CrossMark

JFM PAPERS

Special JFM volume in celebration of the
George K. Batchelor centenary

Sparse identification of multiphase turbulence closures for coupled fluid–particle flows

S. Beetham[1],†, **R.O. Fox**[2] **and J. Capecelatro**[1,3]

[1]Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA

[2]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50011, USA

[3]Department of Aerospace Engineering, University of Michgian, Ann Arbor, MI 48109, USA

In this work, model closures of the multiphase Reynolds-averaged Navier–Stokes (RANS) equations are developed for homogeneous, fully developed gas–particle flows. To date, the majority of RANS closures are based on extensions of single-phase turbulence models, which fail to capture complex two-phase flow dynamics across dilute and dense regimes, especially when two-way coupling between the phases is important. In the present study, particles settle under gravity in an unbounded viscous fluid. At sufficient mass loadings, interphase momentum exchange between the phases results in the spontaneous generation of particle clusters that sustain velocity fluctuations in the fluid. Data generated from Eulerian–Lagrangian simulations are used in a sparse regression method for model closure that ensures form invariance. Particular attention is paid to modelling the unclosed terms unique to the multiphase RANS equations (drag production, drag exchange, pressure strain and viscous dissipation). A minimal set of tensors is presented that serve as the basis for modelling. It is found that sparse regression identifies compact, algebraic models that are accurate across flow conditions and robust to sparse training data.

Key words: multiphase flow, particle/fluid flow, turbulence modelling

1. Introduction

Many natural and industrial processes involve the flow of solid particles, liquid droplets or gaseous bubbles whose dynamical evolution and morphology are intimately coupled with a carrier fluid. A peculiar behaviour of disperse multiphase flows is their ability to give rise to large-scale structures (hundreds to thousands of times the size of individual particles), from dense clusters to nearly particle-free voids (see figure 1). The emergence of spatial segregation in particles can be attributed to a number of factors, e.g. due to dissipation during inelastic collisions (Hopkins & Louge 1991; Goldhirsch & Zanetti 1993), viscous damping by the fluid (Wylie & Koch 2000), preferential concentration
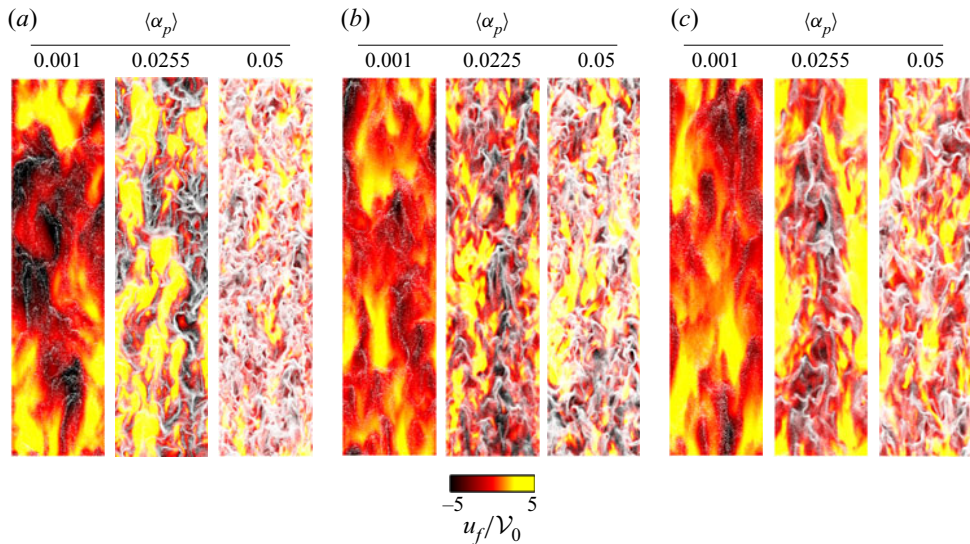
Figure 1. Instantaneous snapshots of fully developed CIT at statistical steady state. A slice at the centreline in the x–y plane is shown, with particle position (white) and normalized vertical fluid velocity $u_f/\mathcal{V}_0$ (colour): (*a*) $Ar = 1.80$; (*b*) $Ar = 5.40$; and (*c*) $Ar = 18.0$.

of particles by the background turbulence (Eaton & Fessler 1994) and instabilities that arise due to interphase coupling (Glasser, Sundaresan & Kevrekidis 1998; Agrawal *et al.* 2001; Capecelatro, Desjardins & Fox 2014, 2015). Such large-scale heterogeneity can effectively 'demix' the underlying flow, reducing contact between the phases and resulting in enormous consequences at scales much larger than the size of individual particles, such as hindered heat and mass transfer (Agrawal *et al.* 2013; Miller *et al.* 2014; Pouransari & Mani 2017; Beetham & Capecelatro 2019; Guo & Capecelatro 2019).

The focus of the present work is on modelling disperse multiphase turbulence at length scales much larger than the particle diameter. Such flows can be categorized into two broad classes: (i) flows where the turbulence originates in the continuous phase with the discrete phase modulating the small scales of the turbulence (see e.g. Balachandar & Eaton (2010), and references therein); and (ii) flows where the turbulence arises due to the coupling between the discrete and continuous phases. The former is mainly focused on how the discrete phase modifies the classical turbulence structures seen in single-phase flows. The latter is the focus of the present work and can be observed in gas–particle flows when the mass of the particles is of the same order or greater than that of the gas phase, or in bubbly flows when the bubble volume fraction is high enough to lead to buoyancy-driven instabilities (e.g. Besnard, Harlow & Rauenzhan 1987; Rzehak & Krepper 2013; Ma, Lucas & Bragg 2020).

To date, turbulent particle-laden flows are most often discussed in the dilute limit where the fluid-phase turbulence interacts with inertial particles without significant feedback from the particles. It is well established that dilute suspensions of heavy particles in isotropic turbulence will preferentially concentrate in regions of high strain rate and low vorticity (Eaton & Fessler 1994). In the presence of a mean body force (e.g. gravity), particles have been observed to experience enhanced settling as a result of preferential sweeping, by which the particles tend toward regions of downward fluid motion when encountering vortical structures in the flow (e.g. Wang & Maxey 1993; Aliseda *et al.* 2002; Ferrante & Elghobashi 2003; Yang & Shy 2003; Bosse, Kleiser & Meiburg 2006).

When the particle concentration is sufficiently high, the background flow is largely controlled by interphase coupling. Under these conditions, particles tend to accumulate in regions of low vorticity, resembling preferential concentration that typically occurs in dilute particle-laden turbulence. However, in the dense limit the vorticity is generated in shear layers between highly concentration regions (clusters), unlike in classical preferential concentration where vorticity would exist even in the absence of the disperse phase (Capecelatro *et al.* 2015). Additional effects contribute to the settling velocity and spatial segregation of particles in denser suspensions when two-way coupling between the phases is non-negligible.

Seminal works by G. K. Batchelor have provided theoretical estimates describing the motion of collections of solid particles suspended in viscous flows (Batchelor 1972, 1982), in addition to important insights into the instabilities present in such systems. For example, Batchelor (1988) demonstrated that small rigid spheres falling under gravity will give rise to long-range hydrodynamic interactions that result in hindered settling (Batchelor 1972). In more recent studies, it was demonstrated that, at higher Reynolds numbers and particle concentrations, momentum exchange between the phases results in enhanced settling when the mean mass loading, φ, defined by the ratio of the specific masses of the particle and fluid phases, is of order one or larger (Capecelatro *et al.* 2015). In statistically homogeneous gravity-driven gas–solid flows, the average particle settling speed, $\mathcal{V}$, can be approximated as

$$\mathcal{V} = \mathcal{V}_0 + \langle u_f \rangle_p \qquad (1.1)$$

for Stokes flow (Capecelatro *et al.* 2015), where $\mathcal{V}_0 = \tau_p g$ is the terminal Stokes settling velocity of an isolated particle with τ_p the particle response time and g gravity. In this expression, the phase-averaged fluid velocity, $\langle u_f \rangle_p = \langle \alpha_p u_f \rangle / \langle \alpha_p \rangle$, is sometimes referred to as the velocity seen by the particles, where u_f is the local fluid velocity aligned with gravity, α_p is the local particle volume fraction and angled brackets denote a spatial and temporal average. At sufficient mass loading, fluctuations in particle concentration can generate and sustain fluid-phase turbulence (as shown in figure 1), referred to here as cluster-induced turbulence (CIT). Because clusters entrain the carrier phase, u_f and α_p are often highly correlated, resulting in $\mathcal{V} > \mathcal{V}_0$.

Due to the breadth of length and time scales present in turbulent fluid–particle mixtures, accurate modelling of industrial and environmental flows remains challenging. Thus, the Reynolds-averaged Navier–Stokes (RANS) equations are the workhorse of industry to inform engineering designs and decisions. Because of the importance of the multiphase physics present in large-scale systems, developing multiphase RANS closures that are accurate under relevant conditions is critically important.

To date, multiphase turbulence models have largely relied upon extensions to single-phase models (e.g. Sinclair & Jackson 1989; Dasgupta, Jackson & Sundaresan 1994; Sundaram & Collins 1994; Cao & Ahmadi 1995; Dasgupta, Jackson & Sundaresan 1998; Cheng *et al.* 1999; Zeng & Zhou 2006; Jiang & Zhang 2012; Rao *et al.* 2012) that were derived directly from the Navier–Stokes equations. It should be noted, however, that multiphase turbulence does share some commonalities with single-phase flows, especially with variable-density turbulence. For example, multiphase flows subject to mean shear can develop velocity fluctuations that strongly modify the mean velocity profiles and transport properties of the flow (Capecelatro, Desjardins & Fox 2018). Single-phase, variable-density flows subject to Rayleigh–Taylor instabilities (see, e.g. Johnson & Schilling 2011; Zhou 2017) develop velocity and density fluctuations similar to those observed in heterogeneous bubbly flows (Mudde 2005). However, the main difference between disperse multiphase and variable-density flows is that the former

has a separate velocity for each phase while the latter has a single fluid velocity. Moreover, momentum coupling by particles introduces velocity fluctuations at small scales that further complicates the energy budget of turbulence. By introducing the slip velocity, i.e. the velocity difference between the discrete and continuous phases, additional dimensionless parameters, such as the Stokes number and mass loading, are needed to describe multiphase turbulence.

In contrast to modelling by analogy with single-phase flow, Fox (2014) developed the exact Reynolds-averaged equations for collisional fluid–particle flows. In that work, it was demonstrated that directly averaging the Navier–Stokes equations fails to capture important two-phase interactions. Instead, it was demonstrated that phase averaging the mesoscale (locally averaged) equations results in a set of equations that explicitly account for two-way coupling contributions. Capecelatro *et al.* (2015) further developed the Reynolds-averaged formulation of Fox (2014) to include transport equations for the volume-fraction variance, drift velocity and the separate components of the Reynolds stresses of each phase and particle-phase pressure tensor. While exact, it does lead to a large number of unclosed terms that require modelling, which is the focus of the present study.

Accurate modelling of the unclosed terms that remain predictive from dilute to dense regimes remains an outstanding challenge. Fox (2014) proposed closures of the phase-averaged (PA) terms based largely on single-phase turbulence models without extensive validation. Capecelatro, Desjardins & Fox (2016*b*) extended these models to account for near-wall effects in particle-laden channel flows. Agreement with the turbulence statistics obtained from simulation data was found to be satisfactory at first order (e.g. PA velocities) but less so at second order (e.g. PA turbulent kinetic energy). Innocenti *et al.* (2019) drew upon a probability-density-function approach, along with extensions from single-phase turbulence modelling (particularly in the fluid phase), showing satisfactory agreement for statistics up to second order. However, the model was restricted to relatively dilute flows. Due to the large parameter space associated with turbulent multiphase flows, a reliable modelling approach valid across two-phase flow regimes (e.g. dilute to dense limit) remains elusive.

