Preasymptotic Taylor dispersion: evolution from the initial condition

E. Taghizadeh[1], F. J. Valdés-Parada[2] and B. D. Wood[1,†]

[1]School of Chemical, Biological and Environmental Engineering, Corvallis, OR 97331, USA

[2]Departamento de Ingeniería de Procesos e Hidráulica. División de ciencias Básicas e Ingeniería. Universidad Autónoma Metropolitana-Iztapalapa. Av. San Rafael Atlixco 186, Col. Vicentina, 09340, CDMX, Mexico

Although the process of hydrodynamic dispersion has been studied for many years, the description of solute spreading at early times has proved to be challenging. In particular, for some kinds of initial conditions, the solute evolution may exhibit a second moment that decreases (rather than increases, as is typically observed) in time. Most classical approaches would predict a negative effective hydrodynamic dispersion coefficient for such a situation. This creates some difficulties: not only does a negative dispersion coefficient lead to a violation of the second law of thermodynamics, but it also creates a mathematically ill-posed problem. We outline a set of four desirable qualities in a well-structured theory of unsteady dispersion as follows: (i) positivity of the dispersion coefficient, (ii) non-dependence upon initial conditions, (iii) superposability of solutions and (iv) convergence of solutions to classical asymptotic results. We use averaging to develop an upscaled result that adheres to these qualities. We find that the upscaled equation contains a source term that accounts for the relaxation of the initial configuration. This term decreases exponentially fast in time, leading to correct asymptotic behaviour while also accounting for the early-time solute dynamics. Analytical solutions are presented for both the effective dispersion coefficient and the source term, and we compare our upscaled results with averaged solutions obtained from numerical simulations; both averaged concentrations and spatial moments are compared. Error estimates are quantified, and we find good correspondence between the upscaled theory and the numerical results for all times.

Key words: mixing and dispersion, general fluid mechanics

---

1. Introduction

Taylor dispersion, the process of solute spreading in a capillary tube, is the archetype for hydrodynamic dispersive processes. The primary reason for this is that it has all of the essential physical features important to hydrodynamic dispersion, and yet it has a geometry that is simple enough that an analytical approach is feasible.

† Email address for correspondence: brian.wood@oregonstate.edu

As such, there have been literally thousands of papers on the topic since the early 1950s.

The particular focus of this paper is to address the case of non-uniform initial distributions of a solute in Taylor dispersion under steady flow conditions. Non-uniform distributions of solute have a number of applications. A few concrete examples are (i) the design of exchangers for transient heat transfer (Heris, Esfahany & Etemad 2007); (ii) the diffusion of solute from an initially spherical input to model drug delivery in blood flow (Gekle 2017); (iii) pulsed radially non-uniform inputs to tubular reactors used to promote mixing (Stonestreet & Harvey 2002; Abbott *et al.* 2013); (iv) the use of phosphorescently tagged particles for fluid velocimetry (Mohand *et al.* 2017); and (v) non-uniform injections of chemical solutes in electrophoretic separations (Liu & Ivory 2013).

The primary issue in unsteady solute dispersion is the transition of the solute distribution from an initial disequilibrium state, to a state that represents the quasi-steady Gaussian distribution conditions that are necessary for the Taylor dispersion regime to be valid. Rather than setting up the problem for a variety of special cases, we develop a general method that applies to a wide variety of initial conditions. Our particular interest is to develop a model for unsteady solute dispersion where the contribution from the initial configuration can be clearly identified. This contribution can be expected to relax over time in order to recover the classical Taylor dispersion model.

There are a few useful concepts that can be proposed to help ensure that a theory for dispersion is well structured (although it is possible to develop theories that do not adhere to these guidelines). These are as follows:

(i) The effective dispersion coefficient should be positive. Although negative dispersion coefficients have been proposed, they suffer from a few significant problems. These include the ill posedness of the resulting balance equation (i.e. the inverse heat equation (Weber 1981)) and the fundamental incompatibility with macroscale thermodynamics (Gray & Miller 2009; Gray & Miller 2014, chap. 10). Negative dispersion coefficients lead to an equation that is no longer even of the diffusion type because it no longer obeys a strong maximum principle (Olver 2014).

(ii) The effective dispersion coefficient should be independent of the particular initial conditions imposed. In other words, the effective dispersion coefficient should be a function of only the molecular diffusion coefficient, and the structure of the fluid velocity field. This allows the dispersion coefficient to be defined strictly on a molecular and hydrodynamic basis, without having to resort to the incorporation of particular initial configurations of otherwise passive solutes. If the dispersion coefficient were to be dependent upon particular initial configurations, one would then lose the principle of linear superposition for an (otherwise) entirely linear problem. This has substantial ramifications for the interpretation of initial-configuration-dependent formulations.

(iii) Solutions to the effective convection–dispersion equation should be superposable in the sense defined by Taylor (1959). This criterion is based on a physical desideratum outlined by Taylor (1959). Essentially, the argument was that if a new source was injected after some time had elapsed (the 'two release' problem), the effective dispersion coefficient should be single valued at points where the two sources overlap. The time-dependent dispersion theories would predict two different values for the dispersion coefficient at such overlapping points, a situation to which Taylor (1959) objected.

(iv) Whatever the theory for preasymptotic dispersion is, the dispersion coefficient should approach over time the classical asymptotic values that have been developed in the literature. These classical asymptotic values have been scrutinized rigorously for decades; there is little uncertainty in those results.

Our goal in this work is to develop a theory that addresses each of these issues; we seek to make the theory as concrete as possible by expressing results in analytical form when feasible. We neglect reaction, although the methods proposed are extendable to that case (cf. Wang *et al.* 2015). Ultimately, our approach leads to an upscaled equation for a single chemical species that contains a non-conventional source term as follows:

$$\frac{\partial \langle c \rangle}{\partial t} = D^*(t) \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} + s^*(z, t). \tag{1.1}$$

Here, c is the concentration of the solute, the angled brackets indicate cross-section average, D^* is the effective dispersion coefficient, U is the cross-sectional-averaged longitudinal velocity and s^* is a non-conventional source term that is exponentially decaying in time; z and t are the independent variables representing space and time, respectively. While this form for a macroscale balance equation may seem unusual, the source term s^* is an essential component that arises directly from the upscaling analysis. This source term creates an overall balance that meets the set of criteria defined above for a well-structured dispersion process.

The remainder of the manuscript is outlined as follows. In the next section, we give an overview of the literature on Taylor dispersion, with a specific focus on methods that have been developed to handle early time evolution from the initial conditions. This is followed by a presentation of the microscale physics in § 3. In § 4, we define the upscaling process. In § 5, the problem of closure is presented and integral closure solutions are described. The closed problem is described in § 6; in this section we also define the effective dispersion coefficient, D^*, and source term, s^*, and provide explicit analytical solutions for these quantities. In § 7, we compute solutions to the dispersion problem for three different initial configurations, and compare the results derived from the upscaled model with those derived from numerical simulations (NS) computed at the microscale. Both averaged concentrations and spatial moments are compared. In § 8 we provide some discussion of the results of the theory applied to several interesting initial conditions. In § 9, we specifically address the application of the upscaled balance equation to conditions where the problem superposition is of interest. Finally, in § 10 a list of conclusions is presented. Supplementary materials available at https://doi.org/10.1017/jfm.2020.56 have been generated to support this work; these materials include specific details regarding error analysis, correspondences between the theory developed here and infinite-order expansion methods and a list of notation.

2. Background

In this section, we review a subset of the work conducted on the Taylor dispersion problem. An exhaustive literature review on the subject is challenging because of the volume of literature represented by this topic; therefore, our review is focused specifically on studies that seek to provide solutions for early-time behaviour. In the material that follows, we define the Péclet number by

$$Pe = \frac{Ua}{\mathcal{D}}, \tag{2.1}$$

where $\mathcal{D}$ is the molecular diffusion coefficient, and a is the radius of the tube. Most studies of Taylor dispersion have adopted this definition for the Péclet number; it can be thought of as the ratio of the convective to the (radial) diffusive fluxes. This is slightly different from the work of Taylor (1953), who used the ratio of the diffusive to convective time scales (this difference is discussed in subsequent sections).

2.1. *Review of the literature*

For the purposes of the review below, early time is qualitatively defined as the transient time period during which the relaxation of the initial configuration of the solute is important to the dynamics of the Taylor dispersion process. It is possible to broadly characterize the different mathematical approaches for describing Taylor dispersion in the transient, early-time regime as outlined below.

(i) *Kramers–Moyal-like expansions*. By far the most popular approach has been the proposition that the macroscale solution can be expressed as an infinite series of spatial derivatives of the average concentration. Several variations of this approach have been attempted, and they are represented in the works of Gill *et al.* (Gill & Sankarasubramanian 1970, 1971), Chatwin (O'Hara 1969; Chatwin 1970, 1972), Degance & Johns (1978*a*,*b*) and Mauri (1991). Balakotaiah, Chang & Smith (1995) used centre manifold theory to develop a macroscale formulation that is identical to that of Gill & Sankarasubramanian (1970). The resolution of the centre manifold model was improved by Watt & Roberts (1996) by constructing a multi-mode model for dispersion in channels. These authors considered a two-component model that follows the same trend of thought of the zonal models by Chikwendu & Ojiakor (1985), Chikwendu (1986*a*,*b*). The work by Yu (1976, 1979) is interesting in that it began as a formal eigenfunction expansion for an unsimplified version of the convection–diffusion equation by assuming that expansions in Bessel functions were possible. After a formal integral solution was developed, the term arising from the initial conditions was discarded, and the remaining integrations of the Bessel functions were expanded in a power series. The result was a solution in an infinite series of derivatives of all orders, much like the Kramers–Moyal-type expansions discussed above. Although originally promoted as being more capable than the models of Gill & Sankarasubramanian (1970), a sequence of comments (Gill & Subramanian 1980; Yu 1980) established that the two methods are, at least to order 3, identical. The works by Westerterp, Dil'man & Kronberg (1995) and Kronberg, Benneker & Westerterp (1996) are of this type, but in that work the iterative process was truncated at order 2. One unique feature of those works is that mixed derivative terms were maintained, and this ultimately led to a macroscale form that had hyperbolic (rather than parabolic) features. We note that the work by Gill & Sankarasubramanian (1970) also derived hyperbolic terms in their approach, but ultimately eliminated those terms on the basis of their assumed macroscale form. More recently, Wu & Chen (2014) used an expansion involving spatial derivatives up to ninth order. Although their results compared well with previously reported numerical results, they did not recognize that their solutions included significant negative concentrations outside of a limited domain. This kind of error was corrected in a later paper (Wang & Chen 2016). Other works (Mercer & Roberts 1990, 1994; Balakotaiah *et al.* 1995) also considered higher-order expansions and addressed the issue of negative concentrations.

(ii) *Direct solutions*. In the direct solutions, the focus has been to solve the convection–diffusion mass balance equation in the tube directly using an eigenfunction

expansion. This is a challenging task, because the conventional methods (e.g. separation of variables) used to determine closed-form expressions for the Green's function do not work with this problem, primarily because the convection term mixes both radial and longitudinal independent variables. The first direct solutions appear to be those of Philip (1963). For that work, however, the solution sought was initially of the form of a diffusion equation (similar to the more general approach of Gill & Sankarasubramanian (1970)), which automatically excluded certain forms of the initial condition (with roughly the same reasoning discussed by Degance & Johns (1980)). The paper proposed a solution as a series of the confluent hypergeometric functions $_1F_1$. The papers by Tseng & Besant (1970, 1972) assumed the same solution as did Philip (1963) (and were subject to the same restrictions), but managed to find simplifications that yielded solutions in terms of Bessel J_0 functions, the roots of the Bessel J_1 function and the error function. In the work by Stokes & Barton (1990), a Fourier transformation method was used, but the results had to be inverted numerically; as such, it is difficult to compare these results with other efforts.

(iii) *Asymptotic (perturbation) solutions*. A number of perturbation-type solutions have been proposed, and it is not possible to cleanly categorize solutions from other approaches as being free from perturbation-like arguments. However, more classical perturbation-type expansions have been investigated by Vrentas & Vrentas (1988, 2000). The paper by Lighthill (1966) is of this form, as was the extension proposed by Chatwin (1976, 1977); both of these solutions required the time to be small compared to a characteristic diffusive time scale. Fife & Nicholes (1975) conducted a conventional perturbation analysis, but, importantly, recognized the potential influence of initial source terms, which is somewhat unique. Phillips and Kaye examined the asymptotic (but direct) solutions for $Pe \rightarrow \infty$ for both large z (1996), and short times (1997); both solutions neglected longitudinal diffusion. Such solutions have generally been very successful when used in their range of validity (which generally involves either long or short times, and/or large or small Péclet numbers).

(iv) *Moment methods*. The early paper by Aris (1956) presented the first moment-based method for Taylor dispersion. Although the paper purported to eliminate the constraints proposed by Taylor (1954), in fact the method is technically only suitable for asymptotic estimates of the dispersion coefficient, although some transient results were presented for particular initial conditions. Moment methods focus on determining the spatial moments of the concentration field rather than its mean value (which, quite usefully, simplifies the analysis), often with the tacit assumption that the effective parameters of the averaged mass balance equation can be determined directly from the moments. Again, it is difficult to categorize approaches as uniquely moment based, since many approaches compute moments regardless of the underpinning mathematical methods used. A number of researchers have extended moment method into the preasymptotic regime, including the works of Horn (1971), Chatwin (1977), Degance & Johns (1978b), Latini & Bernoff (2001) and Dentz & Carrera (2007). The moment technique has become particularly popular in describing Taylor-like flows in stratified media (e.g. Fried & Combarnous 1971; Lake & Hirasaki 1981; Valocchi 1989), downstream contaminant release in rivers (Smith 1984), among many others. The method inherently assumes that a finite number of moments suitably defines the transport behaviour of the system, an assumption that is technically only true for specific initial conditions or in the near-asymptotic regime (this issue is discussed further in § 2.2).

(v) *Non-local formulations.* Few non-local formulations (containing terms with either time or space integrations in the transport equation) have been proposed for this problem. Although not specific to the Taylor problem, the work of Deng & Cushman (1995) is of this type, and somewhat set the standard for non-local equations of the convection–dispersion type (however, in a more general context, the work of Eringen (1978) was ground-breaking for non-local balance equations). The paper by Wood & Valdés-Parada (2013) also proposed non-local formulations from the volume averaging perspective, although they were also not specific to the Taylor dispersion problem. In the report of Smith (1981*a*), one of the first non-local formulations to the Taylor problem was outlined. The developments of that paper were also guided by a Kramers–Moyal-type expansion, and thus lead to hyperbolic expressions with mixed derivatives as well as more general convolution-integral expressions (depending on the order of the truncation, and where in the analysis the truncations are performed). The work by Jones & Young (1994) also used a variation of this approach, but in an asymptotic framework (also capitalizing on the centre manifold theory). In this theory, the results were able to capture some features of the early-time behaviour, but were not able to capture exponentially decaying-in-time modes; most likely, this was because the initial condition was not represented as a source term in the averaged equation.

(vi) *Formulations that include an initial condition source.* Most studies on Taylor dispersion have not been overly interested in the relaxation of the initial condition; often, the only initial conditions examined are either uniform pulses, or delta-like impulses (Vedel, Hovad & Bruus 2014). Few papers have been developed with the realization that the initial condition can manifest as a decaying source term in the averaged mass balance when more complex initial conditions are considered. Philip (1963) was aware that essentially forcing the upscaled mass balance to be of a convection–dispersion form would limit the choices available for the initial condition. The importance of the initial condition was also apparent to Lighthill (1966), although he was not able to complete a general analysis for his solution. In a series of papers, Gill and Sankarasubramanian derived macroscale models for unsteady solute dispersion considering uniform (Gill & Sankarasubramanian 1970) and non-uniform (Gill & Sankarasubramanian 1971; Sankarasubramanian & Gill 1973) slugs as initial conditions. Connections between these works and the developments presented here are derived in additional detail in the supplementary material. The work by Yu (1979) began by including the initial condition in the macroscale balance, but later this term was dropped resulting in a final theory that does not have a source arising from the initial configuration of the system. In the papers by Degance and Johns (1978*b*, 1980), the influence of the initial condition appeared as part of the theory. However, this work specifically considered a subset of possible initial conditions that work with the theory; in particular, separable product forms for the initial condition function were allowable, but sums of any two such functions were excluded (Degance & Johns 1978*a*, § 4). Smith (1981*b*) identified and described three stages of unsteady dispersion from a point source using a ray-tracing method. Haber & Mauri (1988) used a stochastic Markovian particle approach to examine the role of the initial condition; explicit asymptotic solutions were derived that illustrated that the observed dispersion coefficient depended upon the initial location of the particles. From a moment-based approach, the paper by Dentz & Carrera (2007) similarly considered delta-type impulses placed near the centre versus the wall initially; in that work, a distinction between local and global dispersion was defined to distinguish between the dispersion of a single delta impulse versus the dispersion observed for an extended

source. Meng & Yang (2018) adopted a perspective similar to that of Dentz & Carrera (2007), but solved for the effective dispersion coefficient via eigenfunction expansions. They also found that the dispersion coefficient was a function of the initial configuration. In a sequence of two papers, Wood (2009, equations (17) and (29)) and Wood & Valdés-Parada (2013) developed theories that explicitly included the influence of the initial condition as a source term in the averaged mass balance equation; this process was then used by Ostvar & Wood (2016) to describe the average diffusion and reaction from an initial (unmixed) configuration. Independently, Balakotaiah & Ratnakar (2010) also developed balance expressions specifically for the Taylor dispersion problem in which source terms arising from the initial conditions were employed; a Kramers–Moyal-like expansion was used, but truncated at second order. This model had the ability to predict significant skewness in the distribution (Ratnakar & Balakotaiah 2011). While not unique to the models discussed, the prediction of skewness is an important check on the physics of the solution, since skewness should become non-zero at early times (even if the initial condition is symmetric), and it should asymptotically approach zero in the long-time regime.

   In the remainder of the background section, we elaborate a little more on two specific issues that arise in the previous work summarized above: (i) The use of infinite-order (Kramers–Moyal-type) expansions, and (ii) the neglect of source terms arising from the initial condition.

2.2. *Comments regarding Kramers–Moyal-type expansions*

There are two significant problems with the infinite Kramers–Moyal-type expansions. First, because of the specific macroscale representation that is chosen, the resulting theory generates dispersion coefficients that may depend upon the initial condition (Degance & Johns 1980). Second, research on the positivity of such series was apparently not known to progenitors of this approach. It is now known that such series either (i) converge exactly at second order and are strictly positive, or (ii) require all terms in order to converge and remain strictly positive (Marcinkiewicz 1939; Pawula 1967). Although Kramers–Moyal expansions do (under appropriate conditions) converge, they do not converge in a highly physical way when strictly positive results are sought (as in the case of concentration). In other words, any finite truncation of a Kramers–Moyal expansion of order greater than two generates results that are negative in some parts of the domain. This problem has been discussed in the literature (e.g. Risken & Vollmer 1979), but it is not generally well recognized in many disciplines outside of physics. The fact that negative concentrations are essentially guaranteed to occur for truncated expansions does not negate the usefulness of such expansions (Risken & Vollmer 1979), but it does demand careful analysis and cautiousness when interpreting results. The problem of negative concentrations is most acute at early times (cf. Wood & Valdés-Parada (2013, § 7)). Some of the problems associated with the Kramers–Moyal expansion can be eliminated by the recognition that such expansions can often be represented by non-local convolution expressions (Mercer & Roberts 1990, 1994). It is easy to show that a general non-local transport formulation can be written as an infinite-order expansion of local derivatives (this is illustrated in the supplementary material), although the converse of this is not always true. However, given that the macroscale equation form adopted in most of these examples is posited as an ansatz (e.g. Gill & Sankarasubramanian (1970, equations (5) and (8))), then it might be reasonable to start directly with the convolution rather than the series form. For the specific problem of Taylor dispersion investigated in this

paper, we are able to show that the solution is generally of a convolution form, giving further support that this kind of approach might have a stronger physical basis. Regardless, the non-local formulations generally avoid the problems associated with negative concentrations, and, in appropriate limits, reduce to more familiar local macroscale equations.

2.3. *Influence of the initial conditions*

As the review of the literature above suggests, the importance of the initial configuration at early times has been identified (although not resolved) previously. The problem can be most easily thought about in the context of layered media which, although not a tube, is still frequently represented as a Taylor dispersion process, (e.g. Fried & Combarnous 1971; Gelhar, Gutjahr & Naff 1979; Lake & Hirasaki 1981; Chikwendu & Ojiakor 1985; Chikwendu 1986*a*,*b*). Consider an initially uniform slug of solute in a system of horizontal layers of alternating high and low conductivity that is transported by a forward-flow/reverse-flow transport cycle. If a constant pressure gradient is applied for a short period of time such that the mean flow is right to left, the solute is displaced more in the high-velocity layers than in the low ones, leading to a spatially staggered final condition. Now, suppose the flow is stopped, and the pressure gradient reversed. The staggered configuration is the initial condition for a new transport process. Assuming that the total time interval is small enough so that transverse diffusion does not spread the solute very much between layers, the resulting motion is largely reversible. Upon reversing the flow field, a seemingly spread-out initial condition converges to become more organized (and less spread out). In the limit of zero diffusion, the final result of this first forward and then reversed transport would be exactly identical to the initial condition.

This kind of process creates some conceptual problems. Seemingly, the second moment of the process described above first increases and then, upon flow reversal, decreases in time. If one interprets the effective dispersion coefficient as being the classical one half of the time rate of change of the second spatial moment, then one predicts a negative dispersion coefficient for the second half of the cycle. This problem has been recognized in part as one that involves the distinction between spreading and mixing (e.g. Dentz & Carrera 2007). However, one is still left with a macroscale equation that contains an apparent dispersion coefficient that can be negative under some conditions. The problem is of the inverse heat equation type, which is known to be ill posed (Weber 1981). The solutions to such problems are lossy in general and can lead to significant problems in interpretation. Of more importance, however, is that such expressions are inconsistent with macroscale thermodynamics (Gray & Miller 2009; Gray & Miller 2014, chap. 10; Miller *et al.* 2018). Thus, if one hopes to solve problems in a manner consistent with the thermodynamics appropriate to the upscaled system, a negative dispersion coefficient is not a suitable choice. The use of higher-order diffusive-type equations (such as the method of Gill & Sankarasubramanian (1970)) does not entirely remediate this problem. Although there are more degrees of freedom in these kinds of expansions to accommodate the influence of initial conditions, the fundamental problem is that the initial conditions enter the problem as source terms that are independent of the spatial derivatives of the average concentration. Thus, some initial configurations, when forced to fit a transport equation that contains no term representing a source, may predict non-physical behaviour for the dispersion coefficient, depending upon the underlying theory.
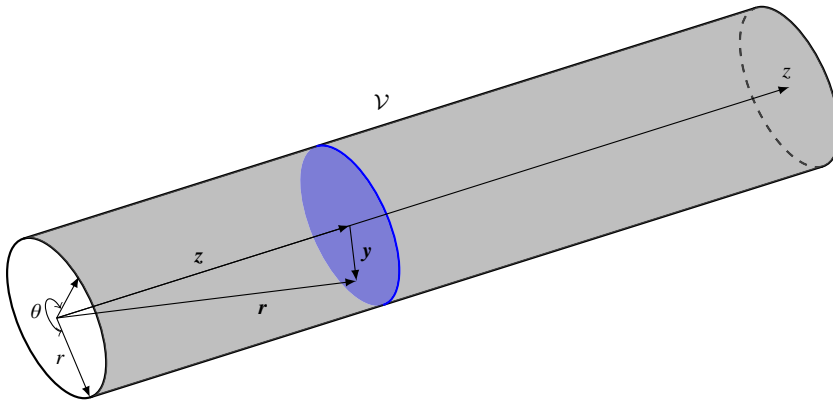
FIGURE 1. A subset of the domain, $\mathcal{V}$, geometry indicating coordinate directions and example vectors. Each vector (e.g. $\boldsymbol{r}$, $\boldsymbol{z}$ and $\boldsymbol{y}$) can be expressed in a cylindrical coordinate system for analytical or numerical computations. Note the distinction between the coordinates r and z versus the vectors $\boldsymbol{r}$ and $\boldsymbol{z}$. In this figure and the figures that follow, the aspect ratio of the radial, r, to longitudinal, z, coordinates has been increased to improve the presentation of results.

3. Problem formulation

One of the primary purposes of this work is to examine the early-time dispersion process in the context of an averaging theory, but maintaining the influence of the initial configuration throughout the averaging process. Our proposed model seeks to remedy the problems identified above by assuring that both the concentration and the dispersion coefficient remain positive over the entire (time and space) domain of the problem.

We consider passive transport of a chemical species in a cylindrical domain specified by $\boldsymbol{r} \in \mathcal{V}$, where $\boldsymbol{r} \equiv (r, \theta, z)$ (figure 1). For the cases considered in this analysis, we assume that the solute does not interact significantly with the external boundaries perpendicular to the axis of the tube. Thus, for concreteness, we set the external planes perpendicular to the tube axis at locations $z \to \pm\infty$, respectively. The initial solute distribution is assumed to be a known function of position denoted by $\varphi(r, \theta, z)$. The transport equations at the microscale for a solute in terms of molar concentrations can be written in cylindrical coordinates as

Microscale mass balance

$$\frac{\partial c}{\partial t} + v_z(r)\frac{\partial c}{\partial z} = \frac{\mathcal{D}}{r}\frac{\partial}{\partial r}\left(r\frac{\partial c}{\partial r}\right) + \frac{\mathcal{D}}{r^2}\frac{\partial^2 c}{\partial \theta^2} + \mathcal{D}\frac{\partial^2 c}{\partial z^2}, \tag{3.1a}$$

$$\text{B.C.1} \quad -\mathcal{D}\frac{\partial c}{\partial r} = 0 \quad \text{at } r = 0, a, \tag{3.1b}$$

$$\text{B.C.2a} \quad c \to 0 \quad z \to \pm\infty, \tag{3.1c}$$

$$\text{B.C.2b} \quad -\mathcal{D}\frac{\partial c}{\partial z} \to 0 \quad z \to \pm\infty, \tag{3.1d}$$

$$\text{I.C.} \quad c(r, \theta, z, 0) = \varphi(r, \theta, z). \tag{3.1e}$$

Here, and for the equations that follow, 'B.C.' indicates a boundary condition, and 'I.C.' an initial condition. For the formulation given by (3.1a)–(3.1e), it has been

assumed that the mole fraction of the solute is small enough compared to unity so that Fick's law is applicable. In addition, the following reasonable assumptions have been made in order to simplify the volume averaging process: (i) the fluid flow is incompressible, Newtonian, and steady; and (ii) the fluid pressure at the external boundaries is uniform leading to a θ-symmetric velocity field given by the well-known expression depending only on the distance, r, from the centreline of the tube

$$v_z(r) = 2U \left(1 - \frac{r^2}{a^2} \right). \tag{3.2}$$

With this information, the problem is fully specified.

