J. Fluid Mech. (2012), *vol.* 700, *pp.* 459–489.
doi:10.1017/jfm.2012.143

The pollution of pristine material in compressible turbulence

Liubin Pan[1]†, Evan Scannapieco[1] and John Scalo[2]

[1] School of Earth and Space Exploration, Arizona State University, PO Box 871404, Tempe, AZ 85287, USA

[2] Department of Astronomy, University of Texas, Austin, TX 78712, USA

(Received 1 October 2011; revised 6 January 2012; accepted 10 March 2012; first published online 1 May 2012)

The first generation of stars had very different properties than later stellar generations, as they formed from a 'pristine' gas that was completely free of heavy elements. Normal star formation took place only after the first stars had polluted the surrounding turbulent interstellar gas, increasing its local heavy-element mass concentration, Z, beyond a 'critical' threshold value, Z_c ($10^{-8} \lesssim Z_c \lesssim 10^{-5}$). Motivated by this astrophysical problem, we investigate the fundamental physics of the pollution of pristine fluid elements in statistically homogeneous and isotropic compressible turbulence. Turbulence stretches the pollutants, produces concentration structures at small scales, and brings the pollutants and the unpolluted flow in closer contact. The pristine material is polluted when exposed to the pollutant sources or the fluid elements polluted by previous mixing events. Our theoretical approach employs the probability distribution function (p.d.f.) method for turbulent mixing, as the fraction of pristine mass corresponds to the low tail of the density-weighted concentration p.d.f. We adopt a number of p.d.f. closure models and derive evolution equations for the pristine fraction from the models. To test and constrain the prediction of theoretical models, we conduct numerical simulations for decaying passive scalars in isothermal turbulent flows with Mach numbers of 0.9 and 6.2, and compute the mass fraction, $P(Z_c, t)$, of the flow with $Z \leqslant Z_c$. In the Mach 0.9 flow, the evolution of $P(Z_c, t)$ is well-described by a continuous convolution model and goes as $\dot{P}(Z_c, t) = P(Z_c, t) \ln[P(Z_c, t)]/\tau_{con}$, if the mass fraction of the polluted flow is larger than $\approx$0.1. If the initial pollutant fraction is smaller than $\approx$0.1, an early phase exists during which the pristine fraction follows an equation derived from a nonlinear integral model: $\dot{P}(Z_c, t) = P(Z_c, t)[P(Z_c, t) - 1]/\tau_{int}$. The time scales τ_{con} and τ_{int} are measured from our simulations. When normalized to the flow dynamical time, the decay of $P(Z_c, t)$ in the Mach 6.2 flow is slower than at Mach 0.9 because the time scale for scalar variance decay is slightly larger and the low tail of the concentration p.d.f. broadens with increasing Mach number. We show that $P(Z_c, t)$ in the Mach 6.2 flow can be well fitted using a formula from a generalized version of the self-convolution model.

Key words: compressible turbulence, turbulence modelling, turbulent mixing

† Email address for correspondence: liubin.pan@asu.edu

1. Introduction

Big bang nucleosynthesis produced helium efficiently, but it was halted by the expansion of the universe before it was able to make stable elements heavier than lithium (Walker *et al.* 1991). On the other hand, even the most pristine stars observed today (Cayrel *et al.* 2004; Frebel *et al.* 2008; Caffau *et al.* 2011) have substantial mass fractions of heavier elements, indicating that they have been polluted with the nucleosynthesis products of an as-yet-undetected first generation of stars. This early stellar generation had an enormous impact on the evolution of later forming stars, and such stars are likely to have been much more massive (Abel, Bryan & Norman 2000; Bromm, Coppi & Larson 2002) and much hotter (Schaerer 2002) than present-day stars, due to the important role that heavy elements play in star formation and evolution.

When and where this remarkable early stellar generation formed is a question of fundamental astrophysical importance. On cosmological scales, the key issue is the time it takes for heavy elements to propagate from one galaxy to another. As shown in Scannapieco, Schneider & Ferrara (2003), the distances between these regions of early star formation are so vast that the universe was divided into two regions: one in which galaxies formed out of material that was already polluted with heavy elements and one in which galaxies were formed from initially pristine material.

This second set of initially pristine galaxies is especially interesting, as the first stars formed in these galaxies may be observable (Scannapieco *et al.* 2005; Jimenez & Haiman 2006; Nagao *et al.* 2008). As star formation continued in these objects, the interstellar gas became enriched with heavy elements released by the explosions of the first stars. This self-enrichment process increased the abundance or mass fraction, Z, of heavy elements, and finally led to a transition to normal star formation in regions where Z exceeds a critical value, Z_c. This critical value is expected to lie in the range $10^{-8} \lesssim Z_c \lesssim 10^{-5}$ (or 10^{-6}–10^{-3} times the heavy-element abundance in the Sun), depending on whether the cooling of the interstellar gas is dominated by dust grains (Omukai *et al.* 2005) or by the fine structures of carbon and oxygen (Bromm & Loeb 2003).

In a given galaxy, the key quantity to characterize the transition to normal star formation is the fraction, $P(Z_c, t)$, of the interstellar gas with Z below Z_c as a function of time. The temporal behaviour of this fraction depends not only on the rate at which new sources of heavy elements are released to the interstellar gas, but, more importantly, on the transport and mixing process of these elements in the galaxy (Pan & Scalo 2007). For example, a high mixing efficiency would result in a rapid decrease in $P(Z_c, t)$, and hence in a sharp transition as the average concentration of heavy elements exceeds the threshold Z_c. On the other hand, a low mixing efficiency would lead to a gradual transition. The interstellar gas in these galaxies is expected to be turbulent and highly compressible, and the turbulent motions are likely to be supersonic (Greif *et al.* 2008; Wise, Turk & Abel 2008). Therefore, understanding mixing in supersonic turbulence is crucial to answering the question of how the pristine gas in early galaxies was polluted.

In the present paper, we do not intend to directly model the complicated mixing process in a realistic galactic environment. Instead, we investigate the fundamental physics of turbulent mixing in compressible flows using idealized analytical and numerical tools. The primary goal is to understand the pollution of pristine material in statistically homogeneous and isotropic turbulence. This underlying physics is prerequisite for modelling the mixing of primordial gas in realistic interstellar turbulence. In a future work, we will apply the results of the current study to build

a subgrid model for large-scale simulations of the formation and evolution of early galaxies. These simulations account for the complexities in the interstellar medium, but cannot resolve the scales at which true mixing occurs. The subgrid model will provide a crucial step toward predicting the transition from primordial to normal star formation in the first generation of galaxies.

A systematic numerical study of passive scalar physics in supersonic turbulence has been recently conducted by Pan & Scannapieco (2010, 2011), who simulated scalar evolution in six compressible turbulent flows with Mach number ranging from 1 to 6. In those papers, a detailed analysis of various statistical measures for the scalar field was performed, including the scalar dissipation, the scalar probability distribution, the power spectrum, the structure functions and intermittency. It was found that the classic cascade picture for passive scalars in incompressible turbulence is generally valid also for mixing in supersonic turbulent flows. The effect of compressible modes in supersonic turbulence and their modifications to the classic picture for passive scalar turbulence were examined by analysing the Mach number dependence of the scalar statistics. The conclusions of those studies provide general theoretical guidelines for understanding the mixing process in interstellar turbulence.

To explore how the pollutant-free mass is contaminated in turbulent flows, we make use of the probability distribution method for turbulent mixing. The fraction of unpolluted or slightly polluted flow mass corresponds to the far-left tail of the probability distribution function (p.d.f.) of the concentration field, as Z_c is typically much smaller than the average value. This fraction can be evaluated by integrating the concentration p.d.f. from zero to the threshold, Z_c. We will generally refer to the fraction $P(Z_c, t)$ as the pristine fraction. Note that our approach here is general, and is not limited to mixing in early galaxies.

The p.d.f. equation for passive scalars cannot be solved exactly because of the closure problem, and various closure approximations have been developed to model the p.d.f. evolution. In this work, we consider several existing closure models and derive equations for the fraction $P(Z_c, t)$ for each of them. The far-left p.d.f. tail corresponds to high-order moments of the p.d.f., and thus it is quite uncertain whether the closure models can capture the high-order statistics with sufficient accuracy. In order to test the reliability of the adopted models and constrain their parameters, we perform numerical simulations for turbulent mixing in a transonic flow and a highly supersonic flow.

The structure of this paper is as follows. In § 2, we present the general p.d.f. formulation for mixing in compressible turbulence. Section 3 gives a brief description of several existing closure models for the diffusivity term in the p.d.f. equation. The predictions of these models for the mass fraction of unpolluted or slightly polluted flow are derived in § 4. We describe our numerical simulations in § 5, which are used to test and constrain the theoretical models in § 6. Our main conclusions are summarized in § 7.

2. The p.d.f. formulation for mixing in compressible flows

The p.d.f. formulation was first developed for the probability distribution of the turbulent velocity field by Lundgren (1967) and Monin (1967), and for the p.d.f. of the flow vorticity by Novikov (1967). The derivation of Monin (1967) was based on the equation for characteristic functions of the velocity field, while Lundgren (1967) started directly from the conservation laws of the flow. The two methods were later extended to derive p.d.f. equations for scalar fields convected by a turbulent flow, such as the flow temperature or enthalpy, and the concentration fields of passive or reactive

species in the flow (Ievlev 1973; Dopazo & O'Brien 1974; Pope 1976; Obrien 1980; Pope 1985; Kollmann 1990; Dopazo, Valino & Fueyo 1997). Recent discussions of p.d.f. equations for passive or active scalar turbulence can be found in the monograph by Fox (2003) and the thorough reviews by Veynante & Vervisch (2002) and Haworth (2010).

In the Appendix, we derive the p.d.f. equation for a passive or active scalar in compressible turbulence using the method of Lundgren (1967). The derivation is based on the equation of the concentration field, $C(\boldsymbol{x}, t)$, for passive tracers advected in a turbulent flow with density $\rho(\boldsymbol{x}, t)$ and velocity $\boldsymbol{v}(\boldsymbol{x}, t)$:

$$\frac{\partial C}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} C = \frac{1}{\rho} \boldsymbol{\nabla} \cdot (\rho \kappa \boldsymbol{\nabla} C) + S(\boldsymbol{x}, t), \tag{2.1}$$

where the concentration field is defined as the ratio of the local tracer density to the flow density. In the diffusion term, κ denotes the kinematic molecular diffusivity, and the dynamic diffusivity, $\rho\kappa$, is basically independent of ρ (i.e. $\kappa \propto \rho^{-1}$). The term $S(\boldsymbol{x}, t)$ represents the sources of new pollutants.

Our derivation in the Appendix adopts a density-weighting scheme, which is appropriate for passive scalar mixing in compressible flows (Pan & Scannapieco 2010). We define a density-weighted concentration p.d.f., $p(Z; \boldsymbol{x}, t) \equiv \langle \tilde{\rho} \delta[Z - C(\boldsymbol{x}, t)] \rangle$, where $\langle \cdots \rangle$ denotes the ensemble average, the density-weighting factor $\tilde{\rho} \equiv \rho(\boldsymbol{x}, t)/\bar{\rho}$ is the ratio of the local flow density to the average density $\bar{\rho}$, and Z is the sampling variable. Using the advection–diffusion equation (2.1), and the continuity equation for the evolution of the density-weighting factor, we obtain

$$\begin{aligned} &\frac{\partial p(Z; \boldsymbol{x}, t)}{\partial t} + \boldsymbol{\nabla} \cdot \left(p \frac{\langle \rho \boldsymbol{v} | C = Z \rangle}{\langle \rho | C = Z \rangle} \right) \\ &\quad = -\frac{\partial}{\partial Z} \left(p \frac{\langle \boldsymbol{\nabla} \cdot (\rho \kappa \boldsymbol{\nabla} C) | C = Z \rangle}{\langle \rho | C = Z \rangle} \right) - \frac{\partial}{\partial Z} \left(p \frac{\langle \rho S | C = Z \rangle}{\langle \rho | C = Z \rangle} \right), \end{aligned} \tag{2.2}$$

where $\langle \cdots | C = Z \rangle$ denotes the ensemble average under the condition that the concentration field $C(\boldsymbol{x}, t)$ is equal to Z (see the Appendix). The equation is essentially a Liouville equation for the conservation of the concentration probability. To our knowledge, this equation for the scalar p.d.f. with density weighting has not been derived before.

The density-weighting scheme is preferred in our study for two reasons. First, rather than the volume fraction, we are interested in the mass fraction of pristine gas in early galaxies, which corresponds to the left tail of the density-weighted p.d.f. Second, the advection term in the equation for the density-weighted p.d.f. takes the form of a divergence, and thus conserves the global p.d.f. (i.e. the integral of the local p.d.f., $p(Z; \boldsymbol{x}, t)$, over the flow domain). This provides a formal and rigorous proof for the physical intuition that the turbulent velocity field itself does not homogenize the distribution of pollutants. The advecting velocity transports, redistributes and deforms the concentration field, but does not change the *mass* fraction of fluid elements with a given concentration level. Furthermore, the advection term vanishes if the flow and the concentration fluctuations are statistically homogeneous.

In contrast, if one derives an equation for the volume-weighted p.d.f. for a passive scalar in compressible turbulence, the advection term would not be a divergence term. The term reflects the effect of flow compressions and expansions, which can change the volume fraction of fluid elements at a given concentration (Pan & Scannapieco 2010). This effect on the p.d.f. is clearly different from scalar homogenization, and can be avoided by adopting a density-weighting factor. We thus argue that it is more

appropriate to use the density-weighted p.d.f. equation for the study of mixing in compressible turbulence.

Molecular diffusion is the only process that homogenizes, and the molecular diffusivity term in the p.d.f. equation continuously reduces the p.d.f. width. This term can be rewritten as

$$-\frac{\partial}{\partial Z}\left(p\frac{\langle \boldsymbol{\nabla}\cdot(\rho\kappa\boldsymbol{\nabla}C)|C=Z\rangle}{\langle\rho|C=Z\rangle}\right)$$
$$=-\boldsymbol{\nabla}\cdot\left[\frac{\partial}{\partial Z}\left(p\frac{\langle\rho\kappa\boldsymbol{\nabla}C|C=Z\rangle}{\langle\rho|C=Z\rangle}\right)\right]-\frac{\partial^2}{\partial Z^2}\left(p\frac{\langle\rho\kappa\,(\boldsymbol{\nabla}C)^2\,|C=Z\rangle}{\langle\rho|C=Z\rangle}\right), \tag{2.3}$$

where both terms on the right-hand side depend on the ensemble average of the concentration gradients conditioned on $C(\boldsymbol{x},t)=Z$. As it is a divergence term, the first term in (2.3) conserves the global concentration p.d.f. The scalar homogenization is achieved through the second term, which is essentially a diffusion term with a negative coefficient in concentration space. This term keeps narrowing the concentration p.d.f., and the physics of turbulent mixing can be viewed as an anti-diffusion process in concentration space.