Broadly speaking, extracting new models and understanding of physics from data has a long history in many diverse areas of science and engineering (see e.g. Jordan & Mitchell 2015). In the last decade, these data-driven techniques have been applied to turbulence modelling in several ways, including uncertainty prediction and quantification, model calibration and augmentation and the generation of entirely new models. Several recent works have utilized machine learning (neural networks are particularly popular; Milano & Koumoutsakos 2002; Lu 2010; Rajabi & Kavianpour 2012; Duraisamy & Durbin 2014; Duraisamy, Zhang & Singh 2015; Tracey, Duraisamy & Alonso 2015; Ling, Kurzawski & Templeton 2016; Ma, Lu & Tryggvason 2016; Bode *et al.* 2019; Liu & Fang 2019) in order to translate large amounts of experimental or computational data into model closures. Neural networks have shown relatively exceptional performance outside the region in which they were trained. As a departure from more traditional modelling techniques, these methods are inserted modularly, as a 'black box', into an existing flow solver. Thus, while they have displayed a high level of performance on a wide range of flow conditions, the closure does not satisfy the interpretability condition necessary for making physical inferences. Further, a large number of neural network approaches attempt to augment or correct existing models. However, as discussed above, in the context of multiphase flows appropriate existing models in which to augment do not exist.

Rather than relying on a best-fit strategy, as done in neural networks, Brunton, Proctor & Kutz (2016) developed a strategy based on sparse regression that identifies the underlying

functional form of the nonlinear physics by optimizing a coefficient matrix that acts upon a matrix of trial functions. While this method requires knowledge about the physics of the system under configuration (in order to make informed selections of the trial functions), it can be reasonably assumed that the modeller is not entirely naive. In fact, traditional modelling techniques have relied nearly exclusively on this notion. Schmelzer, Dwight & Cinnella (2020) and Beetham & Capecelatro (2020) recently extended the sparse identification framework of Brunton *et al.* (2016) to infer algebraic stress models for the closure of RANS equations. In Schmelzer *et al.* (2020) the models are written as tensor polynomials and built from a library of candidate functions. In Beetham & Capecelatro (2020) Galilean invariance of the resulting models are guaranteed through thoughtful tailoring of the feature space.

In this work, the sparse identification modelling framework of Beetham & Capecelatro (2020) is employed to develop multiphase closure models for homogeneous, gravity-driven gas–solid flows. Eulerian–Lagrangian simulations are performed across a range of Archimedes numbers and volume fractions to provide training data. The terms appearing in the multiphase RANS equations recently derived in Capecelatro *et al.* (2015) are extracted. We then build a minimally invariant basis set of tensors (i.e. a set of functional groups that serve as candidate terms in the desired model). Such basis sets are well established for single-phase turbulent flows (Pope 1975; Speziale, Sarkar & Gatski 1991; Gatski & Speziale 1993); however, an analogous basis has not yet been determined for multiphase flows. Using this basis and the sparse regression methodology, the compact functional form of the physics-based closures are inferred. As we consider exclusively statistically stationary and homogeneous systems, model realizability (Pope 2000) is left for future work.

2. System description

2.1. *Configuration under study*

In the present study, rigid spherical particles of diameter d_p and density ρ_p are suspended in an unbounded (triply periodic) domain containing an initially quiescent gas of density ρ_f and viscosity ν_f. Gravity g acts in the negative x-direction. As particles settle, they spontaneously form clusters. Due to two-way coupling between phases, particles entrain the fluid, generating turbulence therein. A frame of reference with the fluid phase is considered, such that the mean streamwise fluid velocity is null. Given the relative simplicity of the configuration, only a few non-dimensional groups arise. An important non-dimensional number is the Archimedes number, defined as

$$Ar = (\rho_p/\rho_f - 1)d_p^3 g/\nu_f^2. \tag{2.1}$$

Alternatively, a Froude number can be introduced to characterize the balance between gravitational and inertial forces, defined as $Fr = \tau_p^2 g/d_p$, where $\tau_p = \rho_p d_p^2/(18\rho_f \nu_f)$ is the particle response time. The Stokes settling velocity for an isolated particle is given by $\mathcal{V}_0 = \tau_p g$. From this a characteristic cluster length can also be estimated *a priori* as $\mathcal{L} = \tau_p^2 g$. To ensure the hydrodynamics is independent of the domain size, the simulation configurations are equal or larger than Case 4 reported in Capecelatro, Desjardins & Fox (2016*a*).

To sample the parameter space typical of turbulent fluidized bed reactors (Sun & Zhu 2019), the mean particle-phase volume fraction is varied in the range $0.001 \leq \langle \alpha_p \rangle \leq 0.05$ and the Archimedes number is varied in the range $1.8 \leq Ar \leq 18.0$ by adjusting gravity. Due to the large density ratios under consideration, the mean mass loading ranges from

Dimensional Quantities

$\mathcal{V}_0$	Stokes settling velocity	[m s^{-1}]	0.02	0.06	0.2
$\mathcal{L}$	Characteristic cluster length	[m]	5.0×10^{-4}	1.5×10^{-3}	5.0×10^{-3}
τ_p	Drag time	[s]		0.025	
ρ_p	Particle density	[kg m^{-3}]		1000	
d_p	Particle diameter	[m]		90×10^{-6}	
ρ_f	Fluid density	[kg m^{-3}]		1	
ν_f	Fluid viscosity	[m^2 s^{-1}]		1.8×10^{-5}	
g	Gravity	[m s^{-2}]	0.8	2.4	8.0

Non-dimensional Quantities

N_p	Number of particles		(610 370, 15 564 442, 30 518 514)	
$\langle \alpha_p \rangle$	Mean volume fraction $\times 10^{-2}$		(0.1, 2.55, 5.0)	
φ	Mean mass loading		(1.0, 26.2, 52.6)	
Fr	Froude number	5.6	16.7	55.6
Ar	Archimedes number	1.8	5.4	18.0

Computational Quantities

Domain size	[m]		$0.158 \times 0.038 \times 0.038$	
Grid size			$512 \times 128 \times 128$	
$L_x / \mathcal{L}$		316	105	32

Table 1. Summary of parameters for the configurations under consideration.

$O(10)$ to $O(10^2)$, and consequently two-way coupling between the phases is expected to be important. Here, angled brackets denote both a spatial and a temporal average (since the flow under consideration is triply periodic and statistically stationary in time). A list of relevant non-dimensional numbers and other important simulation parameters are summarized in table 1.

2.2. *Volume-filtered equations*

In this section, we present the volume-filtered Eulerian–Lagrangian equations used to formulate the Reynolds-averaged equations in § 3 and generate the simulation data that will be applied to the sparse regression methodology in § 4. The position and velocity of the ith particle is calculated according to Newton's second law

$$\frac{\mathrm{d}\boldsymbol{x}_p^{(i)}}{\mathrm{d}t} = \boldsymbol{v}_p^{(i)} \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{v}_p^{(i)}}{\mathrm{d}t} = \mathcal{A}^{(i)} + \boldsymbol{F}_c^{(i)} + \boldsymbol{g}, \quad (2.2a,b)$$

where $\boldsymbol{x}_p^{(i)}$ is the centre position of particle i and $\boldsymbol{v}_p^{(i)}$ is its velocity at time t and $\boldsymbol{g} = (-g, 0, 0)^{\mathrm{T}}$ is the acceleration due to gravity. The force due to inter-particle collisions, $\boldsymbol{F}_c$, is accounted for using a soft-sphere collision model originally proposed by Cundall & Strack (1979). Particles are treated as inelastic and frictional with a coefficient of restitution of 0.85 and coefficient of friction of 0.1 (Capecelatro & Desjardins 2013). Momentum exchange between the phases is given by

$$\mathcal{A}^{(i)} = \frac{F_d}{\tau_p} \left(\boldsymbol{u}_f - \boldsymbol{v}_p^{(i)} \right) - \frac{1}{\rho_p} \nabla p_f + \frac{1}{\rho_p} \nabla \cdot \boldsymbol{\sigma}_f, \quad (2.3)$$

where $\boldsymbol{u}_f$, p_f and $\boldsymbol{\sigma}_f$ are the fluid-phase velocity, pressure, and viscous-stress tensor evaluated at the particle location, respectively, and F_d is the non-dimensional drag correction of Tenneti & Subramaniam (2011) that takes into account local volume fraction

and Reynolds number effects given by

$$F_d(\alpha_f, Re_p) = \frac{1 + 0.15Re_p^{0.687}}{\alpha_f^2} + \alpha_f F_1\left(\alpha_f\right) + \alpha_f F_2\left(\alpha_f, Re_p\right), \qquad (2.4)$$

where the particle Reynolds number is defined as

$$Re_p = \frac{\alpha_f |\boldsymbol{u}_f - \boldsymbol{v}_p^{(i)}| d_p}{\nu_f}. \qquad (2.5)$$

The remaining two terms in (2.4) are given by

$$F_1(\alpha_f) = \frac{5.81\alpha_p}{\alpha_f^3} + \frac{0.48\alpha_p^{1/3}}{\alpha_f^4}, \qquad (2.6)$$

$$F_2(\alpha_f, Re_p) = \alpha_p^3 Re_p \left(0.95 + \frac{0.61\alpha_p^3}{\alpha_f^2}\right), \qquad (2.7)$$

where $\alpha_f = 1 - \alpha_p$ is the fluid-phase volume fraction.

To account for the presence of particles in the fluid phase without resolving the boundary layers around individual particles, a volume filter is applied to the incompressible Navier–Stokes equations (Anderson & Jackson 1967). This procedure replaces the point variables with smooth, locally filtered fields. These volume-filtered equations are given by

$$\frac{\partial \alpha_f}{\partial t} + \boldsymbol{\nabla} \cdot \left(\alpha_f \boldsymbol{u}_f\right) = 0 \qquad (2.8)$$

and

$$\frac{\partial \alpha_f \boldsymbol{u}_f}{\partial t} + \boldsymbol{\nabla} \cdot \left(\alpha_f \boldsymbol{u}_f \otimes \boldsymbol{u}_f\right) = -\frac{1}{\rho_f}\boldsymbol{\nabla} p_f + \boldsymbol{\nabla} \cdot \boldsymbol{\sigma}_f - \frac{\rho_p}{\rho_f}\alpha_p \mathcal{A} + \alpha_f \boldsymbol{g}, \qquad (2.9)$$

where $\mathcal{A}$ is the locally averaged momentum exchange term, evaluated at each Lagrangian particle and projected to the Eulerian mesh. The fluid-phase viscous-stress tensor is defined as

$$\boldsymbol{\sigma}_f = \nu_f \left[\boldsymbol{\nabla}\boldsymbol{u}_f + \left(\boldsymbol{\nabla}\boldsymbol{u}_f\right)^{\mathrm{T}} - \tfrac{2}{3}\boldsymbol{\nabla} \cdot \boldsymbol{u}_f \mathbb{I}\right], \qquad (2.10)$$

where $\mathbb{I}$ is the identity matrix.

The Eulerian–Lagrangian equations are solved using NGA (Desjardins *et al.* 2008), a fully conservative, low-Mach-number finite volume solver. A pressure Poisson equation is solved to enforce continuity via fast Fourier transforms in all three periodic directions. The fluid equations are solved on a staggered grid with second-order spatial accuracy and advanced in time with second-order accuracy using the semi-implicit Crank–Nicolson scheme of Pierce (2001). Lagrangian particles are integrated using a second-order Runge–Kutta method. Fluid quantities appearing in (2.2a,b) are evaluated at the position of each particle via trilinear interpolation. Particle data are projected onto the Eulerian mesh using the two-step filtering process described in Capecelatro & Desjardins (2013).

