J. Fluid Mech. (2017), *vol.* 829, *pp.* 153–189. © Cambridge University Press 2017
doi:10.1017/jfm.2017.567

A Lagrangian fluctuation–dissipation relation for scalar turbulence. Part I. Flows with no bounding walls

Theodore D. Drivas[1,†] **and Gregory L. Eyink**[1,2]

[1]Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD 21218, USA

[2]Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA

(Received 2 June 2016; revised 13 July 2017; accepted 9 August 2017;
first published online 14 September 2017)

An exact relation is derived between scalar dissipation due to molecular diffusivity and the randomness of stochastic Lagrangian trajectories for flows without bounding walls. This 'Lagrangian fluctuation–dissipation relation' equates the scalar dissipation for either passive or active scalars to the variance of scalar inputs associated with initial scalar values and internal scalar sources, as these are sampled backward in time by the stochastic Lagrangian trajectories. As an important application, we reconsider the phenomenon of 'Lagrangian spontaneous stochasticity' or persistent non-determinism of Lagrangian particle trajectories in the limit of vanishing viscosity and diffusivity. Previous work on the Kraichnan (*Phys. Fluids*, 1968, vol. 11, pp. 945–953) model of turbulent scalar advection has shown that anomalous scalar dissipation is associated in that model with Lagrangian spontaneous stochasticity. There has been controversy, however, regarding the validity of this mechanism for scalars advected by an actual turbulent flow. We here completely resolve this controversy by exploiting the fluctuation–dissipation relation. For either a passive or an active scalar advected by any divergence-free velocity field, including solutions of the incompressible Navier–Stokes equation, and away from walls, we prove that anomalous scalar dissipation requires Lagrangian spontaneous stochasticity. For passive scalars, we prove furthermore that spontaneous stochasticity yields anomalous dissipation for suitable initial scalar fields, so that the two phenomena are there completely equivalent. These points are illustrated by numerical results from a database of homogeneous isotropic turbulence, which provide both additional support to the results and physical insight into the representation of diffusive effects by stochastic Lagrangian particle trajectories.

Key words: turbulent flows, turbulent mixing, turbulence theory

1. Introduction

A fundamental feature of turbulent flows is the enhanced dissipation of kinetic energy. It was suggested by Taylor (1917) that kinetic energy can 'be dissipated in fluid of infinitesimal viscosity'. This idea that turbulent dissipation might become

† Email address for correspondence: tdrivas2@jhu.edu

independent of molecular viscosity at sufficiently high Reynolds numbers was pursued further by Kolmogorov (1941*a*,*b*,*c*) and Onsager (1945, 1949) in developing their theories of turbulence. The physical phenomenon is sometimes called the turbulent 'dissipative anomaly' or even the 'zeroth law of turbulence', although it is, of course, no 'law' but rather an experimentally observed phenomenon which is still only partially understood theoretically. For current empirical evidence from laboratory experiments and numerical simulations, see, e.g., Sreenivasan (1998), Pearson, Krogstad & van de Water (2002) and Kaneda *et al.* (2003), whose data are all consistent with energy dissipation in the bulk of turbulent flows being essentially independent of viscosity. Similar phenomena are expected for other turbulent systems, in particular for scalar fields advected by a turbulent fluid, such as concentrations of dyes or aerosols, temperature fluctuations, etc. It was suggested by Taylor (1922) that diffusion by turbulence should depend 'little on the molecular conductivity and viscosity of the fluid', and the asymptotic independence of the dissipation rate of scalar fluctuations from the molecular transport coefficients was a fundamental assumption in the Kolmogorov-style theories of scalar turbulence developed by Obukhov (1949) and Corrsin (1951). A very comprehensive review of the empirical evidence for this hypothesis on scalar dissipation is contained in the paper of Donzis, Sreenivasan & Yeung (2005), whose compilation of data is again consistent with scalar dissipation in the bulk of turbulent flows being insensibly dependent on molecular transport coefficients at sufficiently high Reynolds and Péclet numbers. This phenomenon still requires a complete theoretical explanation.

Fundamental new ideas on the Lagrangian origin of turbulent scalar dissipation arose from mathematical work of Bernard, Gawędzki & Kupiainen (1998), which was carried out in the Kraichnan (1968) model of synthetic turbulence. In this model, the advecting velocity is a Gaussian random field that has Kolmogorov-type scaling of increments in space but a white-noise correlation in time. It was shown in the Kraichnan model that the dissipative anomaly for a decaying passive scalar is due to a remarkable phenomenon called spontaneous stochasticity (Chaves *et al.* 2003). Simply stated, Bernard *et al.* (1998) showed that Lagrangian particle trajectories become non-unique and stochastic in the infinite-Reynolds-number limit for a fixed initial particle position and a fixed velocity realization, due to the spatial roughness of the advecting velocity field. More precisely, they showed that at very large Reynolds and Péclet numbers, when the velocity field is smooth but approximates a 'rough' field over a long range of scales, small stochastic perturbations on Lagrangian trajectories due to molecular diffusivity lead to persistent randomness over any finite times even as the perturbations vanish. This effect is due to the explosive (superballistic) dispersion of particle pairs in a turbulent flow predicted by Richardson (1926), which leads to loss of memory of initial particle separations or of amplitudes of stochastic perturbations. For excellent reviews of this and related studies on the Kraichnan model, see Falkovich, Gawędzki & Vergassola (2001), Kupiainen (2003) and Gawędzki (2008).

Since this pioneering work, however, recurrent doubts have been expressed concerning the validity of these results for real hydrodynamic turbulence. For example, Tsinober (2009, § 5.4.5) has argued that in real fluids 'The flow field is smooth. In such flows "phenomena" like "spontaneous stochasticity" and "breakdown of Lagrangian flow" do not arise and one has to look at different more realistic possibilities.' This is a simple misunderstanding, because spontaneous stochasticity is a phenomenon that appears for smooth velocity fields that merely appear 'rough'

over a long range of scales. More serious questions have been raised concerning the approximation of a white-noise temporal correlation in the Kraichnan model. In a recent detailed comparison of passive scalars in the Kraichnan model and in fluid turbulence, Sreenivasan & Schumacher (2010) have remarked that 'It is still unclear in the Kraichnan model as to which qualitative and quantitative differences arise from the finite-time correlation of the advecting flow.' This latter paper also discussed some of the challenges in extending results for the Kraichnan model to an understanding of the energy cascade in Navier–Stokes turbulence.

The principal contribution of the present paper is a new approach to the theory of turbulent scalar dissipation based upon an exact fluctuation–dissipation relation for scalars. Our new relation expresses an equality between the time-averaged scalar dissipation and the input of scalar variance from the initial data and interior scalar sources, as these are sampled by stochastic Lagrangian trajectories. This relation makes it intuitively clear that scalar dissipation requires non-vanishing Lagrangian stochasticity. In fact, using our new relation, we can prove the following fact rigorously. Away from boundaries and for any advecting velocity field whatsoever, spontaneous stochasticity of Lagrangian particle trajectories is sufficient for anomalous dissipation of passive scalars, and necessary for anomalous dissipation of both passive and active scalars. Thus, there is no possible mechanism for a scalar dissipative anomaly in such situations other than spontaneous stochasticity. In this way, we completely resolve the controversies on the applicability of the dissipation mechanisms in the Kraichnan model to scalars in hydrodynamic turbulence, at least away from walls. The importance of our exact fluctuation–dissipation relation (FDR) is not limited to analysis of anomalous scalar dissipation and it is valid even when scalar dissipation may vanish as $\nu, \kappa \to 0$. In general, our relation gives a new Lagrangian viewpoint on dissipation of scalars, both active and passive. As such, it generalizes some previously derived relations, such as that of Sawford, Yeung & Borgas (2005) and Buaria, Yeung & Sawford (2016) for scalars forced by a mean scalar gradient and the exact balance relations for stochastic scalar sources which are Gaussian white in time (Novikov 1965). In two companion papers (Drivas & Eyink 2017; Eyink & Drivas 2017; hereafter denoted Parts II and III), we show how the FDR extends also to wall-bounded domains, with either fixed-scalar (Dirichlet) or fixed-flux (Neumann) conditions for the scalar field, and we apply the FDR to the concrete problem of Nusselt–Rayleigh scaling in turbulent Rayleigh–Bénard convection.

The detailed contents of the present paper are as follows. In § 2, we first derive the stochastic representation of scalar advection and our FDR in the case of flows in domains without walls. In § 3, we review the notion of spontaneous stochasticity, with numerical verifications from a database of homogeneous isotropic turbulence. In § 4, we establish the connection of spontaneous stochasticity with anomalous scalar dissipation. In the summary and discussion § 5 we discuss both the implications for turbulent vortex dynamics and other Lagrangian aspects of turbulence, and also the outstanding challenges, including that of relating spontaneous stochasticity to anomalous dissipation of kinetic energy. Three appendices give further details, including rigorous mathematical proofs of all of the results in the main text. These deal with the connection between spontaneous stochasticity and anomalous dissipation (appendix A), the relation of our scalar FDR to previous results in the literature (appendix B) and discussion of numerical methods employed (appendix C).