Taking the second-order moment of the last term in (2.3) gives the scalar dissipation rate, $-2\langle\tilde{\rho}\kappa\,(\boldsymbol{\nabla}C)^2\rangle$. Using this rate, we define a mixing time scale,

$$\tau_m\equiv\frac{\langle\tilde{\rho}\delta C^2\rangle}{2\langle\tilde{\rho}\kappa\,(\boldsymbol{\nabla}C)^2\rangle}, \tag{2.4}$$

where $\delta C=C-\langle\tilde{\rho}C\rangle$ is the fluctuating part of the concentration field. The time scale, τ_m, corresponds to the scalar variance decay by mixing, and thus characterizes the rate at which the diffusivity term reduces the p.d.f. width. Although the diffusivity terms in (2.3) and (2.4) do not have an explicit dependence on the flow velocity, the mixing time scale is determined primarily by the turbulent velocity field. This is because the turbulent velocity produces progressively smaller structures and thus strongly amplifies the scalar gradients, $(\boldsymbol{\nabla}C)^2$, in (2.3) and (2.4). By feeding molecular diffusivity with large-gradient structures, turbulent motions greatly accelerate the scalar dissipation/homogenization.

In the classic phenomenology for mixing in incompressible turbulence, the generation of small-scale concentration structures is through a cascade process similar to that of kinetic energy (Obukhov 1949; Corrsin 1951). The cascade is caused by continuous turbulent stretching, and it starts from the scale where the pollutant sources are injected into the flow, and proceeds to the diffusion scale where the molecular diffusion efficiently homogenizes the scalar fluctuations. The diffusion scale is essentially the scale where the action of molecular diffusivity becomes faster than turbulent stretching. From this picture, the mixing time scale, τ_m, is determined by the cascade time, which is essentially the eddy turnover time at the injection scale of the scalar sources because the cascade becomes faster and faster with decreasing length scale.

Pan & Scannapieco (2010) showed that the cascade picture also applies for mixing in supersonic turbulence. They found that the mixing time scale, τ_m, was close to the eddy turnover time at the pollutant injection scale in all their simulated flows with Mach numbers in the range from 1 to 6. The existence of compressible modes in supersonic flows causes only a slight Mach number dependence of the mixing time scale, and the primary 'mixer' is the stretching by solenoidal modes even at very high Mach numbers.

Translating the physical discussion above to the mixing process of the unpolluted fluid elements in a turbulent flow gives the following picture. The turbulent velocity stretches the pollutants into smaller and smaller structures, and brings them into closer contact with the unpolluted flow. When the separation between the pollutant structures and the unpolluted fluid elements becomes close to or smaller than the diffusion scale, molecular diffusivity efficiently mixes them, reducing the unpolluted mass fraction. This suggests that the time scale for turbulent mixing to contaminate the unpolluted mass is also of the order of the scalar cascade time scale.

The diffusivity term in the p.d.f. equation has to be approximately modelled because of the closure problem (e.g. Dopazo & O'Brien 1974). Extensive efforts have been made to develop closure models for this term, and we will use several existing models in the current study, as described in § 3. The advection term also has a closure problem, but modelling this term is not necessary if the flow and the scalar field are statistically homogeneous. The last term in the p.d.f. equation (2.2) corresponds to the effect of the pollutant sources. In reacting turbulent flows, the source term due to chemical reactions has a closed form in the p.d.f. formulation (e.g. Pope 1976), and this has led to the wide use of the p.d.f. method in studies of chemical reactions in turbulent flows. In the present study, the pollutant source is merely an initial scalar condition, and we do not discuss modelling the source term further (see Pan & Scalo 2007, for an example with a persistent source).

3. The p.d.f. modelling

3.1. *General approach*

We employ both theoretical and numerical tools in the present work. The p.d.f. formulation in § 2 provides a general theoretical framework, and, to complete the theoretical approach, we will consider several existing closure models for the diffusivity term in (2.2). We will compare the predictions of these models for the scalar p.d.f. evolution with that measured from numerical simulations. From our simulation data, we compute the concentration p.d.f. as $p(Z;t) = (1/V)\int_V \tilde{\rho}\delta[Z - C(\boldsymbol{x},t)]\,\mathrm{d}\boldsymbol{x}$, where V is the total volume of the simulation box. This p.d.f. measures the concentration fluctuations over the entire flow domain, and thus should be viewed as a global p.d.f. As pointed out in § 2, the advection term conserves the global density-weighted p.d.f., and thus need not be considered in our tests of the theoretical models against simulations. The global p.d.f. is expected to be equal to the local p.d.f., $p(Z;\boldsymbol{x},t)$, defined in the ensemble context under the assumption of statistical homogeneity.

In our simulations, we only evolve decaying scalars with the source term $S(\boldsymbol{x},t)$ set to be zero. Neglecting the advection and source terms, the p.d.f. equation becomes

$$\frac{\partial p(Z;t)}{\partial t} = -\frac{\partial}{\partial Z}\left(p\frac{\langle\nabla\cdot(\rho\kappa\nabla C)|C=Z\rangle}{\langle\rho|C=Z\rangle}\right). \tag{3.1}$$

The only term that contributes to the p.d.f. evolution in our simulations is the diffusivity term, and modelling this term is the main task of the p.d.f. approach to turbulent mixing. In incompressible turbulence, the flow density is constant, and (3.1) reduces to

$$\frac{\partial p(Z;t)}{\partial t} = -\frac{\partial}{\partial Z}\left(p\langle(\kappa\nabla^2 C)|C=Z\rangle\right), \tag{3.2}$$

which has been extensively studied and modelled.

The second-order moment of equation (3.1) corresponds to the scalar variance equation,

$$\frac{\mathrm{d}\langle \delta Z^2 \rangle}{\mathrm{d}t} = -\frac{\langle \delta Z^2 \rangle}{\tau_m}, \tag{3.3}$$

where $\langle \delta Z^2 \rangle \equiv \langle \tilde{\rho} \delta C^2 \rangle$ denotes the density-weighted variance, and we have used the definition, (2.4), of the mixing time scale. In terms of the p.d.f., the variance is given by $\langle \delta Z^2 \rangle = \int (Z - \tilde{Z})^2 \, p(Z, t) \, \mathrm{d}Z$ with $\tilde{Z} = \int Z p(Z; t) \, \mathrm{d}Z$ being the mean concentration. In general, τ_m may be a function of time. But if it is constant, the scalar variance decreases exponentially, which is the case at the late evolution stage of a decaying scalar (see § 6.3).

In analogy to the enrichment of pristine gas by the first generation of stars, the initial condition of the decaying scalars in this study will be set to be bimodal: consisting of pure pollutants ($Z = 1$) and completely unpolluted flow ($Z = 0$). This corresponds to a double delta-function form for the initial concentration p.d.f.,

$$p(Z; 0) = P_0 \delta(Z) + P_1 \delta(Z - 1), \tag{3.4}$$

where P_1 and P_0 are the initial mass fractions of the pollutants and the unpolluted flow, respectively, and we have $P_0 + P_1 = 1$ from the normalization of the p.d.f.

The rest of this section is devoted to modelling the diffusivity term in the p.d.f. equation. A variety of closure models have been proposed for this term, and the interested reader is referred to Dopazo *et al.* (1997) and Haworth (2010) for reviews. Here we will consider three of the models proposed for the diffusivity closure: the mapping closure model by Chen, Chen & Kraichnan (1989); the nonlinear integral models by Curl (1963), Dopazo (1979) and Janicka, Kolbe & Kollmann (1979); and the self-convolution models by Villermaux & Duplat (2003), Venaille & Sommeria (2007) and Duplat & Villermaux (2008).

We point out that, in compressible turbulence, the diffusivity term has an explicit dependence on the density field, or more precisely, on the joint statistics of the density and the concentration fields. Therefore, an ideal p.d.f. model for mixing in supersonic turbulence needs to account for the effect of density fluctuations on the diffusivity term, and to predict the dependence of the concentration p.d.f. on the flow compressibility. However, to our knowledge, this has not been considered in existing models, which were usually tested against simulation results for mixing in incompressible turbulence. We will compare the predictions of the closure models mentioned above with our simulation data, and examine whether, by adjusting their parameters, these models can be successfully applied to the contamination process of pollutant-free mass in compressible turbulent flows at different Mach numbers. Future studies are motivated to develop closure models that explicitly address the effects of shocks and the Mach number dependence of the passive scalar p.d.f. in supersonic turbulence.

3.2. *The mapping closure model*

We first discuss the mapping closure model developed by Chen *et al.* (1989) for mixing in incompressible turbulence. We give a brief introduction to the model, and a detailed derivation can be found in e.g. Pope (1991). The model is based on a surrogate field, $\phi(\boldsymbol{x}, t)$, obtained from the mapping of a Gaussian reference field $\theta(\boldsymbol{x}, t)$,

$$\phi(\boldsymbol{x}, t) = X[\theta(\boldsymbol{x}, t), t], \tag{3.5}$$

where X is an ordinary, non-stochastic function. The Gaussian field, $\theta(\boldsymbol{x}, t)$, is assumed to be statistically homogeneous and have zero mean and unit variance, i.e. the probability of finding $\theta(\boldsymbol{x}, t)$ equal to a given value η is given by $g(\eta) \equiv (1/\sqrt{2\pi})\exp(-\eta^2/2)$ (see Girimaji 1992 for a generalized version of the mapping closure where the reference field p.d.f. is time-evolving and not limited to Gaussian). The main idea of the model is to pursue a mapping function, $X(\eta, t)$, with which the p.d.f. of the surrogate field obeys exactly the same equation (i.e. (3.2)) as the actual field, $C(\boldsymbol{x}, t)$. This is indeed achieved if the mapping function evolves as

$$\frac{\partial X(\eta, t)}{\partial t} = \frac{\kappa}{\lambda_\theta^2(t)}\left(\frac{\partial^2 X(\eta, t)}{\partial \eta^2} - \eta\frac{\partial X(\eta, t)}{\partial \eta}\right), \tag{3.6}$$

where $\lambda_\theta(t) = \langle(\boldsymbol{\nabla}\theta)^2\rangle^{-1/2}$, and $\lambda_\theta^2(t)/\kappa$ is a time scale that controls the rate at which the mapping function and hence the p.d.f. evolve. The time scale is unspecified in the original model, and can be calibrated by a comparison of the variance decay in the model with simulation results (see He & Zhang 2004 for a theoretical evaluation of this time scale using a two-point closure strategy). The evolution equation for the mapping function was derived in the incompressible limit, and thus the model is intended for mixing in incompressible turbulence only. By a comparison with our simulation data, we will examine whether the mapping closure model may also give acceptable predictions for mixing in compressible turbulence.

With the desired mapping function, one can approximate the p.d.f. of the actual field by that of the surrogate field. Using (3.6), the p.d.f. of the surrogate field can be converted from the Gaussian p.d.f. of the reference field. The conversion gives

$$p(Z; t) = g(\eta)\left(\frac{\partial X(\eta, t)}{\partial \eta}\right)^{-1}, \tag{3.7}$$

where η is the solution of $X(\eta, t) = Z$.

The linear equation for the mapping function, (3.6), can be solved analytically, given the initial condition $X(\eta, 0)$ (Pope 1991). For a double-delta initial p.d.f., the initial mapping is a Heaviside step function $X(\eta, 0) = H(\eta - \eta_0)$, where η_0 satisfies $\int_{-\infty}^{\eta_0} g(\eta)\,\mathrm{d}\eta = P_0$. With this initial condition, $X(\eta, t)$ is solved by

$$X(\eta, t) = G\left(\frac{\eta}{\Sigma(t)} - \frac{\eta_0\left(\Sigma(t)^2 + 1\right)^{1/2}}{\Sigma(t)}\right), \tag{3.8}$$

where $\Sigma(t)^2 = \exp[\int_0^t \kappa/\lambda_\theta^2(t')\,\mathrm{d}t'] - 1$, and $G(\eta) \equiv \int_{-\infty}^{\eta} g(\eta')\,\mathrm{d}\eta'$ is the cumulative function of the Gaussian function. Combining (3.7) and (3.8) gives the predicted p.d.f. evolution by the mapping closure model.

3.3. *The nonlinear integral models*

In this subsection, we consider a class of closure models that use an integral form to approximate the diffusivity term in the p.d.f. equation. This type of models originates from the equation introduced by Curl (1963),

$$\frac{\partial p(Z; t)}{\partial t} = \gamma(t)\left\{\left[\int_0^1 \mathrm{d}Z_1 p(Z_1; t)\int_0^1 \mathrm{d}Z_2 p(Z_2; t)\delta\left(Z - \frac{Z_1 + Z_2}{2}\right)\right] - p(Z; t)\right\}, \tag{3.9}$$

where $\gamma(t)$ is the turbulent stretching rate. A physical interpretation of this equation is as follows (e.g. Dopazo 1979). The turbulent velocity field stretches the concentration

field and produces structures at small scales. As shown by various studies, these structures are primarily in the form of two-dimensional sheets (e.g. Pan & Scannapieco 2011 and references therein). The scalar sheets are brought closer to each other over time by the turbulent velocity. When the typical width and separation of the sheets decrease to the diffusion scale, molecular diffusivity can operate efficiently and homogenize. The time scale for this process is $\gamma(t)^{-1}$, which is expected to be of the order of the scalar cascade time scale or the mixing time scale τ_m. Two scalar sheets of different concentrations brought to close contact are assumed to mix perfectly, resulting in a concentration value equal to their average prior to the mixing event (see the delta function in (3.9)). The last term in (3.9) corresponds to the 'destruction' of the previous p.d.f. by the mixing event.

One problem of Curl's model for turbulent mixing is that, if the initial concentration p.d.f. consists of two delta functions (see (3.4)), the predicted p.d.f. shows unphysical spikes in between the initial delta functions. To avoid this problem, Dopazo (1979) and Janicka *et al.* (1979) independently generalized Curl's model by replacing the delta function in (3.9) by a smooth function $J(Z; Z_1, Z_2)$,

$$\frac{\partial p(Z;t)}{\partial t} = \gamma(t)\left\{\left[\int_0^1 p(Z_1;t)\int_0^1 p(Z_2;t)J(Z;Z_1,Z_2)\,\mathrm{d}Z_1\,\mathrm{d}Z_2\right] - p(Z;t)\right\}, \quad (3.10)$$

where $J(Z; Z_1, Z_2)$ represents the effect of mixing between two nearby scalar sheets with concentration values of Z_1 and Z_2. The function $J(Z; Z_1, Z_2)$ is zero for Z outside the range (Z_1, Z_2) (or (Z_2, Z_1) if $Z_1 > Z_2$), and its normalization is $\int_{Z_1}^{Z_2} J(Z; Z_1, Z_2)\,\mathrm{d}Z = 1$. A simple assumption for $J(Z; Z_1, Z_2)$ is that it is uniform between Z_1 and Z_2, leading to

$$\frac{\partial p(Z;t)}{\partial t} = \gamma(t)\left\{\left[\int_0^Z p(Z_1;t)\int_Z^1 \frac{2}{Z_2 - Z_1}p(Z_2;t)\,\mathrm{d}Z_1\,\mathrm{d}Z_2\right] - p(Z;t)\right\}, \quad (3.11)$$

where $J(Z; Z_1, Z_2)$ was set to $1/|Z_2 - Z_1|$ (Dopazo 1979; Janicka *et al.* 1979).