4. Upscaling process

4.1. *Preliminaries*

In this paper, we upscale (or coarse grain) the system concentration field dynamics by simple spatial averaging against a weighting function. The underlying assumption in this approach is that the system can be sensibly represented by governing equations at more than one scale of resolution due to the multiscale (in space, time or both) characteristics of the system. Practically, this means that there exist averaging operators that sufficiently smooth spatial fluctuations in concentration such that an averaged behaviour is useful for understanding the system evolution. In this work, the averaging process is a somewhat modified form of volume averaging theory (which is frequently applied to porous materials, (Whitaker 1999)), but we do not stress this point further.

In the remainder of this paper, we adopt cylindrical coordinates (figure 1); thus, for any function ψ, we have $\psi(\boldsymbol{r}) = \psi(r, \theta, z)$. The symbol $\boldsymbol{z}$ (set in bold font) is used exclusively to indicate the centroid of the averaging domain; therefore, this vector always has the component form $\boldsymbol{z} = (0, 0, z)$. The averaging operator $\langle \cdot \rangle$ for Taylor dispersion is often defined with respect to weighting function w by (Degance & Johns 1978a)

$$\langle \psi \rangle|_{(z,t)} = \int_{r \in \mathcal{V}(z)} \psi(\boldsymbol{r}, t) w(\boldsymbol{r} - \boldsymbol{z}) \, \mathrm{d}V(\boldsymbol{r}). \tag{4.1}$$

Or, in the cylindrical coordinate system

$$\langle \psi \rangle|_{(z,t)} = \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} \psi(r, \theta, \zeta, t) w(r, \theta, \zeta - z) r \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\zeta. \tag{4.2}$$

For this work, we adopt the classical area average, so that $w(\boldsymbol{r} - \boldsymbol{z}) = w(r, \theta, \zeta - z) = 1/(\pi a^2)\delta(z - \zeta)$, which yields the conventional area average (note, other weighting functions have been adopted for Taylor dispersion, cf. Degance & Johns (1978a))

$$\langle \psi \rangle|_{(z,t)} = \frac{1}{\pi a^2} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} \psi(r, \theta, z, t) r \, \mathrm{d}r \, \mathrm{d}\theta. \tag{4.3}$$

We refer to the filtered functions $\langle \psi \rangle$ as macroscale quantities to distinguish them from their microscale counterparts, ψ.

In addition to the definition of the averaging operator, it is also useful to define the following decompositions:

$$c(\boldsymbol{r}, t) = \langle c \rangle|_{(z,t)} + \tilde{c}(\boldsymbol{r}, t) \tag{4.4a}$$

$$v_z(r) = U + \tilde{v}_z(r). \tag{4.4b}$$

Here, $U = \langle v_z \rangle|_z$ whereas $\tilde{c}$ and $\tilde{v}_z$ denote spatial deviations from their corresponding averages. Note that these deviations are defined in a manner that is different from the conventional definition used in the volume averaging theory. The above decompositions are the same as those proposed by Gray (1975); that is, the deviations in each averaging domain are defined relative to the value of the average at the longitudinal centroid of the tube. There are some advantages to this definition; in particular, we have the following identities:

$$\langle \tilde{c} \rangle|_{(z,t)} = 0, \tag{4.5a}$$

$$\langle \tilde{v}_z \rangle|_z = 0. \tag{4.5b}$$

We will occasionally use the term cross-sectional average (CSA) to indicate averaged properties.

4.2. *Averaging the microscale balance equation*

To obtain the macroscale mass balance equation, we apply the averaging operator defined above to the microscale balance. Note that, although we have been careful to explicitly list the independent variables in the definitions above, we do so in the remainder of the paper only when it is necessary for emphasis or clarity. Upon applying the averaging operator (4.3) to the microscale mass balance equations, the upscaled problem takes the form

$$\frac{\partial \langle c \rangle}{\partial t} = \mathcal{D} \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} - \left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle, \tag{4.6a}$$

$$\text{B.C. Macro 1a} \quad \langle c \rangle|_{(z,t)} = 0 \quad z \to \pm\infty, \tag{4.6b}$$

$$\text{B.C. Macro 1b} \quad -\mathcal{D} \left. \frac{\partial \langle c \rangle}{\partial z} \right|_{(z,t)} = 0 \quad z \to \pm\infty, \tag{4.6c}$$

$$\text{I.C.1} \quad \langle c \rangle|_{(z,0)} = \langle \varphi \rangle|_z. \tag{4.6d}$$

Several steps have been taken into account in order to derive (4.6a). First, interchange of the time derivative and the averaging operator is allowed within the accumulation term due to the fact that the averaging domain does not change with time. Thus we have

$$\left\langle \frac{\partial c}{\partial t} \right\rangle = \frac{\partial \langle c \rangle}{\partial t}. \tag{4.7}$$

Directing our attention to the diffusion term, the microscale boundary conditions lead to the identity

$$\left\langle \mathcal{D} \frac{\partial^2 c}{\partial z^2} \right\rangle = \mathcal{D} \frac{\partial^2 \langle c \rangle}{\partial z^2}. \tag{4.8}$$

For the convection term, we have (on the basis of (4.4a)–(4.5b))

$$\left\langle v_z \frac{\partial c}{\partial z} \right\rangle = U \frac{\partial \langle c \rangle}{\partial z} + \left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle. \tag{4.9}$$

Finally, for the two macroscopic boundary conditions given in (4.6b) and (4.6c), we have imposed the condition that $\tilde{c}(r, \theta, z) \to 0$ and $\partial \tilde{c}/\partial z(r, \theta, z) \to 0$ as $|z| \to \infty$. Up to this point, we have employed no approximations to the result; the averaged model represented by (4.6a)–(4.6d) is exact, at least to the extent that the original microscale balance equations are valid.

5. Deviation equations

Equation (4.6a) is unclosed because it is not expressed exclusively in terms of the average concentration. To close the problem, a set of ancillary balances are developed for the microscale deviations of $\tilde{c}$. Motivated by the definition of the decompositions, we develop the deviation balance by subtracting the averaged equation (4.6a) from the microscale equation (3.1a). In addition, the boundary conditions can be developed by using the decomposition given by (4.4a), and the initial condition is determined by subtracting the averaged initial condition from its microscale counterpart. The resulting equations can be written as follows:

$$
\underbrace{\frac{\partial \tilde{c}}{\partial t}}_{accumulation} + \underbrace{U \frac{\partial \tilde{c}}{\partial z}}_{\substack{mean \\ convection}} + \underbrace{\tilde{v}_z \frac{\partial \tilde{c}}{\partial z}}_{\substack{deviation \\ convection}} - \underbrace{\left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle}_{\substack{non\text{-}local \\ convection}}
$$
$$
\underbrace{- \frac{\mathcal{D}}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \tilde{c}}{\partial r} \right) - \frac{\mathcal{D}}{r^2} \frac{\partial^2 \tilde{c}}{\partial \theta^2} - \mathcal{D} \frac{\partial^2 \tilde{c}}{\partial z^2}}_{diffusion} = - \underbrace{\tilde{v}_z \frac{\partial \langle c \rangle}{\partial z}}_{\substack{local \\ convective\ source}}, \tag{5.1a}
$$

$$
\text{B.C.1} \quad -\mathcal{D} \frac{\partial \tilde{c}}{\partial r} = 0 \quad \text{at } r = 0, a, \tag{5.1b}
$$

$$
\text{B.C.2b} \quad \tilde{c} \to 0 \quad z \to \pm\infty, \tag{5.1c}
$$

$$
\text{B.C.2b} \quad -\mathcal{D} \frac{\partial \tilde{c}}{\partial z} \to 0 \quad z \to \pm\infty, \tag{5.1d}
$$

$$
\text{I.C.1} \quad \tilde{c}(r, \theta, z, 0) = \underbrace{\tilde{\varphi}(r, \theta, z)}_{source} \quad (r, \theta, z) \in \mathcal{V}. \tag{5.1e}
$$

For the developments that follow, it is convenient to adopt a Lagrangian perspective based upon an inertial frame of reference that moves with the centre of mass with the system. To this end, let us define $z' = z - Ut$, $t' = t$, so that

$$
\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} + U \frac{\partial}{\partial z'}. \tag{5.2}
$$

We omit the explicit writing of the prime coordinates in the developments that follow to keep the notation simple. Making the translation to Lagrangian coordinates, the governing differential equation for the concentration deviations becomes

$$
\frac{\partial \tilde{c}}{\partial t} + \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} - \left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle - \frac{\mathcal{D}}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \tilde{c}}{\partial r} \right) - \frac{\mathcal{D}}{r^2} \frac{\partial^2 \tilde{c}}{\partial \theta^2} - \mathcal{D} \frac{\partial^2 \tilde{c}}{\partial z^2} = -\tilde{v}_z \frac{\partial \langle c \rangle}{\partial z}. \tag{5.3}
$$

In (5.3), the term $\langle \tilde{v}_z (\partial \tilde{c}/\partial z) \rangle$ is non-local in the variable r. In this context, we use the term non-local to denote quantities that involve time or space locations in addition to the local ones. Solutions to this balance equation that maintain the non-local term are mathematically tractable, but such solutions remain an active area of research (Chasseigne, Chaves & Rossi 2006; Ignat & Rossi 2007; García-Melián & Rossi 2009).
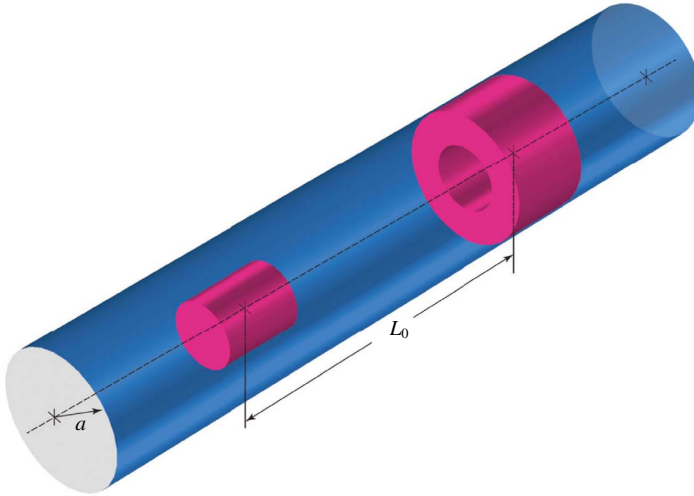
FIGURE 2. Length scales for the Taylor dispersion problem. The characteristic length associated with the initial solute configuration (illustrated in red) is L_0. Because of solute spreading, the generic length scale $L = L(t)$ (defined for all times) is asymptotically larger than L_0. This figure is meant to be primarily schematic; the actual initial conditions examined in this work are smoothed versions of similar configurations.

5.1. *Elimination of the non-local term*

Rather than maintaining the non-local term, our approach is to develop reasonable constraints that allow us to neglect its influence. We do this by establishing characteristic time and length scales, and then developing constraints that indicate when the desired simplification is valid. Although these time and length scales can be given formal, computable definitions (e.g. such as defining them by the integral scale of the fields involved (Wood & Valdés-Parada 2013)), formal evaluation of scales is usually not necessary.

For the Taylor dispersion problem, there are several distinct length scales that can be defined. We have identified two important length scales in figure 2. The size of the macroscale length, L_0, is described as being approximately the size of the initial condition; however, the validity of this estimate depends upon the particular structure of the initial condition (this is discussed further in § 8.2). The characteristic length of the radial diffusion process is of the order of the tube radius, a, for this particular application.

The non-local term and the convection term on the left-hand side of (5.3) are of the same order of magnitude. Thus, our desired restriction at this juncture would be to impose

$$\tilde{v}_z \frac{\partial \tilde{c}}{\partial z} - \left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle \ll \frac{\mathcal{D}}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \tilde{c}}{\partial r} \right) + \cdots. \tag{5.4}$$

For complex problems where one wishes to find all possible asymptotically simplified models, there are specific methods and algorithms that can be followed to construct the simplified models (see Yip (1996) for a fuller discussion). Because our restriction is rather simple (and because a full analysis is tedious), we only sketch out the result here. To start, we note that the left-hand side of (5.4) is always less than or equal to

$\tilde{v}_z \partial \tilde{c}/\partial z$. Then, we non-dimensionalize and rearrange the problem as follows (where $Z = z/L_0$, $R = r/a$, $\tilde{V}_z = \tilde{v}_z/U$, $\tilde{C} = \tilde{c}/C_0$ and $C_0 = \max(\tilde{\phi})$):

$$\left(\frac{a^2 U}{L_0 \mathcal{D}}\right) \tilde{V}_z \frac{\partial \tilde{C}}{\partial Z} = \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial \tilde{C}}{\partial R}\right) + \cdots . \tag{5.5}$$

Now, we set $\varepsilon = a^2 U/(L_0 \mathcal{D})$. By assumption, $\tilde{V}_z \partial \tilde{C}/\partial Z = \boldsymbol{O}(1)$ and the first term on the right-hand side of (5.5) is of $\boldsymbol{O}(1)$. Examining

$$\varepsilon \tilde{V}_z \frac{\partial \tilde{c}}{\partial Z} = \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial \tilde{c}}{\partial R}\right) + \cdots , \tag{5.6}$$

it is clear that the left-hand side of this expression (and, hence, the left-hand side of (5.4)) can be neglected under the conditions

$$\frac{Ua}{\mathcal{D}}\frac{a}{L_0} \ll 1 \quad \text{or, equivalently,} \quad Pe\frac{a}{L_0} \ll 1. \tag{5.7}$$

The symbols '$\ll$' and '$\boldsymbol{O}$' have the conventional meanings adopted in perturbation theory (cf. Bender & Orszag (1978, chapters 3.4 and 7)).

A few additional comments are helpful here. First, we note that this approximation may fail at early times if the concentration gradients in the initial condition are large (e.g. step functions); for these early-time solutions, the longitudinal derivative term would need to be maintained. In this work, we consider initial conditions that are reasonably smooth at $t = 0$ (i.e. the first two derivatives are continuous, and correspond to a value of L_0 that is not much smaller than a). Second, this restriction is almost certainly overly severe once the relaxation of the initial condition has progressed for even a relatively small time. Often an approximation such as the inequality (5.4) is valid, even when the two sides are of the same order of magnitude. We describe the results of the error involved in this approximation in additional detail in §8.2, where we compute it directly to show that the approximation is reasonable for the Péclet numbers investigated in this paper. It is worth noting that the constraint given in (5.7) is identical (except for a factor of 4) to the one imposed by Taylor (1954)

$$Pe_T = \frac{a^2}{4\mathcal{D}}\frac{U}{L} \ll 1. \tag{5.8}$$

Here, Pe_T is the Taylor-type Péclet number, which arises naturally when we non-dimensionalize the transport equation. The diffusion time scale arises from considerations of the conventional free-field diffusion relationship with variance $\sigma^2 = 4\mathcal{D}t$ (and, hence, $L \sim \sqrt{4\mathcal{D}t}$). The convective time scale comes from a simple estimate of the z-derivative of the initial condition. With the exception of Taylor (1954) and Aris (1956), the literature has tended to use the definition that is based on a simple ratio of the convective to diffusive fluxes, of the form (assuming both convective and diffusive fluxes are characterized by a length scale equal to the tube radius, a)

$$Pe = \frac{Ua}{\mathcal{D}} = \frac{4L}{a}Pe_T. \tag{5.9}$$

Note that for the simulations reported below with the parameters in table 1, $Pe_T = 1$ corresponds to $Pe = 80$.

Parameters	Value	Definition
L_0	0.20 m	Separation distance between centres of mass of initial concentration distribution
a	0.01 m	Radius of the tube
$\mathcal{D}$	1×10^{-9} m^2 s^{-1}	Molecular diffusion coefficient for the solute
c_0	1 mol m^{-3}	Initial concentration constant in (7.10)
α_1	1.1 m	Constant for (7.11a)
α_2	1.1 m	Constant for (7.11b)
β_1	0.125 m	Constant for (7.11a)
β_2	0.325 m	Constant for (7.11b)
σ_1	0.03 m	Constant for (7.11a)
σ_2	0.03 m	Constant for (7.11b)

TABLE 1. Parameters used in the simulations.

With the approximation given by the inequality (5.4) and the use of Lagrangian coordinates, the resulting set of equations for $\tilde{c}$ becomes

$$\frac{\partial \tilde{c}}{\partial t} = \frac{\mathcal{D}}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{c}}{\partial r}\right) + \frac{\mathcal{D}}{r^2}\frac{\partial^2 \tilde{c}}{\partial \theta^2} + \mathcal{D}\frac{\partial^2 \tilde{c}}{\partial z^2} - \underbrace{\tilde{v}_z(r)\left.\frac{\partial \langle c \rangle}{\partial z}\right|_{(z,t)}}_{source} \quad (r,\theta,z) \in \mathcal{V}, \quad (5.10a)$$

$$\text{B.C.1} \quad -\mathcal{D}\frac{\partial \tilde{c}}{\partial r} = 0 \quad \text{at } r = 0, a, \quad (5.10b)$$

$$\text{B.C.2b} \quad \tilde{c} \to 0 \quad z \to \pm\infty, \quad (5.10c)$$

$$\text{B.C.2b} \quad -\mathcal{D}\frac{\partial \tilde{c}}{\partial z} \to 0 \quad z \to \pm\infty, \quad (5.10d)$$

$$\text{I.C.1} \quad \tilde{c}(r,\theta,z,0) = \underbrace{\tilde{\varphi}(r,\theta,z)}_{source} \quad (r,\theta,z) \in \mathcal{V}. \quad (5.10e)$$

Notice that there are only two source terms in the initial and boundary-value problem. Thus, it is the interplay of the initial configuration and the local convective source term appearing in the deviation balance equation that essentially drives the Taylor dispersion process.

The above simplified closure problem is a local, linear, non-homogeneous parabolic equation which has well-known solutions. With the conventional assumptions (e.g. boundedness of the initial condition, integrability of the initial condition, smoothness of the boundary conditions, etc.), the solution to the problem for solute transport can be put in an integral formulation as (cf. Polyanin & Nazaikinskii 2015)

$$\begin{aligned}
\tilde{c}(r,\theta,z,t) &= \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r,\rho,\theta,\vartheta,z,\zeta,t)\tilde{\varphi}(\rho,\vartheta,\zeta)\rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \,\mathrm{d}\zeta \\
&\quad - \int_{\tau=0}^{\tau=t} \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r,\rho,\theta,\vartheta,z,\zeta,t-\tau)\tilde{v}_z(\rho)\rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \left.\frac{\partial \langle c \rangle}{\partial z}\right|_{(\zeta,\tau)} \,\mathrm{d}\zeta \,\mathrm{d}\tau,
\end{aligned}$$
(5.11)

where $G(r,\rho,\theta,\vartheta,z,\zeta,t-\tau)$ is the Green's function for this problem, which solves the initial and boundary-value problem given in appendix A by (A 2a)–(A 2e).

6. Closed problem

6.1. *Non-local solution*

Using the unsimplified integral solution defined above by (5.11), we can develop a non-local but closed macroscale problem that takes the form

$$\frac{\partial \langle c \rangle}{\partial t} = \mathcal{D}\frac{\partial^2 \langle c \rangle}{\partial z^2} + \frac{\partial}{\partial z} \int_{\tau=0}^{\tau=t} \int_{\zeta=-\infty}^{\zeta=\infty} \kappa(z, \zeta, t-\tau) \left.\frac{\partial \langle c \rangle}{\partial z}\right|_{(\zeta,\tau)} \mathrm{d}\zeta\, \mathrm{d}\tau - U \left.\frac{\partial \langle c \rangle}{\partial z}\right|_{(z,t)} + s^*(z, t),$$
(6.1)

where, for simplicity in notation, we have introduced

$$\kappa(z, \zeta, t-\tau) = \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} \langle \tilde{v}_z G \rangle|_{(\rho,\vartheta,z,\zeta,t-\tau)} \tilde{v}_z(\rho)\rho\, \mathrm{d}\rho\, \mathrm{d}\vartheta \qquad (6.2a)$$

and

$$s^*(z, t) = -\int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} \left\langle \tilde{v}_z \frac{\partial G}{\partial z} \right\rangle\bigg|_{(\vartheta,\zeta,t)} \tilde{\varphi}(\rho, \vartheta, \zeta)\rho\, \mathrm{d}\rho\, \mathrm{d}\vartheta\, \mathrm{d}\zeta. \qquad (6.2b)$$

Non-local solutions have been proposed to describe a number of interesting transport phenomena (cf. Eringen 1978; Smith 1981*a*; Cushman & Ginn 1993; Deng, Cushman & Delleur 1993; Neuman 1993; Deng & Cushman 1995). While they have some advantages, they also come with significant costs. Among these are that (i) analytical solutions are virtually non-existent, and (ii) because the solution at every point depends upon every other point in the domain, numerical schemes often have to deal with full rather than sparse matrices.

6.2. *Localized solution*

At this juncture in the analysis, we consider the effects of imposing two additional approximations: (i) the length scale constraint $a/L_0 \ll 1$, and (ii) a time scale constraint $t^*/T^* \ll 1$ (where t^* is a characteristic microscale process time, and T^* is a characteristic macroscale process time). In appendix A we show that, for the Taylor problem in the range of Péclet numbers that we examine, the single constraint

$$\frac{a}{L_0} \ll 1 \qquad (6.3)$$

is sufficient to ensure that both the length and time scale constraints are met (a detailed derivation can be found in Wood & Valdés-Parada (2013)). Note that this constraint is generally less restrictive (at least for $Pe > 1$) than the constraint already imposed by the inequality (5.7). Technically, it is not necessary to impose any constraints on the problem. However, if we do not impose the inequalities (5.7) and (6.3), the result is a fully time- and space-non-local theory whose Green's functions would be determined by an integro-differential equation with no known analytical solutions. Such a result would be unlikely to be simpler than directly solving the microscale balance given by (3.1*a*)–(3.1*e*).

With the constraints given by (5.7) and (6.3) in place, the localized solution for the set of balance equations associated with the solute ((5.10*a*)–(5.10*e*)) is given by (appendix A, (A 22))

$$\tilde{c}(r, \theta, z, t) = b(r, t) \left.\frac{\partial \langle c \rangle}{\partial z}\right|_{(z,t)} + \Phi(r, \theta, z, t), \qquad (6.4)$$

where b and Φ are known as closure variables. These functions are defined in terms of the Green's function by

$$b(r, t; \tilde{v}_z) = - \int_{\tau=0}^{\tau=t} \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r, \rho, \theta, \vartheta, z, \zeta, t-\tau) \tilde{v}_z(\rho) \rho \, d\rho \, d\vartheta \, d\zeta \, d\tau,$$

(6.5a)

$$\Phi(r, \theta, z, t; \tilde{\varphi}) = \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r, \rho, \theta, \vartheta, z, \zeta, t) \tilde{\varphi}(\rho, \vartheta, \zeta) \rho \, d\rho \, d\vartheta \, d\zeta.$$

(6.5b)

We note that both b and Φ are functionals (of $\tilde{v}_z$ and $\tilde{\varphi}$, respectively); ordinarily, the explicit dependence on the function will be supressed in the notation unless it is needed for clarity. In § A.4 we show that Φ is linear in the initial condition deviation function $\tilde{\varphi}$. We also note that Φ is an exponentially decaying function of time. At early times, the magnitude of this function can be significant, and can create (apparent) deviations from Fickian behaviour (especially in the second moment), even though the spreading process is a Fickian one.

The solution form given by (6.4) is similar in form to the solution proposed by Paine, Carbonell & Whitaker (1983, equation (30)) for Taylor dispersion. Paine *et al.* (1983) correctly predicted that the solution for the deviation concentration should involve an independent function of time, they did not understand that this function is related to the initial condition for the problem (see appendix A). A fully local averaged balance equation can be developed, taking the form:

Local macroscale equation

$$\underbrace{\frac{\partial \langle c \rangle}{\partial t}\bigg|_{(z,t)}}_{accumulation} = \underbrace{D^*(t)\, \frac{\partial^2 \langle c \rangle}{\partial z^2}\bigg|_{(z,t)}}_{diffusive\ transport} - \underbrace{U\, \frac{\partial \langle c \rangle}{\partial z}\bigg|_{(z,t)}}_{convective\ transport} + \underbrace{s^*(z, t)}_{\substack{non\text{-}conventional \\ source}}\,,$$

(6.6a)

$$\text{B.C. Macro 1a} \quad \langle c \rangle|_{(z,t)} = 0 \quad z \to \pm\infty,$$

(6.6b)

$$\text{B.C. Macro 1b} \quad -\mathcal{D}\, \frac{\partial \langle c \rangle}{\partial z}\bigg|_{(z,t)} = 0 \quad z \to \pm\infty,$$

(6.6c)

$$\text{I.C.1} \quad \langle c \rangle|_{(z,0)} = \langle \varphi \rangle|_z \quad z \to \pm\infty.$$

(6.6d)

Here, the effective parameters are defined by

$$s^*(z, t) = -\left\langle \tilde{v}_z \frac{\partial \Phi}{\partial z} \right\rangle,$$

(6.7a)

$$D(t) = -\langle \tilde{v}_z b \rangle,$$

(6.7b)

$$D^*(t) = \mathcal{D} + D(t).$$

(6.7c)

The effective parameters D and D^* are the hydrodynamic dispersion and the total dispersion coefficients, respectively. This latter quantity is also called the Taylor dispersion coefficient, as proposed by Aris (1956). Note that both our non-local and local averaged equations are unusual in that they contain a source term, s^*. We also note that the effective dispersion coefficient, D^* depends only upon the closure variable b (and, is therefore independent of the initial condition, as shown in § A.3), whereas the non-conventional source term, s^* depends only upon the closure variable

Φ. It is interesting to compare these effective parameters with the results of Mercer & Roberts (1994, equation (3.2)); the correction terms derived by Mercer & Roberts (1994) show similarities with the effective parameters defined in (6.7a)–(6.7c).