2. Lagrangian FDR

We consider in this paper turbulent fluid flows in finite domains without walls. A relevant numerical example is direct numerical simulation (DNS) of turbulence in a periodic box. A set of examples from nature is provided by large-scale flows in thin planetary atmospheres, which can be modelled as 2D flows on a sphere. Mathematically speaking, the results in this section apply to fluid flows on any compact Riemannian manifold without boundary, merely replacing the Wiener process with the Brownian motion on the manifold whose infinitesimal generator is the Laplace–Beltrami operator (Ikeda & Watanabe 1989). For simplicity of presentation, we derive the relation only for periodic domains.

Scalar fields θ (such as temperature, dye or pollutants) transported by a fluid with velocity $\boldsymbol{u}$ are described by the advection–diffusion equation

$$\partial_t\theta + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta = \kappa\Delta\theta + S, \tag{2.1}$$

where $S(\boldsymbol{x}, t)$ is a source field and $\kappa > 0$ is the molecular diffusivity of the scalar. The work of Bernard *et al.* (1998) employed a stochastic representation of the solutions of this equation, which is known as the Feynman–Kac representation in the mathematics literature (Oksendal 2013) and as a stochastic Lagrangian representation in the turbulence modelling field (Sawford 2001). This stochastic approach is the natural extension to diffusive scalars of the Lagrangian description developed for 'ideal' scalar fields without diffusion advected by smooth velocities. We presently discuss this representation only for domains Ω without boundaries, as assumed also by Bernard *et al.* (1998), and in Part II we describe the extension to wall-bounded domains. We shall further discuss in these papers only advection by an incompressible fluid satisfying

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{2.2}$$

so that the ideal advection term formally conserves all integrals of the form $I_h(t) = \int_\Omega \mathrm{d}^d x\ h(\theta(\boldsymbol{x}, t))$ for any continuous function $h(\theta)$. It should be noted that the representation applies in any space dimension d, with most immediate physical interest for $d = 2, 3$, of course.

The stochastic representation of non-ideal scalar dynamics involves stochastic Lagrangian flow maps $\widetilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,s}(\boldsymbol{x})$ describing the motion of particles labelled by their positions $\boldsymbol{x}$ at time t to random positions at earlier times $s < t$. The physical relevance of the backward-in-time particle trajectories can be anticipated from the fact that the advection–diffusion equation (2.1) mixes (averages) the values of the scalar field given in the past and not, of course, the future values. Mathematically, the relevant stochastic flows are governed by the backward Itô stochastic differential equations,

$$\hat{\mathrm{d}}_s\widetilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,s}(\boldsymbol{x}) = \boldsymbol{u}^\nu(\widetilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,s}(\boldsymbol{x}), s)\mathrm{d}s + \sqrt{2\kappa}\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s, \quad \widetilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,t}(\boldsymbol{x}) = \boldsymbol{x}. \tag{2.3}$$

Here, $\widetilde{\boldsymbol{W}}_s$ is a standard Brownian motion and $\hat{\mathrm{d}}_s$ denotes the backward Itô stochastic differential in the time s. For detailed discussions of backward Itô equations and stochastic flows, see Kunita (1997) and Friedman (2006). For those who are familiar with the more standard forward Itô equations, the backward equations are simply the time reverses of the forward ones. Thus, a backward Itô equation in the time variable s is equivalent to a forward Itô equation in the time $\hat{s} = t_r - s$ reflected around a chosen reference time t_r. (We note that the difference between forward and backward Itô equations is not essentially the direction of time in which they

are integrated. Rather, the difference has to do with the time direction in which those equations are adapted (Kunita 1997; Friedman 2006). Thus, a forward Itô differential $b(\widetilde{\boldsymbol{W}}_t)\,\mathrm{d}\widetilde{\boldsymbol{W}}_t$ is discretized in time as $b(\widetilde{\boldsymbol{W}}_{t_n})(\widetilde{\boldsymbol{W}}_{t_{n+1}} - \widetilde{\boldsymbol{W}}_{t_n})$ for $t_{n+1} > t_n$, with the increment $\widetilde{\boldsymbol{W}}_{t_{n+1}} - \widetilde{\boldsymbol{W}}_{t_n}$ statistically independent of $\widetilde{\boldsymbol{W}}_t$ for $t \leqslant t_n$. Instead, a backward Itô differential $b(\widetilde{\boldsymbol{W}}_t)\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_t$ is discretized as $b(\widetilde{\boldsymbol{W}}_{t_n})(\widetilde{\boldsymbol{W}}_{t_n} - \widetilde{\boldsymbol{W}}_{t_{n-1}})$ for $t_n > t_{n-1}$, with $\widetilde{\boldsymbol{W}}_{t_n} - \widetilde{\boldsymbol{W}}_{t_{n-1}}$ statistically independent of $\widetilde{\boldsymbol{W}}_t$ for $t \geqslant t_n$. The distinction only matters when, as in our (2.4), the differential of $\widetilde{\boldsymbol{W}}_s$ is multiplied by a stochastic function of $\widetilde{\boldsymbol{W}}$.) The noise term involving the Brownian motion in (2.3) is proportional to the square root of the molecular diffusivity κ. The velocity field $\boldsymbol{u}^\nu$ is assumed to be smooth so long as the parameter $\nu > 0$. In the case of greatest physical interest when $\boldsymbol{u}^\nu$ is a solution of the incompressible Navier–Stokes equation, then ν represents the kinematic viscosity and we assume, for simplicity of presentation, that there is no blowup in these solutions. (See Rezakhanlou (2014) for weak solutions.) Because (2.3) involves both ν and κ, its random solutions $\tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,s}$ have statistics that depend upon these parameters, represented by the superscripts. To avoid too heavy a notation, we omit these superscripts and write simply $\tilde{\boldsymbol{\xi}}_{t,s}$ unless it is essential to refer to the dependence upon ν, κ. It should be noted that when $\kappa = 0$ and $\boldsymbol{u}^\nu$ remains smooth, then $\boldsymbol{\xi}^{\nu,0}_{t,s}(\boldsymbol{x})$ is no longer stochastic and gives the usual reverse Lagrangian flow from time t backward to the earlier time $s < t$.

The stochastic representation of the solutions of the advection–diffusion equation follows from the backward differential

$$\begin{aligned}\hat{\mathrm{d}}_s\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s) &= [(\partial_s + \boldsymbol{u}^\nu\cdot\nabla - \kappa\Delta)\theta](\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s + \sqrt{2\kappa}\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s\cdot\nabla\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\\ &= S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s + \sqrt{2\kappa}\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s\cdot\nabla\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s),\end{aligned} \tag{2.4}$$

using the backward Itô formula (Kunita 1997; Friedman 2006) in the first line and (2.1) in the second. Integration over time s from 0 to t gives

$$\theta(\boldsymbol{x}, t) = \theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})) + \int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s + \sqrt{2\kappa}\int_0^t \hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s\cdot\nabla\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s), \tag{2.5}$$

where θ_0 represents the initial data for the scalar at time 0. Because the backward Itô integral term in (2.5) averages to zero, one obtains

$$\theta(\boldsymbol{x}, t) = \mathbb{E}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})) + \int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right], \tag{2.6}$$

where $\mathbb{E}$ denotes the average over the Brownian motion. Equation (2.6) is the desired stochastic representation of the solution of the advection–diffusion equation (2.1). It should be noted that the reverse statement is also true, that the field $\theta(\boldsymbol{x}, t)$ defined *a priori* by (2.6) is the solution of (2.1) for the initial data θ_0. For a simple proof, see Eyink & Drivas (2015*b*, § 4.1), which gives the analogous argument for the Burgers equation.

To see that this stochastic representation naturally generalizes the standard Lagrangian description to non-ideal scalars, we observe that the scalar values along stochastic Lagrangian trajectories $\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)$ are, for $S \equiv 0$, martingales backward in time. This means that

$$\mathbb{E}[\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)|\{\widetilde{\boldsymbol{W}}_\tau, r < \tau < t\}] = \theta(\tilde{\boldsymbol{\xi}}_{t,r}(\boldsymbol{x}), r), \quad s < r < t, \tag{2.7}$$

where the expectation is conditioned upon knowledge of the Brownian motion over the time interval $[r, t]$. Thus, the conditional average value is the last known value (going backward in time). This is the property for diffusive flow, which corresponds to the statement for diffusionless smooth advection that θ is conserved along Lagrangian trajectories, or that $\theta(\boldsymbol{\xi}_{t,s}(\boldsymbol{x}), s)$ is constant in s. The proof is obtained by integrating the differential (2.4) over the time interval $[s, t]$ to obtain

$$\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s) = \theta(\boldsymbol{x}, t) - \sqrt{2\kappa} \int_s^t \hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_\tau \cdot \boldsymbol{\nabla}\theta(\tilde{\boldsymbol{\xi}}_{t,\tau}(\boldsymbol{x}), \tau) \tag{2.8}$$

and then exploiting the corresponding martingale property of the backward Itô integral (Kunita 1997; Friedman 2006). It is important to emphasize that a martingale property like (2.7) does not hold forward in time, which would instead give a solution of the negative-diffusion equation with κ replaced by $-\kappa < 0$. Thus, the backward-in-time martingale property (2.7) expresses the arrow of time arising from the irreversibility of the diffusion process.

