider this component. After multiplying it by τ_η , we will obtain an adequate measure of displacement, due to the hydrodynamic forces.

¹²This ballistic regime should not be confused with the 'ballistic regime' that we refer to later, when speaking about the particle's response to the turbulent eddies. The notion is nonetheless somewhat analogous, in that here it refers to the particle's response to the random bombardment to the molecules. Furthermore, the relaxation time of the particle to the random impulses from Brownian motion is defined identically, since the drag law of Stokes is still valid in this regime.

¹³Here we briefly get ahead of ourselves, since this aspect will be treated in more detail in the following section. However, it can also be interpreted as a motivation example for the analysis that is done therein.

Let us consider a turbulent flow. If the flow of interest were laminar, the same reasoning applies by replacing the Kolmogorov microscales with the typical scales of the problem at hand. If the main assumptions of the MRE are met and the particles are very small, so as to make Brownian motion potentially relevant, the scaling analysis of Balachandar [19] applies and the order of magnitude of the slip velocity can be estimated as:

$$W \sim (1 - \beta_\rho) \frac{\tau_p}{\tau_\eta} U = (1 - \beta_\rho) St U \quad (2.74)$$

where

$$\beta_\rho = \frac{3}{2\rho + 1} \quad (2.75)$$

and τ_η is the Kolmogorov time scale, while U is the magnitude of the local fluid velocity. The estimate is derived under the hypothesis that the particle is most of the time close to steady state; that is, that the drag force dwarfs the other terms in the MRE most of the time. Such situation can be expected for very small Stokes numbers (for $St < 1$, according to Balachandar, as we will discuss in the following section¹⁴).

The estimate above does not include the effect of buoyancy (and weight). Balachandar takes this into account separately by keeping track of the point at which the effect of gravity dominates the other forces. That is, when the settling velocity exceeds the slip velocity from equation Eq. (2.74), which happens beyond the point where the Kolmogorov acceleration ($\equiv \epsilon^{3/4}/\nu^{1/4}$, with ϵ the turbulent energy dissipation rate) is smaller than $\|\mathbf{g}\|$. In order to derive our order-of-magnitude relation, it is sufficient to separately consider the steady settling velocity, W_g , which is obtained by balancing the drag force with the submerged weight, giving

$$W_g = (1 - \beta_\rho) St \tau_\eta g \quad (2.76)$$

Once the order of magnitude of the slip velocity has been identified, we can construct the following time-dependent distance scale

$$\delta_f(t) = W t \quad (2.77)$$

which is a measure of the inertial deviation from the perfect tracer particle trajectory of a given particle. Similarly, we define $\delta_g(t) = W_g t$.

Now, the question is, in what circumstances is the effect of Brownian motion in deviating the particles' trajectories from the fluid stream lines of an order comparable to that caused by

¹⁴Note that this argument applies to the forces induced by the turbulent fluctuations alone. The superposed Brownian fluctuations will induce additional hydrodynamic forces that we consider independently, for which the above comments must not hold.

the particles' inertia (or buoyancy force)? We claim that an answer to this question can be given in terms of the quotient $\delta(\tau_\eta)/\delta_f(\tau_\eta)$ (or $\delta(\tau_\eta)/\delta_g(\tau_\eta)$); that is, the ratio of contributions to the deviation from a perfect tracer particle over one Kolmogorov eddy turnover. We propose to consider that the contribution of Brownian motion may be reasonably neglected whenever

$$\frac{\delta(\tau_\eta)}{\delta_f(\tau_\eta)}, \frac{\delta(\tau_\eta)}{\delta_g(\tau_\eta)} \ll 1 \quad (2.78)$$

The Langevin theory [345] provides the right tool to calculate $\delta(t)$. This theory models the movement of the particle as a result of Newton's second law of motion, where the forces are given by a macroscopic contribution or resistance plus a random force originated from the molecular collisions which is modelled as Gaussian white noise ¹⁵. In particular, the often-called *modified* Langevin equation has recently been experimentally verified [172, 160] to be accurate all the way down to the ballistic regime. This version of the equation includes, apart from the steady drag, all the other terms in the MRE. From it, Clercx and Schram [75] have derived the exact expression of the MSD, which we have written as a function of the nondimensional time $\tau^* = t/\tau_\nu$:

$$\text{MSD}(\tau^*) = 2D\tau_\nu\tau^* \left(1 - \sqrt{\frac{4}{\pi}} \tau^{*-1/2} + \left(1 - \frac{\tau_p}{\tau_\nu} + \varepsilon(\tau^*) \right) \tau^{*-1} \right) \quad (2.79)$$

where

$$\varepsilon(\tau^*) = \frac{1}{\sqrt{1 - 4\frac{\tau_p}{\tau_\nu}}} \left(\frac{1}{\alpha_+^3} e^{\alpha_+^2 \tau^*} \text{erfc}(\alpha_+ \sqrt{\tau^*}) - \frac{1}{\alpha_-^3} e^{\alpha_-^2 \tau^*} \text{erfc}(\alpha_- \sqrt{\tau^*}) \right) \quad (2.80)$$

with

$$\alpha_\pm = \frac{1 \pm \sqrt{1 - 4\frac{\tau_p}{\tau_\nu}}}{2\frac{\tau_p}{\tau_\nu}} \quad (2.81)$$

The parameter ε is a small correction, relevant only for $\tau^* \ll 1$ [277]. The function defined in Eq. (2.79) has two extreme regimes; those corresponding to $\tau^* \ll 1$ (ballistic regime), and $\tau^* \gg 1$ (diffusive regime). It is a direct calculation to check that their corresponding asymptotic relations are given by

$$\begin{aligned} \text{MSD}(\tau^*) &\sim \frac{k\theta}{m_{p,f}} \tau_\nu^2 \tau^{*2} & \text{as } \tau^* \rightarrow 0 \\ \text{MSD}(\tau^*) &\sim 2D\tau_\nu\tau^* & \text{as } \tau^* \rightarrow \infty \end{aligned} \quad (2.82)$$

¹⁵The hypothesis of the thermal noise being Gaussian (white noise) is actually only an approximation and more sophisticated coloured versions of the theory have been proposed [37].

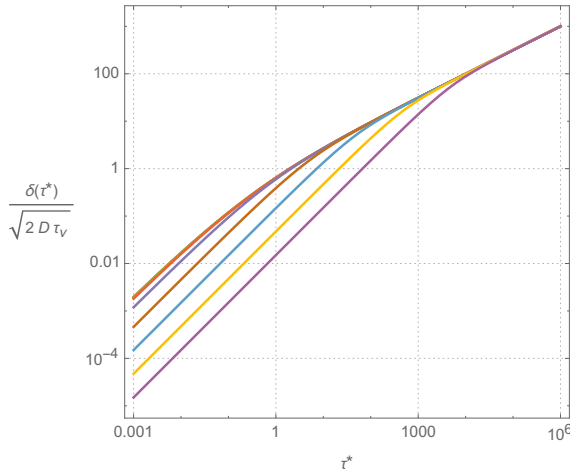


Fig. 2.9 Evolution of δ , normalized by the constant coefficient in Eq. (2.79), as a function of the elapsed nondimensional time for a wide range of relative densities. The two extreme regimes are visible: the ballistic regime corresponds to the 1:1 slope for short times and the diffusive regime to the 2:1 slope for longer times.

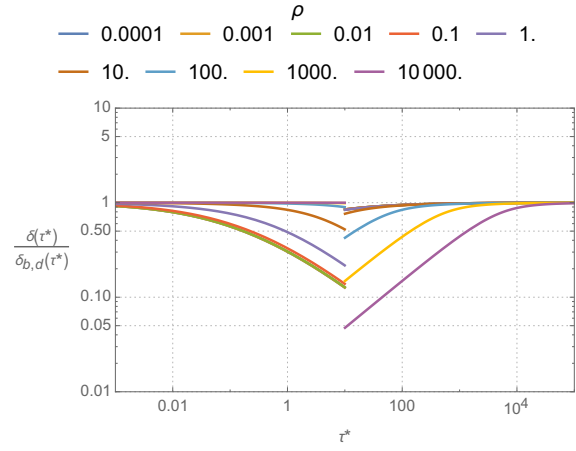


Fig. 2.10 Ratio of δ as calculated by Eq. (2.79) over the value obtained by using its asymptotic approximations. The ballistic approximation is used for $\tau^* < 10$, the diffusive one being used thereon. The error is never more than one order of magnitude, and practically always on the safe side (below one).

Eq. (2.79) can be used to compute $\delta(\tau_\eta)$. However, we would like to provide a simple rule that would serve for most situations. For that, it is enough to realize that the two asymptotic expressions Eq. (2.82) provide sufficient accuracy when used to estimate δ for $\tau^* < 10$ and $\tau^* \geq 10$ as the ballistic and diffusive regimes. By gluing together the two solutions at $\tau^* = 10$ we generate the function $\delta_{b,d}$, that approximates δ . Fig. 2.10 shows that the error made using such expression is at most of one order of magnitude, much less for the less extreme values of ρ . Furthermore, such estimates are on the safe side, in the sense that they are upper bounds for the value of the MSD. An additional observation that follows from Fig. 2.10 and Eq. (2.71) is that the diffusive approximation should suffice for our purposes, since its value becomes of the same order of magnitude as that given by Eq. (2.79) under the hypothesis that the movement of the particle is caused by the turbulent fluctuations of the flow and buoyancy. In these conditions, we are always to the right of $\tau^* = 100$ in Fig. 2.10, and Eq. (2.78) can be expressed as (going back to the dimensional version of the MSD).

$$\begin{aligned} \frac{\delta(\tau_\eta)}{\delta_f(\tau_\eta)} &= \frac{\sqrt{2D\tau_\eta}}{W\tau_\eta} \ll 1 \Leftrightarrow \frac{2D\tau_\eta}{WU|1-\beta_\rho|St\tau_\eta^2} \ll 0.01 \Leftrightarrow \\ &\Leftrightarrow \frac{\frac{D}{\nu}}{\pi\frac{WL}{\nu}|1-\beta_\rho|St} \ll 0.01 \Leftrightarrow \frac{\frac{D}{\nu}}{\pi\frac{W_a}{\nu}|1-\beta_\rho|St} \ll 0.1 \Leftrightarrow \frac{1}{Re_p Sc St} \frac{2\rho+1}{|\rho-1|} \ll 1 \end{aligned} \quad (2.83)$$

where the second order-of-magnitude inequality follows from the first, after substituting Eq. (2.75), using the definition of St and removing the square root; the third inequality follows after considering that an L-sized circular eddy completes a turn in a time τ_η ; the subsequent implication follows from replacing η by a and taking into account the fundamental hypotheses of the MRE theory (i.e. $a \ll \eta$, which we have taken as $a < 0.1\eta$); finally, the last inequality is obtained from Eq. (2.75), the fact that $0.6\pi \approx 1.88 \sim 1$, the definition of Re_p and that of the so-called Schmidt number, $Sc = \nu/D$, which is the ratio of viscous diffusion over molecular diffusion.

Now, it is clear that for the case of neutrally floating particles ($\rho = 1$), the above estimates cannot be used, because the MRE (excluding Faxén terms) predicts the particles to exactly follow the streamlines. The relative importance of Brownian motion for this case becomes unbounded due to the factor involving ρ (see Fig. 2.12) and it should be compared to some other effect, perhaps higher order corrections of the hydrodynamic forces, like the Faxén corrections, or other, relevant forces. We will however leave the discussion at this point.

The settling (or raising) velocity due to the buoyancy force is systematic, just like the long-term effect of Brownian motion, so that any sufficiently long time scale is adequate to compare both effects. It is convenient to use the time it takes for diffusion to cover a distance a . We have seen that for Brownian motion this time is τ_D . Therefore the condition Eq. (2.78) becomes

$$\frac{\delta(\tau_\eta)}{\delta_g(\tau_\eta)} \sim \frac{\delta(\tau_D)}{\delta_g(\tau_D)} = \frac{a}{W_g a^2/D} = \frac{D}{W_g a} = Pe_g^{-1} \ll 1 \Rightarrow Pe_g \gg 1 \quad (2.84)$$

The nondimensional number Pe_g is the Péclet number associated to the settling velocity [35].

In order to put the above estimates in perspective, it is useful to examine Fig. 2.11, that shows the relative importance of the Brownian diffusion in front of the added effects of gravitational settling and inertial drift, for different particle sizes. The different curves correspond to different typical conditions in a variety of environments, where the ranges of parameter values (indicated with error bars) have been determined according to the typical specific energy dissipation rates and the range of values for the density corresponding to the different material combinations from Table 2.1, except the neutrally buoyant cases.

2.2.6 Other effects

The list of effects reviewed in sections 2.2.1 to 2.2.5 cannot be exhaustive. In order to keep the size of our endeavour manageable, we have decided to leave some effects that are

¹⁶Estimated as $\varepsilon = U^3/L$.

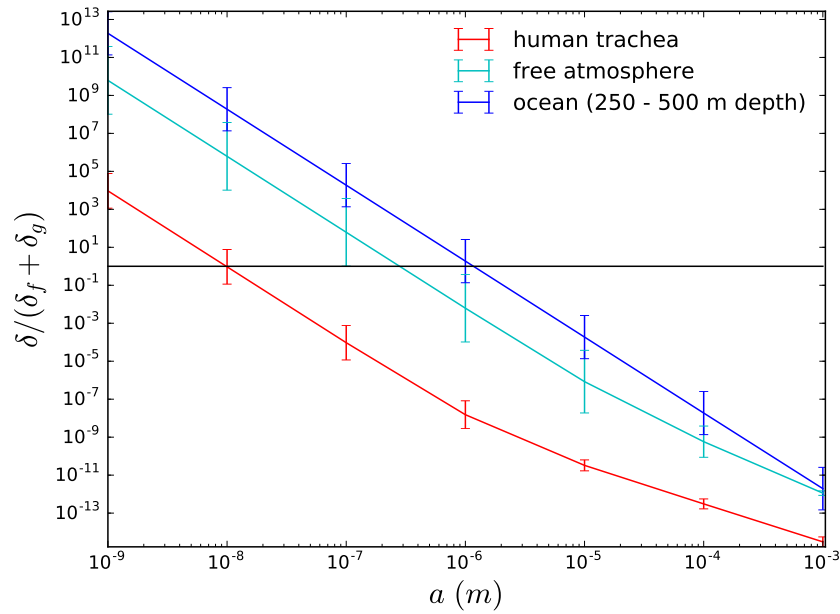


Fig. 2.11 Importance of diffusion relative to gravitational and inertial drift effects for different environments, as a function of the particle size. The curves correspond to the (log)-averaged values corresponding to the extremes of the typical ranges of values found in those environments, including a range of turbulent kinetic energies (human trachea: $3 \times 10^1 \text{ W kg}^{-1}$ to $2 \times 10^3 \text{ W kg}^{-1}$ [115] ¹⁶; free atmosphere: $1 \times 10^{-5} \text{ W kg}^{-1}$ to $1 \times 10^{-1} \text{ W kg}^{-1}$ [190]; ocean, mesopelagic zone: $1 \times 10^{-10} \text{ W kg}^{-1}$ to $1 \times 10^{-8} \text{ W kg}^{-1}$ [228]; and for the material combinations in Table 2.1, except for the neutrally buoyant cases.

indeed important for future work. These effects include the action of common long-range inter-particle forces, such as the van der Waals and electrostatic forces; or the deviations from the Newtonian description of the fluid, for example.

Another notable omission is the analysis of the effect of boundaries in the domain, which has been ignored in this work. Certainly, the MRE is derived under the assumption of an infinite fluid domain. But, in reality, the presence of nearby boundaries often plays a fundamental role in explaining the motion of suspended particles, especially given the slow decay of the particles disturbance flow with distance as discussed in Section 2.2.4. These effects are of especial relevance to internal flows with large specific boundary surface area, such as in transport in tubes [101]; and their understanding is paramount to explain a series of phenomena in microfluidic devices; see [388] for a recent review. The characterization of the first effects of the presence of boundaries on the MRE is therefore of vital importance in order to characterize the applicability of the MRE to many industrial systems and the subject of ongoing research by the author.

With respect to particle collisions, we have commented in Section 2.2.4 that the CFD-DEM method is naturally able to take into account inter-particle collisions. However, this is only true *as long as short-range inter-particle hydrodynamic forces can be neglected*. Indeed, as two small particles approach each other, their Stokesian perturbations start interacting causing the particles to deviate from their non-interacting trajectory. In most cases, these interactions reduce the probability of collision between the particles with respect to their non-interacting trajectories, although they can also increase them under some circumstances [362]. The proportion of particles that collide by taking into account short-range interactions with respect to the number of particles that collide when neglecting them is called collision efficiency [169, 292]. If we can find a theory that characterizes the collision efficiency in a generic turbulent flow, as a function of the flow and particle parameters, we could use it to estimate the first effects of the short-range hydrodynamic interactions on the particles and extend our analysis with them. This task is left for future work.

Another important set of effects relevant to the motion of very small particles are those related to the gradients in temperature, that produce a drift of the microscopic particles towards less energetic zones. A good review on the matter can be found in [249]. The study and characterization of these effects have also been left for future work.

2.3 Scaling analysis

The review in the previous section highlights the breadth of the applicability range of the MRE. Just like other equations that can be applied to a wide range of systems, e.g. the

Navier–Stokes equations, it is interesting to study the possibility of specializing the MRE to more constrained situations. The Stokes equation is an example of such specialization of the Navier–Stokes equations. Its range of applicability is smaller (very small Reynolds number), but its higher simplicity facilitates analysis and very often its numerical solution too, making it preferable under the appropriate conditions. Similarly, the Maxey–Riley equation is often simplified in practice, typically by dropping one or more of its terms.

A necessary condition for being able to do this is that the relative weight of such terms is small compared to that of the ones retained. Scaling analysis permits estimating ranges for the characterizing dimensionless quantities for which such conditions hold. While such analyses always contain a certain (order-of-magnitude) degree of uncertainty, their precision is often sufficient and can later be compared to empirical evidence to reinforce them and refine them if necessary.

In this section we apply an order-of-magnitude analysis to study these questions and quantify the relative importance of the different terms.

2.3.1 Analysis of a simplified MRE

In this section we mainly draw from [79, 19, 211]. We strived to unify the most relevant parts of these works, which had not been done by the most recent author. Our results are very similar but still in some parts different to that of these authors.

Let us consider a simplified version of the MRE in which, for the moment, we neglect the Faxén corrections and the buoyancy terms. For this case Eq. (2.2) reveals that the Stokes number controls, at its limits (i.e. 0 and $+\infty$), the relative importance of the viscous terms with respect to the pressure terms. Specifically, making St large enough (while leaving Re and ρ fixed), for example, by making the size of the particles very big, will make the first two terms negligible and the acceleration of the particle will be completely dictated by the drag and the history terms. Conversely, if St tends to 0 (e.g. by making a small), only the first two terms contribute, while the drag force becomes negligible. Note however that such limit can never be reached without destroying the validity of the equation, since it is valid in the limit of vanishing Re and that for Re to characterize the flow around the particle, it is necessary that $L > a$. Moreover, this form of the equation does not immediately show that the viscous terms vanish for $\rho = 1$, as this case corresponds to a fluid point-particle (that will therefore lead to a vanishing relative velocity with respect to the fluid).

We have seen that by inspection of the nondimensional coefficients of the varying terms in the ODE one can have an idea of the asymptotic behaviour of the ODE. However, it would be naive to directly compare their magnitude in order to assess the relative importance of the different terms. The reason is of course that there is no guarantee that the dependent variables

they multiply are all of a similar order. This is exactly the issue that can be addressed by order-one scaling ($\mathcal{O}(1)$ -scaling) [197]. The method consists in precisely making sure that the dependent variables range the interval between zero and a quantity of order one and thus their coefficients directly reflect their relative weight. The method is iterative, and we will skip the details, although we will include the nondimensionalized equation and justify the order-one boundedness of each term. While the notion of $\mathcal{O}(1)$ is not completely rigorous, it should be understood to mean that the referred value lies *well within an order of magnitude of one*. This notion, while imprecise mathematically, should be useful from an engineering point of view. It should therefore be viewed a method to systematically produce postulates about the size of the error one is making, to be refined by experience and further analysis.

Let us therefore consider an unspecified background flow that nonetheless can be regarded as an oscillatory, periodic flow. A useful mental image could be a sinusoidally varying signal (for each spatial component) or a single eddy inside which the particle is submerged. Let us consider that the period of this flow is given by T and that its velocity's magnitude is bounded by U . We additionally consider that its wavelength is given by L (i.e. the diameter of the eddy). Our goal is to characterize the reaction of the particle to a generic fluid perturbation. We will consider that sufficient time has passed so that all information about the initial conditions (when the particle was first introduced in the flow) has been forgotten. We consider the evolution of the particle during one period, and, thus, the temporal domain is given by $[0, T]$. Let us write down the appropriately simplified version of equation Eq. (2.1):

$$m_p \frac{d\mathbf{v}}{dt} = m_f \frac{D\mathbf{u}}{Dt} + \frac{1}{2} m_f \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right) + 6\pi\mu a(\mathbf{u} - \mathbf{v}) + 6a^2 \sqrt{\pi\rho_f\mu} \frac{d}{dt} \int_{-\infty}^t \frac{1}{\sqrt{t-\tau}} (\mathbf{u} - \mathbf{v}) d\tau \quad (2.85)$$

The lower limit of the integral in the Basset term has been set to minus infinity to account for the fact that initial conditions have been forgotten. We now nondimensionalize the equation by introducing the scalars T , V , W , D_w and $D_{w,1/2}$ such that

$$Tt^* = t; \quad V\mathbf{v}^* = \mathbf{v}; \quad W\mathbf{w}^* = \mathbf{w}; \quad D_w \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right)^* = \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right); \quad (2.86)$$

$$D_{w,1/2} \left(\frac{1}{\sqrt{\pi}} \frac{d}{dt} \int_{-\infty}^t \frac{1}{\sqrt{t-\tau}} (\mathbf{u} - \mathbf{v}) d\tau \right)^* = \frac{1}{\sqrt{\pi}} \frac{d}{dt} \int_{-\infty}^t \frac{1}{\sqrt{t-\tau}} (\mathbf{u} - \mathbf{v}) d\tau$$

where we require that W , D_w and $D_{w,1/2}$ normalize their associated variable, in such a way that its characteristic value is one. Applying these relations to Eq. (2.85), dividing through by the coefficient of $d\mathbf{v}/dt$ and, again for notational simplicity, dropping the asterisks to denote

nondimensional variables we obtain

$$\begin{aligned} \frac{d\mathbf{v}}{dt} = & \frac{1}{\rho} \frac{U}{V} \frac{D\mathbf{u}}{Dt} + \frac{1}{2\rho} \frac{D_w}{V} T \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right) \\ & + \frac{9}{2} \frac{\nu}{\rho a^2} \frac{W}{V} T(\mathbf{u} - \mathbf{v}) + \frac{9}{2} \frac{1}{\rho} \frac{\sqrt{\nu}}{a} \frac{D_{w,1/2}}{V} \int_{-\infty}^t \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{t-\tau}} \frac{d(\mathbf{u} - \mathbf{v})}{dt} d\tau \end{aligned} \quad (2.87)$$

Note that since all the dimensionless variables are now of order one (assuming that none is exactly null), the constant coefficients of the different terms can be interpreted as their characteristic sizes. Since the drag force is (almost) always retained, we will compare the importance of the other effects with respect to it, by evaluating the following quotients.

$$\begin{aligned} \frac{F_U}{F_D} & \sim \frac{2}{9} \frac{a^2}{\nu} \frac{U}{T} \\ \frac{F_A}{F_D} & \sim \frac{1}{9} \frac{a^2}{\nu} \frac{D_w}{W} \\ \frac{F_H}{F_D} & \sim \sqrt{\frac{a^2}{\nu}} \frac{D_{w,1/2}}{W} \end{aligned} \quad (2.88)$$

where we have used the notation $F_D := \|\mathbf{F}_D\|$ and so on for the magnitudes of the different forces.

At this point, it is convenient to distinguish between two regimes: that for which $St < 1$, which we may call the *tracer-particle regime* and that for which $St > 1$, or the *ballistic regime*.

Tracer-particle regime

In the tracer-particle regime the particle relaxation time is smaller than the turnaround time of the smallest eddies in the fluid by definition, and so its velocity has time to become similar to that of the fluid, except perhaps for values of St very close to one. This allows us to consider that $V \sim U$. Evidence supporting this scaling can be found in [392].

We now introduce a simple model to estimate the magnitude of the slip velocity W . Let us consider a particle moving within an eddy of size L . Since, according to Eq. (2.88), in this regime the drag force is never dominated by the history force or the added-mass force, its magnitude must balance that of the rest of terms in the MRE. Approximating the particle velocity to that of the fluid velocity, the drag force must balance inertia and the force due to the unperturbed flow. That is

$$(m_p - m_f) \frac{U^2}{L} = 6\pi a \mu W \Rightarrow W \sim \frac{2}{9} \frac{a^2}{\nu} \frac{U}{T} |1 - \rho| = 2 \frac{|\rho - 1|}{2\rho + 1} St U \quad (2.89)$$

where the LHS is the centripetal force plus the force as would be applied to the displaced volume in an unperturbed fluid (minus its centrifugal force). Note that this agrees with the expression given by Balachandar [19], using a different argument. Furthermore

$$\begin{aligned}
 D_w &= \left| \frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right| = \left| \frac{d\mathbf{w}}{dt} - \mathbf{w} \cdot \nabla \mathbf{u} \right| \sim \left| \frac{d\mathbf{w}}{dt} \right| + |\mathbf{w} \cdot \nabla \mathbf{u}| \\
 &\sim \frac{W}{T} + W \frac{U}{L} = \frac{W}{T} + \frac{W}{T} \sim \frac{W}{T} \\
 D_{w,1/2} &\sim \frac{W}{\sqrt{T}}
 \end{aligned} \tag{2.90}$$

with these estimates we can develop Eq. (2.88) into

$$\begin{aligned}
 \frac{F_U}{F_D} &\sim \frac{1}{|\rho - 1|} \\
 \frac{F_A}{F_D} &\sim \frac{St}{2\rho + 1} \\
 \frac{F_H}{F_D} &\sim 3\sqrt{\frac{St}{2\rho + 1}}
 \end{aligned} \tag{2.91}$$

It is interesting to compare the expression for the importance of the history force to the estimate proposed in [90], where the following estimate is given in the context of direct numerical simulation of turbulence

$$\frac{W_{\text{memory}}}{W_{\text{no memory}}} \approx \frac{1}{1 + \alpha a/L} \tag{2.92}$$

where a good fit to their numerical data was found for $\alpha = 0.69$. It is easy to see that, for the present low inertia case, our estimations are basically equivalent, since (excluding the case of nearly neutrally buoyant particles, which was not addressed in [90])

$$\begin{aligned}
 \frac{W_{\text{memory}}}{W_{\text{no memory}}} &\approx \frac{F_D + F_U + F_A + F_H}{F_D + F_U + F_A} \sim \frac{F_D + F_H}{F_D} \sim \\
 &\sim 1 - 3\sqrt{\frac{St}{2\rho + 1}} \sim 1 - \frac{a}{L} \frac{\sqrt{UL}}{\nu} \sim 1 - \frac{a}{L}
 \end{aligned} \tag{2.93}$$

where we have used the fact that for non-nearly buoyant particles $F_U \lesssim F_D$, that $F_A \ll F_H$ and that the characteristic scales of the flow are given by the Kolmogorov microscales. The negative sign in the RHS indicates that this force tends to diminish the slip velocity. Note

that the RHS of Eq. (2.93) is equivalent to that of Eq. (2.92) to first order in a/L , where the role of α in Eq. (2.92) is played by the order-one quantity $\sqrt{UL}/\nu \sim 1$).

The condition Eq. (2.9), i.e. that the particles must be substantially smaller than the smallest fluid scales, strongly limits the importance of the added mass term and, to a lesser extent, the history term. To see this, let us use Kolmogorov's theory of turbulence, to estimate

$$\frac{a}{L} \sim \frac{a}{\sqrt{\nu T}} \ll 1 \Rightarrow \sqrt{T} \gg \sqrt{\frac{a^2}{\nu}} \quad (2.94)$$

However, if the unsteady force is to be important, the characteristic time has to be small enough. Indeed, \mathbf{F}_A will be important (with the usual one percent relative magnitude threshold) if

$$\frac{1}{9} \frac{a^2}{\nu} \frac{1}{T} > 0.01 \Rightarrow T < \frac{100}{9} \frac{a^2}{\nu} \sim 10 \frac{a^2}{\nu} \quad (2.95)$$

Note that only by relaxing the condition of small particle radius this force can (just about) reach the region in which it is important. For instance, by only requiring that $L > 10a$. But since it has been shown that the validity of the MRE can be extended beyond $a/L \sim 1$ (see Section 2.3.2), it is still of interest to take these forces into account.

Similarly, the analogous condition to Eq. (2.95) for the history force in this regime is

$$T < 10^4 \frac{a^2}{\nu} \quad (2.96)$$

Note that here the history force will be just about important with the strict interpretation ($L > 100a$) of Eq. (2.94) and will have a region of importance spanning two orders of magnitude if the more permissive $L > 10a$ condition is used.

Ballistic regime

For $St > 1$, the particles are too slow to react to the smallest fluid scales and thus here Eq. (2.89) is no longer a good approximation. We use the same reasoning as in [19]. By considering the limit of expression Eq. (2.89) approaches one, one obtains

$$W \sim 2 \frac{|\rho - 1|}{2\rho + 1} U \quad (2.97)$$

Now, considering that this expression is valid down to $St = 1$, we argue that by taking larger values of St one does not increase the slip velocity any further (in order-of-magnitude terms), since the velocities of the particle are already largely uncorrelated. On the other hand, V should not increase to much larger values than U , since its movement is *caused* by the

surrounding fluid in the absence of external forces. Again, we find empirical evidence for this in [392]. The estimates for the derivatives are now

$$\begin{aligned} D_w &\sim \frac{U}{T} \\ D_{w,1/2} &\sim \frac{U}{\sqrt{T}} \end{aligned} \quad (2.98)$$

with these estimates for the slip velocity we obtain

$$\begin{aligned} \frac{F_U}{F_D} &\sim \frac{St}{|\rho - 1|} \\ \frac{F_A}{F_D} &\sim \frac{St}{2|\rho - 1|} \\ \frac{F_H}{F_D} &\sim \frac{3}{2} \frac{\sqrt{(2\rho + 1)}}{|\rho - 1|} \sqrt{St} \end{aligned} \quad (2.99)$$

Note that the relative weight of the history force is now dependent on the Stokes number. For large Stokes numbers, the non-dissipative forces have larger relative weights, diminishing the importance of the history force, that is more heavily penalized for large values of ρ than in Eq. (2.91). This is not captured by Eq. (2.92), leading to an over-prediction of the importance of the history force as St which clearly shows in Figure 4 (b) in [90].

With regards to the limitations imposed by Eq. (2.9), the analogous pairs of conditions to Eq. (2.95) and Eq. (2.96) read

$$\begin{aligned} T &< 4 \frac{a^2}{\nu} \frac{2\rho + 1}{|\rho - 1|} \quad (\text{added mass force}) \\ \sqrt{T} &< 10^4 \sqrt{\frac{a^2}{\nu} \frac{2\rho + 1}{|\rho - 1|}} \quad (\text{history force}) \end{aligned} \quad (2.100)$$

which are essentially the same conditions as in the tracer-particle case with the factor $(2\rho + 1)/|\rho - 1|$, whose behaviour is shown in Fig. 2.12. This quantity is always greater than one and of order one in most of the domain. It is only close to the neutrally buoyant case ($\rho = 1$) that its value explodes asymptotically. But of course in this case the drag force ceases to be a dominant force, since neutrally buoyant particles tend to behave as perfect tracers. In this case one may normalize these forces by F_U instead, which gives

$$\begin{aligned} \frac{F_A}{F_U} &\sim \frac{1}{2} \\ \frac{F_H}{F_U} &\sim \frac{3}{2} \sqrt{\frac{(2\rho + 1)}{St}} \end{aligned} \quad (2.101)$$

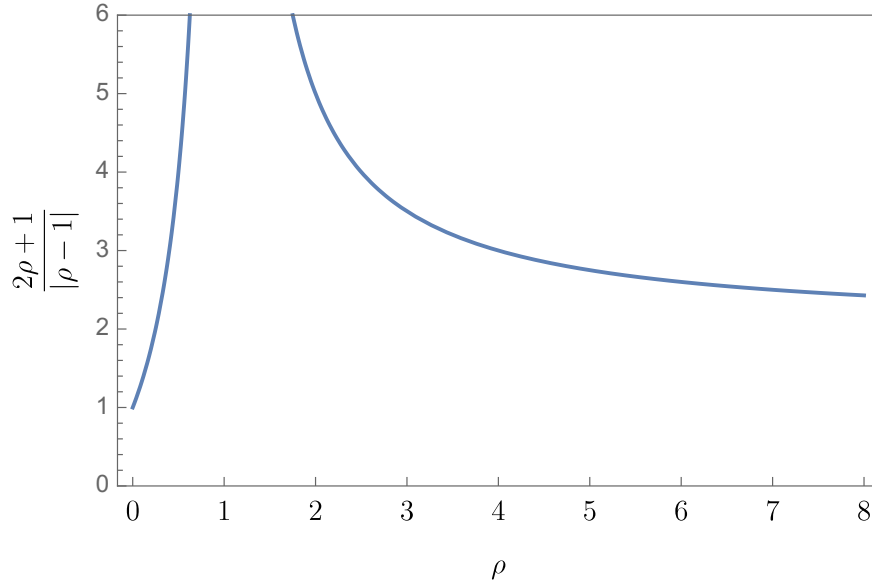


Fig. 2.12 Behaviour of the factor $(2\rho + 1)/|\rho - 1|$.

Nearly neutrally buoyant particles

In this regime, the force that the fluid particles feel is very close to that would be felt by a fluid particle. This means that the most dominant force, by definition, is F_U . The estimates for the slip velocity given for the tracer-particle regime are still valid here, and thus the relative magnitudes between the different forces are preserved. It is nonetheless interesting to present their magnitudes this time normalized by F_U . The results are

$$\begin{aligned}
 \frac{F_D}{F_U} &\sim |\rho - 1| \\
 \frac{F_A}{F_U} &\sim |\rho - 1| \frac{St}{2\rho + 1} \sim \frac{1}{3} |\rho - 1| St \\
 \frac{F_H}{F_U} &\sim 3 |\rho - 1| \sqrt{\frac{St}{2\rho + 1}} \sim |\rho - 1| \sqrt{St}
 \end{aligned} \tag{2.102}$$

where we have used $\rho \approx 1$ and $\sqrt{3} \sim 1$ in the two last estimates. On the other hand, $St > 1$ cannot be reached for $|\rho - 1| < 1$, since this leads to the violation of Eq.(2.94), and the MRE is not applicable.

2.3.2 Importance of the Faxén terms

The Faxén corrections arise as approximations of volume and surface integrals of the fluid field over the particle [238, 222, 53]. These terms are almost universally neglected, although there are very few number of studies to discuss the validity of such assumptions, highlighting their relevance when the particle radius approaches the size of the smallest scales in the fluid [53]. Let us analyse their importance, through scaling arguments.

Whenever the slip velocity is substantial, compared to the velocity of the fluid, their importance is necessarily small within the range of applicability of the MRE, since they are proportional to the radius of the particle squared and thus

$$\frac{\|Ca^2\Delta\mathbf{u}\|}{\|\mathbf{w}\|} \sim Ca^2\frac{U/L^2}{U} \sim C\left(\frac{a}{L}\right)^2 \quad (2.103)$$

which means they are small corrections to the terms they accompany. We thus concentrate on the situations where the slip velocity is expected to be a small fraction of the fluid velocity, such as when St is significantly smaller than one. Then, using the estimate from Eq. (2.89) instead one finds

$$\frac{\|Ca^2\Delta\mathbf{u}\|}{\|\mathbf{w}\|} \sim Ca^2\frac{U/L^2}{\frac{2}{9}\frac{a^2}{\nu}T|1-\rho|} \sim \frac{9}{2}C\frac{\nu T/L^2}{|\rho-1|} \sim \frac{9}{2}C\frac{\nu/UL}{|\rho-1|} \sim \frac{1}{|\rho-1|} \quad (2.104)$$

where we have used the fact that $\frac{9}{2}C \sim 1$ for all the Faxén terms, that $L \sim UT$ and that the Reynolds number of the Kolmogorov scales is of order one. Note that this result is remarkable: in this regime the Faxén terms are important regardless of how small the particle is, compared to the drag force and thus cannot be neglected unless the particle is substantially more dense than the fluid.

However, this estimate cannot be universally valid, since for ρ close to one, the Faxén corrections become dominant in the drag force term and so the assumptions made in the derivation of Eq. (2.89) are incorrect. This means that a different route must be taken to assess the relative importance of these terms. We follow the same procedure as in [171], considering an asymptotic expression with $\varepsilon := a/\lambda$ ¹⁷, where λ represents the Taylor microscale, as the small parameter. The procedure consists in considering the particle velocity as a perturbed version of the fluid velocity, that is

$$\mathbf{v} = \mathbf{u} + \varepsilon^\alpha \mathbf{f}(t) + o(\varepsilon^\alpha) \quad (2.105)$$

¹⁷For the sake of clarity, let us mention that these authors considered 2ε as the small parameter.

where \mathbf{f} is an unknown function and α is an unknown exponent, both to be determined. Note that in the present discussion it is guaranteed that $a/\lambda \ll a/L \ll 1$. Neglecting the added mass and history terms (since those are uniformly smaller than the drag force for small Stokes numbers and $\rho \sim 1$, see Eq. (2.91)) in Eq. (2.1) and dividing through by m_f , one obtains the following equation

$$\frac{d\mathbf{v}}{dt} = m_f \frac{D\mathbf{u}}{Dt} + 6\pi a\mu \left(\mathbf{u} - \mathbf{v} + \frac{1}{6}a^2\nabla^2\mathbf{u} \right) \quad (2.106)$$

Substituting Eq. (2.105) into Eq. (2.106) yields

$$\left(1 - \frac{1}{\rho}\right) \frac{d\mathbf{u}}{dt} + \varepsilon^\alpha \frac{d\mathbf{f}}{dt} = -\frac{1}{\rho} \mathbf{w} \cdot \nabla \mathbf{u} + \frac{9}{2} \frac{1}{\rho} \frac{\nu}{a^2} \left(\frac{1}{6} a^2 \nabla^2 \mathbf{u} - \varepsilon^\alpha \mathbf{f} \right) \quad (2.107)$$

where we have used the relation $\frac{D\mathbf{u}}{Dt} = \frac{d\mathbf{u}}{dt} - \mathbf{w} \cdot \nabla \mathbf{u}$. Let us compare the orders of magnitude of the various terms.

$$\begin{aligned} \left| \left(1 - \frac{1}{\rho}\right) \frac{d\mathbf{u}}{dt} \right| &\sim \left(1 - \frac{1}{\rho}\right) \frac{U}{T} \\ \left| \varepsilon^\alpha \frac{d\mathbf{f}}{dt} \right| &\sim \varepsilon^\alpha \frac{F}{T} \\ \left| \frac{1}{\rho} \mathbf{w} \cdot \nabla \mathbf{u} \right| &\sim \frac{1}{\rho} W \frac{U}{L} \sim \frac{W}{T} \\ \left| \frac{9}{2} \frac{1}{\rho} \frac{\nu}{a^2} \left(\frac{1}{6} a^2 \nabla^2 \mathbf{u} \right) \right| &\sim \frac{3}{4} \frac{\nu U}{L^2} \sim \frac{U}{L} \\ \left| \frac{9}{2} \frac{1}{\rho} \frac{\nu}{a^2} \varepsilon^\alpha \mathbf{f} \right| &\sim \frac{3}{2} \frac{1}{St} \varepsilon^\alpha \frac{F}{T} \sim \frac{1}{St} \varepsilon^\alpha \frac{F}{T} \end{aligned} \quad (2.108)$$

where F represents the characteristic module of \mathbf{f} . Note that the fourth term is of a greater order than the first, due to the assumption of nearly-neutrally buoyant particles, and the third one, for the same reason (reinforced by the assumed smallness of the Stokes number). Furthermore, the fifth term is at least significantly greater than the second one due to the smallness of the Stokes number. This means that the fourth and fifth terms must be of the same order to fulfil Eq. (2.107), which provides an estimate for the first-order correction, since it must hold that

$$\varepsilon^\alpha \mathbf{f} \sim \frac{1}{6} a^2 \nabla^2 \mathbf{u} \sim \frac{1}{6} \left(\frac{a}{L} \right)^2 U \quad (2.109)$$

which means that $\alpha = 2$ and $\mathbf{f} = \frac{\lambda^2}{6} \nabla^2 \mathbf{u}$. Note that this result is not consistent (although of the same order) as the one reported in [171], where the six in the denominator is replaced by a ten. Note that nonetheless, in this work it is stated that “the first-order terms originate from

Faxén corrections to the Stokes drag”, which would be consistent with the one-sixth factor that we have obtained here. In summary, the Faxén corrections become the dominant correction to the particles’ velocities for nearly neutrally-buoyant particles, and this correction is proportional to the particle radius squared. Based on the asymptotic reasoning above Homann and Bec [171] give the following estimate for the averaged effect of this term, based on the assumption of isotropic turbulence

$$\langle \|\mathbf{v}\|^2 \rangle - \langle \|\mathbf{u}\|^2 \rangle \approx -\frac{1}{15} \left(\frac{a}{\lambda} \right)^2 \langle \|\mathbf{u}\|^2 \rangle \quad (2.110)$$

where the $1/15$ would be $1/25$ according to [171]. This implies that this effect is small in the range of validity of the MRE. Indeed these authors study the importance of the Faxén corrections for finite-radii neutrally-buoyant particles and confirm that this effect only becomes significant outside this range. They conclude that tracer-like behaviour is prevalent for radii smaller than $a \approx 1.5L$ and that the effect of the growing radii is well described by the Faxén corrections up until $a \approx 2L$, point at which inertial effects start to become important (note that at this point we are well outside the theoretical range of applicability of the MRE).

2.4 Summary

In this chapter we have applied scaling analysis to the Maxey–Riley equation and to a selection of the hypotheses that support its validity. Our goal was to move closer to a complete characterization of the range of validity of this equation; to be useful as a tool for the assessment of the reliability of the numerical simulations based on it. We are guided by the idea that a physical model equation, such as the MRE, should be tightly attached to its domain of applicability, which is in fact also part of the model. We envision that in the future all the information that is currently dispersedly stored in distributed support (i.e. a combination of human brains and remarks in the literature), will be much more connected and centralized and encoded in a non-ambiguous way. According to such a paradigm, one should be able to program more sophisticated algorithms that automatically analyse the validity of the models used, and that generate estimates of the error according to a variety of measures.

If we model the range of applicability of the MRE as a manifold formed by the combinations of possible values that the relevant non-dimensional quantities may attain (as a function of the desired level of accuracy), our objective is to characterize this manifold by identifying parts of its unknown boundary.

Our analysis has produced a number of conditions on a set of dimensionless quantities that characterize the problem of the motion of a single particle in a Newtonian fluid. A first

subset of these parameters resulted from the analysis of the dimensionless parameters that parametrize the equation itself. These are ρ , Re_p and St (Fr was not explored). A second set of parameters were taken into account explicitly in the derivation of the MRE [238], where it is required that they are small. This is the case of a/L . The rest of the parameters must be introduced from outside the theoretical background of the MRE to analyse its range of validity. Such are Re_s , Re_Ω , related to the inertial effects; ε_p , related to the nonsphericity; and Kn , Pe and Sc related to the limits of the continuum description of the flow, for very small particles. Finally, we also tackled the problem of characterizing the first effects of the presence of neighbouring particles in the flow. Within our limited theoretical framework (uniform suspensions or infinite, isotropic turbulence with heavy particles), the relevant nondimensional quantities have been identified to be α_p and St_{coll} . The main results are summarized in Tables 2.2 to 2.5.

Table 2.2 Summary of estimates for the limits of the applicability range of the MRE except neighbour-related effects. The function E represents the relative error of the function in parenthesis due to neglecting the corresponding effect.

Dimensionless parameter	Applicability condition	Error measure and growth rate	Associated effect
Re_p	$Re_p \ll 1$	$E(F_D) \sim 3/8 Re_p$	Onset of inertial effects on drag force
Re_s	$Re_p \ll 1$	$F_{L,\text{Saff}}/F_D \sim 1/3 Re_s$	Lift force due to shear
Re_Ω	$Re_\Omega \ll 1$	$F_{L,\text{Rub}}/F_D \sim 1/6 Re_\Omega$	Lift due to the Magnus effect
ε_p	$\varepsilon_p \ll 1$	$E(F_D) \leq \varepsilon_p^{18}$	Effect of nonsphericity
Kn	$Kn \ll 1$	$E(F_D) \sim \frac{\beta_{\text{slip}}^{-1}}{1+3\beta_{\text{slip}}^{-1}}$	Breakdown of no-slip B.C. for gases
Sc	$Sc \gg \frac{1}{Re_p St} \frac{2\rho+1}{ \rho-1 }$	$\frac{\delta(T)}{\delta_f(T)} \sim \frac{1}{Re_p Sc St} \frac{2\rho+1}{ \rho-1 }$	Brownian effects vs inertial drift for liquids
Pe_g	$Pe_g \gg 1$	$\delta(\tau_\eta)/\delta_g(\tau_\eta) \sim Pe_g^{-1}$	Brownian effects vs gravitational settling

¹⁸The rate depends on the particular shape of the particle; see Eq.(2.29).

Table 2.3 Summary of estimates for the limits of the applicability range of the MRE regarding the effect of neighbours. The function E represents the relative error of the function in parenthesis due to neglecting the corresponding effect.

Dimensionless parameter	Applicability condition	Error measure and growth rate	Associated effect
α_p	$\alpha_p \ll 1/S$	see Eq.(2.39) and equations leading to Fig. 2.6a	Onset of collective effects in settling
St_{coll}	$St_{\text{coll}} \ll 1$	see Eqs.(2.55) and (2.59)	Collisions become important in the macroscopic flow

Table 2.4 Importance of the different terms in the MRE. The estimates are based on the assumption that the smallest scales in the flow are the principal action driving the motion of the particles.

	F_U	F_A	F_D	F_H
Tracer-particle regime	$\frac{1}{ \rho-1 }$	$\frac{St}{2\rho+1}$	1	$3\sqrt{\frac{St}{2\rho+1}}$
Ballistic regime	$\frac{St}{ \rho-1 }$	$\frac{St}{2 \rho-1 }$	1	$\frac{3}{2}\frac{\sqrt{(2\rho+1)}}{ \rho-1 }\sqrt{St}$
Nearly-neutrally buoyant particles	1	$\frac{1}{3} \rho-1 St$	$ \rho-1 $	$ \rho-1 \sqrt{St}$

Table 2.5 Importance of the Faxén corrections. The function E represents the relative error of the function in parenthesis due to neglecting the corresponding effect.

Tracer-particle regime	Ballistic regime (non-neutrally buoyant)	Nearly-neutrally buoyant particles
$E(F_D) \sim \frac{1}{ \rho-1 }$	$E(F_D) \sim \frac{1}{6}\left(\frac{a}{L}\right)^2$	$E(\langle\ \mathbf{v}\ ^2\rangle) \sim \frac{1}{15}\left(\frac{a}{\lambda}\right)^2$

Chapter 3

The numerical solution of the Maxey–Riley equation

3.1 Introduction

Particle-laden flow simulations typically involve a multitude of particles, even for highly disperse regimes. And while it is sometimes possible to form representative ensembles by instead considering only a small fraction of the actual particles [318], their total number must still be large enough to ensure statistical significance, often a demanding requirement. For methods that track the evolution of each individual particle, as in Euler-Lagrange methods, this translates into solving the single-particle equation of motion (i.e., the MRE) a correspondingly large number of times. As mentioned earlier, the associated computational costs can quickly become cumbersome and thus efficiency becomes of great concern.

Without the history term, the MRE is an ordinary differential equation, whose solution is well known to be regular for a sufficiently smooth background fields [124]. Its numerical solution can be achieved by standard integration schemes like the Runge–Kutta methods [215] or the Adams-Bashforth schemes [330] in a reasonably efficient way. Both explicit and implicit methods can be used. Explicit schemes are conditionally stable and require small enough time steps. This requirement can become quite severe, especially when unsteady effects are included (added mass and history force). This is why many implementations use implicit strategies instead [221].

In the context of (soft-sphere) discrete element methods, explicit schemes are nonetheless the preferred option in the vast majority of implementations. This is because explicit algorithms are already the most common choice for regular, ‘dry’ DEM algorithms. In such case, the contact forces enter the RHS of the MRE simply as additional external forces. These

forces have a very small characteristic duration and often pose an even stricter bound on the maximum allowable time step than any of the hydrodynamic forces.

Other, hybrid options are also possible. For instance, implicit scheme for the MRE can be combined with an explicit treatment of the contact forces to improve stability. Fully implicit schemes, treating both the contact and the fluid interaction implicitly are rare. For a comparative review of various integration schemes for fluid-coupled DEM, see [348].

In summary, whether explicit or implicit, relatively standard and efficient methods have been shown to work with the MRE without the history force term.

On the other hand, the presence of the history term turns the MRE into an integro-differential equation (in fact, to a fractional differential equation, see Section 3.2.1) with qualitatively different mathematical properties (in particular, the dependence of the force on the whole trajectory of the particle)¹. Furthermore, its presence complicates the numerical solution of the MRE, due to the necessity to calculate an (improper) integral over the whole past history of the particle, at every time step. Straightforward extensions of standard methods either fail or become extremely inefficient here.

Perhaps for this reason, the history term has been routinely neglected (see, for example, [212, 286, 241, 235]), although clearly not always on a solid physical basis. For instance, for the numerical study of bubbly turbulent water flow, its effect was neglected a priori in [242], without explicitly commenting on it. However, the range of parameters in the examples examined therein do not justify this, according to Table 2.4 and values $\rho \approx 1$ and $St > 0.1$ that used in that work. Notwithstanding this, awareness of this issue is growing within the scientific community as a number of (mostly numerical) studies provide new evidence of the importance of including this force in a variety of contexts [155, 266, 90, 254].

This chapter is concerned with the discretization and numerical solution of the MRE. We focus on the numerical treatment of the history force term, reviewing the state of the art and making several contributions to the subject. We present a complete algorithm to take advantage of the previous developments and we test its accuracy and efficiency in a sequence of tests of increasing level of complexity. We conclude that the proposed algorithm permits the inclusion of the history force in large-scale particle-laden flow simulations at an acceptable cost.

¹It is perhaps appropriate to mention at this point the recent proof, given by Farazmand and Haller [124], of local existence, uniqueness and regularity of (weak) solutions of the MRE (including the history term) under quite general conditions. In particular, the proof requires the fluid velocity field to be at least four times continuously differentiable in space and time over the whole trajectory of the particle and that its partial derivatives (including mixed partials) are Lipschitz continuous and uniformly bounded up to order three.

3.2 Overview of approaches for the treatment of the Boussinesq–Basset term

In this section, we briefly review the state of the art in the numerical treatment of the history term, introducing the challenges it poses and how researchers currently go about overcoming them. For more information, see the recent review by Moreno-Casas and Bombardelli [254]. Our review will also serve to introduce the methods and notations that are used in this work to treat the history term, some of which have been modified or further developed. This is the subject of the subsequent sections.

The Boussinesq–Basset force term has the form of the derivative of an integral, as presented in Eq. (2.1). In considering the finite-difference discretization of the MRE, the derivative becomes a finite difference and thus one is faced with the question of how to evaluate the integral at a given discrete time. Let us start by recalling its exact form (ignoring the coefficient C_B):

$$\int_{t_0}^t \frac{1}{\sqrt{t-\tau}} \left(\mathbf{u} - \mathbf{v} + \frac{1}{6} a^2 \nabla^2 \mathbf{u} \right) d\tau \quad (3.1)$$

which can be rewritten as

$$\mathbf{I}_B(t) := \int_{t_0}^t K_B(t-\tau) \mathbf{f}(\tau) d\tau \quad (3.2)$$

where $K_B(s) := 1/\sqrt{s}$ defines the Boussinesq–Basset kernel function and $\mathbf{f}(s)$ a differentiable vector field (assuming the unknowns to be smooth enough). Thus, the integral above can be understood as the convolution of \mathbf{f} with the kernel K_B . Note that, with this notation, we have $\mathbf{F}_H = C_B d\mathbf{I}_B(t)/dt$.

It is also possible to consider a different form of the Boussinesq–Basset integral, where the derivative moves under the integral sign; see Appendix C. Specifically

$$\frac{d}{dt} \mathbf{I}_B(t) = \int_{t_0}^t K_B(t-\tau) \frac{d\mathbf{f}(\tau)}{d\tau} d\tau + K_B(t-t_0) \mathbf{f}(t_0) \quad \forall t > t_0 \quad (3.3)$$

where the alternative form is on the right-hand side. The latter was precisely the one used in [353], though its second term vanished because the initial time was taken as $t_0 = -\infty$. Nonetheless, the same method can be applied to both situations; it is only necessary to correctly identify the field that must be considered: either \mathbf{f} , or its derivative ². As pointed

²Such derivative would involve a finite difference representation of the unknowns and the fluid field values under the integral sign, leading to a stencil involving the linear combination of quadratures at different points in time.

out in [89], the first form is very convenient because it eliminates the need to compute either any additional derivatives or the second term in Eq. (3.3). It is for this reason that we have also adopted this form in this work.

A naive approach would be to use a classical quadrature scheme (e.g. Newton-Cotes formula) to express the integral in terms of a finite number of discrete values of the unknowns. An initial difficulty is encountered when trying to evaluate the integrand at the current time, since the integrand is actually not defined at that point. Still, semi-open interval quadrature formulae can be found to overcome this difficulty. Daitche [89] refers to the semi-open Newton–Cotes scheme [129] in this regard. Two alternatives are mentioned in [40]: the Euler-Maclaurin formula [129] and a method that had been apparently proposed by L.M. Brush et al. in 1964. The latter consists in applying the following approximation

$$\int_{t_0}^t \frac{1}{\sqrt{t-\tau}} f(\tau) d\tau \approx \int_{t_0}^{t-h} \frac{1}{\sqrt{t-\tau}} f(\tau) d\tau + 2\sqrt{h} f(\xi) \quad (3.4)$$

where ξ is evaluated in some point in $(t-h, t]$ and a standard quadrature rule is then to be applied to the (now proper) integral on the RHS. The first two methods were shown to yield errors that scale like $h^{1/2}$. In Section 3.5.1 we implement Brush's method, which also shows the same order of accuracy. Such low-order methods lead to excessively small time steps and, we will be interested in finding higher-order quadrature schemes to replace them. First, however, we introduce the concept of a fractional derivative, a generalized definition that allows us to define the history term in an elegant way.

3.2.1 Fractional derivative approach

The theory of fractional derivatives generalizes the classical derivatives and anti-derivatives (indefinite integration) of integer order, by letting this order take arbitrary real values. There exist a rather large number of alternative ways in which this generalization has been realized, leading to a profusion of derivative definitions that has recently been summarized in an extensive survey by de Oliveira and Tenreiro Machado [94]. Letting f be a function defined in an interval $\Omega = (t_0, t_1) \subset \mathbb{R}$, its Riemann-Liouville left-sided derivative is defined, for all $t \in \Omega$, as

$${}^{\text{RL}}D_{t_0^+}^\alpha f(t) := \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_{t_0}^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau \quad (3.5)$$

where $\alpha \in (0, \infty)$ represents the derivative order and Γ is Euler's Gamma function and the subindex t_0^+ indicates that the derivative is defined for $t > t_0$ (Sometimes the alternative notation ${}^{\text{RL}}D_t^\alpha$ is used).

A different definition for the fractional derivative is given by the Grünwald–Letnikov derivative. This derivative is obtained by generalizing the usual finite difference formulas and taking the limit to vanishingly small step size. The definition (left-sided version) reads

$${}^{\text{GL}}D_{t_0^+}^\alpha f(t) := \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{k=0}^{\lfloor n \rfloor} (-1)^k \frac{\Gamma(\alpha + 1)}{\Gamma(k + 1)\Gamma(\alpha - k + 1)} f(t - kh) \quad (3.6)$$

where $nh = t - t_0$.

The following lemma relates both definitions [279, p. 76]:

Lemma 1 *Let f be continuous and its first order derivative be integrable in $[t_0, t]$. Then, for every $0 < \alpha < 1$ both the Riemann–Liouville and the Grünwald–Letnikov derivatives of f exist and are equal.*

The observation that the Boussinesq–Basset force can be understood as a fractional time derivative (times a constant), was apparently first made by Tatom [338]. Specifically, this term is proportional to the Riemann–Liouville fractional derivative of order $\alpha = 1/2$ of the slip velocity (plus the Faxén corrections); that is

$$\mathbf{F}_H = \sqrt{\pi} C_B D_{t_0^+}^{1/2} \left(\mathbf{u} - \mathbf{v} + \frac{1}{6} a^2 \nabla^2 \mathbf{u} \right) \quad (3.7)$$

since $\Gamma(1/2) = \sqrt{\pi}$. where the derivative operates component-wise.

The Grünwald–Letnikov representation of the derivative suggests an approximation method to compute ${}^{\text{GL}}D_{t_0^+}^{1/2} f(t)$: pick a small h and compute the resulting finite sum. And, indeed, under the assumptions of the above lemma, this approximation may also be used to calculate the Boussinesq–Basset force. This possibility was, to the best of our knowledge, first explored by Bombardelli et al. [40], see also [148] (although Tatom already acknowledges the possibility of using the Grünwald–Letnikov representation for numerical computation). These authors tested the summation method (hereafter referred to as the GL method) against a standard method (i.e. the Euler–Mclaurin summation formula, mentioned in the introduction of this section) to deal with the kernel singularity at the time. They found that the method was more efficient, achieving first order scaling of the error, as opposed to the order one half achieved with the standard method. It was not long, however, until new techniques appeared ([353], followed by [89]), achieving second and higher accuracies and somewhat overshadowing this more ‘exotic’ method. In [353], the first order accuracy obtained by Bombardelli is referred as a motivation to look for alternative, higher order methods. Furthermore, in the recent review of Moreno-Casas and Bombardelli [254], it is concluded:

For the accuracy in the computation of the standard term, the intrinsic difference between all three methods is the order of accuracy of each approximation. Whereas the approach of Bombardelli et al. [40] leads to a first-order solution in the computation, the method by van Hinsberg et al. [353] leads to a second-order solution, and the Daitche approach [89] sets grounds to obtain higher order solutions (second-order and higher).

While their remark captures the behaviour of the algorithms as were presented in the cited works, we think it is excessive to talk about an ‘intrinsic’ difference between the methods. In Section 3.5.1 we will argue that Bombardelli’s approach can be modified to achieve higher order accuracies. Similarly, we will view the method of van Hinsberg et al., not as an alternative to the other methods, but actually as an *evolved version* of them. This point is addressed in the Section 3.3.

3.2.2 Hybrid polynomial interpolation/analytic approach

The idea of using a polynomial interpolation to approximate only the nonsingular part of the integrand (f in Eq. (3.2)) was introduced by van Hinsberg et al. [353], who used linear polynomials and obtained a first-order accurate method. That is, they showed the quadrature error to scale with h^2 ; h being the distance between successive quadrature points, assumed to be equal to the time step. This performance beat standard methods (for example, the second order Euler-McLaurin formula, of order one half) but also the fractional derivative approach used by Bombardelli (first order accurate). The method consists of the following steps:

1. Subdivide the integration domain, $[t_0, t]$ in n subintervals $[t_0, t_0 + h], \dots, [t_{n-1}, t_n]$, with $t = t_n$. The boundaries of these intervals are all the scheme’s intermediate time integration points up to the present time, t .
2. Interpolate f in the interior of each interval with a linear polynomial.
3. Replace f by its interpolated approximation back into the integral and find its primitive (it has a simple analytical formula).
4. Reorder the resulting expression as a weighed sum of the values $f_0 := f(t_0), \dots, f_n := f(t_n)$. The result is the discretized version of the Boussinesq–Basset history force, F_H
5. Plug the expression into a suitable time integration scheme. The resulting algorithm expresses the force at every time step as a function of the unknowns at all the past (and current) time steps.

The method was recently extended (and slightly generalized to allow nonzero initial relative velocities) to third and fourth order by Daitche [89]. We describe his approach in Section 3.3, as it is part of our proposed algorithm.

3.2.3 Comparing the accuracies of the different quadrature methods

Let us take look at the performance of the different quadrature methods in approximating an integral with known analytical closed form. As a test function we choose $f(t) = \sin(t)$. We are looking to approximate the value

$$\int_0^t K_B(t - \tau) \sin(\tau) d\tau = \int_0^t \frac{\sin(\tau)}{\sqrt{t - \tau}} d\tau = \frac{1}{2}\pi\sqrt{t} (J_{1/2}(t) - J_{-1/2}(t)) := I(t) \quad (3.8)$$

where the J_ν represent Anger functions [267, p. 295]. Fig. 3.1 shows the performance of different methods for different values of the quadrature time-step, h . The sequence of h is formed by doubling each consecutive value. The error ε is defined as

$$\varepsilon = \max_{t \in T_h} |I(t) - I_h(t)| \quad (3.9)$$

where I_h corresponds to the numerical approximation for the quadrature time step h and T_h has been chosen to be a subset of all the discrete times in which the interval $[0, 10]$ is subdivided. Specifically, T contains a number of equidistant points equal to the minimum between 32 and the total number of points in each case (so as to save computational time). The point $t = 10$ was included in all cases. Numerical experiments show that the number of points considered does not significantly modify the plot. Therefore in subsequent experiments only eight time points will be considered.

Fig. 3.1 is very similar to the one presented in [89] for the same test. the main difference is the inclusion of the fractional derivative method, of order 1. As expected, the numerical behaviour of the error scales as h^1 . The worst performance corresponds to the Newton–Cotes method (trapezoidal rule), with the modification proposed by Brush that we described at the end of Section 3.2.

3.2.4 Addressing memory requirements: window methods

Even though the methods discussed above achieve considerable improvements in accuracy (therefore allowing for larger time steps) with respect to standard techniques, their memory requirements are still large. Especially for systems with many particles, the necessity of keeping track of the complete history of each of them is highly demanding. Let us consider

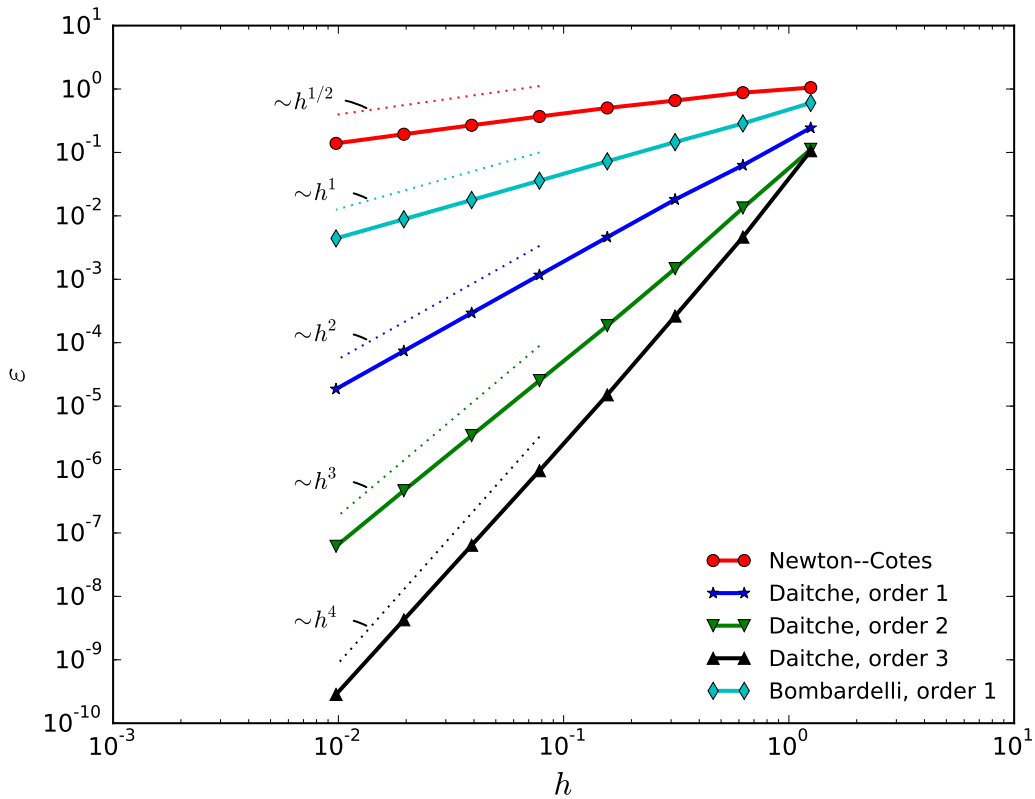


Fig. 3.1 Scaling of the global error of the different methods for the test-case $f(t) = \sin t$

a system with 10^5 particles. If we assume a time step of 0.01s (i.e., value chosen for the accuracy tests in [89]) and a simulation duration of 10s, the necessity of storing one vector per time step and particle leads to a total 10^8 vectors or, roughly, about 1 GB in memory (assuming a vector takes up 3×8 bytes) by the end of the simulation!. Furthermore, the data contained in these vectors must be continually accessed, inside the innermost loop of the algorithm, for each particle, at every time step. Obviously, this tends to dramatically slow down, making it clear why this force has so often been neglected in the past. Window methods deal with this problem directly.

Window methods make use of the decaying influence that the past has on the advancing present value of the history force to avoid having to store the complete history of the particles. The simplest approach consists in neglecting the contribution of the particle's history that is older than some cut-off time $t - t_w$, where t is the current time. One refers to t_w as the *window time*, because only the history within the 'window' $[t - t_w, t]$ is taken into account.

That is, one considers

$$\int_{t_0}^t K_B(t - \tau) \mathbf{f}(\tau) d\tau \approx \int_{t-t_w}^t K_B(t - \tau) \mathbf{f}(\tau) d\tau \quad (3.10)$$

which is equivalent to

$$\int_{t_0}^{t-t_w} K_B(t - \tau) \mathbf{f}(\tau) d\tau \approx 0 \quad (3.11)$$

The integral on the right-hand side of Eq. (3.10) is termed the *window contribution*, while the integral on the left-hand side of Eq. (3.11) is the *tail contribution*. The idea is to employ some method like, for example, any of the ones mentioned above to calculate the window contribution and to simply neglect the rest, thereby saving all the associated memory space (and avoiding the corresponding computations). As such, this method corresponds to what is known as the window method in literature, following its introduction by Dorgan and Loth [108]. These researchers applied it to the case of finite particle Reynolds numbers, taking advantage of the faster decaying nature of the history force kernel in this case ³, see for instance [223]. Here we apply the term *window method* generically to any method that records only a limited part of the particles' most recent history. The main weakness of the original window method is its poor accuracy, precisely due to the slow decay of the convolution kernel in time (as the inverse of the square root for the Basset kernel); because one is thus forced to increase t_w significantly to keep the tail truncation error low, at the cost of sacrificing potential memory savings.

A much more accurate window method was proposed by van Hinsberg et al. [353] (henceforth referred to as MAE for *Method of Approximation by Exponentials*). In it, the tail contribution is not neglected but rather *approximated*. Specifically, while in the window region the kernel is kept exactly equal to the Basset kernel, in the tail region it is only approximated by a different, special kernel. The advantage of this special kernel is that it leads to a recursive expression of the evolution of the Basset integral in time. That is, the evolution of the integral from one time step to the next can be expressed in terms of its value in the previous time step plus a contribution that depends only on the recent history. Despite its relatively recent development, the method has already found application in the study of particle-laden turbulence; see [72], [188] and [352].

³The exact formulation includes an algorithm to approximately determine t_w so as to minimize the truncation error. But since we would like to have t_w as small as possible, we cannot in general apply it without running into a conflict

3.3 Improvements on the MAE

Let us review the MAE in sufficient detail. The goal is to define an algorithm to compute the integral in Eq. (3.2) approximately. We begin by replacing the Boussinesq–Basset Kernel $K_B : (0, t - t_0] \rightarrow \mathbb{R}$ by the modified kernel $K : (0, t - t_0] \rightarrow \mathbb{R}$, defined by

$$K(\tau) = \begin{cases} K_B(\tau) & \text{if } \tau \leq t_w \\ K_T(\tau) & \text{if } \tau > t_w \end{cases} \quad (3.12)$$

where $K_B(\tau) = 1/\sqrt{\tau}$ as before and the tail kernel is defined by

$$K_T(\tau) = \sum_{i=1}^m a_i K_i(\tau) \quad (3.13)$$

with

$$K_i(\tau) = \alpha(t_i) e^{\beta(t_i)\tau} \quad i = 1, \dots, m \quad (3.14)$$

and

$$\alpha(t_i) = \sqrt{\frac{e}{t_i}}, \quad \beta(t_i) = -\frac{1}{2t_i} \quad (3.15)$$

Which introduces $2m$ free parametric constants, a_1, \dots, a_m and t_1, \dots, t_m . In other words: we replace the Basset kernel by a linear combination of exponentials at the tail. As for the form of the functions, van Hinsberg et al. reason as follows: Let us consider the unknown parameters t_1, \dots, t_m to be a set of points that belong to the interval $(0, \infty)$. Let us then impose the condition that the graph of each K_1, \dots, K_m be tangent to that of K_B at the these points t_1, \dots, t_m . This condition looks to align the exponentials with the reference curve, so that a much smaller set of good candidate exponential approximants is considered. Precisely setting α and β as in Eq. (3.15) fulfils this purpose. The hope is that this choice will later facilitate the determination of the sets of parameters a_i and t_i .

Now, instead of Eq. (3.10), one has

$$\begin{aligned} \mathbf{I}_B(t) &= \int_{t_0}^t K_B(t - \tau) \mathbf{f}(\tau) d\tau \approx \int_{t-t_w}^t K_B(t - \tau) \mathbf{f}(\tau) d\tau + \int_{t_0}^{t-t_w} K(t - \tau) \mathbf{f}(\tau) d\tau \\ &= \mathbf{F}_w(t) + \mathbf{F}_t(t) := \mathbf{I}(t) \end{aligned} \quad (3.16)$$

where we assume $t - t_w \geq t_0$. Thus, the whole integral becomes the sum two terms: the first term on the right hand side, $\mathbf{F}_w(t)$, is the window contribution, which can be calculated

with any of the aforementioned methods (In particular, Hinsberg et al. employed the method explained in Section 3.2). The second term, $\mathbf{F}_t(t)$, is the tail contribution and is calculated as explained below.

The use of exponentials to approximate the tail of the kernel becomes the key that permits a recursive calculation of the tail integral, which leads to a radical decrease in memory requirements. Let us now show this recursive property of the tail kernel. It is enough to consider a single exponential term:

$$\begin{aligned}
\Psi(t + \Delta t) &= \int_{t_0}^{t+\Delta t-t_w} e^{\beta(t_i)(t+\Delta t-\tau)} \mathbf{f}(\tau) \\
&= \int_{t-t_w}^{t+\Delta t-t_w} e^{\beta(t_i)(t+\Delta t-\tau)} \mathbf{f}(\tau) + \int_{t_0}^{t-t_w} e^{\beta(t_i)(t+\Delta t-\tau)} \mathbf{f}(\tau) \\
&= \int_{t-t_w}^{t+\Delta t-t_w} e^{\beta(t_i)(t+\Delta t-\tau)} \mathbf{f}(\tau) + e^{\beta(t_i)\Delta t} \int_{t_0}^{t-t_w} e^{\beta(t_i)(t-\tau)} \mathbf{f}(\tau) \\
&= \int_{t-t_w}^{t+\Delta t-t_w} e^{\beta(t_i)(t+\Delta t-\tau)} \mathbf{f}(\tau) + e^{\beta(t_i)\Delta t} \Psi(t)
\end{aligned} \tag{3.17}$$

where we have introduced the auxiliary function Ψ to show the recursive property. Therefore

$$\begin{aligned}
\mathbf{F}_t(t) &= \sum_{i=1}^m a_i \mathbf{F}_i(t) \\
&= \sum_{i=1}^m a_i \left(\int_{t-\Delta t-t_w}^{t-t_w} K_i(t-\tau) \mathbf{f}(\tau) d\tau + \int_{t_0}^{t-\Delta t-t_w} K_i(t-\tau) \mathbf{f}(\tau) d\tau \right) \\
&= \sum_{i=1}^m a_i \left(\int_{t-\Delta t-t_w}^{t-t_w} K_i(t-\tau) \mathbf{f}(\tau) d\tau + e^{\beta(t_i)\Delta t} \mathbf{F}_i(t-\Delta t) \right) \\
&= \sum_{i=1}^m a_i (\mathbf{F}_{d,i}(t) + \mathbf{F}_{r,i}(t))
\end{aligned} \tag{3.18}$$

where

$$\begin{aligned}
\mathbf{F}_{d,i}(t) &= \int_{t-h-t_w}^{t-t_w} K_i(t-\tau) \mathbf{f}(\tau) d\tau \\
\mathbf{F}_{r,i}(t) &= e^{\beta(t_i)h} \mathbf{F}_i(t-h)
\end{aligned} \tag{3.19}$$

and where a suitable temporal discretization of $\mathbf{F}_{d,i}$ must be provided. Note how only recent values of the integrand are involved in the calculation of $\mathbf{F}_i(t)$; all the information regarding older values is extracted, in Eq. (3.19), by referring to the old value of the integral itself.

Of course, it is still left to define parameters a_i and t_i . Let us for this purpose consider the error associated with the approximation of K_B by K , which we would like to have as small as possible:

$$|\mathbf{F}_H(t) - \mathbf{F}_{H,K}(t)| = C_B \left| \int_{t_0}^t (K_B - K)(t - \tau) \frac{d\mathbf{f}(\tau)}{d\tau} d\tau + (K_B - K)(t - t_0)\mathbf{f}(t_0) \right| \quad (3.20)$$

where $\mathbf{F}_{H,K}$ is the approximate history force, obtained using \mathbf{I} in place of \mathbf{I}_B and $|\cdot|$ denotes the usual vector modulus. Note that we have employed the RHS form of Eq. (3.3) also for the kernel K . This is possible, given the regularity of K , see Appendix C. Furthermore, it is straightforward to generalize the bound given in [353], see Appendix D:

$$|\mathbf{F}_H(t) - \mathbf{F}_{H,K}(t)| \leq C_B \frac{\|\mathbf{f}\|_\infty}{\sqrt{t_w}} \left(|K_B(1) - \tilde{K}(1)| + \int_1^{\tilde{t}_0} \left| \frac{d}{d\tau} (K_B(\tau) - \tilde{K}(\tau)) \right| d\tau \right) \quad (3.21)$$

where \tilde{K} is obtained from K by replacing the t_i with their rescaled analogues, $\tilde{t}_i = t_i/t_w$ and $\tilde{t}_0 = (t - t_0)/t_w$ and $\|\cdot\|_\infty$ is defined by

$$\|\mathbf{f}\|_\infty = \inf \{C : |\mathbf{f}| \leq C \text{ a.e.}\} \quad (3.22)$$

In fact, these authors considered only the limit $t_0 \rightarrow -\infty$, at which the first term in the parenthesis in Eq. (3.21) vanishes. Their interest was therefore to find a good approximation for long simulation times. Indeed, by taking this limit in Eq. (3.21), we recover the bound derived by [353]:

$$|\mathbf{F}_H(t) - \mathbf{F}_{H,K}(t)| \leq C_B \frac{\|\mathbf{f}\|_\infty}{\sqrt{t_w}} \left(|1 - \tilde{K}(1)| + \int_1^\infty \left| \frac{d}{d\tau} (K_B(\tau) - \tilde{K}(\tau)) \right| d\tau \right) \quad (3.23)$$

Now, an important observation to be made at this point is that the bound in Eq. (3.23) is also a bound for all values $t_0 > -\infty$. This means it is legitimate to take it as a reference for real, finite-time simulations. Like Hinsberg et al. we now concentrate in this bound, with the idea of minimizing it.

Hinsberg et al. point out that the minimization of the quantity between parenthesis on the right hand side of the inequality Eq. (3.22) is not amenable to the standard Newton–Raphson algorithm. They propose the following substitute objective function, which is differentiable with respect to the unknown parameters a_i and t_i :

$$\left(1 - \tilde{K}(1)\right)^2 + \int_1^\infty \tau \left(\frac{d}{d\tau} (K_B(\tau) - \tilde{K}(\tau)) \right)^2 d\tau \quad (3.24)$$

The first step in the optimization procedure proposed by the same researchers is to ‘make a reasonable choice for \tilde{t}_i ’ and then to calculate the optimal a_i parameters by applying Newton–Raphson’s algorithm. About the choice of \tilde{t}_i , the only indication that is given is to make the choice of \tilde{t}_i so that they cover a great range and they become closer as they become smaller. A single set of \tilde{t}_i values was provided in [353], specifically, the 10-member set $\{0.1, 0.3, 1, 3, 40, 190, 1000, 6500, 50000\}$, for which the corresponding Newton–Raphson-averaged values were also provided.

The method has proved remarkably effective in reducing the computational cost of the simulations [353, 254]. Nonetheless, the method still leaves room for improvement. In particular, the following issues are not properly addressed within it:

- a) the possibility of having a nonzero initial relative velocity, which is not taken into account
- b) the unexplored possibility of using a different time step for the quadrature
- c) the choice of the t_i , which is not well defined (it relies on guessing)

Point a) we have briefly touched upon, showing that the method’s bound is still applicable in this case. Nonetheless, we do not pursue here an adaptable method which should include the dependence in time of the kernel’s free parameters. Item b) will be addressed in Section 3.3.1, where it is discussed in relation to the Discrete Element Method. Item c) is the objective of Section 3.3.2.

3.3.1 Introduction of quadrature substepping

A well-known feature of the soft-sphere DEM is its requirement of very fine temporal discretizations. The reason is the simultaneous time resolution of phenomena whose characteristic time scales have different orders of magnitude, which is inherent to the approach. Indeed, the time discretization of the small time scales (dynamics during contact) determines the minimum resolution of the whole simulation, whose total duration is of a much larger scale. The inevitable result is a very large total number of time steps.

While it is often possible to artificially soften the contact model to increase the smallest time scales without significantly altering the macroscopic response [95], this procedure can only be justified up to a certain point and is fundamentally limited by the necessity to stop the particles going through each other. The immediate consequence is that often the time resolution is excessively great for the description of bigger time scales like, for example, the acceleration of the particle due to the hydrodynamic forces.

Similarly, the time step associated with the quadrature of the Boussinesq–Basset integral might be unnecessarily small for the required precision if it is made equal to the overall integration time step. Accordingly, in this section we make the necessary modifications to the methods proposed by Daitche and van Hinsberg et al. to develop a more general method allowing for larger quadrature steps, while still falling back to the standard version when the quadrature step is made equal to the general time step. We refer to this technique as *quadrature sub-stepping*. Daitche refers to the numerical computation of the Boussinesq–Basset term for a given time as the quadrature method, while saving the term ‘integration’ for the numerical time-integration of the equation of motion. In the same spirit, we speak of ‘quadrature’ every time we refer to the calculation of the Boussinesq–Basset integral (e.g. *quadrature time step*) while the alternative context can be assumed otherwise (e.g., *time step* of the integration scheme).

Up to this point, the time discretization has been defined (both time steps of the scheme and quadrature points) solely by the sequence of discrete times t_0, \dots, t_1, \dots and the time step, h . Now it is necessary to consider an additional sequence τ_0, τ_1, \dots and a second time step Δt . In our notation, the finer time discretization, corresponding to the overall scheme, is defined by t_0, \dots, t_1, \dots and Δt ; while τ_0, τ_1, \dots and h define the coarser discretization used for the quadrature.

In order to make both discretizations conformal, we consider the most recent quadrature step to have a variable length, equal to ϕh , such that $\phi \in (0, 1]$ and $h = N_q \Delta t$, with N_q a positive integer; while keeping the length of the remaining steps, h , fixed. At the initial stages of the simulation, one new evaluation of \mathbf{f} , is added to the growing list every time a whole quadrature step is completed (i.e. when ϕh becomes equal to one). Once the simulated time has increased by t_w , the integrand list stops growing and every new evaluation added to it is compensated by removing the oldest one at the end of every quadrature step. The situation is represented in Fig. 3.2. Note that time-step numbering changes in time according to the following formulas

$$\begin{aligned}
 t_w &= N_w h \\
 n &= \min \left\{ N_w, \left\lfloor \frac{\tau_n + h}{h} \right\rfloor \right\} \\
 \phi &= \frac{\tau_n - \tau_{n-1}}{h} \\
 \tau_0 &= \max \{ \tau_n - t_w + (1 - \phi)h, t_0 \}
 \end{aligned} \tag{3.25}$$

where N_w is the user-defined number of quadrature steps in the window. Note that now t_w is a derived quantity which approximates the variable size of the window region in the present

generalized scheme with sub-stepping; see Fig. 3.2. Both quantities coincide only when $\phi = 1$.

Adapting the window contribution formulas

While any of the methods reviewed in Section 3.2 could be used, we will focus on the hybrid polynomial interpolation/analytic approach. We will first describe it in detail, closely following [89]. We will then proceed to generalize the method.