The role of the source term is to account for the early-time memory of the initial distribution of the spatial deviations of the concentration. Although it is a source term in the balance equation, it is easy to show that the integral of this term over the domain $-\infty < z < \infty$ is identically zero (§ A.3). This indicates that the role of this term is only to redistribute mass within the domain; no mass is created or destroyed by this term.

Finally, a few comments regarding the effective parameters s^* and D^* are in order. First, we note that it can be shown that the effective dispersion coefficient, D^* is a strictly positive function for all time, regardless of the initial conditions (§ A.5). As mentioned previously, this is a necessary condition for the problem if one wants the solution to be consistent with macroscopic scale thermodynamics (Miller *et al.* 2018). This has important repercussions for previous works that have suggested the use of negative dispersion coefficients; this is examined in detail in the Discussion section. Second, we note that the source term, s^*, like the function Φ, is an exponentially decreasing function of time (§ A.5). Therefore, Φ decays toward zero as time grows large enough, recovering the conventional Taylor–Aris theory at asymptotic times. Although this term does decay to zero in time, the effect of this source term on the early stages of the diffusion–convection transport phenomenon is significant in some cases. This is discussed in additional detail in the following sections.

In the remainder of this paper, we adopt a strictly local representation; ultimately, we are able to show that this is an appropriate approximation for our system and range of parameters. Readers interested in examining how the non-local formulation can be solved numerically are directed to the details outlined by Deng *et al.* (1993).

7. Analytical solutions for the closure variables

Equations (6.5a)–(6.5b) provide the general integral form of the solutions; however, the Green's functions for particular cases must be determined and subsequently integrated to obtain explicit (series-form) solutions for b and Φ. Because the b-field depends only upon $\tilde{v}_z$, the solution for this problem is relatively straightforward. For Φ, the solution depends explicitly upon the particular initial condition selected.

7.1. *Analytical solution for the b-field*

In appendix A, the Green's function for the general problem is identified, and integrated. The result is (§ A.3)

$$b(r, t) = \frac{1}{4} \frac{Ua^2}{\mathcal{D}} \left(\frac{r^2}{a^2} - \frac{r^4}{2a^4} - \frac{1}{3} \right) + \frac{2Ua^2}{\mathcal{D}} \sum_{n=1}^{\infty} \frac{J_3(\lambda_n) J_0 \left(\lambda_n \frac{r}{a} \right)}{\lambda_n^3 J_0^2(\lambda_n)} \exp \left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t \right). \quad (7.1)$$

Note that b is not a function of θ or z. The eigenvalues λ_n can be computed by solving

$$J_1(\lambda_n) = 0, \quad n = 1, 2, 3, \ldots. \quad (7.2)$$

Equation (7.1) is equivalent to equation (16) in the work of Gill & Sankarasubramanian (1970). We note additionally here that in this derivation, it has been explicitly assumed that the initial time is represented by $t = 0$.
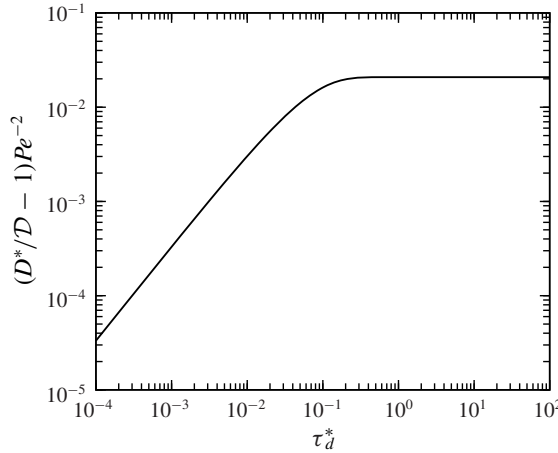
FIGURE 3. Evaluation of the dynamics of the dispersion coefficient using the analytical solution given in (7.5).

The effective parameters, D^* and s^* associated with the upscaled equation are given by (6.7a) and (6.7c). Substituting the two solutions above for b and Φ as required into the general expressions for the effective parameters gives the following results (§ A.5):

$$\frac{D^*(t)}{\mathcal{D}} = \left(1 + \frac{1}{48}\frac{U^2a^2}{\mathcal{D}^2}\right) - 4\frac{U^2a^2}{\mathcal{D}^2}\sum_{n=1}^{\infty}\left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)}\right)^2 \exp\left(-\lambda_n^2\frac{\mathcal{D}}{a^2}t\right), \qquad (7.3)$$

where we note that

$$\frac{1}{48}\frac{U^2a^2}{\mathcal{D}^2} = 4\frac{U^2a^2}{\mathcal{D}^2}\sum_{n=1}^{\infty}\left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)}\right)^2. \qquad (7.4)$$

We can also express this result in a form that is independent of Pe

$$\left(\frac{D^*(t)}{\mathcal{D}} - 1\right) Pe^{-2} = \frac{1}{48} - 4\sum_{n=1}^{\infty}\left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)}\right)^2 \exp(-\lambda_n^2\tau_d^*). \qquad (7.5)$$

Here, we have defined the dimensionless diffusion time variable

$$\tau_d^* = \frac{\mathcal{D}}{a^2}t, \qquad (7.6)$$

where $a^2/\mathcal{D}$ is the diffusive time scale as defined in (B 32). The dispersion coefficient predicted from this expression is plotted in figure 3. Note that (7.5) is an evolving function of time, reaching more than 99 % of the expected asymptotic value of $1/48$ for $\tau_d^* \gtrsim 0.3$.

7.2. *Analytical solution for the Φ-field*

For the Φ field, the final results depend on the particular initial condition that is investigated. However, we can provide the following general integral solution (the

solutions for particular initial conditions are discussed in the sections following):

$$\Phi(r, \theta, z, t) = B_{0,0}(z, t) + \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} B_{n,m}(\theta, z, t) \mathrm{J}_n\left(\frac{\lambda_{nm}r}{a}\right) \exp\left(-\frac{\lambda_{nm}^2 \mathcal{D}}{a^2}t\right), \qquad (7.7)$$

where, for convenience, we introduced the functions $B_{0,0}(z, t)$ and $B_{n,m}(\theta, z, t)$, which are defined by

$$B_{0,0}(z, t) = \frac{1}{\pi a^2} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z, \zeta, t) \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} \tilde{\varphi}(\rho, \vartheta, \zeta)\rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \,\mathrm{d}\zeta \qquad (7.8a)$$

$$B_{n,m}(\theta, z, t) = \frac{A_n \lambda_{nm}^2}{\pi a^2 (\lambda_{nm}^2 - n^2) \mathrm{J}_n^2(\lambda_{nm})} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z, \zeta, t) \int_{\vartheta=0}^{\vartheta=2\pi} \cos[n(\theta - \vartheta)]$$

$$\times \int_{\rho=0}^{\rho=a} \mathrm{J}_n\left(\lambda_{nm}\frac{\rho}{a}\right) \tilde{\varphi}(\rho, \vartheta, \zeta)\rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \,\mathrm{d}\zeta. \qquad (7.8b)$$

Here, G_2 is the Green's function in the axial direction and it is defined in appendix A (A 4b). The constants A_n are given by $A_1 = 1$ and $A_n = 2$, $n > 1$. The eigenvalues λ_{nm} are the positive roots of the transcendental equation $\mathrm{J}_n'(\lambda_{nm}) = 0$.

The results for the s^* field depend upon the particular initial configuration that is specified; therefore, we cannot provide explicit expressions until the initial condition is identified. In general, the expression is given by (6.7a), which, using (7.7), can be put in the form

$$s^*(z, t) = -\sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{\partial B_{n,m}}{\partial z} \left\langle \tilde{v}_z \mathrm{J}_n\left(\frac{\lambda_{nm}r}{a}\right) \right\rangle \exp\left(-\frac{\lambda_{nm}^2 \mathcal{D}}{a^2}t\right). \qquad (7.9)$$

Because s^* is a source term, the characteristics of the source can have a large impact on the solution. Explicit analytical expressions for a few specific initial conditions are derived in the next section. We note that for radially symmetric initial conditions, these general solutions are simplified somewhat due to symmetry (see § A.4).

As a final note, the functions D^* and s^* were developed explicitly for the case where the initial time is defined by $t = 0$. If the initial time is some positive value, T_1, then the functions D^* and s^* should be specified as functions of the time since injection of the initial condition, t', where $t' = t - T_1$. The important point is that D^* and s^* depend on time since injection of the initial condition, not on absolute time.

7.3. Analytical solution for specific initial conditions

For this work, we consider three different initial conditions (illustrated in figure 4, where flow in the tube is from left to right), and provide the analytical solutions for s^* for these three cases. Each initial condition is radially symmetric; this is not necessary for the general solution developed to this point, but to date we have found explicit series solutions only for these symmetry conditions (cf. the radially symmetric solution form for Φ given in § A.4). Each case consists of initial conditions in which two concentration distributions are separated spatially (specifically, the centre of mass is separated by $L_0 = 20$ cm, see table 1). In order to avoid the singular points that arise when differentiating discontinuous initial distributions, these were weighted in the z-direction by a non-compact smoothing function. For each initial condition
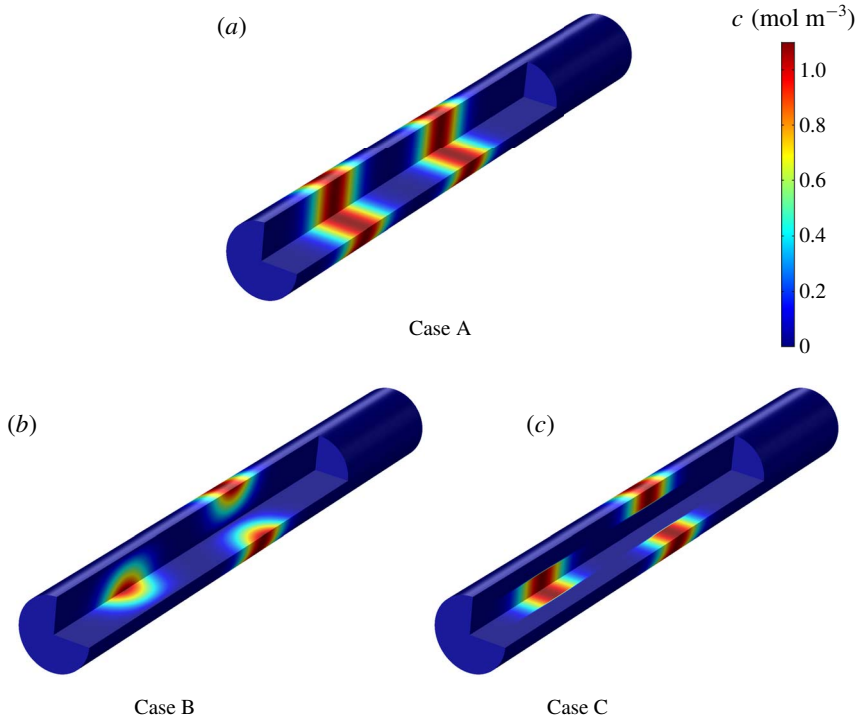
FIGURE 4. Three-dimensional representations for the three types of the initial concentrations. The flow direction is left to right. Case B and C initial configurations have radial dependence. For visualization purposes, only a portion of the domain is shown.

considered, the configuration could be decomposed into multiplicative radial and longitudinal components as follows:

$$\varphi(r, z) = c_0 R_1(r) Z_1(z) + c_0 R_2(r) Z_2(z). \tag{7.10}$$

This multiplicative decomposition is not a necessary part of the solution, but it does aid in the subsequent determination of analytical solutions to the closure problem. In the longitudinal direction, the initial condition function was the same for each of the three cases, and was specified by

$$Z_1(z) = \alpha_1 \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2}\right), \tag{7.11a}$$

$$Z_2(z) = \alpha_2 \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2}\right). \tag{7.11b}$$

The coefficients α_1, α_2, β_1, β_2, σ_1 and σ_2 are given in table 1. The functions R_1 and R_2 vary for each of the three cases considered, and are detailed below.

We note that a more accurate value for L_0 might be of the order of σ_1 and σ_2 rather than the spacing between the pulses. Note that with these estimates we have $a/\sigma_1 = a/\sigma_2 = 1/3$. While this quantity is less than one, it is in a region of the

constraint $a/L_0 \ll 1$ that is somewhat unclear in terms of validity. However, we explicitly compute errors associated with these approximations in §8.2; those results suggest that the approximation of separation of length scales is reasonably well met for the three problems investigated here.

7.3.1. *Case A: radially uniform slugs*

For this case, the initial condition is defined by two radially uniform slugs whose centre of mass is separated by L_0. The functions R_1 and R_2 are given by

$$R_1(r) = 1, \tag{7.12a}$$
$$R_2(r) = 1 \tag{7.12b}$$

i.e. there is no variation in the radial direction. For initial conditions with no radial variation, we have shown in appendix A that the solution for the source term is $s^*(z, t) = 0$. This can also be seen through the closure problem for Φ specified above. When initial conditions contain radially constant concentrations, the deviation concentrations are identically zero initially. This means that there are no source terms in the closure for Φ, and the problem generates only the zero solution.

7.3.2. *Case B: linear radial distributions*

For this case, each part of the initial condition was specified by a linear function. On the left (centred at $z = 12.5$ cm) the concentration was maximum in the centre, and decreased linearly toward the wall. On the right (centred at $z = 32.5$ cm) the concentration was greatest at the wall, and decreased linearly toward the centre. The functions R_1 and R_2 are given by

$$R_1(r) = 1 - \frac{r}{a}, \tag{7.13a}$$
$$R_2(r) = \frac{r}{a}. \tag{7.13b}$$

For this case, the source term has been derived in appendix B. The result is

$$
\begin{aligned}
s^*(z, t) \;=\; -4\pi c_0 U &\left[\frac{\sigma_1 \alpha_1 (z - \Xi_1(t) - \beta_1)}{\left(\sigma_1^2 + 4\mathcal{D}t\right)^{3/2}} \exp\left(-\frac{(z - \Xi_1(t) - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t}\right) \right. \\
&\left. - \frac{\sigma_2 \alpha_2 (z - \Xi_2(t) - \beta_2)}{\left(\sigma_2^2 + 4\mathcal{D}t\right)^{3/2}} \exp\left(-\frac{(z - \Xi_2(t) - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t}\right) \right] \\
&\times \sum_{n=1}^{\infty} \frac{\boldsymbol{H}_1(\lambda_n)}{\lambda_n^3} \frac{\mathrm{J}_3(\lambda_n)}{\mathrm{J}_0(\lambda_n)} \exp\left(-\frac{\lambda_n^2 \mathcal{D}}{a^2} t\right),
\end{aligned}
\tag{7.14}
$$

where $\boldsymbol{H}_1$ is the Struve H function (cf. equation (11.1.7) in Abramowitz & Stegun 1965). The quantities Ξ_1 and Ξ_2 are the functions defining the motion of the centre of mass of the two portions (left and right) of the initial condition. They are specified by

$$
\Xi_1(t) = \begin{cases} \dfrac{7}{5} Ut - \dfrac{4}{15} U \dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^* \\[2mm] \Xi_1(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^* \end{cases}
\tag{7.15a}
$$

$$\Xi_2(t) = \begin{cases} \dfrac{4}{5}Ut + \dfrac{2}{15}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\[2ex] \Xi_2(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*, \end{cases} \tag{7.15b}$$

where $t_d^* \equiv a^2/4\mathcal{D}$ is the diffusive time scale.

7.3.3. *Case C: step radial distributions*

For the third case, each part of the initial condition is specified by radial step functions that are complements of one another. The result is an initial condition that resembles a cylinder (with radius $a/2$) concentrated on the axis centred at $z = 12.5$ cm, and a hollow cylinder (with radial thickness equal to $a/2$) centred at $z = 32.5$ cm. The functions R_1 and R_2 are given by

$$R_1(r) = \begin{cases} 1, & 0 \leqslant r \leqslant \dfrac{a}{2}, \\[2ex] 0, & \dfrac{a}{2} < r \leqslant a, \end{cases} \tag{7.16a}$$

$$R_2(r) = \begin{cases} 0, & 0 \leqslant r \leqslant \dfrac{a}{2}, \\[2ex] 1, & \dfrac{a}{2} < r \leqslant a. \end{cases} \tag{7.16b}$$

These initial conditions were investigated separately by Degance & Johns (1978b). The solution for the source term is derived in appendix B, and the result is

$$\begin{aligned} s^*(z, t) = 4c_0 U &\left[\frac{\sigma_1\alpha_1(z - \Xi_1(t) - \beta_1)}{(\sigma_1^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_1(t) - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t} \right) \right. \\ &\left. - \frac{\sigma_2\alpha_2(z - \Xi_2(t) - \beta_2)}{(\sigma_2^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_2(t) - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t} \right) \right] \\ &\times \sum_{n=1}^{\infty} \frac{J_1(\lambda_n/2)J_3(\lambda_n)}{\lambda_n^2 J_0^2(\lambda_n)} \exp\left(-\frac{\lambda_n^2 \mathcal{D}}{a^2} t \right). \end{aligned} \tag{7.17}$$

Examples of this function are plotted in figure 5. Similar to the developments for the radially linear case, the centre of mass for the left and right components of the initial condition are given by

$$\Xi_1(t) = \begin{cases} \dfrac{7}{4}Ut - \dfrac{1}{2}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\[2ex] \Xi_1(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*, \end{cases} \tag{7.18a}$$

$$\Xi_2(t) = \begin{cases} \dfrac{3}{4}Ut + \dfrac{1}{6}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\[2ex] \Xi_2(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*. \end{cases} \tag{7.18b}$$

Following the material in appendix B as a guide, it should be relatively easy to develop analytical solutions for s^* for other initial conditions of multiplicative form, assuming that the integration defined by (7.8b) can be computed.
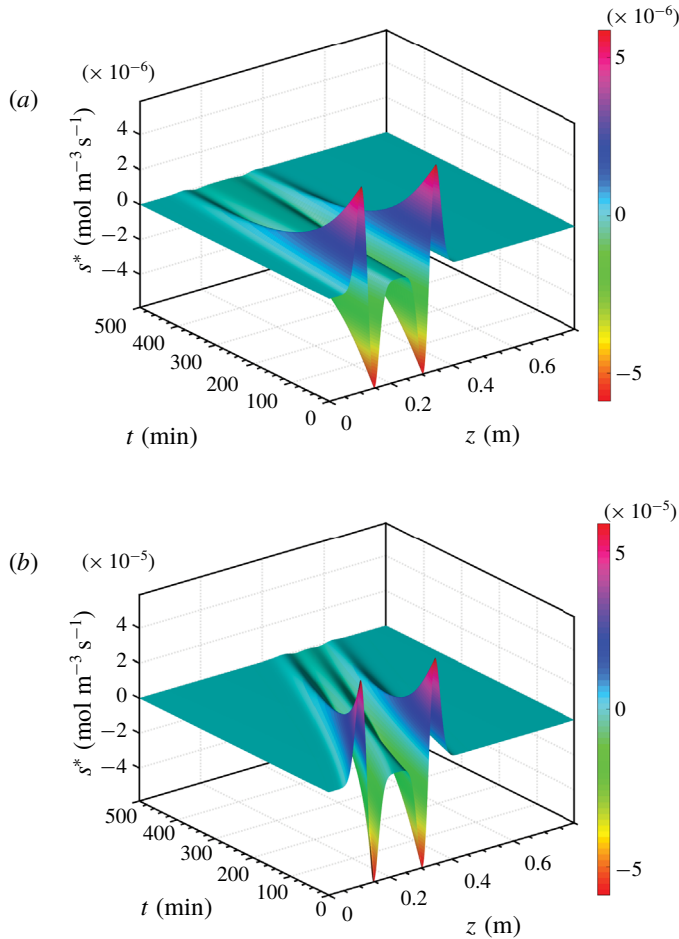
FIGURE 5. Surfaces representing $s^*(z, t)$ for Case C, $Pe = 10$ (*a*) and $Pe = 100$ (*b*). Note that the vertical scales of the left and right figures differ by a factor of 10. As time increases, the source term is exponentially damped. The magnitude of the source term is larger for increasing Pe.

8. Results and discussion

The finite element software COMSOL Multiphysics5.3® was used to solve the microscale partial differential equations (PDEs); we refer to these as NS. Post-processing of data was done primarily in MATLAB R2017®. The physical parameters for the simulations are reported in table 2. The spatial dimension configuration was the two-dimensional axisymmetric case, and the 'transport of diluted species' was selected as the physics package. The software was used to generate the domain and boundary conditions necessary to solve (3.1*a*)–(3.1*e*). An adaptive mesh refinement was performed in order to increase the mesh resolution near the area under higher convection. Backward differentiation formulas (BDF) of order 3 or 4 were used with linear multistep methods in the PARDISO solver. Stability of the solutions was found to be sensitive to the order of the BDF. Both streamline and cross-wind diffusion were employed as two consistent stabilization methods. The

Pe	Parameter	Value	Definition
	L_T	1.60 m	Length of the tube
	a	0.01 m	Radius of the tube
	$M_1(Z_1)$	0.125 m	Centre of mass of initial condition (left portion)
	$M_1(Z_2)$	0.325 m	Centre of mass of initial condition (right portion)
10	Δr	4.5×10^{-4} m	Radial grid dimension (average)
	Δz	4.5×10^{-4} m	Axial grid dimension (average)
	BDF order	4	Backward differentiation formula order (for time)
100	Δr	2.5×10^{-4} m	Radial grid dimension (average)
	Δz	2.5×10^{-4} m	Axial grid dimension (average)
	BDF order	3	Backward differentiation formula order (for time)

TABLE 2. Numerical parameters used in the simulations.

streamline diffusion method recovers the streamline upwind Petrov–Galerkin method and the Galerkin least-squares method. Periodic boundary conditions were applied at the external boundaries perpendicular to the tube axis, although the domain was sufficiently long that no significant mass approached the domain ends over the entire simulation period. A no-flux condition was imposed at the cylinder walls. Computing averages and moments was done by extracting the results on a grid using MATLAB®. A convergence analysis based on Richardson extrapolation was performed following Roache (1994, 2003) in order to quantify the stability of numerical results. The details of the convergence study are summarized in the supplementary material. The grid convergence index, which provides a relative error estimate of the calculated solutions, was found to be of the order of 10^{-5}. The estimated convergence errors were below 2% for all simulations.

8.1. *Microscale concentration fields from numerical solution*

Visualizations of the microscale concentration field at early times provide some direct evidence as to why studying Taylor dispersion at early times has been a challenging process. In figures 6 and 7, visualizations of the concentration field as it evolves from each of the three initial conditions are shown. To compare the results at different Péclet numbers, we have chosen the dimensionless time variable

$$\tau^* = \frac{t\mathcal{D}}{a^2}Pe, \quad \text{or, equivalently,} \quad \tau^* = \frac{Ut}{a}. \tag{8.1}$$

This facilitates the direct comparison of plots that would, otherwise, have dramatically different time scales. The downside is that the actual time scale is somewhat obscured; we provide absolute times parenthetically to help maintain a feeling for the actual time scales involved. Some caution is necessary when interpreting these plots because the aspect ratio has been greatly increased for visualization purposes (i.e. the radial direction is magnified by a factor 10). In addition, to help with the visualization of the concentration at longer times (where the concentrations would be otherwise visually undetectable) the concentration field has been normalized and logarithmically transformed by $c' = \log_{10}(c/c_{A0} + \varepsilon)$, where $\varepsilon = 0.1$, and $c_{A0} = \max[c(r, z, t)]$. For reference, the area-averaged concentrations for each case are presented to the right of the microscale concentration plots.
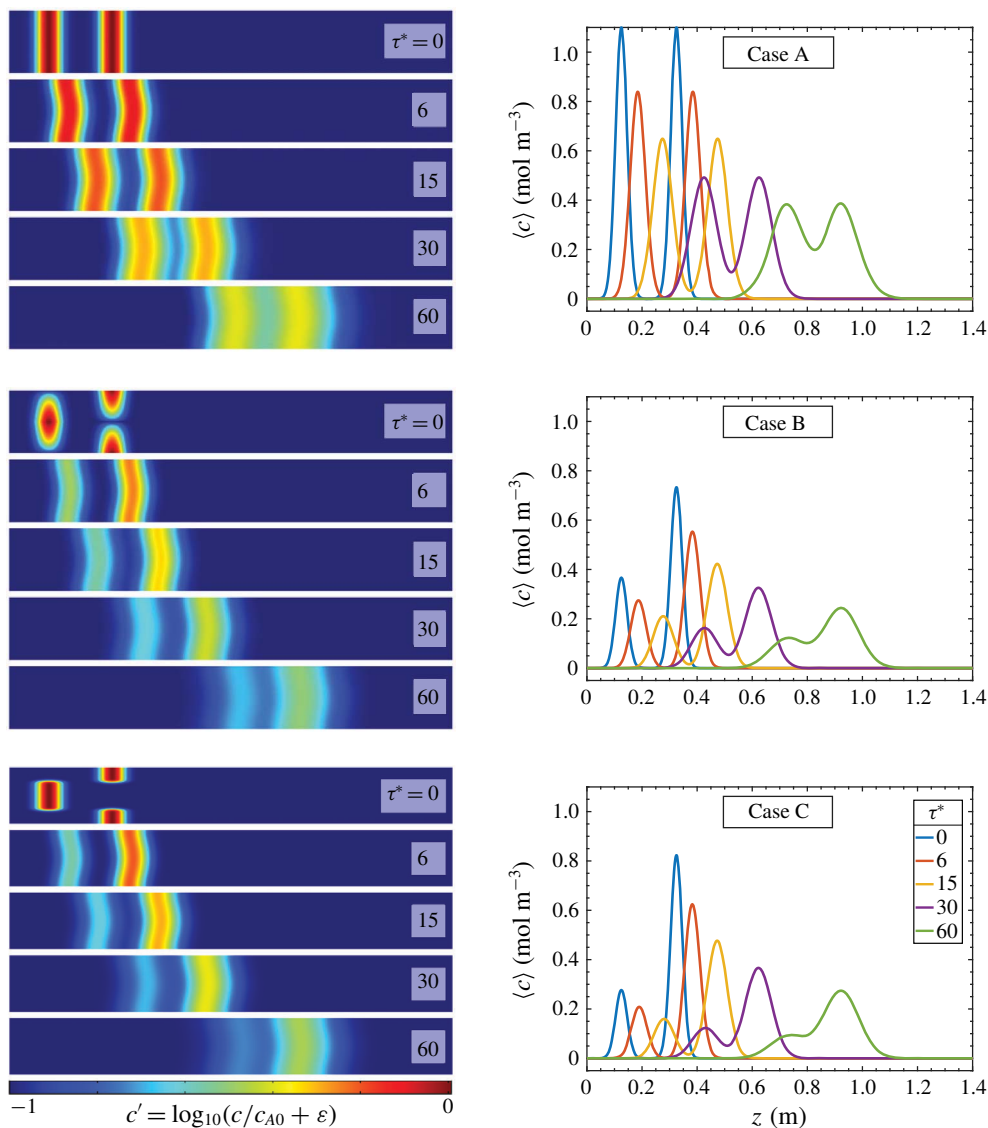
FIGURE 6. Dynamics of the microscale and macroscale concentration profiles for the three cases of initial configuration, $Pe = 10$ (all plots are presented in terms of the dimensionless time variable τ^*). Note that the aspect ratio, a/L, has been increased by a factor of 10 to aid visualization.