The main result of this paper is a new exact FDR between scalar dissipation due to molecular diffusivity and fluctuations associated with stochastic Lagrangian trajectories. To state the result, we introduce a stochastic scalar field

$$\tilde{\theta}(\boldsymbol{x}, t) \equiv \theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})) + \int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\, \mathrm{d}s, \tag{2.9}$$

which, according to (2.6), satisfies $\theta(\boldsymbol{x}, t) = \mathbb{E}[\tilde{\theta}(\boldsymbol{x}, t)]$ when averaged over Brownian motions. (It should be noted that, for all $s < t$, the quantity $\tilde{\theta}(\boldsymbol{x}, t; s) = \theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s) + \int_s^t S(\tilde{\boldsymbol{\xi}}_{t,r}(\boldsymbol{x}), r)\, \mathrm{d}r$ is a martingale backward in time, by the same argument as used above for $S = 0$.) Thus, $\tilde{\theta}(\boldsymbol{x}, t)$ in (2.9) represents the contribution to $\theta(\boldsymbol{x}, t)$ from an individual stochastic Lagrangian trajectory as it samples the initial data θ_0 and scalar source S backward in time. Using this definition and (2.6), we can rewrite (2.5) as

$$\tilde{\theta}(\boldsymbol{x}, t) - \mathbb{E}[\tilde{\theta}(\boldsymbol{x}, t)] = -\sqrt{2\kappa} \int_0^t \hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s \cdot \boldsymbol{\nabla}\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s). \tag{2.10}$$

Squaring this equation and averaging over the Brownian motion gives

$$\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)] = 2\kappa \int_0^t \mathrm{d}s\, \mathbb{E}[|\boldsymbol{\nabla}\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)|^2], \tag{2.11}$$

where 'Var' on the left-hand side denotes the stochastic scalar variance in the average over the Brownian motion, and on the right-hand side we have used the Itô isometry (see Oksendal 2013, § 3.1) to evaluate the mean square of the backward Itô integral. If we now average in $\boldsymbol{x}$ over the flow domain Ω, use the fact that the stochastic flows $\tilde{\boldsymbol{\xi}}_{t,s}$ with condition (2.2) preserve volume, and divide by $1/2$ we obtain

$$\frac{1}{2}\langle \mathrm{Var}\, \tilde{\theta}(t)\rangle_\Omega = \kappa \int_0^t \mathrm{d}s \langle |\boldsymbol{\nabla}\theta(s)|^2\rangle_\Omega. \tag{2.12}$$

This is our exact FDR. The quantity on the right-hand side is just the volume-averaged and cumulative (time-integrated) scalar dissipation and the quantity on the left-hand side is (half) the stochastic scalar variance. The relation (2.12) thus represents a

balance between scalar dissipation and the input of scalar fluctuations from the initial scalar field and the scalar sources, as sampled by stochastic Lagrangian trajectories backward in time.

It is important to emphasize that the origin of statistical fluctuations in our relation (2.12) is not that assumed in most traditional discussions of turbulence, i.e. random ensembles of initial scalar fields, of advecting velocity fields or of stochastic scalar sources. Our FDR (2.12) is valid for fixed realizations of all of these quantities. The fluctuating quantity $\tilde{\theta}(\boldsymbol{x}, t)$ which is defined in (2.9) and which appears in our (2.12) is an entirely different object from the conventional 'turbulent' scalar fluctuation $\theta'(\boldsymbol{x}, t)$. The latter is usually defined by $\theta' := \theta - \langle\theta\rangle$, where the scalar mean $\langle\theta\rangle$ is taken to be an ensemble or space/time average. Instead, the origin of randomness in $\tilde{\theta}(\boldsymbol{x}, t)$ is the Brownian motion in the stochastic flow equation (2.3). In special cases, e.g. a dye passively advected by a turbulent flow, this mathematical Wiener process has direct significance as the description of a physical Brownian motion of individual dye molecules in the liquid (Saffman 1960; Buaria *et al.* 2016). In general, however, the Wiener process is simply a means to model the effects of diffusion in a Lagrangian framework. For example, for a temperature field, there are no 'thermal molecules' undergoing physical Brownian motion.

Because our FDR is valid for fixed realizations of initial scalar fields, of advecting velocity fields or of scalar sources, we are free to average subsequently over random ensembles of these objects. In this manner, we recover from (2.12) as special cases some known results. For example, when the scalar source is a random field with zero mean and delta-correlated in time,

$$\langle\tilde{S}(\boldsymbol{x}, t)\tilde{S}(\boldsymbol{x}', t')\rangle = 2C_S(\boldsymbol{x}, \boldsymbol{x}')\delta(t - t'), \tag{2.13}$$

then we recover the steady-state balance equation for the scalar dissipation,

$$\langle\kappa|\boldsymbol{\nabla}\theta|^2\rangle_{\Omega,\infty,S} = \frac{1}{V}\int_{\Omega} \mathrm{d}^d x\, C_S(\boldsymbol{x}, \boldsymbol{x}), \tag{2.14}$$

where the average on the left-hand side is over space domain Ω, an infinite time interval and the random source $\tilde{S}$. This is the standard result usually derived for Gaussian random source fields as an application of the Furutsu–Donsker–Novikov theorem (Novikov 1965; Frisch 1995). We derive it instead as a consequence of a general steady-state FDR,

$$\langle\kappa|\boldsymbol{\nabla}\theta|^2\rangle_{\Omega,\infty} = \int_{-\infty}^{0} \mathrm{d}t\langle\langle\tilde{S}_L(0)\tilde{S}_L(t)\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}\rangle_{\infty}, \tag{2.15}$$

where the random variable $\tilde{S}_L(\boldsymbol{x}, s) = S(\tilde{\boldsymbol{\xi}}_{0,s}(\boldsymbol{x}), s)$ arises by sampling a single realization of the source S along stochastic Lagrangian trajectories, $\langle\cdot\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}$ denotes the truncated correlation function (covariance) in the average over Brownian motion and space domain, and $\langle\cdot\rangle_{\infty}$ an infinite-time average with respect to the release time 0 of stochastic particles. Further averaging of (2.15) over random ensembles of S with delta covariance (2.13) then gives the steady-state balance (2.14). For details, see appendix B. Similar relations hold for freely decaying scalars with no sources but random initial scalar fields. For example, when the initial scalar has a uniform random space gradient, $\tilde{\theta}_0(\boldsymbol{x}) = \tilde{\boldsymbol{G}}\cdot\boldsymbol{x}$ with isotropic statistics

$$\langle\tilde{\boldsymbol{G}}\tilde{\boldsymbol{G}}^{\mathrm{T}}\rangle_G = G^2\boldsymbol{I}, \tag{2.16}$$

then we recover a relation of Sawford *et al.* (2005) and Buaria *et al.* (2016),

$$\kappa \int_0^t \mathrm{d}s \langle |\nabla\theta(s)|^2\rangle_{\Omega,\theta_0} = \frac{1}{4} G^2 \mathbb{E}^{1,2} \langle |\tilde{\boldsymbol{\xi}}^{(1)}_{t,0} - \tilde{\boldsymbol{\xi}}^{(2)}_{t,0}|^2\rangle_{\Omega}, \tag{2.17}$$

where the 1, 2 averages are taken over two independent ensembles of Brownian motion. We delay the derivation of the special cases (2.14), (2.15), (2.17) to appendix B, since the proofs require additional material which will be introduced in subsequent sections.

It should be noted, finally, that the result (2.11) provides a spatially local FDR, which we may write in the form

$$\frac{1}{2t}\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)] = \langle \mathbb{E}[\kappa |\nabla\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)|^2]\rangle_t, \tag{2.18}$$

where, on the right-hand side, $\langle\cdot\rangle_t$ denotes an average over s in the time interval $[0, t]$, carried out along stochastic Lagrangian trajectories moving backward in time from space–time point $(\boldsymbol{x}, t)$. It follows that at short times the local scalar variance exactly recovers the local scalar dissipation,

$$\lim_{t\to 0} \frac{1}{2t}\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)] = \kappa |\nabla\theta(\boldsymbol{x}, 0)|^2. \tag{2.19}$$

A substantial spatial correlation between $(1/2t)\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)]$ and $\varepsilon_\theta(\boldsymbol{x}, t) = \kappa|\nabla\theta(\boldsymbol{x}, t)|^2$ should persist for relatively short times t. On the other hand, in the long-time limit, the local scalar variance becomes space–time-independent and equals

$$\lim_{t\to\infty} \frac{1}{2t}\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)] = \langle \kappa |\nabla\theta|^2\rangle_{\Omega,\infty} \quad \text{for all } \boldsymbol{x} \in \Omega. \tag{2.20}$$