By following the steps outlined in Section 3.2.2 using Lagrange polynomials of a given order, one obtains particular versions of the method. The generic quadrature formula for any order reads, after conveniently reordering and grouping the different terms (see [89])

$$\int_{\tau_0}^t \frac{\mathbf{f}(\tau)}{\sqrt{t-\tau}} d\tau \approx \sqrt{h} \sum_{j=0}^n \mu_j^n \mathbf{f}(\tau_{n-j}) \quad (3.26)$$

where the coefficients μ_j^n are constants. The use of an extra index indicating the polynomial order of the interpolation is avoided. Instead, Daitche replaces μ_j^n with α_j^n (first order), with β_j^n (second order) or with γ_j^n (third order) when particularizing. We will restrict our attention to the first and the second-order versions of the method. The formulas for the α_j^n and β_j^n coefficients are

$$\alpha_j^n = \frac{4}{3} \begin{cases} 1 & j = 0 \\ (j-1)^{3/2} + (j+1)^{3/2} - 2j^{3/2} & 0 < j < n \\ (n-1)^{3/2} - n^{3/2} + \frac{3}{2}\sqrt{n} & j = n \end{cases} \quad (3.27)$$

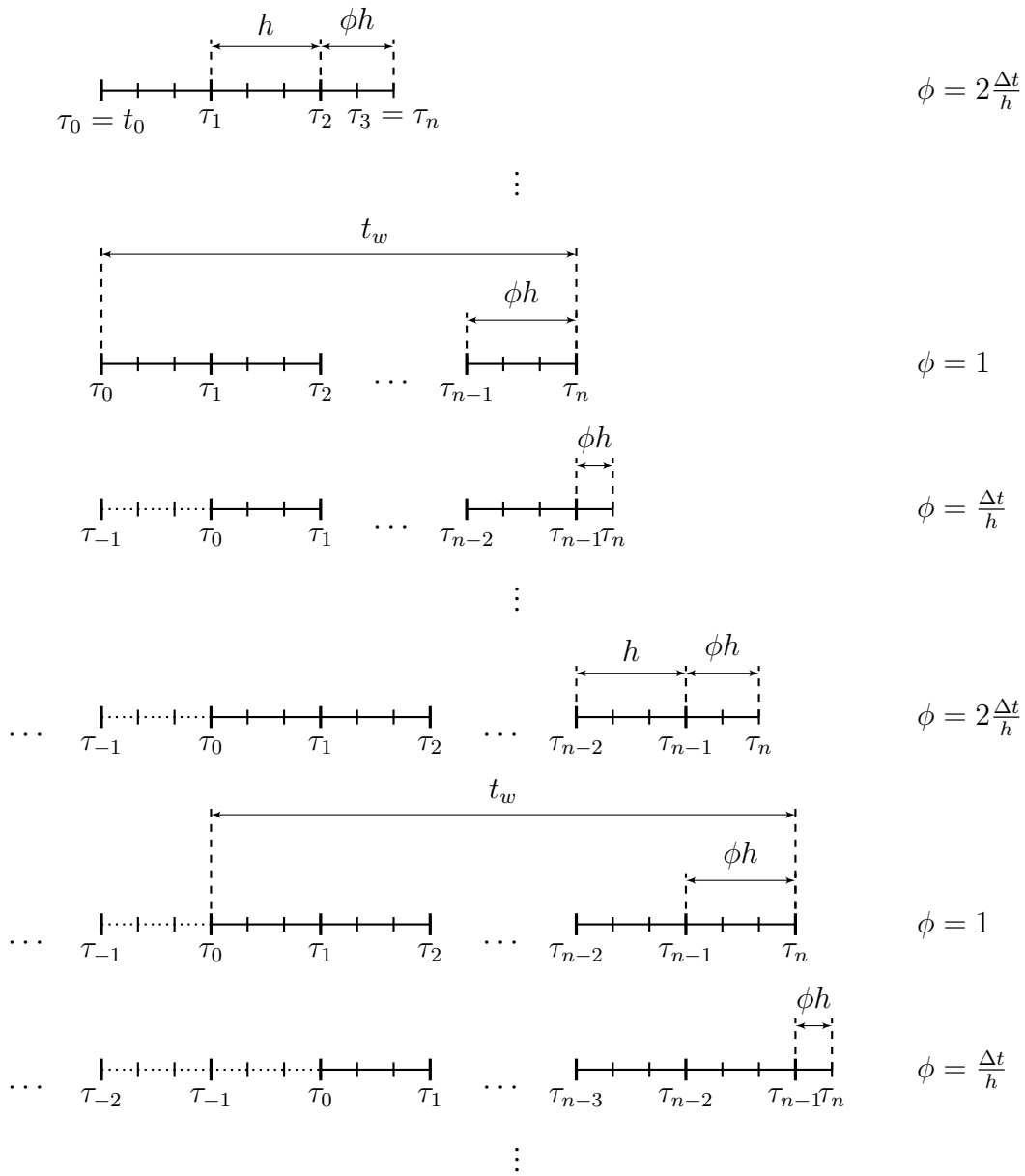


Fig. 3.2 Time-stepping procedure and designation of time points for a time integration step Δt and a quadrature step $h = 3\Delta t$; τ_n represents the current time; t_w is a fixed, user-defined quantity representing the window time (the actual discrete window time varies between t_w and $t_w - h + \Delta t$); time points on the dotted-line have ceased being kept track of.

$$\beta_j^n = \begin{cases} \begin{cases} \frac{12}{15}\sqrt{2} & j = 0 \\ \frac{16}{15}\sqrt{2} & j = 1 \\ \frac{2}{15}\sqrt{2} & j = 2 \end{cases} & n = 2 \\ \begin{cases} \frac{4}{5}\sqrt{2} & j = 0 \\ \frac{14}{5}\sqrt{3} - \frac{12}{5}\sqrt{2} & j = 1 \\ -\frac{8}{5}\sqrt{3} + \frac{12}{5}\sqrt{2} & j = 2 \\ \frac{4}{5}\sqrt{3} - \frac{4}{5}\sqrt{2} & j = 3 \end{cases} & n = 3 \\ \begin{cases} \frac{4}{5}\sqrt{2} & j = 0 \\ \frac{14}{5}\sqrt{3} - \frac{12}{5}\sqrt{2} & j = 1 \\ \frac{176}{15} - \frac{42}{5}\sqrt{3} + \frac{12}{5}\sqrt{2} & j = 2 \\ \frac{8}{15} \left((j+2)^{5/2} - 3(j+1)^{5/2} + 3j^{5/2} - (j-1)^{5/2} \right) \\ + \frac{2}{3} \left(-(j+2)^{3/2} + 3(j+1)^{3/2} - 3j^{3/2} + (j-1)^{3/2} \right) & 2 < j < n-1 \\ \frac{8}{15} \left(-2n^{5/2} + 3(n-1)^{5/2} - (n-2)^{5/2} \right) \\ + \frac{2}{3} \left(4n^{3/2} - 3(n-1)^{3/2} + (n-2)^{3/2} \right) & j = n-1 \\ \frac{8}{15} \left(n^{5/2} - (n-1)^{5/2} \right) + \frac{2}{3} \left(-3n^{3/2} + (n-1)^{3/2} \right) + 2\sqrt{n} & j = n \end{cases} & n \geq 4 \end{cases} \quad (3.28)$$

In our proposed quadrature sub-stepping scheme, the most recent quadrature step is shorter than the rest, its length being ϕh , according to Eq. (3.25). The formulas in Eq. (3.27) and Eq. (3.28) are however only valid for the special case $\phi = 1$, so they need to be modified. The length change of the most recent step has two effects. First, it affects the definition of the Lagrange polynomials that interpolate \mathbf{f} at the most recent point, τ_n . Second, it changes the formula for τ_n itself, appearing in the argument of the kernel in every integral. A quite detailed derivation of the formulas is included in Appendix B, where an inconsistency found in Daitche's paper (though not affecting the correctness of the order-specific formulas Eq. (3.27) and Eq. (3.28)) is discussed and fixed. The generalized formula for α_j^n is

$$\alpha_j^n = \frac{4}{3} \begin{cases} \phi & j = 0 \\ (j-2+\phi)^{3/2} + (j+\phi)^{3/2} - 2(j-1+\phi)^{3/2} & 0 < j < n \\ (n-2+\phi)^{3/2} - (n-1+\phi)^{3/2} + \frac{3}{2}\sqrt{n-1+\phi} & j = n \end{cases} \quad (3.29)$$

and the one for β_j^n reads

$$\beta_j^n = \begin{cases} \begin{cases} \frac{4}{15\phi}(1+\phi)(4\phi-1) & j=0 \\ \frac{4}{15\phi}(1+\phi)^{5/2} & j=1 \\ \frac{2}{15}\sqrt{1+\phi}(3-2\phi) & j=2 \end{cases} & n=2 \\ \begin{cases} \frac{4}{15\phi}(\sqrt{1+\phi}(4\phi-1)) & j=0 \\ \frac{2}{15}\left(\frac{2}{\phi}(1+\phi)^{5/2} - (1+\phi)^{3/2}(4\phi+9) + (2+\phi)^{3/2}(4\phi+3)\right) & j=1 \\ \frac{4}{15}\left((2+\phi)^{3/2}(2-4\phi) + \sqrt{1+\phi}(4\phi^2+7\phi-2)\right) & j=2 \\ \frac{2}{15}\left((1+\phi)^{3/2}(1-4\phi) + \sqrt{2+\phi}(4\phi^2+\phi+1)\right) & j=3 \end{cases} & n=3 \\ \begin{cases} \frac{4}{15\phi}\sqrt{1+\phi}(4\phi-1) & j=0 \\ \frac{2}{15}\left(\frac{1}{\phi}(\phi+1)^{5/2} + (2+\phi)^{3/2}(3+4\phi) + (3+4\phi)^{3/2}(9+4\phi)\right) & j=1 \\ \frac{2}{15}\left((\phi+1)(8\phi^2+14\phi-4) - 3(\phi+2)^{3/2}(3+4\phi) + (\phi+3)^{3/2}(7+4\phi)\right) & j=2 \\ \frac{8}{15}\left((j+1+\phi)^{5/2} - 3(j+\phi)^{5/2} + 3(j-1+\phi)^{5/2} - (j-2+\phi)^{5/2}\right) \\ + \frac{2}{3}\left(- (j+1+\phi)^{3/2} + 3(j+\phi)^{3/2} - 3(j-1+\phi)^{3/2} + (j-2+\phi)^{3/2}\right) & 2 < j < n-1 \\ \frac{8}{15}\left(-2(n-1+\phi)^{5/2} + 3(n-2+\phi)^{5/2} - (n-3+\phi)^{5/2}\right) \\ + \frac{2}{3}\left(4(n-1+\phi)^{3/2} - 3(n-2+\phi)^{3/2} + (n-3+\phi)^{3/2}\right) & j=n-1 \\ \frac{8}{15}\left((n-1+\phi)^{5/2} - (n-2+\phi)^{5/2}\right) + \\ \frac{2}{3}\left(-3(n-1+\phi)^{3/2} + (n-2+\phi)^{3/2}\right) + 2\sqrt{n-1+\phi} & j=n \end{cases} & n \geq 4 \end{cases} \quad (3.30)$$

Adapting the tail contribution formulas

In addition to modifying the window contribution formulas, the introduction of quadrature sub-stepping also requires adapting the tail contribution formulas of the MAE. From Eq. (3.3), we have

$$\mathbf{F}_{d,i}(t) = \int_{t-t_w-\Delta t}^{t-t_w} K_i(t-\tau) \mathbf{f}(\tau) d\tau \quad (3.31)$$

while

$$\mathbf{F}_{r,i}(t) = e^{\beta(t_i)\Delta t} \mathbf{F}_i \quad (3.32)$$

Let us see how this method generalizes to include sub-steps. According to the previous discussion, the actual time window is not in general constant (see Fig. 3.2), but equal to

$t_w + (\phi - 1)h$. Therefore, its value ranges

$$\left\{ t_w - h + \Delta t, t_w - h + 2\Delta t, \dots, t_w - h + \frac{N_q - 1}{N_q} \Delta t, t_w \right\} \quad (3.33)$$

where $N_q = h/\Delta t$, which by construction is an integer value. The original MAE corresponds to the case $\phi = 1$, for which the window time is constant and both a new quadrature point is added and the oldest one is forgotten at each time step. With sub-stepping, the situation is different. Now, changes in the set of quadrature points only take place every N_q steps, specifically, for $\phi = 1/N_q$ (see Fig. 3.2). The calculation of the tail contribution in this case is similar to Eq. (3.31), although the integration limits must be adapted appropriately:

$$\mathbf{F}_{d,i}(t) = \int_{t-t_w-\phi h}^{t-t_w+h-\phi h} K_i(t-\tau) \mathbf{f}(\tau) d\tau \quad (3.34)$$

On the other hand, for $\phi \neq 1/N_q$, there is no change in the quadrature points and only the integration limits change. In this case, there is no $\mathbf{F}_{d,i}$ and the whole tail contribution can be updated as $\mathbf{F}_{r,i}$, that is:

$$\mathbf{F}_i(t) = \mathbf{F}_{r,i}(t) = e^{\beta(t_i)\Delta t} \mathbf{F}_i(t - \Delta t) \quad (3.35)$$

The integral on the right-hand side of Eq. (3.3.1) can be approximately computed using the same technique as in Daitche's method, since the kernel K_i , now of exponential form, also leads to analytically computable integrals when convoluted with polynomials. This is, in fact, what was done in [353], where the first-order-polynomial version was considered. The generalized first-order version of the quadrature is given by (see Appendix Appendix B)

$$\begin{aligned} \mathbf{F}_{d,i}(t) = & \frac{e^{\beta(t_w+(\phi-1)h)}}{\beta^2 h} \left(\mathbf{f}(\tau_{-1}) (e^{\beta h} - h\beta - 1) + \mathbf{f}(\tau_0) (e^{\beta h}(\beta h - 1) + 1) \right) \\ & + \mathcal{O}(h^2) \end{aligned} \quad (3.36)$$

where the explicit dependence of $\beta(t_i)$ on t_i has been omitted for brevity. For $\phi = 1$, the expression becomes equivalent to the one presented in [353].

The extension to order two can also be done in the same manner. Nonetheless, the resulting formula turns out to be numerically unstable due to cancellation errors and thus we followed an alternative path. Since the kernels K_i are well behaved near $t_n - t_w$, it is possible to use a standard quadrature method there. By interpolating, not only \mathbf{f} , but the

product $K_i \mathbf{f}$, with second order polynomials one obtains the following quadrature rule:

$$\begin{aligned} \mathbf{F}_{d,i}(t) = \frac{e^{\beta(t_w + (\phi-1)h)}}{\beta^3 h^2} & \left(\mathbf{f}_i(\tau_{-1}) (e^{\beta h} (2 - 3\beta h + 2\beta^2 h^2) + \beta h - 2) \right. \\ & + \mathbf{f}_i(\tau_0) (-4e^{\beta h} (1 - \beta h) - 2\beta^2 h^2 + 4) \\ & \left. + \mathbf{f}_i(\tau_1) (e^{\beta h} (2 - \beta h) - \beta^2 h^2 - 2) \right) + \mathcal{O}(h^3) \end{aligned} \quad (3.37)$$

where $\mathbf{f}_i(\tau) := K_i(\tau_n - \beta) \mathbf{f}(\tau)$.

3.3.2 How to choose the t_i parameters

In this section we generalize the optimization problem considered in [353] in order to fix the free parameters a_i and t_i . First, we pose the problem in a mathematically sound setting. We then present several alternative options to circumvent some of the difficulties brought about by the original formulation. Next, we explore the behaviour of the different options, which show significant differences in behaviour. Finally, we present the results for the different alternatives. The best ones (listed in Appendix F) are used in the subsequent chapters.

Posing the optimization problem

Let us start by making explicit the dependence of the functions involved. Re-expressing the modified kernel Eq. (3.12) in terms of a_i and t_i , we obtain

$$\tilde{K}_T(\tilde{\tau}, a_i, \tilde{t}_i) = \sum \frac{a_i}{\sqrt{\tilde{t}_i}} e^{\frac{1}{2}(1 - \frac{\tilde{\tau}}{\tilde{t}_i})}. \quad (3.38)$$

and the Boussinesq–Basset kernel is

$$K_B(\tilde{\tau}) = \frac{1}{\sqrt{\tilde{\tau}}} \quad (3.39)$$

For the sake of a compact notation, we define the kernel approximation error as the following function:

$$e_K(\tilde{\tau}, a_i, \tilde{t}_i) = K_B(\tilde{\tau}) - \tilde{K}_T(\tilde{\tau}, a_i, \tilde{t}_i). \quad (3.40)$$

We would like to minimize the error in the calculation of the history force, \mathbf{F}_B . However the force depends on the unknown relative flow and thus we must settle for minimizing its

bound, given by Eq. (3.23). This is equivalent to minimizing

$$I_1(a_i, \tilde{t}_i) = |e_K(1, a_i, \tilde{t}_i)| + \int_1^\infty \left| \frac{\partial}{\partial \tilde{\tau}} e_K(\tilde{\tau}, a_i, \tilde{t}_i) \right| d\tilde{\tau}. \quad (3.41)$$

Anticipating the numerical difficulties associated with this cost function, we set up the minimization problem leaving the cost function unspecified so that alternative cost functions can be later considered as follows

$$\begin{aligned} & \underset{a_i, \tilde{t}_i}{\text{minimize}} && I(a_i, \tilde{t}_i) \\ & \text{subjected to:} && 0 \leq a_i \leq 1 \\ & && 0 \leq \tilde{t}_i \end{aligned} \quad (3.42)$$

where i takes integer values from 1 to m and $I(a_i, \tilde{t}_i)$ stands for a general objective function. For example, one of the alternatives is to consider $I \equiv I_1$.

Let us discuss this set-up in some detail. Regarding the design variables a_i , the zero lower bound is imposed in order to avoid exponentials (with negative exponent) weighted by a negative value. This would give rise to concave functions with which to approximate a convex one (one over the square root), which is not in the spirit of the MAE, which is based on the idea of restricting the set of exponentials to those resembling the Boussinesq–Basset kernel as much as possible. On the other hand, the upper bound of parameter a_i is arbitrary. However, this bound helps the optimizer to find the solution in a smaller space. Of course, sub-optimal values could be expected by this simplification, but in our numerical experience, when we do not use the upper bound constraint, the optimal solution always fulfils it anyway.

Likewise, the positivity of \tilde{t}_i is required to keep the values of the modified Boussinesq–Basset kernel real. Furthermore, requiring $\tilde{t}_i > 1$ may seem reasonable. This variable represents the non-dimensional history time, thus taking values on the window part may appear as unnatural. However, we do not add this constraint in order to get better solutions. In fact, the numerical results explained in the following sections support this idea. [353] had also made this point.

The box constraints can be directly handled by means of the line search method [261]. The problem contains a small number of design variables ($2m$), which represent at most 20 unknowns for the cases considered. Thus, the complexity in solving the optimization problem will depend mainly on the nature of the cost function.

We next present the different cost functions that we have considered. I_1 is included along three additional alternatives leading to more tractable optimization problems.

A **I_1 cost function:** The first option is defined by the use of I_1 itself (see Eq. (3.23)), which we repeat here for completeness:

$$I_1(a_i, \tilde{t}_i) := |e_K(1, a_i, \tilde{t}_i)| + \int_1^\infty \left| \frac{\partial}{\partial \tilde{\tau}} e_K(\tilde{\tau}, a_i, \tilde{t}_i) \right| d\tilde{\tau}. \quad (3.43)$$

Two terms are involved: the absolute value of the error at the point where the window and the tail meet ($t - t_w$) plus the absolute value of the error derivative integrated along the whole tail. The second term can be interpreted as the $W^{1,1}([1, \infty); \mathbb{R})$ semi-norm. The latter penalizes the outliers weakly, so the error is expected to be small in most of the points but remain large in few points.

Since the L^1 norm is linear with respect to its argument, the convexity and non-linearity properties of the problem are essentially determined by the dependence of the error derivative with respect to \tilde{t}_i (it also linear with respect to the a_i).

The abs-value function is not continuously differentiable, so a non straightforward treatment on the computation of the gradient will be required.

B **I_2 cost function:** Alternatively, we can consider replacing the absolute-value function in the I_1 cost function by the square function. That is

$$I_2(a_i, \tilde{t}_i) := e_K(1, a_i, \tilde{t}_i)^2 + \int_1^\infty \left(\frac{\partial}{\partial \tilde{\tau}} e_K(\tilde{\tau}, a_i, \tilde{t}_i) \right)^2 d\tilde{\tau} \quad (3.44)$$

This option leads with a stronger penalization of the outliers, i.e. points with large errors in the derivative tend to be eliminated more easily. In this case, the second term in the cost function can be interpreted as the $H^1([1, \infty); \mathbb{R})$ semi-norm, or the $L^2([1, \infty); \mathbb{R})$ norm of the error derivative.

With the L^2 norm we gain convexity but we add nonlinearity to the already nonlinear dependence on \tilde{t}_i . However we add a quadratic non-linearity which is relatively weak. In addition, this modification also contributes to improve regularity (I_2 is continuously differentiable) and no special treatment of the gradient is required.

C **I_{2t} cost function:** In this case, we replace the abs-value of the cost function by taking the square, to which we add an extra weight $\tilde{\tau}$.

$$I_{2t}(a_i, \tilde{t}_i) := e_K(1, a_i, \tilde{t}_i)^2 + \int_1^\infty \tilde{\tau} \left(\frac{\partial}{\partial \tilde{\tau}} e_K(\tilde{\tau}, a_i, \tilde{t}_i) \right)^2 d\tilde{\tau}. \quad (3.45)$$

This extra weight $\tilde{\tau}$ was included in [353] 'to correct for the change in norm'. With it the values at the end of the tail will be more strongly penalized than the values at the beginning of the tail. The convexity and non-linear aspects are very similar properties to those from Option B.

D I_{2tH} cost function: This option can be understood as a restricted version of Option C, were the $\tilde{t}_i = \tilde{T}_i$ are given. That is

$$I_{2tH}(a_i) := I_{2t}(a_i, \tilde{T}_i) = e_K(1, a_i, \tilde{T}_i)^2 + \int_1^\infty \tilde{\tau} \left(\frac{\partial}{\partial \tilde{\tau}} e_K(\tilde{\tau}, a_i, \tilde{T}_i) \right)^2 d\tilde{\tau}. \quad (3.46)$$

So that the dependence with respect to \tilde{t}_i has been removed. Basically, the \tilde{t}_i values (represented by \tilde{T}_i) must be provided as data. This is the cost function that was explored in [353], where an increasing separation between successive values was suggested for the T_i . Its properties are similar to those of Options C, but since the space of possible solutions has been reduced, we should expect higher optimal costs in this case.

Note that after a few manipulations (detailed in Appendix E) and taking $x_i := a_i$, the optimization problem can re-expressed in matrix notation as

$$\begin{aligned} & \underset{x}{\text{minimize}} && \frac{1}{2}x^T Ax - b^T x \\ & \text{subjected to:} && 0 \leq x \leq 1. \end{aligned} \quad (3.47)$$

where A and b are defined as in Appendix E. Eq. (3.47) thus has the form of standard quadratic programming problem. More specifically, with a quadratic cost function and box constraints. This kind of problem can be approached effectively by standard optimization algorithms [261].

Exploring the character of the cost functions

In this section, we explore the behaviour of the different options introduced above as a function of the variables a_i and \tilde{t}_i at low dimensionality ($m = 2$), to facilitate visualization.

Section 3.3.2 shows the cost function for Options A and C in terms of a_1 and \tilde{t}_1 . Note the discontinuity on the derivative of the cost function in Fig. 3.3a. In contrast, Fig. 3.3b shows the smoothness of the cost function I_{2t} , which greatly facilitate the search for minima. Its graph is otherwise very close to that of I_1 (Option A). We will show that the gain in regularity makes up for the difference.

Note that the cost function is, in both cases non-convex, specifically with respect to \tilde{t}_1 . Therefore local minima are expected to appear. Indeed, in practice we observed the

appearance of multitude of local minima, whose number grew very quickly with dimensionality. Nonetheless, for this low-dimensional case it is possible to exactly determine the optimum point, which is marked in Section 3.3.2.

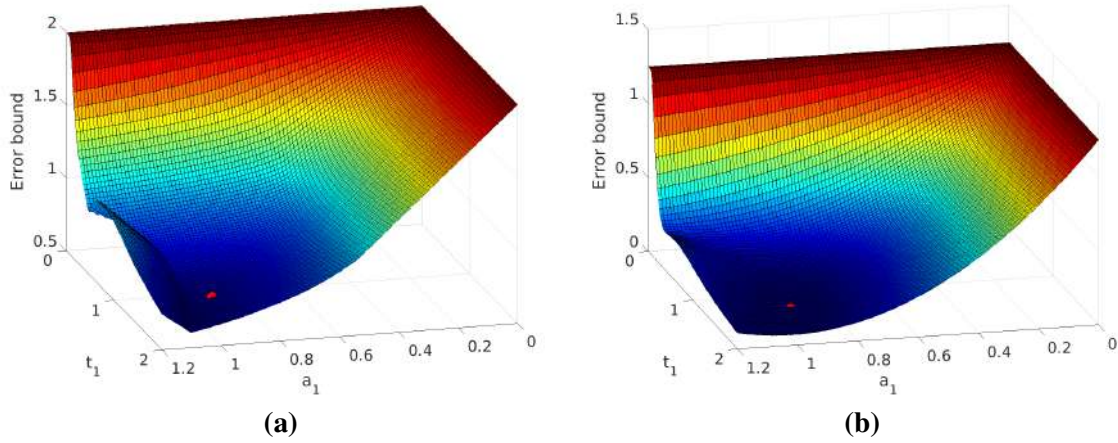


Fig. 3.3 Error bound for one exponential approximation of the Boussinesq–Basset kernel versus a_1 and t_1 for both Options A and C. The respective minima are marked with a red dot.

Next we study the dependence of the different cost function with respect as a_1 and a_2 are varied, while \tilde{t}_1^* and \tilde{t}_2^* are held fixed to their optimal values. Note from Fig. 3.4 the convexity of both cost functions with respect to the dependence with respect to the a_i , as compared to \tilde{t}_i . Certainly, to consider only the dependence on a_i makes the optimization problem much more tractable, especially as the dimensionality grows.

Finally, let us look at the dependence on \tilde{t}_1 and \tilde{t}_2 alone. Fig. 3.5 shows a contour plot of the I_1 bound in terms of \tilde{t}_1 and \tilde{t}_2 where, this time, a_1^* and a_2^* are being held fixed (at their optimal values). The strong gradients and the nonlinearity of the dependence on the \tilde{t}_i can be clearly appreciated. The cost function is almost flat close to the optimum, but away from it the gradient grows very quickly. This could indicate that the selection of the \tilde{t}_i might not need to be determined with extremely high precision. However, obtaining a fair approximation should be essential. Indeed, we have found in practice this to be the case (see next section).

Numerical solution of the optimization problem

In this work we have developed a code in Matlab[®] and solved the problem with a standard PC (3.40GHz processor in a 64-bit architecture). We have nonetheless taken advantage of symbolic coding mode in the Matlab environment to calculate all the derivatives and integrations.

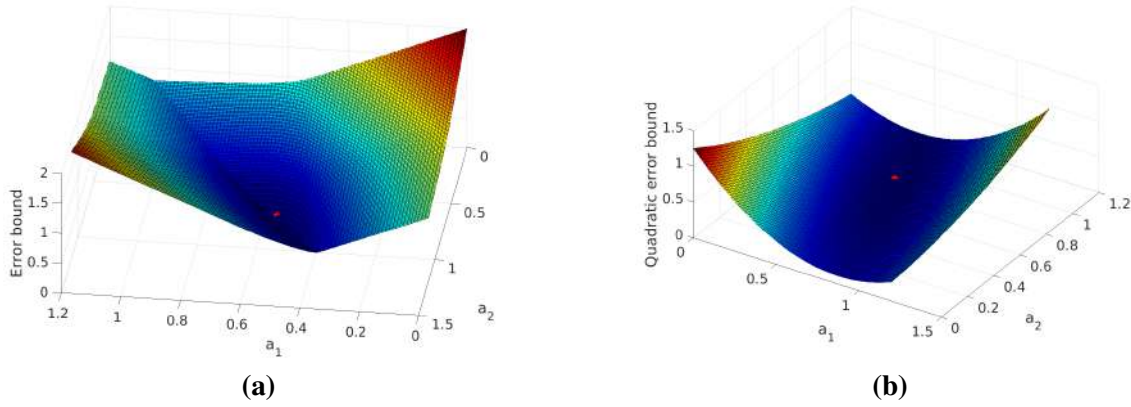


Fig. 3.4 Error bound for two exponentials as a function of a_1 and a_2 , with fixed (optimal) \tilde{t}_1^* and \tilde{t}_2^* for Options A and C. Options C and D are equivalent in this case.

Furthermore, We have also employed the optimization toolbox of Matlab. This tool makes combined use of a variety of methods (Newton-Raphson, Quasi-Newton and steepest descent, among others) for constraint optimization (box constraints in our case). Continuous optimization algorithms helped to obtain the real (local) minima. Furthermore, their convergence is normally faster than the alternative genetic algorithms. Specifically, we have used the interior point algorithm (primal-dual Newton-Raphson), which considers both the primal and dual variables simultaneously (see, e.g., [261]).

Our numerical experiments show that the problem is full of local minima, thus the strategy of initial point seeding becomes crucial. For each case m we must determine a_i and t_i , with $i = 1, \dots, m$. We propose the following heuristic initial point strategy, which has worked well up to $m = 10$, and consists in the following (An upper index m is used to refer to the case the parameters correspond to).

1. Take a_i^m randomly generated in $(0, 1]$.
2. For $m = 1$, take \tilde{t}_1^1 as also randomly generated in $(0, 1]$.
3. For $m = 2, \dots, 10$, define

$$\tilde{t}_i^m = \begin{cases} \tilde{t}_i^m & \text{if } i = 1 \\ \frac{\tilde{t}_{i-1}^{m-1} + \tilde{t}_i^{m-1}}{2} & \text{if } i = 2, \dots, m-1 \\ 10\tilde{t}_{m-1}^{m-1} & \text{if } i = m \end{cases} \quad (3.48)$$

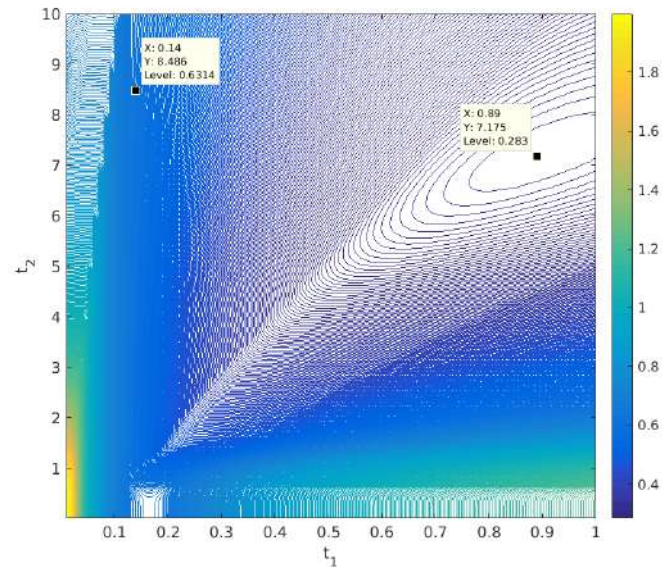


Fig. 3.5 Contour plot of the I_1 bound in terms of \tilde{t}_1 and \tilde{t}_2 . The dependence with respect to the t_i is clearly visible.

This initial strategy is used when solving Options A, B and C. In combination with it, we also employed the Global-search toolbox of Matlab, which is a type of multi-start technique for Global Optimization. This tool granted the possibility of exploring many different initial points and was specifically used with the a_i variables.

Let us make some particularized remarks for the different options:

For Option A the gradient is discontinuous in some regions. So in order to have a robust algorithm, we computed the gradient by perturbations, though at a high computational cost. For Options B and C, the gradient could be computed analytically (symbolically), leading to faster performance. For option D, the gradient is also computed analytically, taking the t_i as specified in 3.3.2.

Results

Fig. 3.6, shows the minimized a_i and \tilde{t}_i values of the upper bound for the different options. We have also included the single result reported in [353]. Case C seems to slightly outperform the others, although option A is very close, except perhaps for the last point, where the computational cost was already very high. note that this implies that, at least for the points where $I_{2t} < I_1$, either only local minima were found or convergence had not been achieved when using I_1 as the objective function. The complete list of optimized values a_i and \tilde{t}_i

corresponding to the I_1 and I_{2t} bounds can be found in Appendix F. Let us now make a few comments about these results, looking at each individual option.

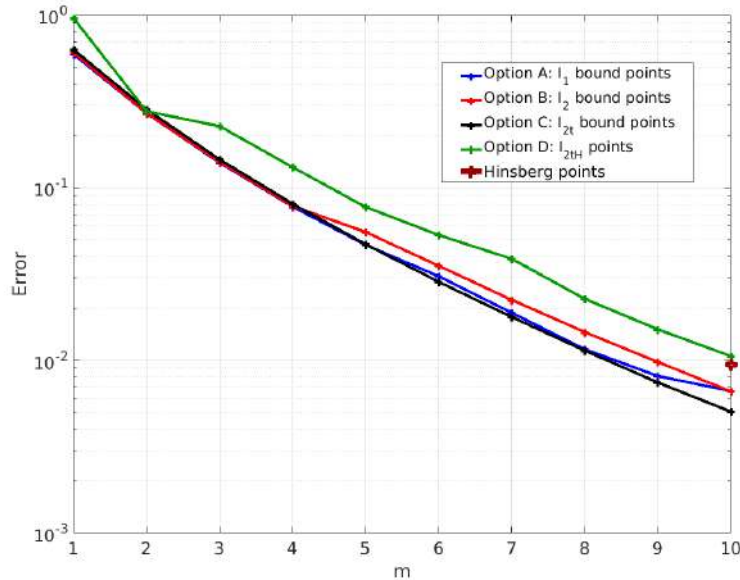


Fig. 3.6 Resulting values of the I_1 function for the optimized parameters using each of the different alternatives. The values provided by van Hinsberg et al. for $m = 10$ are also shown for comparison.

For Option A (I_1), for which the gradient is obtained by perturbations, the algorithm suffered significantly for $m > 4$. This was due large number of computations needed to evaluate the gradient, which increases exponentially with m when employing perturbations. That is, very sub-optimal solutions were found when the initial values were not set adequately. Precisely this fact justified the convenience of trying large number of initial values and, consequently, the use of Global-search with its associated computational costs. The results come out second-best (after Option C), even though this alternative is the only one that uses the original objective function, I_1 .

For Option B and C (I_2 and I_{2t}), the computational cost per initial point was significantly less, thanks to having an analytical expression for the gradient. Still, finding appropriate initial values remained very demanding. We again observed a strong dependence on initial values and a huge number of local minima significantly hampered the search. Nonetheless, thanks to a reduced cost in computing the gradient, a much larger number of initial points could be employed. As a result, Option C came out slightly on top of Option B and clearly beat Option A and Option D by a considerable margin.

Regarding Option D, number of design variables is halved and, consequently, the number of possible initial values decreases. Though there are local minima, a much smaller number of them were found in this case, which considerably sped up the computations. Nonetheless, the optimized costs were significantly higher than for the other options. Note how, encouragingly, with our proposed strategy 3.3.2 a very similar cost is achieved for $m = 10$ compared to that of the parameter set reported in [353].

In general, the bounds monotonically decrease as the number of exponentials increases for all the options, as expected. By looking at particular examples (see Section 3.6), we will see that the actual error also follows the same trend, although not as robustly.

Despite all the technology applied, the problem starts to become unaffordable for $m > 10$ for Options A, B and C. The high computational cost is basically caused by the number of required initial guesses, which strongly increases with m . This affects all the Options, except Option D, where the problem becomes less severe and the process could actually be carried on further. As we will see, this will not turn out to be necessary in many applications.

3.4 Overall algorithm

We next describe a finite difference discretization of the Maxey–Riley Equation. Our integration scheme of choice is largely borrowed from Daitche [89], who, as we have seen, used the same form of the Boussinesq–Basset force (a Riemann–Liouville half order derivative). A part from the quadrature, it can be classified as a semi-implicit Adams–Bashforth-type method. This type of scheme had earlier been used in [330] where inter-particle collisions were also considered. We have generalized the basic algorithm by adding a quadrature substepping cycle and combining it with Hinsberg’s window method (see sections Section 3.3). Furthermore, our algorithm is written in terms of the absolute velocities (instead of relative velocities), as this is the more convenient form for our DEM-compatible implementation. We focus on the second-order version of the scheme, although we use its first order version for the initial step, as it is commonly done with such multi-step methods.

Let us for convenience rewrite Eq. (2.1) as

$$(m_p + \frac{1}{2}m_f) \frac{d\mathbf{v}}{dt} = \mathbf{F}_{NH} + \mathbf{F}_H \quad (3.49)$$

where \mathbf{F}_{NH} stands for the ‘non-memory’ forces

$$\mathbf{F}_{NH} = \mathbf{F}_U^* + \mathbf{F}_D + \mathbf{F}_H + \mathbf{F}_W \quad (3.50)$$

and where now

$$\mathbf{F}_U^* = \mathbf{F}_U + \frac{1}{2}m_f \left(\frac{D\mathbf{u}}{Dt} + \frac{1}{10}a^2 \frac{d}{dt} (\nabla^2 \mathbf{u}) \right) \quad (3.51)$$

that is, \mathbf{F}_A has been split in two: one part has been moved to the LHS of the MRE and the rest has been absorbed, along with \mathbf{F}_U , in \mathbf{F}_U^* , which therefore now holds the non-historic contribution from the fluid acceleration. We obtain

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{1}{m_p + \frac{m_f}{2}} \left(\mathbf{H}(t + \Delta t) - \mathbf{H}(t) + \int_t^{t+\Delta t} \mathbf{F}_{NH} d\tau \right) \quad (3.52)$$

where

$$\mathbf{H}(\tau) := C_B \int_{t_0}^s \frac{1}{\sqrt{s - \tau}} \left(\mathbf{u}(\tau) - \mathbf{v}(\tau) + \frac{1}{6}a^2 \nabla^2 \mathbf{u}(\tau) \right) d\tau \quad (3.53)$$

The discretization of the integral term in Eq. (3.52) can be done for a variety of finite difference schemes. Here we have chosen the Adams–Bashforth Formulas, as in [89]. On the other hand, the \mathbf{H} -terms are partitioned according to the MAE, see Eq. (3.16), as

$$\mathbf{H}(s) = \mathbf{H}_w(s) + \mathbf{H}_t(s) \quad (3.54)$$

and where $\mathbf{H}_w(s)$ is calculated with the Daitche method, using the formulas from Section 3.3.1, while $\mathbf{H}_t(s)$ is calculated according to Eq. (3.16), Eq. (3.19) and Eq. (3.3). Furthermore, and similarly to what has been done with \mathbf{F}_A , the term $\mathbf{H}_{t,n+1}$, which also contains a term proportional to \mathbf{v}_{n+1} , is split and the latter is sent to the LHS of the equation, leaving the modified term $\mathbf{H}_{t,n+1}^*$ on the RHS. By doing so, the equations become *semi-implicit*. In the work of [353] it was found that the accuracy and stability of the resulting algorithm improved greatly, avoiding the need for extremely small time-steps. Our preliminary calculations indeed confirmed the same tendency. Denoting the time stepping index as k , the resulting first-order scheme (used in the initialization) reads

$$\begin{aligned} \mathbf{r}_{k+1} &= \mathbf{r}_k + \Delta t \mathbf{v}_k + \mathcal{O}(\Delta t^2) \\ \mathbf{v}_{k+1} &= \mathbf{v}_k + \frac{\Delta t}{m_p + \frac{1}{2}m_f + 6a^2 \sqrt{\pi \rho \mu \phi h} \alpha_0^n} \left(\mathbf{F}_{NH,k} + \mathbf{H}_{\alpha,k+1}^* - \mathbf{H}_{\alpha,k} \right) + \mathcal{O}(\Delta t^2) \end{aligned} \quad (3.55)$$

while the second order version is

$$\begin{aligned} \mathbf{r}_{k+1} &= \mathbf{r}_k + \frac{\Delta t}{2} (3\mathbf{v}_k - \mathbf{v}_{k-1}) + \mathcal{O}(\Delta t^3) \\ \mathbf{v}_{k+1} &= \mathbf{v}_k + \frac{\Delta t}{m_p + \frac{1}{2}m_f + 6a^2\sqrt{\pi\rho\mu\phi h}\beta_0^n} \left(\frac{3}{2}\mathbf{F}_{NH,k} - \frac{1}{2}\mathbf{F}_{NH,k-1} + \mathbf{H}_{\beta,k+1}^* - \mathbf{H}_{\beta,k} \right) \\ &\quad + \mathcal{O}(\Delta t^3) \end{aligned} \quad (3.56)$$

where the subscripts indicate the time-step at which each term must be evaluated and where and extra index (α or β) has been used to indicate which of the formulas, Eq. (3.29) or Eq. (3.30) must be used in each case. The necessity to move to the LHS the part of the discretized added mass force that is proportional to the unknown \mathbf{v}_{n+1} can be exemplified by the stability analysis of the following simplified equation. Suppose we must discretize

$$\frac{dv}{dt} = av + b \frac{dv}{dt} \quad (3.57)$$

where a and b are negative real values (they resist motion). The naive scheme

$$v_{k+1} = (a + 1)v_k + b \frac{v_k - v_{k-1}}{\Delta t} \quad (3.58)$$

is unstable for small time steps, as can easily be checked by applying von Neumann's method (see, e.g., [168]). Let us consider a harmonic of the exact solution of the difference scheme Eq. (3.58) at time step i to be given by $w_{\phi,i} = V_i \exp(\theta i)$, with i the imaginary unit. Replacing each v_i with $w_{\phi,i}$ in the finite difference scheme and dividing through by $\exp(\theta i)$ we obtain the relation

$$V_{k+1} = (1 + a)V_k + \frac{b}{\Delta t}(V_k - V_{k+1}) \quad (3.59)$$

which by calling $G_k = V_k/V_{k-1}$ can be expressed as

$$G_k^2 = \left(1 + a + \frac{b}{\Delta t}\right)G_k - \frac{b}{\Delta t} \frac{G_k}{G_{k+1}} \quad (3.60)$$

We are interested in the asymptotic behaviour of G_k , so that, unless G_k diverges we can assume $G_k/G_{k-1} \sim 1$ and we can write

$$G_k^2 - \left(1 + a + \frac{b}{\Delta t}\right)G_k + \frac{b}{\Delta t} = 0 \quad (3.61)$$

which is asymptotic to

$$G_k^2 - \frac{b}{\Delta t}G_k + \frac{b}{\Delta t} = 0 \quad (3.62)$$

as Δt tends to zero. But none of the solutions of the above quadratic equation yield $|G_k| < 1$ for $b/\Delta t$ large enough, and thus we conclude that the amplification factor grows without bound, proving that the naive scheme becomes unstable for small enough time steps.

Going back to the MRE, both expressions Eq. (3.55) and Eq. (3.56) can be more conveniently written as (eliminating the higher order truncation error terms)

$$\begin{aligned} \mathbf{r}_{k+1} &= \mathbf{r}_k + \Delta t \mathbf{v}_k \\ \mathbf{v}_{k+1} &= \mathbf{v}_k + \frac{\Delta t}{m_\mu} \mathbf{F}_\mu \end{aligned} \quad (3.63)$$

where μ takes either the value α (first order) or β (second order). The flow of instructions that implements the described time integration scheme with the MAE for the approximation of the tail can be expressed as the algorithms 1 and 2.

3.5 The fractional calculus perspective

In this section we make a connection between current work on the numerical discretization of the MRE and parallel research in fractional calculus, since we believe it has been missing so far in the literature.

Fractional calculus deals with the generalization of the standard (integer order) differential and integral operators to their analogues of arbitrary, real or complex order [307]. As mentioned in Section 3.2.1, there exist many definitions of such generalized operators [94]. Among those, the most popular are probably the Riemann–Liouville (R–L), Caputo and Grünwald–Letnikov (G–L)-type definitions.

The fractional derivative operators have similar properties to that of their integer-order ones. In particular, the three types mentioned above fulfil: i) linearity, ii) identity of the zeroth-order operator (i.e. its application to a function leaves the function unchanged), iii) backward compatibility (i.e. for integer orders one recovers the standard operators) and iv) index law (i.e. they respect the same composition rules as the standard operators); see [270]. Nevertheless, some important particularities make their treatment more difficult. Among these their nonlocality of the fractional operators stands out.

By writing equations relating various fractional order derivatives, one forms *fractional differential equations* (FDEs), in an analogy to their integer-order counterparts. Such equations have been used to model several physical processes such as anomalous diffusion [246],

Algorithm 1 Time integration algorithm for the MRE: single particle

```

1:  $k \leftarrow 0$  ▷ time step index
2:  $n \leftarrow 0$  ▷ quadrature step index
3:  $\tau_n \leftarrow t_0$ 
4:  $\mathbf{r} \leftarrow \mathbf{r}_0$ 
5:  $\mathbf{v} \leftarrow \mathbf{v}_0$ 
6:  $fList \leftarrow []$ 
7:  $F_iList \leftarrow []$ 
8: resize  $F_iList$  to size  $m$  ▷ chosen number of MAE exponentials
9: while  $\tau_n < t$  do
10:    $k \leftarrow k + 1$ 
11:    $\tau_n \leftarrow \tau_n + \Delta t$ 
12:    $\mathbf{u}, \nabla^2 \mathbf{u}, \dots \leftarrow \text{INTERPOLATEFLOW}(\mathbf{r}, \mathbf{v})$  ▷  $\mathbf{v}$  is needed for d/dt-type derivatives
13:   if  $\text{QUADRATURESTEPCOMPLETE}(k)$  then
14:      $\mathbf{f} \leftarrow \mathbf{u} - \mathbf{v} + \frac{1}{6}a^2 \nabla^2 \mathbf{u}$ 
15:     if  $\text{WINDOWISFULL}(n)$  then
16:       for  $f_i$  in  $fList$  except  $f_n$  do
17:          $f_i \leftarrow f_{i+1}$ 
18:        $f_n \leftarrow \mathbf{f}$ 
19:     else
20:        $n \leftarrow n + 1$ 
21:       append  $\mathbf{f}$  to  $fList$ 
22:    $\mathbf{r} \leftarrow \text{UPDATEPOSITION}(k, \mathbf{r}, \mathbf{v})$ 
23:   if  $\text{WINDOWISFULL}(n)$  then ▷ calculate tail term
24:     for  $i = 1, \dots, m$  do ▷ for each MAE exponential
25:       if  $\text{QUADRATURESTEPCOMPLETE}(k)$  then
26:          $F_iList[i] \leftarrow \text{ADDFREANDFDI}(n, F_iList, fList)$  ▷ Eq. (3.31) and
Eq. (3.32)
27:       else
28:          $F_iList[i] \leftarrow \text{ADDFRE}(F_iList, fList)$  ▷ Eq. (3.32)
29:    $\mathbf{F}_\mu \leftarrow \text{ADDRHS CONTRIBUTIONS}(k, n, \mathbf{v}, \mathbf{u}, \dots)$  ▷ window term + other forces
30:    $m_\mu \leftarrow \text{CALCULATEEQUIVALENTMASS}(k, n)$ 
31:    $\mathbf{v} \leftarrow \text{UPDATEVELOCITY}(k, \mathbf{r}, \mathbf{v}, m_\mu, \mathbf{F}_\mu)$  ▷ Eq. (3.56)

```

Algorithm 2 Implementation of some of the functions in Algorithm1

```

1: function WINDOWISFULL( $n$ )
2:   return  $n == N_w$  ▷ this includes  $f_{-1}$ , used in Equation
3: function QUADRATURESTEPCOMPLETE( $k$ )
4:   return  $k \equiv 0 \pmod{N_q}$ 
5: function UPDATEPOSITION( $k, \mathbf{r}, \mathbf{v}$ )
6:   if  $k == 1$  then
7:      $\mathbf{r} \leftarrow \text{UPDATEPOSITIONORDER1}(\mathbf{r}, \mathbf{v})$  ▷ Eq. (3.55)
8:   else
9:      $\mathbf{r} \leftarrow \text{UPDATEPOSITIONORDER2}(\mathbf{r}, \mathbf{v})$  ▷ Eq. (3.56)
10:  return  $\mathbf{r}$ 
11: function UPDATEVELOCITY( $k, \mathbf{r}, \mathbf{v}, m_\mu, \mathbf{F}_\mu$ )
12:  if  $k == 1$  then ▷  $\mu = \alpha$ 
13:     $\mathbf{r} \leftarrow \text{UPDATEVELOCITYORDER1}(\mathbf{r}, \mathbf{v}, m_\mu, \mathbf{F}_\mu)$  ▷ Eq. (3.55)
14:  else ▷  $\mu = \beta$ 
15:     $\mathbf{r} \leftarrow \text{UPDATEVELOCITYORDER2}(\mathbf{r}, \mathbf{v}, m_\mu, \mathbf{F}_\mu)$  ▷ Eq. (3.56)
16:  return  $\mathbf{v}$ 
17: function ADDRHSCONTRIBUTIONS( $k, n, \mathbf{v}, \mathbf{u}, \dots$ )
18:  if  $k == 1$  then ▷  $\mu = \alpha$ 
19:     $\{F_{NH,n}, H_{\beta,n+1}^*, H_{\alpha,n}\} \leftarrow \text{CALCULATERHSORDER1}(\mathbf{v}, \mathbf{u}, \dots)$ 
20:     $F_\mu \leftarrow F_{NH,k} + H_{\alpha,k+1}^* - H_{\alpha,k}$ 
21:  else ▷  $\mu = \beta$ 
22:     $\{F_{NH,n}, F_{NH,n-1}, H_{\beta,n+1}^*, H_{\beta,n}\} \leftarrow \text{CALCULATERHSORDER2}(\mathbf{v}, \mathbf{u}, \dots)$ 
23:     $F_\mu \leftarrow 3F_{NH,k} - F_{NH,k-1} + H_{\beta,k+1}^* - H_{\beta,k}$ 
24:  return  $F_\mu$ 
25: function CALCULATEEQUIVALENTMASS( $n$ )
26:  if  $n == 1$  then
27:     $m_\mu \leftarrow \text{CALCULATEADDEDMASSORDER1}(n)$  ▷ denominator in Equation
    Eq. (3.55)
28:  else
29:     $m_\mu \leftarrow \text{CALCULATEADDEDMASSORDER2}(n)$  ▷ denominator in Equation
    Eq. (3.63)
30:  return  $m_\mu$ 

```

viscoelasticity [230], hydrology [36], economics [357] and biological systems [291]; see also the review [339]. In particular, by using Eq. (3.7), the Maxey–Riley equation can be interpreted as a nonlinear system of fractional differential equations, where the only fractional-order term is the history force.

The theoretical study of FDEs has only started relatively recently [201] and its efforts have been focused in generalizing classical results from the theory of differential equations of integer order to the fractional setting. Some significant advances related to the problems of existence of solutions for initial value problems for FDE have been made, see [1, 201, 102], and, since these works were published, the number of publications in this area has accelerated. It is worth highlighting here the work of Farazmand and Haller [124], who established the existence and uniqueness of weak solutions for the MRE under suitable smoothness conditions; see also [203].

In parallel to the development of the theory, there have been significant advances in the research on numerical methods for the resolution of this type of equations. Most of these consist in adaptations of traditional numerical methods such as finite differences [205], finite elements [319, 76], finite volumes [191], spectral methods [130, 385], etc. The work [147] summarizes the particular challenges that arise in trying to solve FDEs and some strategies used to overcome them. Again, the nonlocality of the fractional-type operator becomes critical, as it leads to full matrices and full history-dependence, as with the case that concerns us.

New interest in physical models based on FDEs has spurred many developments in recent years [103, 69]; see also the broad review by [204]. Specifically, in the field of the numerical resolution of time-fractional differential equations with finite difference methods there have been significant advances, especially geared toward the resolution of multi-dimensional diffusion equations. Many of them focused in obtaining higher-order schemes, since those become more interesting for FDE than they do for standard differential equations, due to their nonlocal nature [69]. The derivative order in most of these works is considered generic, typically in the interval $(0, 2)$. For example, in [136] (see also [141]) an explicit, second order method is developed based on *fractional multi-step methods*; first introduced in [226]. Rigorous convergence order and stability analysis are also provided. A different approach based on truncating the infinite series defining the G–L derivative has been chosen by several authors. This method yields first order accuracy in its standard form, but can be enhanced by several procedures. By combining 'shifted' versions of the series, a second order in time and up to third-order accurate in space method was presented in [320]. The method was based on the existence of superconvergent points for the G–L series. A second order in time and up

to fourth order accuracy was obtained in [137] with a method based on the same principle. Both methods were applied to diffusion-related problems.

Other authors have focused on making approximations to decrease the huge computational and memory requirements that naturally result for these problems. For instance, in [96, 227] the authors make use of the *fading history principle* or *short memory principle* (see, e.g., [279]), which states that longer past points have less influence than more recent ones, which allowed them to simply neglect the influence of the former. Note that this is precisely the principle behind window methods. An alternative approach is presented in [209], where the number of time steps in which the solution needs to be stored is distributed unevenly, so that most of the information concentrates closer to the present time. The same work combines this technique Richardson's extrapolation to accelerate convergence. Other methods employ the same concept as window methods by applying different approximations to the convolution kernel for the window region and the tail. For instance, in [207] a method was used that is very similar to the MAE, in that the convolution kernel was in the tail region ('history part' in that work) by a linear combination of exponentials. A different technique is applied in [18], where a multi-pole approximation of the Laplace transform of the kernel is used instead.

However, even today, very few researchers working with the MRE are taking advantage or even mentioning these numerical methods. In fact, the very nature of the history term as a R–L derivative has been ignored for the majority of works on this field. For example, Dorgan and Loth [108] had to rediscover the standard window method, making no mention of the short memory principle nor of fractional calculus altogether. Notably, Bombardelli et al. [40] did make use of the R–L derivative interpretation, presenting their method based on truncating the G–L series (both definitions are equivalent under the studied hypotheses, see [279]). It is nonetheless revealing that these authors do not refer to previous work such as Yuste and Acedo [379], who had used essentially the same method before, although not specifically applying it to the MRE. Similarly, Diethelm et al. [104] (see also [61]) had employed a version of the polynomial/analytic approach later rediscovered by van Hinsberg et al. [353] and Daitche [89] (see also [61]), neither of whom mention these works (or any other works in numerical FDE). In fact, both techniques had been already described in the classic work by Oldham and Spanier [265], where even higher-than-one order G–L series were discussed already. This just goes to show the disconnection that has prevailed between these disciplines.

While the work of Bombardelli et al. was certainly not ignored, it did not seem to have a major impact in the field, due to the modest first-order accuracy that it delivered (although it still improved on previously used methods at the time), as it was soon surpassed by that of [353] and [89]. Nevertheless, we are convinced that the G–L approach can be

developed to become more competitive for this problem, since higher-order methods based on this method are becoming increasingly common [320, 227]; see also [51], who deals with high-order quadrature formulas. Conversely, we believe that all these works concerned with the numerical treatment of the history force in the MRE can be of value to the discipline of FDE, both in their proposed methodologies and in their role in promoting the visibility of this particular problem. Daitche [89] provided a semi-implicit method with high-order accuracy in time, which can be used to study fractional derivatives of order, not necessarily $1/2$. Similarly, the MAE might be of use in other areas apart from the MRE (despite it requiring, as it does in its current version, the derivative to be of order $1/2$), or at least inspire new methods to approximate the convolution kernel in other contexts.

In summary, the fractional derivative perspective on the analysis of the MRE enriches our understanding of the problem and can inspire future work, although this path has been somewhat neglected in the literature so far. Moreover, we expect the existing approaches to be of value to this broader subject and thus it deserves to be given more visibility. We hope that this section contributes to both goals.

In the next subsection we point to a direction in which the fractional derivative method of Bombardelli et al. could be improved. Admittedly, this section is tentative, but is included to motivate what we interpret to be interesting future lines of work.

3.5.1 Exploring an idea: Richardson’s extrapolation

Richardson’s extrapolation can be used to transform a sequence of approximants to a new sequence with an accelerated rate of convergence. Suppose that we take a sequence of partitions of the interval $[0, t]$ given by $\{h_1, \dots, h_k\}$ over which we wish to calculate the quadrature. The partitions are chosen such that each partition contains twice as many subintervals as the preceding one, that is

$$h_{i+1} = h_i/2 \quad (3.64)$$

The application of a given quadrature rule over each partition generates a corresponding sequence of approximants

$$S_k^0 = \{I_1^0(t), \dots, I_k^0(t)\} \quad (3.65)$$

where the notation $I_i^0(t)$ replaces $I_{h_i}^0(t)$ and the zero superscripts indicate the sequence is the zeroth element of a sequence (of sequences), i.e. the zeroth order Richardson sequence.

Now, let us assume the leading term of the error made by a generic I_i^0 can be expressed as

$$|I(t) - I_i^0(t)| = C^r h_i^r + \mathcal{O}(h_i^{r+1}) \quad (3.66)$$

so that the error scales with order r . Then the first-order Richardson sequence can be obtained by the following rule

$$I_i^1 = \frac{2^n I_i^0(t) - I_{i-1}^0}{2^n - 1} = 0 + \mathcal{O}h^{r+1} \quad (3.67)$$

for $1 \leq i \leq k - 1$, since $h_{i-1} = 2h_i$. Note that the last element of the new sequence S_{k-1}^1 formed following this rule is at least one order more accurate than the most accurate approximant in S_k^0 . Note also that the new sequence has one less element and therefore the process can be iterated a maximum of $k - 1$ times to obtain ever higher-order approximations.

Richardson's extrapolation does not always work. A review of possible difficulties associated with its use can be found in [52]. The issue of numerical stability and convergence are treated in detail in [315]. We are going to proceed empirically, showing how well it performs for the main quadrature methods seen so far. In general, a very important restriction is given by the fact that obtaining a sequence of partitions such as the one described above, formed by successive powers of two, cannot be achieved in practice in a dynamically changing history of a real simulation. Nonetheless, it is possible to do such a thing with the window method, as the window length can be forced to have the desired number of points. We next show that Bombardelli's method is the best suited for the application of Richardson's technique, turning it into a much more competitive alternative.

Let us, first of all, look at the performance of polynomial interpolation/analytic methods with Richardson's extrapolation for comparison. Since the theoretical order of the error was derived in [89], the corresponding exponents r are known and the method can be directly applied. Fig. 3.7 shows the results for a total duration $t = 10$. The method seems to work best for the first order polynomial case (second order accurate). Its performance becomes comparable to that of order two, but not better. The performance with the other methods is even poorer. The same tendency has been observed for different durations. We suspect this may be due to the piecewise definition of the quadrature formulas, which does not lead to very well-behaved asymptotic error expressions, with perfectly stable coefficients, but more work is needed to establish this point with confidence.

It is remarkable that, in any case, the gains only start to be noticeable after the mesh has become already quite fine. This factor goes against our goal of keeping the set of historic points to a minimum, which requires a very good approximation for the coarsest discretizations.

On the other hand, notice the strong and consistent gains that are achieved by applying the same technique to the fractional derivative approach in Figs. (3.8)–(3.11). Note furthermore how the comparison is more and more favourable to this technique as the total time duration increases.

While we will not pursue this line of research further in this work, we recognize such phenomenon to be an opportunity for boosting the performance of the method. Such a method would be a kind of window method that would consider a number of points in the window region equal to some power of 2. The window contribution would be calculated using Richardson’s extrapolation on the Grünwald–Letnikov fractional derivative formula. Let us

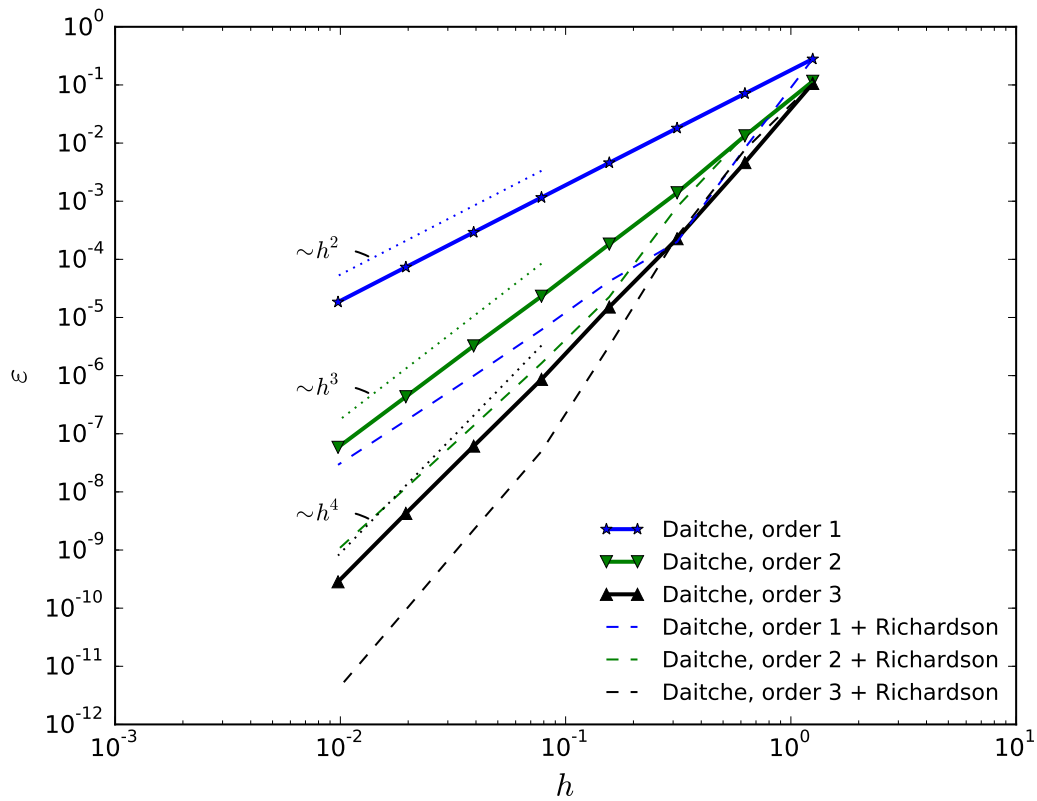


Fig. 3.7 Scaling of the global error of the Daitche method with Richardson extrapolation for $t = 10$

now analyse the performance of Bombardelli’s method with Richardson’s extrapolation.

3.6 Performance of the methodology

In this section we test the performance of the MAE using the optimized a_i and t_i values from Section 3.3.2. We start with a very elementary example, aimed only at measuring the quadrature error. Next, we consider a single-particle example with analytical solution to benchmark the accuracy of the full scheme. The last example features a long-term, 10 000 particle simulation with which we show the remarkable efficiency of the method.

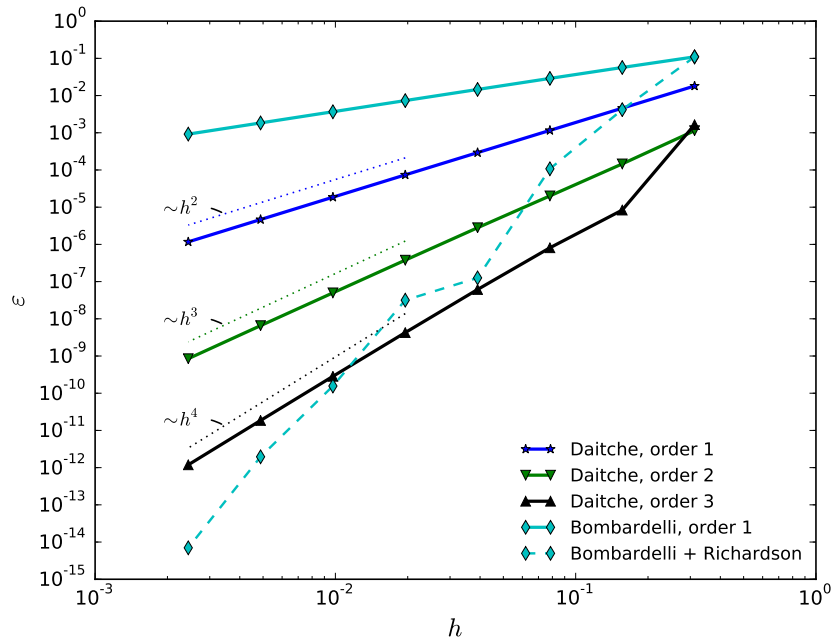


Fig. 3.8 Scaling of global error, $t = 2.5$

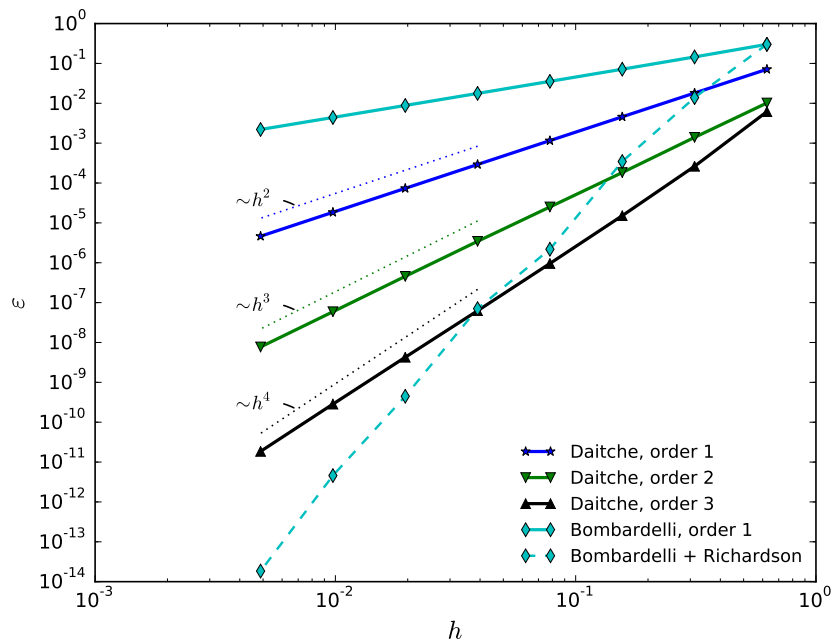


Fig. 3.9 Scaling of global error, $t = 5$

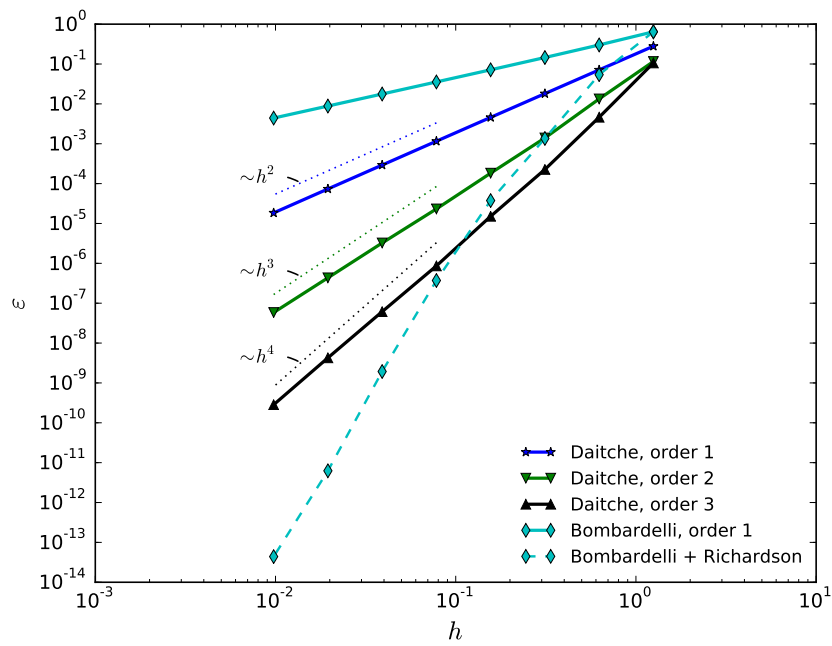


Fig. 3.10 Scaling of global error, $t = 10$

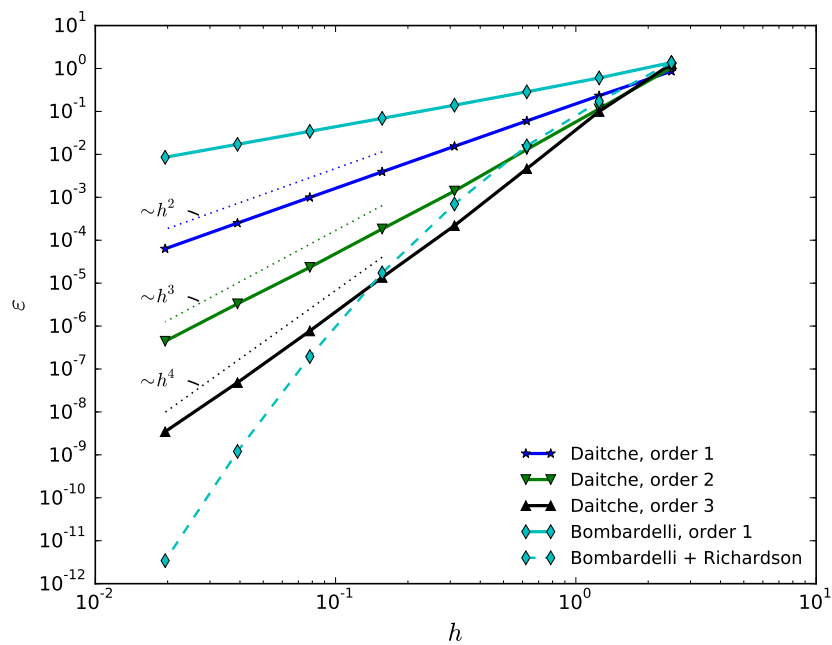


Fig. 3.11 Scaling of global error, $t = 20$

3.6.1 First benchmark: an integral with analytical solution

The error bound Eq. (3.23) can indeed be used for conservative predictions about the expected error when using the MAE. Nonetheless, we wish investigate how this bound in fact relates to the actual error, but of course this can only be done for particular cases. We revisit the same example that we considered in Section 3.2.3. Note that it corresponds to the convolution of the sine function with the Boussinesq–Basset kernel, which may be denoted $\sin * K_B(t)$. This is a very convenient and commonly chosen example, see [40, 353, 89]⁴. Its physical significance is that of a sphere being forced to oscillate in an otherwise quiescent fluid. The error can be expressed, see Eq. (3.23), as

$$E(t) := \left| \sin * (K_B - \tilde{K})(t) \right| \leq \frac{1}{\sqrt{t_w}} \left(\left| 1 - \tilde{K}(1) \right| + \int_1^\infty \left| \frac{d}{d\tau} (K_B(\tau) - \tilde{K}(\tau)) \right| d\tau \right) \quad (3.68)$$

since $\|\cos\|_\infty = 1$. We have computed the bound numerically, by partitioning the integral in Eq. (3.68) in two parts: an integral over the region where the argument of the absolute value may change and an integral over the rest, where the sign is given by that of $dK_B/d\tau$ (which has a slower asymptotic decay than $dK/d\tau$ as $t \rightarrow \infty$). The first part can be calculated (to very high accuracy) using a standard quadrature method, while the second one has an analytical formula. In order to obtain a representative measure of the error, we evaluate the integral over $[-\infty, 2\pi - \phi_i]$, with ϕ_i sampled at 40, evenly distributed, points in $[0, 2\pi]$ and we take the mean absolute error within the sample⁵. We denote this error as $E_{2\pi}(t_w, m)$, making explicit its dependence on both t_w and m . The results are shown in Fig. 3.12, where three different sets of parameters are considered: the single list given in [353] ($m = 10$), resulting from the optimization of I_{2tH} ; the set obtained from the optimization of I_1 , for $m = 1, \dots, 10$; and the set obtained from the optimization of I_{2t} , again for $m = 1, \dots, 10$ (see Section 3.3.2). For the sake of clarity, we include a single set of results in Fig. 3.12a where we have picked $t_w = 2\pi/10$. Fig. 3.12b shows the analogous results for a range of window times, including the single curve from Fig. 3.12a.

Fig. 3.12 shows how indeed the errors fall significantly below their optimized upper bounds, up to more than an order of magnitude so. Overall, it seems that both the I_1 and the I_{2t} methods achieve comparable results, while the parameters by van Hinsberg et al. come out as a bit less accurate. Nonetheless, this difference turns out not to be consistently significant, despite our inclusion of the t_i parameters as variables in the optimization problem. This

⁴Actually, the convolution of the cosine was considered by [40] and by [353].

⁵Note that the convolution of the sine function is periodic, with the same fundamental period as the sine, as it is readily seen with a change of variable.

indicates that the trial-and-error strategy used by these authors yielded a close-to-optimal result.

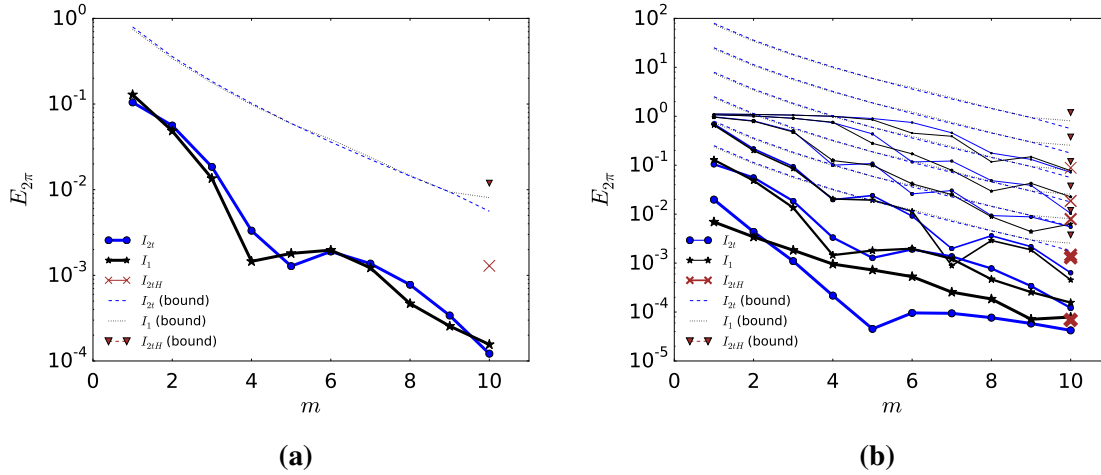


Fig. 3.12 Error produced by the approximation $K \approx K_B$ in calculating $\sin * K_B(t)$ for three different sets of values a_i and t_i , corresponding to different cost functions: Option A, Option C and the point given by van Hinsberg et al. The resulting errors are accompanied by their predicted upper bounds. A window time of $t_w = 0.2\pi$ has been considered in Fig. 3.12a. Fig. 3.12b shows the analogous results for different values of t_w corresponding to $t_w = 2\pi 10^k$ for $k = -5, \dots, 0$ (with thicker lines corresponding to larger values of t_w).

In theory, the greater the window, the smaller the error should be, because the approximation of the kernel is done over a smaller portion of the total domain. This tendency is expected to be especially pronounced for small values of t_w , close to the singularity of K_B at $\tau = t$. There, the well-behaved exponentials have difficulty keeping close to the curve for as it diverges. This is indeed what can be seen in Fig. 3.12b. The time window ranges 1×10^{-5} to 1 fractions of a period and is represented by using greater thickness on the error curves corresponding to larger windows. The tendency of the curves to go horizontal for increasing m signals the breakdown of the kernel approximation for extremely small values of t_w . Note that, for t_w smaller than 1×10^{-3} , none of our optimal a_i, t_i sets manage to keep the error under the 1×10^{-2} mark.

However, a small t_w requires an even smaller integration time step. So as long as the time step is kept large enough, such small time windows become unnecessary, because the memory cost can be afforded. This is why the most effective solution is to use a high accuracy scheme, such as the second or third Daitche schemes for the integration of the window in the MAE.

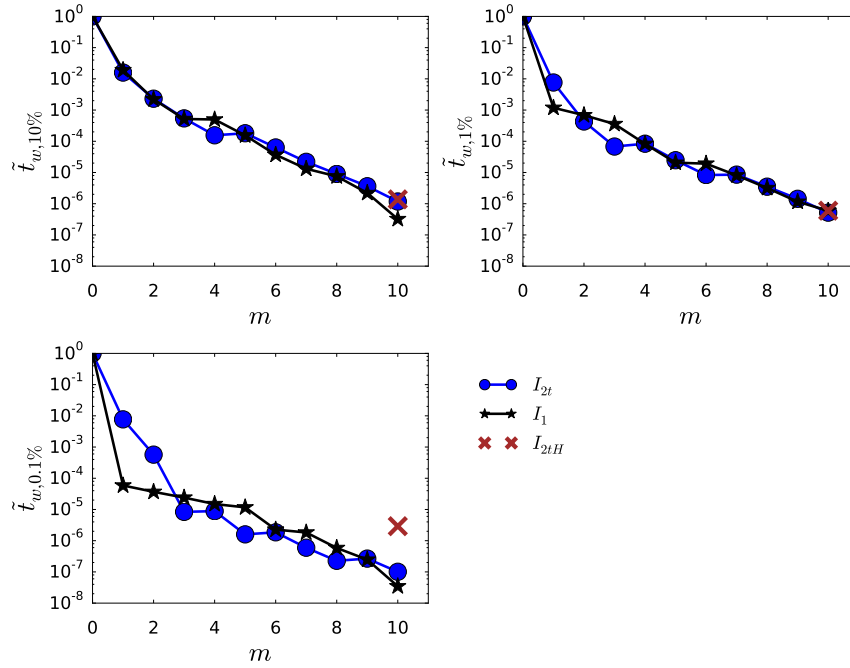


Fig. 3.13 Minimal window time necessary to obtain an error $E_{2\pi} = 1\%$, as a function of the number of exponentials. The time is normalized by the minimal time window required when $m = 0$.

A final remark about Fig. 3.12b: note that although the expected monotonous behaviour of the error with respect to variations in t_w is mostly realized, there can be exceptions: for $m = 7$ there is a crossing between adjacent sets of curves. This fact highlights the complexity of the relations that govern the method.

A different way to characterize the accuracy gains obtained by increasing the number of exponentials is represented by Fig. 3.13. It shows the relation between the number of exponentials, m and the minimum window time necessary to attain a desired level of accuracy. On the horizontal axis we represent m ; on the vertical axis, the normalized window time $\tilde{t}_{w,1\%}(m)$. The latter is defined as

$$\tilde{t}_{w,1\%}(m) := \frac{t_{w,1\%}(m)}{t_{w,1\%}(0)} \quad (3.69)$$

where

$$t_{w,1\%}(m) := t_w \quad \text{such that} \quad E_{2\pi}(t_w, m) = 0.01 \quad (3.70)$$

and similarly we define $t_{w,10\%}$ and $t_{w,0.1\%}$. In other words, Fig. 3.13 addresses the question if $t_{w,x\%}(0)$ is the minimal time window necessary to have less than $x\%$ error when using the standard window method, then what fraction of $t_{w,x\%}(0)$ is instead required when using the

MAE?. The answer will depend on the values of m and x . Strictly speaking, the existence of a unique solution is not even guaranteed, since that would require the strict monotonicity of the dependence of $E_{2\pi}$ on m , which we have seen to only hold approximately. Nonetheless we have applied the bisection algorithm to find such solutions, producing Fig. 3.13. Note the immense memory cutoffs that are generated by the MAE. For instance, suppose the accuracy goal set to one percent. Then just by including a single exponential approximant, the interval of the particles history that must be tracked is reduced by a factor 10^3 , by just using the I_1 -optimized a_1 and t_1 . Or else by more than 10^6 if ten exponentials are considered.

Let us now consider the joint effect of the quadrature algorithm and the kernel approximation, using the same test flow as above. For that, the initial time is taken as $t_0 = 0$, the final time is set to 10. The time interval $[0, 10]$ is initially partitioned in eighty parts. This amount is successively doubled, defining the range of values of the time step, h . For simplicity, the error is now measured at single point ($t = 10$); that is $E(10)$. Fig. 3.14 shows the performance of the algorithm described in Section 3.4 for a fixed $t_w = 1$ and $m = 10$. The method initially follows the third order slope of the second-order accurate Daitche quadrature algorithm. But then, as soon as the time integration error becomes dominated by the error due to the approximation $K \approx K_B$, its accuracy stagnates, rendering further reductions in h futile.

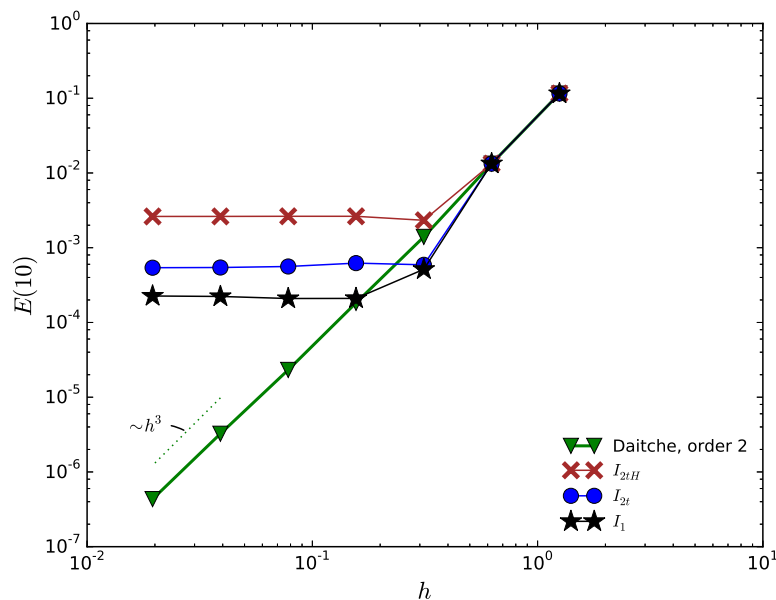


Fig. 3.14 Error vs. time step of Daitche's method as compared to the MAE for different optimized values t_i and a_i . The curves reach a plateau as soon as the kernel approximation error exceeds the quadrature error.

3.6.2 Second benchmark: Candelier's solution

There are a limited number of known closed-form particular solutions for the Maxey-Riley equation. Some of these can be found in the works cited in [124]. Fortunately, the solution obtained by Candelier et al. [57] includes the effect of the Boussinesq–Basset force, as well as all the other forces, though the Faxén terms do not contribute in this case. The solution corresponds to the trajectory of a particle that sediments under gravity in an infinite container which is rotating as a rigid solid around a fixed axis. The same benchmark was considered in [89] (see the same work for the input parameters). Fig. 3.15 shows two spiral trajectories resulting from two different versions of the solution by Candelier and co-workers. The inner spiral corresponds to the full solution, including all the terms in the MRE. The outer spiral is obtained by neglecting the Boussinesq–Basset contribution, while retaining the remaining effects. The asymptotic behaviour of the radial coordinate around the rotation axis is in both cases of exponential form, of which the exponent's coefficient is modified by the presence of Boussinesq–Basset force. This asymptotic solution becomes a very good approximation only after a few turnover times [57]. The exponential form suggests that this test is especially demanding, as it should tend to amplify any systematic inaccuracies over time.

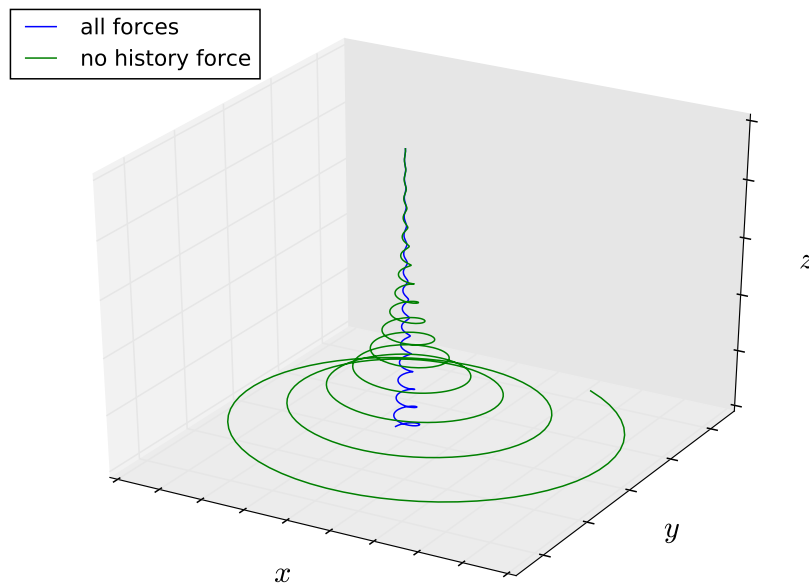


Fig. 3.15 Trajectories described by the particle in the Candelier et al. benchmark after 100s ($50/\pi$ rotation periods). The mass of fluid is rotating around the axis $x = y = 0$. The curve with the largest maximum radial coordinate corresponds to the solution without Boussinesq–Basset force, while the other does include the effect.

For the purpose of measuring the accuracy of the algorithm, let us define the error function as

$$E(t) = \frac{r(t) - r_{\text{num}}(t)}{r(t)} \quad (3.71)$$

where $r(t) = \sqrt{x(t)^2 + y(t)^2}$ is the exact radial coordinate and r_{num} its numerical approximation. We consider $E(100)$ as the measure of the error. This measure has a very similar behaviour to the max-norm over all the time steps. This is because the error tends to accumulate in a monotone way, leading to the maximum value occurring at the final point in most cases; see [89]. For instance, the error introduced by neglecting the Boussinesq–Basset force at $t = 100$ is $\frac{r(100) - r_{\text{no Basset}}(t)}{r(100)} > 14$.

Fig. 3.16 shows the error for different values of m for the MAE as compared to the bare, second order Daitche method. The optimal points were those obtained with the cost function I_1 , since this set gave slightly more accurate results for all values of m in this example. We observe that, despite a few permutations do occur between curves with successive m values, in general terms increasing m yields an increase in accuracy which, on average, amounts to about half an order of magnitude each successive increase.

But, *how many exponentials should be used?*. Based on accuracy, the safest answer should be *as many as possible*, or at least, *as many as necessary*. However, taking efficiency into consideration might complicate the matter. Indeed, adding additional exponentials to the kernel has two detrimental effects on efficiency. First of all, it implies a proportional increase in the total number of operations needed to compute the tail contribution. But more importantly, each new exponential requires an extra vector to be kept in memory per particle, i.e. the value of \mathbf{F}_i in the previous step, see Eq. (3.19). Of course it is still possible to increase the time window, thus improving the effectiveness of a fixed number of exponentials, as Fig. 3.12 suggests. But again, this increase also implies a raise in the memory demand. The situation is summarized in Fig. 3.17, where the error is plotted against the total number of bytes to be kept in memory per particle. The number of bytes is estimated by assuming that 24 bytes are taken up by each vector, as

$$\text{number of bytes} = 24(m + 1 + t_w/h) \quad (3.72)$$

where the time step h is kept constant at 2.5×10^{-3} , close to the saturation time step (below which no further gains are obtained, see Fig. 3.16). This corresponds to an error of around 10^{-3} for 8, 9 or 10 exponentials, according to the same figure. The window time is successively doubled, starting at $t_w = 5h$. Again, Fig. 3.16 shows that only initial increases in the memory demand (via increase in the number of exponentials or the value of t_w) yield significant gains in accuracy. Furthermore, the gains in accuracy per byte are greater when investing them into

more exponentials rather than into additional t_w (except perhaps for the very smallest values of t_w). In other words, within the analysed ranges, it pays off to take t_w as small as possible, while using as many exponentials as necessary.

Some remarks concerning the choice of t_w

It is not quite clear how small is 'as small as possible'. The smallest value considered in Fig. 3.17 is $t_w = 0.05$ s, or $t_w \approx 0.008$ periods. This t_w corresponds to an error of around 10^{-3} , based on Fig. 3.14. Making it smaller might very quickly lead to important approximation errors. Furthermore, since the time step is only a fifth of that amount, only five points enter the window region. Any further decrease in t_w would surely damage accuracy, while not significantly reducing the memory requirements, according to Eq. (3.72), with $m = 10$.

Based on this line of reasoning (and our numerical experience), we conclude that ten points are a reasonable option, which we have adopted by default in conjunction to $m = 10$. This translates into 21 vectors to be stored per particle. Indeed, for a fixed integration time step (which is typically determined by other factors, such as the overall time scheme or the fluid time resolution) the accuracy gain is quite substantial compared to, say, only five points, while still avoiding important memory sanctions. We do recommend this combination as a starting point in other simulations, as long as the frequency of the relative motion is low enough. That is, making sure that at least a few time steps fall within the period corresponding to the highest frequency modes in the motion. For instance, in a turbulent flow, this frequency is that of the Kolmogorov microscales, unless some external, high frequency force is acting on the particle. We test this 10-point, $m = 10$ combination in the following benchmark.