From figures 6 and 7, it is clear that the concentration evolution at early times is dominated by the particular initial configuration. Depending upon how the configuration is distributed across the velocity space, dramatically different kinds of behaviours can occur, and this is influenced by the relative importance of diffusion and convection as well as the local shear rate. It is useful to briefly discuss how this is related to the concept of mixing. In short, mixing is stretching-enhanced diffusion (Villermaux 2019). The process of mixing in frequently divided into two conceptual components: (i) convective mixing (Lacey 1954) (or macromixing or
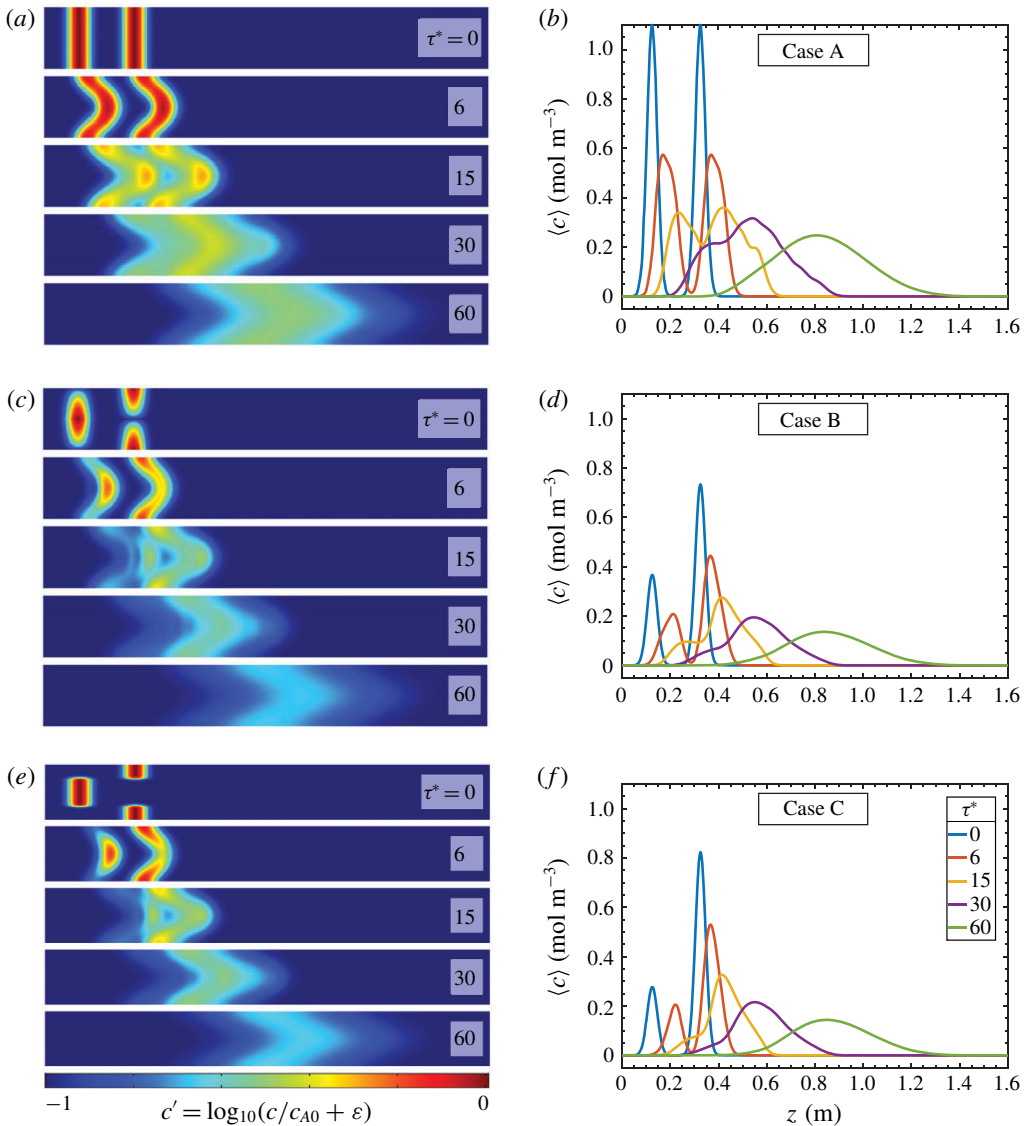
FIGURE 7. Dynamics of the microscale and macroscale concentration profiles for the three cases of initial configuration, $Pe = 100$ (all plots are presented in terms of the dimensionless time variable τ^*). Note that the aspect ratio, a/L, has been increased by a factor of 10 to aid visualization.

stirring) (Villermaux 2019), and (ii) diffusive mixing (or micromixing) (Epstein 1990; Bourne 2003).

From Newton's law of viscosity the shear rate is given by (Bird, Stewart & Lightfoot 2007)

$$\dot{\gamma} = \frac{\partial v_z}{\partial r} = \frac{4Ur}{a^2}, \tag{8.2}$$

where $\dot{\gamma}$ is shear rate. Note that highest convective mixing occurs in places where shear rate is maximum (the region near the wall). Although the convective transport

is highest at the centre of tube, it experiences the lowest shear rate and, consequently, lowest convection-mediated mixing. Radial diffusion conveys solute away from the centreline, creating diffusive mixing. More correctly, convection and diffusion create mixing in concert; shearing of the initial solute distribution ultimately creates a deformed configuration of the solute where diffusion is much more effective. Note that at small times for some of the initial configurations studied here (specifically Cases B and C), convection may actually impose conditions that transiently create less convective mixing before increasing again.

Even for seemingly simple initial conditions (e.g. Case A), we find that the early-time behaviour can be quite complex. For example, for $Pe = 100$ at least three peaks are observable at $\tau^* = 15$ ($t = 250$ min) (figure 7a,b the third peak occurs near $z = 0.6$ m), even though there were only two peaks initially. At the beginning of the process, the concentration was uniform across the cross-section where the initial pulses were placed. Thus as time progresses, the concentration near the centre of the tube experiences more convection, whereas the solute located near the walls is transported mainly by diffusion. For Cases B and C, we find that the two initial peaks converge more rapidly for both Péclet numbers than for Case A; this is because the initial distribution on the left-hand side for these cases is distributed in the highest-velocity portion of the flow field (and the converse is true for the right-hand component of the initial condition). Thus, the centre of mass of the left-hand component catches up with the right-hand component. The effect of this is even more evident when we examine the behaviour of the moments of the solute. For $Pe = 100$, by $\tau^* = 60$ ($t = 1000$ min) we have essentially a single mode concentration field. This is a result, primarily, of the separation distance of the two components of the initial condition. Were they to have been separated by a greater distance, the time required to achieve a single modal concentration distribution in space would be longer.

8.2. *Error analysis*

In addition to issues of convergence, the numerical results also allow making estimates for the amount of error induced by the approximation made for the deviation balance equation. Specifically, we have imposed the following approximations in the analysis to this point:

$$\tilde{v}_z \frac{\partial \tilde{c}}{\partial z} - \left\langle \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \right\rangle \ll \frac{\mathcal{D}}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \tilde{c}}{\partial r} \right) - \frac{\mathcal{D}}{r^2} \frac{\partial^2 \tilde{c}}{\partial \theta^2} - \mathcal{D} \frac{\partial^2 \tilde{c}}{\partial z^2} - \tilde{v}_z \frac{\partial \langle c \rangle}{\partial z} - U \frac{\partial \tilde{c}}{\partial z}, \quad (8.3a)$$

$$a \ll L_0, \quad t^* \ll T^*. \quad (8.3b)$$

Recall that the first of these two constraints was imposed in §5 and led to the constraint $Pe(a/L_0) \ll 1$. Here, we have used the raw restriction (before generating the much simpler constraint) so that we can assess numerically the magnitudes of the terms are being neglected (the left-hand side) versus those retained (the right-hand side). The second of these constraints was associated with approximating the non-local solution by a local one (§6). For the second of these two constraints, we have shown in appendix A that the time constraint is automatically as long as $a/L_0 \ll 1$. It is clear that the length scale constraint itself is reasonably met ($a/L_0 \approx 1/20$), thus we do not pursue the validity of this constraint further.

Therefore, the remaining questions regarding the error associated with the restriction given by the inequality (8.3a). This is best measured by the error observed between the upscaled model and the direct microscale numerical results. We can compute each

of the terms in the inequalities in (8.3*a*)–(8.3*b*); however, the dimension of the result is as large as the numerical solution itself. To generate a reasonable summary metric, we propose to compute the following error (noting that $\widetilde{\tilde{v}_z\tilde{c}} = \tilde{v}_z\tilde{c} - \langle\tilde{v}_z\tilde{c}\rangle$ is the spatial deviation of the quantity $\tilde{v}_z\tilde{c}$).

$$\epsilon_1(z, t) = \left(\frac{\langle|\widetilde{\tilde{v}_z\tilde{c}}|\rangle}{\left\langle |U\tilde{c}| + \left|\mathcal{D}\dfrac{\partial\tilde{c}}{\partial z}\right| + \left|\mathcal{D}\dfrac{\partial\tilde{c}}{\partial r}\right| + |\tilde{v}_z\langle c\rangle| \right\rangle + \epsilon_0} \times 100 \right). \tag{8.4}$$

Here, ϵ_0 is a small number used only to assure that division by zero errors do not occur. A few plots of the metric $\epsilon_1(z, t)$ appear in the supplementary materials. This metric is still of high dimension, so to further reduce dimension (and generate a summary metric) we compute the first moment of $\epsilon_1(z, t)$ about the z-axis as follows:

$$\epsilon_{closure}(t) = \frac{1}{A_\epsilon(t)} \int_{z=-\infty}^{z=+\infty} \frac{1}{2}\epsilon_1^2(z, t)\,\mathrm{d}z, \tag{8.5a}$$

$$A_\epsilon(t) = \int_{z=-\infty}^{z=+\infty} \epsilon_1(z, t)\,\mathrm{d}z. \tag{8.5b}$$

This result measures the average error arising from the approximations imposed by the constraints (8.3*a*) and (8.3*b*). For interested readers, the detailed results of this analysis appear in figure 2 of the supplementary materials. A few general conclusions can be made regarding the error arising from this approximation are summarized as follows:

(i) The error of the approximation (8.3*a*) tends to increase as *Pe* increases.
(ii) The error of the approximation decreases (not necessarily monotonically) with increasing time.
(iii) The observed error is influenced by the particular initial condition examined.
(iv) The maximum value of $\epsilon_{closure}$ for the $Pe = 10$ case was less than 10 %, and the error rapidly decreased (within $\tau^* < 5$) to less than 5 %.
(v) The maximum value of $\epsilon_{closure}$ was of the order of 10 %–12 %; this was observed for $Pe = 100$.

Overall, these results seem encouraging and reasonable. The comparison of the averaged result with the numerical results (described below) provide further evidence that the approximation is valid for the Péclet numbers considered here. For larger Péclet numbers, the error associated with the approximation given by (8.3*a*) tends to increase. For such cases, it may become necessary to solve the closure problems associated with *b* and Φ numerically without neglecting terms (an example of the numerical computation of s^* is discussed in § 9 and in appendix A § A.6).

On the basis of the numerical results and the discussion above, it is possible to specify more useful constraints to increase the range of validity of the analysis. We suggest slightly more liberal constraints of the form

$$\varpi Pe \frac{a}{L_0} = \boldsymbol{O}(1), \tag{8.6a}$$

$$\frac{a}{L_0} < 1, \tag{8.6b}$$

where ϖ is a unitless parameter; for our analysis $\varpi \approx 1$. These constraints should generate results that are at least as accurate as those described in the present paper. We note that (8.6a) is consistent with Mercer & Roberts (1994, equation (2.17)). Our constraint is identical to theirs when the left-hand side of (8.6a) is multiplied by a factor of $\varpi = 1/2.2$, lending further support of these slightly more liberal constraints in practice. We note that in the case of Taylor (1954), the value of $\varpi = 1/4$ was adopted; when fixing the value $a/L_0 \approx 1/10$ as approximated by Taylor (1954), equations (8.6a)–(8.6b) are also consistent with that work.

8.3. *Computation using the upscaled balance equation*

The upscaled problem given by (6.6a)–(6.6d) was solved as a one-dimensional model for the three initial condition cases. Note that the dispersion coefficient is independent of the initial condition, thus it remains identical for each case. Simulations were conducted using COMSOL in a manner consistent with what has been described in § 8.1 for the microscale simulations. The analytical solutions for D^* and s^* were computed using a custom-built code constructed in MATLAB; results were then imported into COMSOL, and interpolated to the grid using built-in routines. Simulations were assured to be convergent (using the approach described for the microscale simulations), and the average grid size was 1×10^{-4} m. Other solver parameters are identical to those reported in § 8.1 and table 2.

In figure 8, the averaged concentrations computed from (i) the numerical solution, and (ii) the upscaled, one-dimensional equation are compared. For $Pe = 10$, the average concentration computed from the numerical solution is nearly indistinguishable from those predicted by the upscaled one-dimensional theory. We defined the following error metric for the averaged spatial concentration at a specified time, t_p by

$$\epsilon_{model}(t_p, z) = \frac{|\langle c \rangle_{NS} - \langle c \rangle|}{\max(\langle c \rangle_{NS})} \times 100, \tag{8.7}$$

where the maximum in the denominator is taken as the maximum average concentration observed at time t_p. We computed the error from pairs of points separated by the minimum distance between the curves (as is done in computing total least squares) so that the error was not skewed by small displacement discrepancies. For the $Pe = 10$ cases, the maximum error observed was less than 1.5 % for all times and all cases. For the $Pe = 100$ cases, the maximum error was less than 7 % for all times and cases. For reference, plots of the errors are available in the supplementary materials. The correspondence between the two methods for computing the averaged concentration provide validation that the approximations given by (8.3a) and (8.3b) are reasonable for these Péclet numbers.

8.4. *Moment analysis*

The spatial moments of the concentration field are frequently used to characterize the behaviour of the concentration field, both because of their appearance in the underlying theory, and because they provide information about the characteristics of the concentration distribution. The spatial moments in the z-direction of order n ($n = 0, 1, 2, 3, \ldots$), M_n, are defined by

$$M_0 = \int_{z=-\infty}^{z=\infty} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} c(r, \theta, z, t) r \, dr \, d\theta \, dz, \tag{8.8a}$$
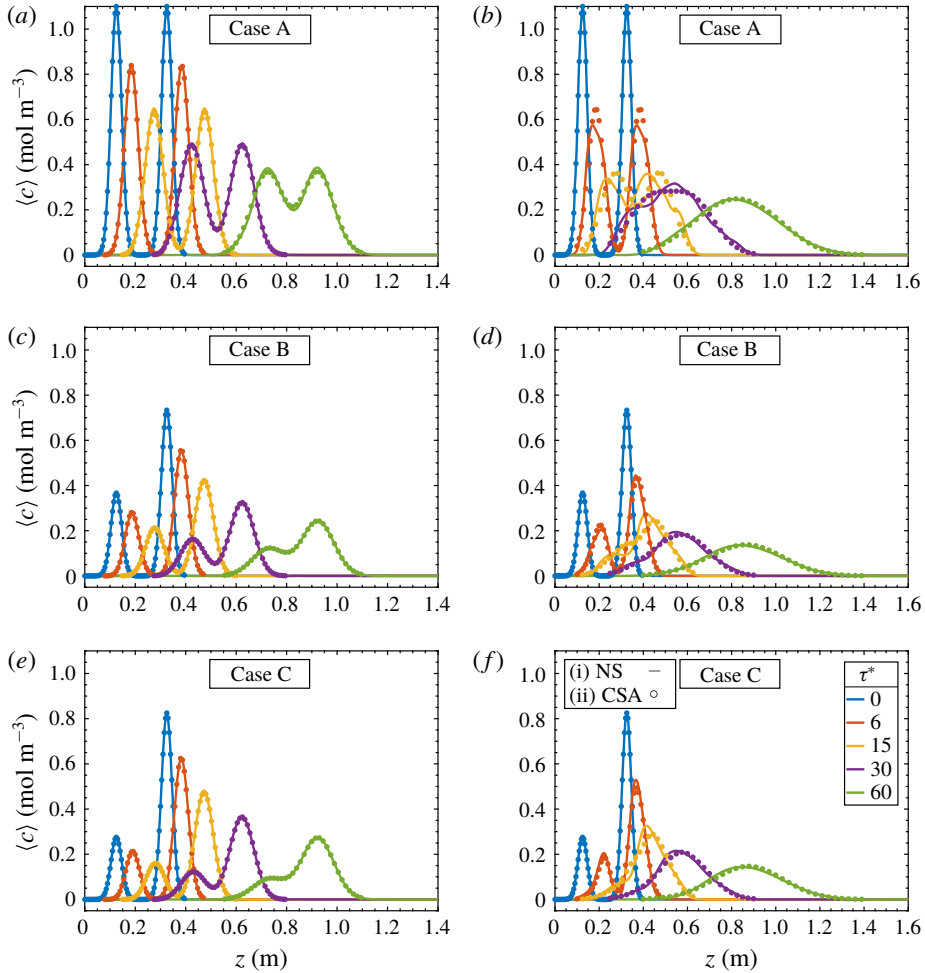
FIGURE 8. Comparisons of the predictions of the average concentration. Profiles were computed by (i) averaging the microscale numerical simulations computed from (3.1a)–(3.1e) (denoted by NS), and (ii) numerically solving the averaged equation with the effective parameters D^* and s^* (equations (6.6a)–(6.7c)) (denoted by CSA). Results are for $Pe = 10$ (a,c,e) and $Pe = 100$ (b,d,f). The five fixed times are expressed in the dimensionless time variable τ^*.

$$M_n(t) = \int_{z=-\infty}^{z=\infty} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} z^n c(r, \theta, z, t) r \, dr \, d\theta \, dz. \tag{8.8b}$$

Often, the centred spatial moments of order $n(n = 0, 1, 2, 3, \ldots)$, μ_n are adopted for characterization. These are given by

$$\mu_0 = M_0, \tag{8.9a}$$

$$\mu_n(t) = \frac{1}{\mu_0} \int_{z=-\infty}^{z=\infty} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} (z - M_1(t))^n c(r, \theta, z, t) r \, dr \, d\theta \, dz. \tag{8.9b}$$

For this work, we are primarily interested in the second and third moments; specifically, we track the centred second moment, which can be expressed by

$$\mu_2(t) = \sigma^2(t) = M_2(t) - M_1^2(t) \tag{8.10}$$

and the skewness, which is a normalized centred third moment, specified by

$$\gamma_1(t) = \frac{\mu_3(t)}{\mu_2^{3/2}(t)}. \tag{8.11}$$

In figures 9 and 10 we plot the second centred moment and the skewness as a function of time for each of the three initial conditions. One interesting feature that can be observed in these data is the distinct differences between the apparent time scales for the second centred moment and the skewness to attain their near-asymptotic behaviour. For the second moment, the asymptotic behaviour is represented by a linear increase in the moment with increasing τ^*; for the skewness, the asymptotic behaviour is the approach to a value of zero (cf. Chatwin 1970). Examining the $Pe = 10$ plots in figure 9, the slope of the second moment ($d\mu_2/dt$) reaches its asymptotic value for τ^* near 2. The skewness, however, continues to show substantial evolution through the entire period plotted (up to $\tau^* = 60$). Similar kinds of behaviour can be observed for the $Pe = 100$ data; the second moment is near asymptotic after approximately $\tau^* = 20$, whereas the skewness continues to evolve for the entire time period plotted. We have plotted the skew for both Péclet numbers on a log–log scale for a time interval up to $\tau^* = 300$ in figure 11; the long-time dynamics of the skew is more easily seen in that figure.

For $Pe = 100$, each of the three initial condition cases more obviously shows substantial transience in the skew. For each of the cases, the skewness starts at a negative value, and then crosses the zero to become positive. The approach to the asymptotic value then occurs slowly from the condition of positive skewness. Overall, these results indicate that skewness generally increases at early times, and this increase may ultimately lead to non-monotonic behaviour as the skewness approaches the asymptotic value. This is consistent with Chatwin (1970), who also found non-monotonic behaviour in third moment terms. Similar results can be seen in the $Pe = 10$ cases (see the inset in figure 11a), although the time scales simulated do not allow a full analysis of the dynamics of the skewness at this value for Pe. Because our simulation times do not extend far enough, we cannot accurately assess the asymptotic behaviour of the skewness (such as those reported by Aris (1956) and Chatwin (1970)).

These results indicate that the approach to normality should not necessarily be measured by, for example, the approach of only the second moment to its asymptotic behaviour (i.e. achieving constant slope). While this is a useful metric, higher-order measures such as the skew may continue to show substantial evolution over a longer characteristic time scale indicating that the system as a whole is not necessarily near its asymptotic state of behaviour.

8.5. *Derivative of the second moment*

In many theories of dispersion, the time derivative of the second moment is used to generate a *de facto* definition of the dispersion coefficient; usually, this is expressed in one dimension by

$$D^{\mu_2} = \frac{1}{2} \frac{d\mu_2}{dt}. \tag{8.12}$$

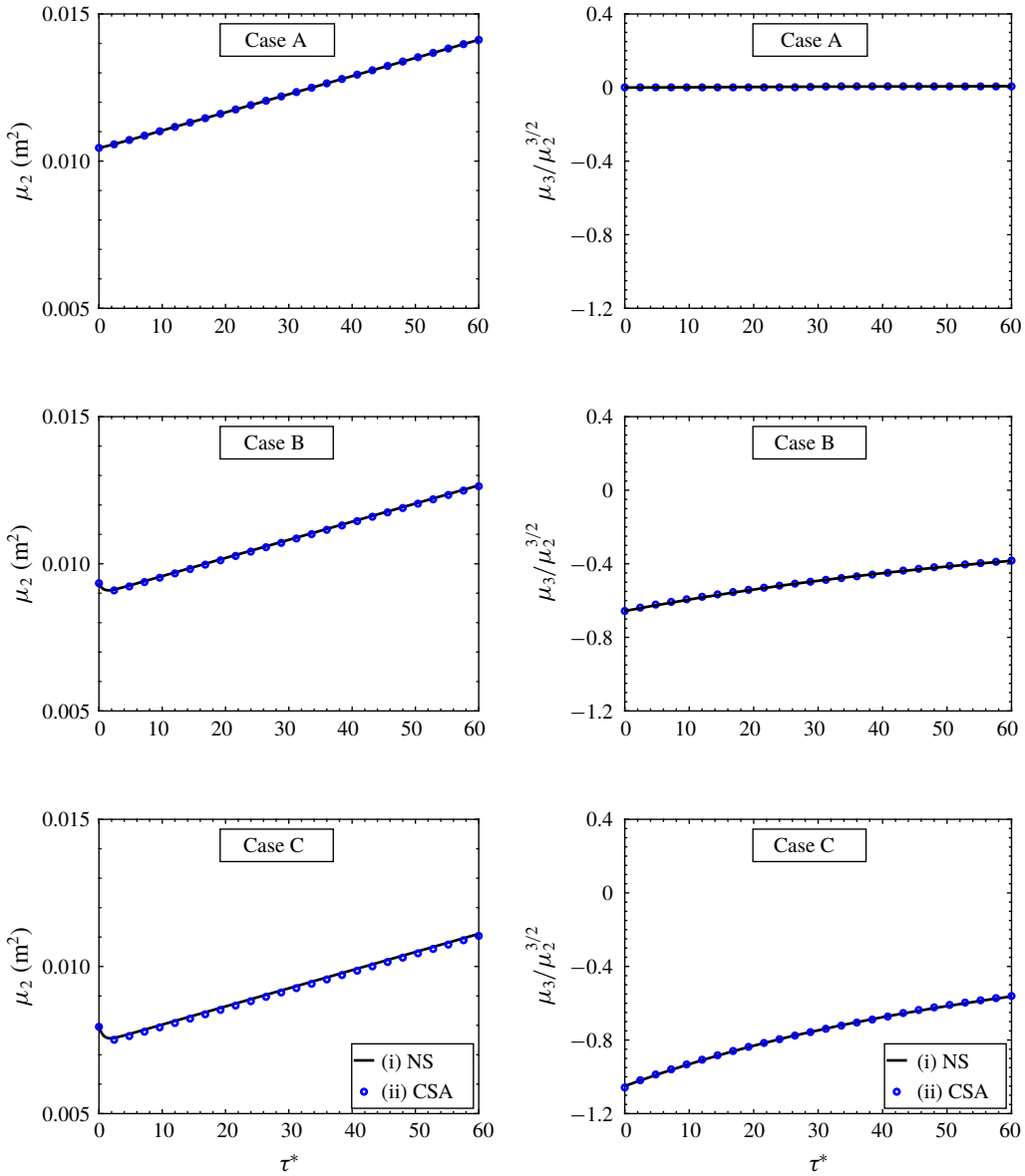FIGURE 9. Comparisons of the moments for different cases of the initial configuration for $Pe = 10$. Moments were computed from two concentration fields as follows: (i) the microscale numerical simulations computed from (3.1a)–(3.1e) (denoted by NS), and (ii) the numerical solution of the averaged equation with the effective parameters D^* and s^* (equations (6.6a)–(6.7c)) (denoted by CSA).