2.3. *Eulerian–Lagrangian training data*

The Eulerian–Lagrangian simulations were initialized with a random distribution of particles and run for approximately $100\tau_p$ until the flow reached a statistically

S. Beetham, R.O. Fox and J. Capecelatro

	$\langle\alpha_p\rangle$	$\dfrac{\sqrt{\langle\alpha_p'^2\rangle}}{\langle\alpha_p\rangle}$	$\dfrac{\langle u_f'''^2\rangle_f}{k_f}$	$\dfrac{\langle v_f'''^2\rangle_f}{k_f}$	$\dfrac{\langle u_p''^2\rangle_p}{k_p}$	$\dfrac{\langle v_p''^2\rangle_p}{k_p}$	$\dfrac{\langle u_p\rangle_p}{\mathcal{V}_0}$	$\dfrac{(\langle u_f\rangle_p+\tau_p^\star g)}{\mathcal{V}_0}$	St_η	Re_λ
$Ar = 1.8$	0.001	0.63	1.49	0.25	1.48	0.26	1.87	1.53	0.63	6.90
	0.0255	0.76	1.58	0.21	1.61	0.20	2.56	2.39	2.57	4.10
	0.05	0.74	1.51	0.24	1.52	0.24	2.46	2.34	2.97	3.29
$Ar = 5.4$	0.001	0.71	1.68	0.16	1.67	0.16	1.63	1.36	1.55	12.54
	0.0225	0.87	1.61	0.19	1.63	0.19	2.28	2.21	5.18	10.23
	0.05	0.84	1.56	0.22	1.57	0.21	2.12	2.07	6.29	8.76
$Ar = 18.0$	0.001	0.72	1.76	0.12	1.76	0.12	1.26	1.15	3.43	24.14
	0.0255	1.00	1.70	0.15	1.73	0.14	1.86	1.91	9.92	32.54
	0.05	0.98	1.64	0.18	1.68	0.16	1.81	1.83	12.97	35.52

Table 2. Statistically stationary Euler–Lagrange (EL) quantities for all nine training cases.

		Streamwise direction					
	$\langle\alpha_p\rangle$	PS	VD	DP	DE	VE	PE
$Ar = 1.8$	0.0100	−0.004	−0.007	0.03	−0.02	0	0
	0.0255	−0.189	−0.122	2.49	−2.75	−0.01	0
	0.5000	−0.406	−0.162	6.04	−6.87	−0.04	0
$Ar = 5.4$	0.0100	−0.021	−0.049	0.15	−0.138	0	0
	0.0255	−1.201	−0.482	15.18	−15.71	−0.06	0
	0.5000	−2.680	−0.709	39.67	−43.00	−0.29	0
$Ar = 18.0$	0.0100	−0.106	−0.264	0.81	−0.732	0	0
	0.0255	−9.988	−2.097	129.99	−131.1	−0.83	0
	0.5000	−22.056	−3.455	317.84	−329.3	−3.32	0

		Cross-stream directions				
	$\langle\alpha_p\rangle$	PS	VD	DE	VE	PE
$Ar = 1.8$	0.0100	0.002	−0.0004	−0.002	0	0
	0.0255	0.096	−0.0079	−0.171	0.002	0
	0.5000	0.219	−0.0140	−0.455	0.012	0.002
$Ar = 5.4$	0.0100	0.011	−0.0016	−0.01	0	0
	0.0225	0.616	−0.0293	−0.83	0.017	0.002
	0.5000	1.425	−0.0537	−2.323	0.083	0.007
$Ar = 18.0$	0.0100	0.053	−0.0052	−0.047	0	0
	0.0255	5.144	−0.1102	−5.686	0.183	0.009
	0.5000	12.020	−0.2305	−17.637	0.672	0.051

Table 3. Averaged terms for each contribution in the fluid-phase Reynolds-stress transport equations (3.2) and (3.3).

stationary state. At this point, statistics are accumulated over $50\tau_p$. Instantaneous snapshots of the streamwise fluid velocity and particle position of each case at steady state are shown in figure 1. It can immediately be seen that clusters of particles are generated and entrain the fluid downward. As a consequence of the frame of reference under consideration, the fluid flows upward in regions void of particles. Clusters are seen to become more distinct with increasing $\langle\alpha_p\rangle$. The effect of Ar on the flow field is less noticeable. As shown in table 2, the standard deviation in volume-fraction

914 A11-8

fluctuations $\langle \alpha_p'^2 \rangle^{1/2}$ increases with increasing Ar, with $\alpha_p' = \alpha_p - \langle \alpha_p \rangle$, indicating enhanced clustering. Perhaps less obvious, the volume-fraction fluctuations normalized by $\langle \alpha_p \rangle$ are maximum for the intermediate volume-fraction case ($\langle \alpha_p \rangle = 0.025$).

Because turbulence in CIT is driven by two-way coupling with the particle phase, there exist few characteristic scales that can be calculated *a priori*. However, the Taylor Reynolds number $Re_\lambda = u_{rms}\lambda/\nu_f$ and the Stokes number $St_\eta = \tau_p/\tau_\eta$, both of which must be computed *a posteriori*, provide insight on the resulting turbulence. Here, u_{rms} is the average root-mean-square velocity, $\lambda = \sqrt{15\nu_f/\varepsilon_f}u_{rms}$ is the Taylor micro-scale with ε_f the viscous dissipation rate and $\tau_\eta = \sqrt{\nu_f/\varepsilon_f}$ the Kolmogorov time scale. The values of these quantities for all nine cases are reported in table 2. It can be seen that high volume fractions correspond to larger Stokes numbers and lower values of Re_λ. Both St_η and Re_λ tend to increase when larger body forces are applied (i.e. larger Ar). Because fluid velocity fluctuations are generated by two-way coupling, $Re_\lambda > 0$ only when $\langle \alpha_p \rangle > 0$. That said, Re_λ is seen to decrease with increasing $\langle \alpha_p \rangle$ for $Ar = 1.8$ and 5.4. This is likely due to increased dissipation through drag exchange (see table 3). However, this reduction is less dramatic as Ar increases, and Re_λ is seen to be approximately constant at $Ar = 18$.

3. PA equations

In this section, we present the PA flow equations in which we seek to model the unclosed terms that arise. This system of equations has been previously derived (Capecelatro *et al.* 2015) and is extended here to take into account nonlinear drag effects due to F_d in (2.3).

The PA is analogous to Favre averaging of variable-density flows and is denoted by $\langle (\cdot) \rangle_p = \langle \alpha_p (\cdot) \rangle / \langle \alpha_p \rangle$. Fluctuations about the PA particle velocity are expressed as $u_p'' = u_p(x, t) - \langle u_p \rangle_p$, with $\langle u_p'' \rangle_p = 0$. This gives rise to the PA particle-phase turbulent kinetic energy (TKE), $k_p = \langle u_p'' \cdot u_p'' \rangle_p/2$. Here, u_p is the particle-phase velocity in an Eulerian frame of reference. It should be noted that $\langle u_p \rangle_p$ is equivalent to the average particle velocity $\langle v_p \rangle$ (with angled brackets here used to represent a particle average). Thus, $\langle u_p \rangle_p$ will be used throughout to characterize the mean settling velocity of the particle phase. In a similar fashion, the PA operator in the fluid phase is defined as $\langle (\cdot) \rangle_f = \langle \alpha_f (\cdot) \rangle / \langle \alpha_f \rangle$. Fluctuations about the PA fluid velocity are given by $u_f''' = u_f(x, t) - \langle u_f \rangle_f$. With this, the fluid-phase TKE is given by $k_f = \langle u_f''' \cdot u_f''' \rangle_f/2$.

For the statistically stationary and homogeneous flows considered herein, continuity implies $\langle \alpha_f \rangle$ is constant and the fluid-phase momentum equation reduces to $\langle u_f \rangle_f = 0$. In the particle phase, the only non-zero component of the averaged momentum equation is in the gravity-aligned direction (the x-direction in this case)

$$\frac{\partial \langle u_p \rangle_p}{\partial t} = \frac{1}{\tau_p^\star} \left(\langle u_f \rangle_p - \langle u_p \rangle_p \right) + \frac{1}{\rho_p} \left(\left\langle \frac{\partial \sigma_{f,xi}}{\partial x_i} \right\rangle_p - \left\langle \frac{\partial p_f}{\partial x} \right\rangle_p \right) + g, \qquad (3.1)$$

noting that for gas–solid flows, the terms involving $\sigma_{f,xi}$ and p_f are small enough to be neglected (Capecelatro *et al.* 2015). This implies that at steady state, $\langle u_p \rangle_p \approx \langle u_f \rangle_p + \tau_p^\star g$. Here, we incorporate the nonlinearities associated with drag in $\tau_p^\star = \tau_p/\langle F_d \rangle_p$, where $\langle F_d \rangle_p(\langle \alpha_f \rangle, \langle Re_p \rangle)$ is the nonlinear drag correction (2.4) defined using averaged flow arguments. This definition does not include the dependencies on drag covariance terms (i.e. $\langle u_f''' F_d'' \rangle_p$ and $\langle u_p'' F_d'' \rangle_p$), however, as shown in table 2, these terms have negligible contributions when describing particle settling, $\langle u_p \rangle_p$, and are thus neglected.

S. Beetham, R.O. Fox & J. Capecelatro

The transport equations for the fluid-phase Reynolds stresses can be reduced to two unique, non-zero components. In the streamwise direction this equation is given as

$$\frac{1}{2}\frac{\partial \langle u_f'''^2\rangle_f}{\partial t} = \underbrace{\frac{1}{\rho_f}\left\langle p_f \frac{\partial u_f'''}{\partial x}\right\rangle}_{pressure\ strain(PS)} - \underbrace{\frac{1}{\rho_f}\left\langle \sigma_{f,1i}\frac{\partial u_f'''}{\partial x_i}\right\rangle}_{viscous\ dissipation(VD)} + \underbrace{\frac{\varphi}{\tau_p^\star}\left(\langle u_f''' u_p''\rangle_p - \langle u_f'''^2\rangle_p\right)}_{drag\ exchange(DE)}$$

$$+ \underbrace{\frac{\varphi}{\tau_p^\star}\langle u_f'''\rangle_p \langle u_p\rangle_p}_{drag\ production\ (DP)} + \underbrace{\frac{\varphi}{\rho_p}\left\langle u_f''' \frac{\partial p_f'}{\partial x}\right\rangle_p}_{pressure\ exchange\ (PE)} - \underbrace{\frac{\varphi}{\rho_p}\left\langle u_f''' \frac{\partial \sigma_{f,1i}'}{\partial x_i}\right\rangle_p}_{viscous\ exchange\ (VE)}. \qquad (3.2)$$

Similarly, both cross-stream equations are given as

$$\frac{1}{2}\frac{\partial \langle v_f'''^2\rangle_f}{\partial t} = \underbrace{\frac{1}{\rho_f}\left\langle p_f \frac{\partial v_f'''}{\partial y}\right\rangle}_{PS} - \underbrace{\frac{1}{\rho_f}\left\langle \sigma_{f,2i}\frac{\partial v_f'''}{\partial x_i}\right\rangle}_{VD} + \underbrace{\frac{\varphi}{\tau_p^\star}\left(\langle v_f''' v_p''\rangle_p - \langle v_f'''^2\rangle_p\right)}_{DE}$$

$$+ \underbrace{\frac{\varphi}{\rho_p}\left\langle v_f''' \frac{\partial p_f'}{\partial y}\right\rangle_p}_{PE} - \underbrace{\frac{\varphi}{\rho_p}\left\langle v_f''' \frac{\partial \sigma_{f,2i}'}{\partial x_i}\right\rangle_p}_{VE}, \qquad (3.3)$$

where the drag production term no longer appears, since it is a gravity-driven phenomenon.