Effect of Substepping

Let us test the performance of the algorithm with sub-stepping. Fig. 3.18 shows the error evolution for different cases in which either the overall time step is modified, or else only the quadrature time step is. Note that the memory requirements per particle associated with the storage of historical values are inversely proportional to the quadrature time step or, equivalently, to N_q . Therefore, scaling either the time step or just the quadrature time step (i.e. N_q) by the same amount imply equal reductions in the memory requirements. However, the accuracy losses are very different. Note that the loss in accuracy for increasing N_q is first only concentrated in the initial stages of the motion, where the whole history is very recent. There the kernel is very large for all the integrated values, amplifying the error. Note that this

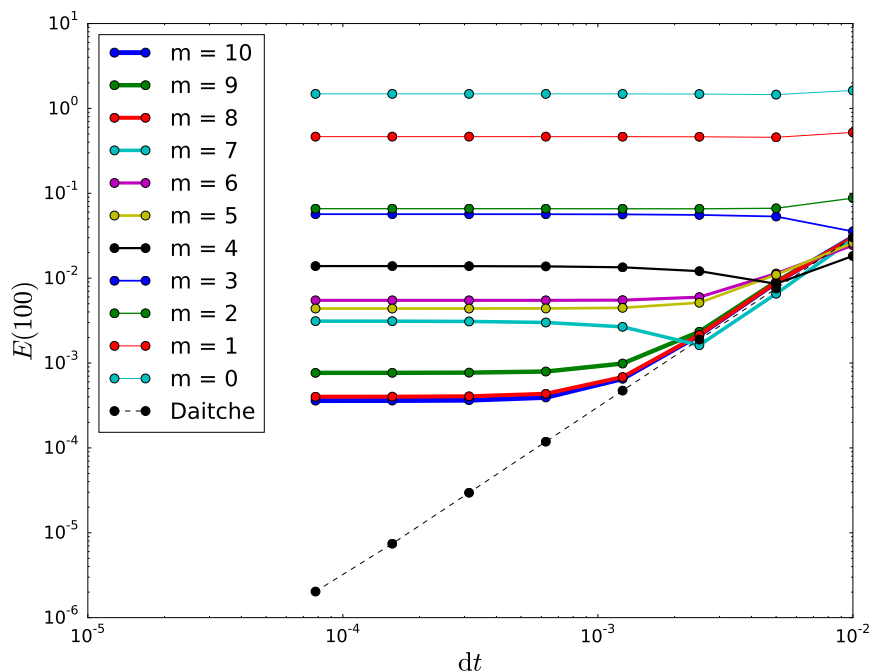


Fig. 3.16 Average relative error of the particle position (final time $t = 100$ s) for different time-steps. Both the raw second-order accurate method of Daitche and the corresponding MAE alternative for different number of exponentials, m . The window time is taken as $t = 0.1$ s

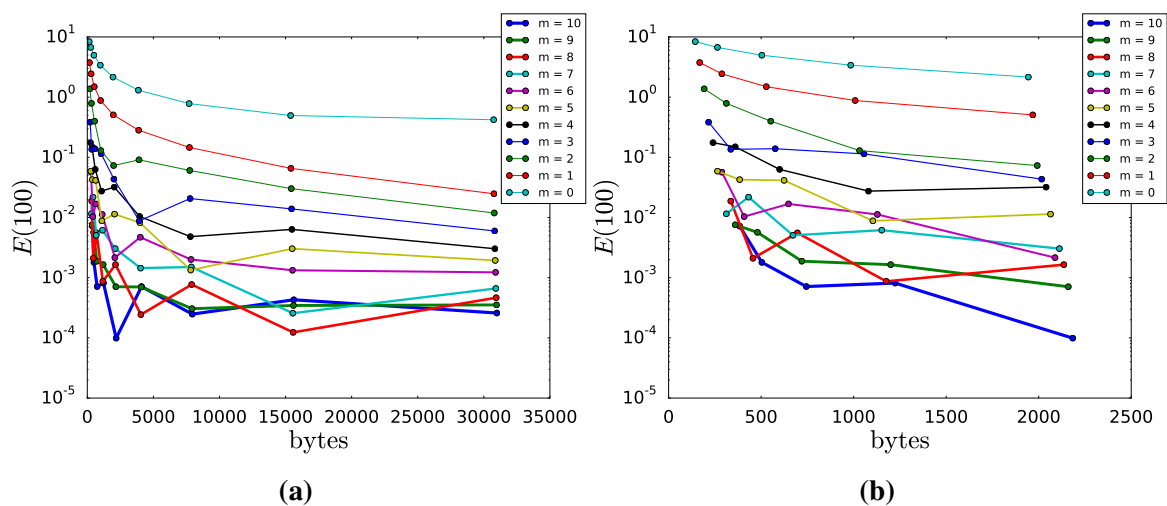


Fig. 3.17 Relative error of the particle position at the final time ($t = 100$ s) for different memory loads per particle. Fig. 3.17b is a zoomed-in version of Fig. 3.17a, showing the initial part with the lesser amount of bytes.

situation only takes place at the start of the simulation, where the initial conditions of the motion are very uncertain anyway ⁶.

Note the appearance of oscillations at the start of the simulation. These seem to correspond to amplifications of the error that are introduced when a new historical value is added to the list, as consistently indicates their constant period. Note that these perturbations take place along the whole trajectory, although they become strongly damped after the initial phase (they are still visible for the larger values of N_q). In any case, the error is always bounded by the error introduced by making the overall time step greater. Note furthermore that the error becomes almost horizontal after the initial phase (joining the reference curve after sufficient time). This indicates that the trajectory of the particle keeps very close to the reference solution ($N_q = 1$), and that the error is mainly accumulated at the start. indeed, any difference in the estimation of the history force would lead to a different rate in the (approximately exponential) outward motion of the spiral, which would in turn lead to upward sloping errors.

It is not as clear how these gains apply when a window method is used. Indeed, let us assume we wish to keep the number of past contributions, n , bounded to ten. To achieve this, we have two parameters to play with: the size of the time window, t_w , and N_q . Making t_w smaller will allow taking smaller quadrature steps, while making N_q bigger will permit increasing t_w and still keep, in both cases, $n = 10$. The increase in N_q brings about a loss in accuracy associated with the quadrature scheme, while a decrease in t_w will require more from the exponential approximation of the kernel which, as we have seen, are a worse approximation nearer the most recent time; thus, some loss of accuracy is expected here too.

The situation suggests the existence of an optimum combination of t_w and N_q , for every given combination of n and m . We explore this possibility in Fig. 3.19, which shows the variation of the relative error at $t = 20$ s as a function of N_q for different numbers of exponentials in the kernel approximation. In each case, the total number of historical data is kept constant at $n = 10$. Note how, as m becomes higher, the optimum slightly tends to move left for most of the curves, which corresponds to smaller t_w sizes (smaller N_q). This is expected, since the greater richness of the exponential kernel makes it better able to approximate the Boussinesq–Basset kernel close to the current time. While all the curves are quite irregular, it is clear that using no sub-stepping at all is the worst option for all but the highest values of m and the largest time steps. Overall it seems much safer to take $N_q = 10$ instead of no substepping at all, especially for the lower m values.

⁶For instance, if the initial conditions are given by the injection of a particle through some mechanical device into the flow, most likely, the presence of such device invalidates the MRE. If, instead, the particle is assumed to have been in the flow before the start of the simulation, then its history is simply unknown and the error will mostly come from this uncertainty in this case.

Clearly, the whole matter of quadrature substepping requires further study, although we have shown the potential of this technique here, especially without the use of a window method. Note that the window method of van Hinsberg et al has only been developed for the Basset kernel and is not a matter of a simple generalization to extend it to other kernels. This is not a limitation for the Daitche method, which can be adapted to other kernels [89], like the ones proposed by Mei and Adrian [243] and Lovalenti and Brady [223], which become more accurate for larger values of Re_p (See Section 2.2.1). Quadrature substepping could certainly be extended to those situations, where its utility would become much more obvious, given the unavailability of the MAE.

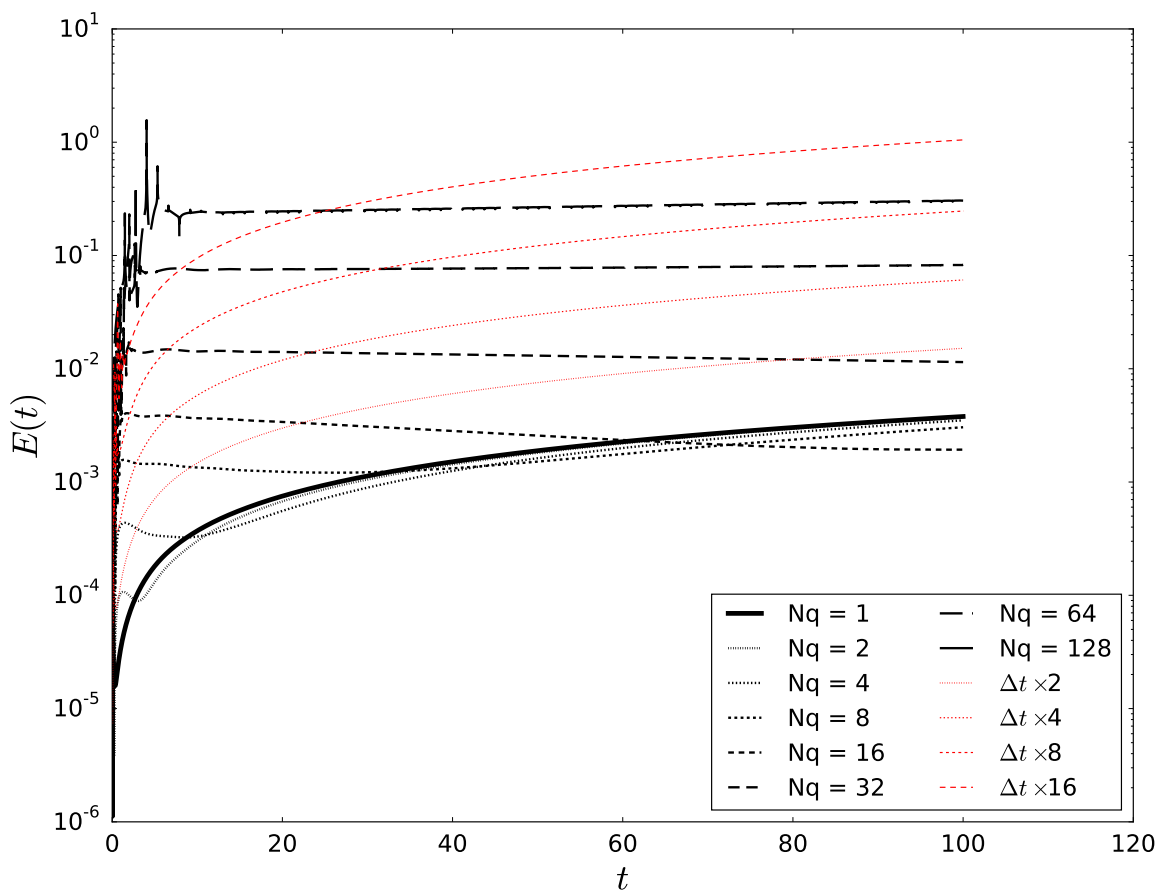


Fig. 3.18 Relative error in the radial coordinate in time using the second-order accurate method of Daitche. The numerical solution with $\Delta t = 0.01$ is represented by a full line. The dashed lines in black represent the results corresponding to successive increases in N_q ; and, in dashed red, the curves corresponding to successive increases in the overall time step Δt while keeping $N_q = 1$ fixed.

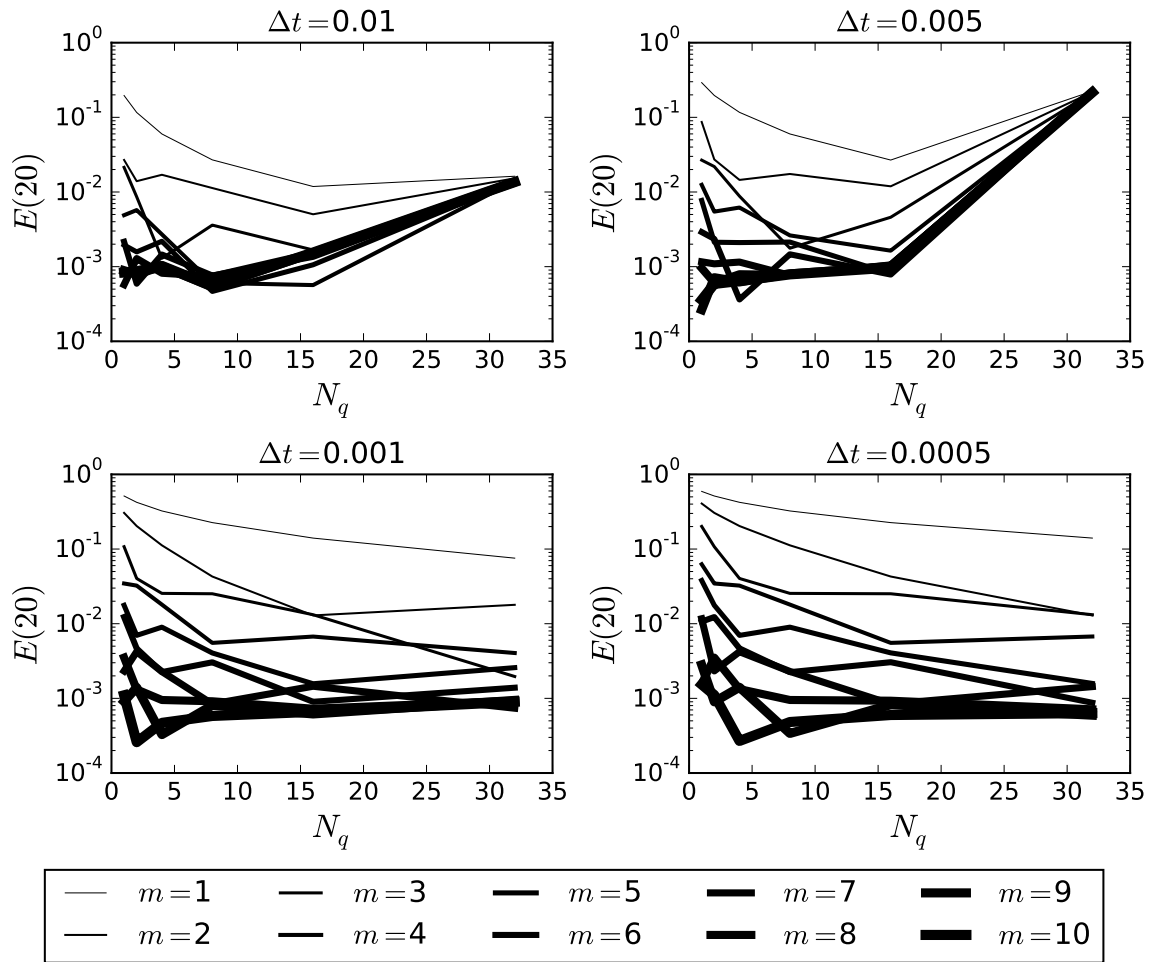


Fig. 3.19 Relative error in the radial coordinate at $t = 20$ for a fixed number of historic points ($n = 10$) and different number of exponential kernel approximants (m). The time step is taken as $\Delta t = 0.005$ s.

Comments about the performance of the coefficients by van Hinsberg et al.

The consideration of a generalized optimization problem, in which the t_i coefficients are considered free to vary, does indeed provide more accurate coefficients. The gains in accuracy are in the order of half an order of magnitude (for the bounds) as compared to the simplified problem, in which the coefficients are set according to the simple rule from Eq. (3.48). We ignore the number of iterations considered or what was the heuristic method employed in [353] to fix the t_i before proceeding with the optimization. But the fact our proposed heuristic yields a very similar error (at least for $m = 10$), as mentioned in 3.3.2, is revealing. It seems that while applying the generalized optimization process improves accuracy, it does so only moderately. We have taken the process up to $m = 10$, and so we cannot be sure this trend still holds beyond that point. However, the distance in accuracy between the two approaches does seem to be quite stable, judging from Fig. 3.6.

The good accuracy of the coefficients by van Hinsberg et al. has been confirmed in the examples considered in Section 3.6.1 and we do not expect this trend to change significantly in other applications. In any case, it is preferable to use the optimal coefficients provided in Appendix F. If the number of exponentials needed was more than ten, using the simplified optimization process (Option C) would be a cost-effective option, perhaps together with the heuristic presented in Eq. (3.48). Due to the accuracy of the MAE, however, in many applications this will not be needed. This is the case of our third benchmark, which we discuss in the following section.

3.6.3 Third benchmark: Sedimentation through synthetic vortices

We now wish to test the efficiency of the full algorithm. For that, we will consider the benchmark test that was considered in Guseva et al. [155] in their work about the influence of the history force on the dynamical properties of a system of sedimenting particles. These authors studied a synthetic bidimensional flow that is a transient variant of the classic cellular flow field previously studied in, e.g., [239]. The flow field is given by ⁷

$$\mathbf{u}(t, x_1, x_2) = \begin{pmatrix} u_1(t, x_1, x_2) \\ u_2(t, x_1, x_2) \end{pmatrix} = U \left(1 + k \sin \left(\omega \frac{U}{L} t \right) \right) \begin{pmatrix} \sin \left(\frac{x_1}{L} \right) \cos \left(\frac{x_2}{L} \right) \\ \cos \left(\frac{x_1}{L} \right) \sin \left(\frac{x_2}{L} \right) \end{pmatrix} \quad (3.73)$$

⁷The flow described in [155] differs from the one described here in that, in it, the arguments of the sine and cosine functions are pre-multiplied by π . We however only managed to obtain similar results to the ones reported in their work upon removing this π factor. We suspected that the difference was simply due to a misprint. In contacting the authors, this possibility was given credibility and thus we have assumed this to be the case. Note that, in particular, this modification changes the meaning of L , which is no longer equal to the diameter of a vortex, but rather 2π times smaller.

where x_i are Cartesian coordinates and U and L are the characteristic velocity and length-scales of the flow; ω controls the frequency of the temporal evolution of the flow and k its amplitude. Such flow covers the Cartesian plane with πL -diameter vortices, each one flowing in the opposite sense than the ones directly next to it.

Guseva and co-workers were interested in studying the long-term evolution of a number of particles subject to the above flow in a double-periodic domain containing four vortices. They monitored 10 000 particles for up to 120 periods employing the first-order accurate version of the Daitche algorithm. They did not extend the simulation duration any longer, since, as they put it: 'A longer interval is not possible to choose due to the numerical cost of recording the history force with small time step for this number of particles'.

We study one of their examples with the MAE and our optimal points and use it as a test for efficiency. The values chosen for the input physical and numerical parameters are summarized in Table 3.1 and are roughly consistent with the nondimensional parameters considered in [155], which correspond to the typical conditions in the phenomenon known as *marine rain*. This term refers to the sedimentation of small agglomerates of mainly organic matter from the surface to the deep ocean, subject to the turbulence found in the upper layers due to the action of wind; see [155] and the references therein. As was done in [155], the effect of Faxén terms will be neglected. Note however that, for the small sized particles considered, such effects are expected to be very small, since

$$\frac{a}{L} \approx 0.004 \ll 1 \quad (3.74)$$

And so the effects of the curvature of the flow field around the particle are expected to be very small, since they are expected to scale as the square of this quotient [238].

Fig. 3.20 shows the position of ten thousand particles suspended in the cellular flow at three different time instants. The particles were initially placed in a uniform lattice covering the whole domain, as was done in [155]. The effect of the history force is apparent: With $\mathbf{F}_H \equiv \mathbf{0}$ the particles become confined to a set of four curves. In contrast, when including all forces, full bands with particles are still visible at the latest stage, although the bands' edges do become more sharply marked. These qualitative trends are consistent with [155].

Let us now turn to the comparison between the second and third rows in Fig. 3.20, which correspond to the Daitche method and the MAE, with the optimum points obtained from the cost function I_1 . No visible effects arise due to the use of the window method at any of the recorded times, taking $m = 10$. Thus, for this example it is enough to keep around thirty integration points in memory per particle (with a time window of 0.1 s) to capture the qualitative behaviour of the system. In fact, Fig. 3.23 shows that even less memory

would have been required for this example. Notice how for $m = 6$, the solution is virtually indistinguishable from the reference, with $m = 10$.

As a consequence, the total computational time is greatly decreased, as is made apparent by Fig. 3.24 (compare the full, black curve to the green curve), where the elapsed time per unit simulated time is represented. The steady increase in memory resources and the number of computations per time step translates into ever higher costs for the Daitche method. The MAE, instead, stabilizes after a few time steps, once the window region is completely filled. Note that in our implementation the cost is of the same order as for the case of neglecting the history force altogether.

Admittedly, our implementation is mounted on top of a 3D discrete element code, where neighbour search has been switched off. The flexibility of the application sacrifices some efficiency and thus one cannot conclude the latter statement to be the case in more optimized implementations. Indeed, more detailed analyses with the help of profilers show that a vast majority of the time spent in computing the hydrodynamic forces is still being spent with the history force. Further analysis and reductions of the history force cost are thus in order. However we have demonstrated that the MAE, as described, already *critically improves the situation*. The perspective of routinely including the Boussinesq–Basset force in the numerical implementation of the Maxey–Riley seems now much closer.

Similarly, Fig. 3.22 shows the effect of using quadrature substepping on the evolution of the same cloud of suspended particles. To simplify the matter, we do not use the window method here. The snapshots correspond to $t = 80$ s, where the changes become most evident. Furthermore, only the top right cell is shown, taking advantage of symmetry, which was preserved in all the studied cases. Note that the overall shape of the different ensembles remain remarkably stable up to $N_q = 40$. At $N_q = 80$ the deterioration of the approximation becomes clearly visible. At $N_q = 160$ the topology of the attractors seems to change and a doubling of the limit lines arises. This corresponds to a quadrature step of 0.72 periods of the cyclic motion.

The effect of quadrature substepping on efficiency is clear: each time the quadrature time step is doubled, the cost of each time step decreases (see Fig. 3.24). Furthermore, the effect is most effective after the first increases, as the memory requirements are reduced enough to relieve the memory bottleneck. A small amount of quadrature substepping can therefore provide a great deal of numerical benefits, while introducing only minimal inaccuracies if it is kept moderate enough. We conjecture that the maximum allowable amount of substepping can be related to the smallest scales in the flow, at least in an order-of-magnitude sense. That is, the smallest time scales must be resolved with a sufficient number of quadrature points

to have an accurate method. More work will however be needed to produce concrete rules, applicable to a wide range of applications.

Table 3.1 Physical parameters considered in the cellular flow example.

Parameter	Value	Description
g	9.81 m/s^2	gravity acceleration
<i>Flow parameters</i>		
U	0.3 m/s	characteristic velocity
L	0.1 m	characteristic length
k	2.72	trans. amplitude parameter
ω	π	trans. frequency parameter
ρ_f	1000 kg/m^3	fluid density
ν	$1 \times 10^{-6} \text{ m}^2/\text{s}$	fluid kinematic viscosity
<i>Particles parameters</i>		
a	$3.9685 \times 10^{-4} \text{ m}$	particle radius
ρ_p	1500 kg/m^3	particle density
W_{ter}	0.172 m/s	terminal settling velocity
<i>Numerical parameters</i>		
h	0.003 s	time step
t_w	0.1 s	time window (for MAE only)
m	10	number of exponentials (for MAE only)

3.7 Summary

In this chapter we discuss the numerical solution of the Maxey–Riley equation, of which we assume the fluid field variables to be available at the particles' locations. Our main interest is the simulation of the large ensembles of suspended particles that commonly arise in the study of particle-laden flows, for which efficiency is paramount. Thus, having identified the Basset–Boussinesq history force as responsible for most of the overall computational cost, we have concentrated on its treatment.

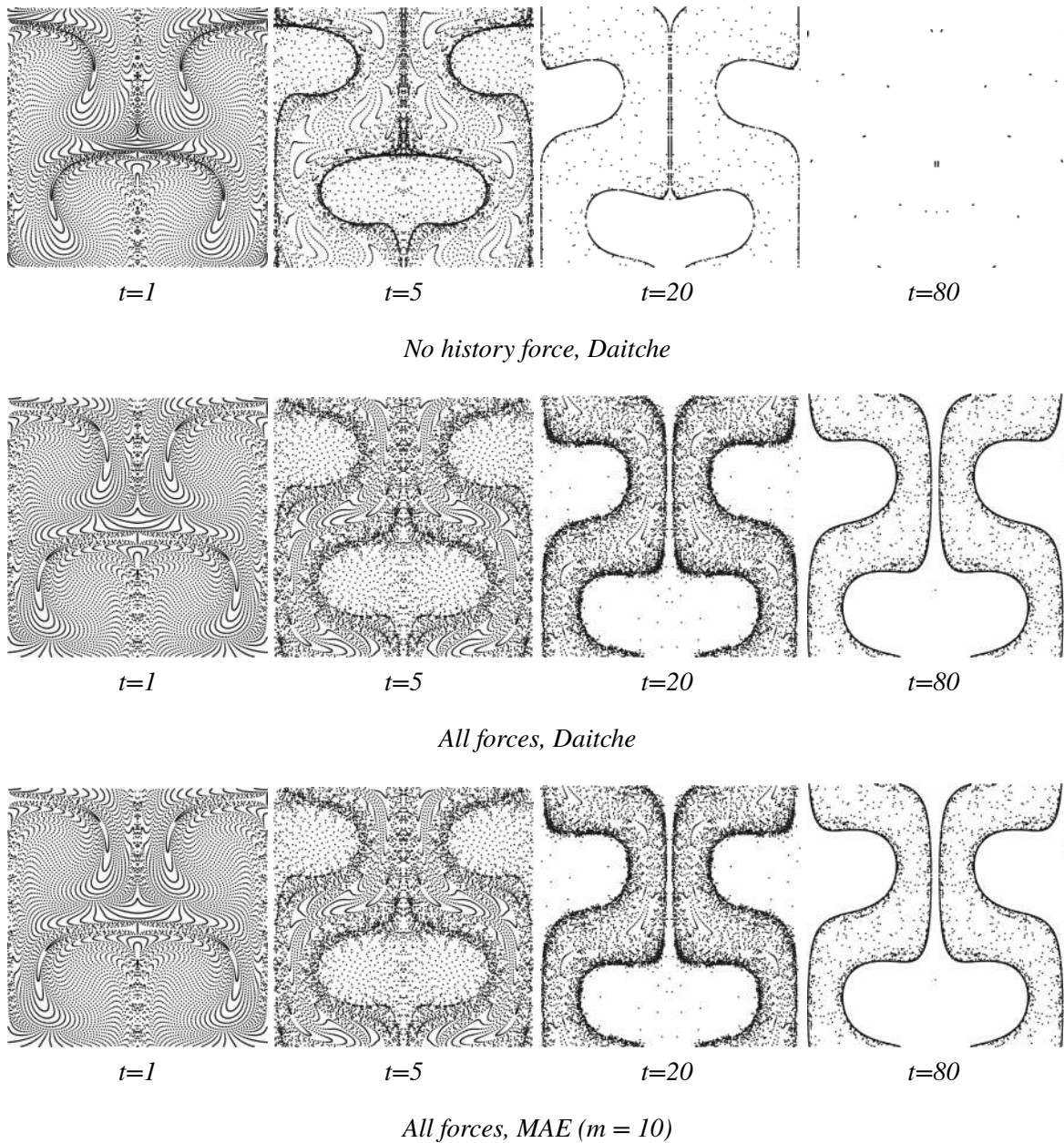


Fig. 3.20 Position of the 10 000 particles at different times in double-periodic spatial representation.

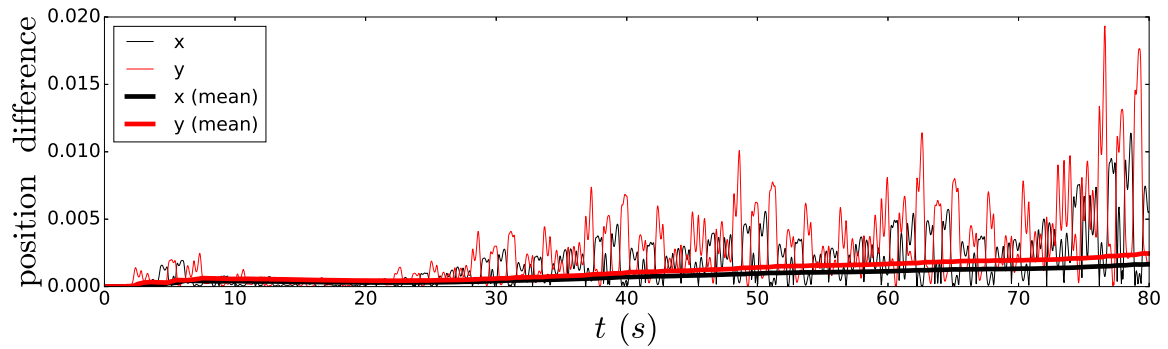


Fig. 3.21 Absolute value of the difference in the x and y components, normalized by L , of the position of a single particle with initial coordinates $(0.20106, 0.20106)$, for the MAE and the Daitche method. The thick curves represent the cumulative averages of these values.

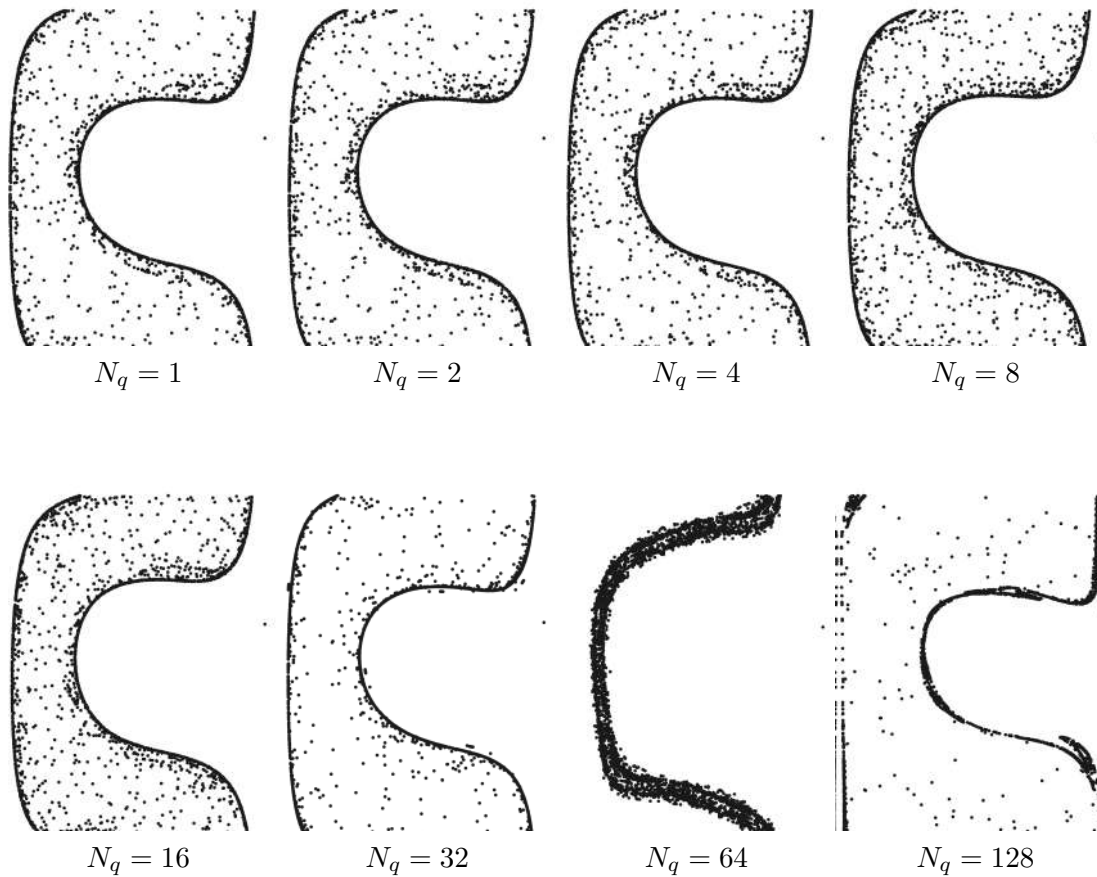


Fig. 3.22 Position of the particles corresponding to the upper-right cell for different quadrature substep sizes at $t = 80$ s. The time step is $\Delta t = 0.003$ s in all cases.

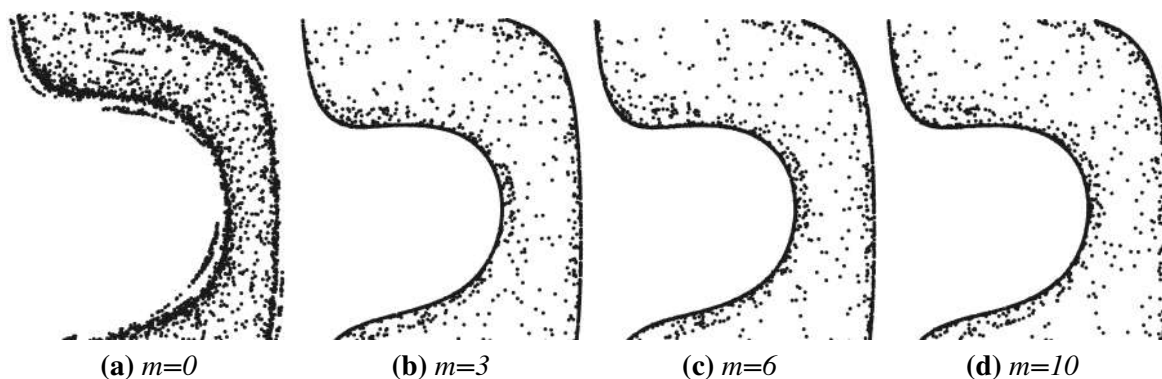


Fig. 3.23 Detail of a quarter of the domain shown in 3.20 at $t = 80$ s for different number of exponentials used to approximate the tail with $t_w = 0.1$ s.

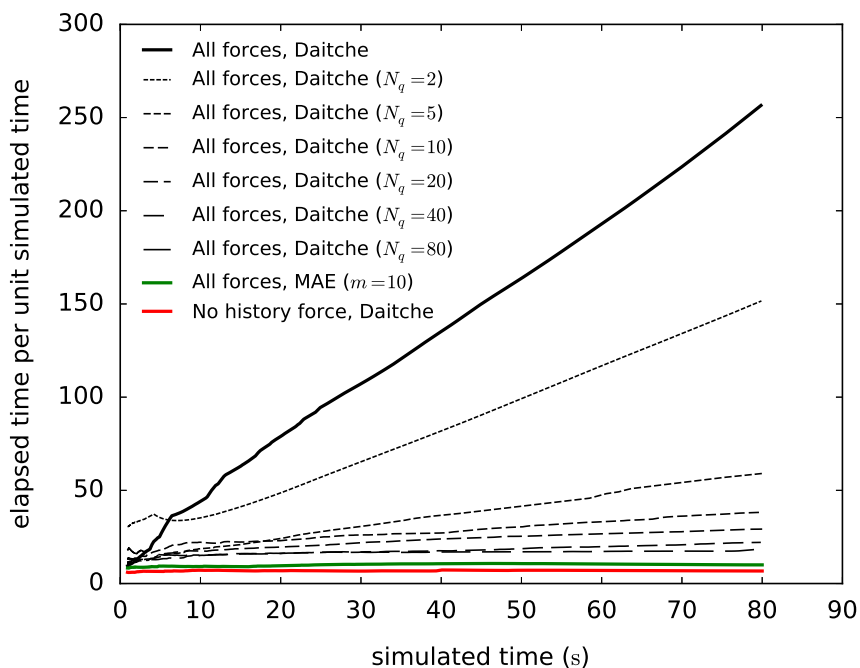


Fig. 3.24 Evolution of the wall-clock time spent in seconds, per unit simulated second, as a function of the simulation duration for the Daitche method, MAE and neglecting the history force. All the runs were performed with the same time step on the same PC (serial implementation). The time step is taken as $\Delta t = 0.003$ s in all cases. The first 100 steps were not taken into account.

We have presented an algorithm that combines the higher order quadrature scheme presented in [89] (Daitche method) with the window method presented in [353] (MAE). This algorithm is designed to minimize the huge memory resources needed to include the history force. This is a common problem encountered in the field of numerical methods for fractional differential equations, due to the intrinsic nonlocality of fractional-order integro-differential operators. This observation motivated the inclusion of the review in Section 3.5, where we establish a connection with parallel research in that field. This had notoriously been missing in the literature dealing with the numerical simulation of the MRE, despite the existing wealth of promising resources in this area.

In order to exemplify the potential of the fractional calculus perspective, we tentatively explore the possibility of introducing Richardson extrapolation to boost its order of approximation in Section 3.5.1. For a fixed final time, this technique works remarkably well with the Grünwald–Letnikov (first-order-accurate) method of Bombardelli et al. [40], yielding slopes in excess of four in the semi-logarithmic error plots. However, in a realistic setting the final time would be constantly being updated and thus the number of history points would change, precluding a literal implementation of the method. Nevertheless, there exists the possibility of using a window method to precisely control the number of points in the window region, although have not pursued this line of research, leaving it for future work.

In Section 3.3 we present several enhancements to the MAE. First, the concept of quadrature substepping is introduced to substantially reduce the number of historical data that must be stored by using different time steps for the quadrature and for the time integration schemes. We generalized the Daitche method to allow for this, of which the corresponding formulas are given in Section 3.3.1. It is shown that a small amount of substepping induces small increases in the numerical error, much less than an equivalent increment of the overall time step. Such strategy might come in very handy when the MAE is not available or lacks accuracy, such as when the history kernel is not of the Boussinesq–Basset form.

Second, we have generalized the original formulation of the MAE by including all the free parameters in the optimization, In the original formulation, half of the parameters (the a_i) were determined by solving an optimization problem, while the other half (the t_i) had to be fixed based on heuristic arguments. In our approach both sets of parameters are treated as unknowns, widening the space of possible solutions. Furthermore, we have considered a range of cost functions, including the norm of the error (I_1) which was replaced by van Hinsberg et al. by a differentiable function (I_{2t}).

It turns out that the resulting problem is significantly more challenging. To address it, We have combined advanced optimization techniques with relatively large computational resources, producing the list of parameters summarized in F and ready for implementation

in particle-laden flow simulations (see Appendix F). By extending the problem, we have introduced a strategy that could inspire future research to continue refining and/or extending our list.

The results show that there is not much to be gained in using the error norm I_1 since it produces results of similar accuracy to that of I_{2t} , a more tractable function. The inclusion of the t_i coefficients in the analysis does improve the accuracy in up to half an order of magnitude (in the optimal error bounds), which is a modest gain that we have shown to depend weakly on the number of exponentials, m .

In any case, using our list of parameters the overall algorithm already delivers a very high performance. We have shown this in several tests, where the method displayed remarkable accuracy. We have demonstrated that the reductions in the memory requirements can consistently be expected to exceed three orders of magnitude, even when the accuracy requirements are strict. As a consequence, the computational cost of including the history force has not changed the order of magnitude of the overall performance of the code in a long simulation with 10 000 particles. We find this result remarkable, since it is an extended view that the inclusion of this force dramatically hinders performance. Furthermore, we have seen that with a very small number of exponentials we may already achieve the necessary accuracy. For example, in the marine rain example, taking $m = 3$ was sufficient to correctly predict the form of the attractors. This fact may become crucial for the viability of many memory-thirsty applications, such as in long-term, large-scale particle-laden flows simulations.

From our work in this chapter we conclude that thanks to the methods we have tested and improved, the possibility of including the history force in practical particle-laden flows is probably close to becoming realistic. The practical implications of this will be particularly relevant to the simulation of liquid-solid flows, were the importance of this force has been proven in a number of studies.

Chapter 4

Forward and backward-coupled particulate flows

4.1 Introduction

The work presented in this chapter has been developed as part of the activities most closely associated to the company's side of the *Doctorat Industrial*. The work has been driven by mixed forces. On the one hand, the pursue of the scientific goal of exploring the possibilities of an FEM-DEM framework, which had been left virtually unexplored in the literature. Indeed, the practical totality of reported simulations of particle-laden flows using the Euler–Lagrange approach model the fluid using one of following techniques: finite volumes[371, 387], lattice-Boltzmann [333] or pseudo-spectral [331] methods.

On the other hand, the need to deal with the constraints of industrial demands strongly shaped the evolution of our application `SwimmingDEMApplication` and the type of problems presented here. As a result, this chapter is mainly application-oriented, with the occasional presentation of related research that has been undertaken to meet the needs arisen during the development of our numerical tool.

The referred applications are thus not always smoothly linked to the theory presented in Chapters 2 and 3. Certainly, we do apply the algorithms developed in Chapter 3 and also draw knowledge from Chapter 2. But industrial demand is rarely fully aligned with one's previous research. As a consequence, for instance, none of the presented applications fall within the range of applicability of the MRE. We do however apply an extended version of the equation, see Section 4.2.

Up to this point, it has always been assumed that the background fluid flow field was known at the particles' positions. In this chapter we will be concerned with how to obtain the

required flow information from the solution of a finite element computation. The field \mathbf{u} in the MRE will therefore be considered an unknown from now on.

Our numerical method is based on the point-particle approach (see e.g. [85]), where the flow disturbance caused by the particles is considered a fine detail and is not resolved by the computational mesh, but at most averaged and spread over the mesh nodes based on a filtering technique (see Section 4.8). The particles are treated as Lagrangian points and their motion is integrated based on some variant of Eq. (4.1) (the MRE). This requires knowledge about the background fluid field (the velocity and perhaps its derivatives) at the location of the particles. Since the fluid is not resolved around the point-particles, the computational mesh needs not be any finer than several times the particles radius, as long as the macroscopic behaviour of the fluid is well-captured; see Fig. 4.1.

Since we use finite elements, the flow velocity field is defined everywhere in the domain and it can be easily calculated as a linear combination of the shape functions. By using a search algorithm very similar to the one employed for the neighbour search [216], every particle is assigned a fluid element, so that only the non-vanishing shape functions need be taken into account in calculating the velocity. The details are presented in Section 4.5. The calculation of the fluid derivatives involves the application of recovery techniques described in Section 4.4.

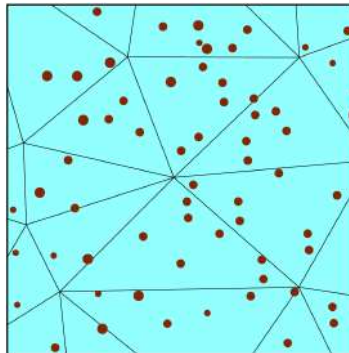


Fig. 4.1 Point-particle approach illustrated. The radius of the particles is only relevant to their contact, not to the fluid interaction, for which only their center point is of relevance.

There are several levels at which the interactions between the different phases or, more generally, *components* [109], could be considered theoretically. Fig. 4.2 shows a conceptual diagram of these interactions. With full line arrows we show the interactions typically considered in the one-way coupled strategy. Those include the forward coupling of the particles to the fluid, whereby the particles are moved according the calculated fluid velocities but the fluid is insensitive to the particles. Optionally, a one-way coupled simulation may also take into account the contacts of the particles with the solid boundaries. These boundaries

often coincide with the Dirichlet boundaries for the fluid problem; i.e. forward fluid-structure interaction (FSI). We also mark the inter-particle contact interactions with full lines, since it is our default setting, although this can be treated separately and is sometimes referred to as four-way coupling [84]. The one way coupled strategy is adopted in the first part of this chapter.

In the second part of the chapter we are concerned with two-way coupled flows, where the movement of the particles affects the fluid phase in two ways. First, the relative motion of the particles with respect to the fluid generates a force that is applied back to the fluid, using a filtering technique. Second, the mass conservation equation is altered to account for the volume occupied by each phase. Both these interactions are indicated with dotted lines in Fig. 4.2.

We have ignored other possible interactions in this work. First, we do not consider the effect of the fluid on the structures (backward FSI) or the self-contacts between structural parts. Second, we have neglected the hydrodynamic interactions that result from the interactions between the sub-scale fluid perturbations and the particle. This type of actions has been termed three-way interaction [84]. It requires the introduction of special methods that take into account the relative position of particles, such as the simple superposition method (see [363] and references therein) and its more sophisticated variants [15, 292] or the expensive (though very accurate) Stokesian dynamics [43, 316], all of which are based on the low particle Reynolds number assumption and most on the Stokes flow assumption. Other, empirical techniques have been proposed to capture a rough estimate of the hydrodynamic interactions, including the case of higher particle Reynolds numbers. Most of these techniques are based on the idea of modifying the drag coefficient using the local average of the particles' volume fraction [31, 378].

Note that it is not clear whether the latter method should be classified as three or two-way coupling, as it involves a calculation of a fluid-average quantity (the volume fraction) that is then projected onto the particles. So, ultimately, it is the fluid field that is being modified if one counts the solid fraction as a variable of the generalized fluid phase. Moreover, this type of modification has the same resolution as the fluid, while the individual particle perturbations remain below this scale. In this sense, it could be seen as two-way coupling. Note that we have already come across this conceptual overlap of the one and three-way coupling interactions in Section 2.2.4. In any case, we do use this technique in our last application example, Section 4.9.

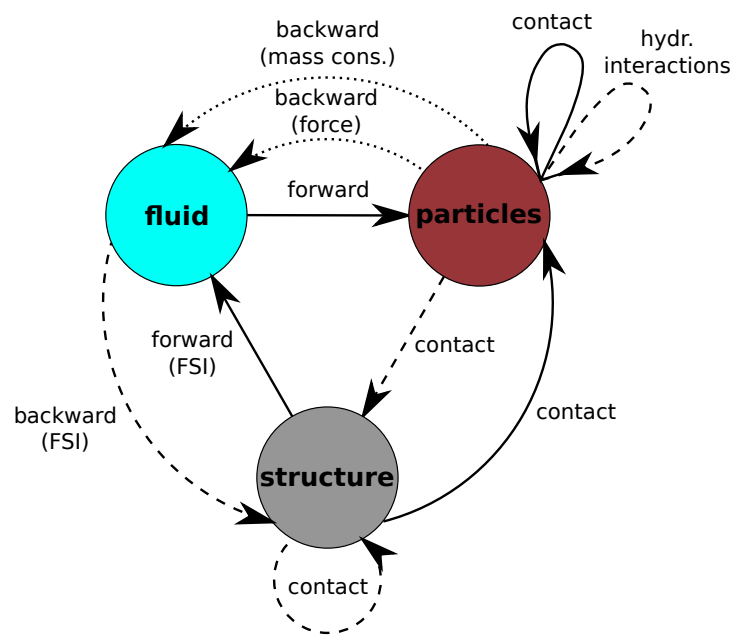


Fig. 4.2 Conceptual diagram of possible interactions in the numerical model. Full lines: interactions in the one-way coupled strategy; dotted lines: interactions specific to the two-way coupled strategy; dashed lines: interactions neglected in the present work. Self-interactions on the particles' phase: the three-way (hydrodynamic interactions) and four-way (contact) interactions.

4.2 Beyond the MRE

As hinted in the introduction, it is possible to modify the MRE to extend its range of applicability beyond that analysed in Chapter 2. For instance, the first-order inertial effects, discussed in Section 2.2.1 may be incorporated, leading to more complicated expressions for the different forces that still fall back to the ones in the MRE in the very low Re_p limit. We mentioned the work of Lovalenti and Brady [223], where such task is undertaken. In [224] the task is generalized to arbitrary time-dependent motions of a sphere in uniform flow, giving the exact formula for the hydrodynamic force to order $O(Re_p)$.

Such analytic models become increasingly complex very quickly and yield increasingly costly and complicated expressions. Furthermore, their generality is often not sufficient and so their practical application is relegated to very specific (though in many cases very relevant) examples. More importantly: their range of application is usually small, as nonlinear effects start to become more and more dominant. For instance, the expressions provided in [224] are only valid for $Re_p < O(1)$. It is for these reasons that we have left their discussion out of the scope of this work. The problem is that there are still many applications where the particle Reynolds numbers (and other nondimensional numbers as well) exceeds the limits studied in Section 2.2 or any of the analytical formulas available.

An alternative to these analytical formulations is provided by *empirical models*. The usual form of these models is analogous to that of the MRE (or its analytical extensions), such that for each term in the latter, there is a modified analogue in the former. The most common way to apply such a modification is to premultiply the original term by a scalar function that is a function of the relevant nondimensional numbers that fall outside the referred range of applicability. The constant parameters are fixed by running experiments that also serve to determine the range of applicability of these expressions. Note that these empirical models are a *new theory* and have to be held as such. That is, they do not logically follow from more fundamental theories (such as the Navier–Stokes equations) just as the MRE equation does. This means that, in particular, it should be checked that

1. Each term in the resulting equation has been validated in its own range of applicability
2. The particular additive decomposition of forces is correct (are the chosen terms mutually compatible?)
3. The uncertainty related to the possible unsteadiness of the flow is compatible with the required accuracy

The first item above is the most obvious: now each force is a model in itself and it should be a validated one. With respect to Item 2, note that the additive combination of the different

forces stems from the mathematical analysis of the problem in the case of the MRE, while here it becomes an assumption. Finally, Item 3 refers to the fact that the chaotic motion of the fluid that arises at higher Reynolds numbers ¹ implies that the deterministic models for the different hydrodynamic forces can only be interpreted as ensemble averages, which means that some dispersion is inherently linked to them and this should be taken into account.

These are not the only issues related to the use of these models. For instance the presence of ambient turbulence [33] and the proximity of neighbouring particles are examples of effects that can play an important role but that are difficult to characterize reliably. The second of these areas has probably received the most attention and is currently a very active area of research, with a large number of new models being continually proposed [31, 378, 99, 317]; see [393, 208] for models proposed to take into account the effect of anisotropy in the distribution of neighbours.

We next summarize the *form* of the expressions of the most relevant extensions for practical purposes (certainly for the present work). We follow closely the most comprehensive exposition that has been given on the subject to date, put forward in a series of reviews by Eric Loth and co-workers [217, 218, 222]. These authors considered the following form of the equation of motion for a small but finite sized particle, ignoring all body forces but the weight, by the equation of the form

$$m_p \frac{d\mathbf{v}}{dt} = \mathbf{F}_{\text{hydr}} + \mathbf{F}_B + \mathbf{F}_{\text{contact}} \quad (4.1)$$

where m_p is the mass of the particle, where $\mathbf{F}_{\text{contact}}$ is the sum of the contact forces due to current particles or solid walls that overlap with the target sphere; and where $\mathbf{F}_{\text{hydro}}$ denotes the hydrodynamic force (discussed below) and \mathbf{F}_B the buoyancy force, given by

$$\mathbf{F}_B = (m_p - m_f)\mathbf{g} \quad (4.2)$$

where m_f is the mass of a volume of fluid displaced by the particle and where the hydrodynamic force is given by the following combination [222]:

$$\mathbf{F}_{\text{hydr}} = \mathbf{F}_U + \mathbf{F}_A + \mathbf{F}_D + \mathbf{F}_H + \mathbf{F}_L \quad (4.3)$$

¹Indeed, above $Re_p = 140$ the wake behind a sphere immersed in uniform flow becomes unsteady, as tested through physical and numerical experiments, see [187, 74] (note that the Reynolds number is based on the diameter in these works).

Furthermore, in some situations it is important to include the rotational degrees of freedom. The angular equation of motion reads

$$I_p \frac{d\Omega_p}{dt} = \mathbf{T}_{\text{hydr}} + \mathbf{T}_{\text{contact}} \quad (4.4)$$

where I_p is the moment of inertia of the particle (a scalar for a sphere, otherwise the whole inertia tensor should be used), where $\mathbf{T}_{\text{contact}}$ is the sum of the moments due to the contact forces (sometimes also moments) between overlapping particles and particles and solid surfaces. The hydrodynamic moment \mathbf{T}_{hydr} is discussed below.

Note that Eqs.(4.1) and (4.4) neglect hydrodynamic interactions between neighbouring particles, the aforementioned three-way coupling. For possible extensions to include these effects see [15, 292]. The different forces and torques above are next detailed.

4.2.1 Unperturbed fluid and added mass forces

The validity of the MRE requires that the particle Reynolds number is smaller than one, as we have seen in Chapter 2. At the opposite extreme, in the inviscid limit, valid for vanishing viscosity, the equation of Auton et al. [14] applies, which is in the form Eq.(4.1) (ignoring the contact terms). In this case the hydrodynamic force reads

$$\mathbf{F}_{\text{hydr}} = \mathbf{F}_U + \mathbf{F}_A + \mathbf{F}_L \quad (4.5)$$

where the forces with the same subindices as in Eq. (2.1) denote forces with the same physical meaning.

Let us now discuss the particular expressions of each of the first two terms above. First, \mathbf{F}_U has exactly the same form in both limits. Its physical meaning is clear: what is the force that the sphere of fluid displaced by the particle would feel if taken as a point-mass? It would correspond to its mass, m_f , multiplied by its acceleration, that is, the material acceleration of the background fluid field, as measured at the center of the sphere. Thus, this force reads

$$\mathbf{F}_U = m_f \frac{D\mathbf{u}}{Dt} \quad (4.6)$$

where the capitalized derivative operator denotes the material derivative. Loth [219] argues that the fact that the expression of \mathbf{F}_U is valid in these two regimes confirms that the expression should be robust generically, and so we will assume so.

Similarly, with respect to the added mass force, the following expression is valid both in the inviscid and the vanishing particle Reynolds number limit

$$\mathbf{F}_A = \frac{1}{2}m_f \left(\frac{D\mathbf{u}}{Dt} - \frac{d\mathbf{v}}{dt} \right) \quad (4.7)$$

Moreover, there is considerable evidence of its accuracy outside the theoretical range of validity [361, 194]. We thus assume this form of the added mass force to be valid for the complete range of Reynolds numbers too.

4.2.2 Drag force

The drag force can be defined as the ensemble-averaged force experienced by a particle submerged in a statistically stationary flow in the direction of the relative velocity between the particle and the far-field averaged flow velocity. It can be expressed as

$$\mathbf{F}_D = \frac{1}{2}A_p C_D \|\mathbf{w}\| (\mathbf{w}) \quad (4.8)$$

where A_p is the cross-sectional area of the particle for a section orthogonal to where the drag coefficient is in general dependent on Re_p , the shape of the particle [218], the local solid fraction [31] and the properties of the fluid (see Section 4.7.3). This formulation is trivially a generalization of the Stokes drag, for which $C_D = 12/Re_p$.

4.2.3 History force

The history force can be defined as the transient component of the viscous force that is parallel to the drag force [223, 222]. This force is associated to the transient development of the flow field in the vicinity of the particle. It depends on the particle's relative velocity history, because of the finite-speed of vorticity to convect from the particle surface and away from it. A generic form of this force, which covers the analytically derived form of \mathbf{F}_H in the MRE and also the empirical model of Mei and Adrian [243], of widespread popularity and valid for particle Reynolds numbers up to about 200; is given by

$$\mathbf{F}_H = 6a^2 \sqrt{\pi \rho_f \mu} \left(\frac{d}{dt} \int_{t_0}^t K(t - \tau) \mathbf{w}(\tau) d\tau - K(t - t_0) \mathbf{w}(t_0) \right) \quad (4.9)$$

where we have used the result proved in Appendix C to take the derivative outside the integral sign. Note that for $K(t) = t^{-1/2}$ we recover the MRE expression for the history force (with no Faxén terms).

4.2.4 Lift force

For flows in which the shear dominates the rigid solid motion ². A generalized formula for the lift force due to shear in the flow is given by [219]

$$\mathbf{F}_{L,\text{shear}} = C_{L,\text{shear}}(Re_p)a^2 \sqrt{\frac{\mu_f \rho_f}{\|\nabla \times \mathbf{w}\|}} (\nabla \times \mathbf{w}) \times \mathbf{w} \quad (4.10)$$

Similarly, for spin-induced flow one has

$$\mathbf{F}_{L,\text{spin}} = C_{L,\text{spin}}(Re_p, \Omega_p^*)a^3 \rho \Omega_p \times \mathbf{w} \quad (4.11)$$

where the nondimensional particle angular velocity is defined as

$$\Omega_p^* := \frac{a}{\|\mathbf{w}\|} \Omega_p \quad (4.12)$$

The linear addition of the two effects (lift due to shear and lift due to spin) has been seen to be reasonable for value as high as $\Omega_p^* \leq 0.4$ and $Re_p \leq 100$ [219].

4.2.5 Torque

If the particle angular degrees of freedom are to be tracked (to calculate the lift, for example), then the hydrodynamic torque must be accounted for in order to predict the angular velocity. Following [34], the expression of such torque can be approximated by

$$\mathbf{T}_{\text{hydr}} = 8\pi a^3 \mu \left(\frac{1}{2} C_{T,\text{vort}}(Re_\Omega) \nabla \times \mathbf{w} - C_{T,\Omega_p}(Re_\Omega) \Omega_p \right) \quad (4.13)$$

where $C_{T,\text{vort}}$ and C_{T,Ω_p} are empirical coefficients that tend to one for $Re_p \ll 1$, recovering the analytical solution of Rubinow and Keller [294], valid in this range. The validity range of this linear addition of effects has been proved to be accurate for $Re_p < 20$ when the particle's angular velocity does not differ too much from the equilibrium velocity, i.e. $\Omega_p \approx 2\nabla \times \mathbf{u}$ [219].

²Note that, according to Section 2.2.1, there is a distinction, for small values of Re_p between vortex-induced lift (due to the flow's solid body rotation) and shear-induced lift. Here we omit this distinction for simplicity given that, in most flows, the shear-induced component is much more important [219]. Furthermore, the distinction disappears at the inviscid limit [13].

4.3 The continuous-phase problem

Let us begin by describing the problem corresponding to the continuous phase when considered uncoupled to the disperse phase. For the moment, we restrict our attention to incompressible, Newtonian fluids, such as air or water at low speeds. Their motion is accurately modelled by the Navier–Stokes equations:

$$\left. \begin{aligned} \rho_f \frac{\partial \mathbf{u}}{\partial t} + \rho_f (\mathbf{u} \cdot \nabla) \mathbf{u} - \mu \nabla^2 \mathbf{u} + \nabla p = \mathbf{f} \\ \nabla \cdot \mathbf{u} = 0 \end{aligned} \right\} \text{in } \Omega \times [0, T) \quad (4.14a)$$

$$(4.14b)$$

where \mathbf{u} is the fluid velocity, p is the pressure and \mathbf{f} is an external body force (for example, gravity). As in the previous chapters, ρ_f is the fluid density and μ its viscosity (both of which we assume to be constant). The corresponding initial, Dirichlet and Neumann boundary conditions are given by

$$\mathbf{u} = \mathbf{u}_0 \quad \text{in } \Omega \times \{0\} \quad (4.15a)$$

$$\mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma_D \times [0, T) \quad (4.15b)$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_N \times [0, T) \quad (4.15c)$$

where \mathbf{u}_0 must fulfil Eq. (4.14b), where the domain's boundary is partitioned as $\partial\Omega = \Gamma_D \cup \Gamma_N$, $\Gamma_D \cap \Gamma_N = \emptyset$, with \mathbf{n} the exterior unitary normal vector on Γ_N ; where \mathbf{t}_N is the imposed surface traction and where the Cauchy stress tensor is defined as

$$\boldsymbol{\sigma} = -p\mathbf{I} + \boldsymbol{\tau} \quad (4.16)$$

with $\boldsymbol{\tau}$ and

$$\boldsymbol{\tau} = \mu \nabla^2 \mathbf{u} \quad (4.17)$$

where $\boldsymbol{\tau}$ is the deviator stress tensor

4.3.1 Variational form of the problem

Let us present the variational (weak) form of Eq. (4.15), from which the FEM is derived. In order to do so it is convenient to fix some notation. The space of square integrable real functions in Ω is denoted by $L^2(\Omega)$ and the space of functions whose first derivative is square integrable is denoted by $H^1(\Omega)$. Let $H_0^1(\Omega)$ be the subspace of functions in $H^1(\Omega)$ vanishing on $\partial\Omega$ (the boundary of Ω). The vector counterparts of these spaces are denoted with bold characters. For example, $\mathbf{H}_0^1(\Omega)$ is the space of vector variables with

zero trace at the boundary. In this context, the dimension of the vectors is given by n_{dim} . Furthermore, let $\mathbf{H}_D^1(\Omega)$ be the space of functions in $\mathbf{H}^1(\Omega)$ that fulfil the Dirichlet condition Eq. (4.15b). Finally, we are interested in time varying functions, and so we define, for a generic space $\mathcal{A}(\Omega)$, $L^2(0, T; \mathcal{A}(\Omega))$ as the space of functions such that, restricted to a fixed time in $[0, T]$, are members of $\mathcal{A}(\Omega)$ and, restricted to a fixed point in Ω , square integrable time functions in $[0, T]$. With these notations, the weak version of the problem formed by Eq. (4.14) and Eq. (4.15) is to find $(\mathbf{u}; p) \in \mathcal{X} := \mathcal{V}_D \times \mathcal{Q}$ with $\mathcal{V}_D := L^2(0, T; \mathbf{H}_D^1(\Omega))$ and $\mathcal{Q} := L^2(0, T; L^2(\Omega) \setminus \mathbb{R})$ such that

$$\begin{aligned} \rho_f \left(\frac{\partial \mathbf{u}}{\partial t}, \mathbf{v} \right) + \rho_f (\mathbf{u} \cdot \nabla \mathbf{u}, \mathbf{v}) + \mu (\nabla \mathbf{u}, \nabla \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) &= \langle \mathbf{f}, \mathbf{v} \rangle \\ (q, \nabla \cdot \mathbf{u}) &= 0 \end{aligned} \quad (4.18)$$

for all $(\mathbf{v}; q)$ in $\mathcal{Y} := \mathcal{V}_0 \times \mathcal{Q}$, where $\mathcal{V}_0 := \mathbf{H}_0^1(\Omega)$; where (\cdot, \cdot) denotes the inner product in $L^2(\Omega)$; and where $\langle \cdot, \cdot \rangle$ the integral of the product of the two functions, defined in the spaces where it makes sense (normally \mathbf{f} is assumed to live in the topological dual of $H_0^1(\Omega)$).

The basic strategy in the FEM is to replace the relevant (infinite-dimensional) spaces of functions above with finite dimensional counterparts in the variational version of the problem, leading to the algebraic system of equations that must be solved computationally. In this work we wish to use the simplest linear simplex elements, both for the pressure and for the velocity. However, not all element combinations lead to converging numerical methods for mixed problems and, in particular, a necessary condition to achieve it is that the particular combination of finite element spaces fulfil the *inf-sup* or Ladyzhenskaya-Babusška-Brezzi (LBB) condition. This condition can be stated as

$$\inf_{q \in \mathcal{Q}, q \neq 0} \sup_{\mathbf{v} \in \mathcal{V}, \mathbf{v} \neq \mathbf{0}} \frac{(q \nabla \cdot \mathbf{v})}{\|q\|_{\mathcal{Q}} \|\mathbf{v}\|_{\mathcal{V}}} \geq C \quad (4.19)$$

where \mathcal{V} and \mathcal{Q} are the spaces containing the velocity and pressure solutions respectively and C is a positive constant. In particular, the equal-order, piecewise linear spaces for the velocity and for the pressure ($\mathcal{P}_1/\mathcal{P}_1$ element) do not fulfil this condition [119]. However, one can resort to stabilization methods to allow for this by modifying the weak form of the problem. We delay the presentation of the weak version of the equations above from which the finite element systems of equations are derived, which is instead presented in Section 4.8 and is not essentially relevant at this point. We thus jump to the discretization associated with our finite element strategy, since the phase coupling is based on it.

4.3.2 VMS-stabilized finite element formulation

Variational multiscale (VMS) methods [174, 77] provide a theoretical framework for the development of stabilized finite element formulations. They are based on the explicit consideration of the decomposition of the continuous solution into a part belonging to the finite element space and its complement in the continuous solution space, or *subscale*. The general strategy consists in modifying the standard FEM solution by taking into account the *effects* of the subscales in the solution, in a way that is reminiscent of LES turbulence models, where the unresolved scales provide terms that alter the large-scales flow, even though these scales remain unresolved [82]. In this subsection we introduce the essential ingredients involved in the formulation of a VMS-stabilised FEM method. This formulation is used for the simulation of the one-way coupled fluid simulations, but is also relevant to the development of the two-way coupled equation discussed in Section 4.8, after modification of the appropriate terms.

Finite element essentials

Let us consider a conforming finite element partition \mathcal{T}_h of the domain Ω . For each element in the domain $\Omega_e \in \mathcal{T}_h$ we denote its diameter as h_e and we define $h = \max \{h_e | \Omega_e \in \mathcal{T}_h\}$. With these tools it is possible to construct the finite element spaces in the usual way, as $\mathcal{X}_h = \mathcal{V}_{D,h} \times \mathcal{Q}_h$, with $\mathcal{V}_{D,h} \subset \mathcal{V}_D$, $\mathcal{Q}_h \subset \mathcal{Q}$. The finite element solution will be a function $U_h = [\mathbf{u}_h, p_h] \in \mathcal{X}_h$. Since we will be using equal-order spaces for the velocity components and for the pressure, the solution can be expressed as (summation is assumed for repeated indices)

$$u_{h,j} = N^b U_j^b, \quad p_h = N^b P^b \quad (4.20)$$

for $j = 1, \dots, n_{\text{dim}}$ and $b = 1, \dots, n_{\text{nodes}}$; where the N^b are the shape functions, n_{dim} is the number of space dimensions (2 or 3) and n_{nodes} is the total number of mesh nodes.

Subscale decomposition for a general convection-diffusion-reaction system

Eq.(4.18) is nonlinear due to the convective term, and it will need to be linearised. Let us thus next consider the already linearised version of the Navier–Stokes system, assuming that the convective velocity is a given \mathbf{a} . In practice, \mathbf{a} will be taken as its value from the previous iteration, but leaving the formulation unspecified has the advantage of allowing the consideration of different definitions of this convective term; see [82]. The formulation presented next is not new, but we need to describe it in sufficient detail and in a way that is general enough to include the formulation that is later presented for the two-way coupled

flow. We are interested in solving the following problem given by

$$\mathbf{M} \frac{\partial}{\partial t} \mathbf{U} + \mathcal{L}(\mathbf{U}) = \mathbf{F} \text{ in } \Omega \times [0, T) \quad (4.21)$$

together with the boundary conditions from Eq.(4.15), and where

$$\mathcal{L}(\mathbf{U}) := \frac{\partial}{\partial x_i} (\mathbf{A}_i \mathbf{U}) - \frac{\partial}{\partial x_i} \left(\mathbf{K}_{ij} \frac{\partial \mathbf{U}}{\partial x_j} \right) + \mathbf{S} \mathbf{U} \quad (4.22)$$

where the usual summation convention is implied upon repeated indices.

For the particular case of the (linearised) Navier–Stokes equations, the different vectors and matrices are defined as

$$\mathbf{U} := \begin{bmatrix} \mathbf{u} \\ p \end{bmatrix}, \quad \mathbf{F} := \begin{bmatrix} \mathbf{f} \\ 0 \end{bmatrix} \quad (4.23)$$

and the matrices \mathbf{M} , \mathbf{A}_i , \mathbf{K}_{ij} and \mathbf{S} are $n_{unk} \times n_{unk}$ matrices ($n_{unk} = d + 1$) which, for the case $d = 3$ are defined as:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_i = \begin{bmatrix} a_i & 0 & 0 & \delta_{i1} \\ 0 & a_i & 0 & \delta_{i2} \\ 0 & 0 & a_i & \delta_{i3} \\ \delta_{i1} & \delta_{i2} & \delta_{i3} & 0 \end{bmatrix}, \quad (4.24)$$

$$\mathbf{K}_{ij} = \nu \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{S} = \mathbf{0} \quad (4.25)$$

where, as is conventionally done, we have divided through by the fluid density and redefined the pressure as itself divided by the density. Assuming the fluid density is constant, the variable disappears from the formulation.

The weak form of the problem can in turn be expressed as *find* $\mathbf{U} := (\mathbf{u}; p) \in \mathcal{X}$ such that $\forall \mathbf{V} := (\mathbf{v}; q) \in \mathcal{Y}$

$$\int_{\Omega} \mathbf{V}^{\top} \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U} + \mathcal{L}(\mathbf{U}) - \mathbf{F} \right) d\Omega = 0 \quad (4.26)$$

which is equivalent to

$$\int_{\Omega} \mathbf{V}^{\top} \mathbf{M} \frac{\partial}{\partial t} \mathbf{U} \, d\Omega + \int_{\Omega} \mathbf{V}^{\top} \frac{\partial}{\partial x_i} (\mathbf{A}_i \mathbf{U}) \, d\Omega - \int_{\Omega} \mathbf{V}^{\top} \frac{\partial}{\partial x_i} \left(\mathbf{K}_{ij} \frac{\partial \mathbf{U}}{\partial x_j} \right) \, d\Omega + \int_{\Omega} \mathbf{V}^{\top} \mathbf{S} \mathbf{U} \, d\Omega = \int_{\Omega} \mathbf{V}^{\top} \mathbf{F} \, d\Omega \quad (4.27)$$

Adding all the terms above and taking into account that the Neumann boundary terms must add up to the known traction on the boundary times the test function (since they arise from integrating by parts an expression of the divergence of the stress tensor), we obtain the following expression for the weak form of the problem

$$\left(\mathbf{M} \frac{\partial \mathbf{U}}{\partial t}, \mathbf{V} \right) + B(\mathbf{U}, \mathbf{V}) = L(\mathbf{V}) \quad \forall \mathbf{V} \in \mathcal{X} \quad (4.28)$$

where

$$B(\mathbf{U}, \mathbf{V}) := (\mathbf{a} \cdot \nabla \mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) + (\nabla \cdot \mathbf{u}, q) + \nu (\nabla^S(\mathbf{u}), \nabla \mathbf{v}) \quad (4.29)$$

and

$$L(\mathbf{V}) := (\mathbf{f}, \mathbf{v}) + \int_{\Gamma_N} \mathbf{t}_N \cdot \mathbf{v} \, d\Gamma \quad (4.30)$$

As mentioned above, the direct consideration of equal-order finite element spaces to approximate the spaces of the velocity and the pressure in Eq.(4.27), would result in an unstable method. We must therefore modify the problem somehow and the VMS method provides an effective answer. The method starts by considering a decomposition of the solution space as $\mathcal{X} = \mathcal{X}_h \oplus \tilde{\mathcal{X}}$, where \mathcal{X}_h is the FE space and $\tilde{\mathcal{X}}$ can be any space to complete the FEM space in \mathcal{X} . We assume that functions in $\tilde{\mathcal{X}}$ already vanish on the Dirichlet boundary (the finite element approximation is perfect there), so that $\tilde{\mathcal{X}} \approx \tilde{\mathcal{X}}_0$. $\tilde{\mathcal{X}}_0$ is called the space of *subgrid scales or subscales*. Note that Eq.(4.26) is equivalent to find $\mathbf{U}_h \in \mathcal{X}_h$ and $\tilde{\mathbf{U}}_h \in \tilde{\mathcal{X}}_0$ such that

$$\begin{aligned} \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + \left(\mathbf{M} \frac{\partial}{\partial t} \tilde{\mathbf{U}}, \mathbf{V}_h \right) + B(\mathbf{U}_h, \mathbf{V}_h) + B(\tilde{\mathbf{U}}, \mathbf{V}_h) &= \\ &= L(\mathbf{V}_h) \quad \forall \mathbf{V}_h \in \mathcal{X}_{h,0} \end{aligned} \quad (4.31)$$

$$\begin{aligned} \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \tilde{\mathbf{V}} \right) + \left(\mathbf{M} \frac{\partial}{\partial t} \tilde{\mathbf{U}}, \tilde{\mathbf{V}} \right) + B(\mathbf{U}_h, \tilde{\mathbf{V}}) + B(\tilde{\mathbf{U}}, \tilde{\mathbf{V}}) &= \\ &= L(\tilde{\mathbf{V}}) \quad \forall \tilde{\mathbf{V}} \in \tilde{\mathcal{X}}_0 \end{aligned} \quad (4.32)$$

by applying Stokes' theorem to the elemental volumes, recognizing the continuity of the exact tested functions and neglecting the contribution of the inter-element boundary integrals that

appear when applying the theorem to $B(\mathbf{U}, \mathbf{V})$, the equations above can be replaced by

$$\begin{aligned} (\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}, \mathbf{V}_h) + (\mathbf{M} \frac{\partial}{\partial t} \tilde{\mathbf{U}}, \mathbf{V}_h) + B(\mathbf{U}_h, \mathbf{V}_h) + \sum_e \int_{\Omega_e} \tilde{\mathbf{U}}^\top \mathcal{L}^*(\mathbf{V}_h) d\Omega = \\ = L(\mathbf{V}_h) \quad \forall \mathbf{V}_h \in \mathcal{X}_{h,0} \end{aligned} \quad (4.33)$$

and

$$\begin{aligned} \sum_e \int_{\Omega_e} \tilde{\mathbf{V}}^\top \mathbf{M} \frac{\partial}{\partial t} \tilde{\mathbf{U}} d\Omega + \sum_e \int_{\Omega_e} \tilde{\mathbf{V}}^\top \mathcal{L}(\tilde{\mathbf{U}}) d\Omega = \\ = \sum_e \int_{\Omega_e} \tilde{\mathbf{V}}^\top (\mathbf{F} - \mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h - \mathcal{L}(\mathbf{U}_h)) d\Omega \quad \forall \tilde{\mathbf{V}} \in \tilde{\mathcal{X}}_0 \end{aligned} \quad (4.34)$$

where the formal adjoint of operator \mathcal{L} , denoted by \mathcal{L}^* , can easily be computed by transposing matrices \mathbf{A}_i , \mathbf{K}_{ij} and \mathbf{S} and multiplying the odd-order derivative terms by -1 [120].

Now, a key insight is to realize that Eq.(4.34) implies there exists a $\mathbf{V}_{h,\text{orth}} \in \tilde{\mathcal{X}}_0^\perp$, where the \perp super-index indicates orthogonal complementarity of the space with respect to the L_2 -norm; such that

$$\left(\mathbf{M} \frac{\partial}{\partial t} + \mathcal{L} \right) \tilde{\mathbf{U}} = \mathbf{R}_h + \mathbf{V}_{h,\text{orth}} \quad \text{in } \Omega_e \times [0, T] \quad (4.35)$$

for every elemental volume Ω_e in the domain; where we have defined $\mathbf{R}_h := \mathbf{F} - \mathcal{L}(\mathbf{U}_h)$. The objective at this point is to provide an approximate solution to Eq. (4.35) in terms of $\mathbf{V}_{h,\text{orth}}$. The resulting expression is then to be introduced in Eq. (4.33) to produce a stabilized method. Different expressions for $\mathbf{V}_{h,\text{orth}}$ produce different variants of the VMS, as we show below.

There are several options. One is to provide a finite difference approximation to the dynamic term of the subscales above. The particular time discretization is not essential and it certainly does not need to coincide with the overall problem's time discretization, which is discussed below. Here we use a backward Euler strategy, although more general schemes are considered elsewhere [78]. Introducing this time discretization into Eq. (4.35) yields

$$\left(\frac{1}{\Delta t} \mathbf{M} + \mathcal{L} \right) \tilde{\mathbf{U}}^{n+1} = \mathbf{R}_h + \frac{1}{\Delta t} \mathbf{M} \tilde{\mathbf{U}}^n + \mathbf{V}_{h,\text{orth}} \quad \text{in } \Omega_e \times [0, T] \quad (4.36)$$

The equation above is then approximately solved by assuming that the operator \mathcal{L} can be approximated by a linear transformation represented by an invertible matrix $\tau_{\Delta t, e}^{-1}$ within the

element Ω_e . In this case an expression for \tilde{U}^{n+1} is given explicitly by

$$\tilde{U}^{n+1} \approx \tau_{\Delta t, e} \left(\mathbf{R}_h + \frac{1}{\Delta t} \mathbf{M} \tilde{U}^n + \mathbf{V}_{h, \text{orth}} \right) \quad (4.37)$$

The design of the matrix $\tau_{\Delta t, e}^{-1}$ is one of the most relevant distinctive features of different stabilization methods. Here we adopt the model proposed in [78], where it is motivated by Fourier analysis of the subscales. Most importantly, optimal convergence can be proved for the present problem under the usual existence and smoothness assumptions (see [77, 78]). The stabilization matrix for the Navier–Stoke problem is given by

$$\tau_{\Delta t, e} = \text{diag}(\tau_{1, t} \mathbf{I}_{n_{\text{dim}}}, \tau_2) \quad (4.38)$$

where

$$\tau_{1, t} := \left(\frac{1}{\Delta t} + \frac{1}{\tau_1} \right)^{-1} \quad (4.39)$$

where

$$\tau_1 = \frac{1}{\sqrt{(c_1 \frac{\nu}{h^2})^2 + (c_2 \frac{\|\mathbf{a}\|}{h})^2}} \quad (4.40)$$

$$\tau_2 = \frac{h^2}{c_1 \tau_1} \quad (4.41)$$

and where the constants c_1 and c_2 are numerical constants that are often taken to be $c_1 = 4, c_2 = 2$ [82].

Determination of the subscale model

Our approximate equation for the subscale Eq. (4.37) depends on the value of $\mathbf{V}_{h, \text{orth}}$. This function is unknown and its determination depends in the specific decomposition $\mathcal{X} = \mathcal{X}_h \oplus \tilde{\mathcal{X}}$ chosen. There are many possibilities, but the specific choice $\tilde{\mathcal{X}} = \mathcal{X}_h^\perp \cap \mathcal{X}$ (*space of orthogonal subscales*) leads, after a number of approximations (see [78]) to $\mathbf{V}_{h, \text{orth}} \in \mathcal{X}_h$ and $\tilde{U} \in \mathcal{X}_h^\perp$. The latter condition means that $\forall \mathbf{V}_h \in \mathcal{W}_h \left(\tilde{U}, \mathbf{V}_h \right) = 0$. Using Eq. (4.37) this condition implies that

$$\sum_e \left(\tau_{\Delta t, e} \left(\mathbf{R}_h + \frac{1}{\Delta t} \mathbf{M} \tilde{U}^n \right), \mathbf{V}_h \right) + \sum_e \left(\tau_{\Delta t, e} \mathbf{V}_{h, \text{orth}}, \mathbf{V}_h \right) = 0 \quad \forall \mathbf{V}_h \in \mathcal{W}_h \quad (4.42)$$

or, equivalently,

$$\mathbf{V}_{h, \text{orth}} = -\Pi_\tau(\mathbf{R}_h) \quad (4.43)$$

where Π_τ is the inner product defined by

$$\Pi_\tau(\mathbf{X}, \mathbf{Y}) := \sum_e (\mathbf{X}, \mathbf{Y}) \quad (4.44)$$

Substituting this expression into Eq. (4.37) we have

$$\tilde{\mathbf{U}}_{n+1} = \tau_{\Delta t, e} \left(\frac{1}{\Delta t} \mathbf{M} \tilde{\mathbf{U}}^n + \Pi_\tau^\perp(\mathbf{R}_h) \right) \quad (4.45)$$

Which defines the subscale model corresponding to the orthogonal subscales (OSS) method by Codina [78]. By instead taking $\Pi_\tau^\perp = I$; i.e. by neglecting $\mathbf{V}_{h, \text{orth}}$ one recovers the algebraic sub-grid scales (ASGS) method [77]. Both methods have similar stabilization properties, with the ASGS adding a little more diffusion.

Final stabilized formulations and simplifications

A possible simplification of the above formulation is to approximate $\Pi_\tau^\perp \approx \Pi^\perp = \mathbf{I} - \Pi$, where Π denotes the L^2 projection within the element interiors. This simplification disregards the effect of differences in the element sizes and is thus expected to work better the more homogeneous the finite element mesh is. Taking this assumption leads to the following equations for the OSS case:

1. $\Pi^\perp(\mathbf{F}) = \mathbf{0}$ (We will assume that \mathbf{F} is a finite element function)
2. $\Pi^\perp(\mathbf{M} \partial \mathbf{U}_h / \partial t) = \mathbf{0}$ (Since $\partial \mathbf{U}_h / \partial t$ is a finite element function)
3. $(\mathbf{M} \delta_t \tilde{\mathbf{U}}^n, \mathbf{V}_h) = 0$ (Since $\tilde{\mathbf{U}}^n$ is orthogonal to $\mathcal{W}_{h,0}$)

Once the model for the subscales has been established, the stabilized method is obtained directly by substituting the corresponding expression in Eq. (4.31), which, for the OSS method and taking into account the above equations, yields

$$\left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + B(\mathbf{U}_h, \mathbf{V}_h) - (\Pi^\perp \mathcal{L}(\mathbf{U}_h), \mathcal{L}^*(\mathbf{V}_h))_\tau = \mathbf{L}(\mathbf{V}_h) - \left(\frac{1}{\theta \Delta t} \mathbf{M} \tilde{\mathbf{U}}^n, \mathcal{L}^*(\mathbf{V}_h) \right)_\tau \quad (4.46)$$

The above equation can equivalently be written as the system

$$\left. \begin{aligned} & \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + B(\mathbf{U}_h, \mathbf{V}_h) - (\mathcal{L}(\mathbf{U}_h) - \boldsymbol{\pi}, \mathcal{L}^*(\mathbf{V}_h))_\tau = \\ & = \mathbf{L}(\mathbf{V}_h) - \left(\frac{1}{\theta \Delta t} \mathbf{M} \tilde{\mathbf{U}}^n, \mathcal{L}^*(\mathbf{V}_h) \right)_\tau \\ & \quad (\boldsymbol{\pi}, \mathbf{V}_h)_\tau = (\mathcal{L}(\mathbf{U}_h), \mathbf{V}_h)_\tau \end{aligned} \right\} \quad (4.47)$$

where $\boldsymbol{\pi} = [\boldsymbol{\pi}_{\text{mom}}, \pi_{\text{mass}}]^\top$ is the τ -projection onto the finite element space (which coincides with the L_2 -projection for uniform meshes). This suggests the possibility of solving the system in an iterative, staggered way, using the values for $\boldsymbol{\pi}$ from the previous nonlinear iteration (see Section 4.3.3). In order to make the formulation clearer let us introduce

$$\mathbf{d} = [\mathbf{d}_{\text{mom}}, d_{\text{mass}}]^\top := \frac{1}{\Delta t} \mathbf{M} \tilde{\mathbf{U}}^n \quad (4.48)$$

$$\begin{aligned} \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + B(\mathbf{U}_h, \mathbf{V}_h) - (\mathcal{L}(\mathbf{U}_h), \mathcal{L}^*(\mathbf{V}_h))_\tau &= \\ &= \mathbf{L}(\mathbf{V}_h) - (\mathbf{d} + \boldsymbol{\pi}, \mathcal{L}^*(\mathbf{V}_h))_\tau \end{aligned} \quad (4.49)$$

The method above is the dynamic-OSS method. Neglecting \mathbf{d} (and also neglecting the term dependent on the time step in the definition of $\boldsymbol{\tau}_{h,\Delta t}$) results in the quasi-static version of the method (Q-OSS). Both methods are viable stabilized numerical methods for the Navier-Stokes equations. However, for very small time steps ($\Delta t \not\gg \tau_1$) it becomes necessary to track the subscales [78]. In practice, it is possible to neglect \mathbf{d} and still consider the terms dependent on the time step in the definition of $\boldsymbol{\tau}_{h,\Delta t}$. However, this results in an inconsistent method, which must be taken into account.

The final expression of the variational problem, assuming a staggered resolution of Eq. (4.47) reads

$$\left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + B_{\text{OSS}}(\mathbf{U}_h, \mathbf{V}_h) = \mathbf{L}_{\text{OSS}}(\mathbf{V}_h) \quad (4.50)$$

where the stabilized bilinear form is defined by

$$B_{\text{OSS}}(\mathbf{U}_h, \mathbf{V}_h) := B(\mathbf{U}_h, \mathbf{V}_h) - (\mathcal{L}(\mathbf{U}_h), \mathcal{L}^*(\mathbf{V}_h))_\tau \quad (4.51)$$

and where

$$\mathbf{L}_{\text{OSS}}(\mathbf{V}_h) := \mathbf{L}(\mathbf{V}_h) - (\mathbf{d} + \boldsymbol{\pi}, \mathcal{L}^*(\mathbf{V}_h))_\tau \quad (4.52)$$

The ASGS method is obtained by taking $\Pi^\perp = I$, which results in the following problem

$$\left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h, \mathbf{V}_h \right) + B_{\text{ASGS}}(\mathbf{U}_h, \mathbf{V}_h) = \mathbf{L}_{\text{ASGS}}(\mathbf{V}_h) \quad (4.53)$$

where the stabilized bilinear form is defined by

$$B_{\text{ASGS}}(\mathbf{U}_h, \mathbf{V}_h) := B(\mathbf{U}_h, \mathbf{V}_h) + \left(\mathbf{M} \frac{\partial}{\partial t} \tilde{\mathbf{U}}, \mathbf{V}_h \right) - \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U}_h + \mathcal{L}(\mathbf{U}_h), \mathcal{L}^*(\mathbf{V}_h) \right)_\tau \quad (4.54)$$

and where

$$\mathbf{L}_{\text{ASGS}}(\mathbf{V}_h) := \mathbf{L}(\mathbf{V}_h) - (\mathbf{F} + \mathbf{d}, \mathcal{L}^*(\mathbf{V}_h))_\tau \quad (4.55)$$

4.3.3 The overall algorithm

The spatial discretization of Eq. (4.18) is obtained through the use of linear tetrahedral elements, stabilized with the variational multiscale stabilization technique [174, 77], which leads to an optimal, second order mesh-size convergence. The particular variant of the method that we use here corresponds to the stationary subgrid scales (Q-ASGS) formulation, which is described in full detail in [82]. After assembling all the elements contributions and imposing the boundary conditions, it leads to a system of equations of the form

$$\mathbf{M} \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix} + \mathbf{C}(\mathbf{U}, \mathbf{P}) \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} = \mathbf{F} \quad (4.56)$$

where \mathbf{U} and \mathbf{P} stand for the nodal unknowns for the velocity ($3n_{\text{nodes}}$ unknowns) and for the pressure (n_{nodes}) for the pressure respectively. For the time discretization we use a second-order, Bossak time integration [370], defined as

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix}^{n+1} = \begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix}^n + \Delta t \left((1 - \gamma_N) \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix}^n + \gamma_N \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix}^{n+1} \right) \quad (4.57)$$

$$(1 - \alpha_B) \mathbf{M} \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix}^{n+1} + \alpha_B \mathbf{M} \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix}^n + \mathbf{C}(\mathbf{U}^{n+1}, \mathbf{P}^{n+1}) \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}^{n+1} = \mathbf{F}^{n+1} \quad (4.58)$$

where n is the time-step index. Combining both equations in Applying this scheme to Eqs.(4.57) and (4.58) one obtains

$$\begin{aligned} \mathbf{R}(\mathbf{U}^{n+1}, \mathbf{P}^{n+1}) = & \mathbf{F} - \frac{1 - \alpha_B}{\gamma_N \Delta t} \mathbf{M} \begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix}^n + \left((1 - \alpha_B) \left(\frac{1}{\gamma_N - 1} \right) + \alpha_B \right) \mathbf{M} \begin{bmatrix} \frac{\partial \mathbf{U}}{\partial t} \\ \mathbf{0} \end{bmatrix}^n + \\ & - \left(\frac{1 - \alpha_B}{\gamma_N \Delta t} \mathbf{M} + \mathbf{C} \right) \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}^{n+1} \end{aligned} \quad (4.59)$$

where we choose $\alpha = -0.3$ and $\gamma_N = 1/2 - \alpha_B$ ³ and where we have defined the residual $\mathbf{R}(\mathbf{U}^{n+1}, \mathbf{P}^{n+1})$. The nonlinearities present in Eq. (4.14) are linearised using a first-order

³This combination of parameters provides good damping properties of the highest frequencies and robust behaviour overall, see [370].