Here, we have used the notation D^{μ_2} to distinguish this value from D^*. While such definitions are true asymptotically, they are not necessarily true for times that are near to the initial configuration time. One of the desirable traits for the dispersion coefficient identified in the Introduction is that it should be positive. This prevents conflicts with macroscale thermodynamics, and also assures that the resulting balance
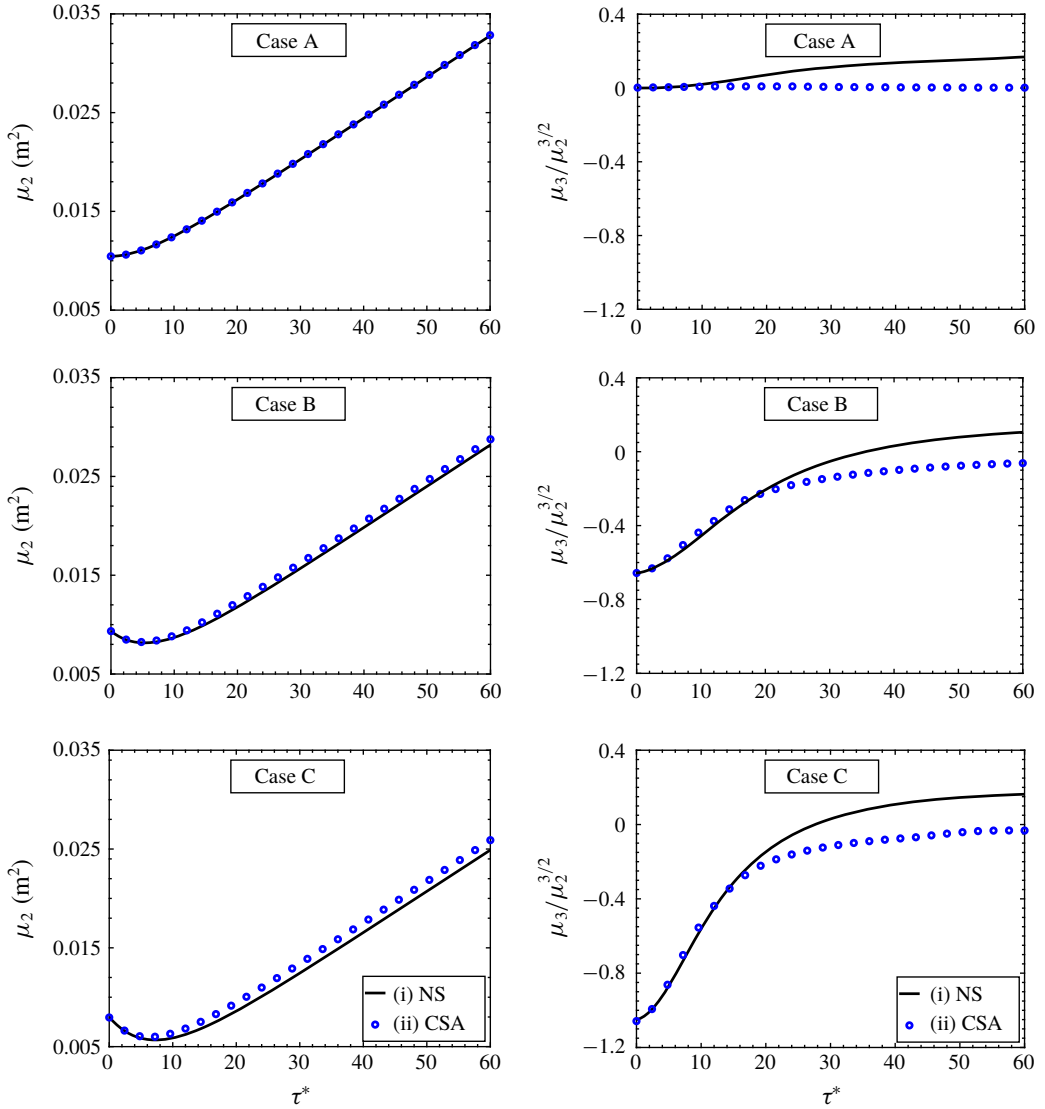
FIGURE 10. Comparisons of the moments for different cases of the initial configuration for $Pe = 100$. Moments were computed from two concentration fields as follows: (i) the microscale numerical simulations computed from $(3.1a)$–$(3.1e)$ (denoted by NS), and (ii) the numerical solution of the averaged equation with the effective parameters D^* and s^* (equations $(6.6a)$–$(6.7c)$) (denoted by CSA).

equations are well posed. However, there are initial configurations for which the second moment decreases in time before increasing again. This can be seen in the curves for μ_2 (Cases B and C) for both $Pe = 10$ and $Pe = 100$ (figures 9 and 10). If one were to use the definition above then, for early times, a negative value for the dispersion coefficient would be produced. As an example, we have computed this quantity for $Pe = 100$ and plotted it in figure 12. Note that for Case A, the definition for D^{μ_2} is identical to that defined for D^*; this is because s^* is identically zero for that
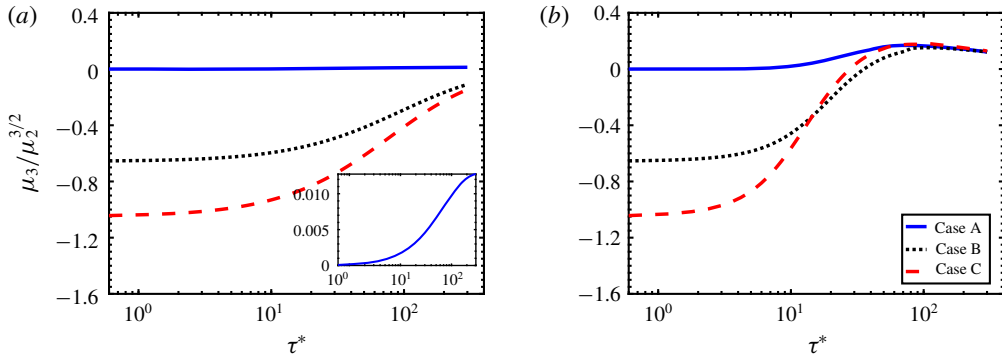
FIGURE 11. Dynamics of the skewness for large values of τ^*. These results were computed directly from numerical simulations. The results represent Cases A, B and C for long-time evolution illustrating near-asymptotic behaviour not observable in figures 9 and 10. (*a*) $Pe = 10$; (*b*) $Pe = 100$. The approach to the asymptotic state is non-monotonic; from these solutions it is not possible to tell much more about the functional form of the approach.
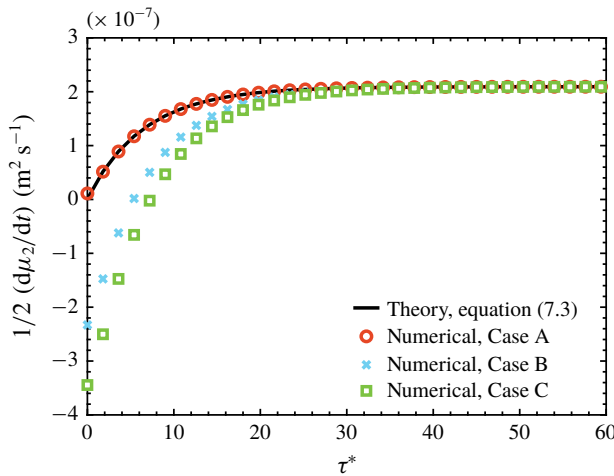

FIGURE 12. Dynamics of the time derivative of the second moment for $Pe = 100$. These results are obtained from the moments of the microscale numerical simulations computed from (3.1*a*)–(3.1*e*). The analytical solution produced by the upscaled theory is shown for comparison. Adopting the conventional definition of dispersion as $1/2 d\mu_2/dt$ is not valid at early times for some initial conditions. In general, the definition based on $1/2 d\mu_2/dt$ is only valid at asymptotic times. In contrast, the definition proposed using the upscaled theory presented is valid at all times and for all initial conditions.

case. However, for the other two cases, we have an initial configuration that transiently reduces the second moment. This occurs because the initial configuration contains two parts: the first part (on the left) contains mass focused in the high-velocity regions, and is upstream from the second part. The second part (on the right) contains mass that is focused near the walls. The net result is that the centre of mass of the left part of the initial distribution moves faster than the right; thus, the left part

eventually catches up to the right. This manifests physically as a decrease in the second moment. Thus the definition given by D^{μ_2} yields non-physical results for those cases because the effective dispersion coefficient is negative at some times; in contrast, $D^*(t)$ presented in (6.6a)–(6.7c) is strictly positive for all times.

These results provide some additional insight into the physical role of the s^* field. Essentially, this field is a source term that assures that the spreading that is encoded by the initial configuration is correctly represented at early times. Because this is independent from the effective dispersion coefficient, there is no need to posit dispersion coefficients that, for example, violate macroscale thermodynamics by attaining negative values.

9. Examples addressing the superposition problem

Taylor himself was not necessarily a proponent of dispersion coefficients that were expressed as functions of time. In particular, he was concerned about possible paradoxes that could occur when pulses were injected into a system at two different times (with the question being, 'which dispersion coefficient applies?'). In his 1959 paper (Taylor 1959) he stated the following about time-dependent dispersion coefficients (using the term diffusion instead of dispersion as is now more common).

> It seems to me that this is an illogical conception. The one thing that seems to be agreed, whatever theory one may have about diffusion, is that diffusing distributions are superposable. If therefore you attempt to analyse the distribution of concentration from two sources which started at different times by this method, it would be necessary to assume, in places where the distributions overlapped, that the diffusion constant had two different values at the same time and at the same point in space.

Because there are many senses in which the concept of superposition may be applied, we offer the following definition for the superposition of non-homogeneous partial differential equations (after Olver (2014, appendix B)):

THEOREM. *If u_1 and u_2 are particular solutions to the non-homogenous linear equation $L[u] = f$, then $u = u_1 + u_2$ solves $L[u] = f_1 + f_2$.*

Implicit in this definition is the idea the linear operators involved must be the same over any time and space intervals considered. Because the problem we are considering involves a parameter, $D^*(t)$, that is a function of time, then the principle of superposition specified above requires that the two solutions, u_1 and u_2 be defined over the same time interval. This is one of the difficulties that has been encountered for the slightly more restrictive conditions that Taylor (1959) required. For the situation outlined by Taylor (1959) there was a stated desideratum (on the basis of physical reasoning) that two solutions defined over different time intervals would be superposable.

While formulations that do not account for a source term, s^*, lose the ability to be sensibly superposed, this is not true for the theory that we have presented. In fact, this is one of the strengths of the proposed theory: superposition is maintained (cf. Mercer & Roberts 1994). In short, the source term, s^*, and the effective dispersion coefficient, D^*, work together in a way as to maintain the principle of superposition. The clearest way to see this is through concrete examples. We present two examples below (with $Pe = 100$) where we compute both the microscale and the corresponding macroscale

solution. Although ordinarily one would not compute the microscale solution, doing so allows us to compare the average computed directly from the microscale solution with the average computed from the upscaled balance equations. It also provides an opportunity to better understand the complicated microscale dynamics involved in the transport process, which provides useful context for interpreting the upscaled results.

We illustrate superposition using two examples, as follows.

(i) For the first example, we provide a problem that makes the purpose of the source term, s^*, more apparent. In this example, we consider a single injection, but break the total time ($0 < t < T_f$) up into two time intervals. During the first interval ($S_1 : 0 < t < T_1$), the initial condition evolves from a uniform slug input to a complex distribution of space (with non-zero radial gradients). The new configuration at time $t = T_1$ is then treated as the initial condition for the second time interval ($S_2 : 0 < t' < T_\Delta$). For the second interval, the time variable is reset such that the initial time for the problem is equal to $t' = 0$. The problem at this juncture is simply a new initial value problem with a complicated initial condition. Importantly, for this second interval, the dispersion coefficient $D^*(t')$ evolves exactly as indicated by (7.3). In other words, the dispersion tensor starts at a value of $D^*(t' = 0) = \mathcal{D}$, and grows in time (in variable t') according to (7.3). Although for this second time interval the value of the dispersion coefficient is reset to its zero-time value, the second moment of the solute continues is shown to grow at the rate that was growing at the end of the first time interval. This illustrates how the memory of the system is encoded in our solution by the source function, s^*, rather than in the dispersion coefficient.

(ii) For the second example, we build off of the discussion of superposability introduced in the first example. We consider specifically the 'two release' case identified by Taylor (1959), where solute is injected first at $t = 0$ (which we refer to as I_1), and the system is allowed to evolve up to the time $t = T_1$. At time $t = T_1$, a second solute injection occurs (I_2), and the system again is allowed to evolve. For this problem, we show that if separate transport equations are written for each release (adopting the notation $\langle c_1 \rangle$ and $\langle c_2 \rangle$ for the first and second injections, respectively), that these two equations can be superposed to derive a single transport equation for $\langle c \rangle = \langle c_1 \rangle + \langle c_2 \rangle$. Notably, for the second time interval, the function D^* is single valued everywhere in space, including locations where the two solute injections overlap. In other words, even though the residence times for the two solute injections are not equal, they are described by a single upscaled dispersion coefficient. As with the first example, this example shows in detail how the history of the solutes are encoded by s^* rather than by the dispersion coefficient, D^*. This example directly addresses the objection raised by Taylor (1959); in addressing this question, we show that our upscaled transport equation is superposable in the sense defined by Taylor (1959), and that the disturbing problem where the dispersion coefficient appears to have multiple values at a single point does not occur.

9.1. *Macroscale dispersion: example 1*

In order to isolate the effect of the s^* field on the overall solution to the macroscale equations, we provide an example where its influence can be more readily understood. In this example, we illustrate that the conventional notion of superposition is still valid for the upscaled theory outlined previously.

For this example, a single solute injection is simulated as a sequence of two initial value problems. It is always possible to break initial value problems into parts like this when superposition is valid. The solution at the end of any one time interval is

simply a concentration field; this concentration field can then be treated as the initial condition that is transported over a second time interval via the same balance equation.

For this example, we examine the transport of a single uniform pulse. We allow the initial condition to evolve until $t = T_1 = 250$ min. The solution at this time, $\langle c \rangle_{(t=T_1,z)}$ is used to define the initial condition for the second time interval, S_2, at starting at $t' = 0$ min, and ending at $t' = T_\Delta = 750$ min. The problem can be stated by the following sequence of two problems. The two problems can be written as

Time interval 1 *solution* $\quad S_1 : 0 < t < 250$ min

$$\frac{\partial \langle c \rangle}{\partial t} = D^*(t) \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} + s_1{}^*(z, t), \quad (9.1a)$$

$$\text{I.C.1} \quad \langle c \rangle|_{(z,t=0)} = c_0 Z_1(z). \quad (9.1b)$$

Time interval 2 *solution* $\quad S_2 : 0 < t' < 750$ min

$$\frac{\partial \langle c \rangle}{\partial t'} = D^*(t') \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} + s_2{}^*(z, t'), \quad (9.1c)$$

$$\text{I.C.2} \quad \langle c \rangle|_{(z,t'=0)} = \langle c \rangle|_{(z,t=T_1)}. \quad (9.1d)$$

Note that the second interval is written in terms of a new time variable, $t' = t - T_1$ such that $0 < t' < T_\Delta$. After the solution $\langle c \rangle|_{(z,t')}$ is computed, it is straightforward to translate this solution to the original time variable via $t = t' + T_1$.

The notable feature of this solution is that the dispersion coefficient has exactly the same dynamics for each of the two time intervals. To be explicit, recall $T_1 = 250$ min, and $T_f = 1000$ min; the dispersion coefficients for the two intervals are given by

Interval 1 $\quad (0 < t < 250$ min$)$

$$D^*(t) = \mathcal{D} \left(1 + \frac{1}{48} \frac{U^2 a^2}{\mathcal{D}^2} \right) - 4 \frac{U^2 a^2}{\mathcal{D}} \sum_{n=1}^{\infty} \left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)} \right)^2 \exp \left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t \right). \quad (9.2a)$$

Interval 2 $\quad (0 < t' < 750$ min$)$

$$D^*(t') = \mathcal{D} \left(1 + \frac{1}{48} \frac{U^2 a^2}{\mathcal{D}^2} \right) - 4 \frac{U^2 a^2}{\mathcal{D}} \sum_{n=1}^{\infty} \left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)} \right)^2 \exp \left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t' \right). \quad (9.2b)$$

This means that $D^*(t = 0) = \mathcal{D}$ at the start of the first time interval, and that $D^*(t' = 0) = \mathcal{D}$ at the start of the second time interval. For both intervals, the dispersion coefficient increases from its initial value with identical dependence upon time according to (9.2a) and (9.2b). In other words, there is no 'memory' for the dispersion coefficient.

We note that for the second time interval, the initial concentration field is not uniform in the radial direction; this means that the s^* field is non-zero. This initial condition is also quite complex (i.e. it involves computing the s^*-field for the initial condition $\langle c \rangle|_{(z,t'=0)} = c_0 Z_1(z) + \langle c_1 \rangle|_{(z,t=T_1)}$), so there is no analytical solution for the s^* field. Thus, s^* is computed numerically, as described in § A.6. The s^* field for the superposed initial condition at time $t' = 0$ min is given in figure 13.

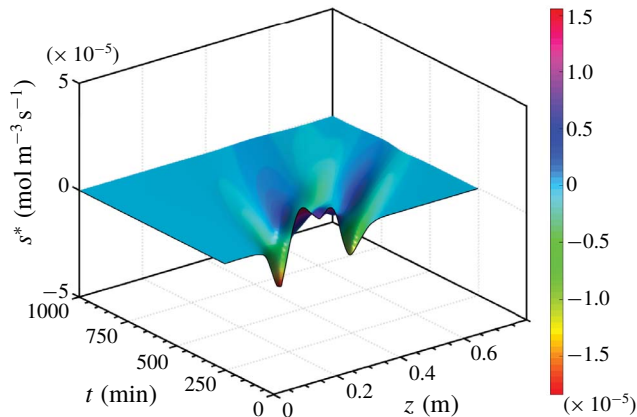This example allows for a unique opportunity to directly assess the effect of the s^*

FIGURE 13. Spatial and temporal dependence of the source term, s^* for $Pe = 100$ at $t' = 0$ min.

field. To do so, we consider the following three cases for obtaining a solution:

(i) The solution to the problem is determined by a single computation, spanning the interval $0 \leqslant t \leqslant 1000$ min.

(ii) The solution is obtained for two time intervals, where the time intervals are given by $S_1 : 0 \leqslant t \leqslant 750$ min and $S_2 : 0 \leqslant t' \leqslant 750$ min. The s^* field is computed as specified previously.

(iii) Strictly for comparison, the solution is obtained as the superposition of the two time intervals described above. The s^* field, however, is set (incorrectly) to zero so that the solutions both with and without the s^* field can be compared. This corresponds to the form of the time-dependent dispersion equation that is used frequently (cf. Sankarasubramanian & Gill (1973, equation (11))), and it exhibits the lack of time translation symmetry pointed out by Taylor (1959).

The second moment is an effective measure for examining the result of these three cases. In figure 14 we have plotted the results of the solutions as computed for each of the three cases listed. In these results, we can consider the results from the first case (the black line in figure 14) to be the ground truth for comparison.

We note that, for the solution computed in two steps with the correct value for s^*, the second moment is continuous, and it matches the single-step solution almost identically. This occurs even though the dispersion coefficient returns to its zero-time value when the problem is restarted at time $t = 250$ min. This behaviour occurs because of the unique interplay between the effective dispersion coefficient, D^*, and the source term given by s^*. Thus, the concern posed by Taylor (1959) about the paradox of which dispersion coefficient to use (the one at the end of the first time interval, or the one corresponding to zero elapsed time for the second time interval) does not arise. The dispersion coefficient is unequivocally defined for both time periods by (9.2a)–(9.2b), and superposition of the solutions for the two time intervals leads to a solution that contains no discontinuities in the slope of the second moment.

For Case 3, we note that without the proper s^* field correction, the second moment develops a cusp at $t = 250$ min. In other words, the rate of spreading for the second time interval is too small, and this occurs because the source term s^* has been neglected. This creates a technical problem, because the derivative of the second
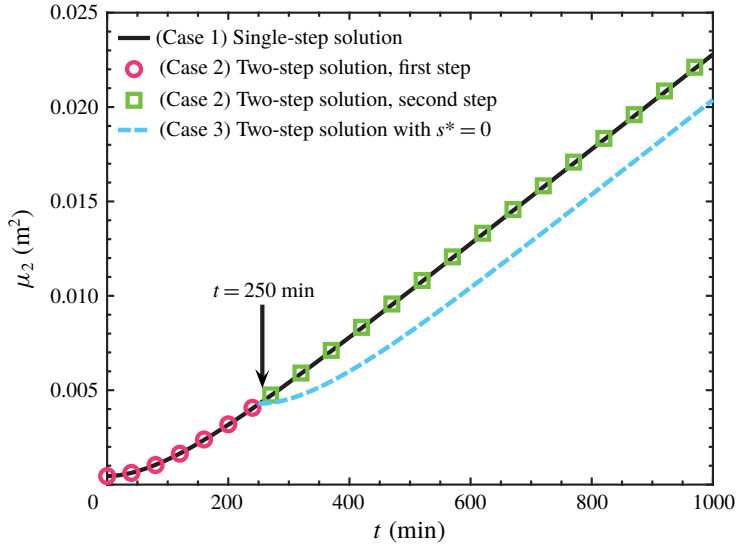
FIGURE 14. Two-step process for $Pe = 100$. (i) Case 1. The solution solved over the whole (1000 min) time period. (ii) Case 2 (first step). The solution, S_1, for $0 < t < 250$ min. (iii) Case 2 (second step) The solution, S_2 for $0 < t' < 750$ min ($250 < t < 1000$ min) with the source term computed numerically, and illustrated in figure 13. (iv) Case 3. The solution, S_2 for $0 < t' < 750$ min with $s^* = 0$.

moment does not exist at that point. Additionally, the second moment then grows too slowly for the remaining time, under-predicting the actual value significantly.

This case not only allows a comparison to better illustrate how the s^* field influences the solution to account for the initial distribution of matter, but it helps to resolve the apparent paradox that troubled Taylor (1959). Specifically, for dispersion in tubes, the dispersion process is not independent of the initial condition at early times; however, this dependence is best represented through s^* rather than D^*. For initial conditions that are uniform in the radial direction, no additional modification from conventional theory is needed. However, for initial conditions that are not uniform in the radial direction, the initial condition itself generates a behaviour that is accounted for by the appearance of the non-conventional source term s^* in the macroscale transport equation. The inclusion of the additional source term is essential for predicting the correct macroscale dynamics of the system.

9.2. Macroscale dispersion: example 2

In this example, we illustrate through direct deviation and computation that our solution is superposable in the sense that was desired by Taylor (1959). For this problem, the system starts with a specified initial condition and progresses for some time ($0 < t < T_1$). Then, a second source is instantaneously injected into the system, and transport continues over the interval $T_1 < t < T_f$. This situation, then, yields a condition where two solutes are in the system for varying amounts of time (and, hence, would experience different time-dependent effective diffusion coefficients). Taylor's 1959 concern was the logical incongruity associated with assigning two different values of the dispersion coefficient at the same point where the solute bodies overlap. In the following, we show that the framework developed above avoids this problem.

9.2.1. *The superposition problem for a two release case*

This example is much like example 1, except now we consider the case where there are two injections of solute at two different times. This represents exactly the conditions that which were addressed by Taylor (1959). In other words, the case of two releases leads to the problem where two solute distributions, each having been in the system for different amounts of time, might overlap. The conventional approach (e.g. Sankarasubramanian & Gill (1973, equation (11))) using a time-dependent dispersion coefficient (without a source term for correction) would lead to a problem where two different values of the dispersion coefficient would occur at points where the solute distributions overlapped. In this example, we illustrate that the balance equation for the average concentration that we derive can be sensibly superposed.