To see that (2.20) should be true, one can note that the random variables $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}) \in \Omega$ for each fixed $\boldsymbol{x}$ are ergodic random processes in the time variable s for $\kappa > 0$. Because of incompressibility of the velocity field and the ergodicity of the stochastic Lagrangian flow, the variables $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})$ will be nearly uniformly distributed over Ω at times $s \leqslant t - \tau$, where τ is a characteristic scalar mixing time. This time τ will be at most of the order L^2/κ, where L is the diameter of the domain, and thus finite for $\kappa > 0$, but usually much shorter because of advective mixing by the velocity field. For any positive integer n,

$$\lim_{t\to\infty} \frac{1}{2t}\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)] = \lim_{t\to\infty} \frac{1}{t} \int_0^{t-n\tau} \mathrm{d}s\, \mathbb{E}[\kappa |\nabla\theta(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)|^2], \tag{2.21}$$

since the corrections are vanishing as $O(n\tau/t)$. By choosing an n sufficiently large but fixed as $t \to \infty$, we can make the right-hand side arbitrarily close to

$$\lim_{t\to\infty} \frac{1}{t - n\tau} \int_0^{t-n\tau} \mathrm{d}s \langle \kappa |\nabla\theta(\boldsymbol{\xi}, s)|^2\rangle_{\Omega} = \langle \kappa |\nabla\theta(\boldsymbol{\xi}, s)|^2\rangle_{\Omega,\infty}, \tag{2.22}$$

where the space–time average $\langle\cdot\rangle_{\Omega,\infty}$ on the right-hand side is over $\boldsymbol{\xi} \in \Omega$ and $s \in [0, \infty)$. Since $\lim_{t\to\infty}(1/2t)\mathrm{Var}[\tilde{\theta}(\boldsymbol{x}, t)]$ is independent of the choice of n, we obtain (2.20). Of course, here we have assumed that all of the various infinite-time averages

exist, as they shall (at least along subsequences of times $t_k \to \infty$) if the space-averaged scalar dissipation remains a bounded function of time.

It should be noted that if the scalar is freely decaying from bounded initial data θ_0, then the variance on the left-hand side of (2.20) is also bounded. In this case, the long-time-averaged scalar dissipation rate tends to zero, which comes as no surprise. In order to have a non-vanishing long-time dissipation, the scalar must be continually supplied to the system so that the variance of $\tilde{\theta}(\boldsymbol{x}, t)$ grows linearly in time. For example, a scalar source $S(\boldsymbol{x}, t)$ within the flow domain can provide the necessary scalar input. In such a case, the variance of $\tilde{\theta}(\boldsymbol{x}, t)$ grows proportionally to time t at long times because of the cumulative contribution from the scalar source S in the time integral $\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}, s)\,\mathrm{d}s$, and the long-time average scalar dissipation rate matches the mean input rate of the scalar. In fact, the expression on the right-hand side of (2.15) arises after dividing by t the variance of this time integral of S and then taking the limit $t \to \infty$. The linear growth of the variance and the expression in (2.15) are central-limit-theorem results based on the statistical independence of the flow maps $\tilde{\boldsymbol{\xi}}_{t,s}$ over widely separated intervals of time $[t, s]$.

3. Spontaneous stochasticity of Lagrangian trajectories

We now specialize in this section to the sourceless case $S \equiv 0$, in order to make contact with the work of Bernard *et al.* (1998) on spontaneous stochasticity and anomalous scalar dissipation. The stochastic representation (2.6) simplifies in this case to

$$\theta(\boldsymbol{x}, t) = \mathbb{E}[\theta_0(\tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,0}(\boldsymbol{x}))] = \int \mathrm{d}^d x_0\, \theta_0(\boldsymbol{x}_0) p^{\nu,\kappa}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{3.1}$$

where we have introduced the backward-in-time transition probability

$$p^{\nu,\kappa}(\boldsymbol{x}', t'|\boldsymbol{x}, t) = \mathbb{E}[\delta^d(\boldsymbol{x}' - \tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,t'}(\boldsymbol{x}))], \quad t' < t \tag{3.2}$$

for the stochastic flow. As already noted, the stochastic flow preserves volume when the velocity field is divergence-free. In terms of the transition probability, this means that

$$\int p^{\nu,\kappa}(\boldsymbol{x}', t'|\boldsymbol{x}, t)\,\mathrm{d}^d x = 1, \tag{3.3}$$

where $\det(\partial \tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,t'}(\boldsymbol{x})/\partial \boldsymbol{x}) = 1$ is used to write $\delta^d(\boldsymbol{x}' - \tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,t'}(\boldsymbol{x})) = \delta^d(\boldsymbol{x} - (\tilde{\boldsymbol{\xi}}^{\nu,\kappa}_{t,t'})^{-1}(\boldsymbol{x}'))$ and perform the integral over $\boldsymbol{x}$. It should be noted that if the limit $\kappa \to 0$ is taken with ν fixed (infinite-Prandtl-number limit), then the stochastic flow (3.2) becomes deterministic and

$$p^{\nu,0}(\boldsymbol{x}', t'|\boldsymbol{x}, t) = \delta^d(\boldsymbol{x}' - \boldsymbol{\xi}^{\nu,0}_{t,t'}(\boldsymbol{x})), \quad t' < t, \tag{3.4}$$

which corresponds to a single deterministic Lagrangian trajectory passing through position $\boldsymbol{x}$ at time t.

In the Kraichnan model of turbulent advection, it was shown by Bernard *et al.* (1998) that the joint limit $\nu, \kappa \to 0$ with $Pr = \nu/\kappa$ fixed is non-deterministic and corresponds to more than one Lagrangian trajectory passing through space–time point $(\boldsymbol{x}, t)$. We remind the reader that the Kraichnan model of turbulent advection replaces the Navier–Stokes solution with a realization drawn from an ensemble of Gaussian random fields $\boldsymbol{u}^\nu$ with mean zero $\langle \boldsymbol{u}^\nu \rangle = \boldsymbol{0}$ and covariance satisfying

$$\langle [u_i^\nu(\boldsymbol{x}+\boldsymbol{r}, t) - u_i^\nu(\boldsymbol{x}, t)][u_j^\nu(\boldsymbol{x}+\boldsymbol{r}, t') - u_j^\nu(\boldsymbol{x}, t')] \rangle = D_{ij}^\nu(\boldsymbol{r})\delta(t - t') \tag{3.5}$$

for a spatial covariance function satisfying $D^{\nu}_{ij}(\boldsymbol{r}) = D^{\nu}_{ji}(\boldsymbol{r})$, $\partial D^{\nu}_{ij}(\boldsymbol{r})/\partial r_j = 0$, and

$$D_{ii}(\boldsymbol{r}) \sim \begin{cases} D_1 r^{\xi}, & \ell_\nu \ll r \ll L, \\ D_2 r^2, & r \ll \ell_\nu, \end{cases} \tag{3.6}$$

for some $0 < \xi < 2$, with the effective 'dissipation length'

$$\ell_\nu = (D_1/D_2)^{1/(2-\xi)}. \tag{3.7}$$

It should be noted that $D_2 \propto \langle |\nabla \boldsymbol{u}^\nu|^2 \rangle$ and in real turbulence would be proportional to ε/ν, where ε is the viscous energy dissipation. Hence, $D_2 \to \infty$ or $\ell_\nu \to 0$ with D_1 fixed is the analogue for the Kraichnan model of the infinite-Reynolds-number limit for Navier–Stokes turbulence. In fact, one can introduce a 'viscosity' parameter ν for the Kraichnan model with units of (length)^2/(time), so that $\ell_\nu = (\nu/D_1)^{1/\xi}$. For any $\nu > 0$, the velocity realizations are spatially smooth, but in the limit $\nu \to 0$, they are only Hölder continuous in space with exponent $0 < \xi/2 < 1$. It is well known that for such 'rough' limiting velocity fields the solutions of the deterministic initial-value problem

$$\mathrm{d}\boldsymbol{\xi}(s)/\mathrm{d}s = \boldsymbol{u}(\boldsymbol{\xi}(s), s), \quad \boldsymbol{\xi}(t) = \boldsymbol{x} \tag{3.8}$$

need not be unique and, if not, form a continuum of solutions (e.g. see Hartman 2002). In the Kraichnan model, it has been proved in the double limit with both $\nu \to 0$ and $\kappa \to 0$ that the transition probabilities tend to a limiting form

$$p^*(\boldsymbol{x}', t'|\boldsymbol{x}, t) = \lim_{\nu,\kappa \to 0} p^{\nu,\kappa}(\boldsymbol{x}', t'|\boldsymbol{x}, t). \tag{3.9}$$

It is important to stress here that no average is taken over $\boldsymbol{u}$ in defining these transition probabilities, but only an average over Brownian motions in the stochastic flow equations (2.3), while the velocity realization is held fixed. (It would be less ambiguous to write them as $p^{\nu,\kappa}_{\boldsymbol{u}}(\boldsymbol{x}', t'|\boldsymbol{x}, t)$, with $\boldsymbol{u}$ denoting the fixed flow realization, but this would lead to an even heavier notation.) Most importantly, the limiting transition probabilities for the Kraichnan model are not delta distributions of the form (3.4), but non-trivial probabilities over an ensemble of non-unique solutions of the limiting ordinary differential equation (3.8). This remarkable phenomenon is called spontaneous stochasticity. See Bernard *et al.* (1998) and the later papers of Vanden-Eijnden & Vanden-Eijnden (2000, 2001), Gawędzki & Vergassola (2000), Falkovich *et al.* (2001) and Le Jan & Raimond (2002, 2004).