Taylor expansion. That is, at each nonlinear iteration one solves for the :

$$\mathbf{R}_{i+1}^{n+1} \approx \mathbf{R}_i^{n+1} + \left(\frac{\partial \mathbf{R}}{\partial \mathbf{U} \partial \mathbf{P}} \right)_{i+1}^{n+1} \begin{bmatrix} \delta \mathbf{U} \\ \delta \mathbf{P} \end{bmatrix}^{n+1} = 0 \quad (4.60)$$

Then the solution and residual are iteratively updated with Picard's method as

$$\begin{aligned} \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}_{i+1}^{n+1} &= \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}_i^{n+1} + \begin{bmatrix} \delta \mathbf{U} \\ \delta \mathbf{P} \end{bmatrix}^{n+1} \\ \mathbf{R}_0^{n+1} &= \mathbf{R}(\mathbf{U}^n, \mathbf{P}^n) \\ \mathbf{R}_{i+1}^{n+1} &= \mathbf{R}(\mathbf{U}_i^{n+1}, \mathbf{P}_i^{n+1}) \end{aligned} \quad (4.61)$$

where the index i represents the nonlinear iteration count and where, in evaluating the derivative of the residual, we use the following approximation

$$\frac{\partial \mathbf{R}}{\partial \mathbf{U} \partial \mathbf{P}} \approx \frac{1 - \alpha_B}{\gamma_N \Delta t} \mathbf{M} + \mathbf{C}_i^{n+1} \quad (4.62)$$

where the indices are applied only to matrix \mathbf{C} , as \mathbf{M} does not depend on the solution. Note that this approximation assumes that the variation of \mathbf{C} is moderate compared to that of the solution vector itself, otherwise convergence problems can appear. Consequently, the final system to be solved is

$$- \left(\frac{1 - \alpha_B}{\gamma_N \Delta t} \mathbf{M} + \mathbf{C}_i^{n+1} \right) \begin{bmatrix} \delta \mathbf{U} \\ \delta \mathbf{P} \end{bmatrix}_i^{n+1} = \mathbf{R}_i^{n+1} \quad (4.63)$$

4.4 Derivative recovery

The procedure by which one obtains approximations to the derivatives of a field, given an approximation of the field itself is called *derivative reconstruction* or, more frequently in the FEM community, *derivative recovery*. The literature on the topic is quite extensive, due to its interest in post-processing [252, 195, 81, 17], error estimation [162, 374], but also as part of the solution process in iterative schemes that use the gradient of the functions that are being solved for, such as in many finite volume codes.

In spite of all this work, there have been relatively few works to study the effect of the recovery method on the quality of the resulting particle-laden flow simulations. This can be explained in part by the fact that often the focus has been placed on problems where only the steady terms of the MRE were relevant, such as those dealing with very small, heavy

particles settling in a gas. There are exceptions though, such as [140, 351]; although these works are concerned with the highly accurate pseudo-spectral methods, used in fundamental research with very simple domains. Their conclusions are of little direct relevance to the low-order finite element applications of our interest. We will thus attempt to partially fill this gap, especially for FEM-based approaches, which are practically absent from the literature.

In this section we explore different possibilities to determine these values from a FEM-based solution of the fluid, focussing on linear tetrahedra finite elements. Notwithstanding this, such methods are not strongly dependent on the underlying CFD methodology. For that reason we will first overview other existing methodologies that have been applied to finite volumes and pseudo-spectral methods, the two preferred methodologies for particle-laden flow simulations.

Let us start with a list of the required derivatives; see Table 4.1⁴. Not all the derivatives collected in Table 4.1 are always needed. Depending on the system of interest there are several possibilities. We have given in Chapter 2 the conditions under which the Faxén can be neglected. Note that under these conditions it is only necessary to recover the first order time derivative and the gradient of the velocity. Extending the MRE with lift or torque terms does not in principle imply any further recoveries, since the vorticity can be obtained algebraically from the gradient. However, depending on the mode of recovery, the whole gradient might not be available and thus a specific approach might be more appropriate in this case.

Note that when the Faxén terms are included, there is a considerable growth in the recovery costs. Not only is the Laplacian of the velocity needed, but also its material derivative, which involves third-order spatial derivatives. We will see later that in such a case it is likely that linear finite elements become too inaccurate.

In other situations, a scaling analysis may justify neglecting specific terms in the equation. The simplest case of all corresponds to neglecting all but the drag force (without Faxén corrections). In this case there is of course no need to recover any derivative at all. Nonetheless this is in fact a very important simplification with multitude of applications in both fundamental [241, 28, 68] and applied [213, 302, 98] research.

Another common scenario allows one to include only the steady drag force and the unperturbed fluid force only. Such case is most commonly encountered in internal flows of liquids, where every time the fluid accelerates due to a spatial change of the container, the particles are affected by a force that cannot be neglected, unless the particles are extremely heavy and the accelerations are mild. In this case the material derivative of the fluid must

⁴There appear to be no Faxén corrections associated with the extended terms. To our knowledge such terms have not been derived yet. In other words, the problem of deriving the lift coefficient in a non-uniform, low Re_p flow is still an open one .

be recovered. This will be the case in two of our application examples (Section 4.6 and Section 4.7).

The motivation to study different techniques of recovery comes from the poor accuracy of the available solution in practice, especially at small scales. Indeed, three-dimensional, transient flow problems are well known to be computationally expensive in general. For particle-laden this issue is typically accentuated, because very often the small scales of the flow crucially affect the movement of the submerged particles, which means that the flow accuracy must be sufficient at these scales too. However, the number of elements available to resolve the small scales is rarely large, due to the need to save computational resources. Naturally, estimating the derivative of these poorly resolved scales can result in very inaccurate estimations of the forces that depend on them, destroying the general accuracy of the problem.

But how small are these small scales? In Chapter 2 the small scales were characterized as the Kolmogorov microscales, since the theory required that the perturbative flow around the particle could be described by the unsteady Stokes equations. In this chapter we have relaxed this requirement and this characterization has become too restrictive. An alternative characterization can be achieved through the *scale-dependent* Stokes number [29], defined as

$$St_L := \frac{\tau_p}{\tau_L} \quad (4.64)$$

where τ_p is the relaxation time of the particle as in Chapter 2 (although to estimate one can use the extended MRE expression for \mathbf{F}_D) and τ_L is the characteristic time of the eddies of size L . One can then define the small scales as those for which $St_L \sim 1$. This will typically be the smallest scales resolved in the fluid simulation; otherwise, for coarser simulations, the interaction model must be enriched with a stochastic turbulent dispersion model (e.g. [296]) to make up for the relevant, unresolved scales.

It is thus important to come up with methodologies to extract as accurate estimates for the local derivatives as possible, preferably with the same asymptotic behaviour of the error as the solution itself. For linear finite elements, this optimal accuracy corresponds to an error that is asymptotically of $O(h^2)$.

4.4.1 Overview of existing approaches

The need to compute derivatives of a pre-computed numerical solution has been part of the field from its very beginnings, and an exhaustive summary would surely be too ambitious. Indeed, The problem of derivative recovery arises in many different fields or computational science, such as image processing [274], error estimation [374] or numerical optimization [83].

Table 4.1 *Derivatives required to calculate the different terms of the equation of motion of the particles. The time derivatives (following the particle) are not considered here, since it is assumed that they will be resolved with finite differences. The derivatives corresponding to the extended version of the MRE are consistent with the forms discussed in Section 4.2*

<i>Force</i>	<i>Non-Faxén terms</i>		<i>Additional Faxén terms</i>	
	<i>Compact</i>	<i>Expanded</i>	<i>Compact</i>	<i>Expanded</i>
\mathbf{F}_U	$\frac{D\mathbf{u}}{Dt}$	$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j}$	–	–
\mathbf{F}_A	$\frac{D\mathbf{u}}{Dt}$	$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j}$	$\frac{D}{Dt} \Delta \mathbf{u}$	$\frac{\partial}{\partial t} \frac{\partial^2 u_i}{\partial x_j \partial x_j} + u_j \frac{\partial^3 u_i}{\partial x_j \partial x_k \partial x_k}$
\mathbf{F}_D	–	–	$\Delta \mathbf{u}$	$\frac{\partial^2 u_i}{\partial x_j \partial x_j}$
\mathbf{F}_H	–	–	$\Delta \mathbf{u}$	$\frac{\partial^2 u_i}{\partial x_j \partial x_j}$
<i>Extended MRE terms</i>				
$\mathbf{F}_{L,\text{Saff}}$	$\nabla \times \mathbf{u}$	$\varepsilon_{ijk} \frac{\partial u_j}{\partial x_k}$	–	–
$\mathbf{F}_{L,\text{Mag}}$	–	–	–	–
\mathbf{T}_{hidr}	$\nabla \times \mathbf{u}$	$\varepsilon_{ijk} \frac{\partial u_j}{\partial x_k}$	–	–

We focus here on the approaches that have been used in particle-laden flow simulations, narrowing the scope of the discussion drastically. The fact that most particle-laden flow simulations have neglected the terms in the MRE that contain derivatives (steady-state approximation, neglecting the Faxén corrections) still furthers this narrowing. In spite of this, there is still a wealth of examples where at least some of the remaining terms have been taken into account.

Pseudo-spectral methods (PSMs), being the standard method for the direct numerical simulation of turbulence [251], have been extensively employed for the study of the interactions between turbulence and small suspended particles. The high accuracy per degree of freedom (DOF) that can be achieved with this class of methods motivates the use of very accurate schemes for the recovery of derivatives, as we discussed in Section 4.5.2.

The great majority of simulations reported in the literature are concerned with the motion of tiny particles of a density much higher than the suspending fluid. In these cases one tends to have $\rho \ll 1$ and $St_L \ll 1$, which justifies (see Eq. (2.91)) the common practice of considering only the drag force \mathbf{F}_D . This means that the interpolation of the fluid velocity is sufficient and there is no need to calculate its derivatives. Moreover, even with the inclusion of \mathbf{F}_U , usually the first other force to become relevant, it is often still possible to avoid this calculation by replacing the material derivative (in a frame following the fluid) by the derivative in a frame following the particle, as both tend to coincide in the limit of very small

Re_p [238, 241, 29]. But the time derivatives in the framework of the particle do not involve a gradient; the time derivative is already material in its frame of reference.

In other occasions, such as in the study of nearly neutrally buoyant particles, the terms with derivatives cannot be neglected. For instance, all (non-Faxén) terms in the MRE were included in [266] and in [90]. The former used quadratic interpolation, while the latter used tricubic interpolation. A sixth-order polynomial interpolation was used in [233]. Only a handful of researchers have included the Faxén corrections in their equations, see [54, 12]. These authors also used polynomial interpolation in a regular mesh.

In pseudo-spectral methods the approximate solution is an infinitely differentiable function. The derivatives can therefore be obtained directly, by taking the derivative of the approximate solution and then proceed as with the velocity, by interpolation. Popular techniques involving this methods are Lagrangian [41] or Hermite polynomial interpolation [117].

However the direct calculation of the derivatives requires additional fast Fourier Transforms (FFT), making the operation much more expensive than mere interpolation [354]. A (faster) alternative is to take the derivatives of the interpoland instead, avoiding all the extra FFT transformations. In [354, 351] several interpolation (and recovery) techniques are compared in detail. A conclusion of this work is that in general higher-order interpolations are preferable in the context of PSM, among which the B-spline method is highlighted, as it achieves comparable accuracy to the Hermite polynomials at a much lower cost .

PSMs are only practical for relatively simple geometries and boundary conditions, making them the preferred approach in areas of fundamental research, especially those involving DNS. On the other hand, the finite volume method (FVM) is the most popular approach in more applied or industrial applications. Here the accuracy per DOF is not as high and, in fact, comparable to finite elements with similar number of unknowns (even if, overall, it is still more efficient, according to recent research [184]) . The complexity of the boundary conditions or the error introduced in the physical models (like turbulence modelling in LES, see [251]) often make higher-order approximations futile in this context. The consensus seems to be that second-order methods are competitive for CFD simulations, but not first order [168].

Accordingly, the order of approximation of the recovered derivatives should not be worse than $O(h^2)$, if possible, especially if its value has an important weight in the calculation of important terms in the MRE, as is often the case with \mathbf{F}_U .

There are two generally cited theoretical approaches that are implemented in most general-purpose commercial and open source codes for CFD with the FVM. The first one is the Green–Gauss theorem (G–G) approach, based on a discrete form of this theorem. This method takes advantage of the face-oriented formulation of finite-volume methods.

Specifically, the value of the gradient at a volume centroid C can be approximated by

$$\nabla f(\mathbf{x}_C) \approx \frac{1}{|\Omega_C|} \int_{\Omega_C} \nabla f(\mathbf{x}) \, dV = \frac{1}{|\Omega_C|} \int_{\partial\Omega_C} f \mathbf{n} \, dS \approx \frac{1}{|\Omega_C|} \sum_{\Phi_C} f_{\Phi} \mathbf{n}_{\Phi} \quad (4.65)$$

where $|\Omega_C|$ is the measure of the finite volume Ω_C , $\partial\Omega_C$ is the boundary of the volume, \mathbf{n} denotes the outward normal and the Φ_C indices run over the set of faces of the boundary, on which the corresponding f_{Φ} and \mathbf{n}_{Φ} are assumed to be constant. The values of the gradient on each face can then be obtained by linear interpolation of the centroid value of the two volumes that share the face.

The particular version of this method described by Eq. (4.65) is the default option in ANSYS–Fluent (Release 17.0) [9] code and one of the two available options in the popular open-source code OpenFOAM [151]

The second type of methods are the least squares-type (LS) algorithms. They are based on the idea of optimizing the first-order Taylor expansion approximation of the gradient based on the local variation of the quantity of interest. Specifically, let us consider again the centroid C corresponding to the finite volume Ω_C surrounded by a collection of adjacent finite volumes with centroids C_k . The gradient approximation is obtained by requiring that, to first order, the Taylor expansion of the quantity of interest based on the candidate gradient minimizes the mean squared error of the quantities of interest at neighbouring locations when comparing its prediction to the value actually stored there. The number of neighbours to consider is not fixed but will always be more than the number of unknowns (three for a scalar quantity) to determine the gradient. In mathematical terms, the function to be minimized is:

$$E_C^2(\mathbf{g}_C) = \sum_{C_k} w_k (f_{C_k} - (f_C + \mathbf{g}_C \cdot \Delta \mathbf{x}_k))^2 \quad (4.66)$$

where E_C^2 is the error squared, \mathbf{g}_C is the candidate gradient, $\Delta \mathbf{x}_k := \mathbf{x}_{C_k} - \mathbf{x}_C$ and the w_k are positive weights that can be tuned to control the locality of the approximation (by giving more or less importance to closer neighbours, for example). Taking the derivatives of function in Eq. (4.66) and imposing that they must vanish at C leads to the following system:

$$\frac{\partial E(\mathbf{g}_C)}{\partial g_{C,i}} = 0 \Leftrightarrow \sum_{i=1}^3 \left(\sum_{C_k} w_k \Delta \mathbf{x}_{k,i} \Delta \mathbf{x}_{k,j} \right) g_{C,j} = \sum_{i=1}^3 w_k \Delta \mathbf{x}_{k,i} \Delta f_k \quad (4.67)$$

where $\Delta f_k := f(\mathbf{x}_{C,k}) - f(\mathbf{x}_C)$. Solving the system above gives the approximation of the gradient.

While this sort of method is considered more costly than the GG-type method, it does allow for a higher flexibility in the accuracy order [255] and is also available in both ANSYS–Fluent and OpenFOAM. Being more accurate than the G–G method in most cases, this method has however shown to be less robust with regular, highly stretched meshes, such as the ones commonly used along boundary layers in CFD.

There have recently appeared a few works analysing the respective accuracies and comparing both methods in their standard form. However, most of these works did not deal with particle-laden flows specifically. In [321] it was shown that the standard G–G method and other variants, very much used in practice, were actually inconsistent (with asymptotic errors of order zero) although this was only revealed in deep enough convergence studies, which explains why this had (apparently) not been detected before. Nonetheless, their remarkable accuracy compared to other alternatives in stretched meshes of the type for the boundary layer region was Highlighted.

In [336] it is concluded that the ‘widely believed’ notion of both methods being second-order accurate only held true with regular meshes. The comparative analysis between the G–G and LS methods presented in this work shows that the orders of the commonly applied versions of the two methods depend on the type of mesh, being at most two for structured meshes; but that it drops to zero and first orders respectively with unstructured meshes of the type produced by common meshing algorithms. Furthermore, the G–G method is also shown to outperform in practice the LS method in the kind of extremely elongated meshes used for boundary layers was also stressed, as observed in [321] (although this difference can be mitigated with correct cell alignment and adequate choice of the weights w_k) [237] .

Notwithstanding this, there exist solutions to the low accuracy of the G–G method. In fact, some variations on the two methods were proposed in [237] and in [336] to avoid an inconsistency in the G–G method when applied to irregular meshes, achieving first-order accuracy consistently on them . Moreover, it is possible to construct a G–G-based method with second-order accuracy in the gradient (and first-order in the Hessian) by making the method *global* [38]; that is, requiring the resolution of a system of equations of the same size as the one to be solved to obtain the solution. Of course, this implies an increase in computational cost.

In summary, it seems that none of the widely used methods is robustly of second-order accuracy for unstructured meshes. However, this is desirable, since this is frequently the accuracy of the underlying numerical method and the existence of a single term with lower accuracy in the MRE might potentially compromise the overall accuracy of the method. In any case, the matter is clearly not completely settled, especially with regards to its implications for

the computation of particle-laden simulations in which field a systematic study for common methods in unstructured meshes is still lacking.

Projection techniques

In the context of the FEM, the most natural approximation of the gradient of the solution is to take the gradient of the finite element approximation itself. The problem is that, inevitably, the resulting function has its smoothness order reduced by one after differentiation. For the case of linear elements, this means that the gradient approximation obtained in this way will in general be discontinuous at the element boundaries. Such jagged function is not an adequate representation of the flow that will surely lead to an important loss of accuracy if it is used in the numeric integration of the MRE without post-processing. Since continuity is important, an elegant option is to project such discontinuous function onto the piecewise linear finite element space, formed by the same shape functions that model the solution itself. This is what we mean by 'projection techniques'. That is, consider

$$\nabla \mathbf{u} \approx \Pi(\nabla \mathbf{u}_h) \quad (4.68)$$

where Π is the L^2 projection operator onto the finite element space $\mathcal{V}_h \times \mathcal{V}_h \times \mathcal{V}_h$ with \mathcal{V}_h formed by the same d -dimensional linear functions from Section 4.3.2, but with no prescribed boundary conditions. The linearity of the projection allows to treat it componentwise, so that each component of the projection of the gradient is equal to the projection of the same component of the gradient. Thus, for simplicity we next focus in the gradient recovery of a scalar field, bearing in mind that such scalar could represent a single component of the velocity, and that the repetition of the process for each component would yield the whole gradient. Let us thus consider the scalar field $u(\mathbf{x})$, of which its finite element approximation is $u_h(\mathbf{x})$. The definition of the projector operator of a square-integrable function (such as ∇u_h) is given by the following finite element problem: find $\mathbf{g}_h \in \mathcal{V}_h$, such that

$$(\mathbf{g}_h, \mathbf{v}_h) = (\nabla u_h, \mathbf{v}_h) \quad \text{for all } \mathbf{v}_h \in \mathcal{V}_h \quad (4.69)$$

where \mathbf{g}_h represents $\Pi(\nabla u_h)$, that is, the recovered gradient of u_h and the set of test functions is given by $\mathcal{V}_h = \{\mathbf{v}_h^{aj} \mid [\mathbf{v}_h^{aj}]_i = N^a \delta_{ij}\}$. Using the finite element expansion from Eq. (4.20)

in Eq. (4.69), one obtains

$$\int_{\Omega} (N^b G_j^b) (N^a \delta_{ij}) = \int_{\Omega} \left(\frac{\partial}{\partial x_j} N^b U^b \right) (N^a \delta_{ij});$$

$$\underbrace{\int_{\Omega} (N^a N^b \delta_{ij})}_{\mathbf{M}_{(ai),(bj)}} \underbrace{G_j^b}_{\mathbf{G}_{(bj)}} = \int_{\Omega} \underbrace{\left(N^a \frac{\partial N^b}{\partial x_i} \right)}_{\mathbf{g}_{(ai)}} U^b \quad (4.70)$$

where the i, j indices run through all the space dimensions and a, b through all the nodes; and where, as suggested by grouping of indices above, the LHS can be arranged into a matrix M called the *consistent mass matrix*, which is positive-definite [263]. And since such problem has a unique solution, the projection operator given by $\Pi(\nabla u_h) = \mathbf{g}_h$ is well defined.

Unfortunately, the accuracy of the gradient recovered in this way is in general of a lower order than that of the solution itself, u_h [394]. In fact, one would expect that one full order is dropped for the derivative in general, unless one takes advantage of certain superconvergence property of particular points, see [46]. This is because even the optimal interpolant of the solution, when differentiated, formally loses an order of approximation. This can be understood if one expands the solution and the approximant in Taylor series. Assuming enough smoothness, the difference in the derivative of both series would be the Taylor expansion of the residual of the derivative, which would have a lower order due to differentiation. This means that the terms containing first derivatives of the fluid velocity in the MRE will have a much lower accuracy than, for example, the drag force, that depends on the velocity only.

Standard approach

The simplest way to compute the nodal gradient of the velocity when using finite elements is to directly differentiate the shape functions inside the element and then to average the elemental contributions onto the node. Let us assume the derivative of the shape functions is not continuous inter-elementary, but rather suppose jump continuities exist at the interfaces. The derivative of the discretized solution is not defined, but one possibility is to take an average of the different derivatives obtained from the interior of the different elements as we approach the node in question. The average weights can be taken as the measure of the corresponding element. That is

$$\nabla u \approx G_j^b N_j^b \quad (4.71)$$

where, like we did before, assume that summation is implied for repeated indices, which are to be assumed to run over all the nodes in the system unless otherwise stated.

$$G_j^b = \frac{1}{|\Omega_b|} \sum_{\Omega_e|b \in \Omega_e} |\Omega_e| (\nabla u_h)_j|_{\Omega_e} = \frac{U^a}{|\Omega_b|} \sum_{\Omega_e|b \in \Omega_e} |\Omega_e| \frac{\partial N^a}{\partial x_i} \Big|_{\Omega_e} \quad (4.72)$$

and where

$$|\Omega_b| = \sum_{\Omega_e|b \in \Omega_e} |\Omega_e| \quad (4.73)$$

can be termed the *nodal volume* of node b . Note that the RHS in Eq. (4.70) is

$$\int_{\Omega} \left(N^a \frac{\partial N^b}{\partial x_i} \right) U^b = \sum_{\Omega_e|a \in \Omega_e} \int_{\Omega_e} \left(N^a \frac{\partial N^b}{\partial x_i} \right) U^b = U^b \sum_{\Omega_e|a \in \Omega_e} \frac{|\Omega_e|}{4} \frac{\partial N^b}{\partial x_i} \Big|_{\Omega_e} \quad (4.74)$$

Now, by replacing the (consistent) mass matrix in Eq. (4.70) by its lumped counterpart, i.e. by the matrix obtained by adding all the column (or row) contributions, making the global matrix diagonal; one obtains the following LHS:

$$\begin{aligned} \mathbf{M}_{(ai),(bj)}^L G_j^b &= \delta_{ab} \sum_b \mathbf{M}_{(ai),(bj)} G_j^b = \delta_{ab} \int_{\Omega} \sum_b (N^a N^b \delta_{ij}) G_j^b = \\ &= \delta_{ab} \delta_{ij} \int_{\Omega} N^a G_j^b = \frac{|\Omega_a|}{4} \delta_{ab} \delta_{ij} G_j^b = \frac{|\Omega_a|}{4} G_i^a \end{aligned} \quad (4.75)$$

where we have made the summation signs explicit selectively for the sake of clarity. Note that by equating the RHSs of Eq. (4.74) and Eq. (4.75) and renaming indices ($a \rightarrow b$), we recover Eq. (4.72). The standard approach can thus be seen as a simplified version of the L_2 -projection technique.

Similarly, the material derivative is defined as

$$\left(\frac{D\mathbf{u}}{Dt} \right)_i = \left(\frac{\partial \mathbf{u}}{\partial t} \right)_i + (\mathbf{u} \cdot \nabla \mathbf{u})_i = \frac{\partial u_i}{\partial t} + u_j (\nabla \mathbf{u})_{ij} \quad (4.76)$$

The first term on the RHS can be computed, for example, as

$$\frac{\partial \mathbf{u}}{\partial t} \approx \frac{1}{\Delta t} (\mathbf{u}_h^{n+1} - \mathbf{u}_h^n) \quad (4.77)$$

where we have made the temporal indices explicit. Meanwhile, the convective term can be obtained as

$$u_j (\nabla \mathbf{u})_{ij} \approx \frac{U_i^b U_j^b}{|\Omega_b|} \sum_{\Omega_e|b \in \Omega_e} |\Omega_e| \frac{\partial}{\partial x_j} N^b \Big|_{\Omega_e} (\mathbf{x}_b) \quad (4.78)$$

Since we are mainly interested in linear elements, the Laplacian cannot be obtained by analogy, taking second derivatives, since they vanish in the elements interiors. The simplest procedure is to perform a sequence of two steps, first obtaining the nodal gradient, and differentiating the interpolated gradient. That is, $\Delta \mathbf{u} \approx L_i^b N^b$, with

$$L_i^b = \frac{G_{i,j}^b}{|\Omega_b|} \sum_{\Omega_e | b \in \Omega_e} |\Omega_e| \frac{\partial}{\partial x_j} N^b |_{\Omega_e}(\mathbf{x}_b) \quad (4.79)$$

where $\frac{G_{i,j}^b}{|\Omega_b|}$ must have been previously calculated, e.g. according to Eq. (4.72). Note that this procedure gathers information from the area corresponding to all the nodes contained in the elements containing the node of interest. This effectively means that the Laplacian is determined by a *coarsened* discretization. In general, it is to be assumed that this is the way in which we will obtain derivatives of order higher than one, unless otherwise specified.

Patch techniques and the method of Zhang and Naga (PPR)

By the term ‘patch techniques’ we loosely refer to recovery methods for which the determination of the derivatives at a point in the domain is based, solely, on information gathered from the points found in a patch around the target point. The term can be used in opposition to ‘global techniques’, where the information of any point depends in general on information from all the other points. By this definition, therefore, the standard approach is also patch technique, whereas the L_2 -projection method is a global techniques.

Here we review the method proposed by Zhang and Naga [391], closely related to the well known method by Zienkiewicz and Zhu [394] but preferable to our interests for two reasons: its nodal-based approach, as opposed to the Gauss point approach of the latter; and its greater accuracy [391, 282], despite its marginally higher cost, which is not important compared to the overall cost of the simulations.

In essence, it is also very similar to the LS-type methods that we have briefly overviewed in Section 4.4.1. It consists in constructing, for each node, a least-squares best fit of a $k + 1$ -order polynomial to the nodal values in a neighbourhood of the node. The gradient is then obtained by taking the derivative of the polynomial, evaluating it at the position of the node. The full algorithm can be broken down into the following steps. Here we also point out any differences we have introduced with respect to the original method:

1. For each node, assign a set of nodes surrounding it. In Zhang and Naga [391] the iterative method in Algorithm 3 is proposed. We have instead used the method described by Algorithm 4. In it, nodes are added, one by one, from the nodes contained in the adjacent elements. For each new node added, the resulting system is solved for a known

test result and the error compared to a pre-established tolerance. If the test is passed, the list of nodes is accepted; otherwise a new node is added to the cloud and the process is repeated until convergence.

The reason for choosing this approach is our observation that the theoretical minimum number of neighbours⁵, given by $m_{PPR} = 10$, produces very ill-conditioned systems in many cases, leading to inaccurate results. Furthermore, we understood there is no need to increase the number of neighbours indirectly (by increasing the radius of the ball that contains it) and so we proceed node by node instead, using the nodes from successive layers of adjacent elements. This simple approach has the advantage of automatically respecting the topology of the domain, avoiding situations like the one depicted in Fig. 4.3.

Let the Voronoi graph of the finite element mesh be $\text{Vor} = (V, E)$, with V the set of vertices and E the set of edges. We use the following recursive definition of the neighbourhood function, that acts on sets of nodes of the tetrahedra mesh and returns the set formed by the union of the original vertices with the set of adjacent neighbours:

$$\begin{aligned} N_{\text{Vor}}: \mathcal{P}(V) &\rightarrow \mathcal{P}(V) \\ \{w\} &\mapsto \{v \in V \mid \{v, w\} \in E\} \\ S &\mapsto \{v \in V \mid \exists w \in S \mid v \in N_{\text{Vor}}(\{w\})\} \end{aligned} \quad (4.80)$$

where $\mathcal{P}(V)$ is the power set of V .

- For each node a and spatial component i , consider the polynomial $p_a(\mathbf{x}) = a_1 + a_2x_1 + a_3x_2 + a_4x_3 + a_5x_1^2 + a_6x_2^2 + a_7x_3^2 + a_8x_1x_2 + a_9x_1x_3 + a_{10}x_2x_3$. The coefficients a_i are obtained from the system

$$\mathbf{A}^\top \mathbf{A} \mathbf{a} = \mathbf{A}^\top \mathbf{b} \quad (4.81)$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,2}x_{1,3} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,2}x_{2,3} \\ 1 & x_{3,1} & x_{3,2} & \cdots & x_{3,2}x_{3,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n_a,1} & x_{n_a,2} & \cdots & x_{n_a,2}x_{n_a,3} \end{pmatrix}; \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{10} \end{pmatrix}; \quad \mathbf{b} = \begin{pmatrix} u_h(\mathbf{x}_1) \\ u_h(\mathbf{x}_2) \\ u_h(\mathbf{x}_3) \\ \vdots \\ u_h(\mathbf{x}_{n_a}) \end{pmatrix} \quad (4.82)$$

⁵It is equal to the number of monomials of a polynomial of order $k + 1$ in dimension dim . By a simple combinatorial argument one derives the formula $N_{\min}(\text{dim}, k) = \binom{\text{dim}+k}{k}$ which in our case yields $N_{\min}(3, 2) = 10$.

where $x_{i,j}$ is the j -component of the coordinates of node i and n_a is the number of nodes in the neighbourhood of a (including a). Moreover, it is convenient to work with nondimensionalized variables to improve the condition number of the resulting matrices [391]. A simple way to do so is to take the \mathbf{x} coordinates to be normalized by the radius of the cloud of points, i.e.

$$h_a = \frac{1}{2} \max_{i,k=1,\dots,n_a} \{\|\mathbf{x}_i - \mathbf{x}_k\|\} \quad (4.83)$$

where the subindices on \mathbf{x} indicate the node to which they belong. In that case the units of p_a must be restored by multiplying the coefficients a_i by h_a the appropriate number of times after solving the system. That is, $a_1 \rightarrow a_1$, $a_2 \rightarrow a_2 h_a$, $a_4 \rightarrow a_4 h_a^2$, and so on.

3. The nodal gradient of the i -velocity component at node a is then approximated by

$$g_{h,i} = \left. \frac{\partial p_a}{\partial x_i}(\mathbf{x}) \right|_{\mathbf{x}=\mathbf{0}} \quad (4.84)$$

Algorithm 3 Creation of clouds (lists) of neighbours (original algorithm)

```

1: for node in nodes do
2:   cloud  $\leftarrow$  [node] ▷ Init. cloud of nodes
3:   contiguous  $\leftarrow$  GETLISTOFCONTIGUOUSNODES(node)
4:    $h \leftarrow$  GetMaxLongEdges(node) ▷ Longest edge containing the node
5:   while Size(cloud) <  $m_{PPR}$  do
6:     cloud  $\leftarrow$  GETLISTOFNODESINBALL(node,  $h$ )
7:      $h \leftarrow 2h$ 

```

The resulting nodal gradient \mathbf{g}_h may then be evaluated at the position of each particle, so that it can be used to calculate the necessary terms in the MRE. In the case of the velocity, it is possible to calculate the full gradient, and then evaluate at the particle position, or else to do so after each component of the gradient is obtained. The second option might be interesting in order to limit the amount of allocated memory by reusing the same container for each component of the gradient.

On the other hand, note that it is possible to avoid repeating the algorithm above every time one does forward coupling by storing the relevant nodal information. Specifically, note that only coefficients a_2 , a_3 and a_4 are needed to calculate the gradient. Therefore only the second, third and fourth rows of the matrix $(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$ need to be stored for each node. To avoid the search operation at each forward coupling step it would also be necessary to keep a list of the neighbouring nodes, for each node. As we mentioned, the minimum number

Algorithm 4 Creation of clouds (lists) of neighbours (modified algorithm)

```

1: for node in nodes do
2:   cloud  $\leftarrow$  [node] ▷ Init. list of neighbour nodes
3:   elements  $\leftarrow$  GETELEMENTSCONTAININGNODE(node) ▷ Init. set of elements
4:   candidates  $\leftarrow$  GETNODES(elements) ▷ Init. list of candidate nodes
5:   success  $\leftarrow$  False
6:   while  $\neg$ success do
7:     if SIZE(cloud)  $\geq m_{PPR}$  then
8:       success  $\leftarrow$  RUNTEST(node)
9:     if  $\neg$ success then
10:      if SIZE(cloud) + 1 > SIZE(candidates) then ▷ Not enough candidates avail.
11:        elements  $\leftarrow$  GETNEIGHBOURELEMENTS(elements) ▷ Eq. (4.80)
12:        candidates  $\leftarrow$  GETNODES(elements)
13:        ORDERBYDISTANCE(cand)
14:      next  $\leftarrow$  SIZE(cloud)
15:      APPEND(cloud, candidates[next]) ▷ Append closest new node to cloud

```

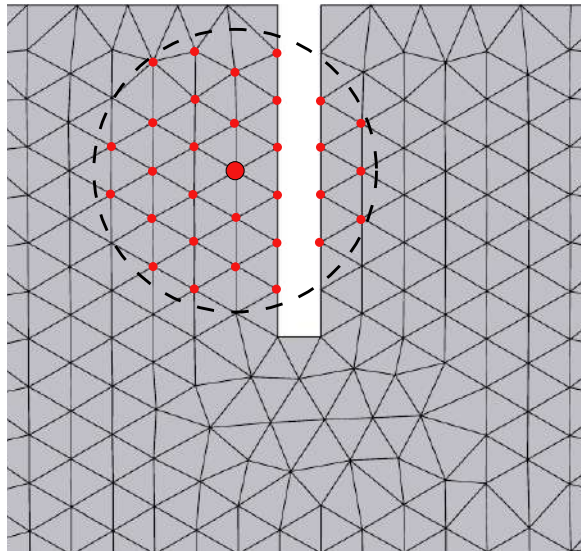


Fig. 4.3 Illustration of the kind of erroneous neighbours cloud that is avoided by taking only nodes from adjacent layers of elements.

of neighbours is ten, although it is often necessary to use a larger number of neighbours to avoid ill-conditioned matrices. We have observed the average number of neighbours to be closer to 25 in practice. Similar numbers have been observed by Ortega [269] in applications of the finite-point method, that employs a similar approach to compute the spatial derivatives. Assuming this list is kept as a list of pointers, one may estimate the total memory requirements to be about 440 B per node. By performing the search at each step this number would fall to about 240 B per node.

This is still a large number that could seriously hamper performance. However, this really depends on the balance between the number of particles and the number of elements to discretize the fluid. If the number of particles is higher than the number of nodes the fluid memory extra requirements will not dramatically affect the overall performance, since the number of variables per particle is often already quite high in practice, especially when using all the terms in the MRE, which requires storing historic information. These memory requirements for the particles become even higher when collisions become relatively common. Nonetheless, this is certainly an important draw-back of the method.

In our code we have implemented the possibility of fixing a maximum number of nodes in a cloud to a predefined number, say, 20. The method is then modified to default to the standard method for the nodes where this is not enough. This option makes the method more robust.

Finally, note that the second order polynomial derived allows us to differentiate it twice to obtain the Laplacian directly, where this time the coefficients a_5 , a_6 and a_7 are needed. This requires some extra memory, but avoids the application of the method in succession, as discussed with the previous methods. This was indeed the approach proposed (and analysed) by Guo et al. [153], who obtained interesting results concerning the ultraconvergence of the method in a class of quasi-regular meshes.

The method of Pouliot et al. (FFC)

The method introduced by Pouliot et al. [282] provides an interesting alternative to the PPR. It results from imposing that the recovered nodal gradient must make *directional derivatives along all element edges equal to that obtained from the FEM approximation u_h* . Note that, while the derivatives are in general discontinuous across elements for piecewise linear elements, the directional derivatives *along the edges* are in fact continuous.

Let us describe the method. Consider an element edge formed by two nodes a and b . Once again, assume that the recovered gradient is denoted by \mathbf{g} . Since it is a linear field, its

value at the edge midpoint is given by

$$g_{h,i}\left(\frac{1}{2}(\mathbf{x}_a + \mathbf{x}_b)\right) = \frac{1}{2}(G_i^a + G_i^b) \quad (4.85)$$

The directional derivative of u_h along the edge is trivially equal to

$$\frac{U^a - U^b}{\|\mathbf{l}_{ab}\|} \quad (4.86)$$

where \mathbf{l}_{ab} is the normalized vector joining a to b . Therefore, the condition we are seeking can be written as

$$\mathbf{g}_h(\mathbf{x}_{\text{midpoint}}) \cdot \mathbf{l}_{ab} = \frac{1}{2}(G_i^a + G_i^b) l_{ab,i} = \frac{U^a - U^b}{\|\mathbf{l}_{ab}\|} \quad (4.87)$$

Multiplying both sides (with the dot product) of Eq. (4.87) by \mathbf{l}_{ab} we obtain a system of three equations with six unknowns for the edge (in 3D). But since in the global system there must be three equations per node, we just assign the system to both nodes in the usual assembly process. This can be expressed as saying that the edge contribution to the global matrix is given by the following system:

$$\underbrace{\begin{bmatrix} \mathbf{L}_{ab} & \mathbf{L}_{ab} \\ \mathbf{L}_{ab} & \mathbf{L}_{ab} \end{bmatrix}}_{\mathbf{M}_{FFC}^{ab}} \underbrace{\begin{bmatrix} \mathbf{G}^a \\ \mathbf{G}^b \end{bmatrix}}_{\mathbf{G}_{FFC}^{ab}} = 2 \underbrace{\frac{U^b - U^a}{\|\mathbf{l}_{ab}\|}}_{\mathbf{f}_{FFC}^{ab}} \underbrace{\begin{bmatrix} \mathbf{l}_{ab} \\ \mathbf{l}_{ab} \end{bmatrix}}_{\mathbf{f}_{FFC}^{ab}} \quad (4.88)$$

where $\mathbf{L}_{ab} = \mathbf{l}_{ab} \mathbf{l}_{ab}^\top$, $\mathbf{G}^a = [G_a^1, G_a^2, G_a^3]^\top$ and $\mathbf{G}^b = [G_b^1, G_b^2, G_b^3]^\top$; and that the assembly is done in the usual way. The assembled global matrix is invertible for most irregular simplex matrices but becomes singular for regular matrices [282], due to the redundancy of a subset of the equations.

In the latter case, instead of adding extra equations to the system, Pouliot et al. propose to *stabilize* the system. They give two alternatives of how to modify Eq. (4.88) to stabilize the assembled system:

$$(\mathbf{M}_{FFC}^{ab} + \epsilon h \mathbf{S}^{ab}) \mathbf{G}_{FFC}^{ab} = \mathbf{f}_{FFC}^{ab} \quad (4.89)$$

where

$$\mathbf{S}^{ab} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\mathbf{I}_{3 \times 3} \\ -\mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \quad (4.90)$$

where $\mathbf{I}_{3 \times 3}$ is the three-by-three identity matrix and ϵ is a 'small parameter', that must be made small enough to minimize the error while maintaining the invertibility of the global matrix. Pouliot et al. recommend $\epsilon = 10^{-6}$. Note that the addition of the matrix \mathbf{S}^{ab} is

equivalent to adding the system $\mathbf{S}^{ab} = \mathbf{0}$, which can be interpreted as requiring that the directional derivatives of the gradient to be all zero along the edges.

The other stabilization option they proposed can be expressed as

$$(\mathbf{M}_{FFC}^{ab} + \epsilon \mathbf{I}_{6 \times 6}) \mathbf{G}_{FFC}^{ab} = \mathbf{f}_{FFC}^{ab} + \epsilon \mathbf{G}_M^{ab} \quad (4.91)$$

where \mathbf{G}_M^{ab} are the nodal values obtained by an alternative recovery method. Pouliot et al. proposed to use the method presented in [32], although they accepted the possibility of using other methods, such as the PPR.

This method is attractive for the following reasons. First, it has been shown to be perhaps the most accurate recovery method to date in many cases, surpassing the PPR in most tests reported in [282], who remarked its superiority for approximations near boundaries. Second, it requires virtually no extra memory, since all the data structures needed were already in existence for solving the continuous phase problem. And third, its edge-based approach is fully compatible with the finite element framework, which means that we could automatically take advantage of all the pre-existing tools for the FEM, such as MPI-parallelism etc.

In the following sections the most important alternatives reviewed above are tested. We discuss their adequacy to calculate the derivatives required for a particle-laden simulation code and pick the most suitable ones for the examples that follow.

4.4.2 Comparison of the different recovery approaches

In order to check the accuracy and robustness of the different alternatives, in this subsection we test the different methods on different example functions and meshes. Here we focus on their performance on cubic meshes, representative of their suitability in bulky domains, where most of the points are far from the boundaries.

Our approach consists in imposing the nodal values to match that of analytic fields and apply the derivative recovery using these nodal values. Since we are mainly interested in the material derivative to calculate \mathbf{F}_U , we will measure the error in calculating the convective term $\mathbf{u} \cdot \nabla \mathbf{u}$. We will also be looking at the error of the recovered Laplacian $\Delta \mathbf{u}$.

We will consider the following measures of the error:

$$E_2 = \frac{\|\mathbf{f}_h - \mathbf{f}\|_2}{\|\mathbf{f}\|_2} \quad (4.92)$$

where \mathbf{f}_h is the approximated field, \mathbf{f} its analytical counterpart and where the approximate L_2 -norm is defined as

$$\|\mathbf{f}\|_2 = \left(\sum_{i=1}^{n_{\text{nodes}}} \|\mathbf{f}(\mathbf{x}_i)\|^2 \Omega_i \right)^{1/2} \quad (4.93)$$

We additionally consider the maximum relative error, defined as

$$E_{\text{max}} = \frac{\max_{\Omega} \{\mathbf{f}_h - \mathbf{f}\}}{\|\mathbf{f}\|_2} \quad (4.94)$$

We consider the domain $\Omega_4 = [0, 4] \times [0, 4] \times [0, 4]$ in all cases. Three families of meshes will be tested: an irregular mesh and two regular meshes, one of conformal and the other of non-conformal type. Each family of meshes will be subdivided three times, starting from the coarsest, to the finest.

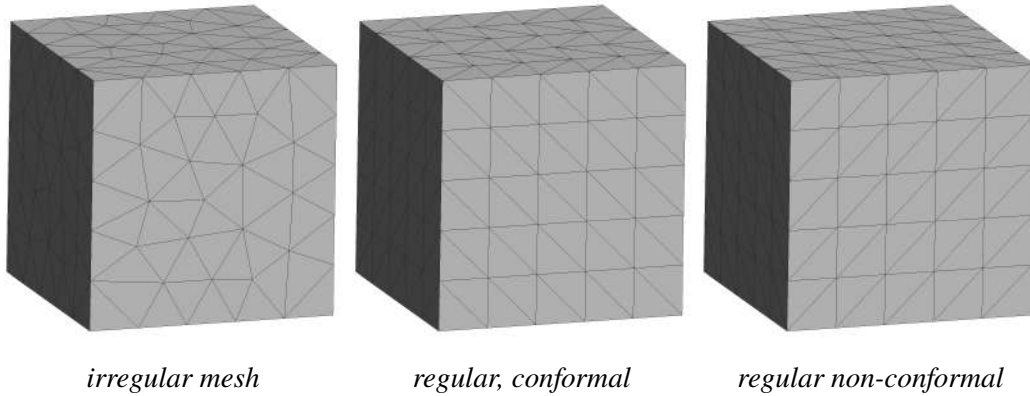


Fig. 4.4 Coarsest meshes considered for each of the three families.

We have made a selection of four method combinations for each differential operator. Those are:

- Standard method: Exactly as described in Section 4.4.1. For the Laplacian, we applied the gradient recovery twice to each component, retrieving the trace of the resulting matrix.
- L_2 -projection: As described in Section 4.4.1. This technique was applied twice as in item above for the Laplacian.
- lumped L_2 -projection: Variation of the method in Section 4.4.1, where directly $\mathbf{u}_h \cdot \nabla \mathbf{u}_h$ is projected, using the lumped mass matrix. For the Laplacian, instead, the lumped matrix is used to project the gradient, using it again to obtain the divergence of its

diagonal components. This approach is mathematically equivalent to the standard method, as shown in Section 4.4.1, so it serves as a verification test.

- FFC: Method described in Section 4.4.1 with the stabilization from Eq. (4.89). The method was applied twice as in first two items for the Laplacian.
- PPR: Method described in Section 4.4.1 (modified algorithm). The Laplacian was obtained by applying the second derivative to the polynomial fit, as explained at the end of Section 4.4.1; and not by applying the method twice, as proposed in [153].

We have used two analytic fields to compare the different combinations above. Next we present the results for each of the two fields analysed.

Product of sines field

The first field $\mathbf{u} = [u_1, u_2, u_3]^\top$ is defined in Ω_4 by

$$\begin{aligned} u_1 &= \sin(\omega x_1) \sin(\omega x_2) \sin(\omega x_3) \\ u_2 &= 0 \\ u_3 &= 0 \end{aligned} \tag{4.95}$$

where ω is taken to be equal to $\pi/2$.

The results of the derivative recoveries are shown in Fig. 4.5. Both E_2 and E_{\max} are shown, the former with full symbols and the latter with hollow symbols. On the left column we find the results corresponding to the convective derivative, for the different types of meshes and, on the right column, the analogous results for the Laplacian.

We will focus on the E_2 error as it is more representative of the majority of elements and is clearly much better behaved (more systematic) in all cases. One must however keep in mind that the maximum error is often more than an order of magnitude greater than E_2 for most methods, as this can have serious consequences in particle-laden flow simulations. For instance, one must be particularly attentive to the existence of narrow locations, where a large proportion of particles may cross a certain region with a potentially large error in the derivative.

In general terms, the FFC and PPR methods are the only methods that exhibit close to $O(h^2)$ behaviour in the error in most situations for the first-order derivatives, both for regular and irregular meshes. For the irregular meshes, the FFC presents slightly higher convergence rate, at an average for the first column of 2.07 compared to 1.93. Additionally, the FFC presents slightly lower errors for the coarser meshes, which leads to an improvement of around half an order of magnitude when compared to the PPR for the finest meshes. A

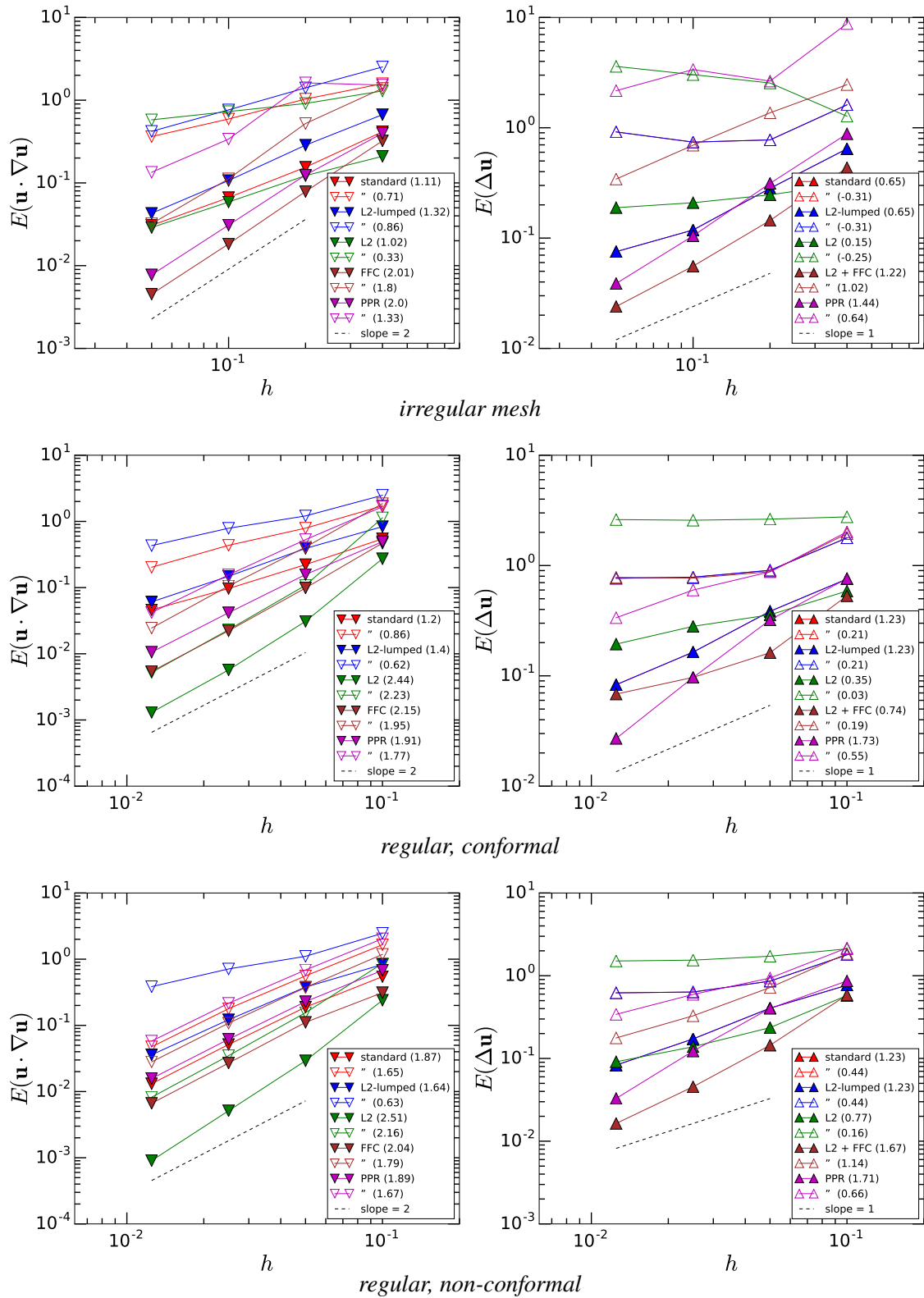


Fig. 4.5 Recovery errors for the product of sines field, Eq. (4.95) (left column: convective derivative; right column: Laplacian), as a function of the mesh size h , normalized by the characteristic scale $\pi/\omega = 2$. Full symbols are used for E_2 measure of the error, with the corresponding E_{max} values in hollow symbols. The obtained order of the different methods, calculated with the two most refined meshes are indicated in parenthesis.

similar picture is found on right column (Laplacian) where we observe the expected decrease in the order of convergence for both methods, although for both the FFC and PPR methods the order of convergence remains higher than one. These results are coherent with the results presented in [282], which showed a moderate advantage of the FFC method with respect to the PPR.

The general behaviour of the FFC and PPR methods in regular meshes is broadly speaking similar for the convective derivative. Nonetheless, note that for the conformal-type mesh, the FFC shows a stagnation in its convergence rate for the Laplacian. This result is interesting, although compatible in principle with the results from [282], where the second-order derivatives recovery was not discussed. We will come back to this point with the next example field, nonetheless.

About the other methods, focussing on the first column of Fig. 4.5, it is clear that the standard approach and the lumped L_2 -projection approach have similar accuracies for the convective derivatives. The consistent L_2 -projection shows comparable accuracy for the irregular mesh but much better accuracy for the regular meshes, where it even surpasses the high-order methods (PPR and FFC), yielding extremely high accuracy, also for E_{\max} .

The second column however shows a different picture: the consistent L_2 -projection performs poorly, compared to the standard method and the lumped L_2 -projection, which here become identical, as expected, because they are mathematically equivalent in this case (the nonlinearity of the convective term breaks the equivalence for the first-derivatives case).

Ethier field

This field corresponds to a particular case of a family of analytic solutions of the Navier-Stokes equations found by Ethier and Steinman [121]. These equations are fully three-dimensional and resemble a pair of interlocking vortices.

$$\begin{aligned} u_1 &= -a (e^{ax_1} \sin(ax_1 + dx_3) + e^{ax_3} \cos(ax_1 + dx_2)) e^{-d^2t} \\ u_2 &= -a (e^{ax_2} \sin(ax_2 + dx_1) + e^{ax_3} \cos(ax_1 + dx_3)) e^{-d^2t} \\ u_3 &= -a (e^{ax_3} \sin(ax_3 + dx_2) + e^{ax_3} \cos(ax_1 + dx_1)) e^{-d^2t} \end{aligned} \quad (4.96)$$

where we have taken $a = \pi/4$ and $d = \pi/2$.

For this field the results are in general worse (larger relative error), as is perhaps to be expected due to the relatively stronger nonlinearities associated to this field. Let us analyse the particular differences observed in Fig. 4.6 with respect to those in Fig. 4.5.

For the irregular meshes and the regular, non-conformal meshes the higher-order methods (FFC and PPR) behave similarly, though now the PPR shows a slightly higher rate of

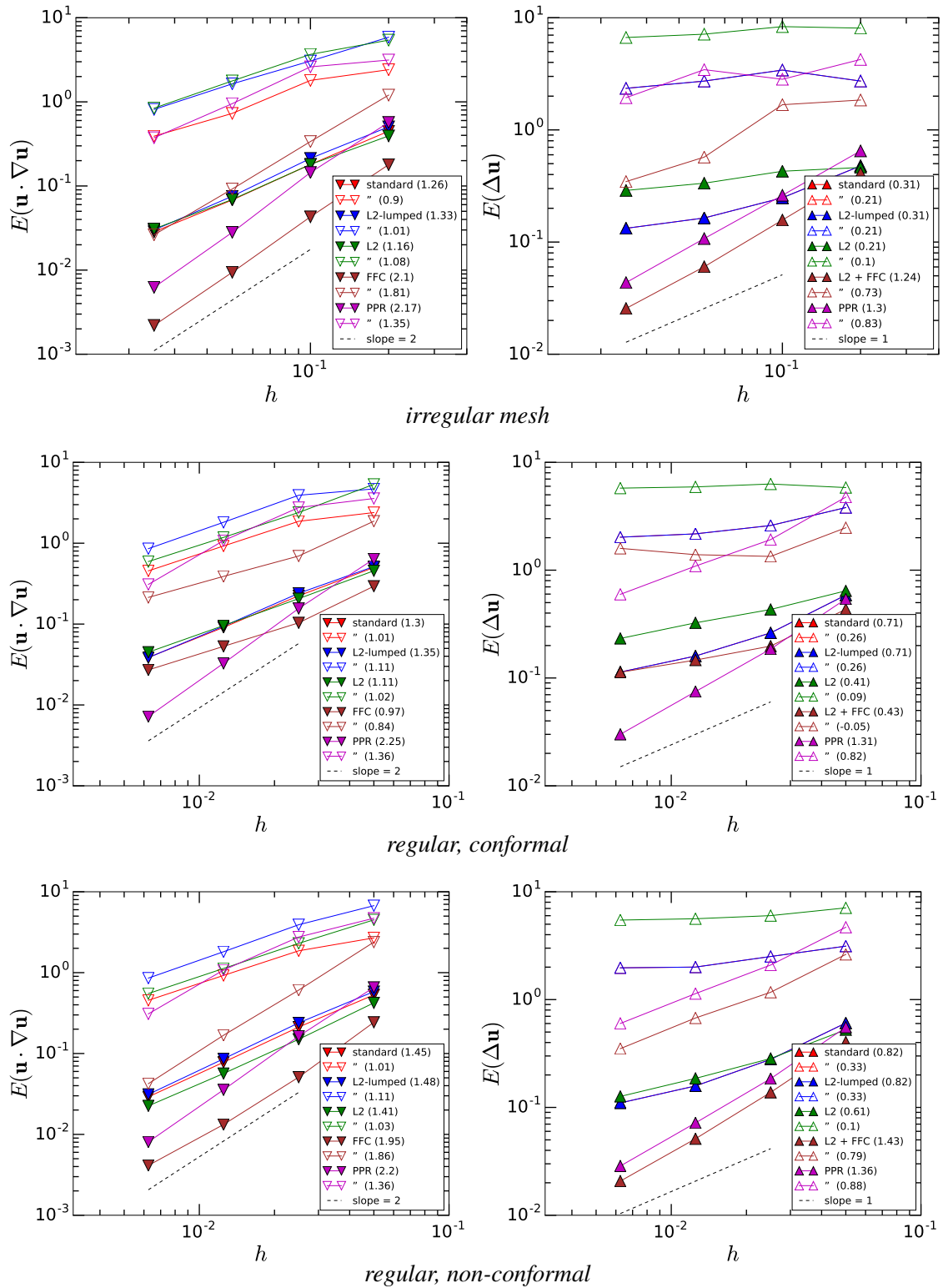


Fig. 4.6 Recovery errors for the Ethier field at $t = 0$ Eq. (4.96) (left column: convective derivative; right column: Laplacian) as a function of the mesh size h , normalized by the characteristic scale $\pi/a = 4$. Full symbols are used for E_2 measure of the error, with the corresponding E_{max} values in hollow symbols. The obtained order of the different methods, calculated with the two most refined meshes are indicated in parenthesis.

convergence. This however does not make up for its lower accuracy for the coarse meshes compared to the FFC.

A more important difference is observed for the regular, conformal mesh. Note that the FFC by far fails to attain its promised second order and becomes much less precise. Note that this loss of accuracy already manifested in the previous example field for the Laplacian. Here, the stagnation is confirmed and takes place both for the convective derivative and the Laplacian.

After very careful scrutiny, we could not find any errors in our implementation. We even compared the explicit matrices obtained for a test example proposed by Dr. Pouliot himself in a private communication, obtaining identical results. Later on, we found that a colleague [105] had run a test on the method by Pouliot, obtaining the desired order. However, we found out that these authors had used non-conformal meshes, which inspired our conjecture that the method is not robust for conformal meshes, precisely the type normally used in finite element simulations. We tried applying small random perturbations to the mesh and the results persisted, suggesting that the problem is related to the connectivity of the matrix.

An additional difference observed is the loss of the ultraconvergent (strictly more than one order over the expected accuracy) character of the L_2 -projection method, which moves closer to its expected order-one behaviour [394].

Discussion

We have reviewed several methods as viable alternatives for the derivative recovery operations required in a FEM-based particle-laden simulation code. From the results discussed above we can conclude that

- The standard approach offers a very efficient alternative with attractive properties such as small radius of influence (it does not require a large patch to work) and a predictable and robust behaviour that often exceeds the expected first-order behaviour for the convective derivative.
- the standard approach is not as accurate and robust for the calculation of the Laplacian, for which a higher-order method would be recommended
- The method of Pouliot et al. (FFC) is a very accurate method, perhaps the most accurate method when it works. This is especially true for irregular meshes, for which it shows very good properties, as previously found in the literature.

- However, the FFC suffers from low accuracy for some types of regular meshes, even when stabilized according to the original formulation. More work is needed to fully understand the origin of this issue, but we regard it as unsafe until it can be fully fixed.
- The method of Zhang and Naga (PPR) is accurate and robust in all situations. Its main drawback is its high memory requirements, as it is a patch method that only requires a moderate number of multiplications per node.

We therefore adopt as a default strategy the PPR with the standard approach fall-back for failed polynomial fits. When the number of unknowns is very big, we simply stick to the standard approach though acknowledging its lower accuracy.

Finally, we come to the conclusion that an accurate calculation of the Laplacian would probably require higher-order FEM methods in the first place, unless an extremely fine mesh is used. Fortunately, there are many situations where the Faxén terms can be neglected and we will therefore not consider these corrections in the forthcoming sections.

4.5 Forward coupling

Once the continuum phase solution has been determined (including the recovered derivatives) there must exist an *interpolation process* to update the information at the particles' locations (their centres), as they will in general fall at arbitrary, intermediate, positions within the domain ⁶. This process generally consists in the following steps

1. Locate the relevant computational cells for each particle
2. Calculate values from the cells

Let us next discuss each of these.

4.5.1 Search

For a given particle, the data that is needed to reconstruct the flow at its center is usually only a small subset of the total available information: that related to a neighbourhood of the point in question. Since the brute-force survey of all the discrete information (for each particle) is out of the question, it is necessary to implement a suitable *search algorithm* to relate each particle with the relevant data in an efficient way.

Our algorithm is based on the same technology used for the inter-particle search. Taking advantage of the abstract syntax offered by Kratos, we create a bin data structure [369] that

⁶Unless the particles' locations are used as computational points, see for instance [67]

pre-classifies objects, in this case tetrahedra, so that the search of the host element for points can be achieved then in constant time. Since we are dealing with constant fluid meshes here (except when using an ALE method, see Section 4.7), there is no need to update this bin data base and it can be reused at every time step.

The search can be further optimized by using the fact that particles cannot jump several elements between time steps, see [215]. In our examples, we have not found the search related to the coupling to ever dominate other costs and thus we have not implemented such methods. Nonetheless, this is an interesting topic for future research.

4.5.2 Interpolation

For some methods, such finite volumes or finite differences, the solution is a discrete set of values associated with a discrete set of spatial points. Since the position of the particle will surely not coincide with any of these points, the information must be interpolated, incurring in an additional error that should be taken into account. Consequently, the approximation order of the interpolation scheme is usually taken to be of the same order as the numerical method's own [11, 62]. Sometimes it is taken higher [236]. The goal is to preserve the order of accuracy of the velocity once transferred to the particle.

In other cases, such as in the FEM, the approximated solution is defined in the totality of the domain and it is a matter of simply calculating the needed values at particular positions. For instance, we use linear tetrahedra, for which the velocity at the location of a given particle, labelled P , is calculated as

$$u_j(\mathbf{x}_P) \approx u_{h,j}(\mathbf{x}_P) = N^{b_P}(\mathbf{x}_P)U_j^{b_P} \quad (4.97)$$

for $j = 1, \dots, n_{\text{dim}}$ and for b_P taking values from the set of four global indices corresponding to the four nodes of the tetrahedron that contains the particle P , which is known thanks to the search algorithm. and similarly one could interpolate the velocity gradient if the nodal values were known. This operation entails no further errors apart from those associated to finite-precision arithmetic.

On the other hand, even when the approximate solution is defined everywhere, there might be reasons to employ interpolation instead. For instance, with pseudo-spectral methods interpolation (using the coefficients of the shape functions as the discrete values) is preferred over evaluation because of the large computational cost of the latter [20]. Note that in fact in this case the order of approximation is often *lower* than that of the solution (spectral methods can achieve exponential approximation order under optimal conditions). In practice however, this error can be made sufficiently small by considering high-order polynomials, see [20].

See also [354, 351] for a recent and detailed review of several interpolation schemes used for particle-laden flow simulations with pseudo-spectral methods for regular grids. In any case, the interpolation order is not recommended to be lower than third-order ($O(h^3)$ behaviour of the error) to obtain very accurate results [377, 330] in the context of DNS using PS methods. Note that this reinforces the notion that it is important to strive for achieving at least this accuracy in the derivative recovery.

There is a ‘temporal discretization’ associated to forward coupling. This is because it is common to have a different time stepping scheme for the particles and for the fluid, since these two phases have different smallest characteristic scales. We employ an *updated-fluid* strategy, by which the fluid is updated first, and then the particles motion is integrated until they meet the fluid, in one or (typically) more sub-steps. As the particles change location during these sub-steps, we use a linear interpolation of the past and present values of the fluid, interpolated at the present location of the particles. For the case of pre-calculated fluid, where the fluid has been previously calculated and stored and the particles are moved with the read fluid data, it would be possible to use higher-order temporal schemes, as the fluid at all time steps is known. Such schemes would become particularly relevant for higher Stokes numbers and less so for small Stokes numbers, as the quasi-stationarity becomes a better and better model for the fluid motion, as seen by the particles, between successive fluid time steps. We have not pursued this here, and we leave the study of such schemes for future works

4.6 Application example: T-junction bubble trapping

We report in this section a first application example to test the coupled DEM-CFD approach, where the fluid is calculated using linear finite elements and a one-way strategy is used to move the particles according to an extended version of MRE for higher Reynolds numbers. The problem is taken from [356], where it was studied both experimentally and numerically, using finite-volume based commercial software.

The system consists in a T-type tube junction through which a steady water flux is imposed; see Fig. 4.7. The flow is laden at the entry with a number of bubbles, which are idealized as spheric, rigid particles (empty glass spheres were also used in part of the experiments). The phenomenon to be studied here is the bubble trapping that was observed to occur around the junction point for a remarkably wide range of parameters.

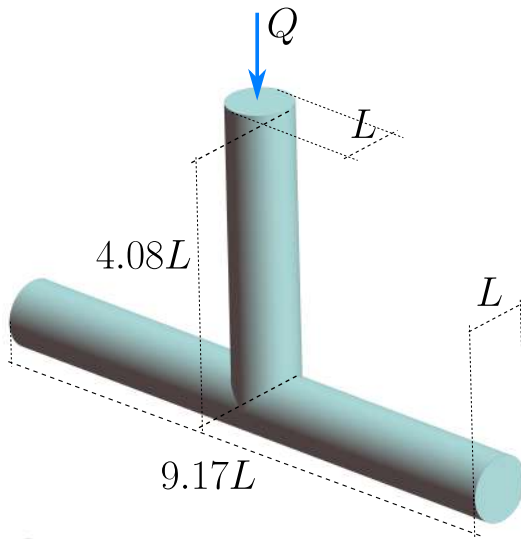


Fig. 4.7 Geometry of the T-junction considered, with characteristic dimensions. The inflow flux Q is indicated with a blue arrow.

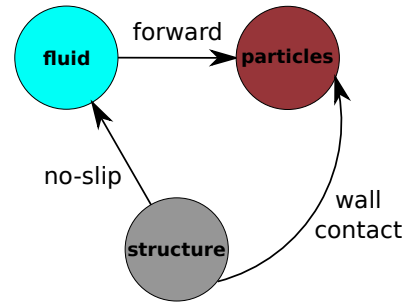


Fig. 4.8 Conceptual diagram of interactions taken into account in the t-junction example; to be compared to Fig. 4.2.

4.6.1 Flow simulation

Our goal is to run a fluid flow simulation that is sufficiently accurate for a sufficiently long time to give a realistic representation of the real flow. The flow solution (the nodal values of the velocity field) is stored in a n HDF5 format file [341] and read afterwards, post-processed and used to integrate the particles positions. The advantage of this strategy is that many particle simulations can be run with the same flow. In this case the flow simulation was clearly the most expensive part, given that the numbers of particles required were moderate and that the particles could be modelled as relatively soft spheres.

The flow data is read from a file, but the derivatives are recovered during the particles simulations. This is acceptable, as the derivative recovery is much less costly than solving the fluid system. In this way we were able to keep the size of the fluid post-process files manageable.

Inputs

The boundary conditions were set to no-slip on the tube walls, uniform velocity at the inlet and constant and equal pressure at the two outlets. Gravity was neglected as in [356].

The particles are represented by spheres whose motion is calculated according to the soft-sphere discrete element model described in Appendix A. The particles are introduced

Table 4.2 Material parameters considered in the T-junction example

Parameter	Value	Description
Fluid parameters		
ρ_f	1000 kg/m ³	density of fluid
ν	1×10^{-3} s/m ²	kinematic viscosity
Particles parameters		
ρ_p	150 kg/m ³	density of bubbles
COR	0.2	coefficient of normal restitution

in the domain through an *inlet* surface, which is a flat triangular mesh where the nodes are the inception location of the particles. The initial velocity of the particles is set to be equal to the average flow velocity, so that a small relative velocity is present from the beginning. This initial slip velocity is given enough time to dissipate though, by the time the particles reach the t-junction through the action of the hydrodynamic interaction forces. The principal particle parameters are summarized in Table 4.2. The coefficient of normal restitution (COR) is the ratio between the incident impact velocity and the reflected impact velocity, such that the proportion of kinetic energy remaining after the impact is roughly given by COR^2 for direct impacts. For air bubbles in water it is related to the Stokes number of the bounce, which can be estimated as [386]

$$St_{\text{bubble}} = \frac{2}{9} \frac{\rho + 1}{\nu} a U_{\text{bounce}} \quad (4.98)$$

where U_{bounce} can be (safely, since we want to show it is small) estimated as U , the characteristic scale of the flow. Substituting in the biggest sizes considered for the particles ($a_{\text{max}} = 2.4 \times 10^{-4}$ m) and Reynolds numbers $Re = 960$ we obtain $St_{\text{bubble}} \lesssim 10$, for which it is realistic to take $COR = 0.2$; see [386]. The stiffness of the sphere was set to 100 Pa, which was calculated to avoid excessive indentations during the simulation. Particle-surface contacts are not prevalent in any case and do not influence the motion of the particles significantly. However, it is advantageous to have a soft contact, as it allows taking larger time steps for the integration of the trajectories.

Hydrodynamic model

To model the motion of the particles we will start by using the same hydrodynamic model proposed in [356], which is a simplified version of Eqs.(4.1) and (4.4), where it is assumed

that $\mathbf{F}_A, \mathbf{F}_H, \mathbf{F}_L, \equiv 0$ and the model of Schiller and Naumann (1935) accurate for $Re_p < 1000$ (see [222]), is used for the drag coefficient

$$C_D = \frac{12}{Re_p} (1 + 0.15 Re_p^{0.687}) \quad (4.99)$$

For \mathbf{F}_U these authors use the formulation based on the pressure gradient and the divergence of the deviator stress tensor, equivalent to our formulation in the continuum limit. The equivalence is obtained by expressing \mathbf{F}_U in terms of the stress tensor using the Navier–Stokes equations:

$$\mathbf{F}_U = m_f \frac{D\mathbf{u}}{Dt} = \frac{4}{3} V_p \rho_f \frac{D\mathbf{u}}{Dt} = \frac{4}{3} V_p (\nabla \cdot \boldsymbol{\tau} - \nabla p) \quad (4.100)$$

where the weight is neglected in this case, following the criterion in [356], where it is argued that $Fr^2 \ll 1$. Note that the data in Table 4.2 lead to $Fr \sim 0.3$; and $Fr^2 \sim 0.1$, which casts some doubt on the validity of such assumption. We will nonetheless keep this assumption to simplify the discussion.

Furthermore, the deviator stress tensor contributions are deemed small in [356] and neglected. This is a common practice that is nonetheless unnecessary if one uses the alternative formulation based on the material derivative, although the latter practice does require two more derivative recoveries (we must recover the convective derivative instead of just the gradient of the pressure), plus an additional time derivative. We thus keep this term, as we are using the formulation based on the material derivative.

Finally, no rotational degrees of freedom are considered. In this system the rotational velocity of the particles is not expected to vary greatly from half the local fluid vorticity. This can be justified based on the very small rotational Stokes number, $St_\Omega \sim 1 \times 10^{-4}$, based on Eq.(2.23) and the data in Table 4.2. Such small value implies that the bubbles angular velocity relaxes very rapidly to that of their surrounding fluid. Eq.(2.23) is strictly valid only under the small particle Reynolds number hypothesis, but here it should be sufficiently accurate to provide *an order-of-magnitude* estimate.

4.6.2 Trapping probability

We attempt to reproduce in this paragraph a numerical result presented in [356], where the proportion of particles that become trapped is measured, for a fixed Re and several particle sizes. In order to do this, we run the simulation for a total time that exceeds the characteristic time scale about thirty times, see Table 4.3. A fixed flux of particles is imposed using the inlet mesh depicted in Fig. 4.10, which provides a sufficiently fine discretization of the whole cross-section. The particles are considered trapped when they remain inside the control

Table 4.3 Simulation parameters for the trapping probability calculation

Parameter	Value	Description
N_{total}	36 204	total number of particles injected
$z_{\text{init, max}}$	0.468	maximum value attained by z_{init}
T_{final}	31.42	dimensionless final time
$T_{\text{inj, max}}$	26	dimensionless max. injection time considered in statistics

volume Ω_{trap} , which is the parallelepiped comprised between two cross-sections located at opposite sides of the junction, at a distance of $3L$ from the plane of symmetry. Then the proportion of particles trapped is computed as

$$P = \frac{N_{t,\text{trap}}}{N_t} \quad (4.101)$$

where $P \in [0, 1]$ represents the proportion of particles trapped, N_t is the total number of particles injected before time t and $N_{t,\text{trap}}$ is the number of these particles present by the end of the simulation inside the control volume Ω_{trap} that were injected before time t (i.e. newer particles are not counted). The results shown in Fig. 4.11 correspond to the numerical predictions for $P(z_{\text{init, max}})$ for several particles' radii; where $z_{\text{init, max}}$ is the distance from the geometry's plane of symmetry at which the particles are injected, normalized by L .

4.6.3 Discussion

We have applied our CFD-DEM strategy to simulate the phenomenon of bubble trapping in T-shaped junctions. We used a one-way strategy that was sufficient to reproduce the trapping and to obtain a qualitative picture similar to that obtained in the earlier work of Vigolo et al. [356].

Our results confirm that the overall approach can be used for the detection of bubble trapping in T-junction systems and other, similar, moderate- Re systems. The one-way coupling strategy allows the application of an efficient methodology, whereby a good fluid simulation is only run once, while a large number of cases with particles can be run at a reduced cost by reading the pre-computed results. The derivative recovery may or may not be calculated during the fluid solution, allowing for extra freedom in later simulations. The elimination of inter-particle interactions allows to further speed up the simulations, since it is

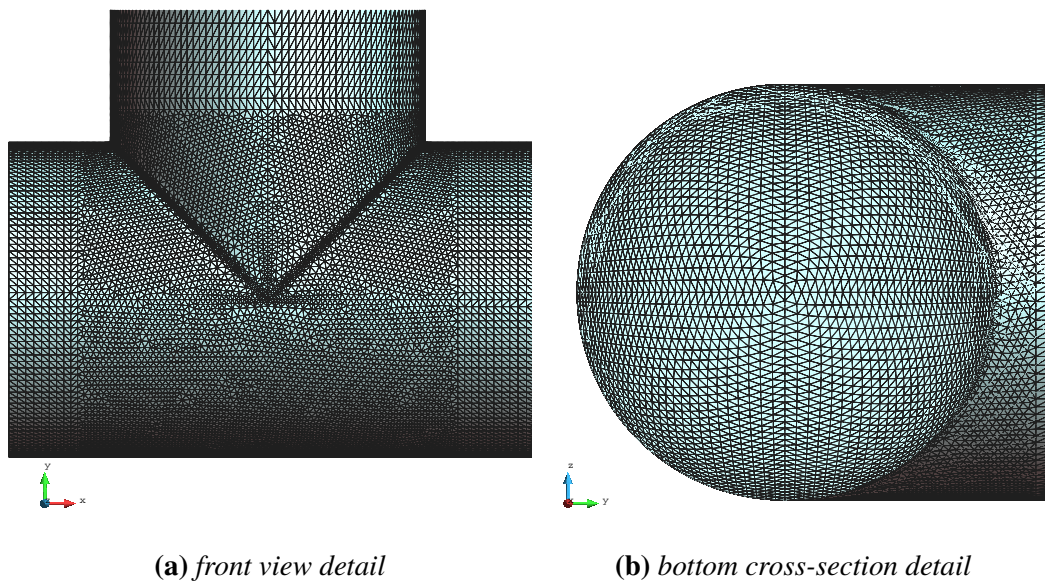


Fig. 4.9 Mesh used for the trapping probability calculations (6 057 731 linear tetrahedra elements). The mesh is structured for most of the tubes length. Near the junction it switches to an unstructured mesh.

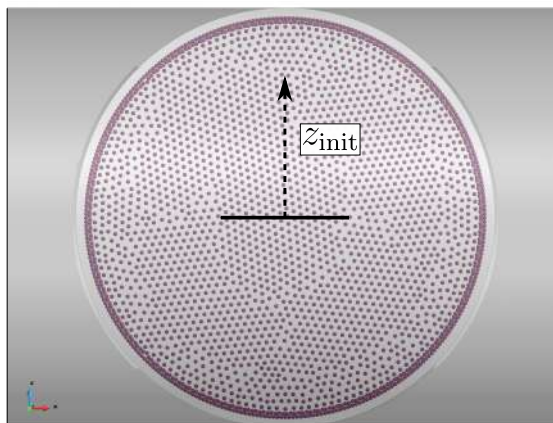


Fig. 4.10 Inlet mesh used to introduce particles into the domain. The red points are the fixed locations of injection. The lighter grey crown is a gap between the outermost injection points and the walls, provided so that enough room is provided to avoid erroneous initial indentations at the wall.

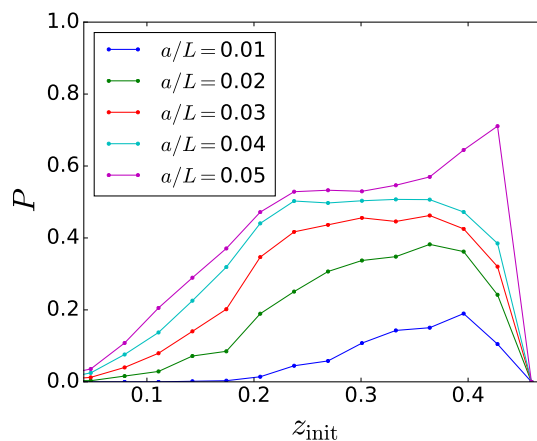


Fig. 4.11 Trapping probability for different particle sizes as a function of z_{init} . Only particles that remained more than fourteen times the characteristic time are taken into account.

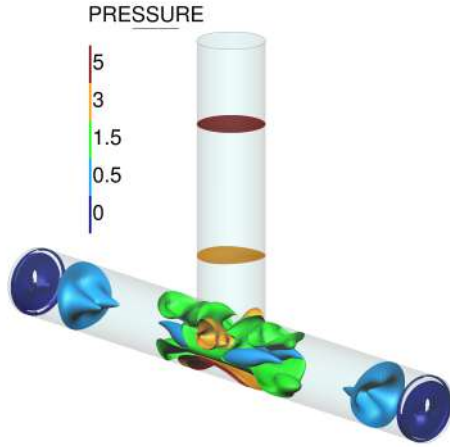


Fig. 4.12 Pressure isosurfaces at nondimensional time $t/T = 26$.

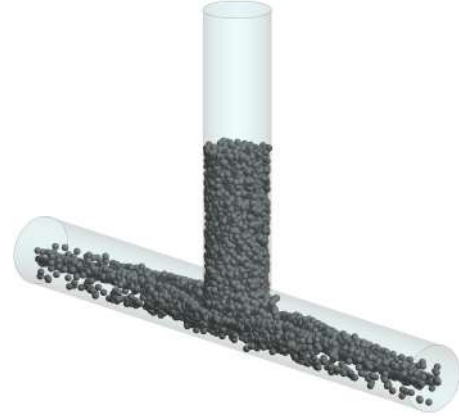


Fig. 4.13 Snapshot of non-interacting particles at $t/T = 26$ (about 20 000 particles).

possible to use an artificially increasing number of particles at a time, decreasing the total simulation time required to reach statistical significance.

Our numerical results do present difference with respect to the ones presented in [356] were the trapping probability was observed to be (quite consistently) higher for all the radii. Furthermore, their predicted maximum trapping was found to be closer to $z_{\text{init, max}} = 3.5$, whereas our prediction is closer to 4. The difference persisted even as we refined our mesh and so there remains some uncertainty regarding the accuracy of the method that should be further tested, perhaps against physical experiments or fully-resolved simulations; despite the high computational cost associated to the latter.

We can identify a number of factors that could explain the observed differences. The most obvious difference is that our formulation includes the deviator stress tensor term $\mathbf{F}_{U,\tau}$ (first term of the RHS of Eq.(4.100)) implicitly, while it was neglected in [356] as it was deemed small. But in fact, a simple scaling analysis confirms this hypothesis, since (using $\mathbf{F}_U \propto \frac{D\mathbf{u}}{Dt} = -\nabla p + \nu\Delta\mathbf{u}$)

$$\frac{|\mathbf{F}_{U,\tau}|}{|\mathbf{F}_U|} = \frac{\nu\Delta\mathbf{u}}{\frac{D\mathbf{u}}{Dt}} \sim \frac{\nu\frac{U}{L^2}}{\frac{U^2}{L}} \sim \frac{\nu}{UL} \sim Re^{-1} \ll 1 \quad (4.102)$$

as the Reynolds number is close to 400. Note that in Eq. (4.102) we have assumed that the material derivative is of the same order as the convective derivative.

Another possibility is the existence of differences in the accuracies of flow solution itself, which could exist between the two works, including the derivative recovery. This is difficult

to test though, since the detailed flow information is not available for comparison. More work is needed to produce robust guidelines for the application of our CFD approach safely; including mesh tolerances and convergence criteria for the velocities and for their derivatives.

Other possible sources of differences include Stokes drag due to the time integration of the particles trajectories (although we have proved the accuracy of our algorithm extensively in Chapter 3) and differences in the initial injection positions (it is not detailed in [356]), the level of statistical convergence (the number of particles is not given either).

Finally, the influence of other neglected effects requires further study, especially if precise quantitative predictions are the goal, such as

- the effect of the history force
- the inter-particle effects (three and four-way coupling)
- the influence of the Faxén terms

We have nonetheless run a simulation to check the magnitude of the effects of including the history force. A slip velocity estimate based on a similar argument as that used in Eq.(2.89), but this time using the drag force formulation in Eq.(4.99). Doing this and solving for the slip velocity we obtain $W \in [0.018, 0.042]$ for all the radii studied, which results in $Re_p \in [0.09, 0.2]$, for which the Basset–Boussinesq history model can be safely applied. In case higher particle Reynolds numbers had occurred, the approximate method by [108] could have been used instead (without major modifications of the algorithm described in Chapter 3), which showed good predictions in the range $Re_p \in [9, 853]$. The results shown in Fig. 4.14 indicate that the effects might not be negligible in this range, and so it is probably preferable to include the effect in cases where accurate quantitative results are required.

4.7 Application example: Particle impact drilling

In this section we use a one-way coupled strategy to analyse an engineering system that has been studied very little either with numerical methods or experimentally: the particle impact drilling (PID) method. Among the scarce exceptions we find the numerical works [87, 185, 289], all based on finite volume method, and the experimental work [375], that studied the rate of erosion of the substrate under different impact conditions.

This drilling technology is used in the oil and gas industries to achieve greater rates of penetration than the traditional alternatives. In its basic features, a PID drill-bit is similar to a conventional one: it consists of a cutter fixed at the end of a rotating tube through which the drilling mud is pumped from the surface. The bit has apertures that allow the mud to

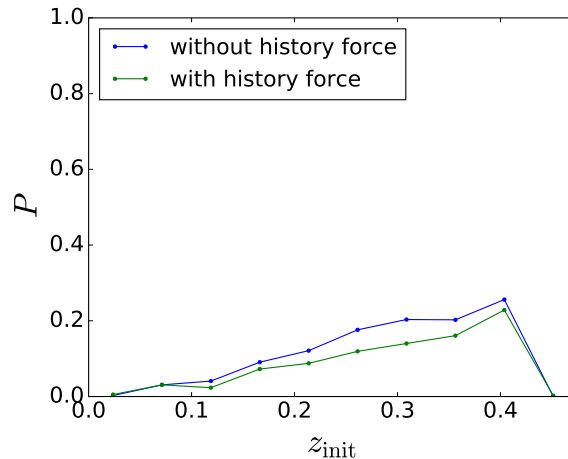


Fig. 4.14 Trapping probability for $a/L = 0.01$ as function of z_{init} , comparing the results with and without the effect of the history force.

flow out of its tip, cleaning the cuttings and dragging them back to the surface through the annulus contained between the outer surface of the tube and the hole's casing. A PID system is unique in that the mud flow is laden at the surface with small steel balls to help erode the rock and increase the rate of penetration. When the balls reach the tip of the drill-bit, they are violently accelerated as they pass through any of the particularly narrow apertures (nozzles) at the tip of the drill bit, where they acquire the necessary kinetic energy to effectively erode off material from the rocky bed. Inside the nozzles the fluid can attain velocities in excess of 200m s^{-1} . A schematic diagram of the workings of the system is shown in Fig. 4.15.