To start the discussion, we first need to determine mathematically how to handle a second pulse injected at some time $t = T_1$ after progression from a defined initial condition at $t = 0$. The sudden injection of solute into the system at $t = T_1$ introduces a discontinuity of the solute field in time. Although there are several ways that this might be represented, the simplest one is to use superposition in the following way. For the first time interval $(0 < t < T_1)$, the system evolves from the specified initial condition (for these examples, we will use the function $Z(z)$ defined previously in (7.11a), such that $\langle c \rangle|_{(z,t=0)} = c_0 Z_1(z)$). At time $t = T_1$, a new solute source is instantaneously injected into the system. The new initial condition for the second time interval is then the superposition of the second pulse configuration with the configuration of the existing solute distribution field, $\langle c \rangle|_{(z,T_1)}$. This new field forms an initial condition for the second time interval, beginning at time $t = T_1$ and ending at time $t = T_f$. In the comments by Taylor (1959), this is ostensibly what is meant by the use of the word 'superposition'. Thus, we consider two problems that together provide the necessary solution.

To make this concrete, again let $T_1 = 250$ min and $T_f = 1000$ min. The two problems can be written as follows. We define the associated independent transport equations for the solute injections (indicated by I_1 and I_2) for the entire time interval $(0 < t < 1000$ min$)$ to be described by

Problem $\langle c_1 \rangle$ $I_1 : 0 < t < 1000$ min

$$\frac{\partial \langle c_1 \rangle}{\partial t} = D^*(t) \frac{\partial^2 \langle c_1 \rangle}{\partial z^2} - U \frac{\partial \langle c_1 \rangle}{\partial z} + s_1^*(z, t), \qquad (9.3a)$$

$$\text{I.C.1} \quad \langle c_1 \rangle|_{(z,t=0)} = c_0 Z_1(z). \qquad (9.3b)$$

Problem $\langle c_2 \rangle$ $I_2 : 0 < t' < 750$ min

$$\frac{\partial \langle c_2 \rangle}{\partial t'} = D^*(t') \frac{\partial^2 \langle c_2 \rangle}{\partial z^2} - U \frac{\partial \langle c_2 \rangle}{\partial z} + s_2^*(z, t'), \qquad (9.3c)$$

$$\text{I.C.2} \quad \langle c_2 \rangle|_{(z,t'=0)} = c_0 Z_1(z). \qquad (9.3d)$$

For the second injection, the time coordinate for the second interval is started at the time $t' = t - T_1$, as described in example 1. Without this translation, the second injection would not start with the correct dispersion coefficient, which is $D^*(t' = 0) = \mathcal{D}$, as indicated by (9.2$b$).

The boundary conditions are identical to those posed in (3.1a)–(3.1e). For reference, the microscale solution to this problem is provided for a number of time points in figure 15. We note that problem $\langle c_2 \rangle$ must be put in the time coordinates for which the
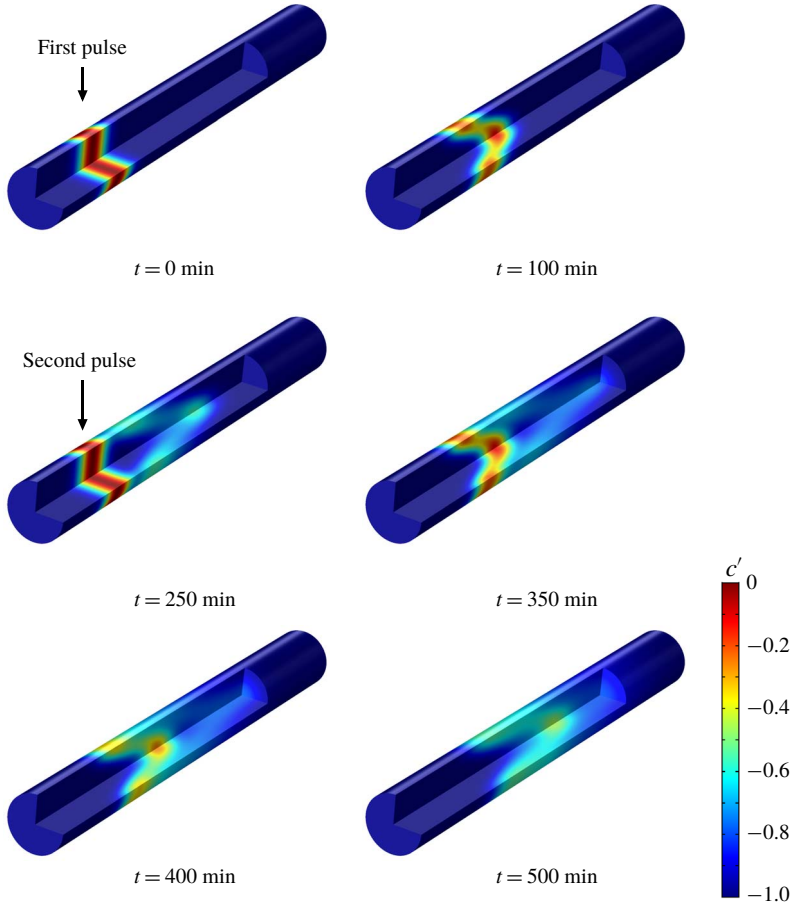
FIGURE 15. Microscale solute distributions in space at selected time points for the problem consisting of two injected pulses; Péclet number is $Pe = 100$. The second pulse is injected at the same location where the first pulse was injected at $t = 250$ min. These images are provided only for visualization of the microscale processes (aspect ratio is $1 : 10$ as for figures 6 and 7), and for validation of the macroscale equations by comparison of the directly averaged microscale fields.

injection happens at time equal to zero (this was noted in §7.2) to properly define D^* and s^*.

We now explore how these two solutions can be superimposed. First, note that the classical principle of superposition (Polyanin 2000; Olver 2014) for linear partial differential equations with time-dependent coefficients requires that the superposition be defined over the same time interval. Thus, problems I_1 and I_2 can be equivalently written as (recalling that for the general problem, the initial condition is represented by $\varphi(z)$; cf. (3.1e))

Problem $\langle c_1 \rangle$ $I_1 : 0 < t < 1000$ min

Interval 1 $0 < t < 250$ min

$$\frac{\partial \langle c_1 \rangle}{\partial t} = D^*(t) \frac{\partial^2 \langle c_1 \rangle}{\partial z^2} - U \frac{\partial \langle c_1 \rangle}{\partial z} + s_1{}^*(z, t), \tag{9.4a}$$

$$\text{I.C.1} \quad \langle c_1 \rangle |_{(z,t=0)} = c_0 Z_1(z). \tag{9.4b}$$

Interval 2 $0 < t' < 750$ min

$$\frac{\partial \langle c_1 \rangle}{\partial t'} = D^*(t') \frac{\partial^2 \langle c_1 \rangle}{\partial z^2} - U \frac{\partial \langle c_1 \rangle}{\partial z} + s_1^*(z, t'), \tag{9.4c}$$

$$\text{I.C.2} \quad \langle c_1 \rangle |_{(z,t'=0)} = \underbrace{\langle c_1 \rangle |_{(z,t=T_1)}}_{\varphi_1}. \tag{9.4d}$$

Problem $\langle c_2 \rangle$ $I_2 : 0 < t < 1000$ min

Interval 1 $0 < t < 250$ min

$$\langle c_2 \rangle = 0. \tag{9.4e}$$

Interval 2 $0 < t' < 750$ min

$$\frac{\partial \langle c_2 \rangle}{\partial t'} = D^*(t') \frac{\partial^2 \langle c_2 \rangle}{\partial z^2} - U \frac{\partial \langle c_2 \rangle}{\partial z} + s_2^*(z, t'), \tag{9.4f}$$

$$\text{I.C.2} \quad \langle c_2 \rangle |_{(z,t'=0)} = \underbrace{c_0 Z_1(z)}_{\varphi_2}. \tag{9.4g}$$

We make the following important notes about the formulation above.

(i) We have divided Problem 1 into two intervals $0 < t < 250$ min and $250 < t < 1000$ min. This is not an arbitrary choice; it is necessary to obtain solutions that involve multiple initial conditions.

(ii) In our formulation, the dispersion coefficient D^* does not have to 'remember' its prior history. Hence, exactly the same function D^* is used in both intervals. Being explicit, for the problem for $\langle c_1 \rangle$ in the first time interval, the dispersion coefficient starts at $D^*(t = 0) = \mathcal{D}$, and grows as predicted by (7.3). When time is restarted in the second interval, exactly the same behaviour is repeated. The dispersion coefficient for the second time interval starts at $D^*(t = 0) = \mathcal{D}$, and again grows as predicted by (7.3). During the second time interval, the source function s_1^* automatically accounts for the spreading inherent in the second initial condition. This function is essentially the 'memory' in the system, but one that encodes that memory in a local rather than non-local in time representation.

(iii) The source function, s^*, is linear in the initial condition; this is proved in § A.5. In other words, we have the following decomposition due to linearity:

$$s^*(\varphi_1 + \varphi_2) = s^*(\varphi_1) + s^*(\varphi_2). \tag{9.5}$$

(iv) These features make the two problems superposable; the two problems can be added to generate one equation that suffices to describe the dispersion process over the two intervals.

9.2.2. *Superposition of the two concentrations*

We now show how superposition can be accomplished for the problem of two releases of solute at two different times. We begin by defining the superposed concentration $\langle c \rangle = \langle c_1 \rangle + \langle c_2 \rangle$. For the first time interval, the superposition is trivial, noting that $\langle c_2 \rangle \equiv 0$. Thus, $\langle c \rangle = \langle c_1 \rangle$. Superposition for the first interval yields

Superposition for interval 1 $0 < t < 250$ min

$$\frac{\partial \langle c \rangle}{\partial t} = D^*(t) \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} + s_1^*(z, t), \tag{9.6a}$$

$$\text{I.C.1} \quad \langle c \rangle|_{(z, t=0)} = c_0 Z_1(z). \tag{9.6b}$$

Note that the concentration at the final time $t = T_1 = 250$ min, forms the initial condition for $\langle c_1 \rangle$ during the second interval. For reference, we denote this concentration by $\langle c_1 \rangle|_{(z, t=T_1)}$.

Because (9.4d) and (9.4g) are (i) linear, and (ii) defined over the same time interval $0 < t' < (T_f - T_1)$, the superposition of these equations is straightforward. Adding (9.4c) and (9.4f) yields

Superposition for interval 2 $0 < t' < 750$ min

$$\frac{\partial \langle c \rangle}{\partial t'} = D^*(t') \frac{\partial^2 \langle c \rangle}{\partial z^2} - U \frac{\partial \langle c \rangle}{\partial z} + s^*(z, t'). \tag{9.6c}$$

Here, we have used the linearity of the source term: $s^*(z, t'; \varphi) = s_1^*(z, t'; \varphi_1) + s_2^*(z, t'; \varphi_2)$ to generate the single source term field, s^*, for the superposed problem. The initial condition for the second interval is found by adding the final concentration from time interval 1, $\langle c_1 \rangle|_{(z, t=T_1)}$, to the initial condition for the second solute release.

$$\text{I.C.2} \quad \langle c \rangle|_{(z, t'=0)} = \underbrace{\langle c_1 \rangle|_{(z, t=T_1)} + c_0 Z_1(z)}_{\varphi}. \tag{9.6d}$$

The Taylor dispersion problem is frequently represented by an equation of the form of (9.4d) and (9.4g), but with $s_1^* = s_2^* = 0$ (cf. Smith 1981a). Superposition for this case fails because even if one divides the problem into two time intervals. Upon restarting time (setting $t' = 0$) for the second interval, the concentrations $\langle c_1 \rangle$ and $\langle c_2 \rangle$ are represented by two different linear operators (one for $\langle c_1 \rangle$ where $D^*(t)$ begins at $t = T_1$, and a second for $\langle c_2 \rangle$ where $D^*(t')$ begins at $t' = 0$). In the formulation we provide, $D^*(t)$ is always given by a single function of time everywhere in the domain; any 'time history' associated with multiple injections at different times is accounted for strictly through the source term field, s^*. This avoids the possibility of having D^* being multi-valued at a single spatial location.

9.2.3. *Proof of validity of superposition by direct computation*

The superposition of $\langle c_1 \rangle$ and $\langle c_2 \rangle$ is given by (9.6a)–(9.6d). For clarity, we re-emphasize the following regarding the interpretation of these equations:

(i) During the first time interval ($0 < t < T_1$) the dispersion coefficient for the time interval starts at $D^*(t = 0) = \mathcal{D}$, and again grows as predicted by (7.3). This is described by the single transport equation given by (9.6a,b). Because the radial derivative of the initial condition for this interval is zero, we have that $s_1^*(z, t)$ is zero.

(ii) For the second time interval, the initial condition is the solution obtained from the end of the first time interval (at $t = T_1$) plus the initial condition representing the second solute release. This is described by the single transport equation given by (9.6c,b).
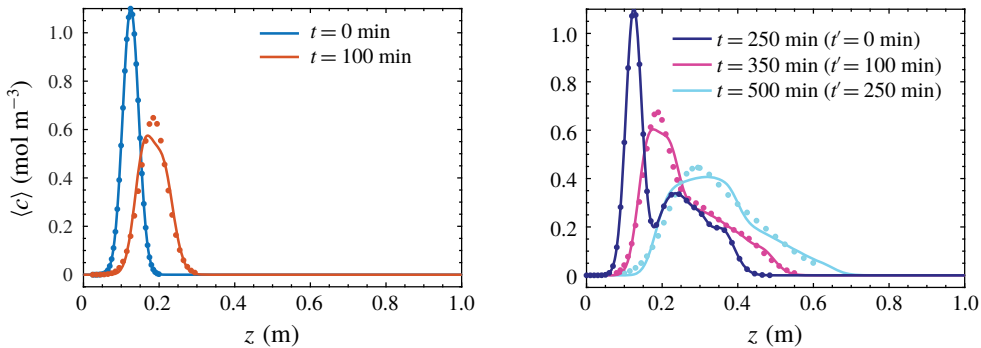
FIGURE 16. The macroscopic concentration field predicted by the upscaled problem statements given by I_1 and I_2. Continuous lines represents the averaged microscale simulation results; points represent the solutions to the upscaled equations given by (9.3a)–(9.3d). The concentration curves are provided as two plots for ease in visualization. Note that the curve in dark blue at $t = 250$ min represents the macroscale initial condition for problem I_2 computed by superposition.

(iii) During the second time interval ($0 < t < T_1$) the dispersion coefficient for the time interval starts at $D^*(t = 0) = \mathcal{D}$, and again grows as predicted by (7.3). Note that the radial derivative of the initial condition for the second time interval is non-zero, thus there is a contribution from the s^*-field for the second time interval.

The initial condition for this problem is similar to that for Problem 1. In fact, the s^* field is identical to that for Problem 1, and is computed numerically as described for Problem 1.

In figure 16 we have illustrated the solutions for this problem computed by (i) averaging the microscale numerical solutions directly, and (ii) via the two-step (superposed) macroscale simulation (solving (9.6a)–(9.6d)) as described above. We note that the two solutions are in good agreement, and, in particular, the amount of dispersion predicted by the upscaled model (both before and after the second pulse) is consistent with the averaged microscale data (for which there are no approximations of any kind). There is a slight mismatch in between the two solutions near the peaks for times $t = 100$, 350 and 500 min. Most likely, these errors are associated with the approximations in the upscaled equation, as discussed previously (cf. the peaks for $Pe = 100$, Case A in figure 8).

In summary, we have provided two examples that meet the desideratum of Taylor (1959). For each solution, we find that (i) we can define a sensible notion of superposition for the upscaled equation of the form that we have derived, (ii) the associated dispersion coefficient is always a single-valued function of space and (iii) the source terms are superposable, and they generate a field correction that assures that the proper rate of spreading (or, equivalently, second moment) is achieved. Because the source terms are responsible for correcting the rate of spreading at early times, this function is not imposed on the dispersion coefficient. It is this component of our solution that generates the necessary conditions for the upscaled equation to be superposable.

10. Conclusions

In this study, we were able to address a long-standing problem of early-time behaviour from arbitrary initial conditions. The approach we adopted generates

an effective macroscale balance equation that is consistent with the conventional one-dimensional convection–dispersion-type equation. One notable difference is that the formulation presented here also contains an exponentially decaying-in-time source term that represents memory of the initial condition. Because our results were compared with numerical simulations, we were able to compute the errors associated with the approximations made in our development. In short, these approximations are represented by

$$\text{(i)} \quad Pe\frac{a}{L_0} = \boldsymbol{O}(1),$$

$$\text{(ii)} \quad \frac{a}{L_0} < 1.$$

In addition, the four guidelines that we proposed in the introduction for generating a well-structured dispersion theory were met. Specifically, the dispersion coefficient for Taylor dispersion was shown to be positive, independent of initial conditions, superposable and converges to the classical result at large times. Of these, the most important result was illustrating that the non-conventional source term is necessary to assure that solutions are superposable. This addresses a long-standing problem about maintaining the principle of superposition when a time-dependent dispersion coefficient has been defined.

Declaration of interests

The authors report no conflict of interest.

Supplementary materials

Supplementary materials are available at https://doi.org/10.1017/jfm.2020.56.

Appendix A. Solutions to the closure problem
A.1. *General integral solutions*

The balance for the deviations developed in the body of the paper is given by the following set of equations:

$$\frac{\partial \tilde{c}}{\partial t} + \underbrace{\tilde{v}_z \frac{\partial \langle c \rangle}{\partial z}}_{source} = \frac{\mathcal{D}}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{c}}{\partial r}\right) + \frac{\mathcal{D}}{r^2}\frac{\partial^2 \tilde{c}}{\partial \theta^2} + \mathcal{D}\frac{\partial^2 \tilde{c}}{\partial z^2}, \tag{A 1a}$$

$$-\mathcal{D}\frac{\partial \tilde{c}}{\partial r} = 0, \quad \text{at } r = 0, a, \tag{A 1b}$$

$$\tilde{c} = 0, \quad \text{as } z \to \pm\infty, \tag{A 1c}$$

$$-\mathcal{D}\frac{\partial \tilde{c}}{\partial z} = 0, \quad \text{as } z \to \pm\infty, \tag{A 1d}$$

$$\tilde{c}(r, \theta, z, 0) = \underbrace{\tilde{\varphi}(r, \theta, z)}_{source}, \tag{A 1e}$$

where the sum of terms $\tilde{v}_z \partial \tilde{c}/\partial z - \langle \tilde{v}_z \partial \tilde{c}/\partial z \rangle$ has been neglected with respect to radial diffusion and a Lagrangian frame of coordinates is being used. This problem has exactly two source terms: one due to macroscopic convection and the second one due to the initial condition. The solution of this problem can be carried out by means of integral equations formulations in terms of Green's functions (Wood & Valdés-Parada 2013; Polyanin & Nazaikinskii 2015) with the associated Green's function (G) satisfying the following initial and boundary-value problem:

$$\frac{\partial G}{\partial t} = \frac{\mathcal{D}}{r}\frac{\partial}{\partial r}\left(r\frac{\partial G}{\partial r}\right) + \frac{\mathcal{D}}{r^2}\frac{\partial^2 G}{\partial \theta^2} + \mathcal{D}\frac{\partial^2 G}{\partial z^2} + \delta(r-\rho)\delta(\theta-\vartheta)\delta(z-\zeta)\delta(t-\tau), \tag{A 2a}$$

$$-\mathcal{D}\frac{\partial G}{\partial r} = 0, \quad \text{at } r = 0, a, \tag{A 2b}$$

$$G = 0, \quad \text{as } z \to \pm\infty, \tag{A 2c}$$

$$-\mathcal{D}\frac{\partial G}{\partial z} = 0, \quad \text{as } z \to \pm\infty, \tag{A 2d}$$

$$G = 0, \quad t < \tau. \tag{A 2e}$$

The solution to this problem can be written as

$$G(r, \rho, \theta, \vartheta, z, \zeta, t, \tau) = G_1(r, \rho, \theta - \vartheta, t - \tau)G_2(z - \zeta, t - \tau), \tag{A 3}$$

where, for convenience, the solution is expressed in terms of the Green's functions of unsteady polar (G_1) and axial coordinates (G_2). These Green's functions are given by

$$\begin{aligned}
G_1(r, \rho, \theta - \vartheta, t - \tau) = {} & \frac{1}{\pi a^2} \\
& + \frac{1}{\pi}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}\frac{A_n\lambda_{nm}^2 J_n(\lambda_{nm}r/a)J_n(\lambda_{nm}\rho/a)}{a^2(\lambda_{nm}^2 - n^2)J_n^2(\lambda_{nm})} \\
& \times \cos[n(\theta - \vartheta)]\exp\left(-\frac{\lambda_{nm}^2 \mathcal{D}}{a^2}(t - \tau)\right),
\end{aligned} \tag{A 4a}$$

$$G_2(z, \zeta, t - \tau) = \frac{1}{2\sqrt{\pi\mathcal{D}(t-\tau)}}\exp\left[-\frac{(z-\zeta)^2}{4\mathcal{D}(t-\tau)}\right]. \tag{A 4b}$$

In (A 4a) A_n is a constant whose value is 1 for $n = 0$ and 2 for $n = 1, 2, \ldots$ In addition, λ_{nm} are the positive roots of the transcendental equation $J_n'(\lambda_{nm}) = 0$. The solution of the problem given in (A 1e) can be obtained by using the unsteady version of Green's formula and the resulting expression can be written in the form given in (5.11). Direct substitution of this result into the unclosed macroscale balance equation yields the closed but non-local model given in (6.1).

A.2. *Localized solutions*

The relative advantages and disadvantages of non-local formulations have been discussed in the body of the paper. Regardless, we are ultimately interested in finding analytical solutions for the effective parameters that arise from averaging, so a localized form is desirable. Recall, this is valid when there is separation between the characteristic space and time scales for the microscale and macroscale variables, i.e.

$$a \ll L_0, \tag{A 5a}$$

$$t^* \ll T^*, \tag{A 5b}$$

where L_0 is a measure of the size of the initial configuration in the longitudinal direction, t^* is the characteristic time scale for the microscale processes and T^* is the characteristic time scale for the macroscale processes.

The length scale constraint is fairly clear in that it is expressed in terms of physical properties of the system. For the time scales, the restriction is less clear. We can attempt to (very roughly) characterize these time scales as follows. There are two reasonable time scales associated with the microscale balance given by (A 1a). First, the convection term involved in the material time derivative has a time scale that would be estimated by $T^* = \boldsymbol{O}(L/U)$. The diffusive time scale in the radial direction is estimated by $t^* = \boldsymbol{O}(a^2/(4\mathcal{D}))$. If we assume that $Pe_T \leqslant \boldsymbol{O}(1)$, then the diffusive time scale can be used for estimates. For the macroscale equation, (6.6a), estimates are complicated somewhat by the fact that the length scale of interest, L, is the longitudinal size of the solution as it spreads. Noting that we are attempting to generate estimates for the length scale associated with $\partial \langle c \rangle / \partial z$ in the convolutions given by (5.11), we can neglect the convection term (since we are in a moving coordinate system). The time scale can be estimated from the dispersive time scale, which we can think of as containing an initial time scale plus a transient time scale. In other words, we are making the estimate

$$L^2(t) = \boldsymbol{O}(L_0^2 + 4tD^*(t)) \tag{A 6}$$

from which we generate the time scale

$$T^* = \boldsymbol{O}\left(\frac{L_0^2}{D^*(0)} + \frac{4D^*(t)t}{D^*(0)}\right) = \boldsymbol{O}\left(\frac{L_0^2}{\mathcal{D}} + 4t\right). \tag{A 7}$$

Recalling $t^* = \boldsymbol{O}(a^2/(4\mathcal{D}))$, it is clear that the constraint $t^* \ll T^*$ is automatically met when $a \ll L_0$. Therefore, the single length scale constraint appears to be sufficient to justify the analysis, and this has been adopted in the body of the paper. We note that an alternate analysis of approximations of this sort is available in Mercer & Roberts (1990, appendix A), where the analysis is carried out in Fourier space in the context of a centre manifold theory approach to upscaling.

Additional details (and refinements about specific metrics for the time and length scales) for this approximation can be found in Wood & Valdés-Parada (2013). Regardless, exact conditions for when this approximation can be imposed are difficult to define; instead, we opt for the approach here of making the assumption that this approximation holds, and then examining the fidelity of the results. We do note that the removal of the axial derivative of the concentration is consistent with the second-order truncation presented by Gill & Sankarasubramanian (1970). The work of Gill & Sankarasubramanian (1970) suggests that the higher-order terms are much smaller than the second-order terms, which provides an independent validation of the approximation that we have employed.

A.3. *Analytical solutions for the b-field*

Assuming that the constraints allowing a local solution are valid, we can extract the axial derivative of the average concentration from the second integral in (5.11), we set

$$
\begin{aligned}
b(r, \theta, z, t) &= - \int_{\tau=0}^{\tau=t} \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r, \rho, \theta, \vartheta, z, \zeta, t-\tau) \tilde{v}_z(\rho) \rho \, d\rho \, d\vartheta \, d\zeta \, d\tau \\
&= - \int_{\tau=0}^{\tau=t} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G_1(r, \rho, \theta - \vartheta, t-\tau) \tilde{v}_z(\rho) \rho \, d\rho \, d\vartheta \\
&\quad \times \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t-\tau) \, d\zeta \, d\tau.
\end{aligned}
\tag{A 8}
$$

Note, we can make use of the identity

$$
\int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t-\tau) \, d\zeta = 1, \quad (t - \tau \neq 0).
\tag{A 9}
$$

It is clear from these results that the *b*-field is entirely independent of the particular initial condition adopted. Also, because of symmetry, the velocity deviations do not depend on the angular direction, θ or longitudinal coordinate, z. Direct integration leads to a result that is only a function of r and t

$$
b(r, t) = \frac{1}{4} \frac{U a^2}{\mathcal{D}} \left(\frac{r^2}{a^2} - \frac{r^4}{2a^4} - \frac{1}{3} \right) + \frac{2 U a^2}{\mathcal{D}} \sum_{n=1}^{\infty} \frac{J_3(\lambda_n) J_0\left(\lambda_n \frac{r}{a}\right)}{\lambda_n^3 J_0^2(\lambda_n)} \exp\left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t \right).
\tag{A 10}
$$

A simplification has been used here to decompose the series into the steady and transient parts, noting

$$
\sum_{n=1}^{\infty} \frac{J_3(\lambda_n) J_0(\lambda_n \frac{r}{a})}{\lambda_n^3 J_0^2(\lambda_n)} = -\frac{1}{8} \left(\frac{r^2}{a^2} - \frac{r^4}{2a^4} - \frac{1}{3} \right).
\tag{A 11}
$$

It is easy to verify that $\langle b \rangle = 0$. Similarly, as $t \to \infty$, we find

$$
\begin{aligned}
-\langle \tilde{v}_z b \rangle|_{t \to \infty} &= -\frac{1}{2} \frac{U^2 a^2}{\mathcal{D}} \left\langle \left(\frac{r^2}{a^2} - \frac{r^4}{2a^4} - \frac{1}{3} \right) \left(\frac{1}{2} - \frac{r^2}{a^2} \right) \right\rangle \\
&= \frac{U^2 a^2}{48 \mathcal{D}},
\end{aligned}
\tag{A 12}
$$

as expected.