As shown in these works, spontaneous stochasticity occurs because of the analogue of Richardson (1926) dispersion in the Kraichnan model, which leads to a loss of influence of the molecular diffusivity κ on the separation of the perturbed Lagrangian trajectories after a short time of order $(\kappa^{2-\xi}/D_1)^{1/\xi}$. It is important to emphasize that this result does not mean that randomness in the Lagrangian trajectories suddenly 'appears' only for $\nu, \kappa = 0$, but instead that the randomness persists even as $\nu, \kappa \to 0$. It is thus a phenomenon that can be observed with sequences of positive values, $\nu, \kappa > 0$, for which the velocity field is smooth. For the case of a divergence-free velocity that we discuss here, it is furthermore known that the result does not depend upon the order of limits $\nu \to 0$ and $\kappa \to 0$, which can be taken in either order or together. (The only delicate case is when $\kappa \to 0$ first, so that the Prandtl number goes to infinity, and then $\nu \to 0$ subsequently. Since the Brownian motion disappears from the stochastic equation (2.3) while the velocity field remains smooth, the limiting Lagrangian

trajectories are deterministic. To observe spontaneous stochasticity in that limit, one must additionally allow the initial condition to be random, e.g. with $\tilde{\boldsymbol{\xi}}(t) = \boldsymbol{x} + \epsilon\tilde{\boldsymbol{\rho}}$ for a stochastic perturbation $\tilde{\boldsymbol{\rho}}$ drawn from some fixed distribution $P(\boldsymbol{\rho})$. In that case, spontaneous stochasticity appears in the double limit with $\epsilon \to 0$ and $\nu \to 0$ together, and, for a divergence-free velocity $\boldsymbol{u}$, the limiting transition probabilities are identical to those obtained for the other limits involving $\kappa \to 0$. This infinite-Prandtl case is discussed carefully by Vanden-Eijnden & Vanden-Eijnden (2000) and Gawędzki & Vergassola (2000).) See Bernard *et al.* (1998), Vanden-Eijnden & Vanden-Eijnden (2000, 2001) Gawędzki & Vergassola (2000), Falkovich *et al.* (2001) and Le Jan & Raimond (2002, 2004) for discussions of this point.

There is empirical evidence for such phenomena also in Navier–Stokes turbulence obtained from numerical studies of two-particle dispersion. Eyink (2011) studied stochastic Lagrangian particles whose motion is governed by (2.3) in a 1024^3 DNS at $Re_\lambda = 433$ and found that the mean-square dispersion becomes independent of κ after a short time of order $(\kappa/\varepsilon)^{1/2}$. Bitane, Homann & Bec (2013) studied dispersion of deterministic Lagrangian trajectories ($\kappa = 0$) in a 2048^3 DNS at $Re_\lambda = 460$ and a 4096^3 DNS at $Re_\lambda = 730$, and found that the mean-square dispersion becomes independent of the initial separation r_0 of particle pairs in a short time of order $r_0^{2/3}/\varepsilon^{1/3}$. The results of these studies provide evidence of Lagrangian spontaneous stochasticity for Navier–Stokes solutions. In particular, Bitane *et al.* (2013) found consistent Richardson-dispersion statistics for the two Reynolds numbers studied there. The principal limitation of these previous studies is that they averaged over the release points $\boldsymbol{x}$ of the particles. A particle dispersion averaged over release points which remains non-vanishing in the joint limit $\nu, \kappa \to 0$ is enough to infer spontaneous stochasticity for a set of points $\boldsymbol{x}$ of non-zero volume measure (Bernard *et al.* 1998). However, averaging over $\boldsymbol{x}$ removes information about the effects of spatial intermittency and the local fluid environment on the limiting behaviour of the particle distributions $p^{\nu,\kappa}(\boldsymbol{x}', t'|\boldsymbol{x}, t)$ for specific release locations $\boldsymbol{x}$. There was some previous study of such spatial intermittency in pair dispersion by Biferale *et al.* (2005, 2014) but they studied only deterministic Lagrangian particles at small (Kolmogorov scale) initial separations, and not the stochastic Lagrangian particles relevant to our FDR.

We present here new data obtained from numerical experiments on a high-Reynolds-number turbulence simulation in a 2π-periodic box, for a couple of representative release points. We use simulation data from the homogeneous isotropic dataset in the Johns Hopkins Turbulence Database (Li *et al.* 2008; Yu *et al.* 2012), publicly available online at http://turbulence.pha.jhu.edu. It is ideal for our purposes, since the entire time history of the velocity is stored for a full large-scale eddy-turnover time, allowing us to integrate the flow equations (2.3) backward in time. A significant limitation, however, is that only one Reynolds number is available, $Re \simeq 5058$. One should consider together with the limit $\kappa \to 0$ also a limit $\nu \to 0$, so that the Navier–Stokes solution $\boldsymbol{u}^\nu$ converges to a fixed velocity $\boldsymbol{u}$ that is some sort of weak solution of Euler (as always occurs along a suitable subsequence $\nu_k \to 0$; see Lions (1996, § 4.4)). (The necessity of extracting such a subsequence makes the empirical study of spontaneous stochasticity quite difficult, as a matter of principle. The compactness argument of Lions (1996) establishes existence of subsequences of $\nu_k \to 0$ such that Navier–Stokes solutions $\boldsymbol{u}^{\nu_k}(\boldsymbol{x}, t)$ converge to a fixed limiting velocity field $\boldsymbol{u}(\boldsymbol{x}, t)$ that is a 'dissipative Euler solution', in a suitable sense. Unfortunately, the proof is not constructive and therefore there is currently no concrete computational algorithm to generate any specific convergent subsequence. Amusingly, the direct

experimental observation of spontaneous stochasticity in such a joint limit may be easier in quantum mechanics than in turbulent fluids. See Eyink & Drivas (2015*a*).) Since no such joint limit $\nu,\ \kappa \to 0$ can be considered within the given database at one fixed value of viscosity, our study of spontaneous stochasticity is based on the assumption that the Reynolds numbers is already 'sufficiently large'. More precisely, we assume that an inertial-range superballistic Richardson-type dispersion of particle pairs released at space–time point $(\boldsymbol{x},\ t)$ will occur at times $|t' - t| > t_c$, with a crossover time $t_c = \max\{(\varepsilon/\nu)^{1/2},\ (\varepsilon/\kappa)^{1/2}\}$, and then $p^{\nu,\kappa}(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t) \simeq p^*(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t)$ for $|t' - t| \gg t_c$. It should be noted that, for $Pr < 1$, mean-square dispersion grows diffusively $\propto\kappa|t' - t|$ up to a time $t_\kappa = (\varepsilon/\kappa)^{1/2}$, when relative advection begins to dominate at the length scale $\eta_\kappa = (\kappa^3/\varepsilon)^{1/4}$ within the inertial range. Instead, for $Pr > 1$, particle pairs also separate diffusively initially but then transition to exponential divergence $\sim\exp(t/t_\eta)\eta_\kappa$, with Kolmogorov time $t_\eta = (\varepsilon/\nu)^{1/2}$, until the particles separate to the Kolmogorov dissipation scale $\eta = (\nu^3/\epsilon)$ at time $\sim t_\eta \log Pr$, when superballistic dispersion commences. The particle distributions $p^{\nu,\kappa}(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t)$ at $|t' - t| \simeq t_c$ will be distinct in these different cases and will presumably also depend upon the particular values of ν, κ even as $\nu, \kappa \to 0$. However, Richardson dispersion leads to a very rapid 'forgetting' of the precise initial data, and thus it is reasonable to expect that $p^{\nu,\kappa}(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t) \simeq p^*(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t)$ for $|t' - t| \gg t_c$. With this assumption, we may study the limiting particle distributions $p^*(\boldsymbol{x}',\ t'|\boldsymbol{x},\ t)$ in the database at large but finite Reynolds number. A check on this assumption is provided by the fact that, for incompressible flows, the limiting distributions are also expected to be independent of the Prandtl number Pr (Vanden-Eijnden & Vanden-Eijnden 2000, 2001; Gawędzki & Vergassola 2000). By varying κ for the fixed ν in the database, we can change Pr and verify to what extent the Prandtl independence of limiting distributions holds for our numerical results.

We consider two release points $\boldsymbol{x}$ at time $t_f = 2.048$, the final database time, one chosen in a typical turbulent 'background' region and the other in the vicinity of a strong large-scale vorticity. We study stochastic trajectories with diffusivities κ corresponding to three values of the Prandtl number, $Pr = 0.1,\ 1$ and 10. See appendix C for details about the numerical methods employed in our analysis. Figure 1(*a*,*b*) shows 30 representative particle trajectories for the two release points and for each of the three Prandtl numbers. To illustrate the local fluid environment, we also plot isosurfaces of the vorticity filtered with a box filter of width $L/4$ (where L is the integral scale) at the time $s = (2/3)T_L$ (where T_L is the large-scale turnover time). The isosurfaces are for magnitudes of filtered vorticity equal to $15/T_L$. Figure 1(*a*,*c*,*e*) shows the particles released in a typical 'background' region with spottier weaker vortices and figure 1(*b*,*d*,*f*) shows particles released near a strong vortex. (In the weak background region in figure 1(*a*), $\overline{\omega}_{rms}T_L = 2.98$, and thus the isosurface level is $|\overline{\boldsymbol{\omega}}| \approx 5.0\overline{\omega}_{rms}$, with 0.75 % of the volume in that box carrying filtered vorticity above this threshold. In the strong vortex region in figure 1(*b*), instead, $\overline{\omega}_{rms}T_L = 4.25$, so that the isosurface there is $|\overline{\boldsymbol{\omega}}| \approx 3.5\overline{\omega}_{rms}$, with 2.9 % of the volume above.) The three colourings of the trajectories (green/blue/red) represent the three values of the Prandtl number, $Pr = 0.1,\ 1,\ 10$ respectively. The clearly observable 'splitting' of the bundle of stochastic trajectories into sub-bundles at specific times recalls one proposed mechanism for Richardson dispersion, via a sequence of smooth transport and rapid 'flight-like' departures at fluid separatrices (Shlesinger, West & Klafter 1987; Davila & Vassilicos 2003; Thalabard, Krstulovic & Bec 2014). Most importantly, as one can see by eye, the ensembles of trajectories are quite similar for the three Pr values.