Our coupling strategy is represented by the diagram in Fig. 4.16. We include two new types of interactions with respect to Section 4.6: the inter-particle interactions, as it has been observed that they have a strong influence in the particles' flow, and a particle-to-structure interaction in the form of a qualitative wear. This is a measure of the strength and frequency of the particles' impacts on the solid surfaces and it has been considered as an 'interaction', although is really only a matter of post-processing and does not couple the solid solution with that of the particles. This type of interaction could also be seen as a trivial or degenerate type of interaction.

4.7.1 Problem data

The work related to the present application example was done under the terms of a private consultancy contract with a consultancy firm that also provided the geometry and input parameters.

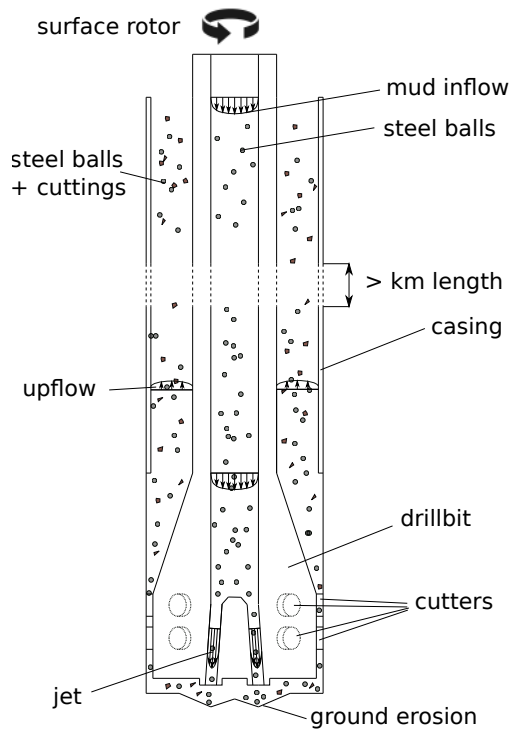


Fig. 4.15 Diagram of workings of the PID wellbore, showing inflow mud with steel particles, and up-flow cuttings wash-up.

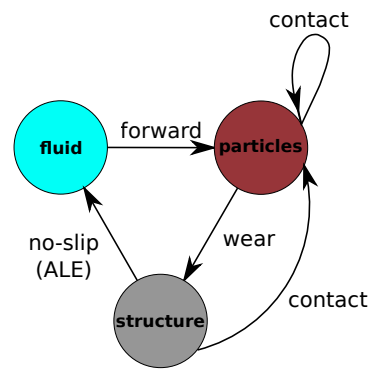
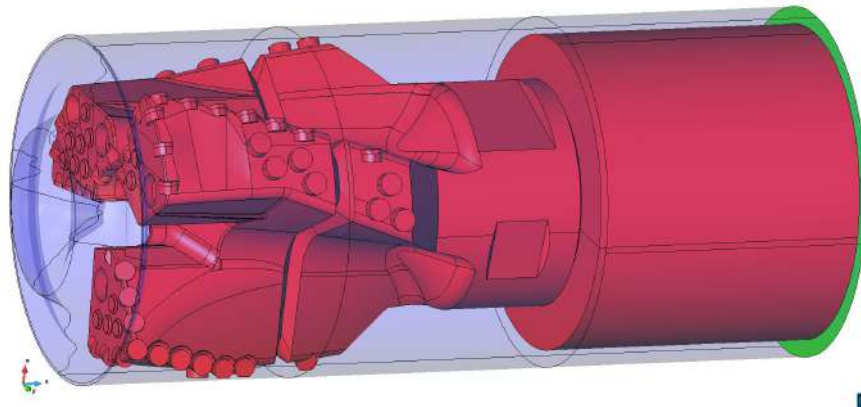


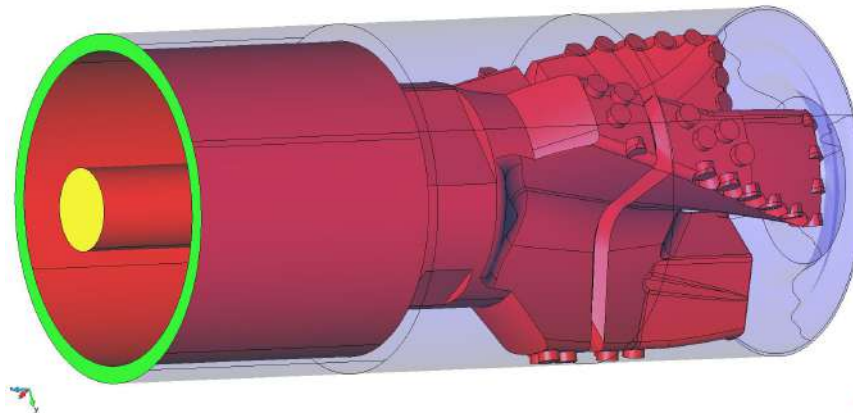
Fig. 4.16 Conceptual diagram of interactions taken into account in the PID example; to be compared to Fig. 4.2.

Domain geometry

The geometry of the drill-bit and the surfaces defining the rocky bed and casing are shown in Fig. 4.17. The surfaces in red belong to the rotating part of the domain, while the casing and ground surfaces are shown in a transparent blue. The inlet surface, at which the inlet flow condition is imposed is marked yellow, while the outlet surface (imposed normal traction) is marked in light green. In Table 4.4 we summarize several useful measurements derived from the geometry.



bottom-up view



top-down view

Fig. 4.17 Depiction of the domain geometry.

Table 4.4 Geometric measurements

Parameter	Value	Description
D_{inlet}	0.053 975 m	interior diameter of inner tube
D_{nozzle}	0.007 69 m	interior diameter of nozzles
L_{nozzle}	0.086 m	average length of the nozzles
V_{total}	0.0111 m ³	total fluid domain considered
V_{internal}	0.001 53 m ³	fluid of internal domain (between inlet surface and tip of nozzles)

Parameters

The values that have been kept fixed throughout this study are summarized in Table 4.5.

Table 4.5 Physical parameters considered in this work

Parameter	Value	Description
Fluid parameters		
ρ_f	1294 kg/m ³	density of fluid
n	0.761	flow behaviour index
K	1.24 Pa s ^{0.761}	flow consistency index
Particles parameters		
d_p	0.001 981 2 m	diameter of steel particles
ρ_p	7850 kg/m ³	density of steel particles
Operation conditions		
Q	0.029 02 m ³ /s	fluid flux
N_{inlet}	$1.5 \times 10^5 \text{ s}^{-1}$	number of particles flux
Ω_{drill}	$-2 \pi \text{ s}^{-1}$	angular velocity of drill-bit

4.7.2 Drilling fluid model

The typical drilling mud used in the oil and gas industries has a non-Newtonian behaviour. This means that its motion is not well-approximated by the standard Navier-Stokes equations, which must be modified. A common type of model used in this context is the Herschel–Bulkley

fluid [214]. This type of fluid departs from the Newtonian behaviour in two fundamental aspects:

1. The fluid has a finite *yield stress*, which is a threshold below which no flow takes place and the fluid behaves essentially as a solid.
2. For shear stresses larger than the yield stress, the viscosity changes with the shear rate.

The constitutive equation for this type of fluids is commonly written as a generalized Newtonian fluid by replacing the constant viscosity by the *effective viscosity*, as a function of the flow:

$$\boldsymbol{\tau} = 2\mu_{\text{eff}}(\dot{\gamma})\mathbf{S} \quad (4.103)$$

where $\dot{\gamma}$ is the local strain rate, given by

$$\dot{\gamma} = \sqrt{2\mathbf{S} : \mathbf{S}} \quad (4.104)$$

where

$$\mathbf{S} = \frac{1}{2} (\nabla \mathbf{u} + (\nabla \mathbf{u})^\top) \quad (4.105)$$

In Herschel–Bulkley fluid the functional dependence of the effective viscosity on the strain rate reads

$$\mu_{\text{eff}} = \begin{cases} \mu_0 & \text{if } \dot{\gamma} \leq \dot{\gamma}_0 \\ K|\dot{\gamma}|^{n-1} + \tau_0\dot{\gamma}^{-1} & \text{if } \dot{\gamma} > \dot{\gamma}_0 \end{cases} \quad (4.106)$$

where μ_0 is such that there is continuity at $\dot{\gamma} = \dot{\gamma}_0$. The material parameters K and n are known as the flow *consistency index* and the *behaviour index* respectively. One distinguishes between *shear thinning* fluids ($n < 1$) and *shear thickening* fluids ($n > 1$). The fluid becomes less viscous with increases of the shear rate in the former, while the opposite is true for the latter. Drilling muds are all of the shear thinning type [214].

We are particularly interested in the special case where $\tau_0 = 0$, that defines the family of *power-law fluids*, whose qualitative behaviour, both for shear-thinning and shear-thickening fluids, is shown in Fig. 4.18. It is quite common to have drilling muds described by this model. Moreover, it was required by the consultancy firm that the flow should be modelled as a power-law fluid.

The continuous problem associated to the fluid phase consists in the generalized Navier–Stokes equations, given by:

$$\left. \begin{aligned} \rho_f \frac{\partial \mathbf{u}}{\partial t} + \rho_f (\mathbf{u} \cdot \nabla) \mathbf{u} - \mu_{\text{eff}} \nabla^2 \mathbf{u} + \nabla p = \mathbf{f} \\ \nabla \cdot \mathbf{u} = 0 \end{aligned} \right\} \text{in } \Omega \times [0, T) \quad (4.107a)$$

$$(4.107b)$$

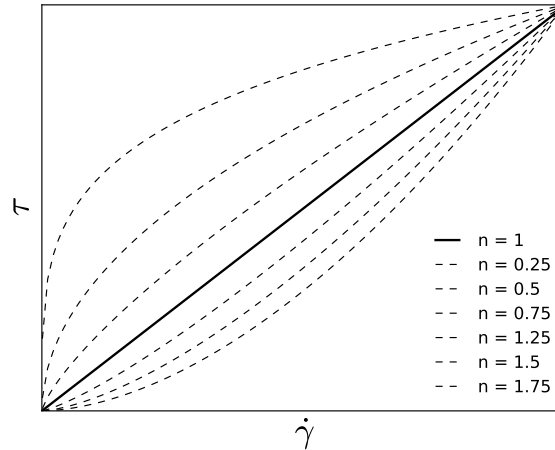


Fig. 4.18 Strain rate-stress relation for a power-law type fluid undergoing pure strain at a constant rate, for different flow behaviour indices ($n = 1$ corresponding to a Newtonian fluid).

where μ_{eff} is the effective viscosity, encoding the non-Newtonian behaviour of the fluid. In order to have a well posed problem, we need a set of initial and boundary conditions of Dirichlet and Neumann type as given in Eq. (4.15).

4.7.3 Hydrodynamic interactions for the particles

Here we again consider a hydrodynamic model of the form of the simplified Eqs.(4.1) and (4.4), where it is assumed that $\mathbf{F}_A, \mathbf{F}_H, \mathbf{F}_L, \equiv 0$. Regarding the viscous forces (\mathbf{F}_D), the particular expression that we have used is described in Section 4.7.3.

To the best of our knowledge, no empirical expressions exist for the history term in non-Newtonian fluids and finite Reynolds number. Given the uncertainties involved and its secondary importance for flows with reasonably high Stokes numbers, we have neglected it.

The importance of the lift force is often secondary as compared to the steady drag force, although in some cases it can become important [219]. Nonetheless, we have neglected its influence in the present work, although we acknowledge its potentially significant importance and further investigation of this issue, including the study of suitable formulations is left for future work.

Finally, no hydrodynamic torque has been considered for the rotational degrees of freedom. Here too, we have opted to keep the formulation simple and so we also leave its consideration, along with that of the lift force, for future work.

Non-Newtonian Drag

In order to obtain an appropriate equation of motion in the non-Newtonian setting, we make the assumption that the same additive superposition of the different forces discussed in Section 4.2 is still applicable in this setting. Other authors have made the same assumption [3]. Furthermore, since the non-Newtonian effects are captured by a varying viscosity, we will assume that the unperturbed flow and added mass forces, which are independent of the viscosity for Newtonian fluids remain unchanged, since all the non-Newtonian effects are concentrated in the varying viscosity. Having neglected the effect of the lift force, it is therefore only left to determine a steady drag law.

The literature on the hydrodynamic forces of a sphere submerged in a non-Newtonian fluid is certainly much more scarce than that for Newtonian fluids. A comprehensive review can be found in [16]. Nonetheless, here we employ the empirical expression proposed by Shah et al. [312] to predict the terminal velocity of particles in power-law fluids in the context of drilling operations. Their formulation has only been tested in stationary conditions but due to our hypothesis of the additive decomposition of the different effects, we will consider it adequate for our purposes. They provide the following expression for the empirical drag coefficient

$$C_{D,Shah} = (A^2 Re_p^{2B-2})^{\frac{1}{2-n}} \quad (4.108)$$

with the empirically determined parameters

$$\begin{aligned} A &= 6.9148n^2 - 24.838n + 22.642 \\ B &= -0.5067n^2 + 1.3234n - 0.1744 \end{aligned} \quad (4.109)$$

and where the particle Reynolds number is defined as

$$Re_{p,Shah} = 2^{n-1} \frac{\|\mathbf{w}\|^{2-n} d_p^n}{K/\rho_f} \quad (4.110)$$

The model above is valid for $n \in [0.281, 1]$ and $Re_{p,Shah} \in [0.001, 1000]$. Note that using the characteristic values calculated in Section 4.7.4 we can check whether the conditions for the validity of this model are met. The value of n is clearly well into the range of validity. In order to estimate the characteristic values of $Re_{p,Shah}$ we can use the estimates for the mean velocity and assume a conservative value for the relative velocity based on it, say 50 % of its value. For instance, in the inlet tube area, this estimate would yield $Re_p \approx 140$, while this number could reach $Re_{p,Shah} \approx 3000$ inside the nozzles. Again, it is expected that the maximum relative velocity occurs inside the nozzles and since the value of $Re_{p,Shah}$ is estimated to

be only slightly greater than the range of validity there, we will accept the associated error nonetheless. The error associated with this choice is unknown but we do not expect it to be too large, especially when taking into account the low degree of non-Newtonian behaviour (n relatively close to one) of the mud.

Final model for the hydrodynamic interactions

Summarizing, we will consider the following the equation of motion for an isolated particle, suspended in the drilling mud:

$$m_p \frac{d\mathbf{v}}{dt} = \xi_U (\mathbf{F}_U + \mathbf{F}_A) + \mathbf{F}_{D,Shah} + \mathbf{F}_B \quad (4.111)$$

where the binary parameter $\xi_U \in \{0, 1\}$ is introduced to easily turn off the pressure-related terms as required.

Summary of limitations of the model adopted

The models adopted for the description of the hydrodynamic interactions inevitably contain a number of assumptions and simplifications that contribute to the final error in the simulations. We next list the most important of these factors. Some of them can be addressed by further developments that will be discussed in Section 5.2.5.

- The drag model is used slightly out of its range of applicability inside the nozzles, which could impact the calculated impact velocity. Furthermore, this drag is based on empirical data looking at terminal velocity data, rather than complex dynamics.
- The lift force has been neglected altogether for simplicity. However, this is not well justified and further work on this issue is needed.
- Similarly, no hydrodynamic torque has been taken into account. Its influence is expected to be rather weak, though, especially when no Magnus effect has been considered (it only affects impacts).
- The additive property of the different effects in the equation of motion requires further testing. However, the same assumption has been made by many authors based on extrapolations and partial evidence in the Newtonian fluid setting.

4.7.4 Characteristic scales of the problem

In order to verify the range of applicability of the different models and the level of numerical resolution required for the simulations it is important to survey the different scales involved in the problem. Let us thus examine them

Characteristic scales of the flow

Let us start by considering some characteristic scales relevant to the continuous problem. From Tables 4.4 and 4.5, one can estimate the average velocities in the different sections of the geometry. Similarly, the fluid residence time (average time spent by the fluid molecules in the domain) can be calculated by dividing the volume over the flux. These and other derived quantities relevant to the fluid are included in Table 4.6.

A Reynolds number can be calculated for power-law fluids, following Metzner and Reed [247], as:

$$Re_{p,PL} = \frac{2^{3-n}}{(3 + 1/n)^n} \frac{U^{2-n} d_p^n}{K/\rho_f} \quad (4.112)$$

who also experimentally derived the following criterion for the transition to turbulent flow:

$$Re_{p,PL} > 2000 \quad (4.113)$$

Note that, according to the estimates discussed above the flow in the inner inlet tube is expected to be just on the verge of turbulence ($Re_{p,PL} \approx 4000$) and only moderately turbulent in the nozzles ($Re_{p,PL} \approx 20800$ assuming equal distribution of flow among the four nozzles). Note that this is the maximum Reynolds number expected in the flow, as this is the most constrained section of the conduct and

$$Re_{p,PL} \sim U^{2-n} d_p^n \sim (d_p^{-2})^{2-n} d_p^n \sim d_p^{3n-4} = d_p^{-1.717} \quad (4.114)$$

which is monotonically decreasing. We thus expect a mostly laminar or transitional flow regime, with some areas presenting weak turbulent or transitional regimes. Relatively moderate turbulence is expected inside the nozzles.

With regards to the near-boundary resolution, one can calculate an analogue to the y^+ distance by Trinh [346]

$$y_{PL}^+ = y_{PL}^+(y) = \frac{\tau_w^{1/n-1} \rho_f}{K^{1/n}} y \quad (4.115)$$

where τ_w is the shear stress at the wall, which can be estimated as

$$\tau_w = \frac{1}{2} f \rho_f U^2 \quad (4.116)$$

and where f is the friction factor $f = 16/Re_{p,PL}$ [247]. It is interesting to calculate the distance from the wall, h^+ , at which $y_{PL}^+ = 1$, since that is the size recommended for the smallest computational cells placed close to it [214]. These values are summarized in Table 4.6.

Table 4.6 Characteristic scales

Parameter	Value	Description
Flow		
U_{inlet}	12.68 m s ⁻¹	average velocity in inlet tube
$U_{nozzles}$	156.2 m s ⁻¹	average velocity in nozzles
t_{res}	0.38 s	total domain residence time
$t_{res, internal}$	0.053 s	internal domain residence time
h_{inlet}^+	2.8×10^{-4} m	recommended size of computational cell adjacent to the wall (inlet tube)
$h_{nozzles}^+$	1.8×10^{-5} m	recommended size computational cell adjacent to the wall (nozzles)
Particles parameters		
d_p	0.001 981 2 m	diameter of steel particles
ρ_p	7850 kg/m ³	density of steel particles
Mixed parameters		
α_p	0.021	solid volume fraction
$\alpha_{p,\rho}$	0.128	solid mass fraction
Operation conditions		
Q	0.029 02 m ³ /s	fluid flux
N_{inlet}	1.5×10^5 s ⁻¹	number of particles flux
Ω_{drill}	-2π s ⁻¹	angular velocity of drill-bit

Characteristic scales of the particles

Stokes number The Stokes number is defined as the quotient between the particle's relaxation time and the typical time scale of the background flow fluctuations, as in Eq. (2.4).

We have seen in Chapter 2 that when the assumption of Newtonian fluid and creeping flow is valid (very small Re_p), this relaxation time is a constant. This definition is still applicable in the context of power-law fluids, although its expression here is dependent on a variable viscosity, which complicates its analytic determination.

In Fig. 4.19 the relaxation times for the steel particles are shown as numerically calculated for different initial relative velocities. The range of values covers all the values of interest in the domain and however the relaxation time is seen to remain quite stable around 0.003 s. This time scale can thus be used as a reference relaxation time in what follows.

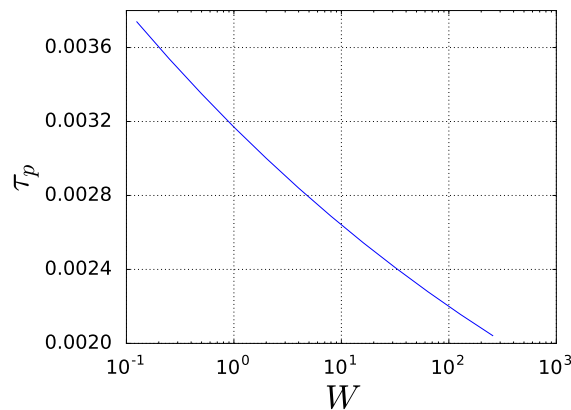


Fig. 4.19 Relaxation times for different initial relative velocities obtained numerically using the drag model of Shah, see Section 4.7.3.

The Stokes number can be used to distinguish particulate flow regimes, classifying them into ballistic ($St > 1$) and tracer-like ($St < 1$) with respect to the fluid motion associated with the time scale T . When St is much greater than one, it is normally assumed that the particles do not have time to respond to the fluid dynamics, see for example the discussion in [19]. In order to estimate the importance of an accurate description of the turbulent structures, it is therefore useful to look at the Stokes number.

For instance, we have seen that inside the nozzles we expect to find the most intense turbulence. But since the particles move at about 100 m s^{-1} there, they only spend less than $1 \times 10^{-3} \text{ s}$ in them. This means that $St > 3$ and most likely $St \gg 3$, as the typical turbulent fluctuations will be significantly smaller spacial amplitude than the length of the nozzles. This means that it is mostly the mean flow that will dictate the trajectory of the particles inside the nozzles. Note however that this is true *as long as the drag force is the/a dominating hydrodynamic force*. We will come back to this question in Section 4.7.8.

Characteristic scales related to the interactions

In order to assess the importance of the influence of the particle in the flow, the most important scale is the typical solid fraction, which is the global proportion of particles volume to fluid volume [85]; see also Section 4.8.4. Another interesting quantity is the solid mass fraction that takes into account the different densities of the two phases. These quantities can be preliminarily estimated assuming a homogeneous distribution of particles as

$$\begin{aligned}\alpha_p &= N_{\text{inlet}} \frac{V_p}{Q} \\ \alpha_{p,\rho} &= \frac{\rho_p}{\rho_f} \alpha_p\end{aligned}\tag{4.117}$$

where $V_p = \frac{1}{6}\pi d_p^3$ is the volume of one steel ball. The characteristic values for the present problem can be found in Table 4.6.

4.7.5 Discussion of the coupling strategy

Forward FSI interactions

The coupling between the fluid phase and the solid phase is given by the imposition of the no-slip condition at the walls. In this case, we have opted for an arbitrary Lagrangian-Eulerian (ALE) formulation [106] with a moving mesh that moves jointly with the drill bit. In this case the mesh velocity is known exactly and coincides with the rotating velocity of the bit.

Forward Hydrodynamic interactions

The fluid velocity field values are determined at the center of each particle using the linear tetrahedral shape functions of the FEM mesh. The derivative recovery tool of choice was the standard method. We chose this method based on two considerations:

1. The large size of the systems to be studied, which meant that it became unmanageable to use the PPR method for the available memory.
2. The existence of very narrow sections, for which it was feared that the larger size of the element patches needed for the recovery lead to inaccuracies in these thin regions, for which the number of elements is limited.

Backward FSI interactions

Neglecting the backward action of the fluid on the structure, the backward fluid-solid interaction (FSI), is reasonable, based on the near-homogeneity of the mass distribution inside the drill-bit and the stiffness of the structure at the pipe-length scales of study.

Inter-particle hydrodynamic interactions

The inter-particle interactions are associated with the fluid field disturbances caused by the presence of nearby particles. Such disturbances are responsible for phenomena such as drag reduction due to the wake effect, which can become important at short distances. Such effects are difficult to take into account due to their dependence on the details of the particles spatial configuration, and only approximated relations such as empirically modified drag coefficients are available at particle Reynolds numbers greater than one; see, e.g. [31]. To our knowledge, there is no expression available in the literature for non-Newtonian fluids, and studying the possibility of adapting formulations derived in the Newtonian setting is left for future work.

The error introduced by this effect is negligible when the inter-particle distance grows substantially above the diameter. In [216] five is suggested as the number of diameters, which leads to a maximum solid volume fraction of $\alpha_p \approx 0.004$. Since we are about five times above this threshold we should expect a significant error, especially wherever particle agglomeration or clustering takes place, although it is not easy to quantify. Qualitatively, the expected results would be overall increased ejection velocities due to the lack of a wake effect.

Backward coupling with the particles

We distinguish between two types of effects of the particles on the fluid phase. The first is associated to the momentum transfer between phases, while the second is related to the conservation of mass. Both effects can be taken into account within the point-particle approach, although we have left their inclusion to future work. Next we discuss each of the effects separately.

Force coupling The hydrodynamics forces that the fluid exerts on the particle are, by Newton's Third Law, the forces that the particles exert on the fluid but with the opposite sign. The importance of these forces on the macroscopic flow (which then in turn is used to move the particles) is negligible if the concentration of particles is small enough. In [216] the criterion $\alpha_{p,\rho} \ll 1$ for negligible influence is suggested. At $\alpha_{p,\rho} \sim 0.1$, see Table 4.6, it is possible that neglecting this effect can introduce a significant error, but we do not expect it to be critical, since unless there are very large inhomogeneities in the flow (especially in

the transverse direction), the streamlines should not be greatly affected. This means that although the pressure drop might not be well captured, the fluid velocity will, leading to a fairly reasonable prediction of the particles' movement.

Mass conservation effects The mass conservation equation Eq. (4.14b) does not take into account the volume displaced by the particles. This simplification would not be expected to lead to very large errors if the expected average value of $\alpha_p \sim 0.02$, see Table 4.6, remained reasonably uniform (the flow through a regular lattice of particles with the same volume fraction would essentially just be sped up by 2%). However, some degree of nonuniform particle agglomeration in some regions of the domain is expected. In these regions, the error could become important. The region inside the nozzles is particularly sensitive to this effect due to their thinness. We will come back to this issue in Section 4.7.8; see also Section 5.2.5. Note however that, at least to some extent, the error introduced by this simplification cancels out with the error described in Section 4.7.5, since this simplification leads to a underestimation of the real velocity (wherever there is particles accumulation), while the drag force is overestimated due to neglecting the wake effect, as we have seen.

4.7.6 Final algorithm

Given that the DEM phase has a much stricter time step requirements than the fluid, due to the very small time scales associated to the contact dynamics that it resolves, we employ a sub-stepping scheme so that, for every fluid time step, many DEM time steps are performed. The fluid is advanced first, and then the DEM catches up in smaller time step increments. The fluid field quantities are evaluated at every DEM step too taking a weighted average between old and new fluid values. The pseudo-code is shown in Algorithm 5, where $N_{f,steps}$ refers to the total number of fluid time steps in the simulation and where t_{DEM} keeps track of the time for the DEM phase.

4.7.7 Time and space resolution

In this section we present the level of discretization that we use for the description of the fluid and the particles.

Discretization of the particles phase: time step selection

The determination of the time step for the DEM must ensure that the maximum indentations that occur during the simulation avoid unrealistic particle positions, such as particles squeezing

Algorithm 5 One-way coupled two-phase algorithm.

```

1:  $t \leftarrow 0.0$ 
2:  $t_{\text{DEM}} \leftarrow 0.0$ 
3:  $m \leftarrow 0$ 
4:  $\mathbf{U}^0 \leftarrow \mathbf{U}_{\text{init}}$ 
5: INITIALIZEPARTICLES()
6: SEARCHNEIGHBOURS()
7: for  $n$  in  $[0, N_{f,\text{steps}}]$  do
8:   ROTATEMESH( $t, n$ ) ▷ rotate the fluid mesh with imposed angular velocity
9:   SOLVEFLUID( $t, n$ ) ▷ Algorithm 6
10:  RECOVERDERIVATIVES() ▷ Section 4.4
11:  while  $t_{\text{DEM}} < t$  do
12:    SOLVEDEM( $t_{\text{DEM}}, m$ ) ▷ Algorithm 9

```

Algorithm 6 SolveFluid function algorithm. These operations are performed at every fluid time step.

```

1:  $t \leftarrow t + \Delta t$ 
2:  $n \leftarrow n + 1$ 
3:  $i \leftarrow 0$ 
4:  $\mathbf{U}_0^{n+1} \leftarrow \mathbf{U}^{n+1}$ 
5: while  $\|\mathbf{R}_{i+1}^{n+1}\| > \text{tolerance}$  do
6:    $i \leftarrow i + 1$ 
7:    $\delta\mathbf{U}, \delta\mathbf{P} \leftarrow \text{ASSEMBLESYSTEMANDSOLVE}()$  ▷ Eq. (4.63)
8:    $\mathbf{U}_{i+1}^{n+1}, \mathbf{P}_{i+1}^{n+1}, \mathbf{R}_{i+1}^{n+1} \leftarrow \text{UPDATEUNKNOWN}(\delta\mathbf{U}, \delta\mathbf{P})$  ▷ Eq. (4.61)

```

through too-narrow passages or even traversing solid boundaries. Furthermore, the numerical stability of the explicit time-integration scheme must be ensured. Finally, the contact duration must be resolved with enough time steps to avoid excessive inaccuracies, taking into account the large uncertainties associated with particle rebounds. To avoid risky iterations (each simulation requires a significant amount of time) we used the methodology explained in Section A.2. The values involved in this process for the two basic cases are summarized in Table 4.7. Note that they are not the same for the total case (which includes the impacts of the particles on the bed) and the internal case, which only presents very skewed impacts as it does not include the bed area.

Discretization of the fluid phase: mesh selection

In order to estimate the mesh sizes required for the simulations we run a mesh convergence study based on a single nozzle (see Figs. 4.20 and 4.22 and section 4.7.7). The nozzle was attached to an idealized cylindrical structure to which an inlet velocity boundary condition

Table 4.7 Time step choices

Parameter	Internal flow	Total flow
number of subdivisions of shortest contact	20	20
security factor	0.325	0.325
maximum expected normal impact velocity	80 m s ⁻¹	200 m s ⁻¹
δt	5.0×10^{-7} s	1.9×10^{-7} s

was imposed so as to match the expected flux. The mesh selected for the internal flow case is shown in Fig. 4.23, while the corresponding mesh for the total flow case is shown in Fig. 4.25.

Discretization of the fluid phase: Time step selection

The criterion for the time step selection responds to a balance between the need for an affordable computation and the accuracy required. In the context of turbulent flows, the CFL (Courant–Friedrichs–Lewy) number is used as a criterion to fix a higher limit to the time step in connection with the spatial resolution used. The CFL number is calculated as

$$\text{CFL} = U \frac{\Delta t}{h} \quad (4.118)$$

where U is the magnitude of the velocity (relative to the mesh velocity). Explicit numerical schemes require that this condition is fulfilled to have numerical stability, but this case does not concern us, since our time scheme is theoretically unconditionally stable. Nonetheless, in order to have an accurate representation of the flow, it is recommended that $\text{CFL} < 1$ anyway. In effect, when using a larger than one, it is impossible for the numerical scheme to resolve the dynamics of the smallest scales, which become filtered out.

The simulation of the PID system involves regions with very different characteristic velocities. Our criterion in choosing the time step has been to use values of the CFL smaller than one everywhere except inside the nozzles, where the CFL has been allowed to reach numbers in the range 20 to 80. We thus expect to capture at most the larger-period instabilities inside the nozzles, while filtering out the fine details of the flow. This need not be a critical inaccuracy, since the particle are expected to respond sluggishly to these instabilities, see Section 4.7.4.

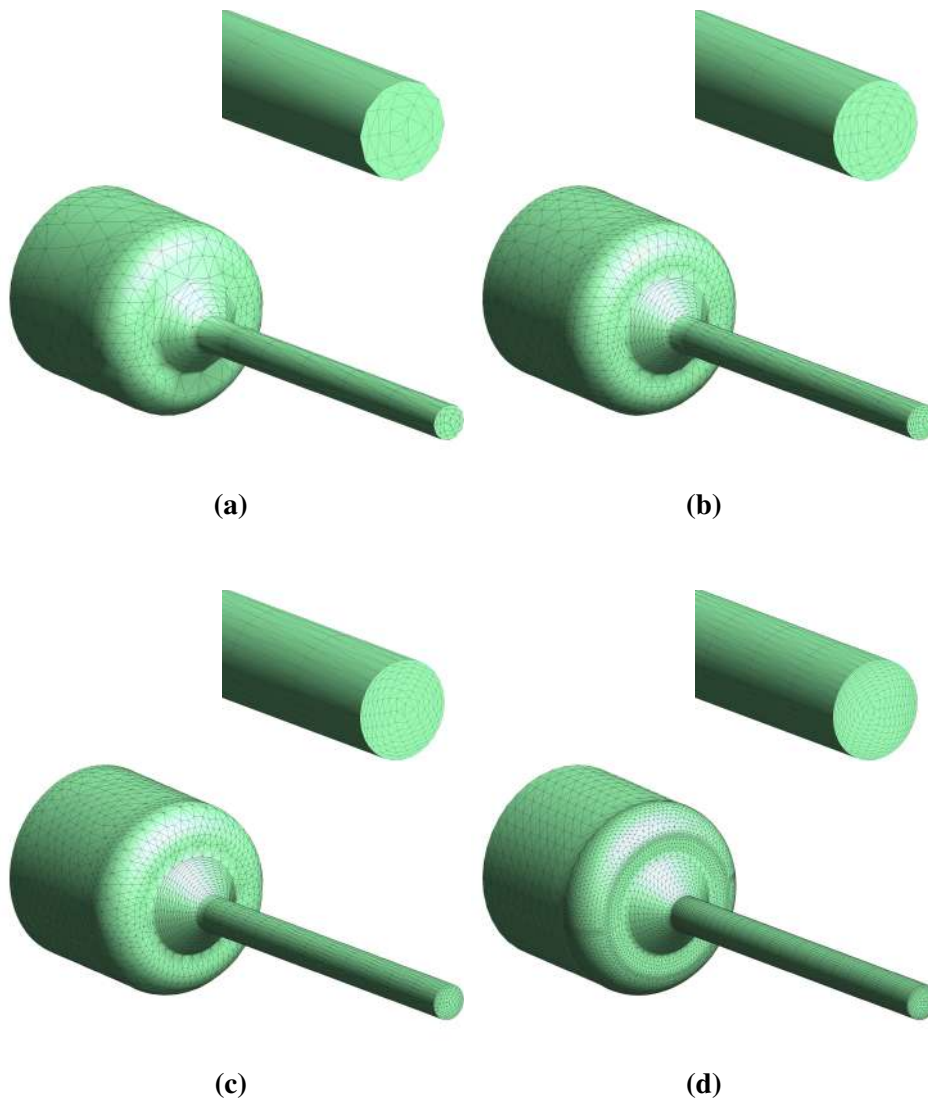


Fig. 4.20 Meshes used in the mesh-convergence study. From coarsest to finest, approximate number of elements: 5×10^3 , 1.6×10^4 , 4.8×10^4 , 2×10^5 . The inserts show a zoom into the tip of the nozzles.

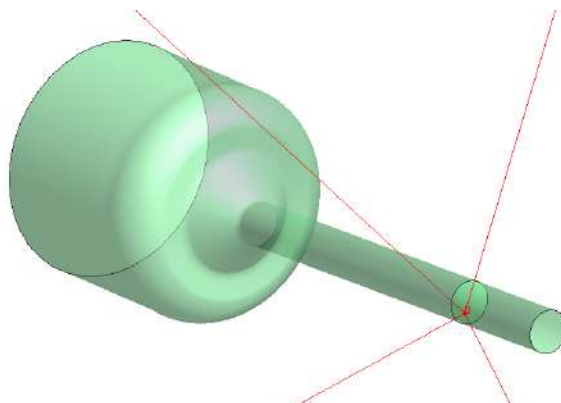


Fig. 4.21 Point of measure of the velocities.

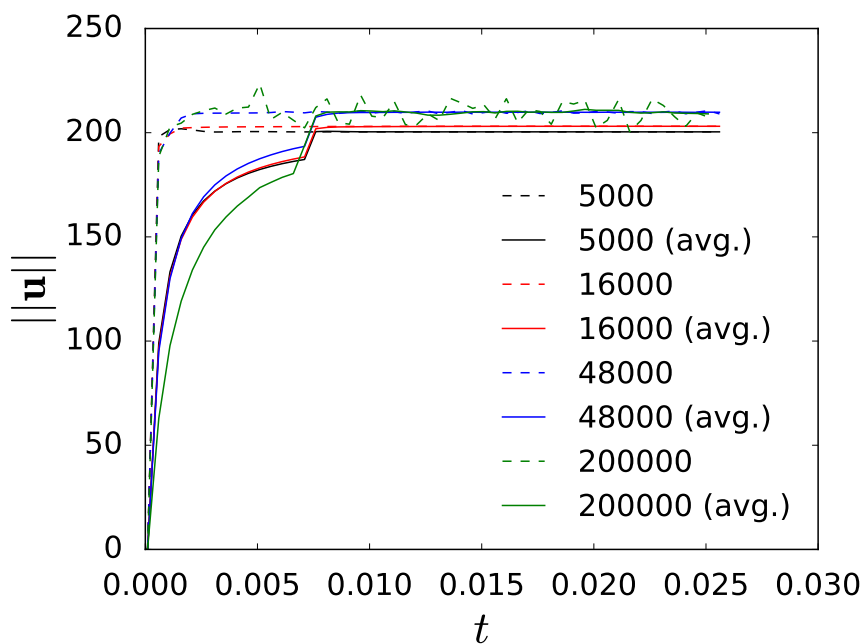


Fig. 4.22 Mesh convergence study. Velocity evolution of a point at mid-shaft, mid-radius (see Fig. 4.21) for the different meshes, identified by the number of nodes. Instantaneous values (dashed line) and moving averages (full line) over a duration of 20 time steps.

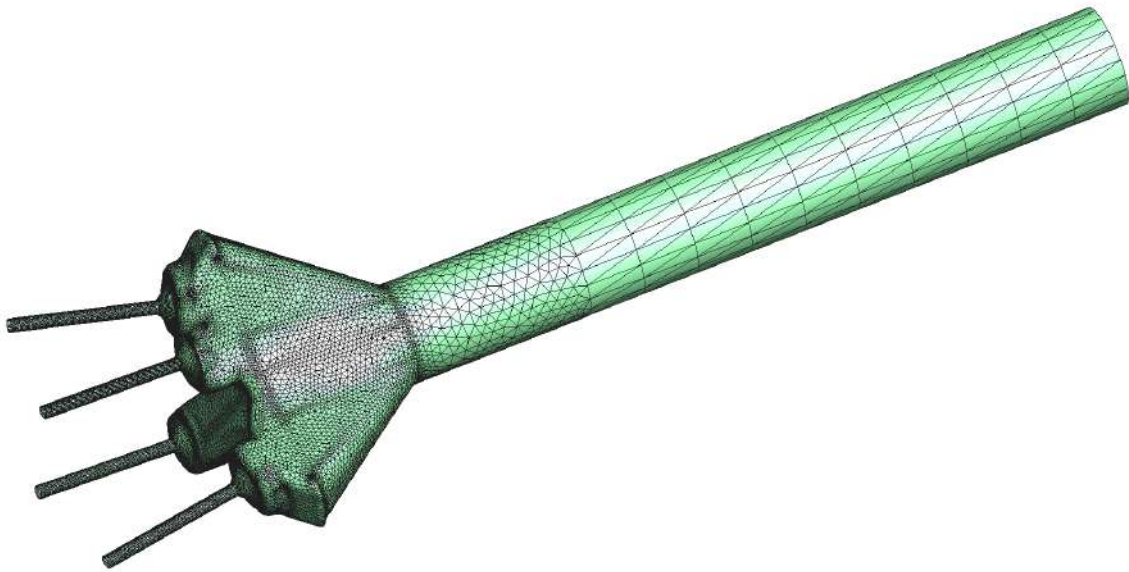


Fig. 4.23 Mesh used for the internal flow problem (814 644 tetrahedra).

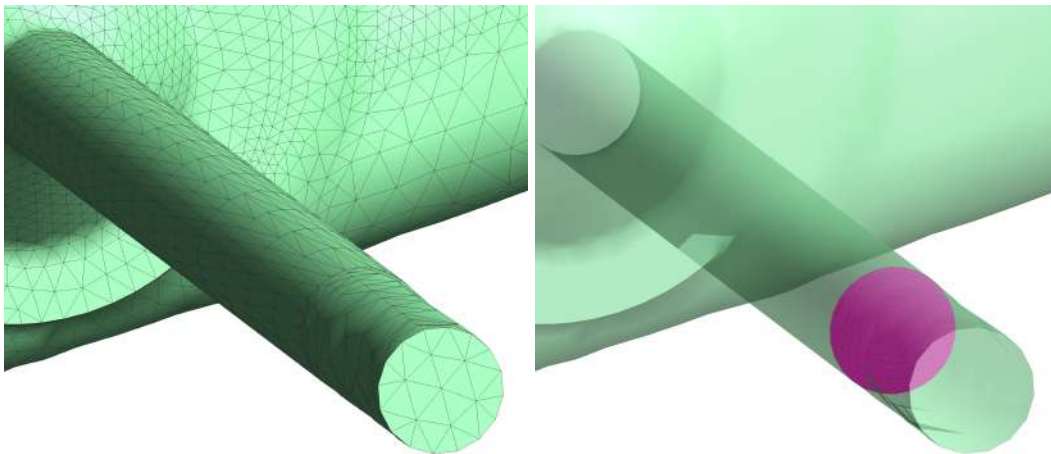


Fig. 4.24 Detail of mesh finally used for the nozzles of the internal case. Note the coarse-mesh portion that was added to the tip of the nozzles in order to avoid back-flow instabilities. The particles were eliminating before entering this area, where the flow resolution is poor.

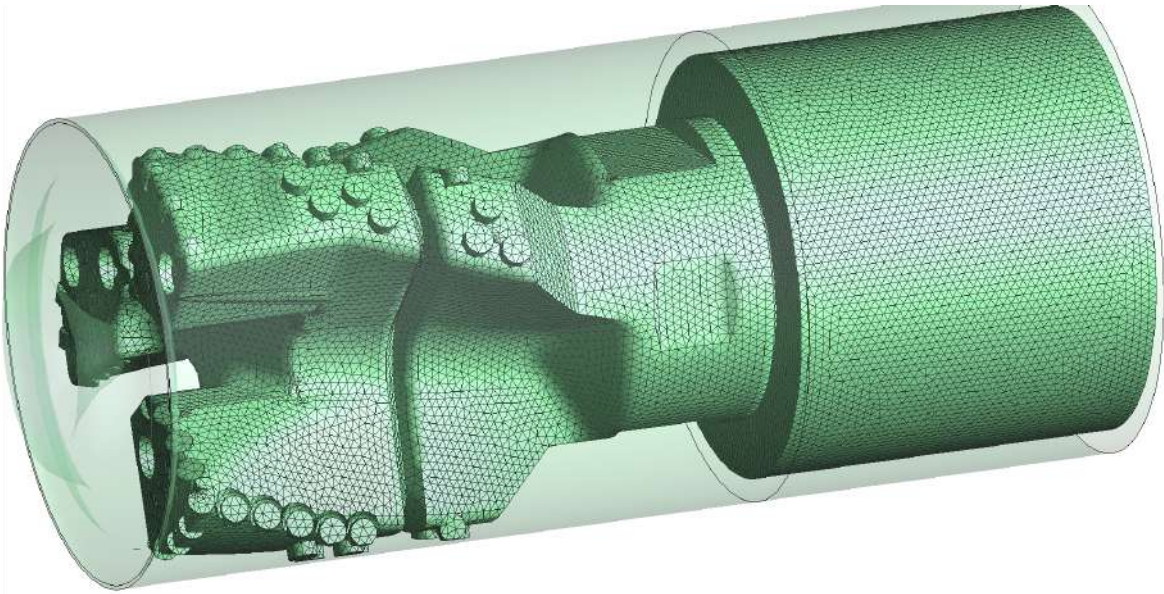


Fig. 4.25 Mesh used for the total flow problem (2 022 811 tetrahedra).

In the previous sections we have presented all the elements used to set up and perform the numerical simulations. In this section we describe the simulations that were finally performed and discuss the simulation results. We have considered two generic scenarios for the simulations: an *internal flow* scenario where we only look at the domain formed by the inner tubing from the inlet down to the tip of the nozzles; and an *total flow* scenario, where we simulate the whole domain. The internal flow simulations allow us to consider the flow with a greater level of detail at reasonable costs.

4.7.8 Simulations

In this subsection we present a selection of the simulation results. We start by listing the most relevant simulation settings.

Standard Settings

Parameters For the simulations we fix a number of parameters and options that we summarize in Table 4.8. These settings correspond to the *standard case* (SC) and any deviations from them will be explicitly remarked in their context.

⁷This number is approximated and is used by the mesh generator (advancing front) in [80] to build the mesh.

Table 4.8 Input parameters for the standard case run

Parameter	Value		Description
Coupling Parameters			
ξ_U	1		Include F_U and F_A (0 or 1)
Contact parameters			
$\mu_{p,w}$	0.42		particle-wall friction coefficient
COR	0.6		coefficient of normal restitution
ξ_C	1		compute inter-particle contact (0 or 1)
Numerical Parameters			
	internal	total	
h	0.0026 m	0.004 m	max. element size of irregular mesh ⁷
$h_{\text{wall}}/h_{\text{nozzles}}^+$	1.1	27	normalized width of elements adjacent to nozzle walls
Δt	1.0×10^{-4} s	1.0×10^{-4} s	time step for the fluid-phase
δt	5×10^{-7} s	1.9×10^{-7} s	time step for the particles-phase

Boundary conditions The boundary conditions imposed in all cases are a combination of strongly-imposed velocity conditions (uniform inlet velocity at the inlets and no-slip at the walls) and weakly imposed Neumann condition at the outlets (zero normal traction).

Internal flow results

We have run the standard internal flow simulations for a simulated time of 0.3 s. This corresponds to more than five times the residence time of this part of the domain. Fig. 4.26 shows a plot of the evolution of the number of particles in time. Notice that, although the number increases at ever lower rates and one would expect it to not grow much above 13 000, this is already almost twice the expected concentration for a uniform distribution of particles at their input rate, as obtained by multiplying α_p by the volume of the domain.

Fig. 4.27 shows a snapshot of the particles phase with different vectorial results represented. Note the ≈ 50 -fold ratio between the largest drag forces and the largest hydrodynamic force, due to the contribution of F_U and F_A around the entrance to the nozzles.

Fig. 4.28 shows the velocity modulus contour maps on several cross-sections. The saturation value of the velocity is varied to highlight either the nozzles or the rest of the domain. Note the regularity of the flow in the inlet tube, where the Reynolds number is on

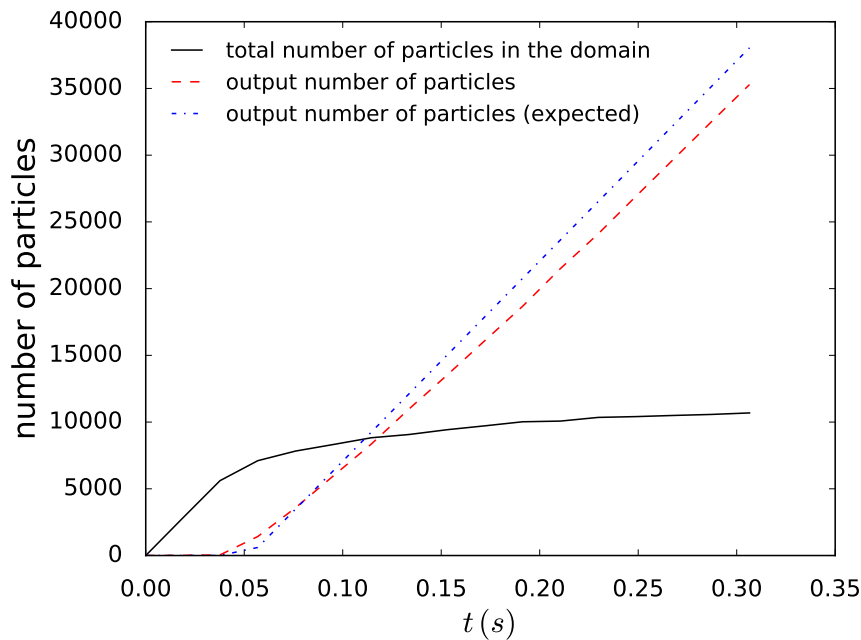


Fig. 4.26 Evolution of the total number of particles in the domain, in time. The expected number of particles leaving the domain through the outlet is also depicted (red dashed line), as well as the expected steady-state expected output estimated using the residence time of the fluid particles (blue, dotted-dashed line). As long as there is a difference in the slopes of the latter curves, there will be accumulation/depletion.

the verge of turbulence. Note also the rotation-triggered vortices inside the three branches of the distribution chamber both in the horizontal and vertical directions.

Fig. 4.29 shows two sets of streamlines at a particular time step. Fig. 4.29a shows a uniformly distributed selection of streamlines that passing through equally spaced points along the inlet tube cross-section. Fig. 4.29b shows a detail of streamlines passing through a segment of points inside a bisecting plane of one of the three branches of the distribution chamber. This figure highlights the convoluted direction of the flow in this recirculation zone.

Similarly, Fig. 4.30 shows a number of particles trajectories in the interval 0.25 s to 0.3 s. Fig. 4.30b shows a detail where the particle temporary trapping in the recirculation zone is highlighted.

Total flow results

The SC runs corresponding to the total flow have been run for a simulated time of 0.38 s, the concentration time of the domain. In this case the mesh is slightly coarser than in the internal case for the inner tube and the focus is more on the annulus flow.

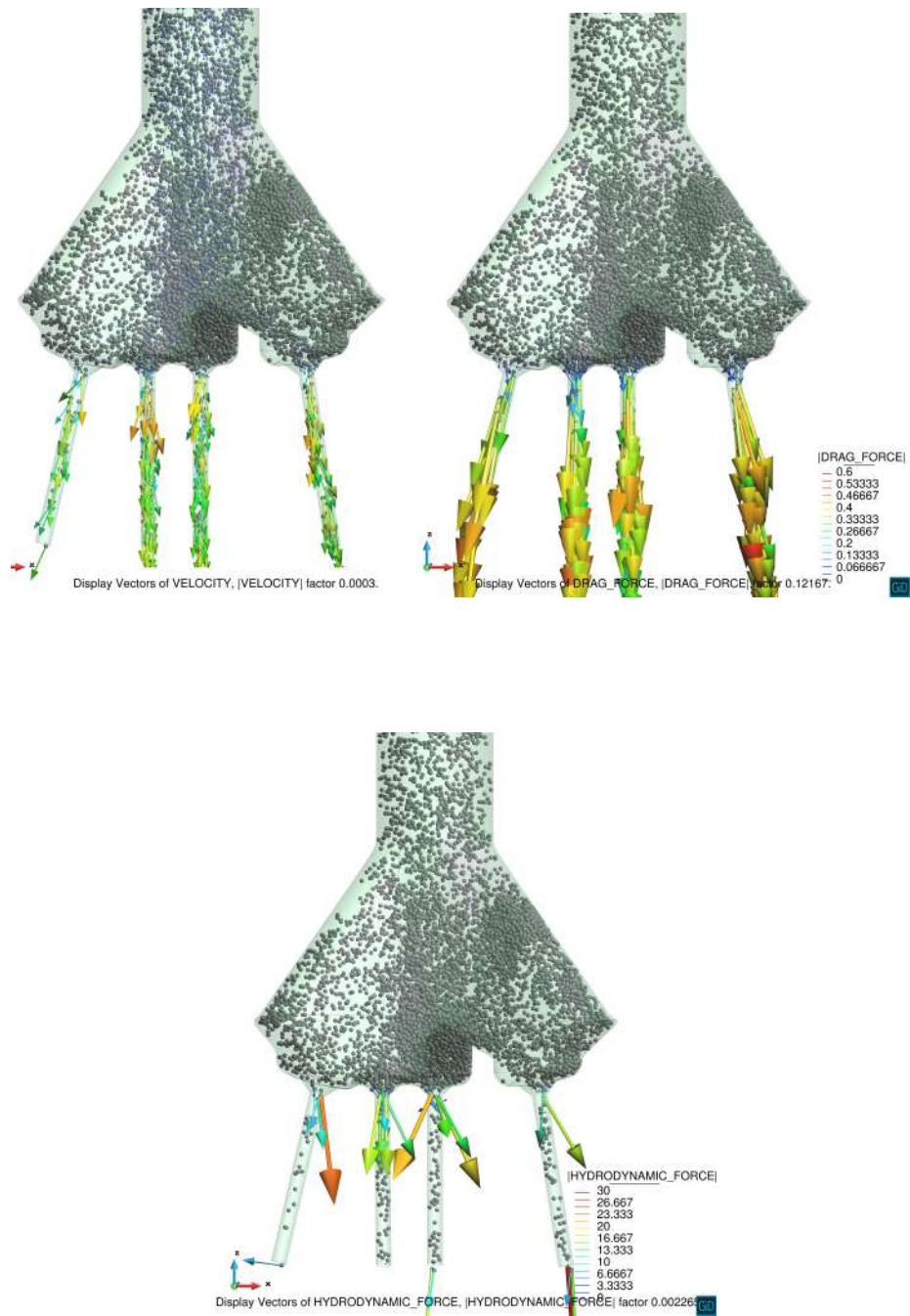
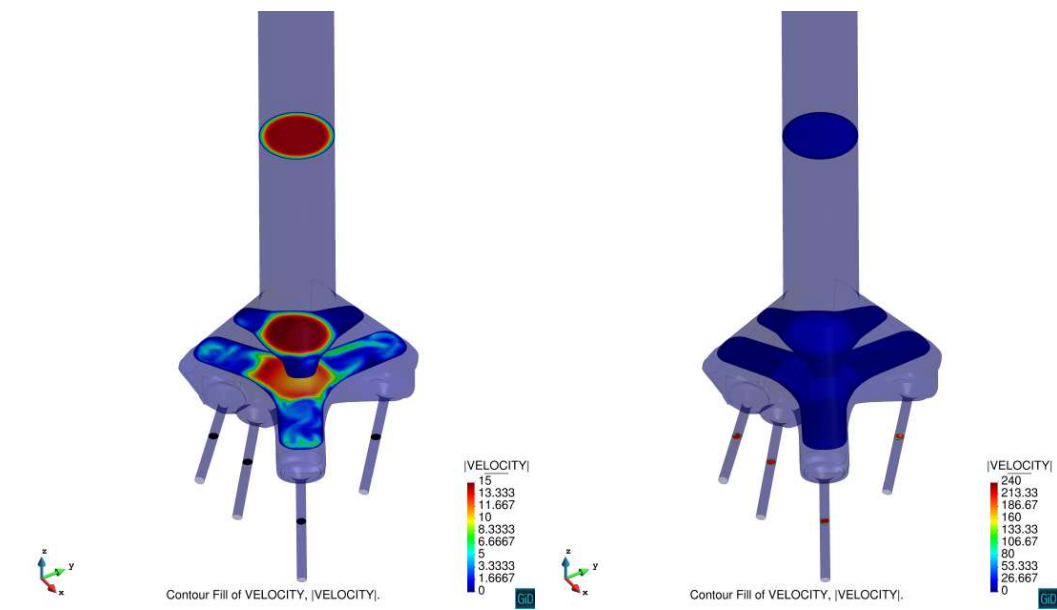
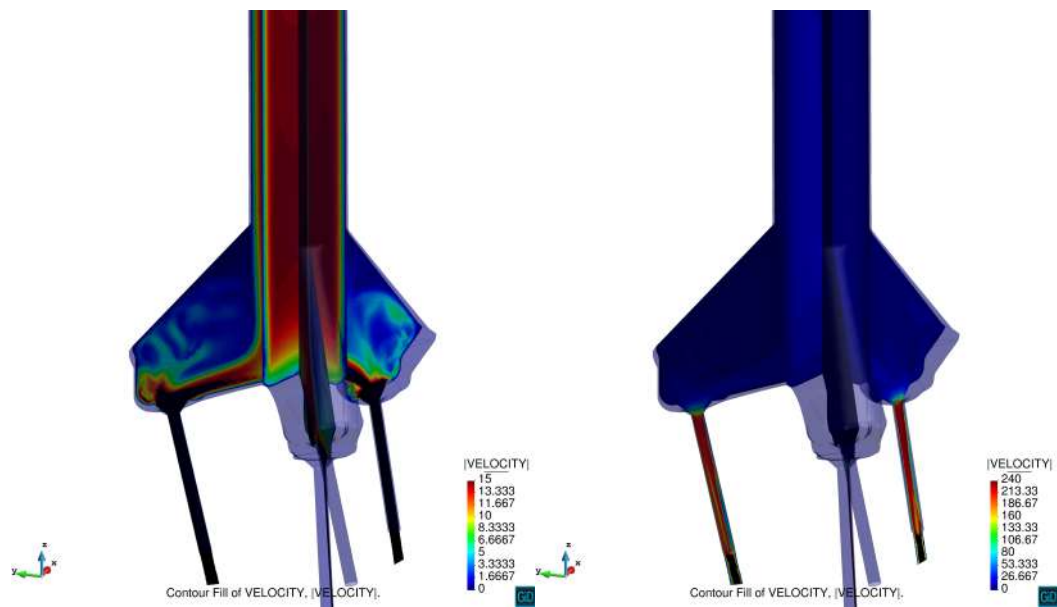


Fig. 4.27 Particles flowing under the action of the flow at $t = 0.25$ s. The particles velocities, drag force and total hydrodynamic forces are shown.



transversal cuts



vertical cuts

Fig. 4.28 Modulus of the velocity field at $t = 0.25$ s.

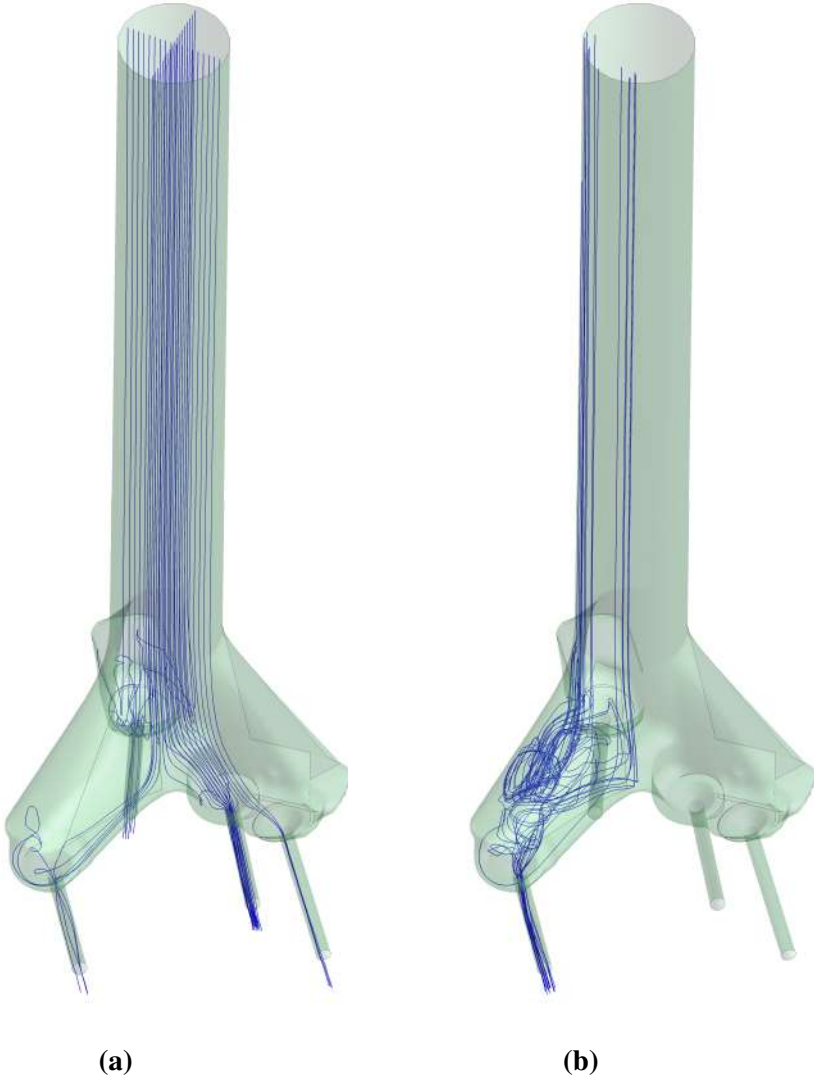


Fig. 4.29 Fluid streamlines at $t = 0.25$ s. Uniformly distributed streamlines (left) and detail of streamlines inside a vortical region (right).

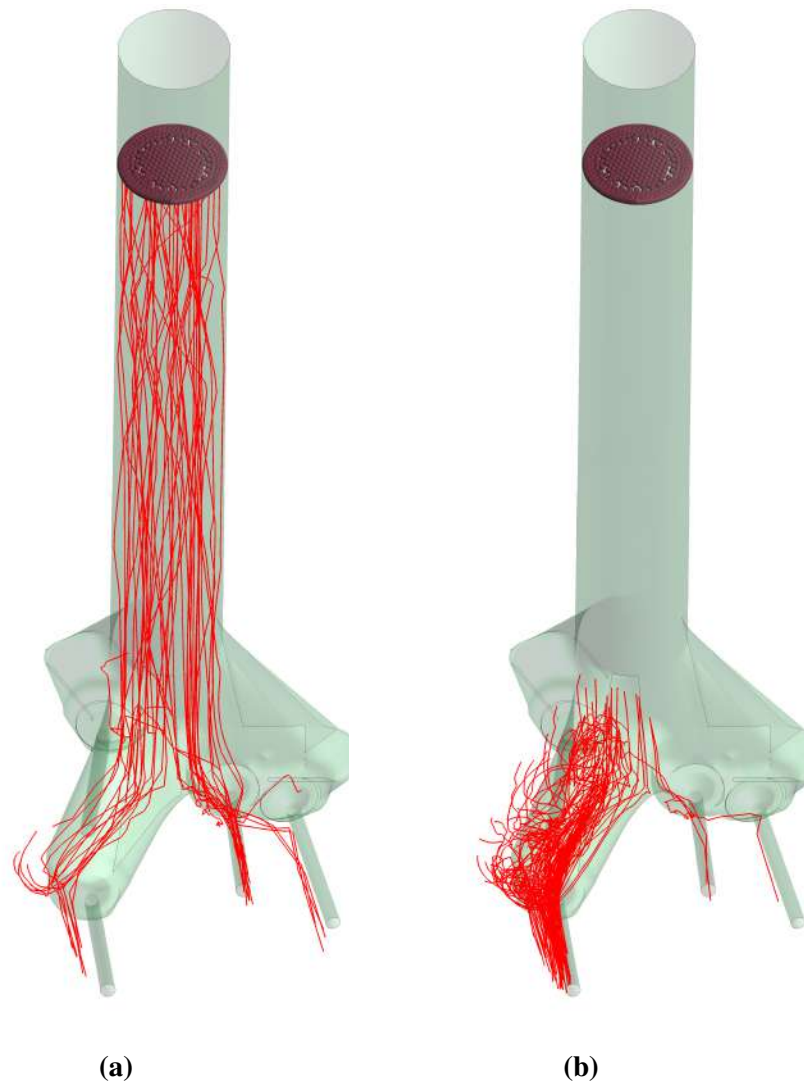


Fig. 4.30 Particle trajectories for the interval 0.25 s to 0.3 s. The reddish spheres represent the points of introduction of particles. Randomly chosen spheres (left), detail of spheres inside a vortical region (right).

Fig. 4.31 shows a sequence of three snapshots where the z -component of the velocity is indicated with color on the particles surface. Note that significant inhomogeneities exist in the solid concentration, which indicates the need for further research to study this effect. Note however that the particle size has been exaggerated by a factor 2.5 to facilitate visualization.

Fig. 4.32 shows a set of contour plots for a series of transversal and vertical cuts. As in Fig. 4.28, two sets of saturation velocities are used to highlight different regions of the flow; see also Fig. 4.33 for the corresponding streamlines. Note that the velocities in the internal flow are above the maximum of 8 m s^{-1} in large parts of the domain in (Fig. 4.32a, Fig. 4.32c).

In Fig. 4.34 a sequence of contour plots of the level of wear on the rock bed is shown, demonstrating the potential for this approach to help in assessing the performance of changes in the design or the operation conditions. The action of the individual jets is clearly visible.

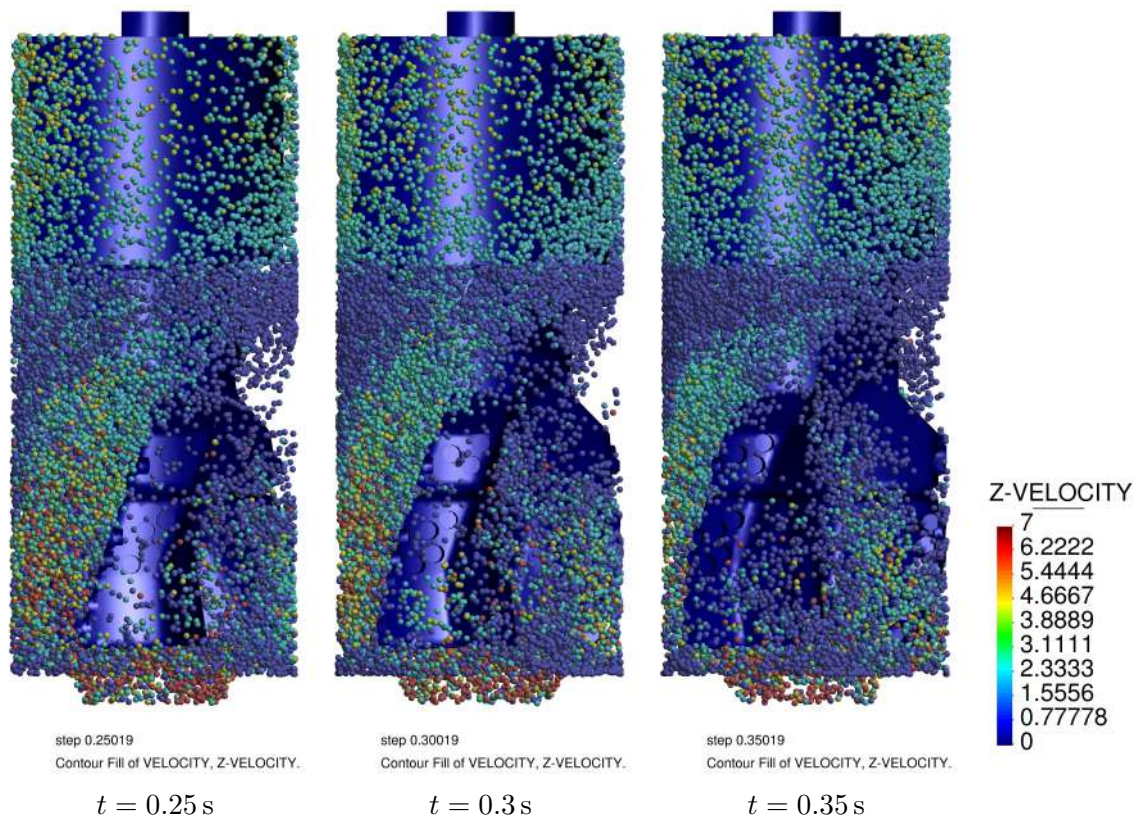
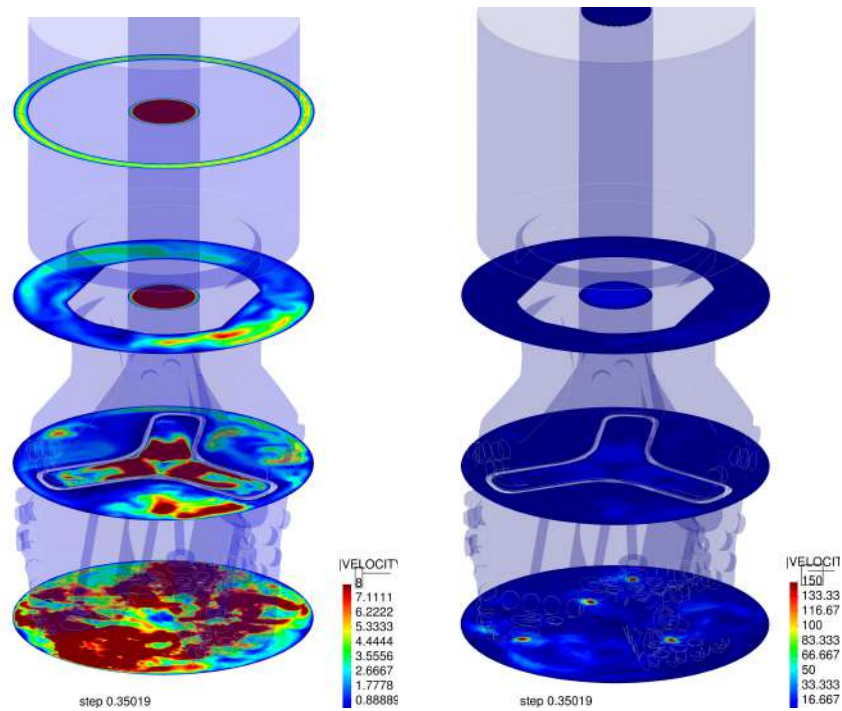
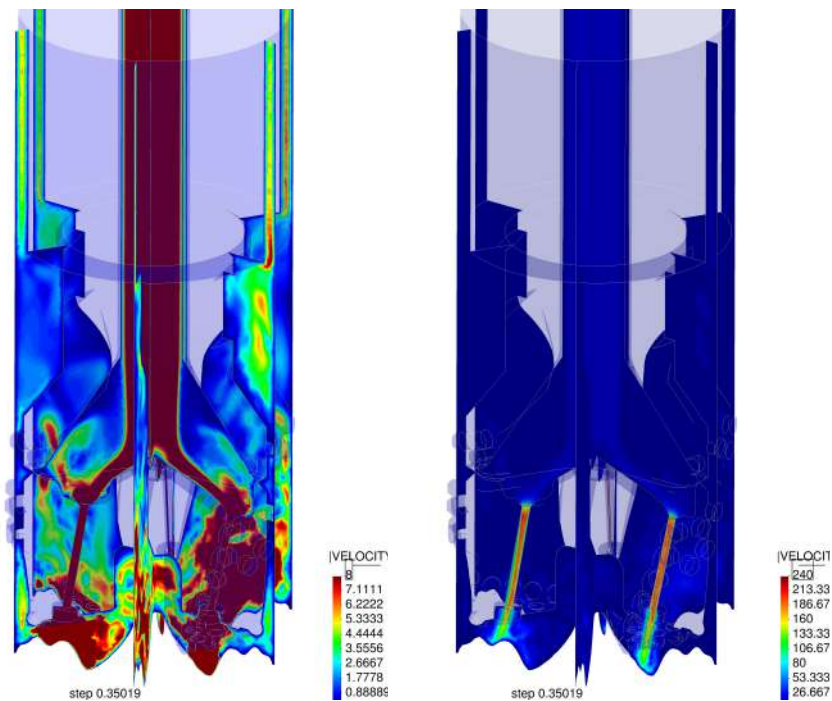


Fig. 4.31 Sequence of snapshots of the particles flowing under the action of the flow at different times. The colors indicate the z -component of the velocity. The size of the particles has been increased by 250% to facilitate visualization.



transversal cuts



vertical cuts

Fig. 4.32 Modulus of the velocity field at $t = 0.35$ s.

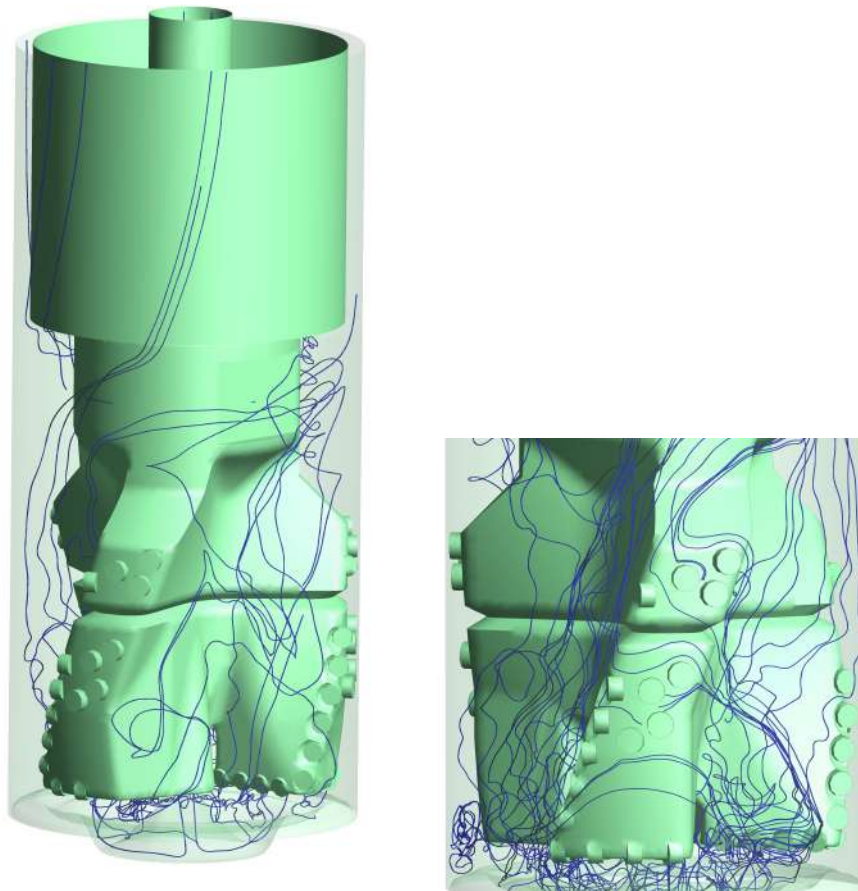


Fig. 4.33 Fluid streamlines at $t = 0.25$ s. Uniformly distributed streamlines (left) and detail of streamlines around the teeth (right).

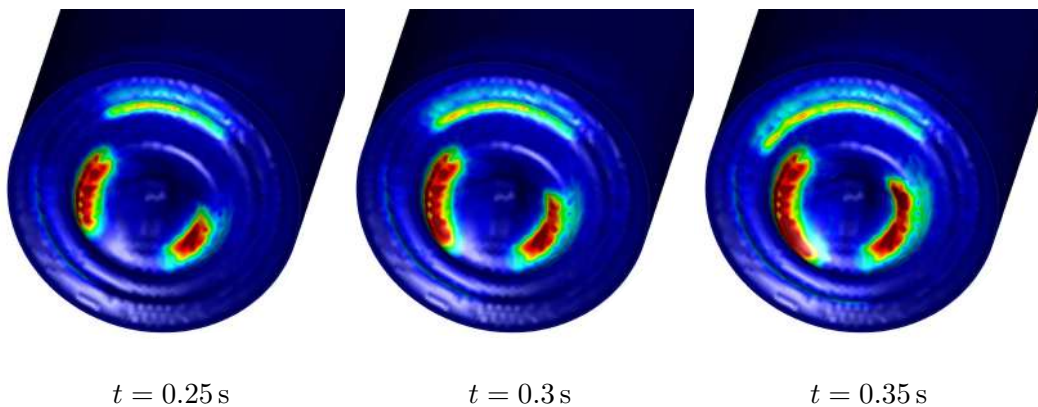


Fig. 4.34 Sequence of snapshots of the wear spread pattern evolution on the bed surface (red: intense wear; blue: light wear).

Sensitivity Analyses

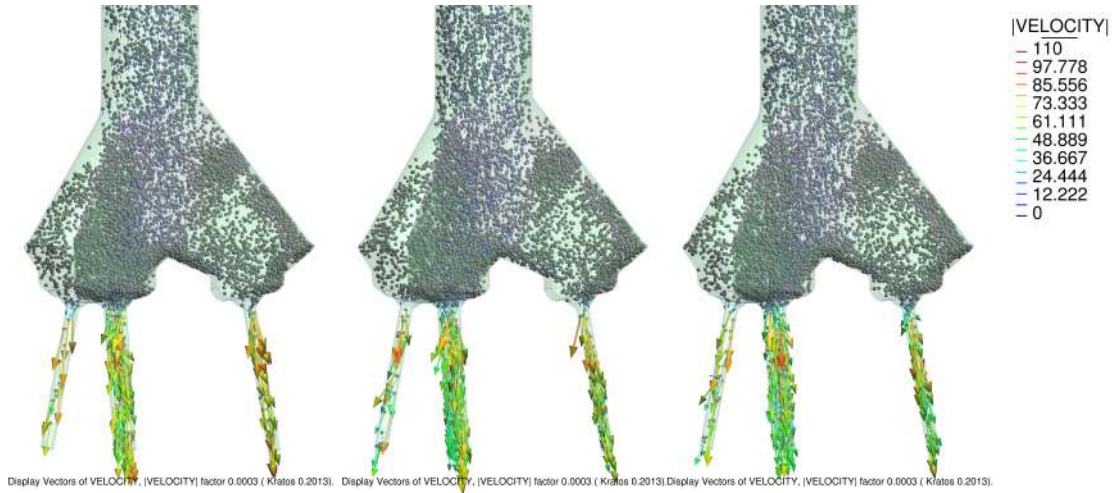
Given the uncertainty in the model parameters, it is interesting to investigate the influence of a few of them separately. This should help in concentrating the research efforts toward the most critical effects.

Coefficient of friction Let us investigate the effect of the most important parameters involved in the contact dynamics of the particles. Fig. 4.35 compares three snapshots corresponding to three different values of the coefficient of friction. The SC is shown in Fig. 4.35c, with a value of $\mu_{p,w} = 0.42$, given as representative of a typical of steel-on-steel contact by the consultancy firm. The other two figures show a snapshot of the same time step for $\mu_{p,w} = 0.2$ and $\mu_{p,w} = 0.6$ respectively, covering a wide range of values for what could be reasonably expected.

As seen in the figures, the effect of this parameter is very weak, and no clear trend can be observed. This can be explained in terms of the fact that the prevalent regime is on impact regime, where the particles are mostly bouncing off each other and off the surfaces, rather than rubbing against each other. It is known that the friction coefficient has a strong incidence on the angle of rebound in individual impacts [342]. However, this does not seem to bear any strong effect on the dynamics of the particles as a whole, probably due to the randomized character of the collective dynamics. Moreover, note that the accumulated number of particles is roughly equal in all cases; and idem with the wear pattern (see Fig. 4.36) indicating that the coefficient of friction plays little to no role in this problem.

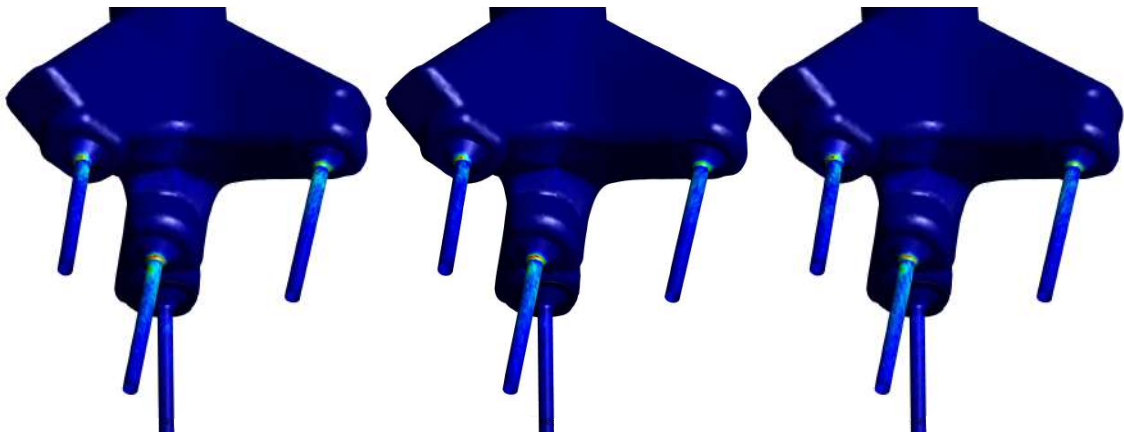
Coefficient of normal restitution On the other hand, Fig. 4.37 shows the effect of variations in the coefficient of normal restitution on the dynamics of the particles. Again, snapshots taken at the same time steps are compared between simulations that only difference being the COR. The central figure corresponds to the SC, with $COR = 0.6$, while Fig. 4.37a and Fig. 4.37c correspond to $COR = 0.2$ (highly dissipative) and $COR = 0.8$ (highly elastic). Clearly, the more dissipative case flows much more easily as the particles concentrate near the path of the dominating streamlines into the nozzles. When the COR is raised, the particles tend to occupy more space and accumulate more, as the number of particles in the domain goes from 8759 (close to the value of 7850 calculated by multiplying the input α_p by the domain volume) for the most dissipative case, to 12 469 for the mildly dissipative case. This represents a 42 % increase after only a quarter of a turn of the drill bit. Given the importance of the COR, therefore, and since (despite common practice) it is known that the COR is in fact not independent of the velocity [210], we must conclude that more work is needed to

characterize the contact more accurately, especially given the large range of characteristic velocities for different regions of the domain.



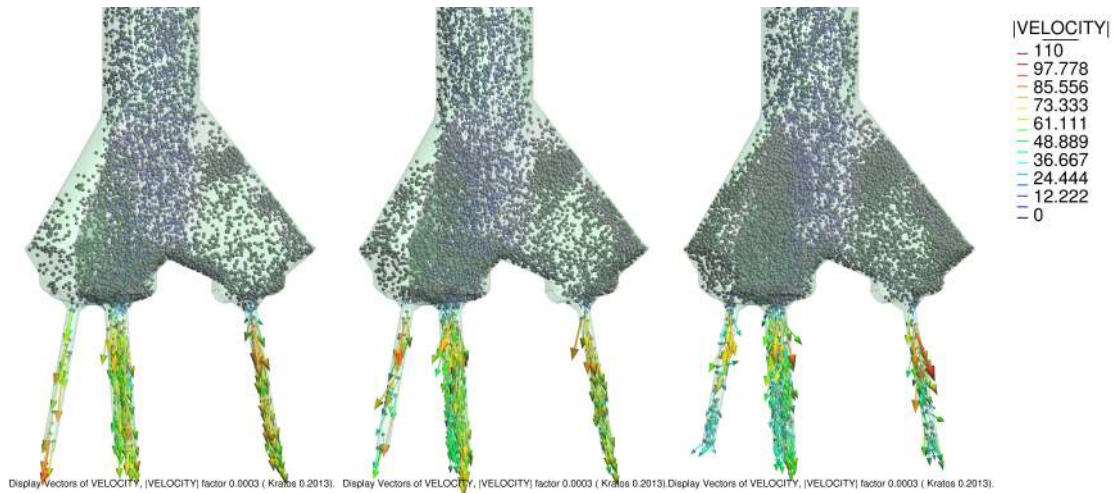
Rough contact ($\mu_{p,w} = 0.6$); 10 229 spheres in the domain at $t = 0.20$ s *Standard case* ($\mu_{p,w} = 0.42$); 10 040 spheres in the domain at $t = 0.20$ s *Lubricated contact* ($\mu_{p,w} = 0.2$); There are 10 304 spheres in the domain at $t = 0.20$ s

Fig. 4.35 Effect of variations in the coefficient of friction on the particles movement



Rough contact ($\mu_{p,w} = 0.6$) *Standard case* ($\mu_{p,w} = 0.42$) *Lubricated contact* ($\mu_{p,w} = 0.2$)

Fig. 4.36 Effect of variations in the coefficient of friction on the wear pattern (red: intense wear; blue: light wear).

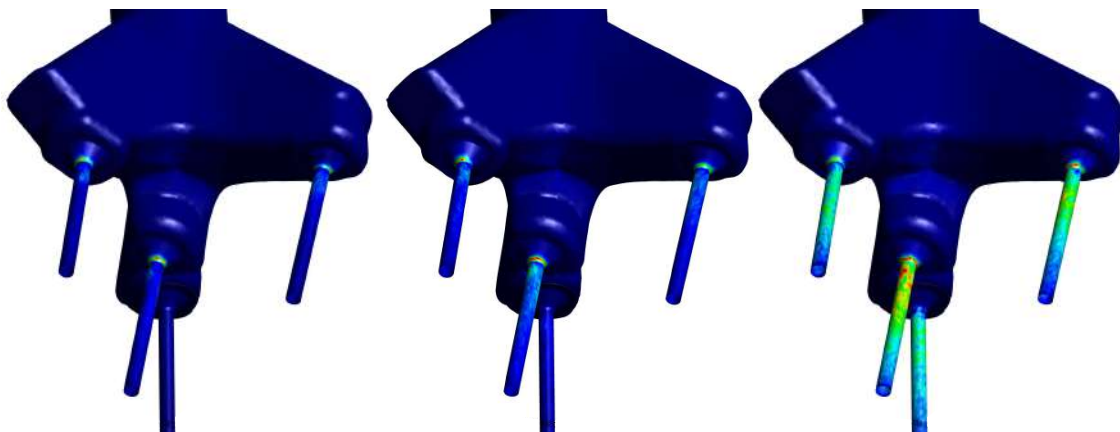


Strongly dissipative impact (COR = 0.4); 8759 spheres in the domain at t = 0.20 s

Standard case (COR = 0.6); 10 040 spheres in the domain at t = 0.20 s

Mildly dissipative impact (COR = 0.8); 12 469 spheres in the domain at t = 0.20 s

Fig. 4.37 Effect of variations in the coefficient of normal restitution (red: intense wear; blue: light wear).