Due to the homogeneity of the flow and symmetry in the directions perpendicular to gravity (y and z directions in this configuration), the unique, non-zero PA Reynolds-stress transport equations in the particle phase are given as

$$\frac{1}{2}\frac{\partial \langle u_p''^2\rangle_p}{\partial t} = \underbrace{\left\langle \Theta \frac{\partial u_p''}{\partial x}\right\rangle_p}_{PS} - \underbrace{\left\langle \sigma_{p,1i}\frac{\partial u_p''}{\partial x_i}\right\rangle_p}_{VD} + \underbrace{\frac{1}{\tau_p^\star}\left(\langle u_f''' u_p''\rangle_p - \langle u_p''^2\rangle_p\right)}_{DE}$$

$$\times \underbrace{\frac{1}{\rho_p}\left\langle u_p'' \frac{\partial \sigma_{f,1i}'}{\partial x_i}\right\rangle_p}_{VE} - \underbrace{\frac{1}{\rho_p}\left\langle u_p'' \frac{\partial p_f'}{\partial x}\right\rangle_p}_{PE}. \qquad (3.4)$$

Similarly, the cross-gravity equations are both (due to symmetry and homogeneity) given as

$$\frac{1}{2}\frac{\partial \langle v_p''^2\rangle_p}{\partial t} = \underbrace{\left\langle \Theta \frac{\partial v_p''}{\partial y}\right\rangle_p}_{PS} - \underbrace{\left\langle \sigma_{p,2i}\frac{\partial v_p''}{\partial x_i}\right\rangle_p}_{VD} + \underbrace{\frac{1}{\tau_p^\star}\left(\langle v_f''' v_p''\rangle_p - \langle v_p''^2\rangle_p\right)}_{DE}$$

$$\times \underbrace{\frac{1}{\rho_p}\left\langle v_p'' \frac{\partial \sigma_{f,2i}'}{\partial y_i}\right\rangle_p}_{VE} - \underbrace{\frac{1}{\rho_p}\left\langle v_p'' \frac{\partial p_f'}{\partial y}\right\rangle_p}_{PE}, \qquad (3.5)$$

where Θ and σ_p are the granular temperature and the particle-phase viscous stress tensor, respectively (Capecelatro et al. 2015).

914 A11-10

Due to the high density ratio in gas–solid flows, PE and VE are often negligible (Capecelatro *et al.* 2015). Therefore, the overall kinetic energy balance for CIT includes DP, PS, VD and DE. As in single-phase turbulence, VD results from the resolved small-scale velocity fluctuations in the fluid phase. In contrast, the drag exchange terms involve (i) DE_2: viscous dissipation of unresolved fluid velocity fluctuations (in the viscous boundary layers around individual particles) and (ii) DE_1: energy transferred to the particles at the fluid–particle interface. Sundaram & Collins (1999) showed that the unresolved dissipation arises in the point-particle model due to the difference between the fluid velocity at the particle location and the particle velocity. Our previous work (Capecelatro *et al.* 2015) showed that the relative contribution of the interphase energy exchange ($DE_1/(DE_1+DE_2)$) is approximately 22 %. Further, the ratio of resolved to unresolved viscous dissipation was found to be less than 6 % and decreases with increasing *Ar*. Therefore we do not expect the unresolved viscous dissipation not captured in the Eulerian–Lagrangian model to impact the overall balance.

4. Closure modelling

4.1. *Sparse regression with embedded invariance*

The focus of this section is modelling the unclosed terms that appear in the fluid-phase Reynolds-stress equations (3.2) and (3.3). The data used to inform these closures, as discussed in § 2.3, are averaged after the flow has become statistically stationary in time. These values are summarized in table 3. In the streamwise direction, DP is mostly balanced by DE. The PS and VD contain fluid-phase residual contributions, while PE and VE contain contributions from both phases. These terms are small compared to DP and DE, but are not negligible in general. In the cross-stream direction, DE is mostly balanced by PS.

Each unclosed term is considered individually and models are learned using the sparse regression methodology described in Beetham & Capecelatro (2020) and summarized here. In this method, it is postulated that any tensor quantity $\mathbb{D}$ can be modelled using an invariant tensor basis, $\mathbb{T}$, and a set of ideal, sparse coefficients, $\hat{\beta}$,

$$\mathbb{D} = \mathbb{T}\hat{\beta}. \tag{4.1}$$

The ideal coefficients are determined by solving the optimization problem

$$\hat{\beta} = \min_{\beta} ||\mathbb{D} - \mathbb{T}\beta||_2^2 + \lambda||\beta||_1, \tag{4.2}$$

where β is a vector of coefficients that varies depending upon the choice of a user-specified sparsity parameter, λ and $|| \cdot ||_2^2$ and $|| \cdot ||_1$ represent the L-2 and L-1 norms, respectively. In the case of single-phase turbulence, this methodology can be used readily with previously derived minimally invariant basis sets (Hastie, Tibshirani & Friedman 2009). It is helpful to note that sparse regression is openly available in several software packages, including PySINDy (de Silva *et al.* 2020). However, to date an analogous basis has not yet been identified for multiphase flows. Due to the relative simplicity of the system under study (i.e. symmetry, homogeneity and stationarity), the parameters that may contribute to such a basis are limited to three tensors: the fluid-phase Reynolds-stress anisotropy tensor, $\hat{\mathbb{R}}_f$, the particle-phase Reynolds-stress anisotropy tensor, $\hat{\mathbb{R}}_p$, and the mean slip tensor, $\hat{\mathbb{U}}_r$ (see table 4). The mean slip tensor is defined as $\boldsymbol{U}_r = \boldsymbol{u}_r \otimes \boldsymbol{u}_r$, where $\boldsymbol{u}_r = \langle \boldsymbol{u}_p \rangle_p - \langle \boldsymbol{u}_f \rangle_f$ is the slip velocity vector. An important property of this vector is that in fully developed CIT it is always aligned with the direction of the body forcing (in this case gravity).

(I) Particle-phase anisotropic stress tensor	$\hat{\mathbb{R}}_p = \frac{\langle u_p''' u_p'''\rangle}{2k_p} - \frac{1}{3}\mathbb{I}$
(II) Fluid-phase anisotropic stress tensor	$\hat{\mathbb{R}}_f = \frac{\langle u_f''' u_f'''\rangle_f}{2k_f} - \frac{1}{3}\mathbb{I}$
(III) Slip velocity tensor	$\hat{\mathbb{U}}_r = \frac{U_r}{\mathrm{tr}(U_r)} - \frac{1}{3}\mathbb{I},$

Table 4. Second-order, symmetric, deviatoric tensors available to the multiphase RANS equations for modelling.

$$\mathcal{T}^{(1)} = \mathbb{I} \qquad \mathcal{T}^{(2)} = \hat{\mathbb{U}}_r$$
$$\mathcal{T}^{(3)} = \hat{\mathbb{U}}_r^2 \qquad \mathcal{T}^{(4)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f)^\dagger$$
$$\mathcal{T}^{(5)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_f)^\dagger \qquad \mathcal{T}^{(6)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_f^2)^\dagger$$
$$\mathcal{T}^{(7)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p)^\dagger \qquad \mathcal{T}^{(8)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p)^\dagger$$
$$\mathcal{T}^{(9)} = (\hat{\mathbb{R}}_f\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_p)^\dagger \qquad \mathcal{T}^{(10)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_f^2\hat{\mathbb{R}}_p)^\dagger$$
$$\mathcal{T}^{(11)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_p)^\dagger \qquad \mathcal{T}^{(12)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p\hat{\mathbb{U}}_r^2)^\dagger$$
$$\mathcal{T}^{(13)} = \hat{\mathbb{R}}_f \qquad \mathcal{T}^{(14)} = \hat{\mathbb{R}}_f^2$$
$$\mathcal{T}^{(15)} = \hat{\mathbb{R}}_p \qquad \mathcal{T}^{(16)} = \hat{\mathbb{R}}_p^2$$
$$\mathcal{T}^{(17)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_p)^\dagger \qquad \mathcal{T}^{(18)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_p)^\dagger$$
$$\mathcal{T}^{(19)} = (\hat{\mathbb{U}}_r\hat{\mathbb{R}}_p^2)^\dagger \qquad \mathcal{T}^{(20)} = (\hat{\mathbb{U}}_r^2\hat{\mathbb{R}}_p^2)^\dagger$$
$$\mathcal{T}^{(21)} = (\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p)^\dagger \qquad \mathcal{T}^{(22)} = (\hat{\mathbb{R}}_f^2\hat{\mathbb{R}}_p)^\dagger$$
$$\mathcal{T}^{(23)} = (\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p^2)^\dagger \qquad \mathcal{T}^{(24)} = (\hat{\mathbb{R}}_f^2\hat{\mathbb{R}}_p^2)^\dagger$$
$$\mathcal{S}^{(1)} = \mathrm{tr}(\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f^2\hat{\mathbb{R}}_p^2) \qquad \mathcal{S}^{(2)} = \mathrm{tr}(\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p^2) \qquad \mathcal{S}^{(3)} = \mathrm{tr}(\hat{\mathbb{U}}_r\hat{\mathbb{R}}_f\hat{\mathbb{R}}_p)$$
$$\mathcal{S}^{(4)} = Ar \qquad \mathcal{S}^{(5)} = \varphi \qquad \mathcal{S}^{(6)} = \langle\alpha_p\rangle$$

Table 5. Minimally invariant set of basis tensors and associated scalar invariants. Here, $(\cdot)^\dagger = (\cdot) + (\cdot)^\mathrm{T}$ denotes the tensor quantity added with its transpose.

Because the sparse regression methodology postulates the model to be a linear combination of the basis tensors, this implies that the basis tensors must take on the same properties as the quantity to be modelled. The four terms under consideration here are all symmetric and thus the basis tensors must also be symmetric. The three tensor quantities shown in table 4 are used in order to formulate a minimally invariant basis by following the procedure described in Spencer & Rivlin (1958). This set of tensors, along with six scalar invariants, denoted $\mathcal{S}^{(i)}$, by definition can exactly describe the Eulerian–Lagrangian data as shown in table 5. In the context of the sparse regression methodology, the ideal coefficients $\hat{\beta}$ may be constants or nonlinear functions of the scalar invariants, $\mathcal{S}^{(i)}$.

4.2. *Results and discussion*

Using the set of basis tensors defined in § 4.1, the sparse regression methodology is employed to identify closures for the terms appearing in the fluid-phase Reynolds-stress equations (3.2) and (3.3), based upon the Eulerian–Lagrangian data described in § 2.3. Since flow data are homogeneous in all three spatial directions and we consider time-averaged data, each case is zero-dimensional (i.e. a single value).