To make the latter observation more quantitative, we plot in figure 1(*c*,*d*) the mean-square dispersion of pairs of stochastic Lagrangian particles with different

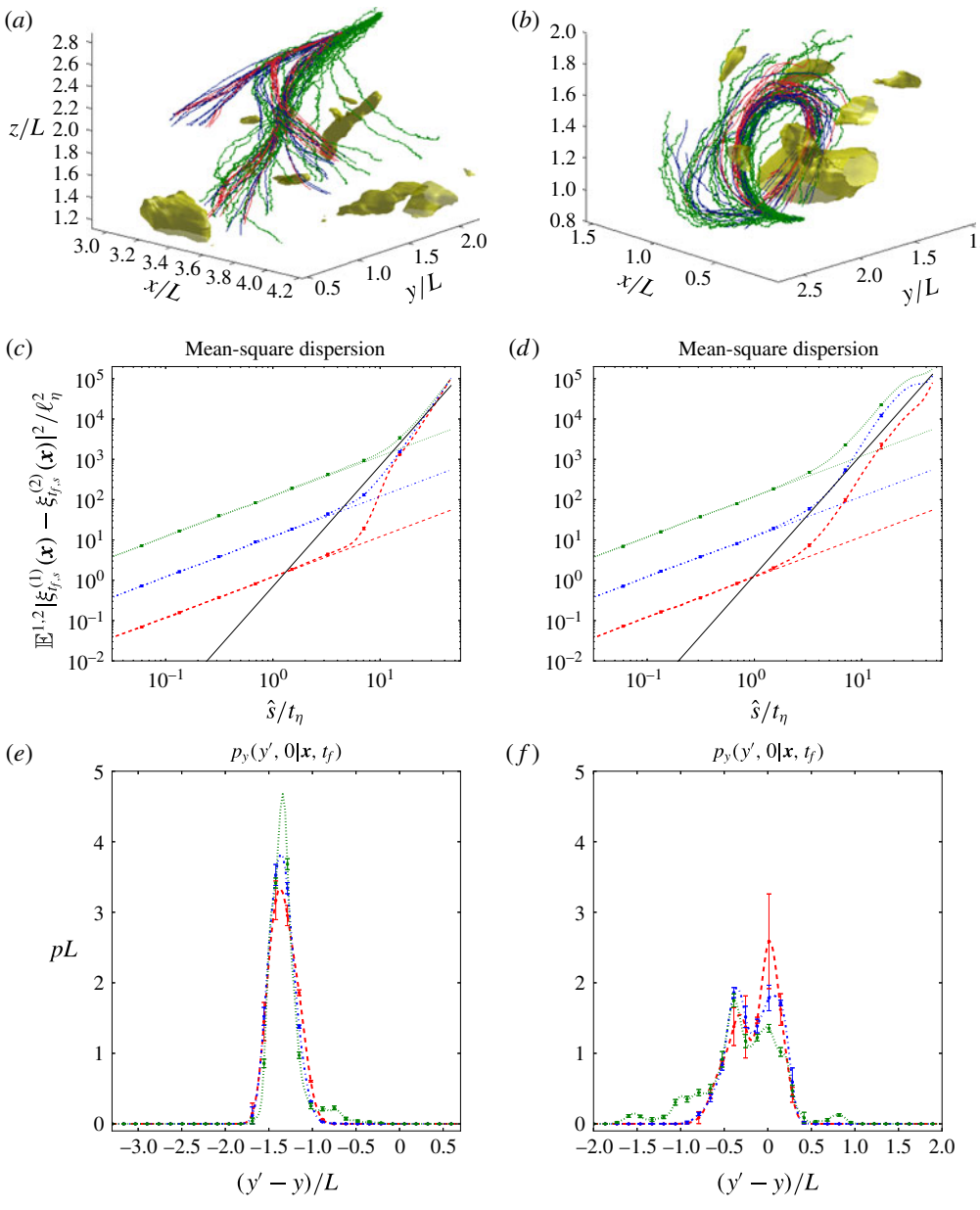

FIGURE 1. (Colour online) (*a*,*c*,*e*) Plots for release at $\boldsymbol{x} = (4.9637, 3.1416, 3.8488)$ in the background region; (*b*,*d*,*f*) plots for release at $\boldsymbol{x} = (0.2610, 3.1416, 1.4617)$ near a strong vortex. (*a*,*b*) Plots of 30 representative stochastic trajectories for $Pr = 0.1$ (green, light), 1.0 (blue, medium) and 10 (red, heavy) together with isosurfaces of coarse-grained vorticity $|\bar{\boldsymbol{\omega}}|T_L = 15$ at time $s = (2/3)T_L$. (*c*,*d*) Plots of particle dispersions (heavy) and short-time results $12\kappa\hat{s}$ (light) for each Pr, with $Pr = 0.1$ (green, dot, $\cdots\cdots$), 1.0 (blue, dash–dot, — · —) and 10 (red, dash, – – –), and a plot in (solid, ——) of $g\varepsilon\hat{s}^3$ with $g = 0.7$ (*c*) and $g = 4/3$ (*d*). (*e*,*f*) Plots of $p_y(y', 0|\boldsymbol{x}, t_f)$ for the three Pr values with the same line styles as in (*c*,*d*).

realizations of the noise, for the two release points and the three Prandtl numbers. The error bars (almost too small to be observed) represent the standard error of the mean (s.e.m.) for averages over $N = 1024$ sample trajectories. For both release points, there is an initial period (going backward in time) where the dispersion grows diffusively as $12\kappa\hat{s}$ with $\hat{s} = t_f - s$, but which then crosses over to a regime of superballistic separation that is close to the $\hat{s}^3$ growth predicted by Richardson (1926) and is approximately independent of Pr. An essential observation from figure 1 is that $\hat{s} \approx t_c$ is indeed the time of crossover to a roughly Richardson t^3 growth. The two release points shown here illustrate behaviour that we have observed also in many other points of the turbulent fluid, where we find that the Richardson $\hat{s}^3$ law is surprisingly robust (albeit imperfect), without the necessity of averaging over release points $\boldsymbol{x}$. This is especially so for points $\boldsymbol{x}$ in 'background' regions, and is at least approximately observed for $\boldsymbol{x}$ located in more intermittent regions.

Finally, we plot in figure 1(*e*,*f*) particle transition probabilities, which provide even more information about the limiting behaviour. We plot at time 0, in the approximate Richardson range, the one-dimensional probability density functions (PDFs) of the y-coordinate, or

$$p_y^{\nu,\kappa}(y', 0|\boldsymbol{x}, t_f) = \int \mathrm{d}x'\,\mathrm{d}z' p^{\nu,\kappa}(\boldsymbol{x}', 0|\boldsymbol{x}, t_f), \tag{3.10}$$

for each of the two release points $\boldsymbol{x}$ and three Prandtl numbers. We observe very similar behaviour also for the x- and z-coordinates. In order to minimize the number of samples required to construct the PDFs numerically, we employed kernel density estimator techniques, which gave us good results with only $N = 6144$ samples. See Silverman (1986) and appendix C, where our numerical procedures are completely described. The error bars represent both the s.e.m.s for the N-sample averages and the effects of variation in the kernel density bandwidth. Consistent with the dispersion plots, we see that the transition PDFs are approximately independent of Pr for times in the superballistic dispersion range. This is especially true for the release point $\boldsymbol{x}$ in the 'background' region, and for the strong vorticity region, such independence holds better for the two largest values of Pr (smallest κ), when the Richardson-like superballistic range is the longest. This approximate independence of the PDFs from the Prandtl number gives some support to the conjecture that $p^{\nu,\kappa}(\boldsymbol{x}', 0|\boldsymbol{x}, t_f) \simeq p^*(\boldsymbol{x}', 0|\boldsymbol{x}, t_f)$ and that the infinite-Re limit is already achieved for such particle transition kernels in the database at finite Re. These numerical studies illustrate the present quality of direct evidence for Lagrangian spontaneous stochasticity in high-Reynolds-number Navier–Stokes turbulence, which is suggestive but far from compelling. As we shall now demonstrate, observations of anomalous scalar dissipation provide further evidence, as the two phenomena are essentially related.