Strongly dissipative impact (COR = 0.4)

Standard case (COR = 0.6)

Mildly dissipative impact (COR = 0.8)

Fig. 4.38 Effect of variations in the coefficient of normal restitution on the wear pattern (red: intense wear; blue: light wear).

Effect of neglecting inter-particle contacts Finally, we have investigated the importance of inter-particle contacts, since it would greatly simplify the analysis if we could do without them at all in the following ways:

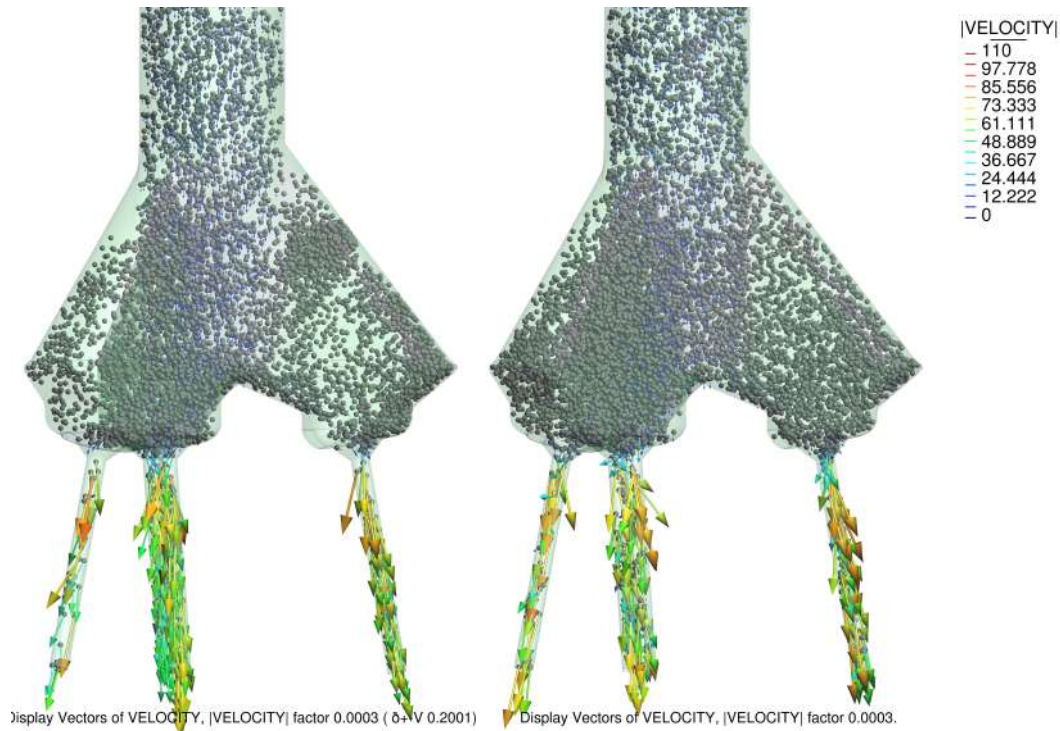
1. The computational cost would be greatly reduced, thanks to avoiding the most expensive parts of the DEM algorithm (force calculation and search)
2. The parameter space would be reduced, since only the particle-wall contact parameters would matter
3. It would be possible to alter (increase) the given concentration of particles to speed up the simulations, since each particle could be seen as an independent statistical test.

However, Fig. 4.39 shows that this simplification is not possible, at least at the current value of α_p . When no interactions are used, the flow becomes more chaotic as the particles do not become entrained in the general flow towards the nozzles. Therefore, there is an increased rate of accumulation. Since the rate of inter-particle momentum transfer is null, slow particles trapped in the recirculation region tend to remain there for longer times, increasing the particles concentration artificially. Note also the conspicuous effect on the wear pattern in Fig. 4.40.

An attempt to simplify the problem: Pseudo-steady-state solution

In order to reduce the computational requirements of the analysis, we attempted to simplify the problem by averaging the fluid field over a short time interval, and take the averaged field as the fluid field for the computations with particles. The fluid-only simulation was run until a pseudo-steady state solution was reached. In this case, a temporal criterion was used, although a criterion based on the variation rate of the flow could have just as easily be used. When this point was reached, the flow was averaged over time, for an interval considered sufficient. Our criterion was to average at least for a period equivalent to half the residence time of the flow. Fig. 4.42 shows contour plots comparing the velocity field for a single snapshot for the SC case with the average field at the same instant. Note that the average field is only rotated with the drillbit, but does not change relative to that solid-body motion. Apparently, the main traits of the flow are well captured.

However, we found out that this strategy (on its own) completely fails when the particles are taken into account. The main reason for this seems to be that the strong fluid accelerations around the entrance to the nozzles, when averaged over time, creates zones of low pressure that act as particle traps. When the flow is dynamic, these low pressure pockets are unstable and the material acceleration changes vigorously in time, so that no trapping occurs. Instead,



Standard case at $t = 0.20$ s; 10 040 spheres in the domain at $t = 0.20$ s

No interactions at $t = 0.20$ s; 12 788 spheres in the domain at $t = 0.20$ s

Fig. 4.39 Effect of inter-particle interactions on the particles movement

for the averaged field, these regions are capable of attracting particles and creating plugs (Fig. 4.43b, see also Fig. 4.44). In Fig. 4.44a, the plug that forms is shown from a zoomed in view. Clearly the system does not work as expected.

A similar phenomenon is observed, even with no inter-particle interactions, see Fig. 4.43c. See also the zoomed in detail of the nozzle entrance in Fig. 4.44b where the pockets of overlapped particles take a toroidal shape.

Measuring particle fluxes

In this section we show results corresponding to an alternative geometry proposed by the consultancy company. Its design is shown in Fig. 4.45. Here the objective was to design a way to monitor any difference in the performance of each of the nozzles. We designed a variant of the usual DEM rigid walls of analytic type (see Section A.4). The surfaces keep the information about all the particles that cross them, their velocity and possibly other data, storing it all in a single HDF5 file for later analysis.

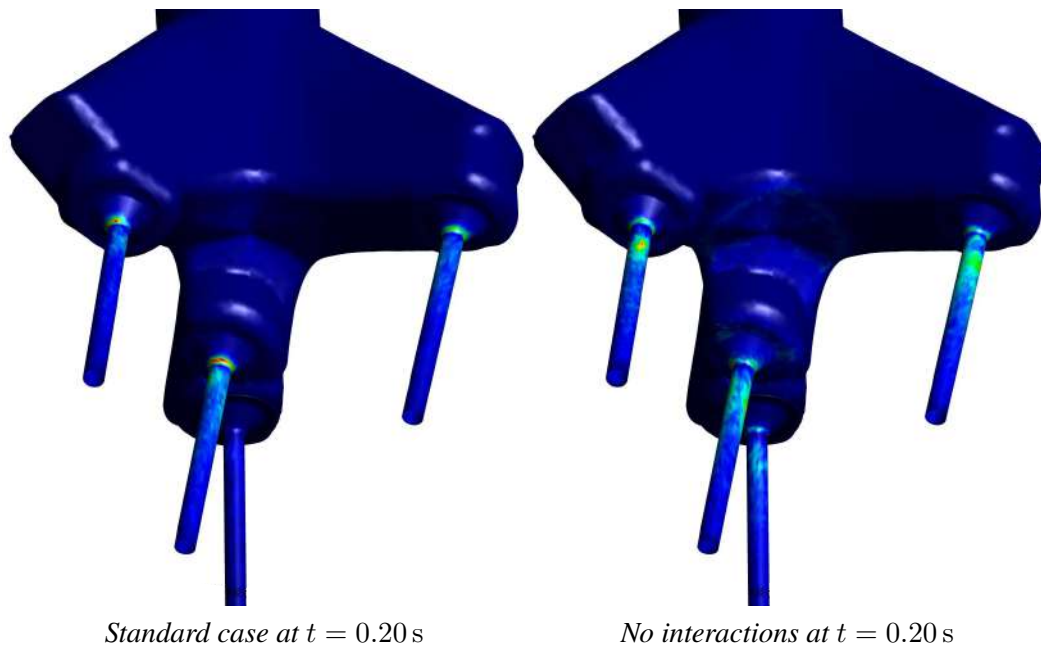


Fig. 4.40 Effect of inter-particle interactions on the wear pattern (red: intense wear; blue: light wear).

Fig. 4.46 shows the resulting measurements for the nozzle surfaces shown in Fig. 4.45b. The method allows to monitor the performance of the nozzles, tracking both velocity and mass. Here we only show the number of particles per second that pass through each of the nozzles, identified by the labels shown in the insert of the same figure. The simulation was run for some time until the averages looked stable enough. Clearly, there are important differences in performance between the different nozzles. This type of result is a very good candidate for future validation, as the flux measurements can be obtained through experiments. Unfortunately, no data was available to us at the moment of writing this work, and such validation is left for the future.

4.7.9 Discussion

Once more, we have applied our CFD-DEM numerical framework to perform simulations of a PID system using a one-way coupled strategy. We have demonstrated the possibilities of the technology by highlighting how a number of effects of relevance to the design of PID drillbits can be analysed for real geometries in operational conditions in a manageable amount of time.

We have demonstrated several interesting uses of such a technology in order to understand the PID systems better and aid in design, including

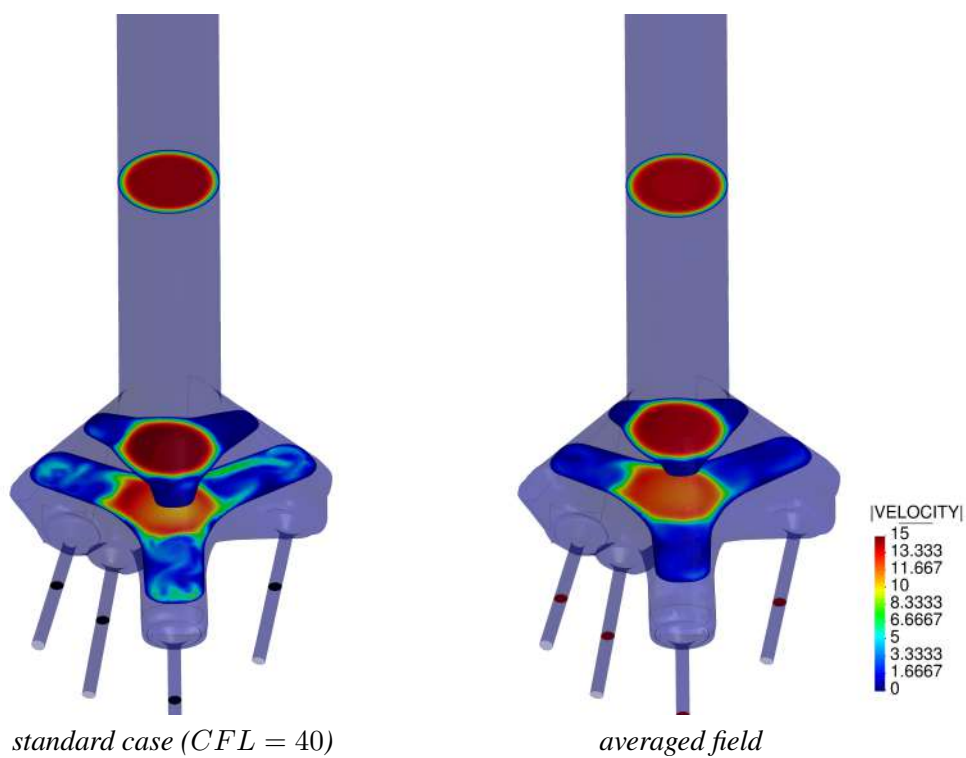


Fig. 4.41 transverse cuts

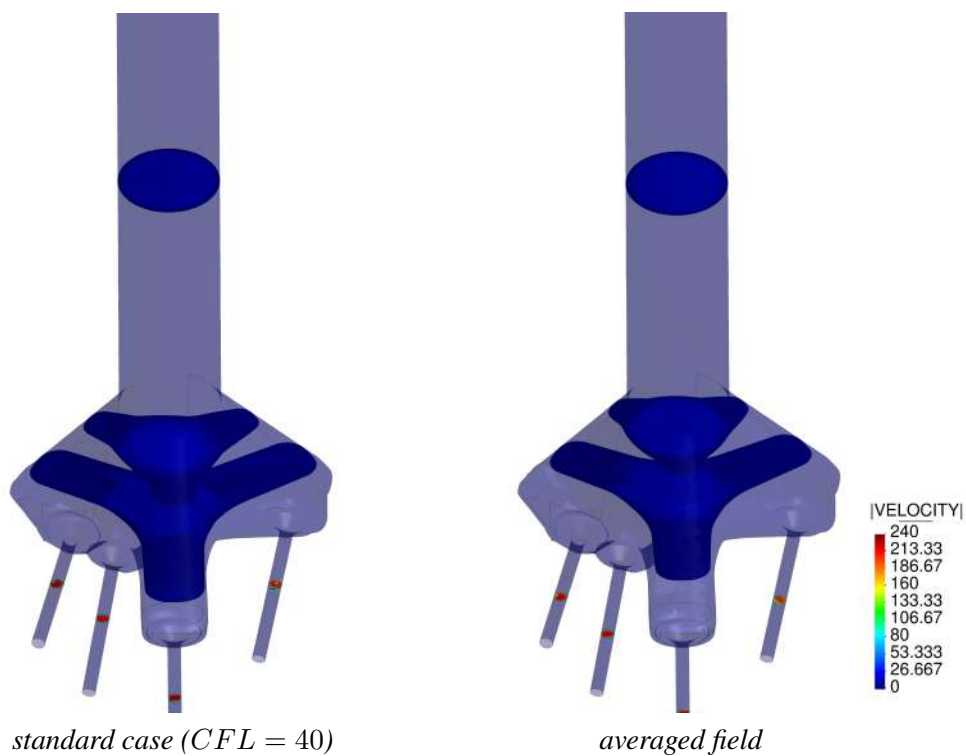


Fig. 4.42 Comparison of the modulus of the velocity field between the standard case at $t = 0.25$ s and the pseudo-steady field averaged velocities.

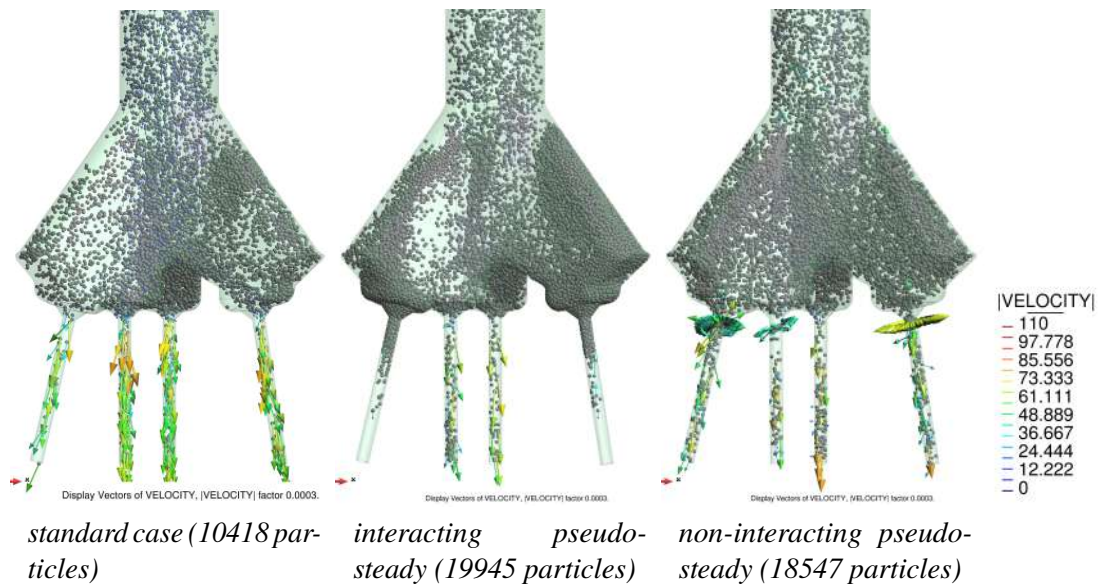


Fig. 4.43 Comparison of three snapshots taken at $t = 0.25$ s for the standard case and the pseudo-static fluid case with and without particles interactions.

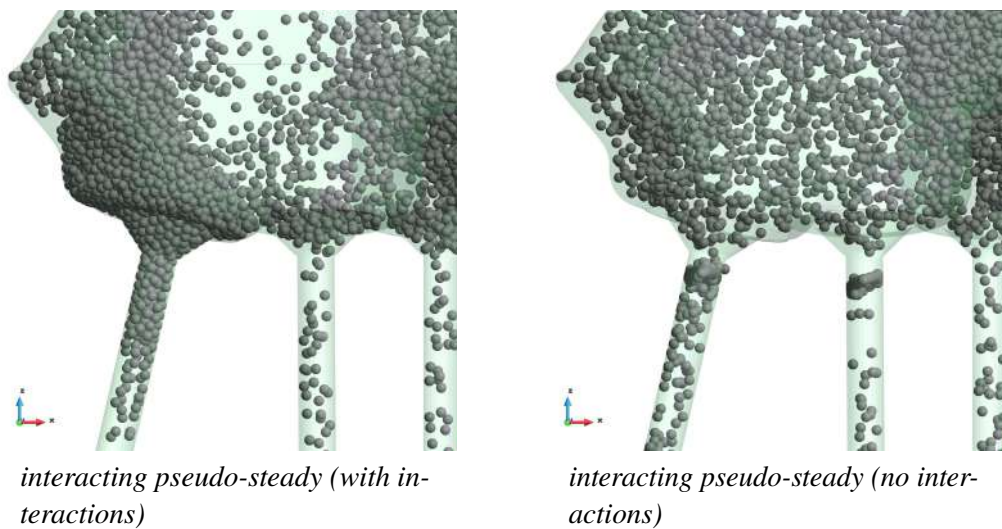


Fig. 4.44 Detail showing the accumulation of particles due to the steadiness of the fluid derivatives around the nozzle entrances with (left) and without (right) inter-particle interactions.

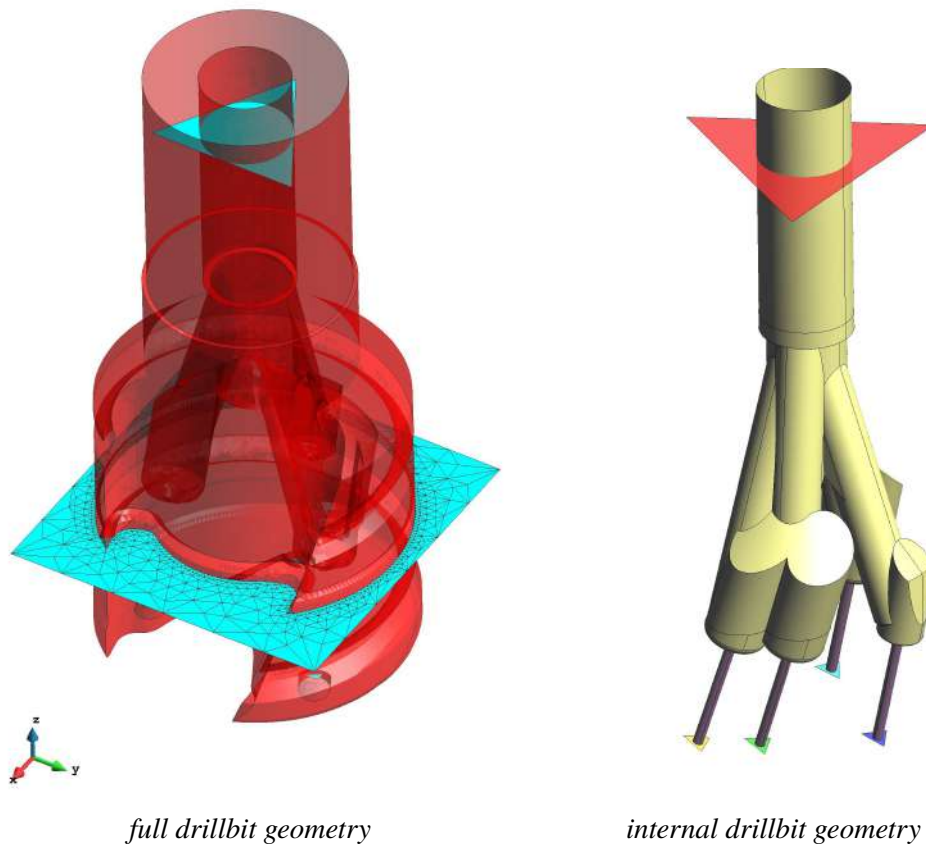


Fig. 4.45 Geometries of the PID drillbit with showing the flux measuring surfaces. Note that the surfaces are triangulations that can adapt to contours as in the left picture. However, it is computationally desirable to keep the surfaces as simple as possible (just one triangle is optimal).

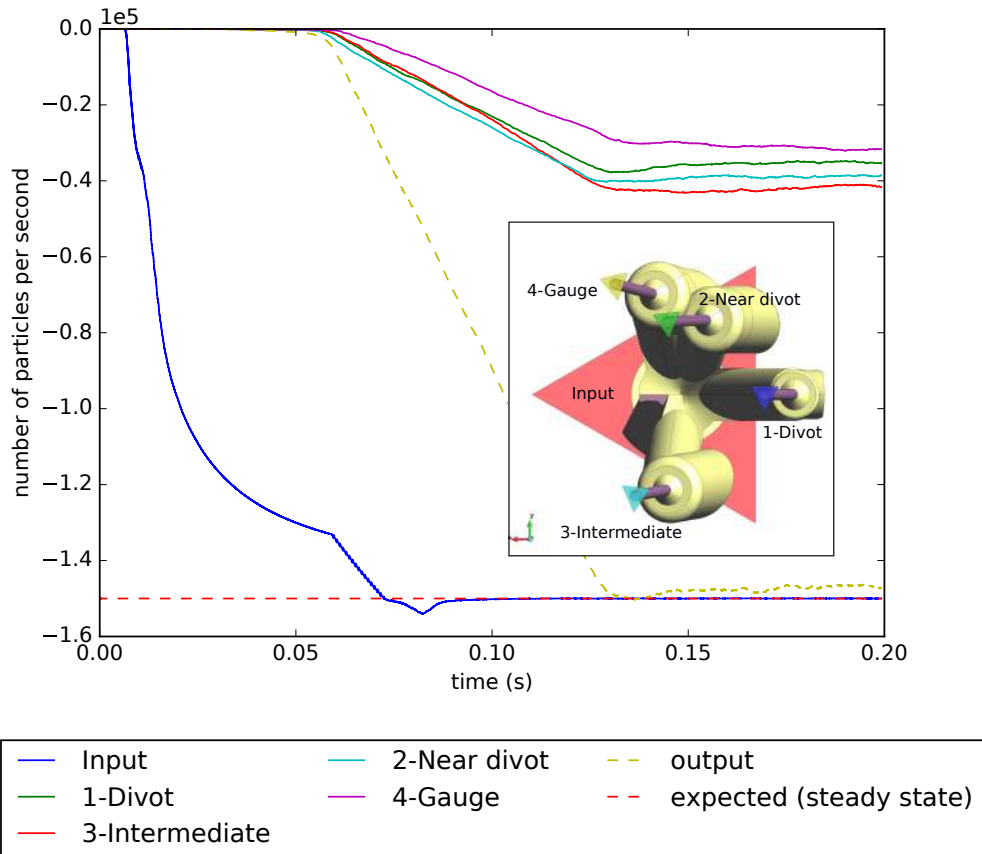


Fig. 4.46 Flux of particles through nozzles calculated with a moving average with an averaging interval of 0.07 s. The ‘output’ refers to the sum of the fluxes through the nozzles, while the expected flux corresponds to the rate of injection of particles $1.5 \times 10^5 \text{ s}^{-1}$. The sign of the measurements corresponds to the particular orientation of the normals that define the measuring surfaces, which are all opposite to the sense of the main fluid flow. The insert identifies the flux-measuring surfaces in a bottom-up perspective.

- The depiction of the internal flow of particles, including their distribution among the several pipes
- The prediction of wear concentrations, which can be used to improve the durability of designs
- The study of the outer flow patterns, which may help understand the movement of particles and possible sources of clogging and undesired accumulation.
- The study of the effect of changes in the number and granulometry of the metal particles

Our analysis suggests possible reductions in the parameter space, such as the elimination of the friction coefficient as an important parameter. It also highlights the importance of other parameters like the COR, which should be determined with sufficient precision in order to build a sufficiently accurate numerical tool. The same analysis also highlights the need for further research to determine the validity of our simplified coupling scheme, and in particular

- Extending the validity of the drag law by Shah for larger Reynolds numbers and testing it in instationary settings
- Studying the importance of lift
- Studying the need for more a sophisticated strategy that includes backward coupling, since the suspension is not lean enough to neglect the influence of the particles on the fluid phase a priori.

4.8 Backward-coupled particle-laden flows

As the solid fraction of particles in a given control volume raises, their influence on the flow becomes more important. This was already mentioned in Section 2.2.4 where, although the main concern was the motion of the particles, it was recognized that their mutual effects required an explanation via coupling to the fluid-phase, essentially through changes in its velocity. The issue was brought up again in Section 4.7, where it was recognized that neglecting these effects was probably causing important inaccuracies. In this section we tackle the issue, extending the formulation to include backward-coupling effect on the continuous phase.

For very lean flows, the study of the effects that the presence of particles have on the fluid is based on simple momentum transfer mechanisms. Specifically, the study of turbulence modulation by suspended particles [165] (see also [298] for a review on the most important variables involved) has been studied both from the theoretical [86, 256] and

numerical [126, 360] points of view. The traditional approach has been to consider point sources to model the effect of the particles [322]. The success of such methods is mixed [113], although they certainly represent an improvement over the simple one-way coupled approach in some situations.

Nonetheless, the utility of point-source approaches are limited to very low solid fractions. For one, they do not take into account the volume-displacement effects due to the presence of a significant proportion of particles, even if only locally. The literature has provided some heuristic estimates for the need to take this volume displacement effects into account [116].

The theory of multicomponent media [109, 179] provides an adequate framework to incorporate both effects into the analysis. This generalization of standard continuum media allows for the presence of an arbitrary number of continuum media superposed over the same spatial region, each with its own set of state variables and constitutive equations though allowing for coupling relations (differential or algebraic equations) to hold. The basic conservation equations are presented in Appendix H and can be taken as postulates to start the analysis [109]. It is however also possible to derive such equations by applying averaging techniques to the lower-scale description, just as the equations of the continuum can often be derived by applying averaging techniques to the microscopic description of the molecules.

We present in Appendix H a brief description of this averaging process that is a slight generalization of the volume averaging technique of Anderson and Jackson [8], see also [366]. This averaging process is of interest here as a tool to motivate the numerical method that is later described. We will focus on the fluid-phase averaged equations, since the particle-phase averaged equations are replaced by the DEM in the final algorithm.

4.8.1 CFD-DEM model

The hybrid CFD-DEM model is a multi-scale approach, where one simultaneously solves for the evolution of two coupled systems described at different levels of spatio-temporal detail. That is, the fluid is described at a coarser scale than the particles, and this scale is larger or equal to the mesh size used to discretize the continuous phase. The method has gained popularity in relatively recent times and its fundamentals can be found, for example in [372].

In order to construct the method, one takes a closed system of the two-fluid multicomponent equations that describe the evolution of the averaged variables, such as the velocity and volume fractions of the two phases involved. Then one replaces the balance equations (momentum and energy) relative to the particles by a DEM problem describing the evolution of the particles with appropriately modified coupling terms.

Our starting point are equations Eq.(H.54). We directly apply a common closure [355, 182] to them, obtaining the following set of equations:

$$\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f \mathbf{u}_f) = 0 \quad (4.119)$$

$$\frac{\partial \alpha_p}{\partial t} + \nabla \cdot (\alpha_p \mathbf{u}_p) = 0 \quad (4.120)$$

$$\alpha_f \rho_f \left(\frac{\partial \mathbf{u}_f}{\partial t} + \mathbf{u}_f \cdot \nabla \mathbf{u}_f \right) = \alpha_f \nabla \cdot \boldsymbol{\sigma}_f - \mathbf{f}_{fp} + \alpha_f \rho_f \mathbf{g} \quad (4.121)$$

$$\alpha_p \rho_p \left(\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p \right) = \alpha_p \nabla \cdot \boldsymbol{\sigma}_f + \nabla \cdot \boldsymbol{\sigma}_p + \mathbf{f}_{fp} + \alpha_p \rho_p \mathbf{g} \quad (4.122)$$

$$\boldsymbol{\sigma}_p : \nabla \mathbf{u}_p - \gamma_p = 0 \quad (4.123)$$

with

$$\boldsymbol{\sigma}_f = -p \mathbf{I} + 2\mu_f \left(\nabla^S \mathbf{u}_f - \frac{1}{3} (\nabla \cdot \mathbf{u}_f) \mathbf{I} \right) \quad (4.124)$$

$$\boldsymbol{\sigma}_p = -p_p \mathbf{I} + 2\mu_p \nabla^S \mathbf{u}_f + \left(\lambda - \frac{2}{3} \mu_p \right) \nabla \cdot \mathbf{u}_f \mathbf{I}$$

and

$$\mathbf{f}_{fp} = \beta^* (\mathbf{u}_f - \mathbf{u}_p) \quad (4.125)$$

with $\nabla^S \mathbf{a} := \frac{1}{2} (\nabla \mathbf{a} + (\nabla \mathbf{a})^\top)$ for any \mathbf{a} and \mathbf{I} is the identity matrix and were we have abused the notation by denoting the averaged stress tensors with subindices instead of overhead indices, such as the ones used in Appendix H. Eq.(4.123) is an algebraic equation of the granular temperature, Θ , where γ_p is the dissipation due to the inelastic nature of the inter-particle impacts, which depends solely on it. This equation is expected to be accurate enough at highly dense flows [355]. The pressure p is taken as an independent variable and the fluid phase viscosity is taken as $\mu_f = \mu$, the viscosity of the fluid itself.

The motivation behind Eq.(4.125) is that \mathbf{f}_{fp} can be interpreted as the averaged force on the particles phase per unit volume applied by the surrounding fluid minus the term $\alpha_p \nabla \cdot \boldsymbol{\sigma}_f$ (see Eq. (H.54)); where the latter term can in turn be interpreted as the averaged unperturbed flow force per unit volume of disperse phase. Thus, it is natural to model this term as the force terms in the equation of motion of a single particle (with the appropriate modifications due to the presence of neighbouring particles) over the volume of the particle, minus \mathbf{F}_U . Eq.(4.125) is one possibility, where the history and lift terms have been neglected. This is a very common practice for dense flows, where the drag force is typically dominant (most applications involve gas-solid flows) and the modelling uncertainties do not justified any more sophisticated models. The coefficient β^* can be determined based on any variant Eq.(4.8).

It is not necessary to provide explicit closures for the particle-phase energy equation Eq. (4.123) since it, along with Eq. (4.122) are replaced by the DEM problem. The resulting coupled system of equations can be formally written as

$$\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f \mathbf{u}) = 0 \quad (4.126)$$

$$\alpha_f \rho_f \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = \alpha_f \nabla \cdot \boldsymbol{\sigma} - \mathbf{f}_{fp} + \alpha_f \rho_f \mathbf{g} \quad (4.127)$$

$$m_p^i \frac{d\mathbf{v}^i}{dt} = \mathbf{F}_{\text{hydr}}(\mathbf{u}, \mathbf{x}^i, \mathbf{v}^i) + \mathbf{F}_B(\mathbf{x}^i) + \mathbf{F}_{\text{contact}}(\mathbf{x}^i, \mathbf{v}^i) \quad \forall i \quad (4.128)$$

$$\alpha_f = 1 - \langle 1 \rangle_p = 1 - \left\langle \sum_i \delta(\|\mathbf{x} - \mathbf{x}^i\|) V_p^i \right\rangle_p \quad (4.129)$$

$$\mathbf{f}_{fp} = \frac{1}{V_p^i} \langle \mathbf{F}_{\text{hydr}}^i - \mathbf{F}_U^i \rangle_p = \left\langle \sum_i \delta(\|\mathbf{x} - \mathbf{x}^i\|) (\mathbf{F}_{\text{hydr}}^i - \mathbf{F}_U^i) \right\rangle_p \quad (4.130)$$

where the i superindices identify variables relative to particle i and where the particle-averaged velocity and stress tensor do not explicitly appear in the system any more, allowing us to rename the fluid-averaged velocity and averaged stress tensor simply to \mathbf{u} and $\boldsymbol{\sigma}$.

4.8.2 The continuous-phase problem for the backward-coupled fluid

Let Ω be an open polyhedral domain of $\mathbb{R}^{n_{\text{dim}}}$, where $n_{\text{dim}} = 2, 3$ is the number of space dimensions, $\Gamma = \delta\Omega$ its boundary and $[0, T]$ the time interval of analysis. Let $\alpha_f = \alpha_f(\mathbf{x}, t)$ be a smooth real field on $\overline{\Omega} \times (0, T)$, where the overline denotes closure, representing the pointwise averaged fluid fraction. Using Eq. (4.125) into the first two equations in Eq. (4.126) yields the following system of equations:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{a} \cdot \nabla \mathbf{u} + \nabla p - 2\nu \nabla \cdot (\nabla^S \mathbf{u}) + \frac{2}{3} \nu \nabla (\nabla \cdot \mathbf{u}) = \mathbf{f} \quad \text{in } \Omega \times (0, T) \quad (4.131)$$

$$\alpha_f \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla \alpha_f = -\frac{\partial \alpha_f}{\partial t} \quad \text{in } \Omega \times (0, T) \quad (4.132)$$

where $\nabla^S \mathbf{u} := \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top)$. In order to have a well-posed problem, it is necessary to provide an adequate set of boundary and initial conditions. Furthermore, the above system can be expressed as a convection-diffusion equation. Let us use both facts to define the

problem we are interested in, i.e. *find* \mathbf{u} and p such that

$$\mathbf{M} \frac{\partial}{\partial t} \mathbf{U} + \mathcal{L}(\mathbf{U}) = \mathbf{F} \quad \text{in } \Omega \times (0, T) \quad (4.133)$$

$$\mathbf{u} = \mathbf{u}_0 \quad \text{in } \Omega \times \{0\} \quad (4.134)$$

$$\mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma_D \times [0, T) \quad (4.135)$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}_N \quad \text{on } \Gamma_N \times [0, T) \quad (4.136)$$

where

$$\mathbf{U} := \begin{bmatrix} \mathbf{u} \\ p \end{bmatrix}, \quad \mathbf{F} := \begin{bmatrix} \mathbf{f} \\ -\frac{\partial \alpha_f}{\partial t} \end{bmatrix} \quad (4.137)$$

and (summation is implied for repeated indices)

$$\mathcal{L}(\mathbf{U}) := \frac{\partial}{\partial x_i} (\mathbf{A}_i \mathbf{U}) - \frac{\partial}{\partial x_i} \left(\mathbf{K}_{ij} \frac{\partial \mathbf{U}}{\partial x_j} \right) \quad (4.138)$$

where, for $d = 3$,

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{K}_{ij} = \begin{bmatrix} \nu \delta_{ij} + \frac{\nu}{3} \delta_{1i} \delta_{1j} & \frac{\nu}{3} \delta_{1i} \delta_{2j} & \frac{\nu}{3} \delta_{1i} \delta_{3j} & 0 \\ \frac{\nu}{3} \delta_{2i} \delta_{1j} & \nu \delta_{ij} + \frac{\nu}{3} \delta_{2i} \delta_{2j} & \frac{\nu}{3} \delta_{2i} \delta_{3j} & 0 \\ \frac{\nu}{3} \delta_{3i} \delta_{1j} & \frac{\nu}{3} \delta_{3i} \delta_{2j} & \nu \delta_{ij} + \frac{\nu}{3} \delta_{3i} \delta_{3j} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.139)$$

$$\mathbf{A}_i = \begin{bmatrix} a_i & 0 & 0 & \delta_{i1} \\ 0 & a_i & 0 & \delta_{i2} \\ 0 & 0 & a_i & \delta_{i3} \\ \alpha_f \delta_{i1} & \alpha_f \delta_{i2} & \alpha_f \delta_{i3} & 0 \end{bmatrix}, \quad \mathbf{S} = \mathbf{0} \quad (4.140)$$

4.8.3 Finite Element Formulation

In this section we modify the formulation presented in Section 4.3.2 to suit the equations presented in Section 4.8.2.

Weak form

Let us come up with a suitable weak formulation of the problem defined by Eq. (4.133). First, the relevant function spaces are defined in a way completely analogous to what was done in Section 4.3.2. The weak form of the problem can be stated as *find* $\mathbf{U} := (\mathbf{u}; p) \in \mathcal{X}$ such

that $\forall \mathbf{V} := (\mathbf{v}; q) \in \mathcal{Y}$

$$\int_{\Omega} \mathbf{V}^{\top} \left(\mathbf{M} \frac{\partial}{\partial t} \mathbf{U} + \mathcal{L}(\mathbf{U}) - \mathbf{F} \right) d\Omega = 0 \quad (4.141)$$

which is of the form Eq. (4.28) with

$$\begin{aligned} B(\mathbf{U}, \mathbf{V}) := & (\mathbf{a} \cdot \nabla \mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) + (\alpha_f \nabla \cdot \mathbf{u}_f, q) \\ & + (\mathbf{u} \cdot \nabla \alpha_f, q) + \nu (\nabla^S(\mathbf{u}), \nabla \mathbf{v}) - \frac{2\nu}{3} ((\nabla \cdot \mathbf{u}), \nabla \cdot \mathbf{v}) \end{aligned} \quad (4.142)$$

and

$$L(\mathbf{V}) := (\mathbf{f}, \mathbf{v}) + \left(-\frac{\partial \alpha_f}{\partial t}, q \right) + \int_{\Gamma_N} \mathbf{t}_N \cdot \mathbf{v} d\Gamma \quad (4.143)$$

According to Eq.(4.50), the terms that must be added to the right hand side of the standard Galerkin formulation for the OSS formulation are the following

$$\begin{aligned} & \sum_e \int_{\Omega_e} \tau_1 (\boldsymbol{\pi}_{\text{mom}} + \mathbf{d}_{\text{mom}}) \cdot \left(\nu \Delta \mathbf{v}_h + \frac{\nu}{3} \nabla (\nabla \cdot \mathbf{v}_h) + (\mathbf{a}_h \cdot \nabla) \mathbf{v}_h + \alpha_{f,h} \nabla q_h \right) d\Omega + \\ & \sum_e \int_{\Omega_e} \tau_2 (\boldsymbol{\pi}_{\text{mass}} + d_{\text{mass}}) (\nabla \cdot \mathbf{v}_h) d\Omega \end{aligned} \quad (4.144)$$

Meanwhile the analogous terms for the ASGS method are obtained by replacing $\mathbf{d} + \boldsymbol{\pi}$ by $\mathbf{d} + \mathbf{F}$ in the integrals above.

On the other hand, let us now expand the term $-(\mathcal{L}(\mathbf{U}_h^{n+\theta}), \mathcal{L}^*(\mathbf{V}_h))_{\tau}$, which is present in both formulations. The formal adjoint of operator \mathcal{L} is obtained by transposing the matrices and changing the sign of odd-order terms; i.e.

$$\begin{aligned} \mathcal{L}^*(\mathbf{a}; \mathbf{V}) \equiv \mathcal{L}^*(\mathbf{V}) = & -\frac{\partial}{\partial x_i} (\mathbf{A}_i^{\top} \mathbf{V}) - \frac{\partial}{\partial x_i} \left(\mathbf{K}_{ij}^{\top} \frac{\partial \mathbf{V}}{\partial x_j} \right) \\ = & \begin{bmatrix} -\nu \Delta \mathbf{v} - \frac{\nu}{3} \nabla (\nabla \cdot \mathbf{v}) - (\mathbf{a} \cdot \nabla) \mathbf{v}_f - \alpha_f \nabla q \\ -\nabla \cdot \mathbf{v} \end{bmatrix} \end{aligned} \quad (4.145)$$

Consequently, for the case of negligible second derivatives (as in piecewise linear elements, where they exactly vanish within elements) we have:

$$\begin{aligned}
& - (\mathcal{L}(\mathbf{U}_h^{n+\theta}), \mathcal{L}^*(\mathbf{V}_h))_\tau \\
& = \sum_e \int_{\Omega_e} \tau_1 (\mathbf{a}_h \cdot \nabla \mathbf{u}_h + \nabla p_h) \cdot ((\mathbf{a}_h \cdot \nabla) \mathbf{v}_h + \alpha_{f,h} \nabla q_h) d\Omega \\
& + \sum_e \int_{\Omega_e} \tau_2 (\mathbf{u}_h \cdot \nabla \alpha_{f,h} + \alpha_{f,h} \nabla \cdot \mathbf{u}_h) (\nabla \cdot \mathbf{v}_h) d\Omega
\end{aligned} \tag{4.146}$$

The resulting elemental matrices resulting from the FEM discretization of the formulations above are detailed in Appendix H for Q-ASGS and Q-OSS.

4.8.4 Backward coupling method

The backward coupling consists in the determination of the discrete counterparts of \mathbf{f}_{fp} and α_f from the DEM solution. According the theory developed in Appendix H, these variables are to be filtered from the disperse phase. Next we describe two alternative methods to do this.

Linear conservative projection

For every particle, its contribution to the discretized fluid fraction field affects its host element's nodes only. Let $\alpha_{p,h}$ be the discrete form of the particles' phase volume fraction with nodal values A_p^b . The linear conservative method is defined by

$$A_p^b = \frac{4}{|\Omega_b|} \sum_{p \in \Omega_b} V_p N^b(\mathbf{x}_p) \tag{4.147}$$

where the same nomenclature as in Eq.(4.73) is used.

This scheme is conservative in the sense that the integrated total volume is conserved after projection. The linearity of the integral of the finite element discretization implies that the effect of many particles is the addition of the effect of each single particle. The conservation property can be readily checked by looking at the solid fraction field produced by a single

particle. We should integrate the generated $\alpha_{p,h}$ over the whole domain:

$$\begin{aligned} \int_{\Omega} N^b(\mathbf{x}) A_p^b d\Omega &= \int_{\Omega} N^b(\mathbf{x}) \frac{4}{|\Omega_b|} \sum_{p \in \Omega_b} V_p N^b(\mathbf{x}_p) d\Omega \\ &= \sum_{b|N^b(\mathbf{x}_p) \neq 0} N^b(\mathbf{x}_p) \frac{4V_p}{|\Omega_b|} \int_{\Omega} N^b(\mathbf{x}) d\Omega = V_p \sum_{b|N^b(\mathbf{x}_p) \neq 0} N^b(\mathbf{x}_p) = V_p \end{aligned} \quad (4.148)$$

where the last equality follows from the partition of unity property. Therefore, the global mass balance is conserved at all times. This conservation property also holds for other projected fields, such as the hydrodynamic forces exchanged with the particles, so that the total exchange forces between both phases fulfil an averaged version of Newton's Third Law.

The method just described is used to compute the averaged fields from a specific DEM solution, which can be assumed to be the latest value calculated. Moreover, it is also possible to smooth the solution in time by averaging the averaged fields over several DEM time steps, as is explicitly done with the technique that follows.

Polynomial filter

When a particle crosses over from one element to another, the fluid fraction field experiments a discontinuous jump proportional to the ratio of volumes of these two elements. To reduce this effect as well as to allow for a larger influence range of each particle, a filtering methodology has been devised in which each particle's volume is smeared over a number of nodes: those found inside a certain sphere concentric to the particle whose radius determines the extend of this smearing. This procedure starts by a search to determine, for each particle, which set of nodes falls inside the particle's influence domain. After that, each node within the domain is assigned a weight which monotonically decreases with distance to the particle, and so that their sum equals 1. This way, conservation is preserved as explained in Section 4.8.4. This procedure can be understood as a direct application of the filtering theory of Section H.1.4. In particular, within the point-particle approach particles are described as points, so that integrals over their volume must be replaced by Dirac deltas.

Let us assume we want to find the filtered value of an extensive quantity α defined inside the particles' volume. Its filtered counterpart would be found by applying Eq.(H.59) to

$\mathbf{A}\delta(\mathbf{x} - \mathbf{x}^i)$:

$$\begin{aligned} \langle \mathbf{a} \rangle_k(t, \mathbf{x}) &= \int_I \int_{\Omega} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \sum_i \mathbf{A}^i(\tau) \delta(\|\mathbf{y} - \mathbf{x}^i(\tau)\|) d\mathbf{y} d\tau = \\ &= \int_I h(|t - \tau|) \sum_i g(\|\mathbf{x} - \mathbf{x}^i(\tau)\|) \mathbf{A}^i(\tau) d\tau \end{aligned} \quad (4.149)$$

where i runs through all the particles in the domain and where \mathbf{A}^i is the integral of \mathbf{a} over the particle centred at \mathbf{x}^i . For the time filter we use a simple quadrature:

$$\int_I h(|t - \tau|) \mathbf{f} d\tau \approx \sum_{n-k}^n \Delta t_{\text{DEM}} \mathbf{f}^n h^n \quad (4.150)$$

where the indices indicate the time of evaluation. This operation can be performed accumulatively, at every DEM time step (or skipping a few at a time). However here we take

$$h(\tau) = \frac{1}{\Delta t_{\text{DEM}}} \delta(\tau) \quad (4.151)$$

So Eq.(4.150) becomes an evaluation at the current time. Thus, we can now concentrate on the space averaging by only assuming that the variables are evaluated at the current time step. Using this methodology, one immediately obtains a formula for the value of the filtered variables at the mesh nodes' locations and then extend their value over the whole domain using the FEM shape functions. For instance, the solid fraction can be calculated as (assuming all the variables are evaluated at the current time step)

$$\alpha_{p,h} = \sum_i \sum_a g(\|\mathbf{x}^a - \mathbf{x}^i\|) N^a V_p^i \quad (4.152)$$

where V_p^i is the volume of the i -th particle and a runs through the mesh nodes.

Note that it would be convenient to have the property

$$\int_{\Omega} \alpha_{p,h} d\Omega = \sum_i V_p^i \quad (4.153)$$

Which is the analogous density-like property assumed to hold in the postulational approach of multicomponent continuum mechanics (see Appendix H). Using Eq. (4.152), such condition can be developed into

$$\sum_i V_p^i \sum_a g(\|\mathbf{x}^a - \mathbf{x}^i\|) |\Omega_a| = \sum_i V_p^i \sum_a \bar{g}(\|\mathbf{x}^a - \mathbf{x}^i\|) = \sum_i V_p^i \quad (4.154)$$

where we define \bar{g} as a modified filter function. The condition above is verified, in particular if \bar{g} is a partition of unity. In our case, we have opted for a hybrid approach, where we use a generic bump function but impose the conservation property Eq.(4.153) a posteriori, by normalizing the contributions for each i . In this way, we guarantee that information is smoothed, that particles closer by are weighed more highly and that Eq.(4.153) holds. We have found this scheme to be quite robust, avoiding the instabilities that might appear when the size of the particles becomes closer to the elemental size when using more standard approaches [276].

4.8.5 Final algorithm

The algorithm used for the two-way coupled model is very similar to the one described in Section 4.7.6, but obviously with the modified finite element formulation of the coupled problem that we discussed in the present section. Algorithm 7 shows the basic layout of this algorithm where the main difference with respect to Algorithm 5 is the addition of a backward-coupling step, designated in the pseudo-code with the name `PERFORMBACKWARDCOUPLING()`. The background coupling does not need to be done at every sub-time step. In fact once every fluid time step is normally enough ⁸ if no time filtering is applied. The variable N_{back} has been introduced to express this new substepping for the background coupling.

4.9 Application example: fluidized bed

In this section we apply the theory presented above to a representative problem in which the backward coupling is unavoidable. This problem was studied numerically recently by Boyce et al. [42] where the limitations of the volume averaging approach in the CFD-DEM were discussed and highlighted. Our objective here is to see if our finite element-based approach can obtain the same qualitative behaviour obtained with a more conventional approach and to check whether the same problems observed there are reproduced here.

The problem at hand is the simulation of a gas-fluidized bed of Geldart-D type [143] particles (poppy seeds), that were experimentally studied previously by Holland et al. [170]. The container is cylindrical, with a circular base. The geometry corresponds to configurations 2 and 3 in Boyce et al. [42], and it is depicted in Fig. 4.47a. The initial bed size is of 50 mm. The most important input parameters are detailed in Table 4.9, and the relevant inter-phase interactions in Fig. 4.50. In the original work hexahedral cells arranged in a structured

⁸It is not logically necessarily so, since it is possible that \mathbf{f}_{fp} depends on α_f . This means that it is conceivable to have the solid fraction calculated more often than the fluid, when its value needs to be repeatedly updated within one fluid time step during the integration of the particles motion.

Algorithm 7 Two-way-coupled two-phase algorithm.

```

1:  $t \leftarrow 0.0$ 
2:  $t_{\text{DEM}} \leftarrow 0.0$ 
3:  $m \leftarrow 0$ 
4:  $\mathbf{U}^0 \leftarrow \mathbf{U}_{\text{init}}$ 
5: INITIALIZEPARTICLES()
6: SEARCHNEIGHBOURS()
7: for  $n$  in  $[0, N_{f,\text{steps}}]$  do
8:   ROTATEMESH( $t, n$ )           ▷ rotate the fluid mesh with imposed angular velocity
9:   SOLVEFLUID( $t, n$ )           ▷ Algorithm 6
10:  RECOVERDERIVATIVES()       ▷ Section 4.4
11:   $n_{\text{substep}} \leftarrow 0$ 
12:  while  $t_{\text{DEM}} < t$  do
13:     $n_{\text{substep}} \leftarrow n_{\text{substep}} + 1$ 
14:    SOLVEDEM( $t_{\text{DEM}}, m$ )     ▷ Algorithm 9
15:    if  $\text{mod}(n_{\text{substep}}, N_{\text{back}} = 0)$  then
16:      PERFORMBACKWARDCOUPLING() ▷ e.g. Eq.(4.152)

```

way were used for the fluid discretization. We instead employed the sequence of irregular tetrahedral meshes shown in Fig. 4.47, where the average sizes (expressed as fractions of the particles' diameter) are calculated as the side of a regular tetrahedron that has the same volume as the average volume of the tetrahedra in the mesh. The average sizes are comparable to the ones used considered in Boyce et al. [42].

Table 4.9 Material parameters considered in the fluidized bed example

Parameter	Value	Description
Fluid parameters		
ρ_f	1 kg/m ³	density of fluid
ν	1.5×10^{-5} s/m ²	kinematic viscosity
Particles parameters		
d_p	1.2×10^{-3} m	particles' diameter
μ_p	0.1	friction coefficient ⁹
ρ_p	900 kg/m ³	apparent density of particles
COR	0.2	coefficient of normal restitution

⁹No difference is made between the particle-wall and the particle-particle friction coefficients in either [355] nor [170]. We have therefore assumed they coincide.

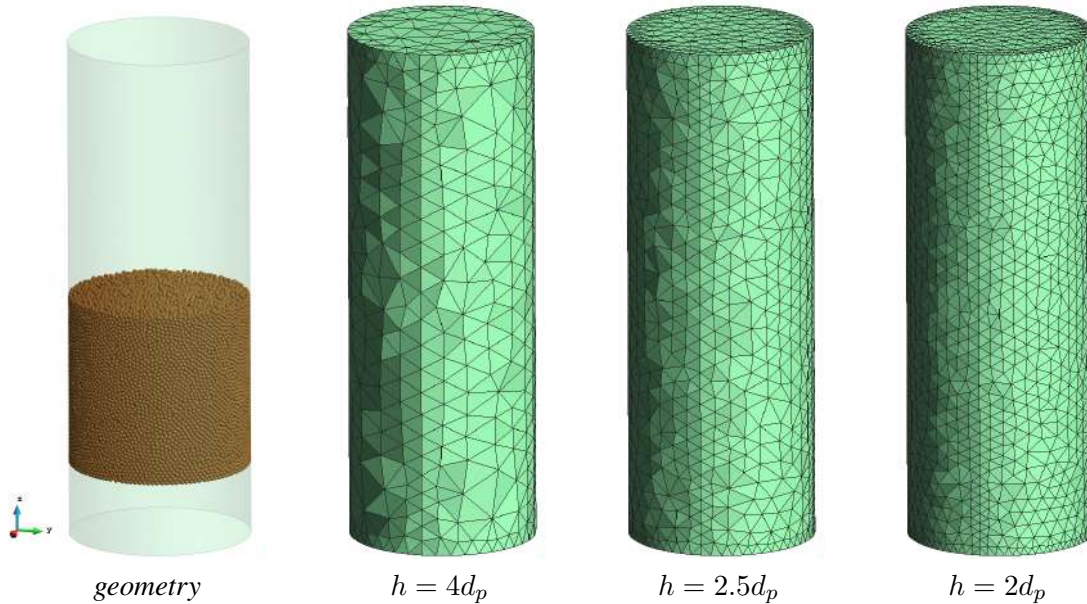


Fig. 4.47 Model geometry with initial particles bed and computational meshes considered.

Fig. 4.49a shows a window of the pressure in time as measured at a single point, for the different meshes considered. The pressure values are made dimensionless by dividing them by the average pressure drop in the first 50 mm from the bottom of the bed, as measured on the symmetry axis of the container. From the graph, it is clear that while the average pressure is not sensitive to the mesh size, but the amplitude of its oscillations does show this dependency. Moreover, Fig. 4.49b shows the dependency of the oscillation frequency with respect to the mesh size. These frequencies were obtained by identifying peaks in the note that this dependency is clear although somewhat less marked than that observed in Boyce et al. [42], where a method analogous to the standard method was used for the coupling. We can conclude therefore that our method suffers from the same mesh-dependent problems reported by others. This may be due to intrinsic limitations of the filtered equations themselves and not the coupling method, although more work is needed to determine this.

4.10 Summary

In this chapter we have described a numerical approach that can be used to simulate a wide range of particle-laden flows. The scope of this chapter has greatly widened with respect to the preceding two, considering a much greater range of physical situations, including larger particle Reynolds numbers, non-Newtonian fluids and dense suspensions, but also

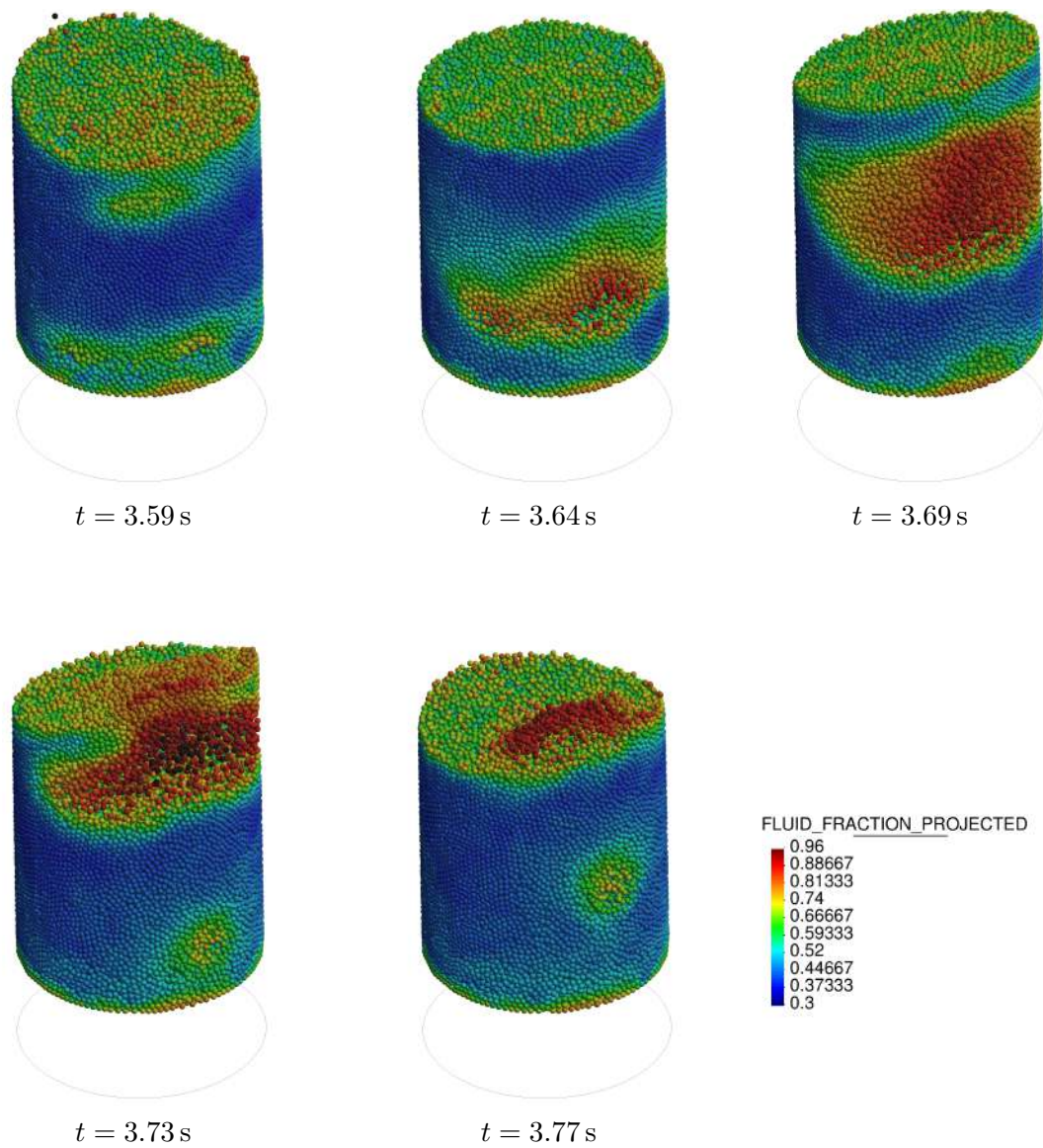
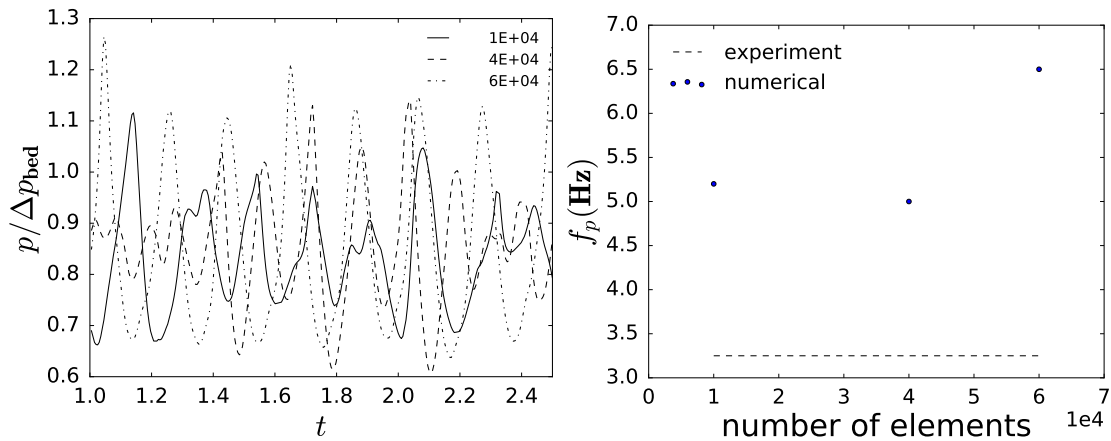


Fig. 4.48 Sequence of snapshots showing the lifetime of a bubble generated close to the walls, as calculated using the finest mesh. The colours show the fluid-fraction calculated at the center of the particles.



Evolution of pressure normalized by the measured pressure drop in the first 50 mm above the bed bottom.

Bubbling frequency as a function of the number of tetrahedral elements. The dashed line indicates the experimental value

Fig. 4.49 Measurements at point (0, 0.1, 0.1), where the axis of the cylinder passes through the origin at the bottom of the bed.

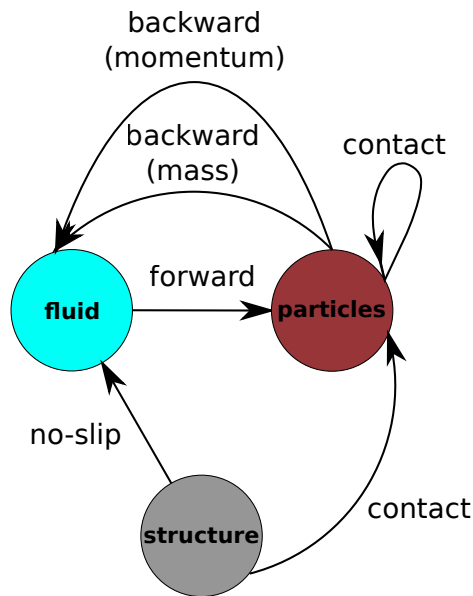


Fig. 4.50 Diagram of phase interactions accounted for in the fluidized bed simulation.

in the number of elements being simulated, including the fluid and the consideration of inter-particle contacts. In general, the core of the discourse has furthermore revolved around a varied set of applications, associated with the industry and the development of a versatile application capable of dealing with a large range of situations.

Nonetheless, we have made some genuine contributions of academic interest:

- A review of the state of the art in derivative recovery including methods associated with the FEM with a practical discussion of their implementation for their use in particle-laden flow simulations. We concluded with the recommendation of using the PPR method with a fall-back to the standard recovery method for difficult areas. When the memory resources are scarce, the standard recovery methodology is robust and reasonably accurate for first-order derivatives.
- The application of the VMS methodology to develop a stabilized FEM formulation for the CFD-DEM
- A backward-coupling methodology based on the notion of filtering the disperse phase that unties the discretization from the filtering scales and largely avoids the limitations associated with the element size to particle diameter ratio.

We have looked at three applications of the described methodology with industrial interest, each one illustrating a different coupling scheme. The first one deals with the phenomenon of bubble-trapping in T-junction bifurcations in liquid piping systems. We showed that the methodology captures the qualitative trends when compared to the literature, although some quantitative discrepancies remain. We have pointed to a possible explanation based on a small difference in the formulation, although it could be due to small differences in the fluid solution. The matter therefore requires further research in order to move toward robust quantitative predictability in such systems.

The second example explores a drilling system for the oil and gas industry. We used a one-way coupled approach to produce simulations of the flow of steel particles within the system. We gave a sensitivity analysis and made conjectures about the parameters that may be of lesser importance. We showed how the tool can be used by engineers to estimate the flux of particles within the system with the use of flux-measuring surfaces and how the system can be used to identify the areas more vulnerable to wear. Unfortunately, no experimental data was made available to us, although we are working toward undertaking this task in the future.

The third example showed a fluidized bed, as a paradigmatic case for the application of the two-way coupled methodology that we have developed, but that can be used in many other settings, including that of the second example, which is currently under study. This example

showed how the system correctly predicted the formation of bubbles, although we reproduced the difficulties encountered by Boyce et al. [42] regarding the estimation of the bubbling frequency. Again, it is our intention to keep studying this problem taking the developments presented here as a foundation.

In summary, we have presented a remarkably general tool that is able to greatly extend the range of application considered in the preceding two chapters. The tool is of interest to many industrial problems and its potential has been demonstrated in a set of tests representative of different coupling schemes.

Chapter 5

Conclusions and Future work

The particular technical advances and findings have already been summarized at the end of each chapter. Consequently, we dedicate the next lines to making only a number of general remarks about the work achieved in relation to the initial goals and some speculations about the future relevance of our contributions. We close our work discussing a selection of the topics that we have left for future work.

5.1 Concluding remarks

It is time to look back to the objectives that we set in Section 1.2 to see how far we went about fulfilling them. Let us go over the list once again:

1. *To develop an algorithm that combines the discrete element method and the finite element method to simulate particle-laden flows with the following list of requisites:*
 - *capability of dealing with a wide range of regimes, including the possibility to have regions with dense and dispersed suspensions simultaneously:* We demonstrated the capabilities of our coupled code in Chapter 4. The backward-coupled formulation is general enough to allow for dense flows and default to the one-way coupled formulation in the limit of small solid volume fractions.
 - *use of the finite element method to discretize the fluid:* Done.
 - *use of the discrete element method to model the particles:* Done.
2. *To study the range of applicability of the Maxey–Riley equation as a model for the motion of the individual particles submerged in a fluid, improving the current knowledge on the subject and generating, where possible, practical estimates of direct application*

to numerical modelling: Chapter 2 was fully dedicated to this task in what, we believe is the most comprehensive analysis of the range of applicability of the MRE to date. Naturally, our analysis is still far from exhaustive, having left some of the most important effects for future work, such as the first effects of the proximity of walls and a more complete study of the presence of lift and also complicated matters such as the effect of a systematic orientation in non-symmetrical particles. Furthermore, the conclusions of our analysis need to be thoroughly tested by experimental and numerical means. Such effort would help to both strengthen the reliability of the estimates and perhaps also discover weaknesses in some of the simplifying assumptions used.

3. *To study current alternatives for the numerical treatment of the history term in the equation of motion and compare them*: We reviewed in Chapter 3 some of the most recent techniques and compared them, updating the state of the art with respect to these novelties. We also made a long-needed connection between the literature on fractional differential equations and the MRE, which has not been sufficiently exploited.
4. *To improve on the method of quadrature (the history term involves a time integral, as will be explained later) proposed by van Hinsberg et al. [353] and provide a detailed study of its efficiency and accuracy, providing convincing evidence that it is not necessary to neglect this term to have an efficient numerical method*: This is the central subject of Chapter 3, where we considered a generalized version of the problem originally posed by van Hinsberg et al. We also provided a detailed study of the efficiency of the method, concluding that it opens the door to a systematic inclusion of the history force in disperse particle-laden flows due to its reasonable cost.
5. *To report an account of relevant application examples of the proposed strategy with interest to the industry, as well as of the different technologies developed for their particular requirements*: In Chapter 4 we have reported three examples representative of different coupling regimes relevant to different industries, including very dispersed, internal flows relevant to microfluidics (Section 4.6); the higher density, although still not too high of relevance to the PID system for the oil and gas industries (Section 4.7) and fully coupled, high density flows, especially relevant to problems where regions with very high density of particles are likely to develop, such as fluidized beds or bed formation in pipes (Section 4.9).
6. *To generalize a stabilized finite element method and use it to discretize the backward-coupled flow equations*: We gave such formulation in Chapter 4, although we left the numerical analysis for future work.

7. *To develop a suitable inter-phase coupling strategy:* We provided two formulations in Section 4.8.4. The filter-based method provided there is robust with respect to local peaks in the solid density, as it smears the variables over several elements, avoiding the appearance of instabilities due to too strong oscillations in the local momentum exchange. However, we have reproduced some of the problems found in the literature regarding the mesh-dependence of the solutions (see Section 4.9) that do require further work.

In general terms, we believe we have been able to fulfil, at least partially, all of the goals we set at the beginning of this work. Looking back, we find that the initial focus, which had been intended to fall upon the development of a two-way coupled algorithm (with a detailed numerical analysis to go with it) shifted toward a study of the single-particle coupling terms. The reason is the unexpected discovery of the large complexity and richness of this area and the need to understand it deeply in order to move toward a macroscopic formulation that, in the end, must be based on it.

The work in Chapter 2 is a reflection of this study. It is our intention to summarize the most valuable results from Chapter 2 in a paper, currently in production.

We also hope that its output is useful as a starting point for a gradual, ‘organic’ evolution of the range of applicability of the MRE. We are convinced that this type of knowledge will in the future be much more tightly associated with the model equation itself thanks to the latest advances in data processing and the internet. We envision that not only the limits of applicability will be associated to the model, but also the information about the certainty of these limits, contradictory data, the sources of each piece of information, associated and related models and experimental data etc.

Generalizing, we see no impediment for this scenario to also arise with many other equations, perhaps in very different fields. For that, it is important to proceed more systematically and get over the current times, where such information is held, in a unified manner, only in the brains of experts.

As a side effect of this careful study, we came across the world of the very disperse suspensions at very low particle Reynolds numbers, which require special care in formulating very precise algorithms. With the idea of building a general tool, we were faced with the necessity to include the history term of the MRE, which resulted in Chapter 3. The work contained in this chapter has been published in [63].

The tests run in Chapter 3 helped creating a generalistic CFD-DEM code with a solid foundation, that we hope will be very much enlarged in its capabilities in the near future. These tests are used as benchmarks to keep the code constantly verified.

Finally, some of the work in Chapter 4 has started interesting lines of research, such as the derivative recovery algorithms, which we plan to make generally available to the users of Kratos, since they can be of interest for many uses, including post-processing and error estimation. Similarly, many hydrodynamic interaction laws have been made available and have already found application in several areas apart from the ones demonstrated here, including pneumatic conveying of seeds and even some applications in the study of expansive clays at colloidal scale.

5.2 Future work

Our work has covered a rather wide range of topics which means that many more doors have been opened than closed. Our study in Chapter 2 was exhaustive by aspiration but many topics were only touched on superficially. The study of the history force from Chapter 3 served as an introduction on several techniques related to the numerical treatment of fractional differential equations that were simply mentioned and even our proposed method was left with several questions to be answered. But it is probably Chapter 4 that introduced the most topics that had to be left for the future.

5.2.1 Range of applicability of the MRE and scaling analysis

The limits collected in Section 2.4 should be tested to give empirical evidence to each of the numbers summarized in the tables we have provided. We are currently planning a campaign to document existing results and identify the weakest estimates. Numerical experiments should be useful in this respect, perhaps using isotropic turbulence with a DNS technique.

The estimates in Section 2.2.4 contain particular uncertainty, due to the complexity of the formulation and the many simplifications. A very interesting experiment which we believe has not been performed yet is a systematic study of the effect of neighbours as

1. Short-range effects, by comparing the motion of particles in a numerically simulated turbulent flow with and without these effects. To simulate these short-range effects one could use simplified formulations such as the one by [292].
2. Long-range effects, by comparing disperse flows with and without two-way coupling, taking into account what part of the interaction is treated indirectly, though the interaction with the fluid; and what part is treated directly. To this respect the work of Huck et al. [173] is relevant.

5.2.2 More on the history force calculation

In Section 3.5.1 we briefly visited the possibility of using extrapolation techniques to enhance the Grünwald–Letnikov formulation of the fractional derivative to produce high-order quadrature techniques to calculate the history term in the MRE. However, we never implemented such techniques to calculate the motion of submerged particles, since this required unspecified changes in the quadrature algorithms that require considerable work. Nonetheless, we think this path can lead to elegant methods to obtain high-order quadrature schemes and therefore the matter deserves further study.

5.2.3 Derivative recovery

The unsuccessful stabilization of the formulation of Pouliot et al. [282] for some regular meshes spoilt a method that otherwise provided a set of very positive properties, including:

- Very high accuracy when it works
- Good properties at the boundary [282]
- A structure naturally apt for an implementation within a typical FEM framework, facilitating parallelization (as the method would inherit all the parallelization work already done in the framework)

We still believe it might be possible to fix the method in a robust way, but the matter is not trivial and we had to leave it for future work.

5.2.4 Future developments for backward coupling

The inaccuracies that we reported in Section 4.9 with respect to the bubbling frequency are a matter of concern. In line with [42], we suspect a modification of the coupling scheme might help. It is equally important to study the mesh-dependence of the method and how it can be dealt with effectively. All these matters require further work.

5.2.5 Future developments for the applications

The various application examples that we have presented in Chapter 4 have also lead to various interesting paths for future work. For instance, ongoing work on the PID systems is requiring further considerations with regards to some of the assumptions taken so far. Here is a selection of the most important if these:

Non-Newtonian drag laws

The hydrodynamic interactions assumed to be valid for the steel particles in power-law type drilling mud are based on a combination of effects including the drag formulation of Shah et al. [312]. This formulation was however only validated for free-falling particles at their terminal velocity and only up to particle Reynolds numbers slightly lower than the most extreme values observed in our simulations. But for quantitatively accurate simulations, it is essential to keep validating the whole hydrodynamic model in general and the drag law in particular. Specifically, it is important to extend the range of applicability of the formulation beyond the current Reynolds numbers and verify its accuracy in instationary flows.

Lift

The importance of lift (hydrodynamic forces perpendicular to the slip velocity) has not been studied in the examples reported in this work. There is currently, to our knowledge, no available formulation for non-Newtonian fluids that can be used beyond very small Reynolds numbers [175]; although some authors have simply used formulations for Newtonian fluids using the local effective viscosity [264, 2]. This matter deserves further research, either to develop a formulation or to provide an estimation of the circumstances where these effects can be neglected or just approximated using the Newtonian formulation.

Backward coupling

Our simulations showed the need to explore the importance of the backward coupling in the PID simulations. Our code should be extended to the non-Newtonian case and a sensitivity study performed to explore this issue.

5.2.6 Backward Coupled CFD-DEM flows

We have barely touched upon the development of FEM-based methods to solve the CFD-DEM fluid-phase equations. Our implementation showed good behaviour and stability in the examples run, but no mathematical proof was provided to back it up. The matter is nontrivial and trying analogous techniques to the ones reported in Codina [77] would not work here, since the resulting stabilised bilinear form Eq.(4.142) is not in fact coercive, property on which is based the stability proof in that work. Our preliminary tests with manufactured solutions (not reported here) did indeed show the right optimally-convergent behaviour, at least in the stationary case, with an arbitrarily-set solid fraction field. In any case, more work

is needed in this area to arrive at solid conclusions on the stability and accuracy properties of the numerical method.

5.2.7 Backward-coupling scheme

Our filtering-based approach has yielded good behaviour that is less prone to instabilities caused by the presence of too-large particles (as compared to the fluid elements) in our experience. However, no systematic work has been done on this matter so far. The examples presented in Section 4.9 showed there are still deficiencies in the overall approach and the coupling scheme is one of the most promising suspects to be responsible for the error. Also, some mesh-dependence was observed and the method should in principle become mesh-independent for large enough filtering radii. The issue of fixing the filtering radii as well as choosing the filtering kernel should therefore be analysed in relation to the problem of mesh-dependence. This question surely provides fecund material for future work.

References

- [1] Agarwal, R. P., Benchohra, M., and Hamani, S. (2010). A survey on existence results for boundary value problems of nonlinear fractional differential equations and inclusions. *Acta Applicandae Mathematicae*, 109(3):973–1033.
- [2] Akhshik, S., Behzad, M., and Rajabi, M. (2015a). CFD–DEM approach to investigate the effect of drill pipe rotation on cuttings transport behavior. *Journal of Petroleum Science and Engineering*, 127:229–244.
- [3] Akhshik, S., Behzad, M., and Rajabi, M. (2015b). CFD–DEM Model for Simulation of Non-spherical Particles in Hole Cleaning Process. *Particulate Science and Technology*, 33(5):472–481.
- [4] Aliseda, A., Cartellier, A., Hainaux, F., and Lasheras, J. C. (2002). Effect of preferential concentration on the settling velocity of heavy particles in homogeneous isotropic turbulence. *Journal of Fluid Mechanics*, 468(October):77–105.
- [5] Aliseda, A. and Lasheras, J. C. (2011). Preferential concentration and rise velocity reduction of bubbles immersed in a homogeneous and isotropic turbulent flow. *Physics of Fluids*, 23(9):093301.
- [6] Anderson, J. D. (2005). Ludwig Prandtl’s Boundary Layer. *Physics Today*, 58(12):42–48.
- [7] Anderson, J. D. J. (1997). *A History of Aerodynamics*. Cambridge University Press, Cambridge.
- [8] Anderson, T. B. and Jackson, R. (1967). Fluid Mechanical Description of Fluidized Beds. Equations of Motion. *Industrial & Engineering Chemistry Fundamentals*, 6(4):527–539.
- [9] ANSYS (2016). ANSYS Fluent Theory Guide–Release 17.0.
- [10] Antypov, D. and Elliott, J. A. (2011). On an analytical solution for the damped Hertzian spring. *EPL (Europhysics Letters)*, 94(5):50004.
- [11] Apte, S., Mahesh, K., Moin, P., and Oefelein, J. (2003). Large-eddy simulation of swirling particle-laden flows in a coaxial-jet combustor. *International Journal of Multiphase Flow*, 29(8):1311–1331.
- [12] Armenio, V. and Fiorotto, V. (2001). The importance of the forces acting on particles in turbulent flows. *Physics of Fluids*, 13(8):2437–2440.
- [13] Auton, T. R. (1987). The lift force on a spherical body in a rotational flow. *Journal of Fluid Mechanics*, 183(-1):199.

- [14] Auton, T. R., Hunt, J. C. R., and Prud'Homme, M. (1988). The force exerted on a body in inviscid unsteady non-uniform rotational flow. *Journal of Fluid Mechanics*, 197(-1):241.
- [15] Ayala, O., Grabowski, W. W., and Wang, L. P. (2007). A hybrid approach for simulating turbulent collisions of hydrodynamically-interacting particles. *Journal of Computational Physics*, 225(1):51–73.
- [16] Azaiez, J. (2008). Bubbles, drops and particles in non-newtonian fluids. R. P. Chhabra. *The Canadian Journal of Chemical Engineering*, 85(2):251–252.
- [17] Babuška, I. and Miller, A. (1984). The post-processing approach in the finite element method—part 1: Calculation of displacements, stresses and other higher derivatives of the displacements. *International Journal for Numerical Methods in Engineering*, 20(6):1085–1109.
- [18] Baffet, D. and Hesthaven, J. S. (2017). A Kernel Compression Scheme for Fractional Differential Equations. *SIAM Journal on Numerical Analysis*, 55(2):496–520.
- [19] Balachandar, S. (2009). A scaling analysis for point-particle approaches to turbulent multiphase flows. *International Journal of Multiphase Flow*, 35(9):801–810.
- [20] Balachandar, S. and Maxey, M. (1989). Methods for evaluating fluid velocities in spectral simulations of turbulence. *Journal of Computational Physics*, 83(1):96–125.
- [21] Barber, R. W. and Emerson, D. R. (2006). Challenges in Modeling Gas-Phase Flow in Microchannels: From Slip to Transition. *Heat Transfer Engineering*, 27(4):3–12.
- [22] Barker, T., Schaeffer, D. G., Bohorquez, P., and Gray, J. M. N. T. (2015). Well-posed and ill-posed behaviour of the μ -rheology for granular flow. *Journal of Fluid Mechanics*, 779:794–818.
- [23] Batchelor, G. K. (1972). Sedimentation in a dilute dispersion of spheres. *Journal of Fluid Mechanics*, 52(02):245.
- [24] Batchelor, G. K. (1982). Sedimentation in a dilute polydisperse system of interacting spheres. Part 1. General theory. *Journal of Fluid Mechanics*, 119(-1):379.
- [25] Bec, J. (2003). Fractal clustering of inertial particles in random flows. *Physics of Fluids*, 15(11):L81–L84.
- [26] Bec, J., Biferale, L., Cencini, M., Lanotte, A., Musacchio, S., and Toschi, F. (2007). Heavy Particle Concentration in Turbulence at Dissipative and Inertial Scales. *Physical Review Letters*, 98(8):084502.
- [27] Bec, J., Biferale, L., Cencini, M., Lanotte, A. S., and Toschi, F. (2006). Effects of vortex filaments on the velocity of tracers and heavy particles in turbulence. *Physics of Fluids*, 18(8):081702.
- [28] Bec, J., Biferale, L., Cencini, M., Lanotte, A. S., and Toschi, F. (2011). Spatial and velocity statistics of inertial particles in turbulent flows. *Journal of Physics: Conference Series*, 333:012003.

- [29] Bec, J., Biferale, L., Lanotte, A. S., Scagliarini, A., and Toschi, F. (2010). Turbulent pair dispersion of inertial particles. *Journal of Fluid Mechanics*, 645:497.
- [30] Bec, J., Celani, A., Cencini, M., and Musacchio, S. (2005). Clustering and collisions of heavy particles in random smooth flows. *Physics of Fluids*, 17(7):073301.
- [31] Beetstra, R., van der Hoef, M. A., and Kuipers, J. A. M. (2007). Drag force of intermediate Reynolds number flow past mono- and bidisperse arrays of spheres. *AIChE Journal*, 53(2):489–501.
- [32] Belhamadia, Y., Fortin, A., and Chamberland, É. (2004). Anisotropic mesh adaptation for the solution of the Stefan problem. *Journal of Computational Physics*, 194(1):233–255.
- [33] Bellani, G. and Variano, E. A. (2012). Slip velocity of large neutrally buoyant particles in turbulent flows. *New Journal of Physics*, 14(12):125009.
- [34] Ben Salem, M. and Oesterle, B. (1998). A shear flow around a spinning sphere: numerical study at moderate reynolds numbers. *International Journal of Multiphase Flow*, 24(4):563–585.
- [35] Benes, K., Tong, P., and Ackerson, B. J. (2007). Sedimentation, Péclet number, and hydrodynamic screening. *Physical Review E*, 76(5):056302.
- [36] Benson, D. A., Meerschaert, M. M., and Revielle, J. (2013). Fractional calculus in hydrologic modeling: A numerical perspective. *Advances in Water Resources*, 51:479–497.
- [37] Berg-Sørensen, K. and Flyvbjerg, H. (2005). The colour of thermal noise in classical Brownian motion: a feasibility study of direct experimental observation. *New Journal of Physics*, 7:38–38.
- [38] Betchen, L. J. and Straatman, A. G. (2009). An accurate gradient and Hessian reconstruction method for cell-centered finite volume discretizations on general unstructured grids. *International Journal for Numerical Methods in Fluids*, 62(9):n/a–n/a.
- [39] Bian, X., Kim, C., and Karniadakis, G. E. (2016). 111 years of Brownian motion. *Soft Matter*, 12(30):6331–6346.
- [40] Bombardelli, F. a., González, A. E., and Niño, Y. I. (2008). Computation of the Particle Basset Force with a Fractional-Derivative Approach. *Journal of Hydraulic Engineering*, 134(10):1513–1520.
- [41] Bosse, T., Kleiser, L., and Meiburg, E. (2006). Small particles in homogeneous turbulence: Settling velocity enhancement by two-way coupling. *Physics of Fluids*, 18(2):027102.
- [42] Boyce, C. M., Holland, D. J., Scott, S. A., and Dennis, J. S. (2015). Limitations on Fluid Grid Sizing for Using Volume-Averaged Fluid Equations in Discrete Element Models of Fluidized Beds. *Industrial & Engineering Chemistry Research*, 54(43):10684–10697.
- [43] Brady, J. (1988). Stokesian Dynamics. *Annual Review of Fluid Mechanics*, 20(1):111–157.

- [44] Bragg, A. D. and Collins, L. R. (2014). New insights from comparing statistical theories for inertial particles in turbulence: I. Spatial distribution of particles. *New Journal of Physics*, 16(5):055013.
- [45] Bragg, A. D., Ireland, P. J., and Collins, L. R. (2015). On the relationship between the non-local clustering mechanism and preferential concentration. *Journal of Fluid Mechanics*, 780:327–343.
- [46] Brandts, J. and Michal, K. (2000). History and future of superconvergence in three-dimensional finite element methods. *Utrecht University Repository (preprint)*, pages 1–10.
- [47] Brennen, C. E. (2005). *Fundamentals of Multiphase Flow*, volume 9780521848. Cambridge University Press, Cambridge.
- [48] Brenner, H. (1996). The Stokes hydrodynamic resistance nonspherical particles. *Chemical Engineering Communications*, 148-150(1):487–562.
- [49] Brinkman, H. C. (1949). A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles. *Flow, Turbulence and Combustion*, 1(1):27.
- [50] Brown, P. P. and Lawler, D. F. (2003). Sphere Drag and Settling Velocity Revisited. *Journal of Environmental Engineering*, 129(3):222–231.
- [51] Brzeziński, D. W. and Ostalczyk, P. (2016). About accuracy increase of fractional order derivative and integral computations by applying the Grünwald–Letnikov formula. *Communications in Nonlinear Science and Numerical Simulation*, 40:151–162.
- [52] Burg, C. and Erwin, T. (2009). Application of Richardson extrapolation to the numerical solution of partial differential equations. *Numerical Methods for Partial Differential Equations*, 25(4):810–832.
- [53] Calzavarini, E., Volk, R., Bourgoïn, M., L ev eque, E., Pinton, J., and Toschi, F. (2009). Acceleration statistics of finite-sized particles in turbulent flow: the role of Fax en forces. *Journal of Fluid Mechanics*, 630(10):179.
- [54] Calzavarini, E., Volk, R., L ev eque, E., Pinton, J.-F., and Toschi, F. (2012). Impact of trailing wake drag on the statistical properties and dynamics of finite-sized particle in turbulence. *Physica D: Nonlinear Phenomena*, 241(3):237–244.
- [55] Candelier, F. (2008). Time-dependent force acting on a particle moving arbitrarily in a rotating flow, at small Reynolds and Taylor numbers. *Journal of Fluid Mechanics*, 608:319–336.
- [56] Candelier, F. and Angilella, J. R. (2006). Analytical investigation of the combined effect of fluid inertia and unsteadiness on low-Re particle centrifugation. *Physical Review E*, 73(4):047301.
- [57] Candelier, F., Angilella, J. R., and Souhar, M. (2004). On the effect of the Boussinesq–Basset force on the radial migration of a Stokes particle in a vortex. *Physics of Fluids*, 16(5):1765–1776.