The parameter $\gamma(t)$ as a function of time can be fixed by comparing the variance equation of these models with the simulation data. The derivation of the variance equation can be found in Janicka *et al.* (1979) or Valino & Dopazo (1990). For Curl's model and the model with uniform $J(Z; Z_1, Z_2)$, the variance decays $\propto \exp[-(1/2)\int_0^t \gamma(t')\,\mathrm{d}t']$ and $\propto \exp[-(1/3)\int_0^t \gamma(t')\,\mathrm{d}t']$, respectively. If the variance decreases exponentially with a constant time scale τ_m, $\gamma(t)$ is constant and equal to $2/\tau_m$ and $3/\tau_m$, respectively, for the two models.

Pope (1982) pointed out a weakness of this class of models: the normalized high-order moments, $\langle \delta Z^m \rangle / \langle \delta Z^2 \rangle^{m/2}$, do not converge with time for $m \geqslant 4$ (see also Valino & Dopazo 1990). This suggests that the predicted p.d.f. by these models has excessively fat tails at late times (Kollmann 1990).

3.4. *The self-convolution models*

The last type of models we consider are those based on the self-convolution of the scalar p.d.f., which can be viewed as extensions of the model by Curl (1963) in Laplace space. A review of the development of these models can found in Duplat & Villermaux (2008).

The Laplace transform $\hat{p}(\zeta; t)$ of the scalar p.d.f. is defined as $\hat{p}(\zeta; t) = \int_0^\infty p(Z; t)\exp(-Z\zeta)\,\mathrm{d}Z$. Using the convolution theorem, the Laplace transform of (3.9)

gives

$$\frac{\partial \hat{p}(\zeta;t)}{\partial t} = \gamma \left[\hat{p}\,(\zeta/2;t)^2 - \hat{p}(\zeta;t)\right]. \tag{3.12}$$

A similar equation in Fourier space was used by Pumir, Shraiman & Siggia (1991). This equation shows that turbulent mixing is essentially treated as a self-convolution process in Curl's model. We rewrite (3.12) in a difference form $\hat{p}(\zeta;t+\delta t) = \epsilon \hat{p}\,(\zeta/2;t)^2 + (1-\epsilon)\hat{p}(\zeta;t)$ where $\epsilon = \gamma \delta t$ with δt an infinitesimal time step. The difference equation can be interpreted as follows: during a time step δt, mixing occurs only in an infinitesimal fraction, ϵ, of the flow, and in this part of the flow the scalar p.d.f. undergoes a complete convolution. The convolution process in (3.12) appears to be 'discrete'.

Following Venaille & Sommeria (2007), we derive a continuous version of Curl's model. We assume that, in each time step δt, the p.d.f. convolution occurs everywhere in the flow, but the number of convolutions is taken to be infinitesimal and equal to ϵ (Duplat & Villermaux 2008). This assumption can be written as $\hat{p}(\zeta;t+\delta t) = \hat{p}\,(\zeta/(1+\epsilon);t)^{(1+\epsilon)}$. Using the Taylor expansion $\hat{p}\,(\zeta/(1+\epsilon);t)^{(1+\epsilon)} \simeq \hat{p}(\zeta;t) + \epsilon[\hat{p}(\zeta;t)\ln(\hat{p}(\zeta;t)) - \zeta\,\partial\hat{p}(\zeta;t)/\partial\zeta]$ and taking the limit $\delta t \to 0$, we obtain

$$\frac{\partial \hat{p}(\zeta;t)}{\partial t} = \gamma \left[\hat{p}\ln(\hat{p}) - \zeta \frac{\partial \hat{p}}{\partial \zeta}\right], \tag{3.13}$$

which represents the model of Venaille & Sommeria (2007). We will refer to this model as the continuous convolution model. In this model, the variance decays $\propto \exp(-\int_0^t \gamma(t')\,\mathrm{d}t')$. Venaille & Sommeria (2007) showed that the p.d.f. predicted by (3.13) evolves toward Gaussian in the long-time limit (in contrast to the integral models in § 3.3). A comparison of this model with experimental data is given in Venaille & Sommeria (2008). We note that, if the initial p.d.f. is two delta functions, the continuous self-convolution model is not applicable for the p.d.f. evolution right from the beginning (Venaille & Sommeria 2007). We thus cannot compare the model prediction for $p(Z;t)$ with our simulation results at the early evolution stage. The model will only be used to study the evolution of the unpolluted mass fraction.

As pointed out by Duplat & Villermaux (2008), a more general extension of Curl's model is

$$\frac{\partial \hat{p}(\zeta;t)}{\partial t} = \gamma n \left[\hat{p}\left(\frac{\zeta}{1+1/n};t\right)^{(1+1/n)} - \hat{p}(\zeta;t)\right]. \tag{3.14}$$

Curl's original model (equation (3.12)) and the model of Venaille & Sommeria (2007) (equation (3.13)) are special cases of (3.14) with $n=1$ and $n\to\infty$, respectively. The parameter n can be a function of time in general. The assumption behind (3.14) is that a fraction, $n\epsilon$, of the flow experiences mixing/convolution events during a time step δt, and the number of convolutions in this fraction of the flow is $1/n$. For (3.14), the variance decay is $\propto \exp(-\int_0^t \gamma(t')n/(n+1)\,\mathrm{d}t')$. We will refer to (3.14) as the generalized convolution model.

We finally consider the model by Villermaux & Duplat (2003), which was motivated by a turbulent mixing picture with three related processes: the generation of pollutant sheets by turbulent stretching; the diffusion of the pollutant sheets by molecular diffusivity; and the merging of the diffused sheets. The merging of the sheets corresponds to a self-convolution process. The model is represented by (Duplat &

Villermaux 2008)

$$\left.\begin{aligned}\frac{\partial \hat{p}(\zeta;t)}{\partial t} &= -\gamma\zeta\frac{\partial \hat{p}}{\partial \zeta} + n\gamma\left[\hat{p}\,(\zeta;t)^{(1+1/n)} - \hat{p}(\zeta;t)\right],\\ \frac{\partial n(t)}{\partial t} &= \gamma n,\end{aligned}\right\} \tag{3.15}$$

where n increases with time and the first equation can be viewed as the expansion of (3.14) at large n. Note that (3.13) and (3.15) approach the same limit as $t \to \infty$. Villermaux & Duplat (2003) showed that (3.15) has an asymptotic solution $\hat{p}(\zeta;t) = (1 + \langle Z\rangle\zeta/n)^{-n}$ at large t, which corresponds to a gamma distribution for the scalar p.d.f. (Duplat & Villermaux 2008),

$$p(Z;t) = \frac{n^n}{\Gamma(n)\langle Z\rangle^n} Z^{n-1}\exp\left(-\frac{nZ}{\langle Z\rangle}\right), \tag{3.16}$$

where $\Gamma(n)$ is the gamma function. The gamma distribution is valid only at late times with $n \gtrsim 1$, and cannot be applied to study the pristine mass fraction at the early evolution stage when the fraction is significant. Therefore, we do not use the model for the pristine mass fraction, but will check whether the scalar p.d.f. in our simulations approaches a gamma distribution at late times.

We point out a fundamental difference between the mapping closure model discussed in § 3.2 and the models presented here and in § 3.3. The mapping closure is established by a direct approximation of the exact, but unclosed, form of the diffusivity term. On the other hand, the nonlinear integral models and the convolution models do not start from the diffusivity term in the p.d.f. equation; instead they are largely based on a physical picture of the mixing process.

4. Mass fraction of unpolluted or slightly polluted flow

As mentioned in the Introduction, we are interested in the mass fraction, $P(Z_c, t)$, of the flow with concentration smaller than a tiny threshold, Z_c, which can be calculated from the concentration p.d.f. as

$$P(Z_c, t) = \int_0^{Z_c} p(Z';t)\,\mathrm{d}Z'. \tag{4.1}$$

The fraction corresponds to the far-left tail of the p.d.f. since the threshold Z_c of interest is typically much smaller than the average concentration. Taking the limit $Z_c \to 0$ in (4.1), we obtain the fraction $P(t)$ of exactly pollutant-free mass, i.e. $P(t) = \lim_{Z_c\to 0}\int_0^{Z_c} p(Z';t)\,\mathrm{d}Z'$. This fraction is zero unless $p(Z;t)$ has a delta-function component, $\delta(Z)$, at $Z = 0$. In this section, we derive equations for $P(Z_c, t)$ and $P(t)$ from the closure models discussed in § 3.

An interesting observation of the action of molecular diffusivity is that it tends to decrease the exactly pollutant-free fraction, $P(t)$, to zero instantaneously. For illustration, we consider a simple situation with a point source diffusing in a static uniform medium. The concentration field obeys the diffusion equation, whose solution is given by a Gaussian function. From this solution, it is clear that, no matter how small the molecular diffusivity, κ, is, the concentration field at a finite time ($t > 0$) becomes non-zero at any finite distance ($r < \infty$) from the initial source, suggesting that all the pollutant-free mass is removed from the system instantaneously.

This acausal behaviour of molecular diffusivity originates from the Laplacian operator in the diffusion equation, which implicitly assumes that the random walk

of some tracer molecules can bring them to an infinite distance during any small (but macroscopic) time interval. This is clearly unrealistic. The thermal motions of tracer molecules must have a finite maximum speed, $\max(v_{th})$, and thus none of them can reach an infinite distance instantaneously. If the size of the system in question is L, there could be exactly pollutant-free mass surviving for a finite time $\sim L/\max(v_{th})$. However, this time is expected to be very small since $\max(v_{th})$ is likely to be much larger than the sound speed. Therefore, the reduction of exactly pollutant-free fraction, $P(t)$, by molecular diffusion may be considered as being essentially instantaneous.

For our astrophysical applications, we need the fraction, $P(Z_c, t)$, of the flow with Z below a finite critical value, Z_c, rather than the exactly pristine fraction. Obviously, it takes finite time for molecular diffusivity to enrich all the fluid elements in the system up to a finite threshold, Z_c. In fact, during a short time interval, the degree of pollution by molecular diffusivity alone is negligible even at small distances from the pollutant source, and the entire system is practically unpolluted. Therefore, the observation of the rapid/immediate erasure of exactly pristine gas by molecular diffusivity is not directly relevant to the astrophysical problem of primordial star formation.

Because κ is usually tiny in practical environments, such as in the interstellar media of galaxies, enriching all the fluid elements to a concentration level of, say, $\gtrsim 10^{-8}$, by the molecular diffusivity alone is very slow (see discussions in Pan & Scalo 2007). The presence of a turbulent velocity field greatly accelerates the mixing process, making the reduction of $P(Z_c, t)$ much faster. We find that the time scale for the reduction of $P(Z_c, t)$ with a small Z_c is basically determined by the rate at which the turbulent stretching produces small-scale structures and is essentially independent of κ.

4.1. *The mapping closure model*

We calculate the fraction $P(Z_c, t)$ predicted by the mapping closure model. From (3.7), it is straightforward to find that

$$P(Z_c, t) = \int_{-\infty}^{\eta_c(t)} g(\eta)\, \mathrm{d}\eta = G(\eta_c(t)), \tag{4.2}$$

where the upper limit $\eta_c(t)$ satisfies $X(\eta_c(t); t) = Z_c$. For a given value of Z_c, the limit $\eta_c(t)$ changes with time as the mapping function evolves, and for our initial bimodal p.d.f. with two delta functions, $\eta_c(t)$ can be computed using (3.8).

From that equation, we see that $Z_c = 0$ corresponds to $\eta_c \to -\infty$ at all times after $t = 0$. Therefore, the mapping closure model predicts that $P(t)$ is zero at any time $t > 0$, or that the fraction of exactly pollutant-free mass deceases to zero instantaneously. This is consistent with our discussion above that the molecular diffusivity alone tends to immediately remove fluid elements with exactly zero concentration. The mapping closure model inherits this particular property of molecular diffusion, because the effect of diffusivity as a Laplacian term is treated directly. The model destroys the initial delta function at $Z = 0$ instantaneously. However, this does not suggest that $p(0; t)$ becomes finite immediately. At the early evolution stage, $p(Z; t)$ does have an infinite peak at $Z = 0$, but the peak is less singular than a delta function ($\delta(Z)$) in the sense that $\int_0^Z p(Z'; t)\, \mathrm{d}Z' \to 0$ in the limit $Z \to 0$.

4.2. *The nonlinear integral models*

Unlike the mapping closure model, the nonlinear integral models preserve the singularities at $Z = 0$ and $Z = 1$. More specifically, the amplitudes of the delta functions at $Z = 0$ and $Z = 1$ decrease with time, but they are never completely

destroyed, such that exactly pollutant-free mass can survive in these models, and $P(t)$ remains finite at any finite time. This is inconsistent with our earlier observation that the molecular diffusivity tends to reduce $P(t)$ to zero immediately. The reason is that the effect of molecular diffusivity is not incorporated directly in these models; instead it is included implicitly through the function $J(Z; Z_1, Z_2)$. Despite the inconsistency, we find that the integral models are useful for understanding the pollution of fluid elements with very low (but non-zero) concentration by turbulent mixing. Below we derive an equation for the fraction, $P(t)$, of exactly pollutant-free mass from these models.