4. Spontaneous stochasticity and anomalous dissipation

The phenomenon of spontaneous stochasticity leads to a simple explanation of anomalous dissipation in a turbulent flow, as was first pointed out by Bernard *et al.* (1998) for decaying scalars (no sources) in the Kraichnan model of random advection. This connection can be understood more directly and more generally using our FDR. In fact, it is intuitively clear from the FDR (2.12) that there can be scalar dissipation that is non-vanishing in the limit $\kappa \to 0$ only if there is a non-vanishing variance in that same limit, implying that Lagrangian trajectories must remain stochastic.

This argument holds in the presence of scalar sources and for a scalar advected by any velocity field $\boldsymbol{u}^\nu$ whatsoever. In particular, the argument holds when $\boldsymbol{u}^\nu$ is a Navier–Stokes solution. Thus, spontaneous stochasticity is the only possible mechanism of anomalous dissipation, for both passive and active scalars, away from walls. Furthermore, we shall show for a passive scalar that does not react back on the flow that spontaneous stochasticity also makes anomalous scalar dissipation possible. Thus, for passive scalars, the two phenomena are completely equivalent. In this section, we shall deduce these conclusions, assuming only that the flow domain is compact (closed and bounded) and without any bounding walls.

We first discuss the technically simpler case with $S \equiv 0$ and then show that the same argument extends easily to the case with a non-zero scalar source. When $S \equiv 0$, we can rewrite the left-hand side of the FDR (2.12) using

$$\begin{aligned}\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))] \;=\; & \int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' \, \theta_0(\boldsymbol{x}_0)\theta_0(\boldsymbol{x}_0') \\ & \times [p_2^{\nu,\kappa}(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^{\nu,\kappa}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^{\nu,\kappa}(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)], \end{aligned} \tag{4.1}$$

where we have introduced the two-time (backward-in-time) transition probability density

$$p_2^{\nu,\kappa}(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) = \mathbb{E}[\delta^d(\boldsymbol{y} - \tilde{\boldsymbol{\xi}}_{t,s}^{\nu,\kappa}(\boldsymbol{x}))\delta^d(\boldsymbol{y}' - \tilde{\boldsymbol{\xi}}_{t,s'}^{\nu,\kappa}(\boldsymbol{x}))], \quad s, s' < t, \tag{4.2}$$

which gives the joint probability for the particle to end up at $\boldsymbol{y}$ at time $s < t$ and at $\boldsymbol{y}'$ at time $s' < t$, given that it started at $\boldsymbol{x}$ at the final time t (moving backward from final to earlier times). At equal times $s = s'$,

$$p_2^{\nu,\kappa}(\boldsymbol{y}, s; \boldsymbol{y}', s|\boldsymbol{x}, t) = \delta^d(\boldsymbol{y} - \boldsymbol{y}')p^{\nu,\kappa}(\boldsymbol{y}, s|\boldsymbol{x}, t). \tag{4.3}$$

We now consider the limit $\nu, \kappa \to 0$, so that the transition probabilities approach limiting values $p^*(\boldsymbol{y}, s; \boldsymbol{y}', s|\boldsymbol{x}, t)$, $p^*(\boldsymbol{y}, s|\boldsymbol{x}, t)$. Such limits exist, at least along suitably chosen subsequences $\nu_n, \kappa_n \to 0$, whenever the flow domain is compact. This can be shown using Young measure methods similar to those that have been employed previously to study statistical equilibria for 2D Euler solutions (Robert 1991; Robert & Sommeria 1991; Sommeria, Staquet & Robert 1991). Because the proof of this result is a little technical, we give it in appendix A.1. When the Lagrangian particles move according to a deterministic flow $\boldsymbol{\xi}_{t,s}^*$, one easily sees that the two-time transition probability factorizes as

$$p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) = \delta^d(\boldsymbol{y} - \boldsymbol{\xi}_{t,s}^*(\boldsymbol{x}))\delta^d(\boldsymbol{y}' - \boldsymbol{\xi}_{t,s'}^*(\boldsymbol{x})) = p^*(\boldsymbol{y}, s|\boldsymbol{x}, t)p^*(\boldsymbol{y}', s'|\boldsymbol{x}, t). \tag{4.4}$$

Hence, non-factorization in the limit $\nu, \kappa \to 0$ is an unequivocal sign of spontaneous stochasticity. The variance on the left-hand side of the FDR (2.12) can only be non-vanishing in the limit if factorization fails, so that anomalous dissipation clearly requires spontaneous stochasticity. In the other direction, if there is spontaneous stochasticity and thus factorization fails for some positive-measure set of $\boldsymbol{x} \in \Omega$, then the contribution to the volume-integrated variance from that subset must be positive for some suitable smooth choice of θ_0, which implies a positive lower bound to the cumulative volume-integrated scalar dissipation. In short, anomalous scalar dissipation and Lagrangian spontaneous stochasticity are seen to be equivalent. This argument is given as a formal mathematical proof in appendix A.2.

The sufficiency argument works only for a passive scalar. For active scalars, the initial datum θ_0 partially determines the velocity field $\boldsymbol{u}$ and so is not free to vary. In order to conclude sufficiency in this case, one needs to assume that the resulting velocity field does not 'conspire' with the initial scalar to cause the variance to vanish, i.e. for the random trajectories to sample only points on a single level set of θ_0. If this remarkable behaviour did happen to occur for some choice of θ_0, then one would not expect it to persist for a small perturbation of θ_0. Thus, it is highly likely also for active scalars that spontaneous stochasticity implies anomalous dissipation, but we have not proved that with the FDR. We can, however, conclude rigorously both for passive and for active scalars that anomalous dissipation implies spontaneous stochasticity. The above proposition shows that any evidence for anomalous scalar dissipation in the free decay of an active or passive scalar (no sources) obtained from DNS in a periodic box is also evidence for spontaneous stochasticity. The argument in this section is a strong motivation to perform DNS studies to verify anomalous dissipation in the free decay of a scalar, since this would provide additional confirmation of spontaneous stochasticity. All of the DNS cited by Yeung, Donzis & Sreenivasan (2005, § 2.1) employed sources (e.g. a mean scalar gradient coupled to the velocity field) that maintained a statistical steady state for the scalar fluctuations.

Inclusion of a non-zero scalar source involves only minor changes to the previous argument. First, it should be noted that

$$\begin{aligned}\mathrm{Var}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))+\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)\,\mathrm{d}s\right]&=\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))]\\&+2\,\mathrm{Cov}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})),\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)\,\mathrm{d}s\right]+\mathrm{Var}\left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)\,\mathrm{d}s\right].\end{aligned}\tag{4.5}$$

Furthermore, one has for the variance of the time-integrated source sampled along the stochastic particle trajectory that

$$\begin{aligned}\mathrm{Var}\left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)\,\mathrm{d}s\right]&=\int_0^t\mathrm{d}s\int_0^t\mathrm{d}s'\int\mathrm{d}^dy\int\mathrm{d}^dy' S(\boldsymbol{y},s)S(\boldsymbol{y}',s')\\&\times[p_2^{v,\kappa}(\boldsymbol{y},s;\boldsymbol{y}',s'|\boldsymbol{x},t)-p^{v,\kappa}(\boldsymbol{y},s|\boldsymbol{x},t)p^{v,\kappa}(\boldsymbol{y}',s'|\boldsymbol{x},t)]\end{aligned}\tag{4.6}$$

and for the covariance between the sampled initial data and the integrated source that

$$\begin{aligned}\mathrm{Cov}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})),\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)\,\mathrm{d}s\right]&=\int_0^t\mathrm{d}s\int\mathrm{d}^dx_0\int\mathrm{d}^dy\,\theta_0(\boldsymbol{x}_0)S(\boldsymbol{y},s)\\&\times[p_2^{v,\kappa}(\boldsymbol{x}_0,0;\boldsymbol{y},s|\boldsymbol{x},t)-p^{v,\kappa}(\boldsymbol{x}_0,0|\boldsymbol{x},t)p^{v,\kappa}(\boldsymbol{y},s|\boldsymbol{x},t)].\end{aligned}\tag{4.7}$$

Clearly, anomalous scalar dissipation requires spontaneous stochasticity. For a passive scalar, we can also argue in the other direction. Indeed, we can repeat the previous argument to conclude that, if there is spontaneous stochasticity for a positive-measure set of $\boldsymbol{x}$, then not only is there a smooth choice of θ_0 so that the variance associated with the initial condition in (4.1) is positive when integrated over this set of $\boldsymbol{x}$, but also there is a smooth choice of source field S so that the contribution of the variance (4.6) is positive. This is already enough to conclude that there must be anomalous dissipation for the scalar with initial condition 0 and with the chosen source S. We can also conclude that there is anomalous dissipation for the initial condition θ_0 and the source S. Indeed, if the total variance contribution in (4.5) is not positive,

then it must vanish, which implies that the covariance term in (4.7) provides a negative contribution. In this case, we simply take $S \to -S$ to make the contributions of all three terms (4.1), (4.6), (4.7) positive. We thus conclude that, also for the passive scalar rejuvenated by a source, there is equivalence between anomalous scalar dissipation and Lagrangian spontaneous stochasticity. The argument is given more carefully in appendix A.2.