As seen in table 3, the contributions from VE and PE are either null, or relatively small even as mass loading is increased. For this reason, modelling efforts are directed toward
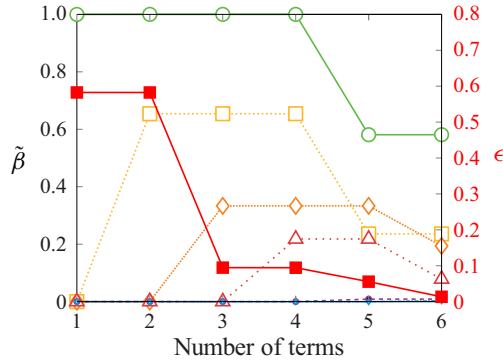
Sparse identification of multiphase turbulence closures

Figure 2. Normalized coefficients, $\tilde{\beta}$ (left axis) and associated model error, ϵ, (($\blacksquare$, lined solid square, red), right axis) for DP. The three-term and six-term models are described in (4.4) and (4.5), respectively. Terms 1–6, as denoted in (4.5), are represented as ($\bullet$, lined solid circle, light green), ($\blacksquare$, dotted square, light orange), ($\blacklozenge$, densely dotted diamond, orange), ($\blacktriangle$, loosely dotted triangle, red), ($\bullet$, dashed solid circle, purple), ($\blacklozenge$, densely dash-dotted diamond, light blue), respectively. These colours also correspond with figure 4.

the four remaining terms: DP, PS, VD and DE. Each is modelled separately, beginning with DP as it is the sole source of fluid-phase TKE in the absence of mean shear. As seen in (3.2), it is proportional to $\langle u_f''' \rangle_p$, which is zero in the absence of particle clusters.

4.2.1. *Drag production*

As input to the sparse regression algorithm, drag production is non-dimensionalized using the square of the PA particle velocity, $\langle u_p \rangle_p^2$ and the drag time, τ_p. Because drag production is symmetric and also contains zero off-diagonal components, the basis set was restricted to only include terms that are functions of $\hat{\mathbb{U}}_r$ and $\mathbb{I}$, which also exhibit this property. While the Reynolds stresses have null off-diagonal components for this particular configuration, this does not hold in a general sense.

During optimization, as λ is decreased, additional terms are added to the learned model and model error decreases (see figure 2), where model error is defined as

$$\epsilon = \frac{||\mathbb{D} - \mathbb{T}\hat{\beta}||_2^2}{||\mathbb{D}||_2^2}. \tag{4.3}$$

In examining the relationship between model error and model complexity, we observe that a significant reduction in error is achieved with three model terms and the error is drastically reduced when considering a six-term model. It is also notable that the most important terms to overall model performance appear in the models with lesser complexity and remain dominant as subsequent terms are added. This is indicated by the behaviour of the normalized coefficients $\tilde{\beta}$, given as $\hat{\beta}^{(p)}/\max \hat{\beta}^{(1)}$, where p denotes the number of terms in the model.

The resultant learned models with three terms and six terms are given, respectively, as

$$\mathcal{R}^{DP} = \frac{\langle u_p \rangle_p^2}{\tau_p} \left[\underbrace{1.11\varphi\hat{\mathbb{U}}_r}_{term\ 1} - \underbrace{0.73\varphi^{-2}\hat{\mathbb{U}}_r}_{term\ 2} + \underbrace{0.37\varphi\mathbb{I}}_{term\ 3} \right] \tag{4.4}$$
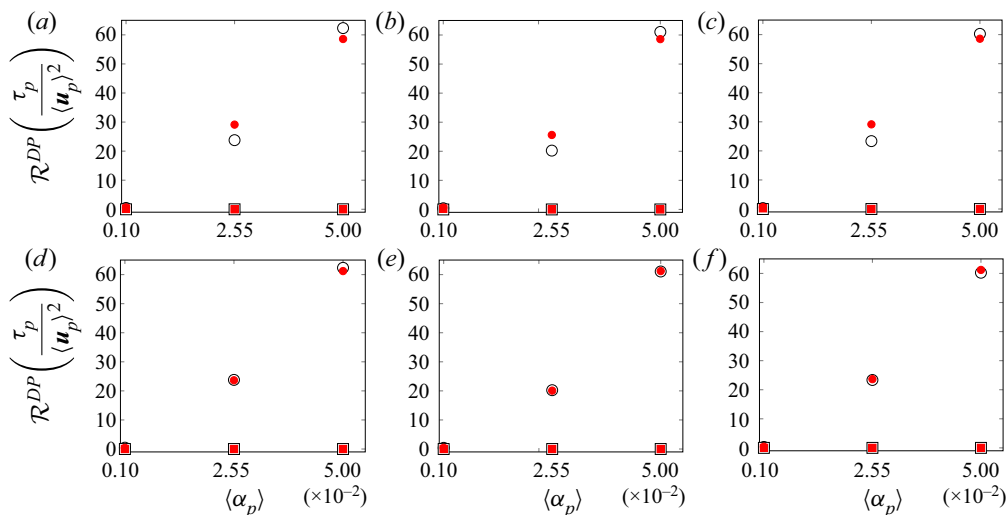
914 A11-13

Figure 3. Drag production obtained from Eulerian–Lagrangian results ($\square$, cross-stream component and $\circ$, streamwise components) and model prediction (($\blacksquare$, red), cross-stream component and ($\bullet$, red), streamwise components). The model corresponds to (4.5) with $\lambda = 0.01$. The associated model error is $\epsilon = 0.01$. (*a*) $Ar = 1.80$; (*b*) $Ar = 5.40$; (*c*) $Ar = 18.0$; (*d*) $Ar = 1.8$; (*e*) $Ar = 5.4$; and (*f*) $Ar = 18.0$.

and

$$\mathcal{R}^{DP} = \frac{\langle u_p \rangle_p^2}{\tau_p} \left[\underbrace{0.65\varphi\hat{\mathbb{U}}_r}_{term\ 1} - \underbrace{0.26\varphi^{-2}\hat{\mathbb{U}}_r}_{term\ 2} + \underbrace{0.22\varphi\mathbb{I}}_{term\ 3} - \underbrace{0.09\varphi^{-2}\mathbb{I}}_{term\ 4} + \underbrace{0.01\varphi^2\hat{\mathbb{U}}_r}_{term\ 5} + \underbrace{0.003\varphi^2\mathbb{I}}_{term\ 6} \right].$$
(4.5)

To illustrate the interplay between model complexity and interaction, we consider the highly accurate, six-term model (see figures 3(*d*)–3(*f*) and (4.5)) and the simpler, three-term model (see figures 3(*a*)–3(*c*) and (4.4)). In comparing the performance of these models, we observe that the general scaling and spread of the data are captured reasonably well with the three-term model, but that the complexity added in the six-term model makes smaller adjustments that drive down model error. As shown in figure 4, the accuracy of the three-term model is primarily centred on the streamwise component of drag production (see figure 4(*a*)); however, it over predicts the cross-stream components (see figure 4*c*). The six-term model, in turn, reduces overall model error by more accurately describing both components; however, this is most pronounced in the cross-stream direction (see figures 4(*b*) and 4(*d*)).

In addition to discovering compact, algebraic models, sparse regression is also robust to sparse training data (Beetham & Capecelatro 2020). To illustrate this, a model was discovered using a sparse training dataset corresponding to $(Ar, \langle \alpha_p \rangle) = [(1.8, 0.05), (5.4, 0.001), (18.0, 2.55)]$ and then tested using the remaining six cases. The resultant model is given as

$$\mathcal{R}^{DP} = \frac{\langle u_p \rangle^2}{\tau_p} \left[\left(0.36\varphi^{-1} + 0.05\varphi^2 - 4 \times 10^{-4}\varphi^3 \right) \hat{\mathbb{U}}_r + \left(0.01\varphi^2 + 0.21\varphi^{-1} \right) \mathbb{I} \right]$$ (4.6)

and shown compared with the trusted Eulerian–Lagrangian data in figures 5(*a*)–5(*c*). It is notable that the sparsely trained model achieves reasonable accuracy using a subset of the
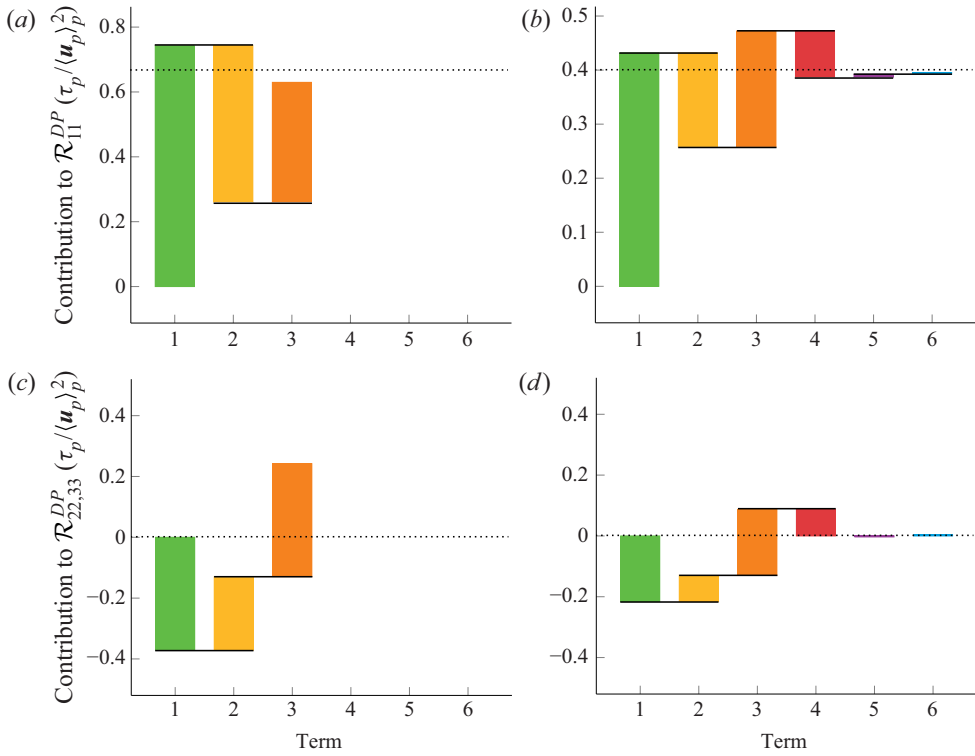
914 A11-14

Figure 4. Term contributions for the streamwise component of drag production for the three-term (4.4) and six-term (4.5) models, shown for the case $Ar = 5.40$ and $\langle \alpha_p \rangle = 0.001$. Drag production obtained from the Eulerian–Lagrangian simulations is shown as the dotted line. Terms 1–6 are represented as (■, green), (■, light orange), (■, orange), (■, red), (■, purple) and (■, light blue), respectively. (*a*) Streamwise three-term model; (*b*) streamwise six-term model; (*c*) cross-stream three-term model; and (*d*) cross-stream six-term model.
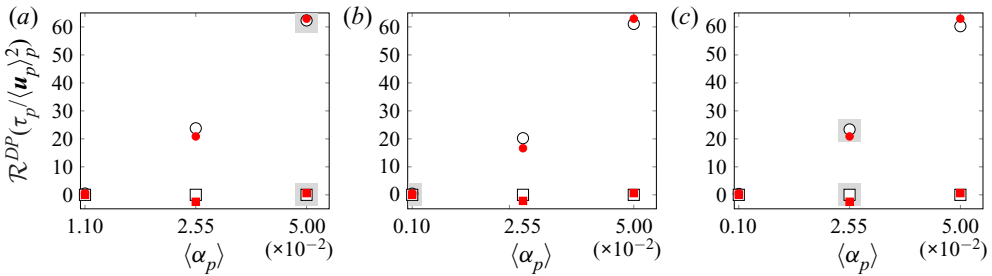
Figure 5. Model learned from sparse training data (denoted with grey shaded boxes). The training and testing errors are 0.07 and 0.08, respectively. Using the convention from previous figures, Eulerian–Lagrangian results (○, streamwise component and □, cross-stream components) and model prediction ((●, red), streamwise component and (■, red), cross-stream components). The sparsely trained model corresponds to (4.6): (*a*) $Ar = 1.8$; (*b*) $Ar = 5.4$; and (*c*) $Ar = 18.0$.

available training points. While additional more challenging tests of the model are required and reserved for future work, this preliminary result may suggest model robustness to variations in Ar and particle volume fraction.