We consider the general model represented by (3.10). Integrating this equation in the range $[0, Z]$ and taking the limit $Z \to 0$, we have

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = \gamma(t)\left(\int_0^1 \mathrm{d}Z_1\, p(Z_1; t)\int_0^1 \mathrm{d}Z_2 p(Z_2; t)\lim_{Z\to 0}\int_0^Z \mathrm{d}Z'\, J(Z'; Z_1, Z_2) - P(t)\right). \quad (4.3)$$

The last integral in the triple-integral term in the limit $Z \to 0$ can be written as $\int_0^{0^+} J(Z'; Z_1, Z_2)\,\mathrm{d}Z'$ where 0^+ represents the upper integral limit approaching zero from the positive vicinity. We first note that this integral is zero if both Z_1 and Z_2 are positive because $J(0; Z_1, Z_2) = 0$ for $Z_1 > 0$ and $Z_2 > 0$ (see § 3.3). We next assume that $J(Z; Z_1, Z_2)$ at $Z = Z_1$ and $Z = Z_2$ is non-singular or less singular than a delta function for $Z_1 \neq Z_2$ (meaning that $\int_{Z_1}^{Z_1^+} J(Z; Z_1, Z_2)\,\mathrm{d}Z = 0$ and $\int_{Z_2}^{Z_2} dJ(Z; Z_1, Z_2)\,\mathrm{d}Z = 0$, where $Z_1 < Z_2$ is assumed without loss of generality). This assumption is clearly satisfied for Curl's model and the model with uniform $J(Z; Z_1, Z_2)$ (equation (3.11)). With this assumption, it is straightforward to see that $\int_0^{0^+} J(Z'; Z_1, Z_2)\,\mathrm{d}Z'$ is finite only if both $Z_1 = 0$ and $Z_2 = 0$. In that case, we have $\int_0^{0^+} \mathrm{d}Z' J(Z'; 0, 0) = 1$ from the normalization of $J(Z; Z_1, Z_2)$. This observation suggests that the contribution to the triple integral in (4.3) comes only from Z_1 and Z_2 values in an infinitesimal range around zero. With such infinitesimal ranges of Z_1 and Z_2, the first two of the three integrals contribute factors of $\int_0^{0^+} p(Z_1; t)\,\mathrm{d}Z_1$ and $\int_0^{0^+} p(Z_2; t)\,\mathrm{d}Z_2$, respectively. As both these factors are equal to $P(t)$, we have the following equation for $P(t)$:

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = -\frac{P(1-P)}{\tau_{int}}, \quad (4.4)$$

where $\tau_{int} \equiv \gamma(t)^{-1}$ is used for the convenience of notation. If the mixing time scale τ_m is constant, the variance decay requirement gives $\tau_{int} = \tau_m/2$ or $\tau_m/3$ for Curl's model and the model with uniform $J(Z; Z_1, Z_2)$, respectively (see § 3.3).

Equation (4.4) gives an interesting physical picture of mixing of the unpolluted mass in turbulent flows: the pristine fraction is reduced when turbulent stretching brings the pollutant-free fluid elements (with a fraction of $P(t)$), and the rest of the flow (with a fraction of $1 - P(t)$) which has been polluted by sources or previous mixing events, close enough for molecular diffusivity to homogenize (Pan & Scalo 2007).

Equation (4.4) has a simple analytic solution,

$$P(t) = \frac{P_0}{P_0 + (1-P_0)\exp\left(\dfrac{t}{\tau_{int}}\right)}, \quad (4.5)$$

where P_0 is the initial fraction of unpolluted mass, and we have assumed τ_{int} is constant with time. Although it is derived for the fraction of exactly pollutant-free mass, we will show in § 5 that, in certain physical regimes, this equation can be used to fit our numerical results for $P(Z_c, t)$ with a finite threshold Z_c.

4.3. *The self-convolution models*

The self-convolution models introduced in § 3.4 also preserve the initial singularities at $Z = 0$ and $Z = 1$, since they are essentially extensions of Curl's model. Again we derive the equations for the fraction, $P(t)$, of exactly pollutant-free mass from the convolution models, which will be used later to understand the mass fraction of nearly pristine, but $Z \neq 0$, flow. We first decompose the concentration p.d.f. as

$$p(Z; t) = P(t)\delta(Z) + p_e(Z; t), \tag{4.6}$$

where $p_e(Z; t)$ denotes the concentration p.d.f. in the enriched/polluted part of the flow, and it satisfies $\lim_{Z\to 0} \int_0^Z p_e(Z'; t)\, \mathrm{d}Z' = 0$. The Laplace transform of (4.6) gives

$$\hat{p}(\zeta; t) = P(t) + \hat{p}_e(\zeta; t), \tag{4.7}$$

where $\hat{p}_e(\zeta; t)$ is the Laplace transform of $p_e(Z; t)$. In the limit $\zeta \to +\infty$, $\hat{p}_e(\zeta; t)$ approaches zero because $\lim_{Z\to 0} \int_0^Z p'_e(Z'; t)\, \mathrm{d}Z' = 0$.

Inserting (4.7) into (3.13) for the model of Venaille & Sommeria (2007), and taking the limit $\zeta \to +\infty$, we find

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = \frac{P \ln(P)}{\tau_{con}}, \tag{4.8}$$

where $\tau_{con} \equiv \gamma^{-1}$, and we used the fact that $\hat{p}_e(\zeta; t) \to 0$ and $\zeta\, \partial_\zeta \hat{p}_e(\zeta; t) \to 0$ as ζ approaches infinity. If τ_{con} is constant, the equation is solved by

$$P(t) = P_0^{\exp(t/\tau_{con})}, \tag{4.9}$$

which can also be obtained from the solution for $\hat{p}(\zeta; t)$ given in Venaille & Sommeria (2007). We will show that (4.9) provides a useful fitting function for our simulation data for $P(Z_c, t)$ with finite Z_c in a transonic flow.

Similarly, we can derive an equation for the pristine fraction from the generalized version, (3.14), of the self-convolution models,

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{n}{\tau_{con}} P(1 - P^{1/n}). \tag{4.10}$$

Assuming both n and τ_{con} are constant with time, the solution of the equation is

$$P(t) = \frac{P_0}{\left[P_0^{1/n} + (1 - P_0^{1/n}) \exp\left(t/\tau_{con}\right)\right]^n}. \tag{4.11}$$

For $n = 1$, the equation reduces to (4.5) for the nonlinear integral models, and in the limit of $n \to \infty$, it approaches (4.9) for the continuous convolution model. We will use (4.11) to fit our simulation results for scalars in a highly supersonic flow, taking τ_{con} and n as fitting parameters.

5. Numerical simulations

To test the theoretical models and fix their parameters, we carried out numerical simulations for mixing in hydrodynamic turbulent flows using the FLASH code

(version 3.2), a multidimensional hydrodynamic code (Fryxell *et al.* 2000) that solves the Riemann problem on a Cartesian grid using a directionally-split piecewise-parabolic method (PPM) solver (Colella & Woodward 1984; Colella & Glaz 1985; Fryxell, Müller & Arnett 1989). We evolved the hydrodynamic equations,

$$\left.\begin{aligned} \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) &= 0, \\ \frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{v} &= -\frac{\boldsymbol{\nabla} p}{\rho} + \boldsymbol{f}, \end{aligned}\right\} \tag{5.1}$$

on a 512^3 grid for a domain of unit size with periodic boundary conditions. We adopted an isothermal equation of state, $p = \rho C_s$, with a constant sound speed, C_s. The isothermal equation of state is commonly used to imitate the nearly constant temperature in some interstellar environments, and is a convenient assumption to investigate the effects of compressibility in interstellar turbulence. Our code does not explicitly incorporate a viscosity term, and the kinetic energy is dissipated by numerical diffusion. A large-scale solenoidal external force, $\boldsymbol{f}$, was applied to drive and maintain the turbulent flows. This driving force was taken to be a Gaussian stochastic vector with an exponential temporal correlation function. We generated $\boldsymbol{f}$ in Fourier space and included all independent modes with wavenumbers in the range from 2π to 6π. Each independent mode was given the same amount of power. We defined a forcing length scale as $L_f \equiv \int (2\pi/k)\mathscr{P}_f(k)\,\mathrm{d}\boldsymbol{k} / \int \mathscr{P}_f(k)\,\mathrm{d}\boldsymbol{k}$, with $\mathscr{P}_f(k)$ being the power spectrum of the driving force, and found that L_f was equal to 0.46 box sizes for our driving scheme.

We adjusted the amplitude of the driving force to obtain a transonic flow with r.m.s. Mach number $M = 0.9$ and a supersonic flow with $M = 6.2$. We refer to the two flows as flow A and flow B. The r.m.s. Mach number was defined as the density-weighted r.m.s. velocity, v_{rms}, divided by the sound speed, C_s, and was computed from the temporal average after the flow reached a statistical steady state. We defined a flow dynamical time scale as $\tau_{dyn} \equiv L_f / v_{rms}$. The simulation setup for the turbulent flows is the same as that in Pan & Scannapieco (2010), to which we refer the interested reader for details.

To study mixing, we solved the advection equation for a number of decaying scalars, which were added to the flow once the turbulence had become fully developed and statistically stationary. The initial concentration field of the decaying scalars was bimodal, consisting of pure pollutants and completely unpolluted flow. The initial pollutant region was chosen to be a single cube located exactly at the centre of the simulation box. Within this cube, we set the concentration field, C, to be unity, i.e. the flow material there was taken to be pure pollutants, and outside the cube we set $C = 0$, i.e. the flow there was completely pollutant free. This initial condition was chosen for its simplicity, and it suffices for the purpose of illustrating the general problem and testing the theoretical models.

An important parameter for the initial condition is the pollutant fraction, P_1, i.e. the ratio of the pollutant mass to the total mass in the simulation box. Clearly, the fraction fixes the initial pristine fraction, $P_0 = 1 - P_1$. We considered three scalars in each of our flows and set the initial pollutant fraction to be $P_1 = 0.5$, 0.1 and 0.01, respectively. In the $M = 0.9$ flow, we name the three scalars with $P_1 = 0.5$, 0.1, and 0.01 as A1, A2 and A3, respectively. The corresponding cases in the $M = 6.2$ flow are named B1, B2 and B3. The exact values for P_1 were achieved by tuning the size of the pollutant regions. Smaller values of P_1 would also be of interest for mixing of heavy elements

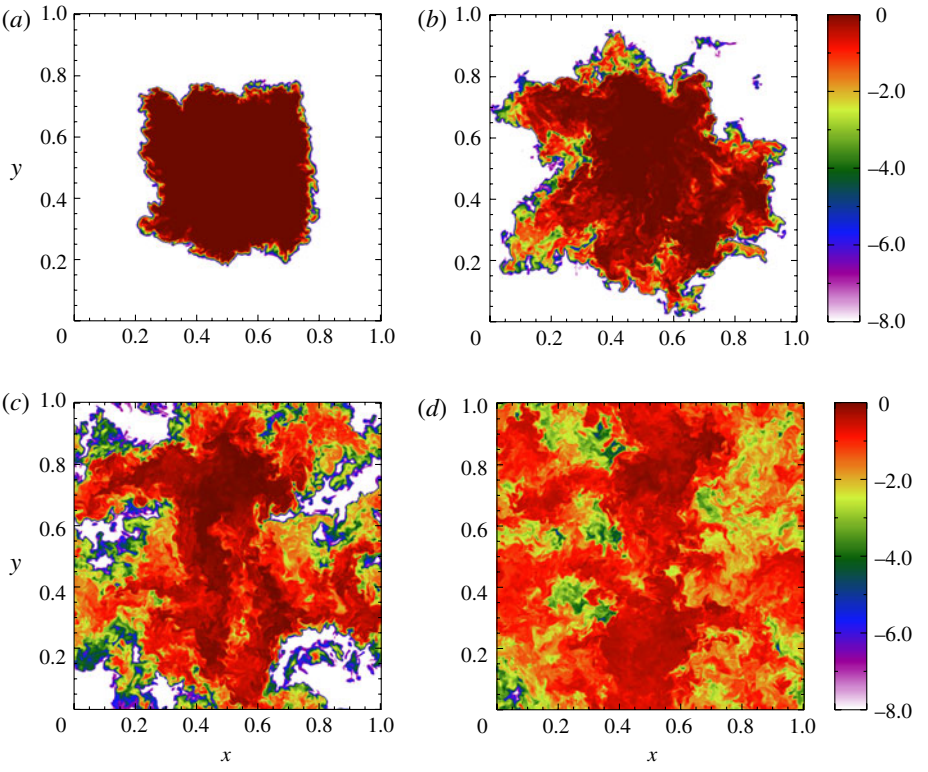

FIGURE 1. The concentration field (in logarithmic scale) of scalar A2 on a slice of the simulation grid at snapshots with $t = 0.12\tau_{dyn}$ (a), $0.5\tau_{dyn}$ (b), $0.9\tau_{dyn}$ (c) and $1.5\tau_{dyn}$ (d). The scalar is advected in the $M = 0.9$ flow. The size of the initial pollutant cube is 0.47 box sizes. The colour table ranges from 10^{-8} to 1, with the white colour representing regions with concentration $Z \leqslant 10^{-8}$.

in the interstellar media of early galaxies. However, for $P_1 \ll 0.01$, the size of the pollutant region becomes smaller than the integral scale of our simulated flows. This gives rise to complications in the evolution of the unpolluted (or slightly polluted) fraction. Smaller initial pollutant fractions will be investigated in a follow-up study.

Similar to the case of kinetic energy dissipation, the scalar dissipation (or homogenization) is also through numerical diffusion in our simulations. The diffusion scale is thus close to the resolution scale. To examine whether our results depend on the amplitude of numerical diffusion, we performed the same runs at a lower resolution, 256^3, and conducted a convergence study. We found that the time scale for the evolution of $P(Z_c, t)$ with $Z_c \sim 10^{-8}$ had already converged at the resolution 512^3.