- [58] Candelier, F., Einarsson, J., Lundell, F., Mehlig, B., and Angilella, J.-R. (2015a). Erratum: Role of inertia for the rotation of a nearly spherical particle in a general linear flow. *Physical Review E*, 92(5):059901.
- [59] Candelier, F., Einarsson, J., Lundell, F., Mehlig, B., and Angilella, J.-R. (2015b). Role of inertia for the rotation of a nearly spherical particle in a general linear flow. *Physical Review E*, 91(5):053023.
- [60] Candelier, F. and Souhar, M. (2007). Time-dependent lift force acting on a particle moving arbitrarily in a pure shear flow, at small Reynolds number. *Physical Review E*, 76(6):067301.
- [61] Cao, J. and Xu, C. (2013). A high order schema for the numerical solution of the fractional ordinary differential equations. *Journal of Computational Physics*, 238:154–168.
- [62] Capecelatro, J. and Desjardins, O. (2013). An Euler–Lagrange strategy for simulating particle-laden flows. *Journal of Computational Physics*, 238:1–31.
- [63] Casas, G., Ferrer, A., and Oñate, E. (2018). Approximating the Basset force by optimizing the method of van Hinsberg et al. *Journal of Computational Physics*, 352:142–171.
- [64] Casas, G., Mukherjee, D., Celigueta, M. A., Zohdi, T. I., and Onate, E. (2017). A modular, partitioned, discrete element framework for industrial grain distribution systems with rotating machinery. *Computational Particle Mechanics*, 4(2):181–198.
- [65] Celigueta, M., Casas, G., Latorre, S., and Arrufat, F. (2018). DEMPack.
- [66] Celigueta, M. A., Deshpande, K. M., Latorre, S., and Oñate, E. (2016a). A FEM-DEM technique for studying the motion of particles in non-Newtonian fluids. Application to the transport of drill cuttings in wellbores. *Computational Particle Mechanics*, 3(2):263–276.
- [67] Celigueta, M. A., Deshpande, K. M., Latorre, S., and Oñate, E. (2016b). A FEM-DEM technique for studying the motion of particles in non-Newtonian fluids. Application to the transport of drill cuttings in wellbores. *Computational Particle Mechanics*, 3(2):263–276.
- [68] Chen, H. and Li, J. (2016). Bubble Collisions in Microchannels Affected by Hydrodynamic Pressures. *Tribology Online*, 11(2):281–287.
- [69] Chen, M. and Deng, W. (2014). Fourth Order Accurate Scheme for the Space Fractional Diffusion Equations. *SIAM Journal on Numerical Analysis*, 52(3):1418–1438.
- [70] Chen, S., Bartello, P., Yau, M. K., Vaillancourt, P. A., and Zwijssen, K. (2016). Cloud Droplet Collisions in Turbulent Environment: Collision Statistics and Parameterization. *Journal of the Atmospheric Sciences*, 73(2):621–636.
- [71] Choi, J.-I., Park, Y., Kwon, O., and Lee, C. (2016). Interparticle collision mechanism in turbulence. *Physical Review E*, 93(1):013112.
- [72] Chong, K., Kelly, S. D., Smith, S., and Eldredge, J. D. (2013). Inertial particle trapping in viscous streaming. *Physics of Fluids*, 25(3):033602.
- [73] Chun, J., Koch, D. L., Rani, S. L., Ahluwalia, A., and Collins, L. R. (2005). Clustering of aerosol particles in isotropic turbulence. *Journal of Fluid Mechanics*, 536(May):219–251.

- [74] Citro, V., Siconolfi, L., Fabre, D., Giannetti, F., and Luchini, P. (2017). Stability and Sensitivity Analysis of the Secondary Instability in the Sphere Wake. *AIAA Journal*, 55(11):3661–3668.
- [75] Clercx, H. J. H. and Schram, P. P. J. M. (1992). Brownian particles in shear flow and harmonic potentials: A study of long-time tails. *Physical Review A*, 46(4):1942–1950.
- [76] Cockburn, B. and Mustapha, K. (2015). A hybridizable discontinuous Galerkin method for fractional diffusion problems. *Numerische Mathematik*, 130(2):293–314.
- [77] Codina, R. (2001). A stabilized finite element method for generalized stationary incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 190(20-21):2681–2706.
- [78] Codina, R. (2002). Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer Methods in Applied Mechanics and Engineering*, 191(39-40):4295–4321.
- [79] Coimbra, C. F. M. and Rangel, R. H. (2001). Spherical particle motion in harmonic stokes flows. *AIAA Journal*, 39(9):2001–1673.
- [80] Coll, A. and Ribó, R. and Pasenau, M. and Escolano, E. and Perez, J. Suit. and Melendo, A. and Monros, A. and Gárate, J. (2016). GiD v.13 Reference Manual. *CIMNE*.
- [81] Correa, C. D., Hero, R., and Kwan-Liu Ma (2011). A Comparison of Gradient Estimation Methods for Volume Rendering on Unstructured Meshes. *IEEE Transactions on Visualization and Computer Graphics*, 17(3):305–319.
- [82] Cotella-Dalmau, J. (2016). *Applications of turbulence modelling in civil engineering*. PhD thesis, Universitat Politècnica de Catalunya.
- [83] Cross, D. M. and Canfield, R. A. (2015). Local continuum shape sensitivity with spatial gradient reconstruction for nonlinear analysis. *Structural and Multidisciplinary Optimization*, 51(4):849–865.
- [84] Crowe, C., Schwarzkopf, J., Sommerfeld, M., and Tsuji, Y. (2005). *Multiphase Flow Handbook*, volume 20052445 of *Mechanical Engineering Series*. CRC Press.
- [85] Crowe, C., Schwarzkopf, J., Sommerfeld, M., and Tsuji, Y. (2012). *Multiphase Flows with Droplets and Particles*. CRC press.
- [86] Crowe, C. T. (2000). On models for turbulence modulation in fluid–particle flows. *International Journal of Multiphase Flow*, 26(5):719–727.
- [87] Cui, M., Zhai, Y.-h., and Ji, G.-d. (2011). Experimental study of rock breaking effect of steel particles. *Journal of Hydrodynamics, Ser. B*, 23(2):241–246.
- [88] Dadvand, P., Rossi, R., and Oñate, E. (2010). An Object-oriented Environment for Developing Finite Element Codes for Multi-disciplinary Applications. *Archives of Computational Methods in Engineering*, 17(3):253–297.
- [89] Daitche, A. (2013). Advection of inertial particles in the presence of the history force: Higher order numerical schemes. *Journal of Computational Physics*, 254:93–106.

- [90] Daitche, A. (2015). On the role of the history force for inertial particles in turbulence. *Journal of Fluid Mechanics*, 782:567–593.
- [91] Dassios, G. (2012). *Ellipsoidal Harmonics*. Cambridge University Press, Cambridge.
- [92] Davis, R. H. and Acrivos, A. (1985). Sedimentation of Noncolloidal Particles at Low Reynolds Numbers. *Annual Review of Fluid Mechanics*, 17(1):91–118.
- [93] Davis, R. H. and Gecol, H. (1994). Hindered settling function with no empirical parameters for polydisperse suspensions. *AIChE Journal*, 40(3):570–575.
- [94] de Oliveira, E. C. and Tenreiro Machado, J. A. (2014). A Review of Definitions for Fractional Derivatives and Integral. *Mathematical Problems in Engineering*, 2014(1940):1–6.
- [95] Deen, N., Van Sint Annaland, M., Van der Hoef, M., and Kuipers, J. (2007). Review of discrete particle modeling of fluidized beds. *Chemical Engineering Science*, 62(1-2):28–44.
- [96] Deng, W. (2007). Short memory principle and a predictor–corrector approach for fractional differential equations. *Journal of Computational and Applied Mathematics*, 206(1):174–188.
- [97] Derksen, J. J. and Sundaresan, S. (2007). Direct numerical simulations of dense suspensions: wave instabilities in liquid-fluidized beds. *Journal of Fluid Mechanics*, 587:303–336.
- [98] Di Benedetto, A., Russo, P., Sanchirico, R., and Di Sarli, V. (2013). CFD simulations of turbulent fluid flow and dust dispersion in the 20 liter explosion vessel. *AIChE Journal*, 59(7):2485–2496.
- [99] Di Felice, R. and Rotondi, M. (2012a). Fluid-particle Drag Force in Binary-solid Suspensions. *International Journal of Chemical Reactor Engineering*, 10(1):1–15.
- [100] Di Felice, R. and Rotondi, M. (2012b). The settling velocity of a single sphere in viscous fluid: The effect of neighboring larger spheres. *Powder Technology*, 217:486–488.
- [101] Diamant, H. (2009). Hydrodynamic Interaction in Confined Geometries. *Journal of the Physical Society of Japan*, 78(4):041002.
- [102] Diethelm, K. (2010). *The Analysis of Fractional Differential Equations*, volume 2004 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [103] Diethelm, K., Ford, N., Freed, A., and Luchko, Y. (2005). Algorithms for the fractional calculus: A selection of numerical methods. *Computer Methods in Applied Mechanics and Engineering*, 194(6-8):743–773.
- [104] Diethelm, K., Ford, N. J., and Freed, A. D. (2004). Detailed Error Analysis for a Fractional Adams Method. *Numerical Algorithms*, 36(1):31–52.
- [105] Díez Rodríguez, D. (2017). *Gradient Based Porosity Calculation in Casting Simulation*. PhD thesis, Universitat Politècnica de Catalunya.

- [106] Donea, J., Huerta, A., Ponthot, J., and Rodríguez-Ferran, A. (2004). Arbitrary Lagrangian-Eulerian Methods. In *Encyclopedia of Computational Mechanics*, number 1969, pages 1–38. John Wiley & Sons, Ltd, Chichester, UK.
- [107] Doostmohammadi, A. and Ardekani, A. M. (2013). Interaction between a pair of particles settling in a stratified fluid. *Physical Review E*, 88(2):023029.
- [108] Dorgan, A. and Loth, E. (2007). Efficient calculation of the history force at finite Reynolds numbers. *International Journal of Multiphase Flow*, 33(8):833–848.
- [109] Drew, D. A. and Passman, S. L. (1999). *Theory of Multicomponent Fluids*, volume 135 of *Applied Mathematical Sciences*. Springer New York, New York, NY.
- [110] Dunatunga, S. and Kamrin, K. (2014). Continuum modelling and simulation of granular flows through their many phases. (Harlow 1964).
- [111] Dusenbery, D. B. (2009). *Living at micro scale: the unexpected physics of being small*. Harvard University Press.
- [112] Eaton, J. and Fessler, J. (1994). Preferential concentration of particles by turbulence. *International Journal of Multiphase Flow*, 20(94):169–209.
- [113] Eaton, J. K. (2009). Two-way coupled turbulence simulations of gas-particle flows using point-particle tracking. *International Journal of Multiphase Flow*, 35(9):792–800.
- [114] Einarsson, J., Candelier, F., Lundell, F., Angilella, J. R., and Mehlig, B. (2015). Rotation of a spheroid in a simple shear at small Reynolds number. *Physics of Fluids*, 27(6):063301.
- [115] Elcner, J., Jedelsky, J., Lizal, F., and Jicha, M. (2013). Velocity profiles in idealized model of human respiratory tract. *EPJ Web of Conferences*, 45:01025.
- [116] Elghobashi, S. (1994). On predicting particle-laden turbulent flows. *Applied Scientific Research*, 52(4):309–329.
- [117] Elghobashi, S. and Truesdell, G. C. (1992). Direct simulation of particle dispersion in a decaying isotropic turbulence. *Journal of Fluid Mechanics*, 242(-1):655.
- [118] Elghobashi, S. (2007). An Updated Classification Map of Particle-Laden Turbulent Flows. In *IUTAM Symposium on Computational Approaches to Multiphase Flow*, pages 3–10. Springer Netherlands, Dordrecht.
- [119] Ern, A. and Guermond, J.-L. (2004). *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer New York, New York, NY.
- [120] Estep, D. J. (2004). *A Short Course on Duality , Adjoint Operators , Green ' s Functions , and A Posteriori Error Analysis*.
- [121] Ethier, C. R. and Steinman, D. A. (1994). Exact fully 3D Navier-Stokes solutions for benchmarking. *International Journal for Numerical Methods in Fluids*, 19(5):369–375.
- [122] Falkovich, G., Fouxon, A., and Stepanov, M. G. (2002). Acceleration of rain initiation by cloud turbulence. *Nature*, 419(6903):151–154.

- [123] Falkovich, G. and Pumir, A. (2007). Sling Effect in Collisions of Water Droplets in Turbulent Clouds. *Journal of the Atmospheric Sciences*, 64(12):4497–4505.
- [124] Farazmand, M. and Haller, G. (2015). The Maxey–Riley equation: Existence, uniqueness and regularity of solutions. *Nonlinear Analysis: Real World Applications*, 22(3):98–106.
- [125] Feng, Y. T. and Owen, D. R. J. (2014). Discrete element modelling of large scale particle systems—I: exact scaling laws. *Computational Particle Mechanics*, 1(2):159–168.
- [126] Ferrante, A. and Elghobashi, S. (2003). On the physical mechanisms of two-way coupling in particle-laden isotropic turbulence. *Physics of Fluids*, 15(2):315–329.
- [127] Fessler, J. R., Kulick, J. D., and Eaton, J. K. (1994). Preferential concentration of heavy particles in a turbulent channel flow. *Physics of Fluids*, 6(11):3742–3749.
- [128] Feuillebois, F. and Lasek, A. (1978). On the rotational historic term in non-stationary stokes flow. *The Quarterly Journal of Mechanics and Applied Mathematics*, 31(4):435–443.
- [129] Flannery, Brian P and Teukolsky, Saul A and Press, William H and Vetterling, W. T. (1988). *Numerical recipes in C: The art of scientific computing*, volume 2.
- [130] Ford, N. J. and Simpson, A. C. (2001). No Title. *Numerical Algorithms*, 26(4):333–346.
- [131] Fouxon, I., Park, Y., Harduf, R., and Lee, C. (2015). Inhomogeneous distribution of water droplets in cloud turbulence. *Physical Review E*, 92(3):033001.
- [132] Fullmer, W. D. and Hrenya, C. M. (2016). Quantitative assessment of fine-grid kinetic-theory-based predictions of mean-slip in unbounded fluidization. *AIChE Journal*, 62(1):11–17.
- [133] Gad-el Hak, M. (1999). The Fluid Mechanics of Microdevices—The Freeman Scholar Lecture. *Journal of Fluids Engineering*, 121(1):5.
- [134] Gad-el Hak, M. (2005). Liquids: The holy grail of microfluidic modeling. *Physics of Fluids*, 17(10):100612.
- [135] Gad-El-Hak, M. (2006). Gas and Liquid Transport at the Microscale. *Heat Transfer Engineering*, 27(4):13–29.
- [136] Galeone, L. and Garrappa, R. (2009). Explicit methods for fractional differential equations and their stability properties. *Journal of Computational and Applied Mathematics*, 228(2):548–560.
- [137] Gao, G.-H., Sun, H.-W., and Sun, Z.-Z. (2015). Stability and convergence of finite difference schemes for a class of time-fractional sub-diffusion equations based on certain superconvergence. *Journal of Computational Physics*, 280:510–528.
- [138] Gao, H., Li, H., and Wang, L.-p. (2013). Lattice Boltzmann simulation of turbulent flow laden with finite-size particles. *Computers & Mathematics with Applications*, 65(2):194–210.

- [139] Gao, J., Lan, X., Fan, Y., Chang, J., Wang, G., Lu, C., and Xu, C. (2009). CFD modeling and validation of the turbulent fluidized bed of FCC particles. *AIChE Journal*, 55(7):1680–1694.
- [140] Garg, R., Narayanan, C., Lakehal, D., and Subramaniam, S. (2007). Accurate numerical estimation of interphase momentum transfer in Lagrangian–Eulerian simulations of dispersed two-phase flows. *International Journal of Multiphase Flow*, 33(12):1337–1364.
- [141] Garrappa, R. (2009). On some explicit Adams multistep methods for fractional differential equations. *Journal of Computational and Applied Mathematics*, 229(2):392–399.
- [142] Gavze, E. (1990). The accelerated motion of rigid bodies in non-steady stokes flow. *International Journal of Multiphase Flow*, 16(1):153–166.
- [143] Geldart, D. (1973). Types of gas fluidization. *Powder Technology*, 7(5):285–292.
- [144] Giacomelli, R. (1930). The Aerodynamics of Leonardo Da Vinci. *Journal of the Royal Aeronautical Society*, 34(240):1016–1038.
- [145] Gibert, M., Xu, H., and Bodenschatz, E. (2012). Where do small, weakly inertial particles go in a turbulent flow? *Journal of Fluid Mechanics*, 698(May 2014):160–167.
- [146] Gmeiner, B., Råde, U., Stengel, H., Waluga, C., and Wohlmuth, B. (2015). Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems. *SIAM Journal on Scientific Computing*, 37(2):C143–C168.
- [147] Gong, C., Bao, W., Tang, G., Jiang, Y., and Liu, J. (2015). Computational Challenge of Fractional Differential Equations and the Potential Solutions: A Survey. *Mathematical Problems in Engineering*, 2015:1–13.
- [148] González, A., Bombardelli, F., and Niño, Y. (2006). Improving the prediction capability of numerical models for particle motion in water bodies. In: *Proceedings of the 7th International Conference on HydroScience and Engineering ICHE 2006, Philadelphia, USA, 2006*(Iche).
- [149] Goto, S. and Vassilicos, J. C. (2008). Sweep-stick mechanism of heavy particle clustering in fluid turbulence. *Physical Review Letters*, 100(5):1–4.
- [150] Gotoh, T., Suehiro, T., and Saito, I. (2016). Continuous growth of cloud droplets in cumulus cloud. *New Journal of Physics*, 18(4):043042.
- [151] Greenshields, C. J. (2015). OpenFOAM–The Open Source CFD Toolbox. Programmer’s Guide.
- [152] Guazzelli, E., Morris, J. F., and Pic, S. (2011). *A Physical Introduction to Suspension Dynamics*. Cambridge University Press, Cambridge.
- [153] Guo, H., Zhang, Z., and Zhao, R. (2016). Hessian recovery for finite element methods. *Mathematics of Computation*, 86(306):1671–1692.
- [154] Guo, Y. and Curtis, J. S. (2015). Discrete Element Method Simulations for Complex Granular Flows. *Annual Review of Fluid Mechanics*, 47(1):21–46.

- [155] Guseva, K., Feudel, U., and Tél, T. (2013). Influence of the history force on inertial particle advection: Gravitational effects and horizontal diffusion. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 88(4):1–11.
- [156] Gus'kov, O. (2017). On the virtual mass of a rough sphere. *Journal of Applied Mathematics and Mechanics*, 81(4).
- [157] Gustavsson, K. and Mehlig, B. (2016). Statistical models for spatial patterns of heavy particles in turbulence. *Advances in Physics*, 65(1):1–57.
- [158] Habib, M., Nemitallah, M., and El-Nakla, M. (2014). Current status of CHF predictions using CFD modeling technique and review of other techniques especially for non-uniform axial and circumferential heating profiles. *Annals of Nuclear Energy*, 70:188–207.
- [159] Hadinoto, K. and Chew, J. W. (2010). Modeling fluid–particle interaction in dilute-phase turbulent liquid–particle flow simulation. *Particuology*, 8(2):150–160.
- [160] Hammond, A. P. and Corwin, E. I. (2017). Direct measurement of the ballistic motion of a freely floating colloid in Newtonian and viscoelastic fluids. *Physical Review E*, 96(4):042606.
- [161] Hansen, J., McDonald, I. R., and Henderson, D. (1988). *Theory of Simple Liquids*, volume 41.
- [162] Hawken, D. M., Townsend, P., and Webster, M. F. (1991). A comparison of gradient recovery methods in finite-element calculations. *Communications in Applied Numerical Methods*, 7(3):195–204.
- [163] Henann, D. L. and Kamrin, K. (2013). A predictive, size-dependent continuum model for dense granular flows. *Proceedings of the National Academy of Sciences*, 110(17):6730–6735.
- [164] Herron, I. H., Davis, S. H., and Bretherton, F. P. (1975). On the sedimentation of a sphere in a centrifuge. *Journal of Fluid Mechanics*, 68(02):209.
- [165] Hetsroni, G. (1989). Particles-turbulence interaction. *International Journal of Multiphase Flow*, 15(5):735–746.
- [166] Hill, R. and Power, G. (1956). Extremum principles for slow viscous flow and the approximate calculation of drag. *The Quarterly Journal of Mechanics and Applied Mathematics*, 9(3):313–319.
- [167] Hill, R. J., Koch, D. L., and Ladd, A. J. C. (2001). The first effects of fluid inertia on flows in ordered and random arrays of spheres. *Journal of Fluid Mechanics*, 448:213–241.
- [168] Hirsch, C. (2007). *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics*. Butterworth-Heinemann, Berlin, Heidelberg, 2nd editio edition.
- [169] Hocking, L. M. (1959). The collision efficiency of small drops. *Quarterly Journal of the Royal Meteorological Society*, 85(363):44–50.

- [170] Holland, D. J., Müller, C. R., Dennis, J. S., Gladden, L. F., and Sederman, A. J. (2008). Spatially resolved measurement of anisotropic granular temperature in gas-fluidized beds. *Powder Technology*, 182(2):171–181.
- [171] Homann, H. and Bec, J. (2009). Finite-size effects in the dynamics of neutrally buoyant particles in turbulent flow. *Journal of Fluid Mechanics*, 651:81.
- [172] Huang, R., Chavez, I., Taute, K. M., Lukić, B., Jeney, S., Raizen, M. G., and Florin, E.-L. (2011). Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid. *Nature Physics*, 7(7):576–580.
- [173] Huck, P. D., Bateson, C., Volk, R., Cartellier, A., Bourgoin, M., and Aliseda, A. (2018). The role of collective effects on settling velocity enhancement for inertial particles in turbulence. *Journal of Fluid Mechanics*, 846:1059–1075.
- [174] Hughes, T. J. (1995). Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387–401.
- [175] Ignatenko, Y., Bocharov, O., and May, R. (2017). Lift and Drag Forces for a Sphere on a Flat Wall in Non-Newtonian Shear Flow. *EPJ Web of Conferences*, 159:00014.
- [176] Ingólfsson, H. I., Lopez, C. A., Uusitalo, J. J., de Jong, D. H., Gopal, S. M., Periole, X., and Marrink, S. J. (2014). The power of coarse graining in biomolecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3):225–248.
- [177] Irazábal, J., Salazar, F., and Oñate, E. (2017). Numerical modelling of granular materials with spherical discrete particles and the bounded rolling friction model. Application to railway ballast. *Computers and Geotechnics*, 85:220–229.
- [178] Ireland, P. J., Bragg, A. D., and Collins, L. R. (2016). The effect of Reynolds number on inertial particle dynamics in isotropic turbulence. Part 1. Simulations without gravitational effects. *Journal of Fluid Mechanics*, 796:617–658.
- [179] Ishii, M. and Hibiki, T. (2011). *Thermo-Fluid Dynamics of Two-Phase Flow*. Springer New York, New York, NY.
- [180] Ishii, Mamoru and Hibiki, T. (2010). *Thermo-fluid dynamics of two-phase flow*. Ishii, Mamoru and Hibiki, Takashi.
- [181] Jabin, P.-E. and Otto, F. (2004). Identification of the Dilute Regime in Particle Sedimentation. *Communications in Mathematical Physics*, 250(2):415–432.
- [182] Jackson, R. (2001). The Dynamics of Fluidized Particles. *Measurement Science and Technology*, 12(6):755–755.
- [183] James, M. and Ray, S. S. (2016). How Violent are the Collisions of Different Sized Droplets in a Turbulent Flow? pages 1–11.
- [184] Jeong, W. and Seong, J. (2014). Comparison of effects on technical variances of computational fluid dynamics (CFD) software based on finite element and finite volume methods. *International Journal of Mechanical Sciences*, 78:19–26.

- [185] Jian, Z., Yiji, X., Jianhua, R., and Deju, H. (2014). Numerical simulation of the bottom hole flow field of particle impact drilling. 8(2):18–23.
- [186] Johnson, A. A. and Tezduyar, T. E. (1999). Advanced mesh generation and update methods for 3D flow simulations. *Computational Mechanics*, 23(2):130–143.
- [187] Johnson, T. A and Patel, V. C. (1999). Flow past a sphere up to a Reynolds number of 300. *Journal of Fluid Mechanics*, 378:19—70.
- [188] Joly, A., Moulin, F., Violeau, D., and Astruc, D. (2012). Diffusion in grid turbulence of isotropic macro-particles using a Lagrangian stochastic method: Theory and validation. *Physics of Fluids*, 24(10):1–25.
- [189] Kaneda, Y. (1986). The drag on a sparse random array of fixed spheres in flow at small but finite Reynolds number. *Journal of Fluid Mechanics*, 167(-1):455.
- [190] Kantha, L. and Hocking, W. (2011). Dissipation rates of turbulence kinetic energy in the free atmosphere: MST radar and radiosondes. *Journal of Atmospheric and Solar-Terrestrial Physics*, 73(9):1043–1051.
- [191] Karaa, Samir and Mustapha, Kassem and Pani, A. K. (2016). Finite volume element method for two-dimensional fractional subdiffusion problems. *IMA Journal of Numerical Analysis*, 37(2):945—964.
- [192] Karanjkar, P. U., Coolman, R. J., Huber, G. W., Blatnik, M. T., Almalkie, S., de Bruyn Kops, S. M., Mountziaris, T. J., and Conner, W. C. (2014). Production of aromatics by catalytic fast pyrolysis of cellulose in a bubbling fluidized bed reactor. *AIChE Journal*, 60(4):1320–1335.
- [193] Karniadakis, G., Beskok, A., and Narayan, A. (2005). *Microflows and Nanoflows*, volume 29 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York.
- [194] Kendoush, A. A., Sulaymon, A. H., and Mohammed, S. A. (2007). Experimental evaluation of the virtual mass of two solid spheres accelerating in fluids. *Experimental Thermal and Fluid Science*, 31(7):813–823.
- [195] Kindlmann, G., Whitaker, R., Tasdizen, T., and Moller, T. (2003). Curvature-based transfer functions for direct volume rendering: methods and applications. In *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, volume d, pages 513–520. IEEE.
- [196] Kolev, N. I. (2012). *Multiphase Flow Dynamics 5*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [197] Krantz, W. B. (2007). *Scaling Analysis in Modeling Transport and Reaction Processes*. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- [198] Kremer, G. M. (2010). *An Introduction to the Boltzmann Equation and Transport Processes in Gases*, volume 53 of *Interaction of Mechanics and Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

- [199] Kruis, F. E. and Kusters, K. A. (1997). The Collision Rate Of Particles In Turbulent Flow. *Chemical Engineering Communications*, 158(1):201–230.
- [200] Labra, C., Ooi, J. Y., and Sun, J. (2013). Spatial and temporal coarse-graining for DEM analysis. In *AIP Conference Proceedings*, volume 1542, pages 1258–1261.
- [201] Lakshmikantham, V. and Vatsala, A. (2008). Basic theory of fractional differential equations. *Nonlinear Analysis: Theory, Methods & Applications*, 69(8):2677–2682.
- [202] Landau, L. D., Lifshitz, E. M. (1989). *Course of Theoretical Physics Volume 6, Fluid mechanics*, volume 1. Butterworth-Heinemann.
- [203] Langlois, G. P., Farazmand, M., and Haller, G. (2015). Asymptotic Dynamics of Inertial Particles with Memory. *Journal of Nonlinear Science*, 25(6):1225–1255.
- [204] Li, C. and Chen, A. (2017). Numerical methods for fractional partial differential equations. *International Journal of Computer Mathematics*, 34(1):1–52.
- [205] Li, C. and Zeng, F. (2013). The Finite Difference Methods for Fractional Ordinary Differential Equations. *Numerical Functional Analysis and Optimization*, 34(2):149–179.
- [206] Li, J., Wang, H., Liu, Z., Chen, S., and Zheng, C. (2012). An experimental study on turbulence modification in the near-wall boundary layer of a dilute gas-particle channel flow. *Experiments in Fluids*, 53(5):1385–1403.
- [207] Li, J.-R. (2010). A Fast Time Stepping Method for Evaluating Fractional Integrals. *SIAM Journal on Scientific Computing*, 31(6):4696–4714.
- [208] Li, T., Wang, L., Rogers, W., Zhou, G., and Ge, W. (2017). An approach for drag correction based on the local heterogeneity for gas-solid flows. *AIChE Journal*, 63(4):1203–1212.
- [209] License, I. and Roy, J. (2002). University of Chester Digital Repository. *European Physical Education Review*, 8(2):157–175.
- [210] Lifshitz, J. and Kolsky, H. (1964). Some experiments on anelastic rebound. *Journal of the Mechanics and Physics of Solids*, 12(1):35–43.
- [211] Ling, Y., Parmar, M., and Balachandar, S. (2013). A scaling analysis of added-mass and history forces and their coupling in dispersed multiphase flows. *International Journal of Multiphase Flow*, 57:102–114.
- [212] Ling, Wei and Chung, JN and Troutt, TR and Crowe, C. (1998). Direct numerical simulation of a three-dimensional temporal mixing layer with particle dispersion. *Journal of Fluid Mechanics*, 358:61—85.
- [213] Liu, Y., Matida, E. A., Gu, J., and Johnson, M. R. (2007). Numerical simulation of aerosol deposition in a 3-D human nasal cavity using RANS, RANS/EIM, and LES. *Journal of Aerosol Science*, 38(7):683–700.
- [214] Löhner, R. (2008). *Applied Computational Fluid Dynamics Techniques*. John Wiley & Sons, Ltd, Chichester, UK.

- [215] Löhner, R., Camelli, F., Baum, J. D., Togashi, F., and Soto, O. (2014). On mesh-particle techniques. *Computational Particle Mechanics*, 1(2):199–209.
- [216] Loth, E. (2000). Numerical approaches for motion of dispersed particles, droplets and bubbles. *Progress in Energy and Combustion Science*, 26(3):161–223.
- [217] Loth, E. (2008a). Compressibility and rarefaction effects on drag of a spherical particle. *Aiaa Journal*, 46(9):2219–2228.
- [218] Loth, E. (2008b). Drag of non-spherical solid particles of regular and irregular shape. *Powder Technology*, 182(3):342–353.
- [219] Loth, E. (2008c). Lift of a Spherical Particle Subject to Vorticity and/or Spin. *AIAA Journal*, 46(4):801–809.
- [220] Loth, E. (2010). *Particles, Drops and Bubbles: Fluid Dynamics and Numerical Methods*. Cambridge University Press.
- [221] Loth, E. (2011). A Discrete Lagrangian Particle Equation of Motion for Significant Reynolds Numbers and Diameters. In *6th AIAA Theoretical Fluid Mechanics Conference*, number June, pages 1–32, Reston, Virginia. American Institute of Aeronautics and Astronautics.
- [222] Loth, E. and Dorgan, A. J. (2009). An equation of motion for particles of finite Reynolds number and size. *Environmental Fluid Mechanics*, 9(2):187–206.
- [223] Lovalenti, P. M. and Brady, J. F. (1993). The hydrodynamic force on a rigid particle undergoing arbitrary time-dependent motion at small Reynolds number. *Journal of Fluid Mechanics*, 256(-1):561.
- [224] Lovalenti, P. M. and Brady, J. F. (1995). The temporal behaviour of the hydrodynamic force on a body in response to an abrupt change in velocity at small but finite Reynolds number. *Journal of Fluid Mechanics*, 293(-1):35.
- [225] Lu, L., Gopalan, B., and Benyahia, S. (2017). Assessment of Different Discrete Particle Methods Ability To Predict Gas-Particle Flow in a Small-Scale Fluidized Bed. *Industrial & Engineering Chemistry Research*, 56(27):7865–7876.
- [226] Lubich, C. (1986). Discretized Fractional Calculus. *SIAM Journal on Mathematical Analysis*, 17(3):704–719.
- [227] MacDonald, C. L., Bhattacharya, N., Sprouse, B. P., and Silva, G. a. (2015). Efficient computation of the Grünwald–Letnikov fractional diffusion derivative using adaptive time step memory. *Journal of Computational Physics*, 297:221–236.
- [228] Macías, D., Rodríguez-Santana, Á., Ramírez-Romero, E., Bruno, M., Pelegrí, J. L., Sangrà, P., Aguiar-González, B., and García, C. M. (2013). Turbulence as a driver for vertical plankton distribution in the subsurface upper ocean. *Scientia Marina*, 77(4):541–549.
- [229] Magnus, G. (1853). Ueber die Abweichung der Geschosse, und: Ueber eine auffallende Erscheinung bei rotirenden Körpern. *Annalen der Physik und Chemie*, 164(1):1–29.

- [230] Mainardi, F. (2010). *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*.
- [231] Malkiel, E., Abras, J. N., Widder, E. A., and Katz, J. (2006). On the spatial distribution and nearest neighbor distance between particles in the water column determined from in situ holographic measurements. *Journal of Plankton Research*, 28(2):149–170.
- [232] Mandø, M., Lightstone, M. F., Rosendahl, L., Yin, C., and Sørensen, H. (2009). Turbulence modulation in dilute particle-laden flow. *International Journal of Heat and Fluid Flow*, 30(2):331–338.
- [233] Marchioli, C., Picciotto, M., and Soldati, A. (2007). Influence of gravity and lift on particle velocity statistics and transfer rates in turbulent vertical channel flow. *International Journal of Multiphase Flow*, 33(3):227–251.
- [234] Marchioli, C. and Soldati, A. (2002). Mechanisms for particle transfer and segregation in a turbulent boundary layer. *Journal of Fluid Mechanics*, 468:283–315.
- [235] Marchioli, C., Soldati, A., Kuerten, J., Arcen, B., Tanière, A., Goldensoph, G., Squires, K., Cargnelutti, M., and Portela, L. (2008). Statistics of particle dispersion in direct numerical simulations of wall-bounded turbulence: Results of an international collaborative benchmark test. *International Journal of Multiphase Flow*, 34(9):879–893.
- [236] Marchioli, C., Soldati, A., Salvetti, M. V., Kuerten, J. G. M., Konan, A., Fede, P., Simonin, O., Squires, K. D., Gobert, C., Manhart, M., Jaszczur, M., and Portela, L. M. (2011). Benchmark test on particle-laden channel flow with point-particle LES. In *ERCOFTAC Series*, volume 15, pages 177–182. Springer Netherlands.
- [237] Mavriplis, D. (2003). Revisiting the Least-Squares Procedure for Gradient Reconstruction on Unstructured Meshes. In *16th AIAA Computational Fluid Dynamics Conference*, number 2003, pages NASA CR–2003–212683, Reston, Virginia. American Institute of Aeronautics and Astronautics.
- [238] Maxey, M. R. (1983). Equation of motion for a small rigid sphere in a nonuniform flow. *Physics of Fluids*, 26(4):883–889.
- [239] Maxey, M. R. (1987). The motion of small spherical particles in a cellular flow field. *Physics of Fluids*, 30(7):1915–1928.
- [240] Maxey, M. R. (1993). The equation of motion for a small rigid sphere in a nonuniform or unsteady flow. *ASME-PUBLICATIONS-FED*, 166:57.
- [241] Maxey, M. R., Patel, B. K., Chang, E. J., and Wang, L.-P. (1997). Simulations of dispersed turbulent multiphase flow. *Fluid Dynamics Research*, 20(1-6):143–156.
- [242] Mazzitelli, I. M., Lohse, D., and Toschi, F. (2003). On the relevance of the lift force in bubbly turbulence. *Journal of Fluid Mechanics*, 488(488):283–313.
- [243] Mei, R. and Adrian, R. J. (1992). Flow past a sphere with an oscillation in the free-stream velocity and unsteady drag at finite Reynolds number. *Journal of Fluid Mechanics*, 237(November):323–341.

- [244] Melendo, A. and Coll, A. and Pasenau, M. and Escolano, E. and Monros, A. (2018). GiD.
- [245] Mendez, P. F. (2010). Characteristic Values in the Scaling of Differential Equations in Engineering. *Journal of Applied Mechanics*, 77(6):061017.
- [246] Metzler, R. and Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77.
- [247] Metzner, A. B. and Reed, J. C. (1955). Flow of non-newtonian fluids—correlation of the laminar, transition, and turbulent-flow regions. *AIChE Journal*, 1(4):434–440.
- [248] Michaelides, E. E. (1997). Review—The Transient Equation of Motion for Particles, Bubbles, and Droplets. *Journal of Fluids Engineering*, 119(2):233—247.
- [249] Michaelides, E. E. (2015). Brownian movement and thermophoresis of nanoparticles in liquids. *International Journal of Heat and Mass Transfer*, 81:179–187.
- [250] Mitsoulis, E. (2004). On creeping drag flow of a viscoplastic fluid past a circular cylinder: wall effects. *Chemical Engineering Science*, 59(4):789–800.
- [251] Moin, P. and Verzicco, R. (2016). On the suitability of second-order accurate discretizations for turbulent flow simulations. *European Journal of Mechanics - B/Fluids*, 55:242–245.
- [252] Moller, T., Mueller, K., Kurzion, Y., Machiraju, R., and Yagel, R. (1998). Design of accurate and smooth filters for function and derivative reconstruction. In *IEEE Symposium on Volume Visualization (Cat. No.989EX300)*, pages 143–151,. IEEE.
- [253] Monchaux, R., Bourgoin, M., and Cartellier, A. (2012). Analyzing preferential concentration and clustering of inertial particles in turbulence. *International Journal of Multiphase Flow*, 40:1–18.
- [254] Moreno-Casas, P. A. and Bombardelli, F. A. (2016). Computation of the Basset force: recent advances and environmental flow applications. *Environmental Fluid Mechanics*, 16(1):193–208.
- [255] Moukalled, F., Mangani, L., and Darwish, M. (2016). *The Finite Volume Method in Computational Fluid Dynamics*, volume 113 of *Fluid Mechanics and Its Applications*. Springer International Publishing, Cham.
- [256] Mukin, R. and Zaichik, L. (2012). Nonlinear algebraic Reynolds stress model for two-phase turbulent flows laden with small heavy particles. *International Journal of Heat and Fluid Flow*, 33(1):81–91.
- [257] Narayanan, C., Lakehal, D., Botto, L., and Soldati, A. (2003). Mechanisms of particle deposition in a fully developed turbulent open channel flow. *Physics of Fluids*, 15(3):763–775.
- [258] Nasato, D. S., Goniva, C., Pirker, S., and Kloss, C. (2015). Coarse Graining for Large-scale DEM Simulations of Particle Flow – An Investigation on Contact and Cohesion Models. *Procedia Engineering*, 102:1484–1490.

- [259] Newton, I. (1671). A letter of mr. isaac newton, professor of the mathematicks in the university of cambridge; containing his new theory about light and colors: sent by the author to the publisher from cambridge. *Philosophical Transactions of the Royal Society of London*, 6(69-80):3075–3087.
- [260] Newton, Sir Isaac, Cohen, I Bernard and Whitman, Anne and Budenz, J. (2016). *the Principia: the authoritative translation and guide: mathematical principles of natural philosophy*. University of California Press.
- [261] Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer New York.
- [262] Obligado, M., Missaoui, M., Monchaux, R., Cartellier, A., and Bourgoïn, M. (2011). Reynolds number influence on preferential concentration of heavy particles in turbulent flows. *Journal of Physics: Conference Series*, 318(5):052015.
- [263] Oden, J. T. and Brauchli, H. J. (1971). On the calculation of consistent stress distributions in finite element approximations. *International Journal for Numerical Methods in Engineering*, 3(3):317–325.
- [264] Ofei, T. N., Irawan, S., and Pao, W. (2014). CFD Method for Predicting Annular Pressure Losses and Cuttings Concentration in Eccentric Horizontal Wells. *Journal of Petroleum Engineering*, 2014:1–16.
- [265] Oldham, K. B. and Spanier, J. (2006). *The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order (Dover Books on Mathematics)*. Dover Publications, Inc.
- [266] Olivieri, S., Picano, F., Sardina, G., Iudicone, D., and Brandt, L. (2014). The effect of the Basset history force on particle clustering in homogeneous and isotropic turbulence. *Physics of Fluids*, 26(4):041704.
- [267] Olver, F. W. J. (2010). *NIST Handbook of Mathematical Functions*.
- [268] Oñate, E., Celigueta, M. A., Latorre, S., Casas, G., Rossi, R., and Rojek, J. (2014). Lagrangian analysis of multiscale particulate flows with the particle finite element method. *Computational Particle Mechanics*, 1(1):85–102.
- [269] Ortega, E. (2014). *Development and applications of the Finite Point Method to compressible aerodynamics problems*. PhD thesis.
- [270] Ortigueira, M. D. and Tenreiro Machado, J. (2015). What is a fractional derivative? *Journal of Computational Physics*, 293:4–13.
- [271] Otto, H., Kerst, K., Roloff, C., Janiga, G., and Katterfeld, A. (2018). behavior of lunar regolith JSC-1A. *Particuology*, pages 1–10.
- [272] Padding, J. T. and Louis, A. A. (2006). Hydrodynamic interactions and Brownian forces in colloidal suspensions: Coarse-graining over time and length scales. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 74(3):1–31.

- [273] Pan, T.-w., Joseph, D. D., Bai, R., Glowinski, R., and Sarin, V. (2002). Fluidization of 1204 spheres: simulation and experiment. *Journal of Fluid Mechanics*, 451:1–28.
- [274] Patel, V. M., Maleh, R., Gilbert, A. C., and Chellappa, R. (2012). Gradient-Based Image Recovery Methods From Incomplete Fourier Measurements. *IEEE Transactions on Image Processing*, 21(1):94–105.
- [275] Pécseli, H. L., Trulsén, J., and Weiland, J. (2009). Predator-prey Encounter Rates in Turbulent Environments: Consequences of Inertia Effects and Finite Sizes. In *AIP Conference Proceedings*, volume 1177, pages 85–95. AIP.
- [276] Pepiot, P. and Desjardins, O. (2012). Numerical analysis of the dynamics of two- and three-dimensional fluidized bed reactors using an Euler–Lagrange approach. *Powder Technology*, 220:104–121.
- [277] Pesce, G., Volpe, G., Volpe, G., and Sasso, A. (2014). Longterm Influence of Inertia on the Diffusion of a Brownian Particle. pages 3–7.
- [278] Pignatelli, F., Nicolas, M., and Guazzelli, É. (2011). A falling cloud of particles at a small but finite Reynolds number. *Journal of Fluid Mechanics*, 671:34–51.
- [279] Podlubny, I. (1999). *Fractional Differential Equations. An Introduction to Fractional Derivatives, Fractional Differential Equations, Some Methods of Their Solution and Some of Their Applications*. Academic Press.
- [280] Pope, S. B. (2001). Turbulent Flows. *Measurement Science and Technology*, 12(11):2020–2021.
- [281] Potter, D., Stadel, J., and Teyssier, R. (2017). PKDGRAV3: beyond trillion particle cosmological simulations for the next era of galaxy surveys. *Computational Astrophysics and Cosmology*, 4(1):2.
- [282] Pouliot, B., Fortin, M., Fortin, A., and Chamberland, É. (2013). On a new edge-based gradient recovery technique. *International Journal for Numerical Methods in Engineering*, 93(1):52–65.
- [283] Pozrikidis, C. and Ferziger, J. H. (1997). Introduction to Theoretical and Computational Fluid Dynamics. *Physics Today*, 50(9):72–74.
- [284] Proudman, I. and Pearson, J. R. A. (1957). Expansions at small Reynolds numbers for the flow past a sphere and a circular cylinder. *Journal of Fluid Mechanics*, 2(3):237–262.
- [285] R. C. Lift, J. R. G. R. and Weber, M. E. (1979). Bubbles, Drops and Particles. *Journal of Fluid Mechanics*, 94(04):795.
- [286] Raju, N. and Meiburg, E. (1997). Dynamics of small, spherical particles in vortical and stagnation point flow fields. *Physics of Fluids*, 9(2):299–314.
- [287] Ramírez, R., Pöschel, T., Brilliantov, N. V., and Schwager, T. (1999). Coefficient of restitution of colliding viscoelastic spheres. *Physical Review E*, 60(4):4465–4472.
- [288] Reade, W. C. and Collins, L. R. (2000). Effect of preferential concentration on turbulent collision rates. *Physics of Fluids*, 12(10):2530–2540.

- [289] Ren, F. S., Ma, R. X., and Cheng, X. Z. (2014). Simulation of particle impact drilling nozzles based on fluent. *Advanced Materials Research*, 988:475–478.
- [290] Renn, J. (2005). Einstein's invention of Brownian motion. *Annalen der Physik*, 14(S1):23–37.
- [291] Rihan, F. A. (2013). Numerical Modeling of Fractional-Order Biological Systems. *Abstract and Applied Analysis*, 2013:1–11.
- [292] Rosa, B., Wang, L.-P., Maxey, M., and Grabowski, W. (2011). An accurate and efficient method for treating aerodynamic interactions of cloud droplets. *Journal of Computational Physics*, 230(22):8109–8133.
- [293] Rovelli, C. (2015). Aristotle's Physics: A Physicist's Look. *Journal of the American Philosophical Association*, 1(01):23–40.
- [294] Rubinow, S. I. and Keller, J. B. (1961). The transverse force on a spinning sphere moving in a viscous fluid. *Journal of Fluid Mechanics*, 11(3):447–459.
- [295] Rudd, R. E. and Broughton, J. Q. (1998). Coarse-grained molecular dynamics and the atomic limit of finite elements. *Physical Review B*, 58(10):R5893–R5896.
- [296] Rybalko, M., Loth, E., and Lankford, D. (2012). A Lagrangian particle random walk model for hybrid RANS/LES turbulent flows. *Powder Technology*, 221:105–113.
- [297] Rycroft, C. H., Kamrin, K., and Bazant, M. Z. (2009). Assessing continuum postulates in simulations of granular flow. *Journal of the Mechanics and Physics of Solids*, 57(5):828–839.
- [298] Saber, A., Lundström, T. S., and Hellström, J. G. I. (2015). Turbulent Modulation in Particulate Flow: A Review of Critical Variables. *Engineering*, 07(10):597–609.
- [299] Saffman, P. G. (1965). The lift on a small sphere in a slow shear flow. *Journal of Fluid Mechanics*, 22(02):385–400.
- [300] Saffman, P. G. (1968). The lift on a small sphere in a slow shear flow - Corrigendum. *Journal of Fluid Mechanics*, 31(03):624.
- [301] Saffman, P. G. and Turner, J. S. (1956). On the collision of drops in turbulent clouds. *Journal of Fluid Mechanics*, 1(1):16–30.
- [302] Safikhani, H., Akhavan-Behabadi, M., Shams, M., and Rahimyan, M. (2010). Numerical simulation of flow field in three types of standard cyclone separators. *Advanced Powder Technology*, 21(4):435–442.
- [303] Sagaut and S. (2004). Large Eddy Simulation for Incompressible Flows: An Introduction. *Journal of Fluid Mechanics*, 501:378–379.
- [304] Sakai, M., Abe, M., Shigeto, Y., Mizutani, S., Takahashi, H., Viré, A., Percival, J. R., Xiang, J., and Pain, C. C. (2014). Verification and validation of a coarse grain model of the DEM in a bubbling fluidized bed. *Chemical Engineering Journal*, 244:33–43.

- [305] Salazar, J. P. L. C., De Jong, J., Cao, L., Woodward, S. H., Meng, H., and Collins, L. R. (2008). Experimental and numerical investigation of inertial particle clustering in isotropic turbulence. *Journal of Fluid Mechanics*, 600(April):245–256.
- [306] Samiei, K., Peters, B., Bolten, M., and Frommer, A. (2013). Assessment of the potentials of implicit integration method in discrete element modelling of granular matter. *Computers & Chemical Engineering*, 49:183–193.
- [307] Samko, S. G., Kilbas, A. A., and Marichev, O. I. (1993). *Fractional integrals and derivatives : theory and applications*. CRC press.
- [308] Santasusana, M., Irazábal, J., Oñate, E., and Carbonell, J. M. (2016). The Double Hierarchy Method. A parallel 3D contact method for the interaction of spherical particles with rigid FE boundaries using the DEM. *Computational Particle Mechanics*, 3(3):407–428.
- [309] Sardina, G., Schlatter, P., Brandt, L., Picano, F., and Casciola, C. M. (2012). Wall accumulation and spatial localization in particle-laden wall flows. *Journal of Fluid Mechanics*, 699(August 2016):50–78.
- [310] Schmidt, L., Fouxon, I., and Holzner, M. (2016). Inertial particles distribute in turbulence as Poissonian points with random intensity inducing clustering and supervoiding. *Physical Review Fluids*, 2(7):1–15.
- [311] Scotta, R., Lazzari, M., Stecca, E., Cotela, J., and Rossi, R. (2016). Numerical wind tunnel for aerodynamic and aeroelastic characterization of bridge deck sections. *Computers & Structures*, 167:96–114.
- [312] Shah, S. N., El Fadili, Y., and Chhabra, R. P. (2007). New model for single spherical particle settling velocity in power law (visco-inelastic) fluids. *International Journal of Multiphase Flow*, 33(1):51–66.
- [313] Sharipov, F. (2016). *Rarefied Gas Dynamics*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- [314] Shaw, R. A. (2003). Particle-turbulence interactions in atmospheric clouds. *Annual Review of Fluid Mechanics*, 35(1):183–227.
- [315] Sidi, A. (2003). *Practical Extrapolation Methods*. Cambridge University Press, Cambridge.
- [316] Sierou, A. and Brady, J. F. (2001). Accelerated Stokesian Dynamics simulations. *Journal of Fluid Mechanics*, 448(2001):115–146.
- [317] Simonin, O., Chevrier, S., Audard, F., and Fede, P. (2016). Drag force modelling in dilute to dense particle-laden flows with mono-disperse or binary mixture of solid particles. *Proceedings of the 9th International Conference on Multiphase Flow*, page 6.
- [318] Sommerfeld, M. and Lain, S. (2015). Parameters influencing dilute-phase pneumatic conveying through pipe systems: A computational study by the Euler/Lagrange approach. *The Canadian Journal of Chemical Engineering*, 93(1):1–17.

- [319] Sorrentino, S. and Fasana, A. (2007). Finite element analysis of vibrating linear systems with fractional derivative viscoelastic models. *Journal of Sound and Vibration*, 299(4-5):839–853.
- [320] Sousa, E. and Li, C. (2015). A weighted finite difference method for the fractional diffusion equation based on the Riemann–Liouville derivative. *Applied Numerical Mathematics*, 90(22):22–37.
- [321] Sozer, E., Brehm, C., and Kiris, C. C. (2014). Gradient Calculation Methods on Arbitrary Polyhedral Unstructured Meshes for Cell-Centered CFD Solvers. In *52nd Aerospace Sciences Meeting*, number January, pages 1–24, Reston, Virginia. American Institute of Aeronautics and Astronautics.
- [322] Squires, K. D. and Eaton, J. K. (1990). Particle response and turbulence modification in isotropic turbulence. *Physics of Fluids A: Fluid Dynamics*, 2(7):1191–1203.
- [323] Squires, K. D. and Eaton, J. K. (1991). Preferential concentration of particles by turbulence. *Physics of Fluids A: Fluid Dynamics*, 3(5):1169–1178.
- [324] Squires, K. D. and Yamazaki, H. (1995). Preferential concentration of marine particles in isotropic turbulence. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(11-12):1989–2004.
- [325] Steele, B. D. (1994). Muskets and Pendulums: Benjamin Robins, Leonhard Euler, and the Ballistics Revolution. *Technology and Culture*, 35(2):348—382.
- [326] Stokes, G. G. (1851). On the effect of the internal friction of fluids on the motion of pendulums. *Cambridge: Pitt Press*, 9:8.
- [327] Struchtrup, H. and Torrilhon, M. (2003). Regularization of Grad’s 13 moment equations: Derivation and linear analysis. *Physics of Fluids*, 15(9):2668–2680.
- [328] Sumbekova, S., Cartellier, A., Aliseda, A., and Bourgoïn, M. (2017). Preferential concentration of inertial sub-Kolmogorov particles: The roles of mass loading of particles, Stokes numbers, and Reynolds numbers. *Physical Review Fluids*, 2(2):024302.
- [329] Sun, Q., Wang, G., and Hu, K. (2009). Some open problems in granular matter mechanics. *Progress in Natural Science*, 19(5):523–529.
- [330] Sundaram, S. and Collins, L. R. (1996). Numerical Considerations in Simulating a Turbulent Suspension of Finite-Volume Particles. *Journal of Computational Physics*, 124(2):337–350.
- [331] Sundaram, S. and Collins, L. R. (1997). Collision statistics in an isotropic particle-laden turbulent suspension. Part 1. Direct numerical simulations. *Journal of Fluid Mechanics*, 335:75—109.
- [332] Sundaresan, S. (2000). Modeling the hydrodynamics of multiphase flow reactors: Current status and challenges. *AIChE Journal*, 46(6):1102–1105.
- [333] Sungkorn, R., Derksen, J., and Khinast, J. (2011). Modeling of turbulent gas–liquid bubbly flows using stochastic Lagrangian model and lattice-Boltzmann scheme. *Chemical Engineering Science*, 66(12):2745–2757.

- [334] Swedish Industrial Association for Multiphase Flows (SIAMUF) and Sommerfeld, M. (2008). *Best Practice Guidelines for Computational Fluid Dynamics of Dispersed Multi-Phase Flows*. European Research Community on Flow, Turbulence and Combustion (ERCOFTAC).
- [335] Swetz, F. J. (1989). An historical example of mathematical modelling: the trajectory of a cannonball. *International Journal of Mathematical Education in Science and Technology*, 20(5):731–741.
- [336] Syrakos, A., Varchanis, S., Dimakopoulos, Y., Goulas, A., and Tsamopoulos, J. (2016). A critical analysis of some popular methods for the discretisation of the gradient operator in finite volume methods. 29(12):1–42.
- [337] Tao, T. (2016). *Analysis II*, volume 38 of *Texts and Readings in Mathematics*. Springer Singapore.
- [338] Tatom, F. B. (1988). The Basset term as a semiderivative. *Applied Scientific Research*, 45(3):283–285.
- [339] Tenreiro Machado, J. a. (2011). And I say to myself: “What a fractional world!”. *Fractional Calculus and Applied Analysis*, 14(4):635–654.
- [340] Tewari, A. and Gokhale, A. (2004). Nearest-neighbor distances between particles of finite size in three-dimensional uniform random microstructures. *Materials Science and Engineering: A*, 385(1-2):332–341.
- [341] The HDF Group (2018). hdf5. <http://www.hdfgroup.org/HDF5/>.
- [342] Thornton, C. (2009). A note on the effect of initial particle spin on the rebound behaviour of oblique particle impacts. *Powder Technology*, 192(2):152–156.
- [343] Thornton, C., Cummins, S. J., and Cleary, P. W. (2013). An investigation of the comparative behaviour of alternative contact force models during inelastic collisions. *Powder Technology*, 233:30–46.
- [344] Tomboulides, A. G. and Orszag, S. A. (2000). Numerical investigation of transitional and weak turbulent flow past a sphere. *Journal of Fluid Mechanics*, 416(September):45—73.
- [345] Tóthová, J., Vasziová, G., Glod, L., and Lisý, V. (2011). Langevin theory of anomalous Brownian motion made simple. *European Journal of Physics*, 32(3):645–655.
- [346] Trinh, K. T. (2010). On The Critical Reynolds Number For Transition From Laminar To Turbulent Flow. page 39.
- [347] Tucker, D. L., Oemler, A., Kirshner, R. P., Lin, H., Shectman, S. A., Landy, S. D., Schechter, P. L., Muller, V., Gottlober, S., and Einasto, J. (1997). The Las Campanas Redshift Survey galaxy–galaxy autocorrelation function. *Monthly Notices of the Royal Astronomical Society*, 285(1):L5–L9.
- [348] Tuley, R., Danby, M., Shrimpton, J., and Palmer, M. (2010). On the optimal numerical time integration for Lagrangian DEM within implicit flow solvers. *Computers and Chemical Engineering*, 34(6):886–899.

- [349] Uhlmann, M. and Chouippe, A. (2016). Clustering and preferential concentration of finite-size particles in forced homogeneous-isotropic turbulence. *Journal of Fluid Mechanics*, 812:991–1023.
- [350] van Aartrijk, M. and Clercx, H. J. (2010). The dynamics of small inertial particles in weakly stratified turbulence. *Journal of Hydro-Environment Research*, 4(2):103–114.
- [351] van Hinsberg, M. A. T., Boonkkamp, J. H. M. T. T., Toschi, F., and Clercx, H. J. H. (2013). Optimal interpolation schemes for particle tracking in turbulence. *Physical Review E*, 87(4):043307.
- [352] van Hinsberg, M. A. T., Clercx, H. J. H., and Toschi, F. (2017). Enhanced settling of nonheavy inertial particles in homogeneous isotropic turbulence: The role of the pressure gradient and the Basset history force. *Physical Review E*, 95(2):023106.
- [353] van Hinsberg, M. A. T., ten Thije Boonkkamp, J. H. M., and Clercx, H. J. H. (2011). An efficient, second order method for the approximation of the Basset history force. *Journal of Computational Physics*, 230(4):1465–1478.
- [354] van Hinsberg, M. A. T., Thije Boonkkamp, J. H. M., Toschi, F., and Clercx, H. J. H. (2012). On the Efficiency and Accuracy of Interpolation Methods for Spectral Codes. *SIAM Journal on Scientific Computing*, 34(4):B479–B498.
- [355] van Wachem, B. G. M., Schouten, J. C., van den Bleek, C. M., Krishna, R., and Sinclair, J. L. (2001). Comparative analysis of CFD models of dense gas–solid systems. *AIChE Journal*, 47(5):1035–1051.
- [356] Vigolo, D., Radl, S., and Stone, H. A. (2014). Unexpected trapping of particles at a T junction. *Proceedings of the National Academy of Sciences*, 111(13):4770–4775.
- [357] Vilela Mendes, R. (2009). A fractional calculus interpretation of the fractional volatility model. *Nonlinear Dynamics*, 55(4):395–399.
- [358] Voßkuhle, M., Pumir, A., Lévêque, E., and Wilkinson, M. (2015). Collision rate for suspensions at large Stokes numbers – comparing Navier–Stokes and synthetic turbulence. *Journal of Turbulence*, 16(1):15–25.
- [359] Vreman, A. (2007). Macroscopic theory of multicomponent flows: Irreversibility and well-posed equations. *Physica D: Nonlinear Phenomena*, 225(1):94–111.
- [360] Vreman, A. W. (2015). Turbulence attenuation in particle-laden flow in smooth and rough channels. *Journal of Fluid Mechanics*, 773:103–136.
- [361] Wakaba, L. and Balachandar, S. (2007). On the added mass force at finite Reynolds and acceleration numbers. *Theoretical and Computational Fluid Dynamics*, 21(2):147–153.
- [362] Wang, L.-P., Ayala, O., and Grabowski, W. W. (2005a). Improved Formulations of the Superposition Method. *Journal of the Atmospheric Sciences*, 62(4):1255–1266.
- [363] Wang, L.-P., Ayala, O., Kasprzak, S. E., and Grabowski, W. W. (2005b). Theoretical Formulation of Collision Rate and Collision Efficiency of Hydrodynamically Interacting Cloud Droplets in Turbulent Atmosphere. *Journal of the Atmospheric Sciences*, 62(7):2433–2450.

- [364] Wang, Q., Squires, K., Chen, M., and McLaughlin, J. (1997). On the role of the lift force in turbulence simulations of particle deposition. *International Journal of Multiphase Flow*, 23(4):749–763.
- [365] Wheeler, J. D., Helfrich, K. R., Anderson, E. J., and Mullineaux, L. S. (2015). Isolating the hydrodynamic triggers of the dive response in eastern oyster larvae. *Limnology and Oceanography*, 60(4):1332–1343.
- [366] Whitaker, S. (1999). *The Method of Volume Averaging*, volume 13 of *Theory and Applications of Transport in Porous Media*. Springer Netherlands, Dordrecht.
- [367] Wilkinson, M. and Mehlig, B. (2005). Caustics in turbulent aerosols. *Europhysics Letters*, 71(2):186–192.
- [368] Williams, J. and O’Connor, R. (1995). A linear complexity intersection algorithm for discrete element simulation of arbitrary geometries. *Engineering Computations*, 12(2):185–201.
- [369] Williams, J. R., Perkins, E., and Cook, B. (2004). A contact algorithm for partitioning N arbitrary sized objects. *Engineering Computations*, 21(2/3/4):235–248.
- [370] Wood, W. L., Bossak, M., and Zienkiewicz, O. C. (1980). An alpha modification of Newmark’s method. *International Journal for Numerical Methods in Engineering*, 15(10):1562–1566.
- [371] Xiao, H. and Sun, J. (2011a). Algorithms in a Robust Hybrid CFD-DEM Solver for Particle-Laden Flows. *Communications in Computational Physics*, 9(02):297–323.
- [372] Xiao, H. and Sun, J. (2011b). Algorithms in a Robust Hybrid CFD-DEM Solver for Particle-Laden Flows. *Communications in Computational Physics*, 9(02):297–323.
- [373] Xu, L., Zhang, Q., Zheng, J., and Zhao, Y. (2016). Numerical prediction of erosion in elbow based on CFD-DEM simulation. *Powder Technology*, 302:236–246.
- [374] Yan, N. and Zhou, A. (2001). Gradient recovery type a posteriori error estimates for finite element approximations on irregular meshes. *Computer Methods in Applied Mechanics and Engineering*, 190(32-33):4289–4299.
- [375] Yan, T., Li, J., and Zhao, L. (2016). Numerical simulation on flow field of particle impact drilling in different drilling parameters. *DEStech Transactions on Computer Science and Engineering*, (Icte):1–5.
- [376] Yap, Y. W. and Sader, J. E. (2016). Sphere oscillating in a rarefied gas. *Journal of Fluid Mechanics*, 794:109–153.
- [377] Yeung, P. and Pope, S. (1988). An algorithm for tracking fluid particles in numerical simulations of homogeneous turbulence. *Journal of Computational Physics*, 79(2):373–416.
- [378] Yin, X. and Sundaresan, S. (2009). Fluid-particle drag in low-Reynolds-number polydisperse gas-solid suspensions. *AIChE Journal*, 55(6):1352–1368.

- [379] Yuste, S. B. and Acedo, L. (2005). An Explicit Finite Difference Method and a New von Neumann-Type Stability Analysis for Fractional Diffusion Equations. *SIAM Journal on Numerical Analysis*, 42(5):1862–1874.
- [380] Zaichik, L., Alipchenkov, V. M., and Sinaiski, E. G. (2008). *Particles in Turbulent Flows*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- [381] Zaichik, L. I. and Alipchenkov, V. M. (2009). Statistical models for predicting pair dispersion and particle clustering in isotropic turbulence and their applications. *New Journal of Physics*, 11(10):103018.
- [382] Zaichik, L. I., Simonin, O., and Alipchenkov, V. M. (2003). Two statistical models for predicting collision rates of inertial particles in homogeneous isotropic turbulence. *Physics of Fluids*, 15(10):2995—3005.
- [383] Zaichik, L. I., Simonin, O., and Alipchenkov, V. M. (2010). Turbulent collision rates of arbitrary-density particles. *International Journal of Heat and Mass Transfer*, 53(9-10):1613–1620.
- [384] Zastawny, M., Mallouppas, G., Zhao, F., and van Wachem, B. (2012). Derivation of drag and lift force and torque coefficients for non-spherical particles in flows. *International Journal of Multiphase Flow*, 39:227–239.
- [385] Zayernouri, M. and Karniadakis, G. E. (2014). Exponentially accurate spectral and spectral element methods for fractional ODEs. *Journal of Computational Physics*, 257:460–480.
- [386] Zenit, R. and Legendre, D. (2009). The coefficient of restitution for air bubbles colliding against solid walls in viscous liquids. *Physics of Fluids*, 21(8):083306.
- [387] Zhang, H., Trias, F. X., Gorobets, A., Oliva, A., Yang, D., Tan, Y., and Sheng, Y. (2015). Effect of collisions on the particle behavior in a turbulent square duct flow. *Powder Technology*, 269:320–336.
- [388] Zhang, J., Yan, S., Yuan, D., Alici, G., Nguyen, N.-T., Ebrahimi Warkiani, M., and Li, W. (2016). Fundamentals and applications of inertial microfluidics: a review. *Lab on a Chip*, 16(1):10–34.
- [389] Zhang, W. and Stone, H. A. (1998). Oscillatory motions of circular disks and nearly spherical particles in viscous flows. *Journal of Fluid Mechanics*, 367:329—358.
- [390] Zhang, W. M., Meng, G., and Wei, X. (2012). A review on slip models for gas microflows. *Microfluidics and Nanofluidics*, 13(6):845–882.
- [391] Zhang, Z. and Naga, A. (2005). A New Finite Element Gradient Recovery Method: Superconvergence Property. *SIAM Journal on Scientific Computing*, 26(4):1192–1213.
- [392] Zhao, L. H., Marchioli, C., and Andersson, H. I. (2012). Stokes number effects on particle slip velocity in wall-bounded turbulence and implications for dispersion models. *Physics of Fluids*, 24(2):021705.

-
- [393] Zhou, G., Xiong, Q., Wang, L., Wang, X., Ren, X., and Ge, W. (2014). Structure-dependent drag in gas–solid flows studied with direct numerical simulation. *Chemical Engineering Science*, 116:9–22.
- [394] Zienkiewicz, O. C. and Zhu, J. Z. (1992). The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique. *International Journal for Numerical Methods in Engineering*, 33(7):1331–1364.

Appendix A

DEM specifics

This appendix discusses a variety of basic topics related to the discrete element method (DEM). The method is not described in any detail and the reader is referred to the standard literature on the topic for generic DEM-related inquiries. The topics included here have been chosen specifically for their relevance in relation to the simulations of Chapter 4.

A.1 Basic Ingredients

Let us briefly review the fundamentals of the Discrete Element Method algorithm. For a recent, widely general review on the subject, see [154]. The DEM consists in the numerical integration of the trajectories of a number of particles, that move according to Newton's laws under the action of both external forces, such as their own weight, and of contact forces, that act between particles when they become close enough. The numerical integration is typically calculated with a finite difference scheme, which in the great majority of implementations is of explicit type. The reason for this has been discussed in the literature [306] and the consensus is that the cost of an implicit implementation would be greater or, at best, not justify the greater complexity and difficulty of implementation associated with it.

In the simplest versions of the method, each particle is modelled as a rigid sphere, and its (rigid solid) movement is determined by the position of its center and its rotation vector, both of which are evolved in time by the integration scheme. The presence of bounding walls can in turn be modelled by a set of flat rigid faces. For instance, in our implementation the bounding surfaces are triangulated in the pre-process step, so that each resulting triangle defines a rigid face.

The most popular variety of the DEM, and the one used here, is called the *soft-sphere* method. In this version the particles are allowed to overlap (slightly) over each other and

also penetrate the walls. A given overlap is characterized by a point inside the overlap region, the contact point, and its magnitude is represented by a scalar δ_{indent} (the indentation or penetration) that measures how far into each neighbour the contact point has moved. Associated with it there is a contact force and, sometimes, also a contact moment, to be added to the total actions being applied to the particle. These contact forces and moments are typically functions of δ_{indent} and its derivatives, and sometimes of their histories too. The simplest versions include a linear spring and dash-pot rheological model, which depend linearly on δ_{indent} and $\dot{\delta}_{\text{indent}}$, although the contact model can become much more complicated, often devised with a particular application in mind. Normally, there exist a number of free parameters that allow to calibrate these micro-scale models by comparing the resulting macroscopic motion with experimental results.

In order to avoid the $O(N^2)$ (where N is the number of particles plus walls) scaling of brute-force check of all the possible overlaps, a suitable search algorithm is always used to determine the correspondence between each particle and its neighbours. State-of-the-art algorithms achieve $O(N \log N)$ scaling or even $O(N)$ [368]. Specifically, we use a binning strategy for both particles and triangular elements for which the hierarchical method is applied [308]. The search is performed in an alternate fashion, between time-integration steps, often with a lower frequency than the time integration steps; see Section A.2.

To summarize, the DEM algorithm looks like the one presented in Algorithms (8, 9).

Algorithm 8 Basic DEM algorithm.

```

1:  $t \leftarrow 0.0$  ▷ time
2:  $m \leftarrow 0$  ▷ time step
3: INITIALIZEPARTICLES( ) ▷ Create particles and give them initial conditions
4: SEARCHNEIGHBOURS( )
5: while  $t < \text{FinalTime}$  do
6:   SOLVEDEM( $t, m$ ) ▷ Algorithm 9

```

The equation of motion of the particles Eq. (4.1) must be modified to take into account the contact forces, as

$$m_p \frac{d\mathbf{v}}{dt} = \mathbf{F}_{\text{hydro}} + \mathbf{F}_{\text{B}} + \xi_C \sum_{k_{p,\text{neigh}}} \mathbf{F}_{k_{p,\text{neigh}}} + \sum_{k_{w,\text{neigh}}} \mathbf{F}_{k_{w,\text{neigh}}} =: \mathbf{F} \quad (\text{A.1})$$

where $k_{p,\text{neigh}}$ runs over all the neighbouring particles, $k_{w,\text{neigh}}$ over all neighbouring triangular rigid walls, and where the binary parameter $\xi_C \in \{0, 1\}$ is introduced to easily turn off inter-particle interactions as required (such as in Section 4.7.8). The force is only *actually* computed if the neighbouring wall or particle center falls within the ball centred at the target particle's center and with radius equal to the *search radius*. The search radius is defined as the

Algorithm 9 Solve DEM function algorithm. These operations performed at every DEM time step.

```

1:  $t \leftarrow t + \delta t$ 
2:  $m \leftarrow m + 1$ 
3: for  $i = 1, \dots, \text{NParticles}$  do ▷ for each particle
4:    $\mathbf{r}_i, \leftarrow \text{UPDATEPOSITION}(\mathbf{r}_i, \mathbf{v}_i)$  ▷ Eq. (A.2a)
5:  $\text{MOVE RIGID PARTS}(t)$  ▷ apply imposed motions to walls
6: if  $\text{mod}(\text{NSearch}, m) == 0$  then ▷ search every NSearch integration steps
7:    $\text{SEARCH NEIGHBOURS}()$  ▷ assign neighbour elements to each particle
8: for  $i = 1, \dots, \text{NParticles}$  do ▷ for each particle
9:    $\mathbf{f}_i \leftarrow 0.0$  ▷ total force for particle  $i$ 
10:   $\mathbf{f}_i \leftarrow \mathbf{f}_i + \text{CALCULATEEXTERNALFORCES}(\mathbf{r}_i, \mathbf{v}_i)$  ▷ RHS of Eq. (4.111)
11:  for  $k_{\text{neigh},i} = 1, \dots, n_{\text{neigh},i}$  do ▷ for each neighbour (walls and other balls)
12:     $\mathbf{f}_i \leftarrow \mathbf{f}_i + \text{CALCULATECONTACTFORCE}(\mathbf{r}_i, \mathbf{v}_i, \mathbf{r}_k, \mathbf{v}_k)$  ▷ from current overlap and
    its derivative
13: for  $i = 1, \dots, \text{NParticles}$  do ▷ for each particle
14:   $\mathbf{v}_i \leftarrow \text{UPDATEVELOCITY}(\mathbf{r}_i, \mathbf{v}_i)$  ▷ Eq. (A.2b)

```

particle radius plus a tolerance that is tuned to optimize the computational cost. Such optimal cost is to be found as an ideal balance between the cost of running the search algorithm at every time step and having a larger number of neighbours per particle as a result of the enlarged search tolerance; see Section A.2.

A.1.1 Contact model

The contact model is the rheological model that relates the kinematics of a contacting pair with the force and moment between both particles involved (a plane can be seen as a degenerate particle). The kinematics in the DEM contain a number of subtleties that we will not cover here, but they can be roughly characterized by the evolution of three degrees of freedom (DOFs): the instantaneous values of the indentation and its derivative (normal motion; one DOF) and the complete history of the relative motion projected onto the plane that passes through the contact point and is orthogonal to the line joining the particles' centres (tangential motion, two DOFs).

Our contact model includes a Hertzian spring-dashpot model with no sticking [287] for the normal motion, with the particularity of having a constant coefficient of normal restitution (COR). This contact element is characterized by a spring constant and the COR. The tangential motion employs a Deresiewicz–Mindlin spring-dashpot model connected in series with a frictional element, characterized by a friction coefficient ($\mu_{p,q}$) (where p and q

identify the material of the two particles involved) in addition to the tangential spring constant and a tangential dissipation coefficient.

The model was proposed by Thornton et al. [343], see also [64]. The COR is the ratio between the impact incident normal velocity (denominator) and the normal velocity immediately after contact (numerator), and is generally bounded between zero and one. It is a useful, engineering parameter to characterize the amount of energy dissipated by the impact of a given material. However, it must be kept in mind that the assumption of it being independent of the incident velocity is actually not entirely correct [210] and so the COR is not, strictly speaking, a *material parameter*, but only approximately so.

A.1.2 Time integration Scheme

To solve Eq. (4.1) (together with a suitable initial condition $\mathbf{x}(0) = \mathbf{x}_0$), we can use a finite difference scheme. Several options are possible but we have run our simulations using the a version of the two-step Adams–Bashforth scheme, which has been extensively tested. The difference equations read as follows:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{\delta t}{2} (3\mathbf{v}_n - \mathbf{v}_{n-1}) \quad (\text{A.2a})$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \frac{\delta t}{2m_p + m_f} (3\mathbf{F}_{n+1}^* - \mathbf{F}_n^*) \quad (\text{A.2b})$$

Where \mathbf{F}^* is the total force \mathbf{F} minus the terms proportional to the acceleration of the particle in the added mass force, which is treated implicitly by increasing the mass of the particle by $m_f/2$. This two-step scheme can be started with the analogue, first step version of the same algorithm. A similar algorithm is used for the integration of the angular motion.

A.2 Scales

The introduction of DEM in point-particle methods implies that a new, (short) time scale is being resolved: the contact time scale. First of all, the fundamental reason ¹ for introducing the DEM method in the first place is that collisions are regarded as playing a non-negligible role in the simulation. Therefore, it is always necessary to ensure the non-penetrability of the spheres through one another and, often, also through the walls.

¹In absolute rigour, there exists the possibility of using the technique, almost unaltered, to simulate other types of close-range interactions, e.g. electric fields, that might not require strict impenetrability and for which the present discussion should be modified. These however fall out of the scope of the present work.

The large majority of DEM simulations employ a contact model composed of an elastic element arranged in parallel with a dissipative element, commonly of viscous nature. This kind of arrangement is known as a spring-dashpot model. The role of the dissipative element is to ensure that a correct amount of energy is lost at each impact, since no real (macroscopic) system is perfectly elastic.

The energy-dissipating mechanism is in principle beneficial from the long-term numerical stability, since it helps to control energy build-ups, although it might also have the opposite effect in some situations, since dissipative contact models often place stronger limitations in the critical time step for numerical stability.

In any case, the intensity of the dissipation element cannot in general be tuned for numerical reasons, since the amount of energy dissipated at each impact is a crucial physical aspect that determines the macroscopic behaviour of the system, specially in collisional regimes (the motion is mostly formed of a succession of collisions alternated with contact-less, ballistic motion).

On the other hand, the action of the elastic element is typically less constrained by the physics. In many DEM simulations, specifically those dealing with moderately dense to disperse regimes, the role of the elastic element can be summarized as

1. ensuring that sufficient repulsion is achieved in all circumstances to avoid excessive penetration and
2. ensuring the separation of scales, in time and space, between those dynamic scales associated with free flight and those associated with the process of rebound.

Note that, in disperse simulations, it is often not relevant to resolve the actual rebound process with accuracy. In fact, in these situations the process is resolved only as a method to obtain a robust numerical technique, just as any penalty model in contact mechanics.

Choosing the time step size

The selection of a suitable time step is crucially important in any DEM simulation, since it has a proportional impact in the numerical cost of the simulations. As in any finite difference calculation, one would like to select the time step based on the required accuracy, so as to minimize the total number of time steps. However, the choice is also restricted by the need to preserve numerical stability and accuracy, which are both always conditional to the size of the time step, that must be kept small enough.

The analytical calculation of the time step can be attempted in some cases, mostly using approximations such as the linearisation of the nonlinear force models. However, such

calculations are too complex in practice or lead to excessively rough estimates of little practical value. We base our choice of the time step on experience, often requiring a certain number of iterations in order to attain a good compromise between accuracy and cost.

Nonetheless, we do not proceed blindly, but actually apply a criterion that we next describe and that will also provide an argument to explain the high computational demands of this numerical approach. Our criterion is based on the following considerations:

- The smallest scales represented in our simulations correspond to the contact dynamics, and consequently it is the contact that dictates the maximum allowable time step, not the interactions with the fluid (only the added mass force, F_A has a response time at a comparable time scale, but we are treating it implicitly).
- As discussed above, the stiffness of the contact model is considered a numerical parameter that can be softened to increase the critical time step for numerical stability. This practice is acceptable as long as the time scale of the contact is still properly separated (see [272] for a discussion on the issue of scales separation) and that the maximum indentations remain small, so as to not significantly affect the packing configuration and avoid going through obstacles.
- The contact must be properly resolved to avoid excessive numerical errors and spurious energy creation. This is guaranteed by dividing the contact duration in, at least 15, to 30 steps, depending on the numerical scheme.

According to this, we thus must first calculate the minimal expected contact duration. In order to do that, we may use the formula that can be found in [64] (see also [10]), valid for Hertzian contact. For Hertzian contact laws, the contact duration decreases with the impact velocity and thus the worse-case scenario is the largest estimated impact velocity. Once we have the minimal expected impact duration, we divide it into a large enough number, say, fifteen, and apply a security factor to further reduce it. We have observed that using this technique leads to robust estimations of the optimal time step with a limited need for trial-and-error.

A.3 Wear

It would be interesting to predict the level of wear on the different surfaces of the drill bit under the repeated impact of the steel particles. While a quantitative prediction is certainly challenging, we believe there is a potential for the prediction of relevant qualitative trends, such as:

- the location of intense wear concentrations
- the sensitivity of the wear spread pattern upon changes in the design and operation parameters
- the identification of unexpected wear mechanisms
- the classification of frictional wear versus impact wear regions

We have implemented a simple wear model to illustrate these points. During the contact of a particle against a wall, an *impact wear* (W_{impact}) contribution is calculated according to

$$\Delta W_{\text{impact}} \sim \rho_p d_p U_n \quad (\text{A.3})$$

where U_n is the normal relative velocity between the particle and the triangular surface. This quantity is then divided by the face area and distributed to the nodes using the triangle's shape functions. Note that while the factors in Eq. (A.3) surely contribute to the wear, their powers have been arbitrarily set to one for simplicity. A realistic model would require further investigations and is thus left for future work.

A.4 Analytic tools

Taking advantage of the object-oriented philosophy of Kratos Multiphysics, we have extended the notion of the discrete element to a generalized *analytic discrete element*. This type of element is designed to collect information during a simulation by increasing its associated data-structures in a trade-off of information vs. computational efficiency.

The key concept is to allow a (small) proportion of the discrete elements to be marked as analytic, behaving as the discrete element they generalize but performing a few extra operations and recording a few extra data. These discrete elements are constantly monitored by an external process that collects the information stored in each of the analytic elements, ignoring the rest. The concrete data stored in the data structures associated with these discrete elements are best understood by example. Here we give a brief account of two variants that we have implemented in our code.

A.4.1 Analytic particles

The analytic particle is a discrete element that interacts exactly identically with the rest of the particles but that keeps a record of the impact data associated with the contacting neighbours at a given time step. This means that the information is only related to a specific time step and

thus must be collected at every time step or else it is lost. This design requires an operation to be added to the DEM solution with the same frequency as the DEM solution itself, but is only (relatively speaking) costly if the proportion of analytic particles is comparable to the total number of particles in the domain. The information collected from all the particles is stored in a database for later analysis. We use HDF5 files to store the information.

The reason of this design is to keep the data structures associated to the discrete elements as small as possible, so as to make the most out of the available cache. Having very heavy particles would result in extremely slow computations overall, spoiling the efficiency of the program. By limiting the total number of possible simultaneous impacts to a few (in our case, only four), the data structures are kept at a fixed size, avoiding allocation/deallocation on the fly. Note that it is extremely unlikely that more than four impacts occur at exactly the same time step.

The precise information kept per impact may vary, but a useful combination is to keep the impact velocity (normal and tangential to the particle surface), the ID of the other particle and the positions at the moment of impact. This information can be used in simulations to keep track of the impact frequency and violence [183] and other interaction such as chemical exchanges etc.

A.4.2 Analytic surfaces

By counting the triangular rigid faces as DEM elements too, one can generalize them in a similar way as we have explained with analytic particles. Such analogous surfaces could be used to track impact locations and wear [373] etc. But we have devised another type of analytic surface with the particularity of *not affecting the contacting particles*. Instead, these surfaces make measurements of the particles as those traverse them. This allows to measure fluxes of particles. The versatility of the DEM rigid faces is inherited by these flux-measuring particles, that may be meshed to cover any complicated cross-section. The information is again stored in appropriate HDF5 files for posterior analysis.

Appendix B

Quadrature substepping formulas

Daitche's method

Formulas Eq. (3.27) and Eq. (3.28) were given in Daitche [89]. A general formula was also derived in the same work which, it was claimed, contained Eq. (3.27) and Eq. (3.28) as special cases. This formula reads

$$\mu_j^{nm} = \sum_{i=0}^{n-1} \sum_{k=0}^m \Theta(o_i^{nm} + k = n - j) \int_{t_i}^{t_{i+1}} K(t - \tau) L_{ik}^{nm}(\tau) d\tau \quad (\text{B.1})$$

where Θ is a function that returns the boolean truth value of the expression it evaluates (i.e., 1 for true statements and 0 otherwise) and where

$$o_i^{nm} = \begin{cases} 0 & 0 \leq i \leq \lfloor \frac{m}{2} \rfloor \\ i - \lfloor \frac{m}{2} \rfloor & \lfloor \frac{m}{2} \rfloor < i < n - m + \lfloor \frac{m}{2} \rfloor \\ n - m & n - m + \lfloor \frac{m}{2} \rfloor \leq i \leq n - 1 \end{cases} \quad (\text{B.2})$$

and where L_{ik}^{nm} is the Lagrange polynomial that is equal to one at $t_{o_i^{nm}+k}$ and zero at the remaining points in $\{t_{o_i^{nm}+0}, \dots, t_{o_i^{nm}+m}\}$.

According to our calculations however, the direct application of Eq. (B.1), together with Eq. (B.2), does not in fact yield Eq. (3.28) for $m = 2$ as it should. The origin of this inconsistency can be traced back to the way in which the interpolation points are positioned around a generic interval $[t_i, t_{i+1}]$ (Note that the notation in Daitche [89] was instead $[\tau_i, \tau_{i+1}]$), which was fixed according to different rules for the generic case and for the special cases in [89]. The derivation of Eq. (B.1) starts by considering the polynomial interpolation

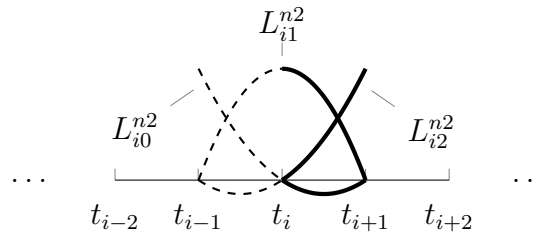


Fig. B.1 Interpolation of f in a generic, away-from-the-boundary interval $[t_i, t_{i+1}]$, according to the general formula proposed by Daitche [89].

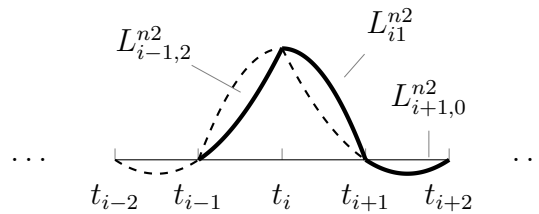


Fig. B.2 Support of the interpolands weighed by $f(t_i)$, according to the general formula proposed by Daitche [89].

of f in Eq. (3.2) with Lagrange polynomials in the interval $[t_i, t_{i+1}]$. Daitche takes

$$\mathbf{f}|_{t_i}^{t_{i+1}}(\tau) \approx \sum_{k=0}^m \mathbf{f}(\tau_{o_i^{nm+k}}) L_i^{nm}(\tau) \quad (\text{B.3})$$

Which, particularized to order two ($m = 2$) reads

$$\mathbf{f}|_{t_i}^{t_{i+1}}(\tau) \approx \sum_{k=0}^2 \mathbf{f}(\tau_{o_i^{n2+k}}) L_i^{n2}(\tau) = \mathbf{f}(\tau_{i-1}) L_{i0}^{n2}(\tau) + \mathbf{f}(\tau_i) L_{i1}^{n2}(\tau) + \mathbf{f}(\tau_{i+1}) L_{i2}^{n2}(\tau) \quad (\text{B.4})$$

Fig. B.1 depicts the interpolation polynomials L_{i0}^{n2} , L_{i1}^{n2} and L_{i2}^{n2} . In full line are represented the segments that are actually used for interpolation in $[t_i, t_{i+1}]$. Fig. B.2, meanwhile, depicts the interpolation term weighed by $f(t_i)$ showing that its support is $[t_{i-1}, t_{i+2}]$. In other words, this interval is the *domain of influence* of $f(t_i)$ for the quadrature.