The remaining terms, PS, VD and DE exhibit similar performance as DP and are summarized here. All three terms are normalized by k_f / τ_p in order to ensure realizability

in the Reynolds stresses. The sparse regression algorithm was given access to constant coefficients as well as coefficients dependent upon the invariants, $\mathcal{S}^{(i)}$, up to order three. It is notable that while complex nonlinear coefficients were accessible to the algorithm, they were ultimately not chosen.

4.2.2. *Pressure strain and viscous diffusion*

Pressure strain and viscous diffusion both redistribute TKE throughout the fluid phase and are present in single-phase flows. The learned models are given as

$$\mathcal{R}^{PS} = \varphi \frac{k_f}{\tau_p} \left[\langle \alpha_p \rangle \left(14.36 \hat{\mathbb{U}}_r - 22.65 \hat{\mathbb{R}}_p \right) + \langle \alpha_f \rangle \left(2.60 \hat{\mathbb{R}}_p - 2.72 \hat{\mathbb{R}}_f \right) \right] \quad (4.7)$$

and

$$\mathcal{R}^{VD} = \frac{k_f}{\tau_p} \Big[-1.62 \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p \hat{\mathbb{R}}_p + \varphi \langle \alpha_p \rangle \left(0.53 \mathbb{I} + 0.72 \hat{\mathbb{U}}_r - 3.14 \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p \right)$$

$$+ \varphi \langle \alpha_f \rangle \left(0.74 \hat{\mathbb{R}}_f - 0.62 \hat{\mathbb{U}}_r \right) \hat{\mathbb{R}}_p \Big], \quad (4.8)$$

respectively (see figures 6 and 7, respectively). The dominant terms important for capturing the behaviour of PS across flow parameters are $\varphi \langle \alpha_p \rangle \hat{\mathbb{U}}_r$ and $\varphi \langle \alpha_p \rangle \hat{\mathbb{R}}_p$ and inclusion of only these two terms results in model error of $\epsilon = 0.15$.

In the case of VD, a four-term model is learned in which the three terms that persist into the six-term model are $\varphi \langle \alpha_p \rangle \hat{\mathbb{U}}_r$, $\varphi \langle \alpha_p \rangle \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p$ and $\hat{\mathbb{R}}_f \hat{\mathbb{R}}_p \hat{\mathbb{R}}_p$. The fourth term, $\varphi \langle \alpha_p \rangle \hat{\mathbb{R}}_p$, is replaced by the three remaining terms that appear in (4.8). This reduces model error from 0.29 to 0.07, in a similar manner as described for DP.

4.2.3. *Drag exchange*

Drag exchange describes the mechanism by which TKE is partitioned between the phases. For this case, all terms in the model are of nearly equal importance. A four-term model is learned which excludes $Ar \hat{\mathbb{U}}_r$ with an error of $\epsilon = 0.16$ as compared with the model error of $\epsilon = 0.15$ in the case of the five-term model (see figure 8). This is due to the minimal dependence of the data on Archimedes number

$$\mathcal{R}^{DE} = \varphi \frac{k_f}{\tau_p} \left[-0.36 \varphi \mathbb{I} - \left(0.02 \varphi^2 \langle \alpha_f \rangle^2 + 0.38 Ar \right) \hat{\mathbb{U}}_r + 0.04 \varphi^3 \langle \alpha_f \rangle^3 \hat{\mathbb{R}}_f \left(\hat{\mathbb{R}}_f - \hat{\mathbb{R}}_p \right) \hat{\mathbb{R}}_p \right]. \quad (4.9)$$

For all of the terms considered, sparse regression is capable of uncovering models with model error to machine precision of zero (associated with $\lambda = 0$); however, these resultant models are substantially more complex and likely would not perform well outside the scope of training due to overfitting subtle nonlinearities. These models, for comparison, contain 18 terms for PS, VD and DE, respectively, and 8 terms for DP.

4.2.4. *Particle-phase closures*

The same procedure as described above was used to formulate closures for each of the terms appearing in the particle-phase Reynolds stress equations. The resultant closures are

given as

$$\mathcal{R}^{PS_p} = \varepsilon_p \left(1.89 \hat{\mathbb{R}}_p + 2.12 \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p - 6.52 \hat{\mathbb{R}}_f^2 \hat{\mathbb{R}}_p \right) \tag{4.10}$$

$$\mathcal{R}^{VD_p} = \varepsilon_p \left(-1.91 \hat{\mathbb{R}}_p + 10.5 \hat{\mathbb{U}}_r \hat{\mathbb{R}}_p - 11.78 \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p + 12.30 \hat{\mathbb{R}}_f^2 \hat{\mathbb{R}}_p \right) \tag{4.11}$$

$$\begin{aligned}
\mathcal{R}^{DE_p} = \varepsilon_f \Big[& \left(-2.69 + 2.50 \varphi \langle \alpha_f \rangle^3 - 1.85 \varphi \langle \alpha_f \rangle \right) \mathbb{I} \\
& + \left(-4.79 + 4.08 \varphi \langle \alpha_f \rangle^3 - 3.05 \varphi \langle \alpha_f \rangle \right) \hat{\mathbb{U}}_r \\
& + \left(9.10 - 5.56 \varphi \langle \alpha_f \rangle \right) \hat{\mathbb{U}}_r \hat{\mathbb{R}}_p + \left(21.40 \varphi \langle \alpha_f \rangle - 18.76 \varphi \langle \alpha_f \rangle^3 \right) \hat{\mathbb{R}}_f \hat{\mathbb{R}}_p \Big].
\end{aligned} \tag{4.12}$$

Here, the particle-phase terms are normalized by the dissipation of TKE in order to achieve appropriate scaling behaviour with respect to time. The associated values of λ and model error for the PS, VD and DE are given as $\lambda = (0.3, 0.3, 1)$ and $\epsilon = (0.08, 0.10, 0.03)$, respectively.

4.2.5. *Summary of the multiphase RANS equations*

In this section, we propose a complete set of transport equations for strongly coupled gas–particle homogeneous flows, with a simpler set of equations for the transport of the total Reynolds Stresses. This allows for improved stability and robustness. The closures for individual balance terms are then useful to algebraically decompose the full Reynolds stresses into individual contributions. This system, along with the particle momentum in (3.1), is given by

$$\frac{\partial \varepsilon_f}{\partial t} = \frac{1}{\tau_p^\star} \left[\frac{C_{\varepsilon_f} \mathcal{V}_0^2}{\tau_p^\star} - \varepsilon_f \right], \tag{4.13}$$

$$\frac{\partial \varepsilon_p}{\partial t} = \frac{1}{\tau_p^\star} \left[\frac{C_{\varepsilon_p} \mathcal{V}_0^2}{\tau_p^\star} - \varepsilon_p \right], \tag{4.14}$$

$$\frac{\partial \langle u_f''' u_f''' \rangle_f}{\partial t} = -\frac{1}{\tau_p^\star} \left(\langle u_f''' u_f''' \rangle_f + 0.5 \langle u_p'' u_p'' \rangle_p \right) + \left(\frac{\mathcal{V}_0^2}{\tau_p^\star} \right) \left[(C_1 - C_2) \frac{U_r}{\mathrm{tr}(U_r)} + C_2 \mathbb{I} \right], \tag{4.15}$$

$$\frac{\partial \langle u_p'' u_p'' \rangle_p}{\partial t} = -\frac{1}{\tau_p^\star} \left(\langle u_p'' u_p'' \rangle_p + 0.5 \langle u_f''' u_f''' \rangle_f \right) + \left(\frac{\mathcal{V}_0^2}{\tau_p^\star} \right) \left[(C_3 - C_4) \frac{U_r}{\mathrm{tr}(U_r)} + C_4 \mathbb{I} \right]. \tag{4.16}$$

Here, the coefficients, C_{ε_f}, C_{ε_p}, C_1, C_2, C_3 and C_4, depend upon flow parameters. These dependencies were learned using the sparse regression algorithm described previously and are all nonlinearly parameterized by the mass loading according to

$$C_{\varepsilon_f} = 0.44 \langle \varphi \rangle - 0.05 \langle \varphi \rangle^{1.5} - 0.21 \langle \varphi \rangle^{-2}, \tag{4.17}$$

$$C_{\varepsilon_p} = 0.01 \langle \varphi \rangle - 0.001 \langle \varphi \rangle^{1.5} + 0.02 \langle \varphi \rangle^{-2}, \tag{4.18}$$
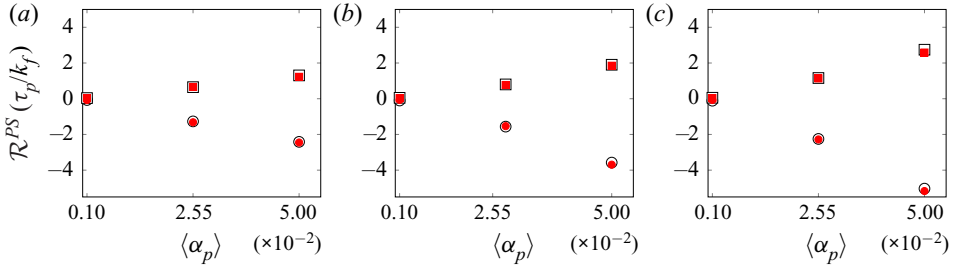
S. Beetham, R.O. Fox and J. Capecelatro

Figure 6. Pressure strain Eulerian–Lagrangian results (∘, streamwise component and □, cross-stream components) and model prediction ((•, red), streamwise component and (■, red), cross-stream components). Model corresponds to (4.7) and results from $\lambda = 0.3$. The associated model error is 0.04: (a) $Ar = 1.8$; (b) $Ar = 5.4$; and (c) $Ar = 18.0$.

$$C_1 = 1.40\langle\varphi\rangle - 0.18\langle\varphi\rangle^{1.5} + 2.17\langle\varphi\rangle^{-2}, \tag{4.19}$$

$$C_2 = 0.18\langle\varphi\rangle - 0.02\langle\varphi\rangle^{1.5} + 0.12\langle\varphi\rangle^{-2}, \tag{4.20}$$

$$C_3 = 1.20\langle\varphi\rangle - 0.16\langle\varphi\rangle^{1.5} + 2.03\langle\varphi\rangle^{-2}, \tag{4.21}$$

$$C_4 = 0.15\langle\varphi\rangle - 0.02\langle\varphi\rangle^{1.5} + 0.11\langle\varphi\rangle^{-2}. \tag{4.22}$$

The fluid drift velocity, defined as $\boldsymbol{u}_d = \langle\boldsymbol{u}_f\rangle_p - \langle\boldsymbol{u}_f\rangle_f$, requires modelling for the fluid velocity seen by the particles, $\langle\boldsymbol{u}_f\rangle_p$, as does the particle-phase momentum equation (3.1). The learned model for $\boldsymbol{u}_d$ is given as

$$\boldsymbol{u}_d = \frac{\sqrt{\langle\alpha_p'^2\rangle}}{\langle\alpha_p\rangle}\left[-7.8\hat{\mathbb{R}}_f + 4.6\hat{\mathbb{R}}_p + 2.2\right]\left(\langle\boldsymbol{u}_p\rangle_p - \langle\boldsymbol{u}_f\rangle_f\right), \tag{4.23}$$

using $\lambda = 1$, resulting in a model error of 0.06. Here, we impose that the model be scaled by the variance of the particle volume fraction, $\langle\alpha_p'^2\rangle$, in order to ensure this quantity approaches zero in the dilute limit.