A.4. *Analytical solutions for the Φ-field*

Following an analysis analogous to that for the *b*-field, we define the following functional, $\Phi(\tilde{\varphi})$:

$$
\Phi(r, \theta, z, t; \tilde{\varphi}) = \int_{\zeta=-\infty}^{\zeta=\infty} \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} G(r, \rho, \theta, \vartheta, z, \zeta, t) \tilde{\varphi}(\rho, \vartheta, \zeta) \rho \, d\rho \, d\vartheta \, d\zeta.
\tag{A 13}
$$

Evidently, in order to draw a specific expression for this closure variable it is necessary to provide a particular initial condition, $\tilde{\varphi}$. Substituting the Green's function G into (A 13), and rearranging, we recover the result given in (7.7) of the main text, which is applicable to a general initial condition

$$\Phi(r, \theta, z, t) = B_{0,0}(z, t) + \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} B_{n,m}(\theta, z, t) J_n\left(\frac{\lambda_{nm} r}{a}\right) \exp\left(-\frac{\lambda_{nm}^2 \mathcal{D}}{a^2} t\right), \quad \text{(A 14)}$$

where the functions $B_{0,0}(z, t)$ and $B_{n,m}(\theta, z, t)$, are defined by

$$B_{0,0}(z, t) = \frac{1}{\pi a^2} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z, \zeta, t) \int_{\vartheta=0}^{\vartheta=2\pi} \int_{\rho=0}^{\rho=a} \tilde{\varphi}(\rho, \vartheta, \zeta) \rho \, d\rho \, d\vartheta \, d\zeta, \quad \text{(A 15a)}$$

$$\begin{aligned} B_{n,m}(\theta, z, t) &= \frac{A_n \lambda_{nm}^2}{\pi a^2 (\lambda_{nm}^2 - n^2) J_n^2(\lambda_{nm})} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z, \zeta, t) \int_{\vartheta=0}^{\vartheta=2\pi} \cos[n(\theta - \vartheta)] \\ &\times \int_{\rho=0}^{\rho=a} J_n\left(\lambda_{nm} \frac{\rho}{a}\right) \tilde{\varphi}(\rho, \vartheta, \zeta) \rho \, d\rho \, d\vartheta \, d\zeta. \end{aligned} \quad \text{(A 15b)}$$

In principle, this expression (combined with (A 15a) and (A 15b)) provides the integral solution for any initial condition. Regardless of the particular initial condition chosen, note that Φ is linear in the initial condition function, i.e.

$$\varphi = \langle \varphi \rangle + \tilde{\varphi}, \quad \text{(A 16)}$$

hence, multiplying the initial condition by a constant K_1 yields

$$K_1 \varphi = K_1 \langle \varphi \rangle + K_1 \tilde{\varphi} \quad \text{(A 17)}$$

or

$$K_1 \varphi \Rightarrow K_1 \tilde{\varphi}. \quad \text{(A 18)}$$

The linearity in the initial condition is then established by

$$\Phi(K_1 \tilde{\varphi}_1 + K_2 \tilde{\varphi}_2) = K_1 \Phi(\varphi_1) + K_1 \Phi(\varphi_2), \quad \text{(A 19)}$$

as determined from (A 13).

As mentioned in the main body of the paper, we have found (non-trivial) particular solutions in analytical form (i.e. no unresolved integrations remain, and the result is given as a series expression) only for the case of θ-symmetric initial conditions. For the case of θ-symmetric initial conditions, the resulting expression for the closure variable Φ simplifies to

$$\Phi(r, z, t) = \sum_{n=1}^{\infty} B_n(z, t) J_0\left(\lambda_n \frac{r}{a}\right) \exp\left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t\right), \quad \text{(A 20)}$$

where $B_n(z, t)$ is defined by

$$B_n(z, t) = \frac{2}{a^2 J_0^2(\lambda_n)} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t) \int_{\rho=0}^{\rho=a} \tilde{\varphi}(\rho, \zeta) J_0\left(\lambda_n \frac{\rho}{a}\right) \rho \, d\rho \, d\zeta. \quad \text{(A 21)}$$

Note that in these results Φ is a time-decaying function that depends on t, r and z. The dependence on the axial direction arises from the initial condition. Furthermore, we note that if the initial condition is chosen to be independent of r, the consequence is that $B_n(z, 0) = 0$ and therefore $\Phi = 0$. As a final note, because $\langle J_0(\lambda_n(r/a)) \rangle \equiv 0$, we have that $\langle \Phi \rangle \equiv 0$ for all admissible initial conditions.

As a final note about the localized solutions, we note that with the developments above the concentration deviations specified by (5.11) can be put in the local formulation

$$\tilde{c}(r, \theta, z, t) = b(r, t) \left. \frac{\partial \langle c \rangle}{\partial z} \right|_{(z,t)} + \Phi(r, \theta, z, t). \tag{A 22}$$

The local macroscale balance equation is easily found to take the form given in (6.6a) of the main body of the paper.

A.5. *Analytical solutions for the effective coefficients*

Recall that the two effective parameters arising in the macroscale equation are

$$D^*(t) = \mathcal{D} - \langle \tilde{v}_z b \rangle, \tag{A 23a}$$

$$s^*(z, t) = - \left\langle \tilde{v}_z \frac{\partial \Phi}{\partial z} \right\rangle. \tag{A 23b}$$

Note that because Φ is a linear functional of the initial condition $\tilde{\varphi}$ (or, equivalently, φ; see § A.4), then so is $s^*(\tilde{\varphi})$. This is easy to check by direct substitution

$$
\begin{aligned}
s^*(K_1\tilde{\varphi}_1 + K_2\tilde{\varphi}_2) &= - \left\langle \tilde{v}_z \frac{\partial \Phi(K_1\tilde{\varphi}_1 + K_2\tilde{\varphi}_2)}{\partial z} \right\rangle \\
&= -K_1 \left\langle \tilde{v}_z \frac{\partial \Phi(\tilde{\varphi}_1)}{\partial z} \right\rangle - K_2 \left\langle \tilde{v}_z \frac{\partial \Phi(\tilde{\varphi}_2)}{\partial z} \right\rangle \\
&= K_1 s^*(\tilde{\varphi}_1) + K_2 s^*(\tilde{\varphi}_2)
\end{aligned}
\tag{A 24}
$$

or, equivalently, this linearity can be expressed in terms of the original initial conditions (see § A.4)

$$s^*(K_1\varphi_1 + K_2\varphi_2) = K_1 s^*(\varphi_1) + K_2 s^*(\varphi_2). \tag{A 25}$$

To compute D^*, it is helpful to note the following

$$\left\langle \tilde{v}_z J_0 \left(\lambda_n \frac{r}{a} \right) \right\rangle = \frac{2U}{\lambda_n} J_3(\lambda_n). \tag{A 26}$$

Using this average and the result of (A 12), the value of the effective dispersion coefficient is found to be

$$\frac{D^*(t)}{\mathcal{D}} = \left(1 + \frac{1}{48} \frac{U^2 a^2}{\mathcal{D}^2} \right) - 4 \frac{U^2 a^2}{\mathcal{D}^2} \sum_{n=1}^{\infty} \left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)} \right)^2 \exp \left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t \right). \tag{A 27}$$

Note that, at $t = 0$, the infinite series can be verified to be absolutely convergent. We also note that

$$4 \frac{U^2 a^2}{\mathcal{D}^2} \sum_{n=1}^{\infty} \left(\frac{J_3(\lambda_n)}{\lambda_n^2 J_0(\lambda_n)} \right)^2 = \frac{1}{48} \frac{U^2 a^2}{\mathcal{D}^2} \tag{A 28}$$

so that at the initial time we have that $D^*(t) = \mathcal{D}$; clearly, this series must be a monotonically increasing function of time. This also serves to show that the effective hydrodynamic diffusion coefficient is positive for all possible values of time. Equation (A 27) is equivalent to equation (19) in Gill & Sankarasubramanian (1970) (the comparison of solutions requires the identity $4/\lambda_n J_2(\lambda_n) = J_3(\lambda_n)$).

For computing s^*, (A 23b) is reworked by gathering the functions of the radial position into the averaging operator; the resulting equation is given in (7.9) of the main text. For the case in which the initial condition is θ-independent, this result simplifies to

$$s^*(z, t) = -\sum_{n=1}^{n=\infty} \left.\frac{\partial B_n}{\partial z}\right|_{(z,t)} \left\langle \tilde{v}_z J_0\left(\lambda_n \frac{r}{a}\right) \right\rangle \exp\left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t\right), \tag{A 29}$$

which can be further reduced, by taking into account the result given in (A 26)

$$s^*(z, t) = -2U \sum_{n=1}^{n=\infty} \left.\frac{\partial B_n}{\partial z}\right|_{(z,t)} \frac{J_3(\lambda_n)}{\lambda_n} \exp\left(-\lambda_n^2 \frac{\mathcal{D}}{a^2} t\right). \tag{A 30}$$

Note that the source term, when integrated over the entire domain, is identically zero. This is straightforward to show. First note

$$\int_{z=-\infty}^{z=\infty} \frac{\partial \Phi}{\partial z} \, dz = \Phi(z)\Big|_{-\infty}^{\infty} = 0. \tag{A 31}$$

Then, we can integrate both sides of s^* to yield

$$\int_{z=-\infty}^{z=\infty} s^*(z) \, dz = -\left\langle \tilde{v}_z(r) \int_{z=-\infty}^{z=\infty} \frac{\partial \Phi}{\partial z} \, dz \right\rangle = 0. \tag{A 32}$$

The condition that the integral of s^* taken over the entire domain be zero is crucial to the problem. Without this condition, the macroscale balance given by (6.6a) would have no well-defined steady-state solution.

A.6. *Numerical solutions for Φ*

In §9 of the body of the paper, a method for numerically computing Φ is needed for examining the problem of superposition. In general, a numerical method for determining the closure variables (and, hence, the effective parameters) can be found by substituting the solution form given by (A 22) into the general statement for the deviation balance ((A 1a)–(A 1e)). The result is the following two initial and boundary-value problems that can be solved numerically for the closure variables b and Φ. Because we do not need to solve the balance equation for b numerically at all for this problem, we present only the result for computing the solution for Φ numerically. In summary, the value of the Φ field can be found by solving the following initial and boundary-value problem:

$$\frac{\partial \Phi}{\partial t} = \frac{\mathcal{D}}{r} \frac{\partial}{\partial r}\left(r \frac{\partial \Phi}{\partial r}\right) + \mathcal{D} \frac{\partial^2 \Phi}{\partial z^2}, \tag{A 33a}$$

$$\left.\frac{\partial \Phi}{\partial r}\right|_{(r,\theta,z,t)} = 0, \quad r = 0, a, \tag{A 33b}$$

$$\Phi(r, \theta, z, t) = 0, \quad \text{for } z \to \pm\infty, \tag{A 33c}$$

$$-\mathcal{D}\frac{\partial \Phi}{\partial z}\bigg|_{(r,\theta,z,t)} = 0, \quad \text{for } z \to \pm\infty, \tag{A 33d}$$

$$\Phi(r, \theta, z, 0) = \tilde{\varphi}(r, z). \tag{A 33e}$$

Appendix B. Particular solution for s^*

As detailed above, the resulting expression for s^* depends on the particular type of initial condition in the system by means of $B_n(z, 0)$, as expressed in (A 21). In this section, an explicit form for $B_n(z, 0)$ (and consequently for s^*) is obtained by considering the case where the initial condition for the concentration of the solute is separable. As a matter of convenience, let the initial condition for the solute concentration be the superposition of two spatial functions.

To start, we impose the following structure on the initial condition, which summarizes the three case studies of interest here as explained in §7.3 of the main text,

$$\varphi(r, z) = R_1(r)Z_1(z) + R_2(r)Z_2(z), \tag{B 1}$$

where R_1 and R_2 are dimensionless piece-wise continuous functions of the radial position that will be specified later. Averaging then yields

$$\langle \varphi \rangle(z) = \langle R_1 \rangle Z_1(z) + \langle R_2 \rangle Z_2(z). \tag{B 2}$$

When subtracted from (B 1), the result gives rise to the following expression for the initial condition of the spatial deviations:

$$\tilde{\varphi}(r, z) = \tilde{R}_1(r)Z_1(z) + \tilde{R}_2(r)Z_2(z). \tag{B 3}$$

For the particular functions adopted here, we consider a linear function in the radial direction, and a Gaussian type function in the longitudinal direction. Thus we have

$$Z_1(z) = c_0\alpha_1 \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2}\right), \tag{B 4a}$$

$$Z_2(z) = c_0\alpha_2 \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2}\right). \tag{B 4b}$$

Here, α_1, and α_2 control the magnitude of the function, σ_1 and σ_2 control the width and β_1 and β_2 control the displacement along the z-axis.

B.1. *Linear function in radial distributions*

For the radial functions R_1 and R_2, we use the simple linear functions

$$R_1(r) = 1 - \frac{r}{a}, \tag{B 5a}$$

$$R_2(r) = \frac{r}{a}. \tag{B 5b}$$

Thus

$$\tilde{R}_1(r) = \frac{2}{3} - \frac{r}{a}, \tag{B 6a}$$

$$\tilde{R}_2(r) = \frac{r}{a} - \frac{2}{3}. \tag{B 6b}$$

It is easy to verify that the deviation function for the initial condition is

$$\tilde{\varphi}(z, r) = \left(\frac{2}{3} - \frac{r}{a}\right) c_0 \alpha_1 \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2}\right) + \left(\frac{r}{a} - \frac{2}{3}\right) c_0 \alpha_2 \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2}\right). \tag{B 7}$$

Noting that $\tilde{R}_2 = -\tilde{R}_1$, then the initial condition can be written as

$$\tilde{\varphi}(z, r) = \left(\frac{2}{3} - \frac{r}{a}\right) c_0 \left[\alpha_1 \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2}\right) - \alpha_2 \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2}\right)\right]. \tag{B 8}$$

Substituting this result into the expression for $B_n(z, t)$ gives

$$B_n(z, t) = \frac{2}{a^2 J_0^2(\lambda_n)} \int_{\zeta = -\infty}^{\zeta = \infty} G_2(z - \zeta, t)[Z_1(\zeta) - Z_2(\zeta)] \, d\zeta \int_{r=0}^{r=a} \tilde{R}_1(r) J_0\left(\lambda_n \frac{r}{a}\right) r \, dr. \tag{B 9}$$

Because $\tilde{R}_1$ is a simple linear function, and because of the properties of Bessel functions, an application of integration by parts is straightforward and leads to substantial simplification. To start, note that the required integral is

$$\int_{r=0}^{r=a} \tilde{R}_1 J_0\left(\lambda_n \frac{r}{a}\right) r \, dr = \frac{2}{3} \int_{r=0}^{r=a} J_0\left(\lambda_n \frac{r}{a}\right) r \, dr - \int_{r=0}^{r=a} \frac{r}{a} J_0\left(\lambda_n \frac{r}{a}\right) r \, dr \tag{B 10}$$

and the first of these two integrals is identically zero by application of (B 21). The second integral can be evaluated in terms of Struve functions, yielding

$$\int_{r=0}^{r=a} \tilde{R}_1 J_0\left(\lambda_n \frac{r}{a}\right) r \, dr = -\frac{a^2 \pi}{2\lambda_n^2} \boldsymbol{H}_1(\lambda_n) J_0(\lambda_n). \tag{B 11}$$

Here, $\boldsymbol{H}_1$ is a Struve H function (cf. equation (11.1.7) in Abramowitz & Stegun (1965), p. 480).

The integral involving G_2 is easily computed to be

$$\int_{\zeta = -\infty}^{\zeta = \infty} G_2(z - \zeta, t)[Z_1(\zeta) - Z_2(\zeta)] \, d\zeta = \frac{c_0 \sigma_1 \alpha_1}{\sqrt{\sigma_1^2 + 4\mathcal{D}t}} \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t}\right)$$
$$- \frac{c_0 \sigma_2 \alpha_2}{\sqrt{\sigma_2^2 + 4\mathcal{D}t}} \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t}\right). \tag{B 12}$$

Inserting this result into (B 9), leads to the following expression for $B_n(z, t)$:

$$B_n(z, t) = -\frac{c_0 \pi}{\lambda_n^2 J_0(\lambda_n)} \boldsymbol{H}_1(\lambda_n) \left[\frac{\sigma_1 \alpha_1}{\sqrt{\sigma_1^2 + 4\mathcal{D}t}} \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t}\right)\right.$$
$$\left. - \frac{\sigma_2 \alpha_2}{\sqrt{\sigma_2^2 + 4\mathcal{D}t}} \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t}\right)\right]. \tag{B 13}$$

Finally, recalling the definition of s^* given by (A 30), the final result of this analysis is obtained

$$
\begin{aligned}
s^*(z, t) = -4\pi c_0 U &\left[\frac{\sigma_1 \alpha_1 (z - \beta_1)}{(\sigma_1^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t} \right) \right. \\
&\left. - \frac{\sigma_2 \alpha_2 (z - \beta_2)}{(\sigma_2^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t} \right) \right] \sum_{n=1}^{\infty} \frac{\boldsymbol{H}_1(\lambda_n)}{\lambda_n^3} \frac{\mathrm{J}_3(\lambda_n)}{\mathrm{J}_0(\lambda_n)} \exp\left(-\frac{\lambda_n^2 \mathcal{D}}{a^2} t \right).
\end{aligned}
$$
(B 14)

B.2. *Step function in radial distributions*

Consider the situation in which the radial variation in the initial condition happens as a step function so that the functions R_1 and R_2 are defined by

$$
R_1(r) = \begin{cases} 1, & 0 \leqslant r \leqslant \dfrac{a}{2} \\ 0, & \dfrac{a}{2} < r \leqslant a \end{cases}
$$
(B 15a)

$$
R_2(r) = \begin{cases} 0, & 0 \leqslant r \leqslant \dfrac{a}{2} \\ 1, & \dfrac{a}{2} < r \leqslant a. \end{cases}
$$
(B 15b)

Applying the spatial averaging operator to (B 15b) leads to $\langle R_1 \rangle$ and $\langle R_2 \rangle$ given by

$$
\langle R_1 \rangle = \tfrac{1}{4}; \quad \langle R_2 \rangle = \tfrac{3}{4},
$$
(B 16)

which when subtracted from (B 15b) give, for the deviations, $\tilde{R}_1$ and $\tilde{R}_2$:

$$
\tilde{R}_1(r) = \begin{cases} \dfrac{3}{4}, & 0 \leqslant r \leqslant \dfrac{a}{2} \\ -\dfrac{1}{4}, & \dfrac{a}{2} < r \leqslant a \end{cases}
$$
(B 17a)

$$
\tilde{R}_2(r) = \begin{cases} -\dfrac{3}{4}, & 0 \leqslant r \leqslant \dfrac{a}{2} \\ \dfrac{1}{4}, & \dfrac{1}{2} < r \leqslant a. \end{cases}
$$
(B 17b)

For this particular initial condition, note that $\tilde{R}_1(r) = -\tilde{R}_2(r)$, consequently the initial condition for the concentration deviations are now

$$
\tilde{\varphi}(r, z) = [Z_1(z) - Z_2(z)]\tilde{R}_1(r).
$$
(B 18)

Recall that we have

$$
B_n(z, t) = \frac{2}{a^2 \mathrm{J}_0^2(\lambda_n)} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t)[Z_1(\zeta) - Z_2(\zeta)]\,\mathrm{d}\zeta \int_{r=0}^{r=a} \tilde{R}_1(r)\mathrm{J}_0\left(\lambda_n \frac{r}{a} \right) r\,\mathrm{d}r.
$$
(B 19)

On the basis of this initial condition, $B_n(z, 0)$ can be expressed as

$$
B_n(z, 0) = \frac{2}{a^2 J_0^2(\lambda_n)} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t)[Z_1(\zeta) - Z_2(\zeta)] \, d\zeta
$$
$$
\times \left[\frac{3}{4} \int_{r=0}^{r=a/2} J_0\left(\lambda_n \frac{r}{a}\right) r \, dr - \frac{1}{4} \int_{r=a/2}^{r=a} J_0\left(\lambda_n \frac{r}{a}\right) r \, dr \right]. \quad \text{(B 20)}
$$

Noting that (Olver *et al.* 2010)

$$
\int J_0\left(\lambda_n \frac{r}{a}\right) r \, dr = \frac{a}{\lambda_n} r J_1\left(\lambda_n \frac{r}{a}\right) + c, \quad \text{(B 21)}
$$

then, (B 20) can be simplified to give the following:

$$
B_n(z, 0) = \frac{J_1(\lambda_n/2)}{\lambda_n J_0^2(\lambda_n)} \int_{\zeta=-\infty}^{\zeta=\infty} G_2(z - \zeta, t)[Z_1(\zeta) - Z_2(\zeta)] \, d\zeta. \quad \text{(B 22)}
$$

Recalling the previous results for the integrals involving Z_1 and Z_2 given by (B 12), and the definition of s^* given by (A 30), the following explicit series solution for s^* is obtained:

$$
s^*(z, t) = 4c_0 U \left[\frac{\sigma_1 \alpha_1 (z - \beta_1)}{(\sigma_1^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t} \right) \right.
$$
$$
\left. - \frac{\sigma_2 \alpha_2 (z - \beta_2)}{(\sigma_2^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t} \right) \right]
$$
$$
\times \sum_{n=1}^{\infty} \frac{J_1(\lambda_n/2) J_3(\lambda_n)}{\lambda_n^2 J_0^2(\lambda_n)} \exp\left(-\frac{\lambda_n^2 \mathcal{D}}{a^2} t \right). \quad \text{(B 23)}
$$

B.3. *Accounting for convection*

For $Pe_T \ll 1$, the solutions above should give sufficient estimates for s^*. However, recall that the closure problem was transformed into the coordinate system moving with the velocity of the centre of mass of the system. At asymptotic times, the distribution of solutes is nearly uniform across the radius; thus, the average fluid velocity is identically equal to the mass average velocity. However, in this work we have been investigating the special condition where initially the solute is distributed in two parts, each with different initial mass averaged velocities. This poses some difficulty, because in general the averaged velocity of the centre of mass of the solute is an unknown function of time. This difference between the mass-averaged and average fluid velocities (and the potential problems that it poses) was also identified in the work of Gill & Sankarasubramanian (1971), Degance & Johns (1978b), Dentz & Carrera (2007) and Ratnakar & Balakotaiah (2011).

For initial-condition-type problems, the mass-averaged velocities can be found from the centre of mass of the moving solute distribution

$$
M_0 = 2\pi \int_{r=0}^{r=a} \int_{z=0}^{z=L} \varphi(r, z) r \, dr \, dz, \quad \text{(B 24a)}
$$

$$
M_{1,z}(t) = \frac{2\pi}{M_0} \int_{r=0}^{r=a} \int_{z=0}^{z=L} z c(r, z, t) r \, dr \, dz. \quad \text{(B 24b)}
$$

In these expressions, the first index indicates the order of the moment, and subsequent indexes indicate for which coordinate the moments are taken. With these definitions in place, the mass-averaged solute velocity is defined by

$$\langle v_{z,M} \rangle = \frac{\partial M_1(t)}{\partial t}. \tag{B 25}$$

Before moving on, it is worth recalling that, among the simplifications adopted in this work, it was assumed that the term $\tilde{v}_z \partial \tilde{c}/\partial z$ and its spatial average are negligible in comparison to the radial diffusive term (valid for $Pe(a/L_0) = \boldsymbol{O}(1)$). In the material that follows, we make corrections to the coordinate used to displace the solutions, but we do not alter the value U that appears in these solutions. The following constraint is adopted to indicate under what conditions this assumption is valid:

$$(\langle v_{z,M} \rangle|_t - U)\frac{\partial \langle c \rangle}{\partial z} \ll \tilde{v}_z \frac{\partial \tilde{c}}{\partial z} \tag{B 26}$$

or, making conventional order-of-magnitude estimates, a reasonable constraint is

$$(\langle v_{z,M} \rangle|_t - U)\frac{a}{L} \ll \boldsymbol{O}(\tilde{v}_z), \quad \forall (z,t). \tag{B 27}$$

Here, it has been assumed that reasonable estimates of the characteristic lengths of variations for $\tilde{c}$ and $\langle c \rangle$ are a and L, respectively. In the solutions below, it will be apparent that this constraint is easily met for the particular initial conditions studied.