6. Results

6.1. *The concentration field*

In figure 1, we plot the evolution of the concentration field of scalar A2 in our simulated flow with $M = 0.9$. The four panels correspond to the log of the concentration field on a slice ($z = 0.5$) of the simulation grid at four snapshots with $t = 0.12\tau_{dyn}$, $0.5\tau_{dyn}$, $0.9\tau_{dyn}$ and $1.5\tau_{dyn}$, respectively. The colour table is in logarithmic scale and the lower limit was chosen to be 10^{-8}, so that the part of the flow with

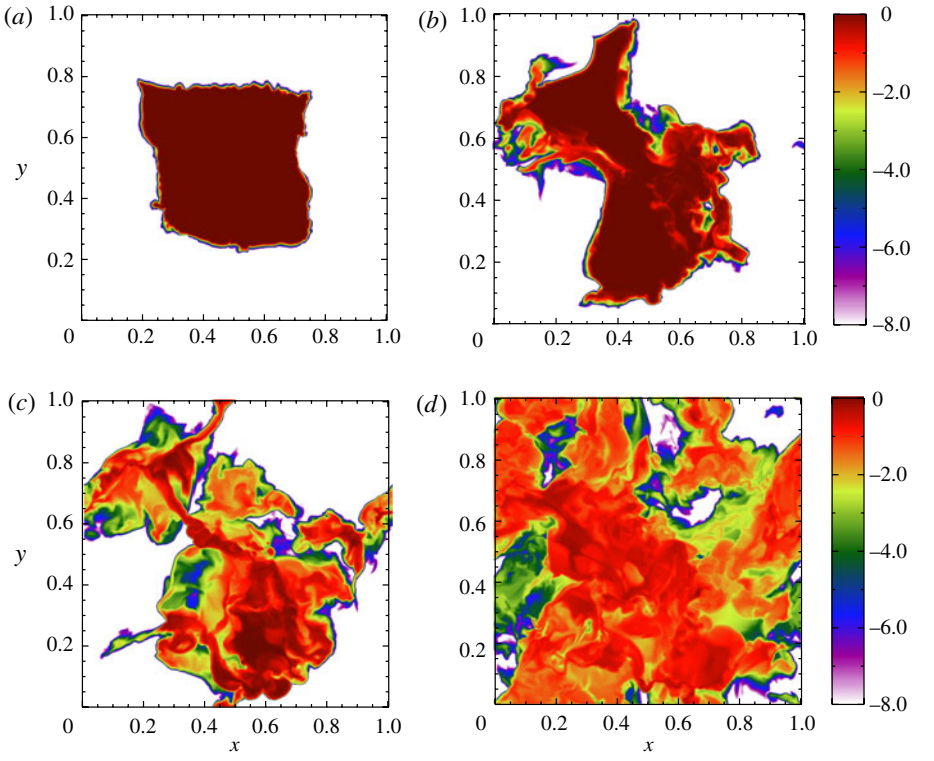

FIGURE 2. Same as figure 1, but for scalar B2 in the $M = 6.2$ flow. The four snapshots, (*a*), (*b*), (*c*) and (*d*), correspond to $t = 0.11\tau_{dyn}$, $0.65\tau_{dyn}$, $1.1\tau_{dyn}$ and $1.7\tau_{dyn}$, respectively. The size of the initial pollutant cube is also 0.47 box sizes, and the initial pollutant fraction is 0.1.

concentration below a small threshold, Z_c, is visible in the figure. The size of initial pollutant at the centre of the box was set to be 0.47 in units of the box size, such that the initial pollutant fraction is 0.1. With time, the turbulent flow transports and spreads out the pollutants, exposing them to more and more pristine fluid elements. Turbulent stretching by vortices and shear continuously produces concentration structures at smaller and smaller scales. Pristine fluid elements are contaminated when encountering a pollutant/polluted structure within a distance smaller than the diffusion scale. At $t = 1.5\tau_{dyn}$, almost the entire flow is polluted, and the mass fraction of the flow with $Z \leqslant 10^{-8}$ becomes negligibly small. Cliff structures typical of passive scalar fields advected in incompressible turbulence are clearly observed in (*c*) and (*d*).

Figure 2 shows the concentration field of scalar B2 in our $M = 6.2$ flow. At the four snapshots selected here, the density-weighted concentration variances are close to those for scalar A2 at the corresponding snapshots in figure 1. It appears that, at similar variances, the unpolluted volume in the $M = 6.2$ flow is significantly larger than in the $M = 0.9$ case. The existence of strong compressions and expansions in a highly supersonic flow gives rise to a very different geometry for the scalar field. Most of the volume in a highly compressible flow is occupied by expanding events, and the flow expansion tends to produce more coherent scalar structures, as a passive scalar follows the expansion. Therefore, scalar B2 appears to be smoother than A2 in figure 1. The edge-like structures in figure 2 are produced by the compression of

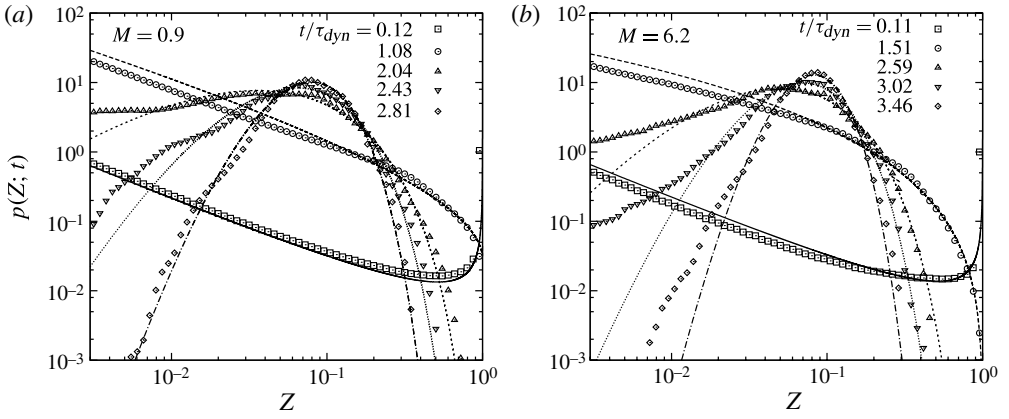

FIGURE 3. Evolution of p.d.f.s of scalars A2 (*a*) and B2 (*b*) in the $M = 0.9$ and $M = 6.2$ flow simulations, respectively. The initial pollutant fraction, P_1, for the two scalars is 0.1. Lines are the predicted p.d.f.s from the mapping closure model with the same values of variance as in the simulations.

velocity shocks, which amplifies the scalar gradient across the shocks. Although the visual impression of figure 2 is dominated by the effect of compressible modes, it is actually the solenoidal modes that contain the majority of kinetic energy in the flow and provide the primary contribution to the scalar cascade even at high Mach numbers (Pan & Scannapieco 2010). The interested reader is referred to Pan & Scannapieco (2010, 2011) for detailed discussions of scalar structures as a function of the flow Mach number and the relative role of solenoidal and potential modes for mixing in supersonic turbulence.

6.2. *The p.d.f. evolution*

Figure 3 plots the p.d.f.s of scalars A2 (*a*) and B2 (*b*) as a function of time. The two scalars were evolved in our simulated flows with $M = 0.9$ and $M = 6.2$, respectively. The initial pollutant fraction, P_1, is 0.1 for the two scalars, meaning that the amplitude of the initial spike at $Z = 0$ is 9 times higher than that at $Z = 1$. Turbulent mixing reduces the heights of the two spikes, and gradually fills the concentration space in between. Eventually a central peak forms around the average concentration, and then the p.d.f. narrows down toward the peak.

The lines in figure 3 are the prediction of the mapping closure model. At each time, the predicted p.d.f. has the same value of variance as that from the simulation (data points). This is equivalent to properly choosing the time scale $\lambda_\theta^2(t)/\kappa$ in (3.6) so that the variance evolution from the model matches the simulation result. The model prediction is in good agreement with the data at the central part and the right (high-Z) tail of the scalar p.d.f. However, the mapping closure considerably underestimates the left p.d.f. tail at intermediate to late times. A detailed discussion of the discrepancy between the prediction of the Gaussian mapping closure and simulation results at late times is given in Girimaji (1992). The weakness of the mapping closure is also discussed by Duplat & Villermaux (2008). It appears that, for scalar A2, the agreement between the mapping closure and the simulation data becomes better in the long-time limit.

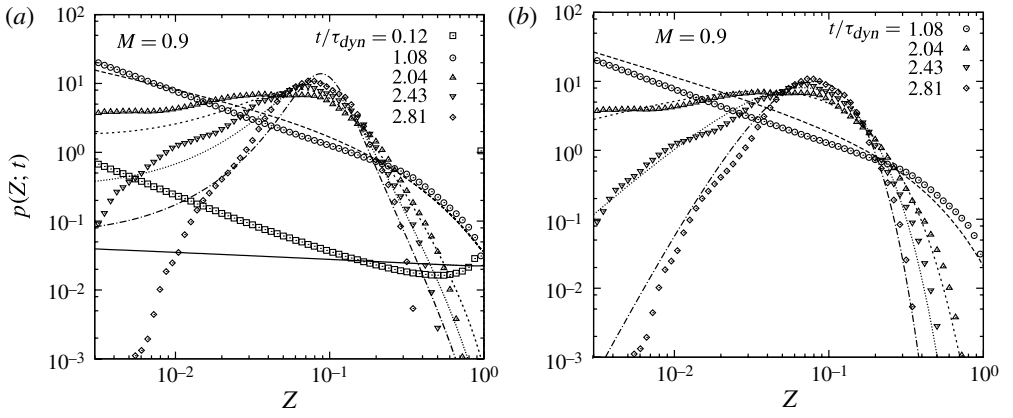

FIGURE 4. Evolution of the p.d.f. of scalar A2 in the $M = 0.9$ flow. Lines are predictions of two models: (*a*) the nonlinear integral model with uniform $J(Z; Z_1, Z_2)$ (equation (3.11)); (*b*) gamma distributions as predicted by the model of Villermaux & Duplat (2003).

In figure 3(*b*) for scalar B2 in the $M = 6.2$ flow, we see that at early times there is also an acceptable agreement between the mapping closure prediction and the simulation results. At later times, the discrepancy at the left tails between the model and the data is larger than the $M = 0.9$ case. This is because the p.d.f. tails broaden with increasing Mach number, as previously found in the simulations of Pan & Scannapieco (2010). The origin of this effect was argued to be related to the increasing degree of intermittency of the velocity field as the Mach number increases.

The mass fraction of unpolluted or slightly polluted flow with $Z \lesssim 10^{-8}$–10^{-5} corresponds to the far left tail of the p.d.f. with Z well below the minimum value shown in figure 3. Therefore, figure 3 does not contain direct information for the part of the p.d.f. of our primary interest. Nevertheless, the left p.d.f. tails shown in figure 3 imply that the mapping closure model is likely to significantly underestimate the unpolluted fraction at intermediate to late times especially for scalar B2. The expectation is confirmed in § 6.4.

In figure 4(*a*), we compare the prediction of the nonlinear integral model with uniform $J(Z; Z_1, Z_1)$ to the simulation data for scalar A2 in the $M = 0.9$ flow. The performance of this model for the p.d.f. evolution is poor. At very early times, the predicted p.d.f. appears to be flat in between the initial spikes, reflecting a 'memory' of the uniform function $J(Z; Z_1, Z_1)$. As mentioned earlier, a problem of the nonlinear integral models is that they predict excessively fat tails in the long-time limit. This is seen in figure 4(*a*), which shows that at large t the model significantly overestimates the left tails. We find that at late times the nonlinear integral model also overestimates the p.d.f. tails for scalars in the $M = 6.2$ flow (not shown). Although the nonlinear integral models do not give good predictions for the p.d.f. evolution, we find that they provide useful fits to the pristine fraction in certain physical regimes (see § 6.4).

In figure 4(*b*), we fit the p.d.f. of scalar A2 in the $M = 0.9$ flow with the gamma distribution, (3.16), predicted by the model of Villermaux & Duplat (2003). For each line, the value of n is chosen such that the variance of the gamma distribution is equal to that from the simulation. The scalar p.d.f. at $t \lesssim \tau_{dyn}$ does not have a gamma distribution shape, and is not shown in this figure. At the four selected times in the figure, however, the gamma distributions fit the simulation data quite well. The

agreement is significantly better than the mapping closure for t between $1.8\tau_{dyn}$ and $2.6\tau_{dyn}$ (corresponding to $1.1 \leqslant n \leqslant 4$). We find that $n(t)$ increases exponentially with time, corresponding to the exponential decay of the scalar variance at late times (see § 6.3), as the variance of the gamma distribution, (3.16), goes like $\langle Z\rangle^2/n$. This is in contrast to the experimental result, $n(t) \propto t^{5/2}$, found by Villermaux & Duplat (2003). The reason for the difference is that our simulated flows are maintained at a steady state by a driving force and are statistically homogeneous, while the experiments by Villermaux & Duplat (2003) are for decaying flows dominated by a mean shear. An exponential decay is also found in a sustained flow by Villermaux, Stroock & Stone (2008).

We point out that the continuous convolution model of Venaille & Sommeria (2007) predicts that, if the scalar p.d.f. is given by a gamma distribution at a given time, then the p.d.f. will remain a gamma distribution at all subsequent times. Therefore, if one starts to use the continuous convolution model at a time when the scalar p.d.f. has evolved to a gamma distribution, its prediction for later times would be the same as the model of Villermaux & Duplat (2003). However, unlike Villermaux & Duplat (2003), the model of Venaille & Sommeria (2007) does not predict that the gamma distribution is an attractive solution that the scalar p.d.f. always reaches at the late evolution stage.

For scalar A2 in the $M = 0.9$ flow, the model of Villermaux & Duplat (2003) starts to apply at $t = 1\tau_{dyn}$–$2\tau_{dyn}$. At this time, the pristine mass fraction has already decreased to very small values, and thus the model is not suitable for studying the pristine fraction. We also tried to fit gamma distributions to the scalar p.d.f.s in the $M = 6.2$ flow, and found they underestimate the left tails, which are broader than the $M = 0.9$ case.

In summary, we found that in the $M = 0.9$ flow the mapping closure gives acceptable fits to the scalar p.d.f. at early times, but significantly underestimates the left p.d.f. tails at late times. Starting from $\simeq 1.8\tau_{dyn}$, the scalar p.d.f. is better fitted by a gamma distribution, which is predicted by the model of Villermaux & Duplat (2003). Since all the models we considered were originally developed to study mixing in incompressible turbulence, they were not expected to perform well in highly compressible flows. In the $M = 6.2$ flow, the left p.d.f. tail is broader, and no models were found to give satisfactory predictions for the scalar p.d.f. at late times.

6.3. *The variance decay*

In figure 5, we show the evolution of the density-weighted concentration variance for scalars A2 and B2 in the $M = 0.9$ and $M = 6.2$ flows, respectively. In figure 5(a), the time is normalized to the flow dynamical time τ_{dyn}, while in figure 5(b) it is normalized to the acoustic time τ_{ac} defined as L_f/C_s, the time for the sound wave to cross the driving scale of the flow, L_f. From the definition, we have $\tau_{ac} = M\tau_{dyn}$. Therefore, the variance decay curves shift to the right by a factor of 1.1 for scalars A2 and to the left by a factor of 6.2 for B2, when the normalization time scale is switched from τ_{dyn} to τ_{ac}. From figure 5, it is clear that the flow dynamical time is a more relevant time scale for the scalar variance decay in supersonic isothermal turbulent flows.