Finally, since the variance of particle volume fraction is not known *a priori*, an additional model was formulated for this quantity, using the same method described previously, given by

$$\frac{\sqrt{\langle\alpha_p'^2\rangle}}{\langle\alpha_p\rangle} = 0.32\hat{\mathbb{R}}_p : \hat{\mathbb{R}}_p + \hat{\mathbb{U}}_r : \left(1.10\hat{\mathbb{R}}_f + 2.67\hat{\mathbb{R}}_p\langle\alpha_p\rangle^{1/2}\right). \tag{4.24}$$

4.3. *Application to transient flow*

To assess model realizability, it is imperative to evaluate the transient behaviour that precedes the stationary state. To this end, we generate transient data by using the statistically stationary solution from the cases described in § 2 as an initial condition and instantaneously reverse the direction of gravity, i.e.

$$\boldsymbol{g} = \begin{cases} (-g, 0, 0), & \text{if } t < 0 \\ (g, 0, 0), & \text{if } t \geq 0 \end{cases}. \tag{4.25}$$

This strategy generates temporally evolving data and allows us to apply our models to a homogeneous flow, while relaxing the assumption of stationarity. Three cases
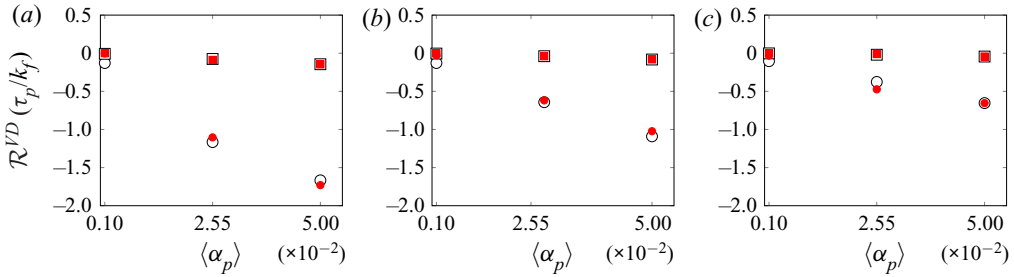
914 A11-18

Figure 7. Viscous diffusion Eulerian–Lagrangian results (○, streamwise component and □, cross-stream components) and model prediction ((●, red), streamwise component and (■, red), cross-stream components). Model corresponds to (4.8) and results from $\lambda = 0.2$. The associated model error is 0.07: (*a*) $Ar = 1.8$; (*b*) $Ar = 5.4$; and (*c*) $Ar = 18.0$.
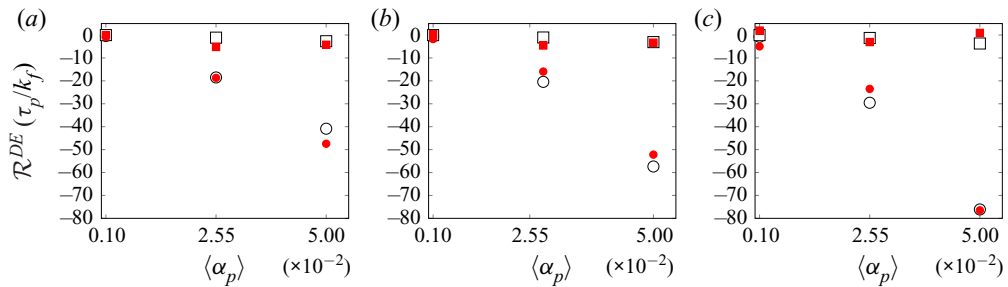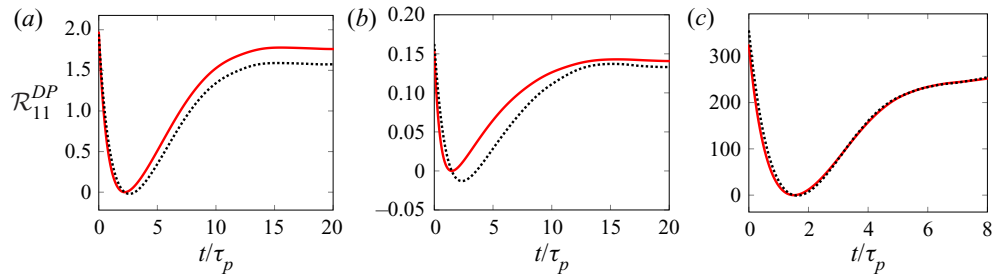

Figure 8. Drag exchange Eulerian–Lagrangian results (○, streamwise component and □, cross-stream components) and model prediction ((●, red), streamwise component and (■, red), cross-stream components). Model corresponds to (4.9) and results from $\lambda = 0.006$. The associated model error is 0.15: (*a*) $Ar = 1.8$; (*b*) $Ar = 5.4$; and (*c*) $Ar = 18.0$.

Figure 9. Temporal evolution of drag production obtained from the learned model 4.5 (——, very thick solid, red) and Euler–Lagrange data (····, very thick dotted, black) for CIT after gravity is reversed instantaneously: (*a*) $Ar, \langle \alpha_p \rangle = (1.8, 0.0255)$; (*b*) $Ar, \langle \alpha_p \rangle = (5.4, 0.001)$; and (*c*) $Ar, \langle \alpha_p \rangle = (18, 0.05)$.

corresponding to $Ar, \langle \alpha_p \rangle = [(1.8, 0.0255), (5.4, 0.05), (18, 0.001)]$ were simulated in order to probe the entire parameter space. The performance of the drag production model, in particular, is highlighted in figure 9 and compared against the transient EL data for all three cases considered. Finally, the forward solution of the system of model equations presented is solved for these cases, and the predicted mean particle velocities are compared with the EL data (see figure 10).

The models described herein are successful for the parameter space studied and perform exceptionally well on transient data, despite being trained using stationary data. While care
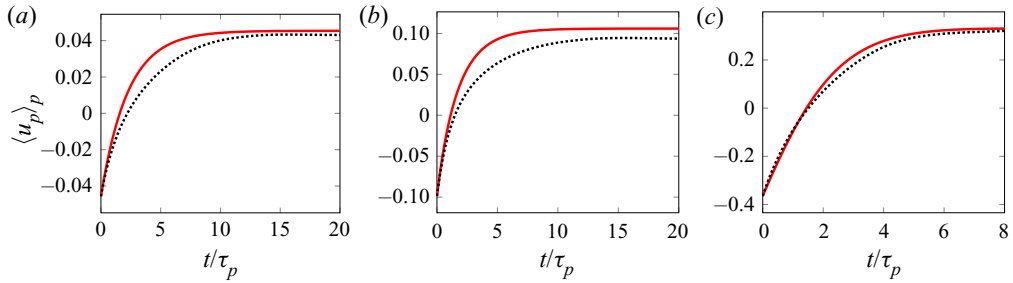
Figure 10. Temporal evolution of mean particle settling velocity obtained from the multiphase RANS equations (——, very thick solid, red) and Euler–Lagrange data (····, very thick dotted, black) for CIT after gravity is reversed instantaneously: (a) $Ar, \langle \alpha_p \rangle = (1.8, 0.0255)$; (b) $Ar, \langle \alpha_p \rangle = (5.4, 0.001)$; and (c) $Ar, \langle \alpha_p \rangle = (18, 0.05)$.

has been taken to ensure asymptotic agreement with the dilute limit, the robustness of the models applied to denser systems or non-homogeneous flows is left for future investigation.

5. Conclusions

In this work, the multiphase RANS equations are presented for two-way coupled gas–solid flows. In this class of flows, the coupling between the phases spontaneously gives rise to coherent particle structures, which in turn generate and sustain turbulence in the carrier phase. This phenomenon has important engineering implications (Shaffer *et al.* 2013; Miller *et al.* 2014; Beetham & Capecelatro 2019; Guo & Capecelatro 2019) and makes the formulation of closure models that are predictive across scales and flow conditions challenging.

We apply a newly formulated modelling methodology, sparse regression with embedded form invariance (Beetham & Capecelatro 2020), to highly resolved Eulerian–Lagrangian data for fully developed CIT. The benefits of this methodology as compared with Neural Networks, which have become increasingly popular, are (i) interpretability of the resultant closures, since they are in a closed algebraic formulation, (ii) ease of dissemination to existing RANS solvers and (iii) robustness to very sparse training sets. The dataset used for model development spans a range of flow parameters, specifically $Ar = (1.8, 5.4, 18.0)$ and $\langle \alpha_p \rangle = (0.001, 0.0255, 0.05)$, in order to formulate models across a range of typical conditions.

Particular attention is paid to the closures for the four dominant unclosed terms that appear in the fluid-phase Reynolds-stress equations – PS, VD, DP and DE. In applying the sparse regression method to each of these terms individually, we discover compact closures containing between four and six terms that are accurate across the scope of training (model error ranges from 0.01 to 0.15). Because of the compact nature of the models developed and the nature of the sparse regression algorithm, we are able to assess the relative importance of each term and its role in reducing model error. Further, we demonstrate that even when training on a subset of the Eulerian–Lagrangian data, the methodology learns a model that remains accurate outside the scope of its training. Additionally, because of the compact, algebraic formulation of the method, resultant models are accessible for interpretation and terms of greater physical significance are easily identified. These findings suggest that the sparse regression methodology holds promise for developing closures for more complicated multiphase flows, such as channel, duct or bubbly flows. Further, since nearly all flows of practical importance involve both walls and strong shear, future modelling

work will focus on near-wall treatments and adaptation of the tensor basis to account for shear and other flow conditions which result in turbulence.

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

Author ORCIDs.

R.O. Fox https://orcid.org/0000-0003-1944-1861;

J. Capecelatro https://orcid.org/0000-0002-1684-4967.

REFERENCES

AGRAWAL, K., HOLLOWAY, W., MILIOLI, C.C., MILIOLI, F.E. & SUNDARESAN, S. 2013 Filtered models for scalar transport in gas–particle flows. *Chem. Engng Sci.* **95**, 291–300.

AGRAWAL, K., LOEZOS, P.N., SYAMLAL, M. & SUNDARESAN, S. 2001 The role of meso-scale structures in rapid gas-solid flows. *J. Fluid Mech.* **445**, 151–186.

ALISEDA, A., CARTELLIER, A., HAINAUX, F. & LASHERAS, J.C. 2002 Effect of preferential concentration on the settling velocity of heavy particles in homogeneous isotropic turbulence. *J. Fluid Mech.* **468**, 77–105.

ANDERSON, T.B. & JACKSON, R. 1967 Fluid mechanical description of fluidized beds. Equations of motion. *Ind. Engng Chem. Fundam.* **6** (4), 527–539.

BALACHANDAR, S. & EATON, J.K. 2010 Turbulent dispersed multiphase flow. *Ann. Rev. Fluid Mech.* **42**, 111–133.

BATCHELOR, G.K. 1972 Sedimentation in a dilute dispersion of spheres. *J. Fluid Mech.* **52**, 245–268.

BATCHELOR, G.K. 1982 Sedimentation in a dilute polydisperse system of interacting spheres. Part 1. General theory. *J. Fluid Mech.* **119**, 379–408.

BATCHELOR, G.K. 1988 A new theory of the instability of a uniform fluidized bed. *J. Fluid Mech.* **193**, 75–110.

BEETHAM, S. & CAPECELATRO, J. 2019 Biomass pyrolysis in fully-developed turbulent riser flow. *Renew. Energy* **140**, 751–760.

BEETHAM, S. & CAPECELATRO, J. 2020 Formulating turbulence closures using sparse regression with embedded form invariance. *Phys. Rev. Fluids* **5**, 084611.