We are going to show that, for second order, the interpolation scheme shown in Figs. B.1 and B.2 cannot lead to Eq. (3.28). Indeed, let us look at the central piece, i.e. for $2 < j < n - 1$.

$$\beta_j^n = \frac{8}{15} \left((j+2)^{5/2} - 3(j+1)^{5/2} + 3j^{5/2} - (j-1)^{5/2} \right) + \frac{2}{3} \left(-(j+2)^{3/2} + 3(j+1)^{3/2} - 3j^{3/2} + (j-1)^{3/2} \right) \quad (\text{B.5})$$

Using $t_{n-i} = (n-i)h$, the formula becomes

$$\begin{aligned} & \frac{8}{15h^{5/2}} \left((t_n - (t_{n-j-2}))^{5/2} - 3(t_n - (t_{n-j-1}))^{5/2} + 3(t_n - (t_{n-j}))^{5/2} - (t_n - (t_{n-j+1}))^{5/2} \right) \\ & + \frac{2}{3h^{3/2}} \left(-(t_n - (t_{n-j-2}))^{3/2} + 3(t_n - (t_{n-j-1}))^{3/2} - 3(t_n - (t_{n-j}))^{3/2} + (t_n - (t_{n-j+1}))^{3/2} \right) \end{aligned} \quad (\text{B.6})$$

Now, remembering that β_j^n is the coefficient of $f(t_{n-j})$ in the overall interpolation (see Eq. (3.26)), it is clear that, according to B.2, the influence of $f(t_{n-j})$ should be restricted to $[t_{n-j-1}, t_{t_{-n+2}}]$, and not to $[t_{n-j-2}, t_{t_{-n+1}}]$ as in Eq. (B.6). This statement alone proves our claim without resorting to the a full calculation using the formula in Eq. (B.1). Our calculations show that however, this problem does not appear for order one or three. The reason is simple: Daitche chose to place the interpolating points as close as possible to the interval $[t_i, t_i + 1]$. For odd-order polynomials (even number of interpolation points), there arises a unique solution; i.e. a symmetric configuration with the interval in the middle. On the other hand, for even-order interpolations there are two possibilities: one having one more point to the left of the interval and the other having one more point to the right. The first option is implemented when using Eq. (B.2). We propose to replace this equation by

$$o_i^{nm} = \begin{cases} 0 & 0 \leq i \leq \lfloor \frac{m}{2} \rfloor \\ i - \lfloor \frac{(m-1)}{2} \rfloor & \lfloor \frac{m}{2} \rfloor < i < n - m + \lfloor \frac{m}{2} \rfloor \\ n - m & n - m + \lfloor \frac{m}{2} \rfloor \leq i \leq n - 1 \end{cases} \quad (\text{B.7})$$

which implements the alternative option, that is, to have one more point *to the right* of the interval. Note that this formula is equivalent to Eq. (B.2) for odd-order cases, while it moves one interval to the right the value of o_i^{nm} for the even-order ones (when possible, that is, away from the present-time boundary). Our calculations show that this solves the problem and that indeed Eq. (B.1), with the new definition of o_i^{nm} , holds α_j^n and β_j^n as special cases. Figs. B.3 and B.4 display the corrected configuration, consistent with Eq. (B.1) with Eq. (B.7) and, in particular, with Eq. (3.27), Eq. (3.28).

First Order

The first-order interpolation polynomials are illustrated with full generality in Fig. B.5. In order to compute the formula for α_j^n , we remember that α_j^n plays the role of the μ_j^n in the quadrature formula Eq. (3.26). Thus, only the Lagrange polynomial that are equal to one at t_{n-j} contribute to this coefficient. Fig. B.6 displays the different variations of the support of these polynomials. Let us calculate α_1^n to illustrate the calculation of the different pieces of

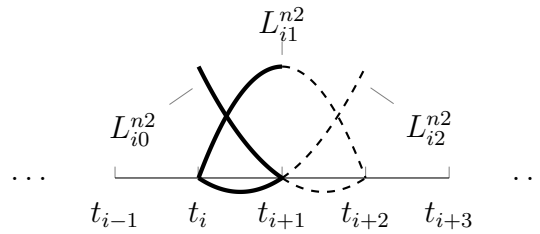


Fig. B.3 Interpolation of f in a generic, away-from-the-boundary interval $[t_i, t_{i+1}]$, according to the corrected formula

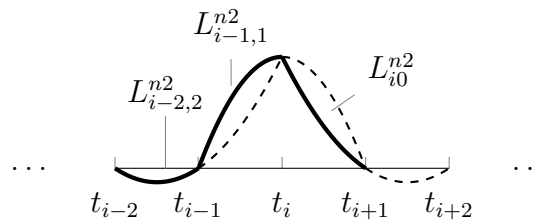


Fig. B.4 Support of the second-order interpolands weighed by $f(t_i)$, according to the corrected formula

α_j^n . The coefficient α_1^n multiplies $f(t_{n-1})$ (see Eq. (3.26)) and so we can rely on Fig. B.6 to right

$$\left(\int_{t_{n-2}}^{t_{n-1}} \frac{L_{n-2,1}^{n1}}{\sqrt{t_n - \tau}} d\tau + \int_{t_{n-1}}^{t_n} \frac{L_{n-1,0}^{n1}}{\sqrt{t_n - \tau}} d\tau \right) f(t_{n-1}) = \alpha_1^n f(t_{n-1}) \quad (\text{B.8})$$

so

$$\alpha_1^n = \frac{4}{3}(1 + \phi)(\sqrt{1 + \phi} - \sqrt{\phi}) \quad (\text{B.9})$$

Similarly the rest of pieces are calculated, yielding the general formula Eq. (3.29).

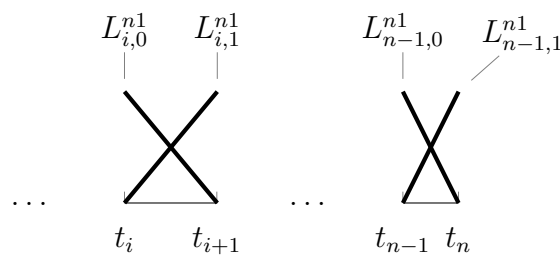


Fig. B.5 Interpolation of f . The distance between successive points is h , except for the pair t_{n-1} and t_n , in which case it is ϕh .

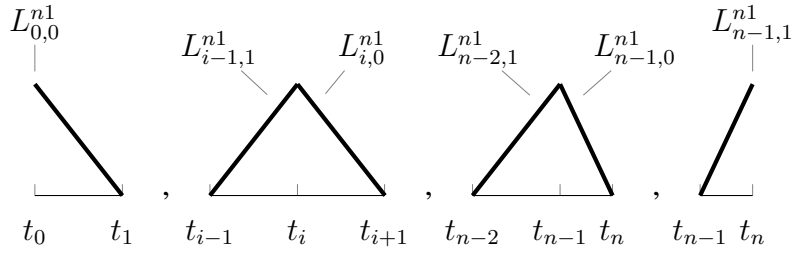


Fig. B.6 Support of the first order interpolands

Second Order

We will illustrate the calculation of the modified formulas by considering two cases: the central piece of the functions, that is, the analogue of Eq. (B.5) with a shortened last interval, which is only affected by the modified argument of the kernel; and the piece corresponding to $n > 3$ and $j = 0$, representative of the formulas affected by the modification of the interpolation polynomials. The rest of the formulas are obtained similarly and only the final results will be included for the sake of brevity. We will proceed without resorting to the general formula Eq. (B.1). Instead, it is sufficient to refer to Fig. B.4. The contribution of $\mathbf{f}(t_{n-j})$ to the total integral is given by

$$\left(\int_{t_{n-j-2}}^{t_{n-j-1}} \frac{L_{n-j-2,2}^{n2}}{\sqrt{t_n - \tau}} d\tau + \int_{t_{n-j-1}}^{t_{n-j}} \frac{L_{n-j-1,1}^{n2}}{\sqrt{t_n - \tau}} d\tau + \int_{t_{n-j}}^{t_{n-1}} \frac{L_{n-j,0}^{n2}}{\sqrt{t_n - \tau}} d\tau \right) \mathbf{f}(t_{n-j}) = \beta_j^n \mathbf{f}(t_{n-j}) \quad (\text{B.10})$$

Substituting the formulas for the Lagrange polynomials and using $t_n = (n - 1 + \phi)h$ and $t_{n-j} = (n - j)h$, we obtain

$$\begin{aligned} \beta_j^n = & \frac{8}{15} \left((j + 1 + \phi)^{5/2} - 3(j + \phi)^{5/2} + 3(j - 1 + \phi)^{5/2} - (j - 2 + \phi)^{5/2} \right) \\ & + \frac{2}{3} \left(-(j + 1 + \phi)^{3/2} + 3(j + \phi)^{3/2} - 3(j - 1 + \phi)^{3/2} + (j - 2 + \phi)^{3/2} \right) \end{aligned} \quad (\text{B.11})$$

Where now Eq. (B.5) is recovered for $\phi = 1$.

Let us now turn to the case where not only the kernel's argument changes, but also the definition of the interpolation polynomials. The case $n > 3$ and $j = 0$ requires the consideration of the situation displayed in B.7. The general interpolation scheme has to be modified to avoid requiring points outside the domain (in this case future values). The contribution of $\mathbf{f}(t_n)$ is given by

$$\left(\int_{t_{n-2}}^{t_n} \frac{L_{n-1,2}^{n2}}{\sqrt{t_n - \tau}} d\tau \right) \mathbf{f}(t_n) = \beta_0^n \mathbf{f}(t_n) \quad (\text{B.12})$$

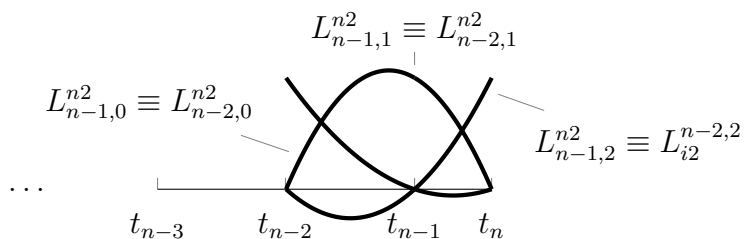


Fig. B.7 Interpolation of f at the interval $[t_{n-2}, t_n]$. The distance between t_{n-1} and t_n is ϕh .

And again, substituting the formula for $L_{n-1,2}^{n2}$ and using $t_n = (n - 1 + \phi)h$

$$\beta_0^n = \frac{4}{15\phi} \sqrt{(1 + \phi)} (4\phi - 1) \sqrt{h} \quad n \geq 4 \quad (\text{B.13})$$

which again, for $\phi = 1$ agrees with Eq. (3.28). Similar calculations can be performed for the rest of the cases, yielding the general expression Eq. (3.30).

Appendix C

Alternative expression for the history force

We will first proof the relationship for kernels that are bounded in $[t_0, t]$. We have that

$$\begin{aligned} \frac{d}{dt} \int_{t_0}^t K(t-\tau) \mathbf{f}(\tau) d\tau &= K(0) \mathbf{f}(t) + \int_{t_0}^t \frac{d}{dt} K(t-\tau) \mathbf{f}(\tau) d\tau \\ &= K(0) \mathbf{f}(t) - \int_{t_0}^t \frac{d}{d\tau} (K(t-\tau) \mathbf{f}(\tau)) d\tau \\ &\quad + \int_{t_0}^t K(t-\tau) \frac{d}{d\tau} \mathbf{f}(\tau) d\tau \\ &= K(t-t_0) \mathbf{f}(t_0) + \int_{t_0}^t K(t-\tau) \frac{d}{d\tau} \mathbf{f}(\tau) d\tau \end{aligned} \tag{C.1}$$

Now, when K is not defined at $\tau = t$, as in the case of the Basset kernel, the derivation above must be altered. We proceed by constructing a sequence of kernels that limit at the singular kernel to derive an analogous result. Consider

$$K_n : [t_0, \infty) \rightarrow \mathbb{R}, \quad t \mapsto K\left(t + \frac{1}{n}\right) \tag{C.2}$$

where n is a positive integer. The sequence of K_n for all positive integers converges pointwise to the desired kernel K . The result we are after will easily follow for K if we can move the limit sign from the integrand on the LHS expression in Eq. (C.1) to outside the derivative, as we will show. Note that we are not interested in validity of the formula at exactly $t = t_0$, since the impulse due to this infinite value at the initial time is zero anyway. Thus, mathematically,

we want to show that for any $t > t_0$

$$\lim_{n \rightarrow \infty} \frac{d}{dt} \int_{t_0}^t K_n(t - \tau) \mathbf{f}(\tau) d\tau = \frac{d}{dt} \int_{t_0}^t \lim_{n \rightarrow \infty} K_n(t - \tau) \mathbf{f}(\tau) d\tau \quad (\text{C.3})$$

First, the limit can be moved outside the integral by the Dominated Convergence theorem. All we must prove is that there exists an integrable function that is greater or equal to all the integrands in the sequence. But

$$|K_n(t - \tau) \mathbf{f}(\tau)| \leq K(t - \tau) \|\mathbf{f}\|_\infty \quad (\text{C.4})$$

and the RHS is an integrable function, because

$$\int_{t_0}^t K \|\mathbf{f}\|_\infty d\tau = 2 \|\mathbf{f}\|_\infty \sqrt{t - t_0} \quad (\text{C.5})$$

On the other hand, to show that the limit can be moved outside the derivative as well, we use the following theorem [337, Theorem 3.7.1]:

Theorem 2 *Let F_n define a sequence of differentiable functions in a closed interval $I = [t_1, t_2]$, for $n \geq 1$, and let F'_n be the corresponding sequence of derivatives, also defined in I . Suppose that F'_n converges uniformly to some function G , also defined in I . Suppose also that there exists a point t_3 where the limit $\lim_{n \rightarrow \infty} F_n(t_3)$ exists. Then $\{F_n\}$ converges uniformly to a differentiable function F , and its derivative is G .*

We want to apply the theorem to the sequence of functions

$$F_n(t) : I \rightarrow \mathbb{R}, \quad t \mapsto \int_{t_0}^t K_n(t - \tau) \mathbf{f}(\tau) d\tau \quad (\text{C.6})$$

where t_1 is an arbitrary number in I . We have already shown that

$$\lim_{n \rightarrow \infty} F_n = \int_{t_0}^t K(t - \tau) \mathbf{f}(\tau) d\tau =: F(t) \quad (\text{C.7})$$

To particularize Theorem 2 to these choices of F and F_n it is enough to show that the following hold:

- a) Each F_n is differentiable in I .

b) There exists a t_3 in I , such that

$$\lim_{n \rightarrow \infty} \int_{t_0}^{t_3} K_n(t_3 - \tau) \mathbf{f}(\tau) \, d\tau \quad (\text{C.8})$$

exists.

c) The sequence $\{F'_n\}$ converges uniformly to some function, G , defined in I .

The requirement a) follows immediately by the differentiability of the integrands. The requirement b) follows from the existence of the Basset force of all $t > t_0$. Finally, let us consider

$$\begin{aligned} \|F'_m - F'_n\| &= \\ &\left\| \mathbf{f}(t_0) (K_m(t - t_0) - K_n(t - t_0)) + \int_{t_0}^t \frac{d\mathbf{f}}{d\tau} (K_m(t - \tau) - K_n(t - \tau)) \, d\tau \right\| \\ &\leq \|\mathbf{f}\|_{W^{1,\infty}} \left| \frac{1}{\sqrt{t - t_0 + \frac{1}{m}}} - \frac{1}{\sqrt{t - t_0 + \frac{1}{n}}} \right| + \\ &2 \left(\left| \sqrt{t - t_0 + \frac{1}{m}} - \sqrt{t - t_0 + \frac{1}{n}} \right| + \left| \sqrt{\frac{1}{m}} - \sqrt{\frac{1}{n}} \right| \right) \end{aligned} \quad (\text{C.9})$$

which tends to zero as m and n tend to zero (assuming the derivative of \mathbf{f} to be well behaved). So the sequence is uniformly Cauchy and, thus, uniformly convergent to some function G , defined in I . We have proven that c) also holds. We can now apply Theorem 2 and write

$$\begin{aligned} F'(t) &= \frac{d}{dt} \int_{t_0}^t K(t - \tau) \mathbf{f}(\tau) \, d\tau = G(t) = \lim_{n \rightarrow \infty} F'_n(t) = \\ &= \lim_{n \rightarrow \infty} K_n(t - t_0) \mathbf{f}(t_0) + \int_{t_0}^t K_n(t - \tau) \frac{d}{dt} \mathbf{f}(\tau) \, d\tau \\ &= K(t - t_0) \mathbf{f}(t_0) + \int_{t_0}^t K(t - \tau) \frac{d}{dt} \mathbf{f}(\tau) \, d\tau \end{aligned} \quad (\text{C.10})$$

for all $t \in I$, where in the last equality we have used the Dominated Convergence theorem again to move the limit under the integral sign.

In summary, we have proved that the formula is valid in an interval $I = [t_1, t_2]$, as long as $t_1 > t_0$; thus proving Eq. (3.3), also for the Basset kernel.

Appendix D

Error bound for the kernel approximation

Let us define

$$e_K := K_B - K \quad (\text{D.1})$$

We want to establish an upper bound for Eq. (3.20). Ignoring the constant C_B , we have that the RHS of this equation is

$$\begin{aligned} I &:= \left| e_K(t - t_0)\mathbf{f}(t_0) + \int_{t_0}^{t-t_w} e_K(t - \tau) \frac{d\mathbf{f}(\tau)}{d\tau} d\tau \right| = \\ &= \left| e_K(t - t_0)\mathbf{f}(t_0) + \int_{t_0}^{t-t_w} \frac{d}{d\tau} (e_K(t - \tau)\mathbf{f}(\tau)) - \mathbf{f}(\tau) \frac{d}{d\tau} e_K(t - \tau) d\tau \right| = \quad (\text{D.2}) \\ &= \left| e_K(t_w - t_0)\mathbf{f}(t - t_w) - \int_{t_0}^{t-t_w} \mathbf{f}(\tau) \frac{d}{d\tau} e_K(t - \tau) d\tau \right| \end{aligned}$$

Applying the change of variables $\tilde{\tau} = (t - \tau)/t_w$, we obtain

$$I = \frac{1}{\sqrt{t_w}} \left| \tilde{e}_K(1)\mathbf{f}(t - t_w) + \int_1^{\tilde{t}_0} \mathbf{f}(t - t_w\tilde{\tau}) \frac{d}{d\tilde{\tau}} \tilde{e}_K(\tilde{\tau}) d\tilde{\tau} \right| \quad (\text{D.3})$$

where $\tilde{t}_0 = (t - t_0)/t_w$ and

$$\tilde{e}_K = K_B - \tilde{K} \quad (\text{D.4})$$

with \tilde{K} defined as

$$\tilde{K}(\tilde{\tau}) = \begin{cases} K_B(\tilde{\tau}) & \text{if } \tilde{\tau} \leq t_w \\ \tilde{K}_T(\tilde{\tau}) = \sum_{i=1}^m a_i \tilde{K}_i(\tilde{\tau}) & \text{if } \tilde{\tau} > t_w \end{cases} \quad (\text{D.5})$$

and

$$\tilde{K}_i(\tilde{\tau}) = \alpha \left(\frac{t_i}{t_w} \right) e^{\beta \left(\frac{t_i}{t_w} \right) \tilde{\tau}} \quad i = 1, \dots, m \quad (\text{D.6})$$

In other words, \tilde{K} is obtained from K by substituting all the t_i by their normalized counterparts. Now, from Eq. (D.3) and Eq. (D.1) we can immediately write

$$I \leq \frac{\|\mathbf{f}\|_\infty}{\sqrt{t_w}} \left(\left| K_B(1) - \tilde{K}(1) \right| + \int_1^{\tilde{t}_0} \left| \frac{d}{d\tau} \left(K_B(\tau) - \tilde{K}(\tau) \right) \right| d\tau \right) \quad (\text{D.7})$$

where we have changed the dummy integration variable back to τ .

Appendix E

Quadratic character of the I_{2tH} problem

In this Appendix, we outline the quadratic character of the minimization problem when I_{2tH} bound is taken as the cost function. The \tilde{t}_i values of I_{2tH} bound must be given and we represent them by \tilde{T}_i in I_{2tH} as

$$I_{2tH}(a_i) = I_{2t}(a_i, \tilde{T}_i) = e_K(1, a_i, \tilde{T}_i)^2 + \int_1^\infty \tilde{\tau} \left(\frac{\partial e_K(\tilde{\tau}, a_i, \tilde{T}_i)}{\partial \tilde{\tau}} \right)^2 d\tilde{\tau} \quad (\text{E.1})$$

The kernel error with respect to a_i can be written as

$$e_K(\tilde{\tau}, a_i, \tilde{t}_i) = K_B(\tilde{\tau}) - \tilde{K}(\tilde{\tau}, a_i, \tilde{t}_i) = \frac{1}{\sqrt{\tilde{\tau}}} - \sum \frac{a_i}{\sqrt{\tilde{t}_i}} e^{\frac{1}{2}(1-\frac{\tilde{\tau}}{\tilde{t}_i})} = K_B(\tilde{\tau}) - \sum \tilde{K}(\tilde{\tau}, 1, \tilde{t}_i) a_i \quad (\text{E.2})$$

Thus, the first term of I_{2tH} bound can be expressed, using Einstein notation, as

$$\begin{aligned} e_K(1, a_i, \tilde{T}_i)^2 &= \left(K_B(1) - \tilde{K}(1, 1, \tilde{t}_i) a_i \right)^2 \\ &= a_i \left(\tilde{K}(1, 1, \tilde{t}_i) \tilde{K}(1, 1, \tilde{t}_j) \right) a_j - 2K_B(1) \tilde{K}(1, 1, \tilde{t}_j) a_j + K_B^2(1) \end{aligned} \quad (\text{E.3})$$

By defining the design variables as $x = a_i$, the $A_{ij}^{(1)}$ matrix, the $b_j^{(1)}$ vector and $c^{(1)}$ as

$$\begin{aligned} A_{ij}^{(1)} &= 2\tilde{K}(1, 1, \tilde{t}_i) \tilde{K}(1, 1, \tilde{t}_j) \\ b_j^{(1)} &= 2K_B(1) \tilde{K}(1, 1, \tilde{t}_j) \\ c^{(1)} &= K_B^2(1) \end{aligned} \quad (\text{E.4})$$

the first term of I_{2tH} is readily rewritten in matrix notation, as

$$e_K(1, a_i, \tilde{T}_i)^2 = \frac{1}{2}xA^{(1)}x - b^{(1)}x + c^{(1)} \quad (\text{E.5})$$

Similarly, the second term of I_{2tH} has the following form

$$\begin{aligned} \int_1^\infty \tilde{\tau} \left(\frac{\partial e_K(\tilde{\tau}, a_i, \tilde{T}_i)}{\partial \tilde{\tau}} \right)^2 d\tilde{\tau} &= \int_1^\infty \tilde{\tau} \left(\frac{\partial K_B(\tilde{\tau})}{\partial \tilde{\tau}} - \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_i)}{\partial \tilde{\tau}} a_i \right)^2 d\tilde{\tau} \\ &= a_i \left(\int_1^\infty \tilde{\tau} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_i)}{\partial \tilde{\tau}} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_j)}{\partial \tilde{\tau}} d\tilde{\tau} \right) a_j \\ &= - \left(2 \int_1^\infty \tilde{\tau} \frac{\partial K_B(\tilde{\tau})}{\partial \tilde{\tau}} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_j)}{\partial \tilde{\tau}} d\tilde{\tau} \right) a_j + \int_1^\infty \tilde{\tau} \left(\frac{\partial K_B(\tilde{\tau})}{\partial \tilde{\tau}} \right)^2 d\tilde{\tau} \end{aligned} \quad (\text{E.6})$$

Again defining $A_{ij}^{(t)}$ matrix, the $b_j^{(t)}$ vector and $c^{(t)}$ as

$$\begin{aligned} A_{ij}^{(t)} &= 2 \int_1^\infty \tilde{\tau} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_i)}{\partial \tilde{\tau}} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_j)}{\partial \tilde{\tau}} d\tilde{\tau} \\ b_j^{(t)} &= 2 \int_1^\infty \tilde{\tau} \frac{\partial K_B(\tilde{\tau})}{\partial \tilde{\tau}} \frac{\partial \tilde{K}(\tilde{\tau}, 1, \tilde{t}_j)}{\partial \tilde{\tau}} d\tilde{\tau} \\ c^{(t)} &= \int_1^\infty \tilde{\tau} \left(\frac{\partial K_B(\tilde{\tau})}{\partial \tilde{\tau}} \right)^2 d\tilde{\tau} \end{aligned} \quad (\text{E.7})$$

The second term of the I_{2tH} bound can be compactly expressed as

$$\int_1^\infty \tilde{\tau} \left(\frac{\partial e_K(\tilde{\tau}, a_i, \tilde{T}_i)}{\partial \tilde{\tau}} \right)^2 d\tilde{\tau} = \frac{1}{2}xA^{(t)}x - b^{(t)}x + c^{(t)} \quad (\text{E.8})$$

Collecting the first and the second terms, we obtain the final expression for the I_{2tH} bound

$$I_{2tH} = \frac{1}{2}x(A^{(1)} + A^{(t)})x - (b^{(1)} + b^{(t)})x + (c^{(1)} + c^{(t)}) \quad (\text{E.9})$$

which stands for a standard quadratic minimization as

$$I_{2tH} = \frac{1}{2}xAx - bx. \quad (\text{E.10})$$

where A and b are defined as $A = A^{(1)} + A^{(t)}$ and $b = b^{(1)} + b^{(t)}$. Note that $c = c^{(1)} + c^{(t)}$ is suppressed from the expression since it does not play any role in the minimization problem.

Appendix F

Optimal a_i and t_i values

We provide the optimal a_i and \tilde{t}_i for each of the cost functions I_1 (Table F.1) and I_{2t} (Table F.2). Note the increasing distance between successive \tilde{t}_i values, as was assumed by van Hinsberg et al. [353]. Note also that for $m = 1$ we obtained $a_1 > 1$.

$m = 1$		$m = 2$		$m = 3$		$m = 4$		$m = 5$	
a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i
1.046347992	1.581186674	0.566192817	0.717656182	0.440072204	0.482318894	0.374397988	0.365083559	0.3450551877	0.3227320427
		0.864298391	8.925153279	0.538287204	3.324763126	0.421322343	1.820334739	0.3762685526	1.4017593843
				0.807797346	38.928376132	0.517872275	11.809488351	0.4383511621	7.3543952717
						0.761539469	127.109159354	0.5502868981	52.9058339347
								0.7701813938	699.4337431732

$m = 6$		$m = 7$		$m = 8$		$m = 9$		$m = 10$	
a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i
0.3227460255	0.2894856389	0.2931405176	0.2413624327	0.2718360249	0.2192620346	0.2570818336	0.1878604572	0.2520642358	0.1878604572
0.3446901326	1.1312690586	0.3053190176	0.8199848671	0.2685924185	0.662026818	0.2610118588	0.5420260992	0.254913066	0.5306382498
0.3924441164	5.1207861657	0.3394616674	3.0838532791	0.2871214552	2.0706383247	0.2799238451	1.6534881587	0.2638832071	1.5524873935
0.471576099	29.6345412934	0.3924532926	13.8047974118	0.3249589764	7.2825402363	0.3051985477	5.5204876302	0.2666445191	4.6517443725
0.5990063177	256.64908268	0.4794140412	80.9779742728	0.3805886345	31.0062809826	0.3418149337	20.8847203692	0.2806268115	14.2413555446
0.7695849793	4254.1241751139	0.5546383969	696.8320792921	0.4469592071	169.6857783353	0.3892337642	93.9005719593	0.344914608	50.7413819742
		0.6207864425	6133.2449027098	0.5474439544	1226.001409491	0.4655655296	532.1532341216	0.4566204962	263.7561507819
				0.7637048975	17271.9375778519	0.6107696402	4683.3937018005	0.5663046247	2146.211201895
						0.784623916	93277.7129340798	0.6253574036	26744.590748687
								0.6932526975	348322.670028861

Table F.1 a_i and t_i optimal values for I_1 cost function.

$m = 1$		$m = 2$		$m = 3$		$m = 4$		$m = 5$	
a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i
0.9384724434	1.4300340551	0.5470597552	0.6666835275	0.430797005	0.4521461414	0.3714051613	0.3505056162	0.3335736291	0.2904610289
		0.8449767491	8.3424872407	0.5319402016	3.0597097311	0.4221306386	1.7525741335	0.3629331173	1.203691574
				0.8046471493	36.769402335	0.5248827638	11.6528756138	0.4197252519	5.9370324806
						0.7814317902	136.8864124598	0.520201698	39.1450598115
								0.7661038702	452.8226228869

$m = 6$		$m = 7$		$m = 8$		$m = 9$		$m = 10$	
a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i	a_i	\tilde{t}_i
0.3065928563	0.2504309713	0.2869667584	0.2229567355	0.2695926115	0.1998084724	0.2560766303	0.1826244466	0.2467020831	0.1711374102
0.3243480187	0.9103056758	0.297777436	0.7379950193	0.2751628513	0.6094217589	0.2580812359	0.5231166336	0.2464749444	0.4695384557
0.3615932545	3.7204976994	0.3249218804	2.6583099103	0.2954155007	1.9746292865	0.2739535236	1.5665809982	0.2597178682	1.3336047234
0.418122689	18.2727422761	0.3631687423	10.9237321521	0.3228159111	7.0527307887	0.2950977014	5.0639463936	0.2773405882	4.038729849
0.5168085735	119.760302387	0.420482447	54.149026921	0.3602702642	28.6942745173	0.3224941082	18.0664363914	0.2995010019	13.2686834339
0.7551149413	1369.9016377844	0.5207711634	360.6375769122	0.4159293673	140.1890961279	0.3598005017	73.4054449448	0.3282822047	48.3505553197
		0.7554318595	4254.1243411105	0.5121568839	911.2555045811	0.4151331109	357.9494752882	0.3678821811	202.2013044128
				0.7402280446	10263.3419763251	0.5104760265	2319.7684648904	0.4276240337	1029.0899279619
						0.7348997012	25980.6116922192	0.5335800139	7177.8752909387
								0.7652665389	93277.7373733731

Table F.2 a_i and t_i optimal values for I_{2t} cost function.

Appendix G

FEM Discretization of the CFD-DEM equations

We provide in this appendix the developed expressions for the elemental matrices for both the Q-OSS and the Q-ASGS formulations.

G.1 OSS FEM discretization

Let us now perform the space discretization of the problem. Putting together the abstract expression of the variational problem, given by Eq. (4.50), with the bilinear form from Eq. (4.142), the RHS Eq. (4.143) and the stabilization terms from Eq. (4.144) we obtain

$$\begin{aligned} & \left(\frac{\partial}{\partial t} \mathbf{u}, \mathbf{v} \right) + (\mathbf{a} \cdot \nabla \mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) + (\alpha_f \nabla \cdot \mathbf{u}_f, q) \\ & + (\mathbf{u} \cdot \nabla \alpha_f, q) + \nu (\nabla^S(\mathbf{u}), \nabla \mathbf{v}) - \frac{2\nu}{3} ((\nabla \cdot \mathbf{u}), \nabla \cdot \mathbf{v}) \\ & + \sum_e \int_{\Omega_e} \tau_1 (\mathbf{a} \cdot \nabla \mathbf{u} + \nabla p) \cdot (-\mathbf{a} \cdot \nabla \mathbf{v} - \alpha_f \nabla q) \, d\Omega \\ & \quad + \sum_e \int_{\Omega_e} \tau_2 (\mathbf{u} \cdot \nabla \alpha_f + \alpha_f \nabla \cdot \mathbf{u}) (-\nabla \cdot \mathbf{v}) \, d\Omega \\ & = (\mathbf{f}, \mathbf{v}) + \left(-\frac{\partial \alpha_f}{\partial t}, q \right) + \int_{\Gamma_N} \mathbf{t}_N \cdot \mathbf{v} \, d\Gamma \\ & \quad + \sum_e \int_{\Omega_e} \tau_1 \boldsymbol{\pi}_{\text{mom}} \cdot ((\mathbf{a} \cdot \nabla) \mathbf{v} + \alpha_f \nabla q) \, d\Omega \\ & \quad \quad + \sum_e \int_{\Omega_e} \tau_2 \pi_{\text{mass}} (\nabla \cdot \mathbf{v}) \, d\Omega \end{aligned} \tag{G.1}$$

Let us now substitute in the standard finite element discretization in Eq. (4.20), taking one term at a time, to construct the elemental matrices. The first term is:

$$\left(\frac{\partial}{\partial t}\mathbf{u}_h, \mathbf{v}_h\right) = \int \frac{\partial}{\partial t}\mathbf{u}_h \cdot \mathbf{v}_h d\Omega \quad (\text{G.2})$$

Introducing the shape functions in the integral above yields the following expression. Here and in similar developments later on, the integration domain and the volume differential will often be suppressed for the sake of conciseness.

$$\int \frac{\partial}{\partial t}\mathbf{u}_h \cdot \mathbf{v}_h = \int N^b \frac{\partial}{\partial t}U_j^b N^a \delta_{ij} = \left(\int N^b N^a \delta_{ij}\right) \frac{\partial}{\partial t}U_j^b \quad (\text{G.3})$$

where the time derivative of the nodal values will in practice generate terms that will depend on the old values, which will be sent to the RHS.

The elemental contribution to be assembled to the global stiffness matrix is obtained by considering the contribution to the above integral due to the elemental domain and considering the nodes belonging to that domain and their corresponding shape functions. The elemental ‘mass’ matrix corresponding to the element e is as follows

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.4})$$

where

$$[\mathbf{M}]_{(ia)(jb)} = \int_{\Omega_e} N^b N^a \delta_{ij} d\Omega \quad (\text{G.5})$$

where the a and b indices run through the element nodes. Let us now look at the elemental contribution of the next term, that is, the convective term

$$\int (\mathbf{a}_h \cdot \nabla \mathbf{u}_h) \cdot \mathbf{v}_h = \int a_{h,k} \frac{\partial}{\partial x_k} N^b U_j^b N^a \delta_{ij} = \left(\int a_{h,k} \frac{\partial}{\partial x_k} N^b N^a \delta_{ij}\right) U_j^b \quad (\text{G.6})$$

The associated elemental matrix is given by

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.7})$$

where

$$[\mathbf{A}]_{(ia)(jb)} = \int_{\Omega_e} a_{h,k} \frac{\partial}{\partial x_k} N^b N^a \delta_{ij} d\Omega \quad (\text{G.8})$$

The next term is of the form $p \nabla \cdot \mathbf{v}$

$$\int (p_h \nabla \cdot \mathbf{v}_h) = \int N^b P^b \partial_j N^a \delta_{ij} = \left(\int N^b \frac{\partial}{\partial x_i} N^a \right) P^b \quad (\text{G.9})$$

The elemental matrix contribution due to the third term is thus

$$\begin{bmatrix} \mathbf{0} & \mathbf{H} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.10})$$

where

$$[\mathbf{H}]_{(ia)b} = - \int_{\Omega_e} N^b \frac{\partial}{\partial x_i} N^a d\Omega \quad (\text{G.11})$$

Let us now look at the fourth term.

$$\int \alpha_{fh} \nabla \cdot \mathbf{u}_h q_h = \int \alpha_{fh} \frac{\partial}{\partial x_j} N^b U_j^b N^a = \left(\int \alpha_{fh} \frac{\partial}{\partial x_j} N^b N^a \right) U_j^b \quad (\text{G.12})$$

so the associated elemental matrix is

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{D} & \mathbf{0} \end{bmatrix} \quad (\text{G.13})$$

with

$$[\mathbf{D}]_{a(jb)} = \int_{\Omega_e} \alpha_{fh} \frac{\partial}{\partial x_j} N^b N^a d\Omega \quad (\text{G.14})$$

The next term is developed as follows:

$$\int (\mathbf{u}_h \cdot \nabla \alpha_{fh}) q_h = \int N^b U_j^b \nabla \alpha_{fh} N^a = \left(\int \nabla \alpha_{fh} N^b N^a \right) U_j^b \quad (\text{G.15})$$

The associated elemental matrix contribution is therefore

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{G} & \mathbf{0} \end{bmatrix} \quad (\text{G.16})$$

with

$$[\mathbf{G}]_{a(jb)} = \int_{\Omega_e} \nabla \alpha_{fh} N^b N^a d\Omega \quad (\text{G.17})$$

The ‘gradient’ viscous terms are of the form (for the fluid equation)

$$\begin{aligned}
2 \int \nabla^S \mathbf{u}_h : \nabla \mathbf{v}_h &= \int \left(\frac{\partial}{\partial x_k} N^b U_j^b + \frac{\partial}{\partial x_j} N^b U_{f,k}^b \right) \left(\frac{\partial}{\partial x_k} N^a \delta_{ij} \right) = \\
&\left(\int \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a \delta_{ij} \right) U_j^b + \int \frac{\partial}{\partial x_j} N^b U_{f,k}^b \frac{\partial}{\partial x_k} N^a \delta_{ij} = \\
&\left(\int \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a \delta_{ij} \right) U_j^b + \left(\int \frac{\partial}{\partial x_i} N^b \frac{\partial}{\partial x_j} N^a \right) U_j^b
\end{aligned} \tag{G.18}$$

And the associated matrix (now including the particle phase contribution)

$$\nu \begin{bmatrix} \mathbf{S}_g & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{G.19}$$

with

$$[\mathbf{S}_g]_{(ia)(jb)} = \int_{\Omega_e} \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a \delta_{ij} + \frac{\partial}{\partial x_i} N^b \frac{\partial}{\partial x_j} N^a d\Omega \tag{G.20}$$

The ‘divergence’ type viscous terms are of the form

$$\begin{aligned}
\int \nabla \cdot \mathbf{u}_h \nabla \cdot \mathbf{v}_h d\Omega &= \int \frac{\partial}{\partial x_j} (N^b U_j^b) \left(\frac{\partial}{\partial x_k} N^a \delta_{ik} \right) d\Omega = \\
&\left(\int \frac{\partial}{\partial x_j} N^b \frac{\partial}{\partial x_i} N^a d\Omega \right) U_j^b
\end{aligned} \tag{G.21}$$

The associated matrix is (now including the particle phase contribution)

$$\frac{2\nu}{3} \begin{bmatrix} \mathbf{S}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{G.22}$$

with

$$[\mathbf{S}_d]_{(ia)(jb)} = - \int_{\Omega_e} \frac{\partial}{\partial x_j} N^a \frac{\partial}{\partial x_i} N^b d\Omega \tag{G.23}$$

Let us now expand the contribution of the stabilization terms. The first term that we will develop is the one of the form $(\mathbf{a} \cdot \nabla \mathbf{u})(\mathbf{a} \cdot \nabla \mathbf{v})$.

$$\int (\mathbf{a}_h \cdot \nabla \mathbf{u}_h)(\mathbf{a}_h \cdot \nabla \mathbf{v}_h) = \left(\int a_{h,k} \frac{\partial}{\partial x_k} N^b a_{h,l} \frac{\partial}{\partial x_l} N^a \delta_{ij} \right) U_j^b \tag{G.24}$$

The elemental contribution looks as follows

$$\begin{bmatrix} \mathbf{AA} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{G.25}$$

where

$$[\mathbf{AA}]_{(ia)(jb)} = \int_{\Omega_e} \tau_1 a_{h,k} \frac{\partial}{\partial x_k} N^b a_{h,l} \frac{\partial}{\partial x_l} N^a \delta_{ij} d\Omega \quad (\text{G.26})$$

Let us now look at the term structurally similar to $(\mathbf{a} \cdot \nabla \mathbf{u}) \cdot (\nabla q)$.

$$\begin{aligned} \int (\mathbf{a}_h \cdot \nabla \mathbf{u}_h) \cdot (\nabla q_h) d\Omega &= \int a_{h,k} \frac{\partial}{\partial x_k} u_{h,j} \frac{\partial}{\partial x_j} q_h d\Omega \\ &= \left(\int a_{h,k} \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_j} N^a d\Omega \right) U_j^b \end{aligned} \quad (\text{G.27})$$

The elemental contribution is as follows

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{AQ} & \mathbf{0} \end{bmatrix} \quad (\text{G.28})$$

where

$$[\mathbf{AQ}]_{a(jb)} = \int_{\Omega_e} \tau_1 \alpha_{fh} a_{h,k} \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_j} N^a d\Omega \quad (\text{G.29})$$

We next look at terms of the form $\nabla p \cdot (\mathbf{a} \cdot \nabla \mathbf{v})$

$$\begin{aligned} \int (\nabla p_h) \cdot (\mathbf{a}_h \cdot \nabla \mathbf{v}_h) &= \int \frac{\partial}{\partial x_j} p_h a_{h,k} \frac{\partial}{\partial x_k} v_{h,j} \\ &= \int \frac{\partial}{\partial x_j} N^b P^b a_{h,k} \frac{\partial}{\partial x_k} N^a \delta_{ij} = \left(\int \frac{\partial}{\partial x_i} N^b a_{h,k} \frac{\partial}{\partial x_k} N^a \right) P^b \end{aligned} \quad (\text{G.30})$$

The corresponding elemental contributions read

$$\begin{bmatrix} \mathbf{0} & \mathbf{GA} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.31})$$

where

$$[\mathbf{GA}]_{(ia)b} = \int_{\Omega_e} \tau_1 \frac{\partial}{\partial x_i} N^b a_{h,k} \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.32})$$

We now turn to the terms of the form $\nabla p \cdot \nabla q$.

$$\int \nabla p_h \cdot \nabla q_h = \int \frac{\partial}{\partial x_k} p_h \frac{\partial}{\partial x_k} q_h = \left(\int \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a \right) P^b \quad (\text{G.33})$$

The elemental matrix turns out to be

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{GG} \end{bmatrix} \quad (\text{G.34})$$

where

$$[\mathbf{GG}_f]_{ab} = \int_{\Omega_e} \tau_1 \alpha_{fh} \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.35})$$

We now turn to the set of terms multiplied by τ_2 . The first is of the form $(\mathbf{u} \cdot \nabla \alpha_f) \nabla \cdot \mathbf{v}$

$$\begin{aligned} \int (\mathbf{u}_h \cdot \nabla \alpha_{fh}) \nabla \cdot \mathbf{v}_h d\Omega &= \int N^b U_k^b \frac{\partial}{\partial x_k} \alpha_{fh} \frac{\partial}{\partial x_l} N^a \delta_{i,l} = \\ &= \int N^b U_k^b \frac{\partial}{\partial x_k} \alpha_{fh} \frac{\partial}{\partial x_i} N^a = \left(\int N^b \frac{\partial}{\partial x_j} \alpha_{fh} \frac{\partial}{\partial x_i} N^a \right) U_j^b \end{aligned} \quad (\text{G.36})$$

The corresponding elemental matrix is

$$\begin{bmatrix} \mathbf{AD} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.37})$$

where

$$[\mathbf{AD}]_{(ia)b} = \int_{\Omega_e} \tau_2 N^b \frac{\partial}{\partial x_j} \alpha_{fh} \frac{\partial}{\partial x_i} N^a d\Omega \quad (\text{G.38})$$

The last term is of the form $(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{v})$, which has already shown up in the viscous term. Therefore the elemental matrix contribution is

$$\begin{bmatrix} \mathbf{DD} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.39})$$

with

$$[\mathbf{DD}]_{(ia)(jb)} = - \int_{\Omega_e} \tau_2 \alpha_{fh} \frac{\partial}{\partial x_j} N^a \frac{\partial}{\partial x_i} N^b d\Omega \quad (\text{G.40})$$

In order to complete the matrix formulation, it is necessary to expand the RHS terms also. Let us now do this, again, taking it term by term and neglecting second-order derivatives within the element domains. The first type of terms to be added are of the form $\mathbf{c} \cdot \mathbf{v}$, where \mathbf{c} is a constant vector (with regards to integration).

$$\int \mathbf{c} \cdot \mathbf{v}_h = \int c_j N^a \delta_{ij} = \int c_i N^a \quad (\text{G.41})$$

The corresponding matrix contributions are

$$\begin{bmatrix} \mathbf{L}_v \\ \mathbf{0} \end{bmatrix} \quad (\text{G.42})$$

where

$$[\mathbf{L}_v]_{(ia)} = \int_{\Omega_e} (g_i + \frac{1}{\delta t} [\mathbf{u}^n]_i^a) N^a d\Omega \quad (\text{G.43})$$

The second standard Galerkin term is of the form q .

$$\int q = \int N^a \quad (\text{G.44})$$

Therefore we have the following matrix contribution:

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{L}_p \end{bmatrix} \quad (\text{G.45})$$

where

$$[\mathbf{L}_p]_{(a)} = - \int_{\Omega_e} \frac{\partial \alpha_{fh}}{\partial t} N^a d\Omega \quad (\text{G.46})$$

Let us now look at the stabilization terms. The first one is of the form $\mathbf{c} \cdot (\mathbf{a} \cdot \nabla \mathbf{v})$

$$\int \mathbf{c} \cdot (\mathbf{a}_h \cdot \nabla \mathbf{v}_h) = \int c_j a_{h,k} \frac{\partial}{\partial x_k} N^a \delta_{ij} = \int c_i a_{h,k} \frac{\partial}{\partial x_k} N^a \quad (\text{G.47})$$

Therefore, we have the following matrix terms

$$\begin{bmatrix} \mathbf{L}\mathbf{A} \\ \mathbf{0} \end{bmatrix} \quad (\text{G.48})$$

where

$$[\mathbf{L}\mathbf{A}]_{(ia)} = \int_{\Omega_e} \tau_1 (\pi_{v,i} + d_{v,i}) a_{h,k} \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.49})$$

The last term multiplying τ_1 is of the form $\mathbf{c} \cdot \nabla q$.

$$\int \mathbf{c} \cdot \nabla q_h = \int c_k \frac{\partial}{\partial x_k} N^a \quad (\text{G.50})$$

The corresponding matrix contribution is

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{LG} \end{bmatrix} \quad (\text{G.51})$$

where

$$[\mathbf{LG}]_a = \int_{\Omega_e} \tau_1 \alpha_{fh} (\pi_{v,k} + d_{v,k}) \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.52})$$

There only remains to develop the the term multiplying either τ_2 . It is of the form $\nabla \cdot \mathbf{v}$.

$$\int \nabla \cdot \mathbf{v}_h = \int \frac{\partial}{\partial x_j} N^a \delta_{ij} = \int \frac{\partial}{\partial x_i} N^a \quad (\text{G.53})$$

Therefore we have the following matrix contribution

$$\begin{bmatrix} \mathbf{LQ} \\ \mathbf{0} \end{bmatrix} \quad (\text{G.54})$$

where

$$[\mathbf{LQ}]_{(ia)} = \int_{\Omega_e} \tau_2 (\pi_p + d_p) \frac{\partial}{\partial x_i} N^a d\Omega \quad (\text{G.55})$$

We will additionally compute the expression corresponding to the Neumann boundary terms:

$$\int_{\Gamma_n} \mathbf{t}_N \cdot \mathbf{v}_h d\Gamma = \int_{\Gamma_n} \bar{t}_j N^a \delta_{ij} d\Gamma = \int_{\Gamma_n} \bar{t}_i N^a d\Gamma \quad (\text{G.56})$$

The matrix contribution is

$$\begin{bmatrix} \mathbf{L}\Gamma \\ \mathbf{0} \end{bmatrix} \quad (\text{G.57})$$

where

$$[\mathbf{L}\Gamma]_{(ia)} = \int_{\Gamma_n \cap \partial K} \bar{t}_i N^a d\Gamma \quad (\text{G.58})$$

Let us now assemble all these contributions as if only one element was present, so as to visualize the relationships between all the matrices.

$$\begin{aligned} \begin{bmatrix} \mathbf{M} + \mathbf{A} + \nu \mathbf{S}_g + \frac{2\nu}{3} \mathbf{S}_d & \mathbf{H} \\ \mathbf{D} + \mathbf{G} & \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} + \begin{bmatrix} \mathbf{AA} + \mathbf{AD} + \mathbf{DD} & \mathbf{GA} \\ \mathbf{AQ} & \mathbf{GG} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} \\ = \begin{bmatrix} \mathbf{L}_v + \mathbf{L}\Gamma \\ \mathbf{L}_p \end{bmatrix} + \begin{bmatrix} \mathbf{LA} + \mathbf{LQ} \\ \mathbf{LG} \end{bmatrix} \quad (\text{G.59}) \end{aligned}$$

where the first terms on both sides LHS are the standard Galerkin contribution, while the second terms are the contribution of the stabilization terms.

G.2 ASGS FEM discretization

Next we derive the matrix contributions corresponding to the ASGS discretization, for the case of quasi-static subscales. The main difference with respect to the OSS method is the presence of the dynamic contribution in the stabilization terms on the LHS. This term can be developed into

$$-\left(\frac{\partial}{\partial t}\mathbf{u}_h, \mathcal{L}^*(\mathbf{V}_h)\right)_\tau = -\sum_e \int_{\Omega_e} \tau_1 \frac{\partial}{\partial t} \mathbf{u}_h \cdot ((\mathbf{a}_h \cdot \nabla) \mathbf{v}_h + \alpha_{f,h} \nabla q_h) d\Omega \quad (\text{G.60})$$

Let us now develop the finite element contributions corresponding to this expression, taking each term one by one. The first term is of the form

$$\begin{aligned} \int \frac{\partial}{\partial t} \mathbf{u}_h \cdot (\mathbf{a}_h \cdot \nabla \mathbf{v}_h) &= \int N^b \frac{\partial}{\partial t} U_j^b a_{h,l} \frac{\partial}{\partial x_l} N^a \delta_{ij} = \\ &= \left(\int N^b a_{h,l} \frac{\partial}{\partial x_l} N^a \delta_{ij} \right) \frac{\partial}{\partial t} U_j^b \end{aligned} \quad (\text{G.61})$$

The elemental contribution is therefore

$$\begin{bmatrix} \mathbf{UA} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{G.62})$$

where

$$[\mathbf{UA}]_{(ia)(jb)} = -\frac{1}{\delta t} \int_{\Omega_e} \tau_1 N^b a_{h,l} \frac{\partial}{\partial x_l} N^a \delta_{ij} d\Omega \quad (\text{G.63})$$

The second term is of the form

$$\int \mathbf{u}_h \cdot \nabla q_h = \int \frac{\partial}{\partial x_k} N^b U_j^b \frac{\partial}{\partial x_k} N^a = \left(\int \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a \right) U_j^b \quad (\text{G.64})$$

The elemental contribution in this case is

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{UG} \end{bmatrix} \quad (\text{G.65})$$

where

$$[\mathbf{UG}]_{ab} = - \int_{\Omega_e} \tau_1 \frac{\partial}{\partial x_k} N^b \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.66})$$

Finally, the RHS stabilization terms must be defined. Their structure is similar to those present in the OSS method, where only the constant vectors differ. The ASGS analogous matrix contributions are next listed

$$\begin{bmatrix} \mathbf{LA}_{ASGS} \\ \mathbf{0} \end{bmatrix} \quad (\text{G.67})$$

where;

$$[\mathbf{LA}_{ASGS}]_{(ia)} = \int_{\Omega_e} \tau_1 f_i a_{h,k} \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.68})$$

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{LG}_{ASGS} \end{bmatrix} \quad (\text{G.69})$$

where; and

$$[\mathbf{LG}_{ASGS}]_a = \int_{\Omega_e} \tau_1 \alpha_{f,h} f_k \frac{\partial}{\partial x_k} N^a d\Omega \quad (\text{G.70})$$

$$\begin{bmatrix} \mathbf{LQ}_{ASGS} \\ \mathbf{0} \end{bmatrix} \quad (\text{G.71})$$

where

$$[\mathbf{LQ}_{ASGS}]_{(ia)} = - \int_{\Omega_e} \tau_2 \frac{\partial \alpha_f}{\partial t} \frac{\partial}{\partial x_i} N^a d\Omega \quad (\text{G.72})$$

Let us now assemble all the above elemental matrix contribution for the ASGS method

$$\begin{aligned} \begin{bmatrix} \mathbf{M} + \mathbf{A} + \nu \mathbf{S}_g + \frac{2\nu}{3} \mathbf{S}_d & \mathbf{H} \\ \mathbf{D} + \mathbf{G} & \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} + \begin{bmatrix} \mathbf{AA} + \mathbf{AD} + \mathbf{DD} + \mathbf{UA} & \mathbf{GA} + \mathbf{PV} \\ \mathbf{AQ} & \mathbf{GG} + \mathbf{UG} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} \\ = \begin{bmatrix} \mathbf{L}_v + \mathbf{L}\Gamma \\ \mathbf{L}_p \end{bmatrix} + \begin{bmatrix} \mathbf{LA}_{ASGS} + \mathbf{LQ}_{ASGS} \\ \mathbf{LG}_{ASGS} \end{bmatrix} \quad (\text{G.73}) \end{aligned}$$

Appendix H

Multicomponent theory fundamentals

H.1 The backward-coupled continuous problem

The theory of multicomponent flows can be understood as a generalization of (single-phase) continuum mechanics. The fundamental concepts defined in the latter are carried over in this generalization in a natural way. For this reason, a very concise summary of the most fundamental notions is next presented, also for the single-component case, so as to fix notation and terminology and facilitate the task of the reader unacquainted with the theory. In the exposition we mainly follow [109].

H.1.1 Single-component continuum theory: basic notions

Let the material body manifold describing the continuum be denoted as \mathcal{B} . A *configuration* is a differentiable map $\chi: \mathcal{B} \subset \mathbb{E}^3 \rightarrow \mathbb{E}^3$, where \mathbb{E}^3 represents the classical affine Euclidean space of continuum mechanics. The *motion* of \mathcal{B} is a differential mapping, $\bar{\chi}$, defined by $\mathcal{B} \times \mathbb{R} \ni (P, t) \mapsto \mathbf{x} \in \mathbb{E}^3$, defined, for each (P, t) , as

$$\mathbf{x} = \bar{\chi}(P, t) := \chi_t(P) \tag{H.1}$$

where the χ_t define a family of configurations parametrized by t , to be interpreted as the time passed since some initial time. The *Lagrangian description* of the body motion is the differential mapping, $\mathbf{x}: \mathbb{E}^3 \times \mathbb{R} \rightarrow \mathbb{E}^3$, obtained as:

$$\mathbf{x}(\mathbf{X}, t) = \chi_t(\chi_{\text{ref}}^{-1}(\mathbf{X})) \tag{H.2}$$

where χ_{ref}^{-1} is the inverse of a configuration, referred to as the *reference configuration*, which is usually taken as

$$\chi_{\text{ref}} = \chi_0 \quad (\text{H.3})$$

In this case, the reference configuration is to be interpreted as a mapping between the points in body \mathcal{B} and their spatial positions at the initial time.

Balance of mass

We suppose the existence of an absolutely continuous function ρ defined pointwise on \mathcal{B} , such that the mass measure is defined as

$$m(\chi_t(\mathcal{B}^*)) = \int_{\chi_t(\mathcal{B}^*)} \rho \, dV \quad (\text{H.4})$$

where \mathcal{B}^* is any V -measurable subset of \mathcal{B} .

The integral version of the postulate of balance of mass is

$$\frac{dm(\chi_t(\mathcal{B}^*))}{dt} = 0 \quad (\text{H.5})$$

The Dubois-Remond lemma implies its local counterpart, e.i.,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (\text{H.6})$$

Balance of momentum

Its integral version can be written as

$$\frac{d\mathcal{M}}{dt} = \mathcal{F}_b + \mathcal{F}_s \quad (\text{H.7})$$

where (omitting explicit mention to functional dependences and volume integration domains, both coinciding with $\chi_t(\mathcal{B}^*)$)

$$\mathcal{M} = \int \rho \mathbf{u} \, dV \quad (\text{H.8})$$

$$\frac{d\mathcal{M}}{dt} = \int \rho \frac{D\mathbf{u}}{Dt} \, dV \quad (\text{H.9})$$

$$\mathcal{F}_b = \int \rho \mathbf{b} \, dV \quad (\text{H.10})$$

$$\mathcal{F}_s = \int_{\partial\chi_t(\mathcal{B}^*)} \mathbf{t} \, dS = \int \nabla \cdot \boldsymbol{\sigma} \, dV \quad (\text{H.11})$$

The point-wise expression of the momentum balance is:

$$\rho \frac{D\mathbf{u}}{Dt} = \rho \mathbf{b} + \nabla \cdot \boldsymbol{\sigma} \quad (\text{H.12})$$

Balance of moment of momentum

Its integral version can be written as

$$\frac{d\mathcal{N}}{dt} = \mathcal{T}_b + \mathcal{T}_s \quad (\text{H.13})$$

where (for non-polar materials with no momentum surface sources) and ting explicit mention to functional dependences and volume integration domains, both coinciding with $\chi_t(\mathcal{B}^*)$)

$$\mathcal{N} = \int \rho \mathbf{x} \times \mathbf{u} \, dV \quad (\text{H.14})$$

$$\frac{d\mathcal{N}}{dt} = \int \rho \mathbf{x} \times \frac{D\mathbf{u}}{Dt} \, dV \quad (\text{H.15})$$

$$\mathcal{T}_b = \int \rho \mathbf{x} \times \mathbf{b} \, dV \quad (\text{H.16})$$

$$\mathcal{T}_s = \int_{\partial\chi_t(\mathcal{B}^*)} \mathbf{x} \times \mathbf{t} \, dS = \int \text{vec } \boldsymbol{\sigma} + \mathbf{x} \times \nabla \cdot \boldsymbol{\sigma} \, dV \quad (\text{H.17})$$

where $\text{vec } \boldsymbol{\sigma}$ is the axial vector of tensor $\boldsymbol{\sigma}$, i.e. $(\text{vec } \boldsymbol{\sigma})_i = \epsilon_{ijk} \sigma_{jk}$.

The point-wise expression of the moment of momentum balance is:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^T \quad (\text{H.18})$$

Balance of energy

Its integral version can be written as

$$\frac{d\mathcal{U}}{dt} = \mathcal{W}_b + \mathcal{W}_s \quad (\text{H.19})$$

where

$$\mathcal{U} = \int \rho \left(u + \frac{1}{2} \|\mathbf{u}\|^2 \right) dV \quad (\text{H.20})$$

$$\mathcal{W}_s = \int_{\partial\chi_t(\mathcal{B}^*)} \mathbf{n} \cdot (-\mathbf{q} + \mathbf{t} \cdot \mathbf{u}) dS \quad (\text{H.21})$$

$$\mathcal{W}_b = \int \rho (r + \mathbf{b} \cdot \mathbf{u}) dV \quad (\text{H.22})$$

The point-wise expression of the energy balance is:

$$\rho \left(\frac{\partial u}{\partial t} + \mathbf{u} \cdot \nabla \right) = -\nabla \cdot \mathbf{q} + \rho r + \mathbf{t} : \nabla \mathbf{u} \quad (\text{H.23})$$

H.1.2 Multi-component continuum theory: basic notions

A sequence of body manifolds \mathcal{B}_k , $k = 1, \dots, N$ (bodies) are considered. A motion is defined for each component:

$$\mathbf{x}_k = \bar{\chi}_k(\mathbf{X}, t) = \chi_{k,t}(\mathbf{X}) \quad (\text{H.24})$$

where each family of configurations $\chi_{k,t}$ is analogous to the corresponding families introduced for single-component motions.

Balance of mass

A mass measure is defined for each component:

$$m_k(\chi_{k,t}(\mathcal{B}^*)) = \int_{\chi_{k,t}(\mathcal{B}^*)} \overset{k}{\rho} dV \quad (\text{H.25})$$

where $\overset{k}{\rho}$ is density-like variable defined as

$$\overset{k}{\rho} = \alpha_k \rho_k \quad (\text{H.26})$$

where each α_k is the volume fraction, a continuous field with values in $[0, 1]$ and such that

$$\sum_k \alpha_k = 1 \quad (\text{H.27})$$

The integral version of the postulate of balance of mass is

$$\frac{dm(\chi_{k,t}(\mathcal{B}^*))}{dt} = 0 \quad (\text{H.28})$$

The Dubois-Remond lemma implies its local counterpart, e.i.,

$$\frac{\partial \alpha_k \rho_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}_k) = 0 \quad (\text{H.29})$$

Balance of momentum

Its integral version can be written as

$$\frac{d\mathcal{M}_k}{dt} = \mathcal{F}_k^b + \mathcal{F}_k^s \quad (\text{H.30})$$

where (omitting explicit mention to functional dependences and volume integration domains, both coinciding with $\chi_t(\mathcal{B}^*)$)

$$\mathcal{M} = \int \rho^k \mathbf{u}_k dV \quad (\text{H.31})$$

$$\mathcal{F}_k^b = \int \rho^k \mathbf{b}_k + \mathbf{F}_k dV \quad (\text{H.32})$$

$$\mathcal{F}_k^s = \int_{\partial\chi(\mathcal{B}_{t,k}^*)} \mathbf{t}^k dS = \int \nabla \cdot \boldsymbol{\sigma}^k dV \quad (\text{H.33})$$

where analogously to Eq. (H.26) we define

$$\mathbf{t}^k = \alpha_k \mathbf{t}_k \quad (\text{H.34})$$

$$\boldsymbol{\sigma}^k = \alpha_k \boldsymbol{\sigma}_k \quad (\text{H.35})$$

and where \mathbf{F}_k is the momentum exchange with other components. The point-wise expression of the momentum balance is then:

$$\alpha_k \rho_k \left(\frac{\partial \mathbf{u}_k}{\partial t} + \mathbf{u}_k \cdot \nabla \mathbf{u}_k \right) = \nabla \cdot \boldsymbol{\sigma}^k + \alpha_k \rho_k \mathbf{b}_k + \mathbf{F}_k \quad (\text{H.36})$$

Balance of moment of momentum

Its integral version is expressed as

$$\frac{d\mathcal{N}}{dt} = \mathcal{T}_b + \mathcal{T}_s \quad (\text{H.37})$$

where, for non-polar materials with no momentum surface sources) and omitting explicit mention to functional dependences and volume integration domains, both coinciding with $\chi_t(\mathcal{B}^*)$

$$\mathcal{N} = \int \rho^k \mathbf{x} \times \mathbf{u}_k \, dV \quad (\text{H.38})$$

$$\mathcal{T}_b = \int \rho^k \mathbf{x} \times \mathbf{b}_k + \boldsymbol{\psi}_k + \mathbf{x} \times \mathbf{F}_k \, dV \quad (\text{H.39})$$

$$\mathcal{T}_s = \int_{\partial\chi_t(\mathcal{B}^*)} \mathbf{x} \times \mathbf{t} \, dS = \int \text{vec } \overset{k}{\boldsymbol{\sigma}} + \mathbf{x} \times \nabla \cdot \overset{k}{\boldsymbol{\sigma}} \, dV \quad (\text{H.40})$$

The point-wise expression of the balance of moment of momentum is:

$$\overset{k}{\boldsymbol{\sigma}} = \overset{k}{\boldsymbol{\sigma}}^T \quad (\text{H.41})$$

Balance of energy

Its integral version can be written as

$$\frac{d\mathcal{U}_k}{dt} = \mathcal{W}_k^b + \mathcal{W}_k^s \quad (\text{H.42})$$

where

$$\mathcal{U}_k = \int \rho^k \left(u_k + \frac{1}{2} \|\mathbf{u}_k\|^2 \right) \, dV \quad (\text{H.43})$$

$$\mathcal{W}_k^s = \int_{\partial} \mathbf{n} \cdot \left(-\overset{k}{\mathbf{q}} + \overset{k}{\mathbf{t}} \cdot \mathbf{u}_k \right) \, dS \quad (\text{H.44})$$

$$\mathcal{W}_k^b = \int \rho^k \left(r_k + \mathbf{b}_k \cdot \mathbf{u}_k \right) + E_k + \mathbf{u}_k \cdot \mathbf{F}_k \, dV \quad (\text{H.45})$$

and where E_k is the source to component k due to the interaction with other components.

The point-wise expression of the energy balance is:

$$\alpha_k \rho_k \left(\frac{\partial u_k}{\partial t} + u_k \nabla \cdot \mathbf{u}_k \right) = -\nabla \cdot \mathbf{q}^k + \alpha_k \rho_k r_k + \overset{k}{\boldsymbol{\sigma}} : \nabla \mathbf{u}_k + E_k \quad (\text{H.46})$$

H.1.3 An important particularization of the balance equations

Let us consider there are only two phases: one is a continuous phase, composed of an incompressible continuous, incompressible fluid and the other is a dispersed phase, also assumed to be incompressible. Let us further assume that the body forces are the effect of a constant gravity field \mathbf{g} . Let the subindex f denote the continuous fluid phase while p denotes the particles phase. The mass, momentum and energy local balance equations may be written as

$$\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f \mathbf{u}_f) = 0 \quad (\text{H.47})$$

$$\frac{\partial \alpha_p}{\partial t} + \nabla \cdot (\alpha_p \mathbf{u}_p) = 0 \quad (\text{H.48})$$

$$\alpha_f \rho_f \left(\frac{\partial \mathbf{u}_f}{\partial t} + \mathbf{u}_f \cdot \nabla \mathbf{u}_f \right) = \nabla \cdot \overset{f}{\boldsymbol{\sigma}} + \alpha_f \rho_f \mathbf{g} - \mathbf{F}_{fp} \quad (\text{H.49})$$

$$\alpha_p \rho_p \left(\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p \right) = \nabla \cdot \overset{p}{\boldsymbol{\sigma}} + \alpha_p \rho_p \mathbf{g} + \mathbf{F}_{fp} \quad (\text{H.50})$$

$$\alpha_f \rho_f \left(\frac{\partial u_f}{\partial t} + u_f \nabla \cdot \mathbf{u}_f \right) = -\nabla \cdot \mathbf{q}^f + \alpha_f \rho_f r_f + \overset{f}{\boldsymbol{\sigma}} : \nabla \mathbf{u}_f + E_f \quad (\text{H.51})$$

$$\alpha_p \rho_p \left(\frac{\partial u_p}{\partial t} + u_p \nabla \cdot \mathbf{u}_p \right) = -\nabla \cdot \mathbf{q}^p + \alpha_p \rho_p r_p + \overset{p}{\boldsymbol{\sigma}} : \nabla \mathbf{u}_p + E_p \quad (\text{H.52})$$

where \mathbf{F}_{fp} denotes the contribution of momentum from the fluid to the particles. Obviously, this set of differential equations must be accompanied by the restriction:

$$\alpha_f + \alpha_p = 1 \quad (\text{H.53})$$

at all points in the domain of definition. The momentum equations may, equivalently, be written, following what is done in [182] as

$$\alpha_f \rho_f \left(\frac{\partial \mathbf{u}_f}{\partial t} + \mathbf{u}_f \cdot \nabla \mathbf{u}_f \right) = \alpha_f \nabla \cdot \overset{f}{\boldsymbol{\sigma}} + \alpha_f \rho_f \mathbf{g} - \mathbf{f}_{fp} \quad (\text{H.54})$$

$$\alpha_p \rho_p \left(\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p \right) = \nabla \cdot \overset{p}{\boldsymbol{\sigma}} + \alpha_p \nabla \cdot \overset{f}{\boldsymbol{\sigma}} + \alpha_p \rho_p \mathbf{g} + \mathbf{f}_{fp} \quad (\text{H.55})$$

where a *buoyancy force* has now been subtracted from \mathbf{F}_{fp} , such that $\mathbf{f}_{fp} := \mathbf{F}_{fp} - \alpha_p \nabla \cdot \overset{f}{\boldsymbol{\sigma}}$ may be referred to *hydrodynamic force density*. Derksen and Sundaresan [97] have called this term the *local fluid-particle interaction force per unit volume*, as it is caused by effects of a small scale nature, as opposed to the larger-scale varying term $\alpha_p \nabla \cdot \overset{f}{\boldsymbol{\sigma}}$.

Eqs.(H.47) and (H.52) are sometimes taken as an axiom (postulational approach, see, e.g. [359]), but they may also be obtained from single-phase continuum balance equations for each phase through averaging [109]. Several averaging methods have been devised for this matter, including time [179], volume [8] and ensemble averages [109], among others. All these approaches arrive at equations that conform to these (although the physical interpretations of the quantities involved are of course quite diverse). In order for the presented equations to define (along a suitable set of boundary conditions) a well-posed problem, closures (extra equations) must be provided for several of their terms, since the present number of scalar equations ($1 + 1 + 3 + 3 + 1 + 1 = 11$) is less than the number of scalar variables (33).

H.1.4 Filtering of a particle-laden flow

In this subsection we slightly generalize the equations in Anderson and Jackson [8] to obtain the analogous relations for a more general averaging procedure. We recast the problem in the language of filters, common in the field of turbulence modelling [303], unifying space and time averages in the same framework. We will focus on the problem of isotropic filters, ignoring for the moment the problem of boundaries, although this is also the case in [8].

The filtering operation is the convolution of a kernel function with the field that one wants to smooth out. We will from the start assume a kernel function defined in \mathbb{R}^4 is of the following form $G(t, \mathbf{x}) = h(|t|)g(\|\mathbf{x}\|)$, where h, g are positive, monotonically decreasing functions that vanish at infinity. A necessary condition for the operation to yield a meaningful description of the system, it is required that the smallest characteristic scales L, T of the flow are such that there is *separation of scales*, that is

$$\begin{aligned} a &\ll L_{\text{filter}} \ll L \\ \tau_p &\ll T_{\text{filter}} \ll T \end{aligned} \tag{H.56}$$

where L_{filter} and T_{filter} are the characteristic scales for the variation of g and h respectively. Anderson and Jackson [8] propose the following condition for the definition of L_{filter}

$$\int_{S_{L_{\text{filter}}}} g dS = \int_{\mathbb{R}^4 - S_{L_{\text{filter}}}} g dS = \frac{1}{2} \tag{H.57}$$

where $S_{L_{\text{filter}}}$ is the ball centred at zero with radius L_{filter} . And, by analogy, one can take

$$\int_{-T_{\text{filter}}}^{T_{\text{filter}}} h dt = \int_{\mathbb{R}-[-T_{\text{filter}}, T_{\text{filter}}]} h dt = \frac{1}{2} \quad (\text{H.58})$$

as the condition that defines T_{filter} .

Our goal is the construction of averaged fields, such as the ones reviewed in Section H.1.2. We do so by applying the following filter to any tensorial quantity \mathbf{a} of arbitrary order

$$\langle \mathbf{a} \rangle_k(t, \mathbf{x}) = \int_I \int_{\Omega} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) I_k(\tau, \mathbf{y}) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} d\tau \quad (\text{H.59})$$

where k is either p (particles) or f (fluid); $I_k(\tau, \mathbf{y})$ is the indicator function of the spatial configuration of the component k . For instance, I_p is zero inside the particles and zero everywhere else. Note that for $h = \delta$, the Dirac delta function, we would recover the usual volume averaging, while $g = \delta$ would yield the usual time averaging.

This filter allows to define the volume fraction fields as

$$\alpha_k = \langle 1 \rangle_k(t, \mathbf{x}) \quad (\text{H.60})$$

Using the volume fractions, we can define the component-averaged quantities as follows

$$\frac{k}{\mathbf{a}} := \begin{cases} \frac{1}{\alpha_k} \langle \mathbf{a} \rangle_k & \text{if } \alpha_k > 0 \\ \mathbf{0} & \text{if } \alpha_k = 0 \end{cases} \quad (\text{H.61})$$

Note that we choose to define the component-averaged variables as derived quantities in the theory, rather than in [8], where they are taken as fundamental. This allows us to define the averaging procedure as a filter and avoids the undefinition of the averaged variables in the regions of the domain where the volume fractions vanish.

Now, in order to derive the conservation equations, we first derive two relations involving derivatives of the filtered quantities.

$$\begin{aligned}
\frac{\partial \langle \mathbf{a} \rangle_k(t, \mathbf{x})}{\partial t} &= \frac{\partial}{\partial t} \int_I \int_{\Omega} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) I_k(\tau, \mathbf{y}) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} d\tau = \\
&= \int_I \int_{\Omega_k(\tau)} \frac{\partial h(|t - \tau|)}{\partial t} g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} d\tau = \\
&= - \int_I \int_{\Omega_k(\tau)} \frac{\partial h(|t - \tau|)}{\partial \tau} g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} d\tau = \\
&= \left\langle \frac{\partial \mathbf{a}(t, \mathbf{x})}{\partial t} \right\rangle_k - \int_I \int_{\Omega_k(\tau)} \frac{\partial h(|t - \tau|) \mathbf{a}(\tau, \mathbf{y})}{\partial \tau} g(\|\mathbf{x} - \mathbf{y}\|) d\mathbf{y} d\tau = \\
&= \left\langle \frac{\partial \mathbf{a}(t, \mathbf{x})}{\partial t} \right\rangle_k - \int_I \frac{\partial}{\partial \tau} \int_{\Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} d\tau + \quad (\text{H.62}) \\
&\quad + \int_I \int_{\partial \Omega_k(\tau)} \mathbf{n} \cdot \mathbf{v}_k h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) dS d\tau = \\
&= \left\langle \frac{\partial \mathbf{a}(t, \mathbf{x})}{\partial t} \right\rangle_k - \int_{\Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) d\mathbf{y} \Big|_{\partial I} + \\
&\quad + \int_I \int_{\partial \Omega_k(\tau)} \mathbf{n} \cdot \mathbf{v}_k h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) dS d\tau \approx \\
&\approx \left\langle \frac{\partial \mathbf{a}(t, \mathbf{x})}{\partial t} \right\rangle_k + \int_I \int_{\partial \Omega_k(\tau)} \mathbf{n} \cdot \mathbf{v}_k h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) \mathbf{a}(\tau, \mathbf{y}) dS d\tau
\end{aligned}$$

where Ω_k is the subset of Ω where we find the component k and where in the third last equality we have used Reynolds' transport theorem and where the last approximate identity is true far from the boundaries. The second identity is (take a to be an arbitrary component of \mathbf{a})

$$\begin{aligned}
\frac{\partial \langle a \rangle_k(t, \mathbf{x})}{\partial x_j} &= \int_I \int_{\Omega_k(\tau)} h(|t - \tau|) \frac{\partial g(\|\mathbf{x} - \mathbf{y}\|)}{\partial x_j} a(\tau, \mathbf{y}) d\mathbf{y} d\tau = \\
&= \left\langle \frac{\partial a(t, \mathbf{x})}{\partial x_j} \right\rangle_k - \int_I \int_{\Omega_k(\tau)} h(|t - \tau|) \frac{\partial g(\|\mathbf{x} - \mathbf{y}\|) a(\tau, \mathbf{y})}{\partial y_j} d\mathbf{y} d\tau = \quad (\text{H.63}) \\
&= \left\langle \frac{\partial a(t, \mathbf{x})}{\partial x_j} \right\rangle_k - \int_I \int_{\partial \Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) n_j a(\tau, \mathbf{y}) dS d\tau
\end{aligned}$$

where in the last equality we have used Gauss' theorem.

The next step is filtering both sides of the equation of interest. Let us apply it to the mass conservation and momentum conservation equations. In order to do so, we must imagine that the values of all the fields extend throughout the domain (e.g. the fluid velocity is defined inside the particles). The value chosen for the fields in this extension is not important, since the indicator function weighs them with zero. A simple possibility is to set them to zero in

the extended areas. Having done this, we can equate the filtered versions of both sides of the conservation equation in question. Let us start with the mass conservation equation

$$\begin{aligned} 0 &= \left\langle \frac{\partial v_{k,j}}{\partial x_j} \right\rangle_k = \frac{\partial \langle v_{k,j} \rangle_k(t, \mathbf{x})}{\partial x_j} - \int_I \int_{\partial\Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) n_j v_{k,j}(\tau, \mathbf{y}) dS d\tau = \\ &= \frac{\partial \alpha_k}{\partial t} + \frac{\partial \langle v_{k,j} \rangle_k(t, \mathbf{x})}{\partial x_j} \end{aligned} \quad (\text{H.64})$$

where we have used Eq.(H.63) in the first equality and Eq.(H.62) in the second. Note that this expression can alternatively be written as

$$\frac{\partial \alpha_k}{\partial t} + \frac{\partial \alpha_k \overline{v_{k,j}}^k(t, \mathbf{x})}{\partial x_j} = 0 \quad (\text{H.65})$$

which has the form of Eq.(H.29) for constant density.

For the momentum equation, let us first, using Eq.(H.62) we have

$$\left\langle \frac{\partial v_{k,i}}{\partial t} \right\rangle_k = \frac{\partial \langle v_{k,i} \rangle_k(t, \mathbf{x})}{\partial t} + \int_I \int_{\partial\Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) n_j v_{k,i} v_{k,j}(\tau, \mathbf{y}) dS d\tau \quad (\text{H.66})$$

While using Eq.(H.63) we have

$$\left\langle \frac{\partial v_{k,i} v_{k,j}}{\partial x_j} \right\rangle_k = \frac{\partial \langle v_{k,i} v_{k,j} \rangle_k(t, \mathbf{x})}{\partial x_j} - \int_I \int_{\partial\Omega_k(\tau)} h(|t - \tau|) g(\|\mathbf{x} - \mathbf{y}\|) n_j v_{k,i} v_{k,j}(\tau, \mathbf{y}) dS d\tau \quad (\text{H.67})$$

Then, filtering the LHS of Section H.1.1 and assuming a constant density we obtain

$$\frac{\partial \langle v_{k,i} \rangle_k(t, \mathbf{x})}{\partial t} + \frac{\partial \langle v_{k,i} v_{k,j} \rangle_k(t, \mathbf{x})}{\partial x_j} \quad (\text{H.68})$$

The second term above is not defined in terms of the filtered variables $\langle v_{k,i} \rangle_k$ and it is convenient to decompose it further to isolate the unclosed quantities. In order to do so, the next step, following [8], is to decompose \mathbf{v}_k into a sum of an average and a sub-scale component, like this

$$\mathbf{v}_k = \overline{\mathbf{v}}_k + \mathbf{v}'_k \quad (\text{H.69})$$

In [8] this decomposition is performed at each point in the domain, and then the filtering operation is performed, leading to

$$\begin{aligned}
\langle v_{k,i} v_{k,j} \rangle_k &= \left\langle \left(\frac{k}{v_{k,i}} + v'_{k,i} \right) \left(\frac{k}{v_{k,j}} + v'_{k,j} \right) \right\rangle_k = \\
&= \left\langle \frac{k}{v_{k,i}} \frac{k}{v_{k,j}} \right\rangle_k + \left\langle v'_{k,i} \frac{k}{v_{k,j}} \right\rangle_k + \left\langle \frac{k}{v_{k,j}} v'_{k,i} \right\rangle_k + \langle v'_{k,i} v'_{k,j} \rangle_k \approx \\
&\approx \alpha_k \frac{k}{v_{k,i}} \frac{k}{v_{k,j}} + \langle v'_{k,i} v'_{k,j} \rangle_k
\end{aligned} \tag{H.70}$$

The approximation above is justified by taking the first averaged quantity out of the average and neglecting the second term above, based on the separation of scales.

Another option is to write

$$\begin{aligned}
&\langle v_{k,i} v_{k,j} \rangle_k(t, \mathbf{x}) \\
&= \int_I \int_{\Omega_k} h g v_{k,i}(\tau, \mathbf{y}) v_{k,j}(\tau, \mathbf{y}) d\mathbf{y} d\tau = \\
&= \int_I \int_{\Omega_k} h g \left(\frac{k}{v_{k,i}}(t, \mathbf{x}) + v''_{k,i}(t, \tau, \mathbf{x}, \mathbf{y}) \right) \left(\frac{k}{v_{k,j}}(t, \mathbf{x}) + v''_{k,j}(t, \tau, \mathbf{x}, \mathbf{y}) \right) d\mathbf{y} d\tau = \tag{H.71} \\
&= \alpha_k \frac{k}{v_{k,i}} \frac{k}{v_{k,j}}(t, \mathbf{x}) + \int_I \int_{\Omega_k(\tau)} h g v''_{k,i} v''_{k,j}(t, \tau, \mathbf{x}, \mathbf{y}) d\mathbf{y} d\tau \\
&= \alpha_k \frac{k}{v_{k,i}} \frac{k}{v_{k,j}}(t, \mathbf{x}) + \left\langle v''_{k,i} v''_{k,j}(t, \tau, \mathbf{x}, \mathbf{y}) \Big|_{\tau=t; \mathbf{y}=\mathbf{x}} \right\rangle_k
\end{aligned}$$

Where some of the dependencies have not been made explicit for the sake of conciseness. Now the approximations have been avoided by splitting the field into the *average evaluated at the target point*. The price to pay is the introduction of the field v'' , which depends on four variables and describes the difference of the unfiltered field with respect to the filtered field, where filter point and point of measure are independent.

In any case, applying the above decomposition to Eq.(H.68) we obtain

$$\frac{\partial \alpha_k \frac{k}{v_{k,i}}}{\partial t} + \frac{\partial \alpha_k \frac{k}{v_{k,i}} \frac{k}{v_{k,j}}}{\partial x_j} + \frac{\partial R_{k,ij}}{\partial x_j} \tag{H.72}$$

which, using Eq.(H.65), can be simplified to

$$\alpha_k \left(\frac{\partial \frac{k}{v_{k,i}}}{\partial t} + \frac{\partial \frac{k}{v_{k,i}} \frac{k}{v_{k,j}}}{\partial x_j} \right) + \frac{\partial R_{k,ij}}{\partial x_j} \tag{H.73}$$

where $R_{k,ij}$ are to be interpreted as the components of the second term in the RHS of equation Eq.(H.70) or Eq.(H.71), depending on the approximation one wishes to work with. Models for this quantity are to be provided to express it in terms of the averaged variables.

Now let us apply the filter to the RHS Eq. (H.12). The result is

$$\rho_k \langle b_i \rangle_k + \frac{\partial \langle \sigma_{ij} \rangle_k}{\partial x_j} - \int_I \int_{\partial \Omega_k} h g n_j \sigma_{ij} dS d\tau \quad (\text{H.74})$$

The first term above offers no difficulties, especially when taking \mathbf{b} as the constant acceleration due to gravity. We can apply Eq.(H.63) for the second term, which yields

$$\begin{aligned} \int_I \int_{\partial \Omega_k} h g n_j \sigma_{ij} dS d\tau &= \int_I \int_{\Omega_i} \frac{\partial h g \sigma_{ij}}{\partial y_j} d\mathbf{y} d\tau = \\ &= \int_I \int_{\Omega_i} \frac{\partial \sigma_{ij}}{\partial y_j} h g d\mathbf{y} d\tau + \int_I \int_{\Omega_i} \frac{\partial h g}{\partial y_j} h g d\mathbf{y} d\tau = \\ &= \int_I \int_{\Omega_i} \frac{\partial \sigma_{ij}}{\partial y_j} h g d\mathbf{y} d\tau - \int_I \int_{\Omega_i} \frac{\partial h g}{\partial x_j} \sigma_{ij} d\mathbf{y} d\tau = \\ &= \int_I \int_{\Omega_i} \frac{\partial \sigma_{ij}}{\partial y_j} h g d\mathbf{y} d\tau - \frac{\partial}{\partial x_j} \int_I \int_{\Omega_i} h g \sigma_{ij} d\mathbf{y} d\tau = \\ &= \int_I \int_{\Omega_i} \frac{\partial \sigma_{ij}}{\partial y_j} h g d\mathbf{y} d\tau - \frac{\partial (1 - \alpha_k) \bar{\sigma}_{ij}^l}{\partial x_j} \end{aligned} \quad (\text{H.75})$$

Now, we may identify the first term as the average force per unit volume on the particles. The second term can be approximated as

$$\frac{\partial (1 - \alpha_k) \bar{\sigma}_{ij}^l}{\partial x_j} \approx \frac{\partial (1 - \alpha_k) \bar{\sigma}_{ij}^k}{\partial x_j} \quad (\text{H.76})$$

That is, by equating the average stress due to the interface force exchange on the particles with that on the fluid. Note that this is reasonable taking into account the small size of the particles. Note that this reasoning concentrates the approximation step in this single last operation, in contrast to what is done in [8]. The final form of the filtered momentum balance is (using Eqs. (H.73)–(H.76))

$$\alpha_k \rho_f \left(\frac{\partial v_{k,i}^k}{\partial t} + \frac{\partial v_{k,i}^k v_{k,j}^k}{\partial x_j} \right) = \frac{\partial \tilde{\sigma}_{k,ij}}{\partial x_j} - \alpha_p \mathbf{f}_p + \alpha_f \rho_f \bar{b}_j^k \quad (\text{H.77})$$

where

$$\tilde{\sigma} = \bar{\sigma} - \mathbf{R}_k \quad (\text{H.78})$$

and where \mathbf{f}_p is the average force per unit volume felt by the disperse phase due to the surface momentum exchange with the continuous phase. Both quantities will need to be modelled, since no analytic closed model exists. Ignoring the stress tensor for a moment, note that is of the form of Eq.(H.49) as is readily seen after identifying the subindex k with f , $\alpha_p \mathbf{f}$ with \mathbf{F} and removing the averaging notation. However, the same analogy does not appear to work with the stress tensor, since one would expect to identify $\tilde{\boldsymbol{\sigma}}_k$ with $\boldsymbol{\sigma}^k$. Though, formally, the analogy works rather with it corresponding to $\alpha_k \boldsymbol{\sigma}^k = \boldsymbol{\sigma}_k$ above.

In reality, the difference appears from an alternative decomposition of the different contributions. Note that if one takes Eq.(H.74) and reinterprets the second term in this equation as \mathbf{F} , one obtains a perfect term-to-term match with Eq.(H.49). However this partition is arguably less consistent than the former. This is because this second interpretation is using the filter with the *surface contributions* and so one is mixing volume averages with surface averages. In the former interpretation it is the volume-averaged momentum exchange between phases that is used, which generates the two terms in Eq.(H.75). The first of these two terms is the consistent (volume-averaged) interpretation of the momentum exchange between phases, while the second is absorbed, after the single approximation Eq.(H.76) into the second term of Eq.(H.49), as it is of the same form. The nature of this difference has not been always recognized as such. See, for example, the oft-cited paper [355]. See also [182] for a well-exposed argument about the often wrongly interpreted differences between two-phase models for the balance equations.

In the end, it will be necessary to provide closures for both the stress tensor and the momentum inter-phase exchange terms, so the relative merit of each decomposition will depend on the adequacy of the closure model provided. Thanks to the analysis of the averaging procedure, we now have a more explicit meaning for each of the referred terms.