As seen in figure 5(a), at early times the variance decreases slowly, corresponding to the initial development of scalar structures toward small scales. In this transient period, the parameter $\gamma(t)$ in the models presented in §§ 3.3 and 3.4 would be time-dependent if the model is required to match the scalar variance decay. At $t \gtrsim 0.5\tau_{dyn}$, the variance decays exponentially, and $\gamma(t)$ would be constant. When normalized to the dynamical

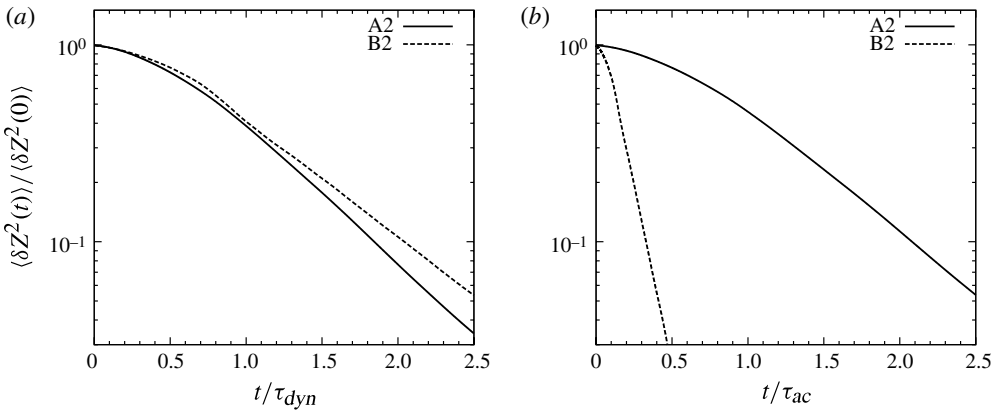

FIGURE 5. Evolution of the density-weighted concentration variance for scalars A2 and B2 in $M = 0.9$ and $M = 6.2$ flows, respectively. The variance is normalized to its initial value. (*a*) Time normalized to the flow dynamical time scale. (*b*) Time normalized to the acoustic time scale.

time scale, the variance decay of scalar A2 is slightly faster than B2, and the decay time scale is measured to be $\tau_m = 0.61\tau_{dyn}$ and $0.73\tau_{dyn}$, for A2 and B2, respectively. These results are consistent with Pan & Scannapieco (2010), who found that the variance decay time scale in units of the flow dynamical time scale has a weak dependence on the flow Mach number, increasing by $\lesssim$20 % as M goes from 1 to 6. This slight increase is due to the fact that compressible modes are less efficient at enhancing the mixing rate than solenoidal modes (Pan & Scannapieco 2010).

We also measured τ_m for the other four scalars included in our simulations, and found that, in each flow, the mixing time scales of the three scalars are close to each other. The time scale for the third scalar (i.e. A3 or B3) is slightly smaller than the other two. On average, the mixing time scale is $\tau_m \simeq 0.6\tau_{dyn}$ for the three scalars in the $M = 0.9$ flow. In the $M = 6.2$ flow, the average mixing time scale is $\simeq 0.7\tau_{dyn}$.

6.4. *The fraction* $P(Z_c, t)$

Next we computed the mass fraction, $P(Z_c, t)$, of fluid elements with Z smaller than different thresholds, Z_c, from the simulation data. We found that the flow with exactly zero concentration (i.e. the special case with $Z_c = 0$) is erased rapidly by numerical diffusion. This is because in each time step the unpolluted computation cells adjacent to those with non-zero Z obtain a finite (but tiny) concentration value due to numerical diffusion. Therefore, after a small number of time steps, no exactly pollutant-free cells were left in the simulation box. However, in such a short time, the degree of pollution in most cells due to numerical diffusion is extremely small, and the concentration level in these cells was much smaller than any threshold of practical interest. The rapid pollution of completely pristine mass by numerical diffusion in the simulations is similar to the effect of molecular diffusivity, which tends to reduce $P(t)$ to zero instantaneously, although the numerical diffusion probably has a different form and a much larger amplitude than the realistic molecular diffusivity.

The threshold of interest for astrophysical applications is $Z_c \gtrsim 10^{-8}$. For this finite threshold, the time scale at which $P(Z_c, t)$ decreases is significant, and is of the order of the flow dynamical time. A comparison of the same simulation runs at two resolutions, 256^3 and 512^3, shows that the time scale for $P(Z_c, t)$ with $Z_c \sim 10^{-8}$ is

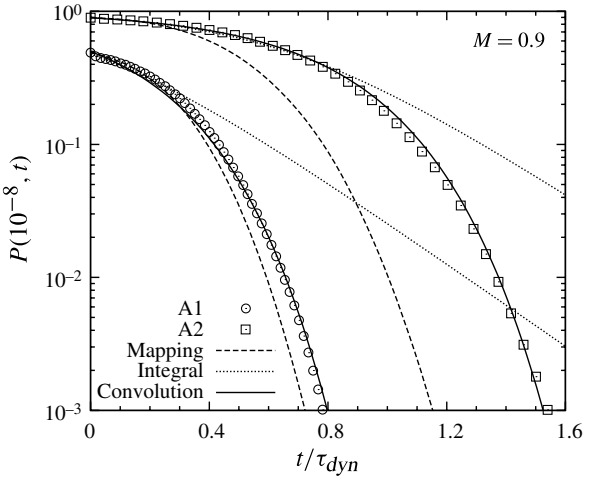

FIGURE 6. Mass fraction of fluid elements with $Z \leqslant 10^{-8}$ for scalars A1 and A2 in the $M = 0.9$ flow. Dashed lines correspond to the prediction of the mapping closure model. Solid lines are the best fits using the continuous convolution model with $\tau_{con} = 0.35\tau_{dyn}$ and $0.37\tau_{dyn}$ for A1 and A2, respectively. The dotted lines are fits by (4.5) from the nonlinear integral models.

independent of the amplitude of numerical diffusion. These suggest that the reduction rate of $P(Z_c, t)$ with $Z_c \gtrsim 10^{-8}$ is mainly determined by the large-scale properties of the flow. The behaviour of $P(Z_c, t)$ as a function of time is similar for different values of Z_c, given that Z_c is much smaller than the average concentration. The evolution time scale of $P(Z_c, t)$ only has a weak logarithmic dependence on Z_c. Below we only present results for $Z_c = 10^{-8}$. Similar results are found for Z_c in the range from 10^{-8} to 10^{-5}.

6.4.1. *The $M = 0.9$ flow*

In figure 6, we plot the mass fraction of fluid elements with $Z \leqslant 10^{-8}$ as a function of time for scalars A1 and A2 in the $M = 0.9$ flow. The initial pollutant fraction, P_1, is 0.5 and 0.1 for the two scalars, respectively. The data points are results from the simulations. The dashed lines correspond to the prediction of the mapping closure model. For this model, the fraction, $P(Z_c, t)$, at a given time is calculated from the predicted p.d.f. with the same scalar variance as that in the simulation. In figure 6, we see that the mapping closure model is in good agreement with the data points for scalar A1. However, the model prediction is well below the data points for scalar A2. This was expected from figure 3(*a*), which shows that the mapping closure model underestimates the left p.d.f. tail of scalar A2 at intermediate to late times. We found that the model also underestimates $P(Z_c, t)$ for scalar A3, which had an initial pollutant fraction of 0.01, and the discrepancy is even larger than the case of A2.

The solid lines in figure 6 are from the continuous convolution model of Venaille & Sommeria (2007), i.e. (4.9), and they match the data points quite well. The time scales τ_{con} used in the fits are $0.35\tau_{dyn}$ and $0.37\tau_{dyn}$, respectively, for scalars A1 and A2. As discussed in § 4.3, (4.9) was originally derived for the fraction, $P(t)$, of exactly pollutant-free mass. The good agreement between (4.9) and the simulation data shows that the model actually provides an excellent fitting function for the fraction, $P(Z_c, t)$, with a finite (but small) threshold Z_c. It also suggests that the continuous

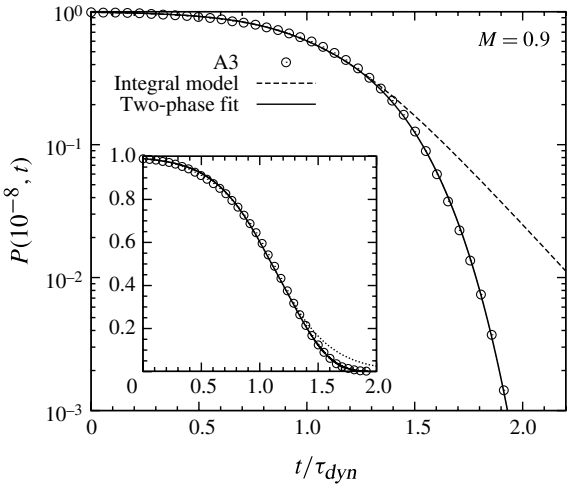

FIGURE 7. Mass fraction of fluid elements with $Z \leqslant 10^{-8}$ for scalar A3 in the $M = 0.9$ flow. The dashed line corresponds to the nonlinear integral model with $\tau_{int} = 0.24\tau_{dyn}$. The solid line is the best fit obtained by combining the nonlinear integral model for the early phase and the continuous convolution model for the later phase. The time scales used in the two phases are $\tau_{int} = 0.24\tau_{dyn}$ and $\tau_{con} = 0.36\tau_{dyn}$, respectively. The inset shows the same data points and lines, but with the vertical coordinate on a linear scale.

convolution process is a good physical description for the erasure of unpolluted (or slightly polluted) flow by turbulent mixing, if the initial pollutant fraction, P_1, is larger than ~ 0.1.

The dotted lines in figure 6 shows the fits using (4.5) from the nonlinear integral model. The time scale τ_{int} was chosen to be $0.28\tau_{dyn}$ and $0.3\tau_{dyn}$ for scalars A1 and A2, respectively. This model predicts an exponential decrease at late times which is much slower than found in the simulation. The overestimate is because the model produces excessively broad p.d.f. tails in the long time limit (see figure 4*a*).

Figure 7 shows the evolution of $P(10^{-8}, t)$ for scalar A3, whose initial pollutant fraction $P_1 = 0.01$. The inset plots the same data points and model fits, but the y-axis is on a linear scale. Unlike the case of scalars A1 and A2, the data points for scalar A3 cannot be well fitted by the continuous convolution model with a single time scale right from the beginning. In fact, $P(10^{-8}, t)$ exhibits different behaviours at early and late times.

The early phase can be well fitted by (4.5) from the nonlinear integral model with $\tau_{int} = 0.24\tau_{dyn}$. This is shown as a dashed line in figure 7, and from the inset we see that the line matches the simulation data over an extended time range before $P(10^{-8}, t)$ decreases to $\simeq 0.3$. The time scale τ_{int} used here corresponds to $\sim 0.4\tau_m$ since the variance decay time scale τ_m for scalar A3 is $\sim 0.6\tau_{dyn}$. This is in between $1/2\tau_m$ and $1/3\tau_m$, the expected values of τ_{int} for Curl's model and the model with uniform $J(Z; Z_1, Z_2)$, respectively (see § 4.2). The dashed line starts to significantly overestimate the simulation results when $P(10^{-8}, t)$ becomes smaller than ~ 0.3. Again, this is because the nonlinear integral models significantly overpredict the p.d.f. tails at late times.

We find that the late-time behaviour of $P(10^{-8}, t)$ can be well described by the continuous convolution model. In fact, this model starts to give a satisfactory fit quite

early, just after $P(10^{-8}, t)$ becomes smaller than $\sim$0.8. The best-fit value of the time scale τ_{con} is $0.36\tau_{dyn}$, which is almost the same as the values used to fit the data for scalars A1 and A2. This, together with the results for scalars A1 and A2, suggests that the continuous convolution model applies if the mass fraction of pollutants or the polluted flow is larger than 0.1–0.2. We point out that there is an extended range of $P(10^{-8}, t)$ (from 0.3 to 0.8) where both the nonlinear integral model and the continuous convolution model can match the data well.

We give a physical speculation for why the early phase of scalar A3 is better fitted by the nonlinear integral model than the continuous convolution model. For this scalar, the amount of pollutants or polluted mass is small at early times. The limited availability of pollution sources leads to a relatively low frequency of contact between the polluted and unpolluted flow. As a consequence, the pollution process would involve fewer convolutions at this stage, and thus may be better captured by the nonlinear integral model, which can be roughly viewed as a discrete self-convolution model in Laplace space. The simulation results presented above suggest that the mixing events between the unpolluted and the polluted fluid elements become frequent enough to trigger the continuous convolution process, when the polluted fraction is larger than 0.1–0.2.

The solid line in figure 7 is obtained by combining the best fits for the early and late phases using the nonlinear integral model and the continuous convolution model, respectively. Clearly, this line is in excellent agreement with the simulation results. We connected the two phases at time, $t_{0.5}$, when $P(10^{-8}, t) = 0.5$, i.e. the second phase is fitted by $0.5^{\exp[(t-t_{0.5})/\tau_{con}]}$. As mentioned earlier, the best-fit time scales for the two phases are $\tau_{int} = 0.24\tau_{dyn}$ and $\tau_{con} = 0.36\tau_{dyn}$, respectively. The choice of connecting the two phases at $P(10^{-8}, t) = 0.5$ is somewhat arbitrary. In fact, combining the two models at any time with $P(10^{-8}, t)$ between $\simeq$0.8 and $\simeq$ 0.3 yields satisfactory results.

The time scales, τ_{int} and τ_{con}, were set to be constant in all our fits to the simulation results for $P(Z_c, t)$. These time scales can be a function of time in general. In fact, from the consideration of the scalar variance decay, these time scales would be time-dependent at early times when the variance decay is slower than exponential, and then become constant at $t \gtrsim 0.5\tau_{dyn}$ when the exponential decay starts (see § 6.3). The reason why assuming constant time scales applies perfectly for the evolution of the pristine mass fraction, but not for the scalar variance decay at all times is probably that the pollution of pristine flow is physically simpler. A fast variance decay relies on the full development of scalar structures toward the diffusion scale, and that is why the decay is slow at early times as the cascade is just starting. On the other hand, the pollution of pristine mass does not need to wait for the scalar structures to be fully developed at small scales: the pollution occurs whenever the unpolluted fluid elements are brought into contact with the pollutants or the polluted flow. This happens as soon as the pollutants are released into the flow. This is perhaps why using constant time scales, τ_{int} and τ_{con}, from the very beginning could give successful fits to the evolution of $P(Z_c, t)$.