It is convenient to denote the portion of the total initial mass per unit length of the tube given by each of the two parts of the initial condition be specified by

$$m_1(z) = 2\pi Z_1(z) \int_{r=0}^{r=a} R_1(r) r \, dr, \tag{B 28a}$$

$$m_2(z) = 2\pi Z_2(z) \int_{r=0}^{r=a} R_2(r) r \, dr. \tag{B 28b}$$

From the above definitions, the mass weightings for the two portions of the initial condition are defined as

$$\omega_1(r) = \frac{Z_1(z)}{m_1(z)} R_1(r) = R_1(r) \left(2\pi \int_{r=0}^{r=a} R_1(r) r \, dr \right)^{-1}, \tag{B 29a}$$

$$\omega_2(r) = \frac{Z_2(z)}{m_2(z)} R_2(r) = R_2(r) \left(2\pi \int_{r=0}^{r=a} R_2(r) r \, dr \right)^{-1}. \tag{B 29b}$$

Note that these weightings are distinctly different from the spatial weighting functions discussed in the main body of the paper. Using these weighting functions, the initial mass-averaged velocities are given by

$$\langle v_{z,M} \rangle_1 (t=0) = 2\pi \int_{r=0}^{r=a} \omega(w)_1(r) v_z(r) r \, dr, \tag{B 30a}$$

$$\langle v_{z,M} \rangle_2 (t=0) = 2\pi \int_{r=0}^{r=a} \omega_2(r) v_z(r) r \, dr. \tag{B 30b}$$

The difficulty now is to determine a time scale for which the initial condition spreads sufficiently to sample the entire velocity distribution. To this end, consider

the following rough estimate for the variance of the initial condition corresponding to an unbounded two-dimensional field:

$$\sigma(t) = \sqrt{4\mathcal{D}t}. \tag{B 31}$$

Hence, the interest is to find a time, t_d^*, for which $\sigma = a$, i.e.

$$t_d^* = \frac{a^2}{4\mathcal{D}}, \tag{B 32}$$

which is referred to as the diffusive time scale. With this time scale established, it is now necessary to develop a model that describes how the average velocity of each of the two non-uniform initial condition contributions (R_1 and R_2) relaxes from its initial average to the average fluid velocity over the time t_d^*. Since the process is diffusion driven, a simple square-root-of-time behaviour seems reasonable. Our proposed model is approximate, but is correct at the initial and asymptotic times. Essentially, we assume that the average solute velocity transitions from its value at $t = 0$ to the average fluid velocity, U, over the time scale t_d^* following a $t^{1/2}$ law.

B.3.1. *Linear distribution*

For the linear radial distribution in the initial condition, the corresponding weighting functions are

$$\omega_1(r) = \frac{3}{a^3\pi}(a - r), \tag{B 33a}$$

$$\omega_2(r) = \frac{3}{2a^3\pi}r. \tag{B 33b}$$

The initial average velocities are given by

$$\langle v_{z,M}\rangle_1(t=0) = \tfrac{7}{5}U, \tag{B 34a}$$

$$\langle v_{z,M}\rangle_2(t=0) = \tfrac{4}{5}U. \tag{B 34b}$$

Assuming once more a square-root-in-time transition to the fluid velocity over the time t_d^*, the transient velocities associated with the linear initial distributions R_1 and R_2 are

$$\langle v_{z,M}\rangle_1(t) = \begin{cases} \dfrac{7}{5}U - \dfrac{2}{5}U\left(\dfrac{t}{t_d^*}\right)^{1/2}, & \text{for } t < t_d^*, \\ U, & \text{for } t \geqslant t_d^*, \end{cases} \tag{B 35a}$$

$$\langle v_{z,M}\rangle_2(t) = \begin{cases} \dfrac{4}{5}U + \dfrac{1}{5}U\left(\dfrac{t}{t_d^*}\right)^{1/2}, & \text{for } t < t_d^*, \\ U, & \text{for } t \geqslant t_d^*. \end{cases} \tag{B 35b}$$

To help visualize this approximation, in figure 17 the predictions of the centre of mass velocity resulting from the approximate solution are compared with those from numerical simulations. The approximate solution qualitatively predicts the dynamics of the centre of mass and, on quantitative grounds, the results are reasonable. This analysis provides some guidance for constructing approximate early-time velocities for other initial conditions.
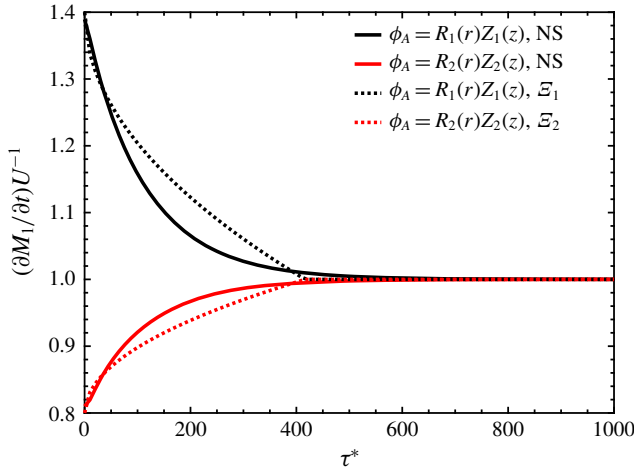
FIGURE 17. Modelled transient centre of mass velocities (equations (B 35a)–(B 35b)) versus the centre of mass velocities extracted from the microscale numerical solutions. $Pe = 100$.

The corresponding expressions for the translations created by these time-varying velocity fields are now given by

$$
\Xi_1(t) = \begin{cases} \dfrac{7}{5}Ut - \dfrac{4}{15}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\[2mm] \Xi_1(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*, \end{cases} \tag{B 36a}
$$

$$
\Xi_2(t) = \begin{cases} \dfrac{4}{5}Ut + \dfrac{2}{15}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\[2mm] \Xi_2(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*. \end{cases} \tag{B 36b}
$$

Finally, the convected solution for the initial condition considered here can be represented by

$$
\begin{aligned}
s^*(z, t) = {}&-4\pi c_0 U \left[\frac{\sigma_1 \alpha_1 (z - \Xi_1(t) - \beta_1)}{(\sigma_1^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_1(t) - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t} \right) \right. \\
&\left. - \frac{\sigma_2 \alpha_2 (z - \Xi_2(t) - \beta_2)}{(\sigma_2^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_2(t) - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t} \right) \right] \\
&\times \sum_{n=1}^{\infty} \frac{\boldsymbol{H}_1(\lambda_n)}{\lambda_n^3} \frac{\mathrm{J}_3(\lambda_n)}{\mathrm{J}_0(\lambda_n)} \exp\left(-\frac{\lambda_n^2 \mathcal{D}}{a^2} t \right).
\end{aligned} \tag{B 37}
$$

B.3.2. *Step distribution*

For the step radial distribution, the weighting functions are $\omega_1 = 4F_1/(\pi a^2)$ and $\omega_2 = 4F_2/(3\pi a^2)$. The associated velocities are given by

$$
\langle v_{z,M} \rangle_1 (t = 0) = \tfrac{7}{4}U, \tag{B 38a}
$$

$$
\langle v_{z,M} \rangle_2 (t = 0) = \tfrac{3}{4}U. \tag{B 38b}
$$

Assuming a square-root-in-time transition to the fluid velocity over the time t_d^*, the transient velocities associated with the initial distributions R_1 and R_2 are found to be

$$\langle v_{z,M}\rangle_1(t) = \begin{cases} \dfrac{7}{4}U - \dfrac{3}{4}U\left(\dfrac{t}{t_d^*}\right)^{1/2}, & \text{for } t < t_d^*, \\ U, & \text{for } t \geqslant t_d^*, \end{cases} \tag{B 39a}$$

$$\langle v_{z,M}\rangle_2(t) = \begin{cases} \dfrac{3}{4}U + \dfrac{1}{4}U\left(\dfrac{t}{t_d^*}\right)^{1/2}, & \text{for } t < t_d^*, \\ U, & \text{for } t \geqslant t_d^*. \end{cases} \tag{B 39b}$$

Denoting the translations created by these time-varying velocity fields by $Z(t)$, it follows that

$$\Xi_1(t) = \begin{cases} \dfrac{7}{4}Ut - \dfrac{1}{2}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\ \Xi_1(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*, \end{cases} \tag{B 40a}$$

$$\Xi_2(t) = \begin{cases} \dfrac{3}{4}Ut + \dfrac{1}{6}U\dfrac{t^{3/2}}{\sqrt{t_d^*}}, & \text{for } t < t_d^*, \\ \Xi_2(t_d^*) + U(t - t_d^*), & \text{for } t \geqslant t_d^*. \end{cases} \tag{B 40b}$$

With these definitions in place, the convected solution for the initial condition considered here can be represented by

$$\begin{aligned} s^*(z, t) = 4c_0U &\left[\frac{\sigma_1\alpha_1(z - \Xi_1 - \beta_1)}{(\sigma_1^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_1 - \beta_1)^2}{\sigma_1^2 + 4\mathcal{D}t}\right) \right. \\ &\left. - \frac{\sigma_2\alpha_2(z - \Xi_2 - \beta_2)}{(\sigma_2^2 + 4\mathcal{D}t)^{3/2}} \exp\left(-\frac{(z - \Xi_2 - \beta_2)^2}{\sigma_2^2 + 4\mathcal{D}t}\right) \right] \\ &\times \sum_{n=1}^{\infty} \frac{J_1(\lambda_n/2)J_3(\lambda_n)}{\lambda_n^2 J_0^2(\lambda_n)} \exp\left(-\frac{\lambda_n^2\mathcal{D}}{a^2}t\right). \end{aligned} \tag{B 41}$$

REFERENCES

ABBOTT, M. S. R., HARVEY, A. P., PEREZ, G. V. & THEODOROU, M. K. 2013 Biological processing in oscillatory baffled reactors: operation, advantages and potential. *Interface Focus* **3** (1), 20120036.

ABRAMOWITZ, M. & STEGUN, I. A. 1965 *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover.

ARIS, R. 1956 On the dispersion of a solute in a fluid flowing through a tube. *Proc. R. Soc. Lond.* A **235** (1200), 67–77.

BALAKOTAIAH, V., CHANG, H.-C. & SMITH, F. T. 1995 Dispersion of chemical solutes in chromatographs and reactors. *Phil. Trans. R. Soc. Lond.* A **351** (1695), 39–75.

BALAKOTAIAH, V. & RATNAKAR, R. R. 2010 On the use of transfer and dispersion coefficient concepts in low-dimensional diffusion–convection–reaction models. *Chem. Engng Res. Des.* **88** (3), 342–361.

BENDER, C. M. & ORSZAG, S. A. 1978 *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill.

BIRD, R. B., STEWART, W. E. & LIGHTFOOT, E. N. 2007 *Transport Phenomena*. John Wiley & Sons.

BOURNE, J. R. 2003 Mixing and the selectivity of chemical reactions. *Organic Process Res. Dev.* **7** (4), 471–508.

CHASSEIGNE, E., CHAVES, M. & ROSSI, J. D. 2006 Asymptotic behavior for nonlocal diffusion equations. *J. Math. Pure Appl.* **86** (3), 271–291.

CHATWIN, P. C. 1970 The approach to normality of the concentration distribution of a solute in a solvent flowing along a straight pipe. *J. Fluid Mech.* **43** (2), 321–352.

CHATWIN, P. C. 1972 The cumulants of the distribution of concentration of a solute dispersing in solvent flowing through a tube. *J. Fluid Mech.* **51** (1), 63–67.

CHATWIN, P. C. 1976 The initial dispersion of contaminant in Poiseuille flow and the smoothing of the snout. *J. Fluid Mech.* **77** (3), 593–602.

CHATWIN, P. C. 1977 The initial development of longitudinal dispersion in straight tubes. *J. Fluid Mech.* **80** (1), 33–48.

CHIKWENDU, S. C. 1986a Application of the slow-zone model to contaminant dispersion in laminar shear flows. *Intl J. Engng Sci.* **24**, 1031–1044.

CHIKWENDU, S. C. 1986b Calculation of longitudinal shear dispersivity using an n-zone model as $n \rightarrow \infty$. *J. Fluid Mech.* **167**, 19–30.

CHIKWENDU, S. C. & OJIAKOR, G. U. 1985 Slow-zone model for longitudinal dispersion in two-dimensional shear flows. *J. Fluid Mech.* **152**, 15–38.

CUSHMAN, J. H. & GINN, T. R. 1993 Nonlocal dispersion in media with continuously evolving scales of heterogeneity. *Trans. Porous Med.* **13** (1), 123–138.

DEGANCE, A. E. & JOHNS, L. E. 1978a The theory of dispersion of chemically active solutes in a rectilinear flow field. *Appl. Sci. Res.* **34** (2–3), 189–225.

DEGANCE, A. E. & JOHNS, L. E. 1978b On the dispersion coefficients for Poiseuille flow in a circular cylinder. *Appl. Sci. Res.* **34** (2–3), 227–258.

DEGANCE, A. E. & JOHNS, L. E. 1980 On the construction of dispersion approximations to the solution of the convective diffusion equation. *AIChE J.* **26** (3), 411–419.

DENG, F.-W. & CUSHMAN, J. H. 1995 Comparison of moments for classical-, quasi-, and convolution-Fickian dispersion of a conservative tracer. *Water Resour. Res.* **31** (4), 1147–1149.

DENG, F.-W., CUSHMAN, J. H. & DELLEUR, J. W. 1993 A fast Fourier transform stochastic analysis of the contaminant transport problem. *Water Resour. Res.* **29** (9), 3241–3247.

DENTZ, M. & CARRERA, J. 2007 Mixing and spreading in stratified flow. *Phys. Fluids* **19** (1), 017107.

EPSTEIN, I. R. 1990 Shaken, stirred – but not mixed. *Nature* **346** (6279), 16–17.

ERINGEN, A. C. 1978 Nonlocal continuum mechanics and some applications. In *Nonlinear Equations in Physics and Mathematics*, pp. 271–318. Springer.

FIFE, P. C. & NICHOLES, K. R. K. 1975 Dispersion in flow through small tubes. *Proc. R. Soc. Lond.* A **344** (1636), 131–145.

FRIED, J. J. & COMBARNOUS, M. A. 1971 Dispersion in porous media. *Adv. Hydrosci.* **7**, 169–282.

GARCÍA-MELIÁN, J. & ROSSI, J. D. 2009 On the principal eigenvalue of some nonlocal diffusion problems. *J. Differ. Equ.* **246** (1), 21–38.

GEKLE, S. 2017 Dispersion of solute released from a sphere flowing in a microchannel. *J. Fluid Mech.* **819**, 104–120.

GELHAR, L. W., GUTJAHR, A. L. & NAFF, R. L. 1979 Stochastic analysis of macrodispersion in a stratified aquifer. *Water Resour. Res.* **15** (6), 1387–1397.

GILL, W. N. & SANKARASUBRAMANIAN, R. 1970 Exact analysis of unsteady convective diffusion. *Proc. R. Soc. Lond.* A **316** (1526), 341–350.

GILL, W. N. & SANKARASUBRAMANIAN, R. 1971 Dispersion of a non-uniform slug in time-dependent flow. *Proc. R. Soc. Lond.* A **322** (1548), 101–117.

GILL, W. N. & SUBRAMANIAN, R. S. 1980 Discussion: 'On laminar dispersion for flow through round tubes'. *J. Appl. Mech.* **47**, 975–976.

GRAY, W. G. 1975 A derivation of the equations for multi-phase transport. *Chem. Engng Sci.* **30** (2), 229–233.

GRAY, W. G. & MILLER, C. T. 2009 Thermodynamically constrained averaging theory approach for modeling flow and transport phenomena in porous medium systems: 5. Single-fluid-phase transport. *Adv. Water Resour.* **32** (5), 681–711.

GRAY, W. G. & MILLER, C. T. 2014 *Introduction to the Thermodynamically Constrained Averaging Theory for Porous Medium Systems.* Springer.

HABER, S. & MAURI, R. 1988 Lagrangian approach to time-dependent laminar dispersion in rectangular conduits. Part 1. Two-dimensional flows. *J. Fluid Mech.* **190**, 201–215.

HERIS, S. Z., ESFAHANY, M. N. & ETEMAD, G. 2007 Numerical investigation of nanofluid laminar convective heat transfer through a circular tube. *Numer. Heat Transfer* A **52** (11), 1043–1058.

HORN, F. J. M. 1971 Calculation of dispersion coefficients by means of moments. *AIChE J.* **17** (3), 613–620.

IGNAT, L. I. & ROSSI, J. D. 2007 A nonlocal convection–diffusion equation. *J. Funct. Anal.* **251** (2), 399–437.

JONES, S. W. & YOUNG, W. R. 1994 Shear dispersion and anomalous diffusion by chaotic advection. *J. Fluid Mech.* **280**, 149–172.

KRONBERG, A. E., BENNEKER, A. H. & WESTERTERP, K. R. 1996 Wave model for longitudinal dispersion: application to the laminar-flow tubular reactor. *AIChE J.* **42** (11), 3133–3145.

LACEY, P. M. C. 1954 Developments in the theory of particle mixing. *J. Appl. Chem.* **4** (5), 257–268.

LAKE, L. W. & HIRASAKI, G. J. 1981 Taylor's dispersion in stratified porous media. *Soc. Petrol. Engng J.* **21** (04), 459–468.

LATINI, M. & BERNOFF, A. J. 2001 Transient anomalous diffusion in Poiseuille flow. *J. Fluid Mech.* **441**, 399–411.

LIGHTHILL, M. J. 1966 Initial development of diffusion in Poiseuille flow. *IMA J. Appl. Maths* **2** (1), 97–108.

LIU, B. & IVORY, C. F. 2013 Isotachophoresis with counterflow in an open capillary: computer simulation and experimental validation. *J. Separation Sci.* **36** (12), 1986–1995.

MARCINKIEWICZ, J. 1939 Sur une propriété de la loi de Gauß ('On a property of Gauss' law'). *Math. Z.* **44**, 612–618.

MAURI, R. 1991 Dispersion, convection, and reaction in porous media. *Phys. Fluids* A **3** (5), 743–756.

MENG, X. & YANG, D. 2018 Dynamic dispersion coefficient of solutes flowing in a circular tube and a tube-bundle model. *J. Energy Resour. Technol.* **140** (1), 012903.

MERCER, G. N. & ROBERTS, A. J. 1990 A centre manifold description of contaminant dispersion in channels with varying flow properties. *SIAM J. Appl. Maths* **50** (6), 1547–1565.

MERCER, G. N. & ROBERTS, A. J. 1994 A complete model of shear dispersion in pipes. *Japan J. Ind. Appl. Math.* **11** (3), 499–521.

MILLER, C. T., VALDÉS-PARADA, F. J., OSTVAR, S. & WOOD, B. D. 2018 *A priori* parameter estimation for the thermodynamically constrained averaging theory: species transport in a saturated porous medium. *Trans. Porous Med.* **122** (3), 611–632.

MOHAND, H. S. H., FREZZOTTI, A., BRANDNER, J. J., BARROT, C. & COLIN, S. 2017 Molecular tagging velocimetry by direct phosphorescence in gas microflows: correction of Taylor dispersion. *Exp. Therm. Fluid Sci.* **83**, 177–190.

NEUMAN, S. P. 1993 Eulerian–Lagrangian theory of transport in space–time nonstationary velocity fields: exact nonlocal formalism by conditional moments and weak approximation. *Water Resour. Res.* **29** (3), 633–645.

O'HARA, K. 1969 Dispersion of soluble matter in solvent flowing through a straight pipe. Master's thesis, University of Liverpool.

OLVER, F. W. J., LOZIER, D. W., BOISVERT, R. F. & CLARK, C. W. 2010 *NIST Handbook of Mathematical Functions.* Cambridge University Press. Available at: https://dlmf.nist.gov/.

OLVER, P. J. 2014 *Introduction to Partial Differential Equations.* Springer.

OSTVAR, S. & WOOD, B. D. 2016 A non-scale-invariant form for coarse-grained diffusion–reaction equations. *J. Chem. Phys.* **145** (11), 114105.

PAINE, M. A., CARBONELL, R. G. & WHITAKER, S. 1983 Dispersion in pulsed systems. I. Heterogenous reaction and reversible adsorption in capillary tubes. *Chem. Engng Sci.* **38** (11), 1781–1793.

PAWULA, R. F. 1967 Approximation of the linear Boltzmann equation by the Fokker–Planck equation. *Phys. Rev.* **162** (1), 186–188.

PHILIP, J. R. 1963 The theory of dispersal during laminar flow in tubes. I. *Austral. J. Phys.* **16** (3), 287–299.

PHILLIPS, C. G. & KAYE, S. R. 1996 A uniformly asymptotic approximation for the development of shear dispersion. *J. Fluid Mech.* **329**, 413–443.

PHILLIPS, C. G. & KAYE, S. R. 1997 The initial transient of concentration during the development of Taylor dispersion. *Proc. R. Soc. Lond.* A **453**, 2669–2688.

POLYANIN, A. D. 2000 *Handbook of Linear Partial Differential Equations for Engineers and Scientists*. Chapman and Hall/CRC.

POLYANIN, A. D. & NAZAIKINSKII, V. E. 2015 *Handbook of Linear Partial Differential Equations for Engineers and Scientists*. CRC Press.

RATNAKAR, R. R. & BALAKOTAIAH, V. 2011 Exact averaging of laminar dispersion. *Phys. Fluids* **23** (2), 023601.

RISKEN, H. & VOLLMER, H. D. 1979 On the application of truncated generalized Fokker–Planck equations. *Z. Phys.* B **35** (3), 313–315.

ROACHE, P. J. 1994 Perspective: a method for uniform reporting of grid refinement studies. *J. Fluids Engng* **116** (3), 405–413.

ROACHE, P. J. 2003 Conservatism of the grid convergence index in finite volume computations on steady-state fluid flow and heat transfer. *J. Fluids Engng* **125** (4), 731–732.

SANKARASUBRAMANIAN, R. & GILL, W. N. 1973 Unsteady convective diffusion with interphase mass transfer. *Proc. R. Soc. Lond.* A **333** (1592), 115–132.

SMITH, R. 1981a A delay-diffusion description for contaminant dispersion. *J. Fluid Mech.* **105**, 469–486.

SMITH, R. 1981b The early stages of contaminant dispersion in shear flows. *J. Fluid Mech.* **111**, 107–122.

SMITH, R. 1984 Temporal moments at large distances downstream of contaminant releases in rivers. *J. Fluid Mech.* **140**, 153–174.

STOKES, A. N. & BARTON, N. G. 1990 The concentration distribution produced by shear dispersion of solute in Poiseuille flow. *J. Fluid Mech.* **210**, 201–221.

STONESTREET, P. & HARVEY, A. P. 2002 A mixing-based design methodology for continuous oscillatory flow reactors. *Chem. Engng Res. Des.* **80** (1), 31–44.

TAYLOR, G. I. 1953 Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. R. Soc. Lond.* A **219**, 186–203.

TAYLOR, G. I. 1954 The dispersion of matter in turbulent flow through a pipe. *Proc. R. Soc. Lond.* A **223** (1155), 446–468.

TAYLOR, G. I. 1959 The present position in the theory of turbulent diffusion. In *Advances in Geophysics*, vol. 6, pp. 101–112. Elsevier.

TSENG, C. M. & BESANT, R. W. 1970 Dispersion of a solute in a fully developed laminar tube flow. *Proc. R. Soc. Lond.* A **317** (1528), 91–99.

TSENG, C. M. & BESANT, R. W. 1972 Transient heat and mass transfer in fully developed laminar tube flows. *Intl J. Heat Mass Transfer* **15** (2), 203–215.

VALOCCHI, A. J. 1989 Spatial moment analysis of the transport of kinetically adsorbing solutes through stratified aquifers. *Water Resour. Res.* **25** (2), 273–279.

VEDEL, S., HOVAD, E. & BRUUS, H. 2014 Time-dependent Taylor–Aris dispersion of an initial point concentration. *J. Fluid Mech.* **752**, 107–122.

VILLERMAUX, E. 2019 Mixing versus stirring. *Annu. Rev. Fluid Mech.* **51**, 245–273.

VRENTAS, J. S. & VRENTAS, C. M. 1988 Dispersion in laminar tube flow at low Péclet numbers or short times. *AIChE J.* **34** (9), 1423–1430.

VRENTAS, J. S. & VRENTAS, C. M. 2000 Asymptotic solutions for laminar dispersion in circular tubes. *Chem. Engng Sci.* **55** (4), 849–855.

WANG, P. & CHEN, G. Q. 2016 Transverse concentration distribution in Taylor dispersion: Gill's method of series expansion supported by concentration moments. *Intl J. Heat Mass Transfer* **95**, 131–141.

WANG, P., LI, Z., WU, X. & AN, Y. 2015 Taylor dispersion in a packed pipe with wall reaction: based on the method of Gill's series solution. *Intl J. Heat Mass Transfer* **91**, 89–97.

WATT, S. D. & ROBERTS, A. J. 1996 The construction of zonal models of dispersion in channels via matched centre manifolds. *J. Austral. Math. Soc.* B **38** (01), 101.

WEBER, C. F. 1981 Analysis and solution of the ill-posed inverse heat conduction problem. *Intl J. Heat Mass Transfer* **24** (11), 1783–1792.

WESTERTERP, K. R., DIL'MAN, V. V. & KRONBERG, A. E. 1995 Wave model for longitudinal dispersion: development of the model. *AIChE J.* **41** (9), 2013–2028.

WHITAKER, S. 1999 *Theory and Applications of Transport in Porous Media: The Method of Volume Averaging*. Kluwer Academic.

WOOD, B. D. 2009 The role of scaling laws in upscaling. *Adv. Water Resour.* **32** (5), 723–736.

WOOD, B. D. & VALDÉS-PARADA, F. J. 2013 Volume averaging: local and nonlocal closures using a Green's function approach. *Adv. Water Resour.* **51**, 139–167.

WU, Z. & CHEN, G. Q. 2014 Approach to transverse uniformity of concentration distribution of a solute in a solvent flowing along a straight pipe. *J. Fluid Mech.* **740**, 196–213.

YIP, K. M.-K. 1996 Model simplification by asymptotic order of magnitude reasoning. *Artif. Intell.* **80** (2), 309–348.

YU, J. S. 1976 An approximate analysis of laminar dispersion in circular tubes. *J. Appl. Mech.* **43** (4), 537–542.

YU, J. S. 1979 On laminar dispersion for flow through round tubes. *J. Appl. Mech.* **46** (4), 750–756.

YU, J. S. 1980 Author's closure: 'On laminar dispersion for flow through round tubes'. *J. Appl. Mech.* **47**, 976–977.