BESNARD, D., HARLOW, F.H. & RAUENZHAN, R.M. 1987 Conservation and transport properties of turbulence with large density variations. *Tech. Rep.* LA-10911-MS. Los Alamos National Laboratory.

BODE, M., GAUDING, M., KLEINHEINZ, K. & PITSCH, H. 2019 Deep learning at scale for subgrid modeling in turbulent flows: regression and reconstruction. arXiv:1910.00928v1.

BOSSE, T., KLEISER, L. & MEIBURG, E. 2006 Small particles in homogeneous turbulence: settling velocity enhancement by two-way coupling. *Phys. Fluids* **18** (2), 027102.

BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci.* **113** (15), 3932–3937.

CAO, J. & AHMADI, G. 1995 Gas–particle two–phase turbulent flow in a vertical duct. *Intl J. Multiphase Flow* **21** (6), 1203–1228.

CAPECELATRO, J. & DESJARDINS, O. 2013 An Euler–Lagrange strategy for simulating particle-laden flows. *J. Comput. Phys.* **238**, 1–31.

CAPECELATRO, J., DESJARDINS, O. & FOX, R.O. 2014 Numerical study of collisional particle dynamics in cluster-induced turbulence. *J. Fluid Mech.* **747**, 1–13.

CAPECELATRO, J., DESJARDINS, O. & FOX, R.O. 2015 On fluid-particle dynamics in fully-developed cluster-induced turbulence. *J. Fluid Mech.* **780**, 578–635.

CAPECELATRO, J., DESJARDINS, O. & FOX, R.O. 2016*a* Effect of domain size on fluid-particle statistics in homogeneous, gravity-driven, cluster-induced turbulence. *Trans. ASME: J. Fluids Engng* **138**, 1–8.

CAPECELATRO, J., DESJARDINS, O. & FOX, R.O. 2016*b* Strongly–coupled gas–particle flows in vertical channels. Part 2. Turbulence modeling. *Phys. Fluids* **28**, 1–22.

CAPECELATRO, J., DESJARDINS, O. & FOX, R.O. 2018 On the transition between turbulence regimes in particle-laden channel flows. *J. Fluid Mech.* **845**, 499.

CHENG, Y., GUO, Y., WEI, F., JIN, Y. & LIN, W. 1999 Modeling the hydrodynamics of downer reactors based on kinetic theory. *Chem. Engng Sci.* **54**, 2019–2027.

CUNDALL, P.A. & STRACK, O.D.L. 1979 A discrete numerical model for granular assemblies. *Geotechnique* **29** (1), 47–65.

DASGUPTA, S., JACKSON, R. & SUNDARESAN, S. 1994 Turbulent gas–particle flow in vertical risers. *AIChE J.* **40** (2), 215–228.

DASGUPTA, S., JACKSON, R. & SUNDARESAN, S. 1998 Gas–particle flow in vertical pipes with high mass loading of particles. *Powder Technol.* **96**, 6–23.

DESJARDINS, O., BLANQUART, G., BALARAC, G. & PITSCH, H. 2008 High order conservative finite difference scheme for variable density low Mach number turbulent flows. *J. Comput. Phys.* **227** (15), 7125–7159.

DURAISAMY, K. & DURBIN, P.A. 2014 Transition modeling using data driven approaches. In *Proceedings of the Summer Program*, p. 427. Center for Turbulence Research, Stanford University.

DURAISAMY, K., ZHANG, Z.J. & SINGH, A.P. 2015 New approaches in turbulence and transition modeling using data-driven techniques. In *53rd AIAA Aerospace Sciences Meeting*, p. 1284.

EATON, J.K. & FESSLER, J.R. 1994 Preferential concentration of particles by turbulence. *Intl J. Multiphase Flow* **20**, 169–209.

FERRANTE, A. & ELGHOBASHI, S. 2003 On the physical mechanisms of two-way coupling in particle-laden isotropic turbulence. *Phys. Fluids* **15** (2), 315–329.

FOX, R.O. 2014 On multiphase turbulence models for collisional fluid–particle flows. *J. Fluid Mech.* **742**, 368–424.

GATSKI, T.B. & SPEZIALE, C.G. 1993 On explicit algebraic stress models for complex turbulent flows. *J. Fluid Mech.* **254**, 59–78.

GLASSER, B.J., SUNDARESAN, S. & KEVREKIDIS, I.G. 1998 From bubbles to clusters in fluidized beds. *Phys. Rev. Lett.* **81**, 1849.

GOLDHIRSCH, I. & ZANETTI, G. 1993 Clustering instability in dissipative gases. *Phys. Rev. Lett.* **70** (11), 1619.

GUO, L. & CAPECELATRO, J. 2019 The role of clusters on heat transfer in sedimenting gas-solid flows. *Intl J. Heat Mass Transfer* **132**, 1217–1230.

HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. 2009 *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.

HOPKINS, M.A. & LOUGE, M.Y. 1991 Inelastic microstructure in rapid granular flows of smooth disks. *Phys. Fluids* **3** (1), 47–57.

INNOCENTI, A., FOX, R.O., SALVETTI, M.V. & CHIBBARO, S. 2019 A Lagrangian probability-density-function model for collisional turbulent fluid-particle flows. *J. Fluid Mech.* **862**, 449–489.

JIANG, Y.Y. & ZHANG, P. 2012 Numerical investigation of slush nitrogen flow in a horizontal pipe. *Chem. Engng Sci.* **73**, 169–180.

JOHNSON, B.M. & SCHILLING, O. 2011 Reynolds-averaged Navier–Stokes model predictions of linear instability. I: Buoyancy-and shear-driven flows. *J. Turbul.* **12**, N36.

JORDAN, M.I. & MITCHELL, T.M. 2015 Machine learning: trends, perspectives, and prospects. *Science* **349** (6245), 255–260.

LING, J., KURZAWSKI, A. & TEMPLETON, J. 2016 Reynolds averaged turbulence modeling using deep neural networks with embedded invariance. *J. Fluid Mech.* **807**, 155–166.

LIU, W. & FANG, J. 2019 Iterative framework of machine-learning based turbulence modeling for Reynolds-averaged Navier–Stokes simulations. arXiv:1910.01232v1.

LU, C. 2010 Artificial neural network for behavior learning from meso-scale simulations, application to multi-scale multimaterial flows. PhD thesis, University of Iowa.

MA, M., LU, J. & TRYGGVASON, G. 2016 Using statistical learning to close two-fluid multiphase flow equations for bubbly flows in vertical channels. *Intl J. Multiphase Flow* **85**, 336–347.

MA, T., LUCAS, D. & BRAGG, A.D. 2020 Explicit algebraic relation for calculating Reynolds normal stresses in flows dominated by bubble-induced turbulence. *Phys. Rev. Fluids* **5**, 084305.

MILANO, M. & KOUMOUTSAKOS, P. 2002 Neural network modeling for near wall turbulent flow. *J. Comput. Phys.* **182**, 1–26.

MILLER, D.C., *et al.* 2014 Carbon capture simulation initiative: a case study in multiscale modeling and new challenges. *Annu. Rev. Chem. Biomol. Engng* **5**, 301–323.

MUDDE, R.F. 2005 Gravity-driven bubbly flows. *Annu. Rev. Fluid Mech.* **37**, 393–423.

PIERCE, C.D. 2001 Progress-variable approach for large-eddy simulation of turbulent combustion. PhD thesis, Stanford University, CA.

POPE, S.B. 1975 A more general effective-viscosity hypothesis. *J. Fluid Mech.* **72** (2), 331–340.

POPE, S.B. 2000 *Turbulent Flows*. Cambridge University Press.

POURANSARI, H. & MANI, A. 2017 Effects of preferential concentration on heat transfer in particle-based solar receivers. *J. Solar Energy Engng* **139** (2), 021008.

RAJABI, E. & KAVIANPOUR, M.R. 2012 Intelligent prediction of turbulent flow over backward-facing step using direct numerical simulation data. *Engng Appl. Comput. Fluid Mech.* **6** (4), 490–503.

RAO, A., CURTIS, J.S., HANCOCK, B.C. & WASSGREN, C. 2012 Numerical simulation of dilute turbulent gas–particle flow with turbulence modulation. *AIChE J.* **58**, 1381–1396.

RZEHAK, R. & KREPPER, E. 2013 CFD modeling of bubble-induced turbulence. *Intl J. Multiphase Flow* **55**, 138–155.

SCHMELZER, M., DWIGHT, R.P. & CINNELLA, P. 2020 Discovery of algebraic Reynolds-stress models using sparse symbolic regression. *Flow Turbul. Combust.* **104** (2), 579–603.

SHAFFER, F., GOPALAN, B., BREAULT, R.W., COCCO, R., KARRI, S.B., HAYS, R. & KNOWLTON, T. 2013 High speed imaging of particle flow fields in CFB risers. *Powder Technol.* **242**, 86–99.

DE SILVA, B.M., CHAMPION, K., QUADE, M., LOISEAU, J. -C., KUTZ, J.N. & BRUNTON, S.L. 2020 Pysindy: a python package for the sparse identification of nonlinear dynamics from data. arXiv:2004.08424.

SINCLAIR, J.L. & JACKSON, R. 1989 Gas–particle flow in a vertical pipe with particle-particle interactions. *AIChE J.* **35** (9), 1473–1486.

SPENCER, A.J.M. & RIVLIN, R.S. 1958 The theory of matrix polynomials and its application to the mechanics of isotropic continua. *Arch. Rat. Mech. Anal.* **2**, 309–336.

SPEZIALE, C.G., SARKAR, S. & GATSKI, T.B. 1991 Modelling the pressure-strain correlation of turbulence: an invariant dynamical systems approach. *J. Fluid Mech.* **227**, 245–272.

SUN, Z. & ZHU, J. 2019 A consolidated flow regime map of upward gas fluidization. *AIChE J.* **65**, 1–15.

SUNDARAM, S. & COLLINS, L.R. 1994 Spectrum of density fluctuations in a particle-fluid system-I. Monodisperse spheres. *Intl J. Multiphase Flow* **20** (6), 1021–1037.

SUNDARAM, S. & COLLINS, L. R 1999 A numerical study of the modulation of isotropic turbulence by suspended particles. *J. Fluid Mech.* **379**, 105–143.

TENNETI, S. & SUBRAMANIAM, S. 2011 Drag law for monodisperse gas-solid systems using particle-resolved direct numerical simulation of flow past fixed assemblies of spheres. *Intl J. Multiphase Flow* **37** (9), 1072–1092.

TOWNS, J., *et al.* 2014 XSEDE: Accelerating scientific discovery. *Comput. Sci. Engng* **16** (5), 62–74. doi:10.1109/MCSE.2014.80

TRACEY, B., DURAISAMY, K. & ALONSO, J.J. 2015 A machine learning strategy to assist turbulence model development. *AIAA Paper* 1287.

WANG, L.P. & MAXEY, M.R. 1993 Settling velocity and concentration distribution of heavy particles in homogeneous isotropic turbulence. *J. Fluid Mech.* **256**, 27–68.

WYLIE, J.J. & KOCH, D.L. 2000 Particle clustering due to hydrodynamic interactions. *Phys. Fluids* **12** (5), 964–970.

YANG, T.S. & SHY, S.S. 2003 The settling velocity of heavy particles in an aqueous near-isotropic turbulence. *Phys. Fluids* **15** (4), 868–880.

ZENG, Z.X. & ZHOU, L.X. 2006 A two-scale second–order moment particle turbulence model and simulation of dense gas–particle flows in a riser. *Powder Technol.* **162**, 27–32.

ZHOU, Y. 2017 Rayleigh–Taylor and Richtmyer–Meshkov instability induced flow, turbulence, and mixing. II. *Phys. Rep.* **723–725**, 1–160.