6.4.2. *The $M = 6.2$ flow*

Figure 8 shows our results for scalars B1, B2 and B3 with initial pollution fractions of 0.5, 0.1, and 0.01, respectively, in the $M = 6.2$ flow. We find that the mapping closure model significantly underestimates $P(10^{-8}, t)$ for all the three scalars, and the discrepancy is much larger than the case of the $M = 0.9$ flow. In figure 8, we attempt to fit the simulation data for scalars B1 and B2 with the continuous convolution model (equation (4.9)) of Venaille & Sommeria (2007), as in the $M = 0.9$ case. The time

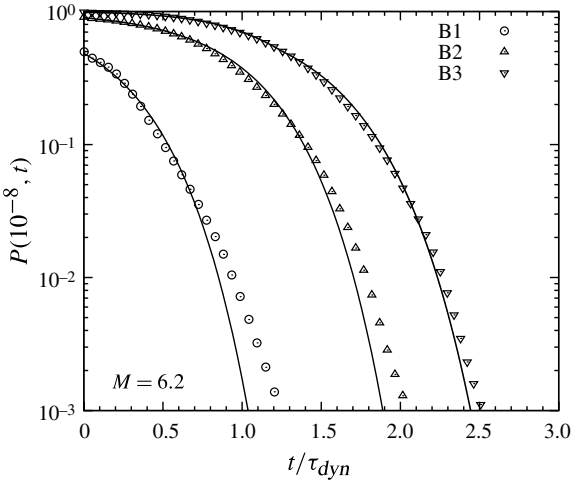

FIGURE 8. Mass fraction of fluid elements with $Z \leqslant 10^{-8}$ for scalars B1, B2 and B3 in the $M = 6.2$ flow. The solid lines for B1 and B2 correspond to the fits by the continuous convolution model (equation (4.9)) with $\tau_{con} = 0.46\tau_{dyn}$ for both cases. The line for B3 combines the nonlinear integral model with $\tau_{int} = 0.30\tau_{dyn}$ for the early phase and the continuous convolution model with $\tau_{con} = 0.51\tau_{dyn}$ for the later phase.

scale, τ_{con}, used in the fitting lines is $0.46\tau_{dyn}$ for both B1 and B2. Again the fitting curve for scalar B3 consists of two phases that connect at $P(Z_c, t) = 0.5$. The early phase is fitted by the nonlinear integral model with $\tau_{int} = 0.3\tau_{dyn}$, and the later phase uses the continuous convolution model with $\tau_{con} = 0.51\tau_{dyn}$. The parameter choice here gives priority to satisfactorily matching the data points at early times.

All the time scales chosen in figure 8 are larger than the corresponding values used for scalars in the $M = 0.9$ flow. This is caused by two effects. First, as mentioned earlier, when normalized to the flow dynamical time, the variance decay time scale in the $M = 6.2$ flow is slightly larger than in the $M = 0.9$ case. Second, as shown in figure 3, the left tail of the scalar p.d.f. broadens with increasing Mach number of the advecting flow. This means that, with the same concentration variance, the scalar p.d.f. in the $M = 6.2$ flow contains a larger probability at low concentration levels. Both effects tend to result in a slower decrease of $P(Z_c, t)$ in the flow with higher M. The second effect appears to be stronger than the first one, and it also explains why the same models that match the results well in the $M = 0.9$ flow significantly underestimate $P(Z_c, t)$ in the $M = 6.2$ flow at late times (see figure 8). The fitting quality of the lines in figure 8 is good at early times when $P(Z_c, t) \gtrsim 0.1$, and generally acceptable before $P(Z_c, t)$ decreases to 0.01. Below that, however, significant discrepancy appears.

We find that the simulation results for B1 and B2 in the $M = 6.2$ flow can be better fitted by (4.11) from the generalized self-convolution model (see (3.14) in § 3.4; Duplat & Villermaux 2008). The parameter n in this equation controls the shape of the fitting curve. Figure 9 shows our results using this equation to fit the simulation data. The data points here are the same as in figure 8. Equation (4.11) with $n = 4.6$ can fit the simulation data well for scalars B1 and B2 at all times and for scalar B3 in the late phase. For B1 and B2, the best-fit time scale τ_{con} is, respectively, $0.40\tau_{dyn}$ and $0.41\tau_{dyn}$. For the early phase of scalar B3, we used the same nonlinear integral model

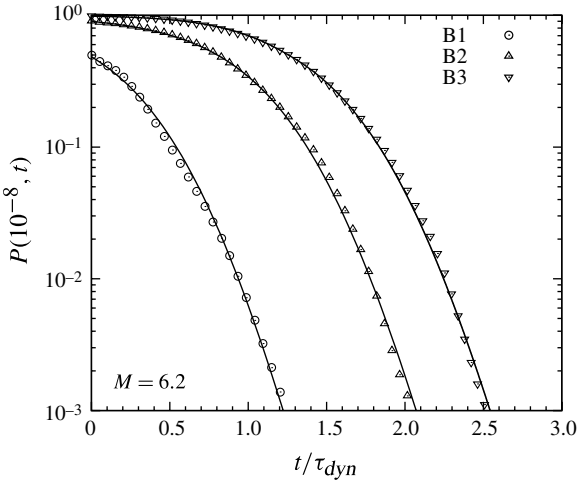

FIGURE 9. Mass fraction of fluid elements with $Z \leqslant 10^{-8}$ for scalars B1, B2 and B3 in the $M = 6.2$ flow. The data points are the same as in figure 8. Equation (4.11) with $n = 4.6$ is used to match the simulation data. The time scale, τ_{con}, in the equation is set to $0.40\tau_{dyn}$ and $0.41\tau_{dyn}$ for scalars B1 and B2, respectively. The line for B3 is a combination of the nonlinear integral model with $\tau_{int} = 0.30\tau_{dyn}$ for the early phase and (4.11) with $\tau_{con} = 0.42\tau_{dyn}$ and $n = 4.6$ for the later phase.

as in figure 8 with $\tau_{int} = 0.3\tau_{dyn}$. The late phase is fitted by (4.11) with $\tau_{con} = 0.42\tau_{dyn}$ and $n = 4.6$. We combined the two phases at $P(Z_c, t) = 0.5$. A comparison of figure 8 and figure 9 shows that, with (4.11) from the generalized convolution model, the fitting quality is significantly improved.

We point out that (4.11) is used simply as a fitting function. The generalized convolution model (§ 3.4) behind this equation does not address the effects of shocks and compressibility on the p.d.f. of passive scalars in supersonic turbulence. There is thus no physical reason why it provides successful fits to the pristine mass fraction in the $M = 6.2$ flow. A physical closure model is motivated to successfully explain and match the evolution of $P(Z_c, t)$ in highly supersonic turbulence.

7. Conclusions

Motivated by the process of primordial star formation in the first generation of galaxies, we investigated the general problem of how the unpolluted flow material in compressible turbulence is contaminated by mixing. We approached this problem using both theoretical modelling and numerical simulations. The theoretical approach is based on the probability distribution method for turbulent mixing, since the fraction of the unpolluted or slightly polluted mass corresponds to the left tail of the concentration p.d.f. We first derived an equation for the concentration p.d.f. with density weighting, where the advection term exactly conserves the global p.d.f. We then considered several existing closure models for the diffusivity term in the p.d.f. equation, including the mapping closure model (Chen *et al.* 1989), the nonlinear integral models (Curl 1963; Dopazo 1979; Janicka *et al.* 1979) and the self-convolution models (Venaille & Sommeria 2007; Duplat & Villermaux 2008), and derived the predictions of these models for the exactly unpolluted fraction, $P(t)$, or for the fraction, $P(Z_c, t)$, of the flow with Z below a small threshold, Z_c.

To test and constrain the model predictions, we carried out numerical simulations evolving decaying scalars in two isothermal turbulent flows with r.m.s. Mach numbers of 0.9 and 6.2. Three passive scalars were included in each flow, and their initial pollutant fractions, P_1 were set to be 0.5, 0.1 and 0.01, respectively. We found that the mapping closure model gives satisfactory predictions for the central part and the high-Z tails of the scalar p.d.f., but underestimates the low-Z tail at large times, especially for scalars with small initial pollutant fraction. The left p.d.f. tails become broader with increasing flow Mach number, and thus the discrepancy between the mapping closure prediction and the simulation results is larger at Mach 6.2. We showed that, in the $M = 0.9$ flow, the scalar p.d.f. is well fitted by gamma distributions at late times, as predicted by Villermaux & Duplat (2003). All the closure models adopted in our study were originally developed for mixing in incompressible turbulence, and they do not provide successful predictions for the scalar p.d.f. in the highly supersonic flow.

Our simulation results for $P(Z_c, t)$ in the Mach 0.9 flow can be well fitted by using (4.5) and (4.9) from the nonlinear integral model and the continuous convolution model of Venaille & Sommeria (2007), respectively. Although these two equations were originally derived for the fraction of exactly pollutant-free mass, they provide useful fitting functions for $P(Z_c, t)$ with a small finite threshold, Z_c. We showed that, for the two scalars with $P_1 \geqslant 0.1$, the evolution of $P(Z_c, t)$ follows the equation $\dot{P}(Z_c, t) = P(Z_c, t) \ln[P(Z_c, t)]/\tau_{con}$ from the continuous convolution model. On the other hand, for the scalar with $P_1 = 0.01$, $P(Z_c, t)$ shows different behaviours at early and late times. In the early phase, the evolution of $P(Z_c, t)$ is consistent with the equation $\dot{P}(Z_c, t) = -P(Z_c, t)[1 - P(Z_c, t)]/\tau_{int}$ from the nonlinear integral model, and the later phase is well fitted by the continuous convolution model. A satisfactory fit to the entire behaviour of $P(Z_c, t)$ was obtained by connecting the two phases. The continuous convolution model starts to apply once the polluted mass fraction is larger than 0.1–0.2.

When normalized to the flow dynamical time (τ_{dyn}), the decay of $P(Z_c, t)$ is slower in the $M = 6.2$ flow for two reasons. First, the mixing time scale in units of τ_{dyn} is $\sim$20 % larger than in the $M = 0.9$ flow. Second, at the same variance, the left tail of the scalar p.d.f. is broader at higher Mach numbers. Owing to the second effect, the shape of the $P(Z_c, t)$ curve as a function of time changes as the Mach number increases. We find that a generalized version of the self-convolution model (§§ 3.4 and 4.3, see Duplat & Villermaux 2008) provides a good fitting function, (4.11), to the evolution of $P(Z_c, t)$ in highly supersonic turbulence. With $n = 4.6$, this equation matches our simulation data well for the two scalars with $P_1 \geqslant 0.1$. Like the case of the $M = 0.9$ flow, we obtained a good fit to the simulation result for the scalar with $P_1 = 0.01$ by combining different behaviours at early and late times. At early times, we used (4.5) from the nonlinear integral model, while the later phase was fitted by (4.11) with $n = 4.6$. We point out that, although it provides a good fitting function to the pristine mass fraction, the generalized convolution model does not have a physical connection to how the flow compressibility affects turbulent mixing in supersonic flows. Physical p.d.f. closure models are motivated to directly incorporate the effects of shocks or the flow compressibility on mixing in supersonic turbulence.

The fitting functions obtained this study can be used to develop a subgrid model for large-scale simulations for mixing of heavy elements in the interstellar media of early galaxies. Such a subgrid model would provide an important step toward predicting the fraction of pristine gas in the first generation of galaxies. In order to apply our results with higher accuracy, we will carry out a systematic numerical study in a future work

covering a broader range of parameters including varying initial scalar conditions and turbulent Mach numbers.

Acknowledgements

L.P. thanks Professor G.-W. He for helpful discussions on the mapping closure model, and the Institute of Mechanics, Chinese Academy of Sciences for its hospitality during his visit. L.P. and E.S. acknowledge support from NASA under theory Grant No. NNX09AD106 and astrobiology institute grant 08-NAI5-0018 and from the National Science Foundation under grant AST 11-03608. J.S. acknowledges support by the NASA Astrobiology Institute, Virtual Planetary Laboratory Lead Team. All simulations were conducted at the Arizona State University Advanced Computing Center and the Texas Advanced Computing Center, using the FLASH code, a product of the DOE ASC/Alliances-funded Center for Astrophysical Thermonuclear Flashes at the University of Chicago.

Appendix. The p.d.f. equation

In this appendix, we derive the equation for the concentration p.d.f. in compressible flows using the technique developed by Lundgren (1967). Similar derivations for the scalar p.d.f. equation can be found in e.g. Pope (1976), Obrien (1980), Dopazo *et al.* (1997) and Pope (2000). Unlike these previous derivations, we adopt a density-weighting scheme, which is needed for passive scalar mixing in compressible turbulence.

Our derivation makes use of a statistical ensemble consisting of many independent realizations. We start with the definition of the fine-grained p.d.f. in a single realization. Because in a specific realization the concentration field is single-valued at given position ($\boldsymbol{x}$) and time (t), the fine-grained p.d.f. is a delta function

$$q'(Z;\boldsymbol{x},t)=\delta[Z-C(\boldsymbol{x},t)], \tag{A 1}$$

where Z is the sampling variable. Considering the existence of significant density fluctuations in supersonic flows, we define a fine-grained p.d.f. with density weighting

$$p'(Z;\boldsymbol{x},t)=\tilde{\rho}(\boldsymbol{x},t)\delta[Z-C(\boldsymbol{x},t)], \tag{A 2}$$

where the density-weighting factor $\tilde{\rho}$ is the ratio of the local flow density $\rho(\boldsymbol{x},t)$ to the average density $\bar{\rho}$. These two fine-grained p.d.f.s are functions of Z, $\boldsymbol{x}$ and t, and their dependence on space and time is though $C(\boldsymbol{x},t)$ and $\rho(\boldsymbol{x},t)$.

We calculate the time-derivatives of $q'(Z;\boldsymbol{x},t)$ and $p'(Z;\boldsymbol{x},t)$. Since $q'(Z;\boldsymbol{x},t)$ depends on t only through the quantity $Z-C(\boldsymbol{x},t)$, we have

$$\frac{\partial q'(Z;\boldsymbol{x},t)}{\partial t}=-\frac{\partial C(\boldsymbol{x},t)}{\partial t}\frac{\partial q'(Z;\boldsymbol{x},t)}{\partial Z}. \tag{A 3}$$

Using (A 2) and (A 3), we obtain the time-derivative of $p'(Z;\boldsymbol{x},t)$,

$$\frac{\partial p'(Z;\boldsymbol{x},t)}{\partial t}=\frac{\partial\tilde{\rho}}{\partial t}q'(Z;\boldsymbol{x},t)-\tilde{\rho}\frac{\partial C(\boldsymbol{x},t)}{\partial t}\frac{\partial q'(Z;\boldsymbol{x},t)}{\partial Z}. \tag{A 4}$$

Similarly, we can derive an equation for $\boldsymbol{\nabla}\cdot(p'\boldsymbol{v})$,

$$\boldsymbol{\nabla}\cdot(p'\boldsymbol{v})=\left[\boldsymbol{\nabla}\cdot(\tilde{\rho}\boldsymbol{v})\right]q'(Z;\boldsymbol{x},t)-\tilde{\rho}\boldsymbol{v}\cdot\boldsymbol{\nabla}C(\boldsymbol{x},t)\frac{\partial q'(Z;\boldsymbol{x},t)}{\partial Z}. \tag{A 5}$$

We add (A 4) and (A 5). From the continuity equation for $\tilde{\rho}(\boldsymbol{x}, t)$, the first terms on the right-hand sides of (A 4) and (A 5) add up to zero. Using the advection–diffusion equation (2.1) of $C(\boldsymbol{x}, t)$ for the sum of the last terms in these two equations, we have

$$\frac{\partial p'(Z;\boldsymbol{x},t)}{\partial t} + \boldsymbol{\nabla}\cdot(p'(Z;\boldsymbol{x},t)\boldsymbol{v}) = -\frac{\partial}{\partial Z}\left[p'(Z;\boldsymbol{x},t)\left(\frac{1}{\rho}\boldsymbol{\nabla}\cdot(\rho\kappa\boldsymbol{\nabla}C)+S\right)\right], \quad \text{(A 6)}$$

where we used the fact that, except $p'(Z;\boldsymbol{x},t)$ or $q'(Z;\boldsymbol{x},t)$, all the quantities in the right-hand-side term are independent of Z. Equation (A 6) is essentially a Liouville equation. In analogy to the kinetic theory, the concentration here corresponds to the particle momentum, and the equation $\mathrm{d}C/\mathrm{d}t = (1/\rho)\boldsymbol{\nabla}\cdot(\rho\kappa\boldsymbol{\nabla}C)+S$ corresponds to the particle equation of motion.

We now consider the coarse-grained p.d.f., defined as the ensemble average of the fine-grained p.d.f.s over independent realizations. The coarse-grained p.d.f.s with volume- and density-weighting are, respectively, $q(Z;\boldsymbol{x},t) = \langle q'(Z;\boldsymbol{x},t)\rangle$ and $p(Z;\boldsymbol{x},t) = \langle p'(Z;\boldsymbol{x},t)\rangle$, where $\langle\cdots\rangle$ denotes the ensemble average. The average is over different values of the concentration, $C(\boldsymbol{x},t)$, the flow velocity and density, $\boldsymbol{v}(\boldsymbol{x},t)$ and $\rho(\boldsymbol{x},t)$, and the scalar source, $S(\boldsymbol{x},t)$, in different realizations. From (A 6), it immediately follows that

$$\frac{\partial p(Z;\boldsymbol{x},t)}{\partial t} + \boldsymbol{\nabla}\cdot\langle p'(Z;\boldsymbol{x},t)\boldsymbol{v}\rangle = -\frac{\partial}{\partial Z}\left\langle p'(Z;\boldsymbol{x},t)\left(\frac{1}{\rho}\boldsymbol{\nabla}\cdot(\rho\kappa\boldsymbol{\nabla}C)+S\right)\right\rangle. \quad \text{(A 7)}$$

The ensemble-average terms in (A 7) can be expressed in more convenient forms. For any stochastic quantity $B(\boldsymbol{x},t)$, we have the following relation (see e.g. Pope 2000):

$$\langle q'(Z;\boldsymbol{x},t)B(\boldsymbol{x},t)\rangle = q(Z;\boldsymbol{x},t)\langle B(\boldsymbol{x},t)|C(\boldsymbol{x},t)=Z\rangle, \quad \text{(A 8)}$$

where $\langle B(\boldsymbol{x},t)|C(\boldsymbol{x},t)=Z\rangle$ denotes the ensemble average of $B(\boldsymbol{x},t)$ conditioned on $C(\boldsymbol{x},t)=Z$. The conditional mean appears here because the factor $q'(Z;\boldsymbol{x},t)$ selects only the realizations where $C(\boldsymbol{x},t)$ is equal to Z. Using (A 8), we have

$$\langle p'(Z;\boldsymbol{x},t)B(\boldsymbol{x},t)\rangle = p(Z;\boldsymbol{x},t)\frac{\langle\rho B|C(\boldsymbol{x},t)=Z\rangle}{\langle\rho|C(\boldsymbol{x},t)=Z\rangle}, \quad \text{(A 9)}$$

where we used $p(Z;\boldsymbol{x},t) = q(Z;\boldsymbol{x},t)\langle\tilde{\rho}|C(\boldsymbol{x},t)=Z\rangle$.

With (A 7) and (A 9), we arrive at the final equation for the coarse-grained p.d.f. with density weighting,

$$\begin{aligned}&\frac{\partial p(Z;\boldsymbol{x},t)}{\partial t} + \boldsymbol{\nabla}\cdot\left(p\frac{\langle\rho\boldsymbol{v}|C=Z\rangle}{\langle\rho|C=Z\rangle}\right)\\ &\quad= -\frac{\partial}{\partial Z}\left(p\frac{\langle\boldsymbol{\nabla}\cdot(\rho\kappa\boldsymbol{\nabla}C)|C=Z\rangle}{\langle\rho|C=Z\rangle}\right) - \frac{\partial}{\partial Z}\left(p\frac{\langle\rho S|C=Z\rangle}{\langle\rho|C=Z\rangle}\right),\end{aligned} \quad \text{(A 10)}$$

where we replaced the condition $C(\boldsymbol{x},t)=Z$ by $C=Z$ for simplicity of notation. Note that the advection term is in a divergence form, which is the motivation for the use of a density-weighting scheme in our derivation. A detailed physical discussion of all the terms in this equation is given in the text.

REFERENCES

ABEL, T., BRYAN, G. L. & NORMAN, M. L. 2000 The formation and fragmentation of primordial molecular clouds. *Astrophys. J.* **540**, 39–44.

BROMM, V., COPPI, P. S. & LARSON, R. B. 2002 The formation of the first stars. I. The primordial star-forming cloud. *Astrophys. J.* **564**, 23–51.

BROMM, V. & LOEB, A. 2003 The formation of the first low-mass stars from gas with low carbon and oxygen abundances. *Nature* **654**, 29–32.

CAFFAU, E., BONIFACIO, P., FRANCOIS, P., SBORDONE, L., MONACO, L., SPITE, M., SPITE, F., LUDWIG, H.-G., CAYREL, R., ZAGGIA, S., HAMMER, F., RANDICH, S., MOLARO, P. & HILL, V. 2011 An extremely primitive star in the Galactic halo. *Nature* **477**, 67–69.

CAYREL, R., DEPAGNE, E., SPITE, M., HILL, V., SPITE, F., FRANCOIS, P., PLEZ, B., BEERS, T., PRIMAS, F., ANDERSEN, J., BARBUY, B., BONIFACIO, P., MOLARO, P. & NORDSTROM, B. 2004 First stars V-abundance patterns from C to Zn and supernova yields in the early Galaxy. *Astron. Astrophys.* **416**, 1117–1138.

CHEN, H., CHEN, S. & KRAICHNAN, R. H. 1989 Probability distribution of a stochastically advected scalar field. *Phys. Rev. Lett.* **425**, 812–814.

COLELLA, P. & GLAZ, H. M. 1985 Efficient solution algorithms for the Riemann problem for real gases. *J. Comput. Phys.* **59**, 264–289.

COLELLA, P. & WOODWARD, P. R. 1984 The Piecewise Parabolic Method (PPM) for gas-dynamical simulations. *J. Comput. Phys.* **54**, 174–201.

CORRSIN, R. L. 1951 On the spectrum of isotropic temperature fluctuations in isotropic turbulence. *J. Appl. Phys.* **22**, 469–473.

CURL, S. 1963 Dispersed phase mixing: I. Theory and effects in simple reactors. *AIChE J.* **9**, 175–181.

DOPAZO, C. 1979 Relaxation of initial probability density functions in the turbulent convection of scalar fields. *Phys. Fluids* **22**, 20–30.

DOPAZO, C. & O'BRIEN, E. E. 1974 An approach to the autoignition of a turbulent mixture. *Acta Astron.* **1**, 1239–1266.

DOPAZO, C., VALINO, L. & FUEYO, N. 1997 Statistical description of the turbulent mixing of scalar fields. *Intl J. Mod. Phys.* B **11**, 2975–3014.

DUPLAT, J. & VILLERMAUX, E. 2008 Mixing by random stirring in confined mixtures. *J. Fluid Mech.* **617**, 51–86.

FOX, R. O. 2003 *Computational Models for Turbulent Reacting Flows*. Cambridge University Press.

FREBEL, A., COLLET, R., ERIKSSON, K., CHRISTLIEB, N. & AOKI, W. 2008 HE 1327-2326, an unevolved star with [Fe/H] < -5.0. II. New 3D-1D corrected abundances from a very large telescope UVES spectrum. *Astrophys. J.* **684**, 588–602.

FRYXELL, B., MÜLLER, E. & ARNETT, D. 1989 Computation of multi-dimensional flows with non-uniform composition. In *Proceedings of the 5th Workshop on Nuclear Astrophysics*, vol. 5. pp. 100–102. ARI.

FRYXELL, B., OLSON, K., RICKER, P., TIMMES, F. X., ZINGALE, M., LAMB, D. Q., MACNEICE, P., ROSNER, R., TRURAN, J. W. & TUFO, H. 2000 FLASH: An adaptive mesh hydrodynamics code for modeling astrophysical thermonuclear flashes. *Astrophys. J. Suppl. Ser.* **131**, 273–334.

GIRIMAJI, S. S. 1992 A mapping closure for turbulent scalar mixing using a time-evolving reference field. *Phys. Fluids* A **4**, 2875.

GREIF, T. H., JOHNSON, J. L., KLESSEN, R. S. & BROMM, V. 2008 The first galaxies: assembly, cooling and the onset of turbulence. *Mon. Not. R. Astron. Soc.* **387**, 1021–1036.

HAWORTH, D. C. 2010 Progress in probability density function methods for turbulent reacting flows. *Prog. Energy Combust. Sci.* **36**, 168–259.

HE, G.-W. & ZHANG, Z.-F. 2004 Two-point closure strategy in the mapping closure approximation approach. *Phys. Rev.* E **70**, 036309.

IEVLEV, V. M. 1973 Equations for the finite-dimensional probability distributions of pulsating variables in a turbulent flow. *Dokl. Akad. Nauk SSSR* **208**, 1044–1047.

JANICKA, J., KOLBE, W. & KOLLMANN, W. 1979 Closure of the transport equation for the probability density funcfion of turbulent scalar fields. *J. Non-Equilib. Thermodyn.* **4**, 47–66.

JIMENEZ, R. & HAIMAN, Z. 2006 Significant primordial star formation at redshifts $z \sim 3 - 4$. *Nature* **440**, 501–504.

KOLLMANN, W. 1990 The pdf approach to turbulent flow. *Theor. Comput. Fluid Dyn.* **1**, 249–285.

LUNDGREN, T. S. 1967 Distribution function in the statistical theory of turbulence. *Phys. Fluids* **12**, 485–497.

MONIN, A. S. 1967 Equations for finte dimensional probability distributions of a field of turbulence. *Dokl. Akad. Nauk SSSR* **177**, 1036–1038.

NAGAO, T., SASAKI, S. S., MAIOLINO, R., GRADY, C., KASHIKAWA, N., LY, C., MALKAN, M. A., MOTOHARA, K., MURAYAMA, T., SCHAERER, D., SHIOYA, Y. & TANIGUCHI, Y. 2008 A photometric survey for Lyα-He II dual emitters: searching for Population III stars in high-redshift galaxies. *Astrophys. J.* **680**, 100–109.

NOVIKOV, E. A. 1967 Kinetic equations for a vortex field. *Dokl. Akad. Nauk SSSR* **177**, 299–301.

OBRIEN, E. E. 1980 The probability density function (PDF) approach to reacting turbulent flows. In *Turbulent Reacting Flows* (ed. P. A. Libby & F. A. Williams). pp. 185–218. Springer.

OBUKHOV, A. M. 1949 The structure of the temperature eld in a turbulent flow. *Izv. Akad. Nauk SSSR Ser. Geogr. Geophys.* **13**, 58–69.

OMUKAI, K., TSURIBE, T., SCHNEIDER, R. & FERRARA, A. 2005 Thermal and fragmentation properties of star-forming slouds in low-metallicity environments. *Astrophys. J.* **626**, 627–643.

PAN, L. & SCALO, J. 2007 Mixing of primordial gas in Lyman Break Galaxies. *Astrophys. J. Lett.* **654**, 29–32.

PAN, L. & SCANNAPIECO, E. 2010 Mixing in supersonic turbulence. *Astrophys. J.* **721**, 1765–1782.

PAN, L. & SCANNAPIECO, E. 2011 Passive scalar structures in supersonic turbulence. *Phys. Rev.* E **83**, 045302.

POPE, S. B. 1976 The probability approach to the modelling of turbulent reacting flows. *Combust. Flame* **27**, 299–312.

POPE, S. B. 1982 An improved turbulent mixing model. *Combust. Sci. Technol.* **28**, 131–145.

POPE, S. B. 1985 PDF methods for turbulent reactive flows. *Prog. Energy Combust. Sci.* **11**, 119–192.

POPE, S. B. 1991 Mapping closures for turbulent mixing and reaction. *Theor. Comput. Fluid Dyn.* **2**, 255–270.

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

PUMIR, A., SHRAIMAN, B. I. & SIGGIA, E. D. 1991 Exponential tails and random advection. *Phys. Rev. Lett.* **66** (23), 2984–2987.

SCANNAPIECO, E., MADAU, P., WOOSLEY, S., HEGER, A. & FERRARA, A. 2005 The detectability of pair-production supernovae at $z \lesssim 6$. *Astrophys. J.* **633**, 1031–1041.

SCANNAPIECO, E., SCHNEIDER, R. & FERRARA, A. 2003 The detectability of the first stars and their cluster enrichment signatures. *Astrophys. J.* **589**, 35–52.

SCHAERER, D. 2002 On the properties of massive Population III stars and metal-free stellar populations. *Astron. Astrophys.* **382**, 28–42.

VALINO, L. & DOPAZO, C. 1990 A binomial sampling model for scalar turbulent mixing. *Phys. Fluids* A **2**, 1204–1212.

VENAILLE, A. & SOMMERIA, J. 2007 A dynamical equation for the distribution of a scalar advected by turbulence. *Phys. Fluids* **19**, 028101.

VENAILLE, A. & SOMMERIA, J. 2008 Is turbulent mixing a self-convolution process? *Phys. Rev. Lett.* **100**, 234506.

VEYNANTE, D. & VERVISCH, L. 2002 Turbulent combustion modelling. *Prog. Energy Combust. Sci.* **28**, 193–266.

VILLERMAUX, E. & DUPLAT, J. 2003 Mixing as an aggregation process? *Phys. Rev. Lett.* **91**, 184501.

VILLERMAUX, E., STROOCK, A. D. & STONE, H. A. 2008 Bridging kinematics and concentration content in a chaotic micromixer. *Phys. Rev.* E **77**, 015301(R).

WALKER, T. P., STEIGMAN, G., KANG, H.-S., SCHRAMM, D. M. & OLIVE, K. A. 1991 Primordial nucleosynthesis redux. *Astrophys. J.* **376**, 51–69.

WISE, J. H., TURK, M. J. & ABEL, T. 2008 Resolving the formation of protogalaxies. II. Central gravitational collapse. *Astrophys. J.* **682**, 745–757.