It has not been generally appreciated that similar conclusions can be reached in the special case of sourceless scalars using the arguments of Bernard *et al.* (1998), which are not at all restricted to the Kraichnan model. To underline this point and, also, to give additional insight, we here briefly summarize their reasoning. It should be noted that the stochastic representation (3.1) of the advected scalar in the limit $\nu, \kappa \to 0$ becomes, using (3.9),

$$\theta^*(\boldsymbol{x}, t) = \int \mathrm{d}^d x_0 \, \theta_0(\boldsymbol{x}_0) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t). \tag{4.8}$$

It is worth noting that $\theta^*(\boldsymbol{x}, t)$ is a kind of 'weak solution' of the ideal advection equation, $\partial_t \theta^* + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta^* = 0$, although this fact is not needed for the argument. It follows from (4.8) that for any strictly convex function $h(\theta)$, e.g. $h(\theta) = \theta^2/2$,

$$h(\theta^*(\boldsymbol{x}, t)) \leqslant \int \mathrm{d}^d x_0 \, h(\theta_0(\boldsymbol{x}_0)) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{4.9}$$

and equality holds if and only if the transition probability is a delta distribution of type (3.4). This is the so-called Jensen inequality (e.g. see Itô 1984). Since the limiting transition probabilities are not delta distributions in the Kraichnan model, the inequality in (4.9) is strict. Furthermore, the limiting transition probabilities for $\nu, \kappa \to 0$ inherit the volume-preservation property (3.3), so that

$$\int p^*(\boldsymbol{x}', t'|\boldsymbol{x}, t) \, \mathrm{d}^d x = 1. \tag{4.10}$$

In this case, integrating (4.9) over $\boldsymbol{x}$ gives

$$\int h(\theta^*(\boldsymbol{x}, t)) \, \mathrm{d}^d x < \int h(\theta_0(\boldsymbol{x}_0)) \, \mathrm{d}^d x_0, \tag{4.11}$$

so that the h-integral is decaying (dissipated) even in the limit $\nu, \kappa \to 0$. The anomalous scalar dissipation in the Kraichnan model thus has an elegant Lagrangian mechanism. Essentially, the molecular diffusivity is replaced by a 'turbulent diffusivity' associated with the persistent stochasticity of the Lagrangian trajectories, which continues to homogenize the scalar field even as the molecular diffusivity vanishes. We give rigorous details of this argument in appendix A.3, where, in the absence of sources, we obtain necessary and sufficient conditions for anomalous dissipation identical to those derived from the FDR.

5. Summary and discussion

This paper has derived a Lagrangian FDR for scalars advected by an incompressible fluid. Our relation expresses an exact balance between molecular dissipation of scalar fluctuations and the input of scalar fluctuations from the initial scalar values and internal sources as these are sampled by stochastic Lagrangian trajectories backward in time. We have exploited this relation to give a simple proof (in domains without walls)

that spontaneous stochasticity of Lagrangian trajectories is necessary and sufficient for anomalous dissipation of passive scalars, and necessary (but possibly not sufficient) for anomalous dissipation of active scalars.

An important outstanding question is the extent to which the results of this paper can be carried over to provide a Lagrangian picture of anomalous energy dissipation in Navier–Stokes turbulence. (The most direct application of our scalar results to Navier–Stokes might appear to be to analyse the viscous dissipation of enstrophy in freely decaying 2D turbulence, where the vorticity is an active (pseudo)scalar field. Unfortunately, all of our analysis assumes that the initial scalar field is square-integrable or L^2, but it has been shown by Eyink (2001) and Tran & Dritschel (2006) that there can be no anomalous enstrophy dissipation for a freely decaying 2D Navier–Stokes solution with finite initial enstrophy. It may still be the case that there is anomalous enstrophy dissipation for more singular infinite-enstrophy initial data and that this dissipation is associated with spontaneous stochasticity (see the further discussion in Eyink 2001). However, we cannot investigate this delicate issue using the FDR of the present paper.) We briefly comment upon this issue here. The formal extension of our FDR to viscous energy dissipation is straightforward. We can exploit the stochastic Lagrangian representation for the incompressible Navier–Stokes equation

$$\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla} p + \nu\Delta\boldsymbol{u}, \tag{5.1}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{5.2}$$

recently elaborated by Constantin & Iyer (2008, 2011), which is valid both for flows in domains without boundaries and for wall-bounded flows. Their results can be most simply derived using a backward stochastic particle flow $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})$ and a corresponding 'momentum' $\tilde{\boldsymbol{\pi}}_{t,s}(\boldsymbol{x}) \equiv \boldsymbol{u}(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)$, which together satisfy the backward Itô equations

$$\hat{\mathrm{d}}\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}) = \tilde{\boldsymbol{\pi}}_{t,s}(\boldsymbol{x})\,\mathrm{d}s + \sqrt{2\nu}\,\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s, \tag{5.3}$$

$$\hat{\mathrm{d}}\tilde{\boldsymbol{\pi}}_{t,s}(\boldsymbol{x}) = -\boldsymbol{\nabla} p(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s + \sqrt{2\nu}\,\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s\cdot\boldsymbol{\nabla}\boldsymbol{u}(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s). \tag{5.4}$$

These are a stochastic generalization of Hamilton's particle equations, making contact with traditional methods of Hamiltonian fluid mechanics (Salmon 1988). See the more detailed discussion of Eyink (2010) and Rezakhanlou (2014). By integrating the second of these Hamilton equations from 0 to t and taking expectations over the Brownian motion, one readily obtains

$$\boldsymbol{u}(\boldsymbol{x}, t) = \mathbb{E}\left[\boldsymbol{u}_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})) - \int_0^t \boldsymbol{\nabla} p(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right], \tag{5.5}$$

using the fact that the stochastic integral $\sqrt{2\nu}\int_0^t \hat{\mathrm{d}}\widetilde{\boldsymbol{W}}_s\cdot\boldsymbol{\nabla}\boldsymbol{u}^\nu(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)$ is a backward martingale and so vanishes under expectation. The formula (5.5) was previously derived by Albeverio & Belopolskaya (2010). Moreover, by exploiting the same Itô-isometry argument as applied earlier for scalars, one can derive

$$\nu\int_0^t \mathrm{d}s\langle|\boldsymbol{\nabla}\boldsymbol{u}(s)|^2\rangle_\Omega = \frac{1}{2}\left\langle \mathrm{Var}\left[\boldsymbol{u}_0(\tilde{\boldsymbol{\xi}}_{t,0}) - \int_0^t \boldsymbol{\nabla} p(\tilde{\boldsymbol{\xi}}_{t,s}, s)\,\mathrm{d}s\right]\right\rangle_\Omega. \tag{5.6}$$

This can be considered as an 'FDR' for viscous energy dissipation in a Navier–Stokes solution.

Unfortunately, this relation does not appear to be particularly useful for analysing the high-Reynolds-number (or inviscid) limit. It has a mixed Eulerian–Lagrangian character, since it involves both the particle trajectories $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})$ and the Eulerian pressure-gradient field $\boldsymbol{\nabla} p(\boldsymbol{x}, t)$. The latter field is furthermore a dissipation-range object, which grows increasingly singular as $\nu \to 0$. For example, using the classical K41 scaling estimates (Obukhov 1949; Yaglom 1949; Batchelor 1951), one expects a root-mean-squared value of the pressure gradient $(\boldsymbol{\nabla} p)_{rms} \sim (\varepsilon^3/\nu)^{1/4}$, and intermittency effects will make this field even more singular. Mathematically speaking, the pressure gradient cannot be expected to exist as an ordinary function in the limit $\nu \to 0$ but only as a distribution. Because of these facts, we cannot derive from (5.6) any relation between anomalous energy dissipation and spontaneous stochasticity for Navier–Stokes turbulence. In particular, even if there were anomalous energy dissipation, the limiting stochastic particle trajectories might become deterministic as $\nu \to 0$. In that case, the variance on the right-hand side of (5.6) could remain non-vanishing, because the smaller fluctuations due to vanishing stochasticity could be compensated by the diverging magnitude of the pressure gradient.

More fundamentally, we believe that (5.6) misses essential physics. It should be noted that this relation holds for freely decaying Navier–Stokes turbulence both in 2D and in 3D, but in the former case there is certainly no anomalous energy dissipation. Furthermore, in forced steady-state 2D turbulence, there is evidence in the inverse energy cascade range for Richardson dispersion and Lagrangian spontaneous stochasticity (Boffetta & Sokolov 2002; Faber & Vassilicos 2009), but this is associated not with small-scale energy dissipation by viscosity but instead with large-scale energy dissipation by Ekman-type damping. A possibly important clue is provided by the fact that Richardson dispersion is faster backward in time for 3D forward energy cascade Sawford *et al.* (2005), Berg *et al.* (2006), Eyink (2011), but faster forward in time for 2D inverse energy cascade (Faber & Vassilicos 2009). By a comparison of these observations for 2D and 3D Navier–Stokes turbulence and by means of exact results for Burgers turbulence, Eyink & Drivas (2015*b*) have argued that anomalous energy dissipation for Navier–Stokes turbulence should be related not simply to the presence of spontaneous stochasticity but instead to time asymmetry of the stochastic Lagrangian trajectories. This is reminiscent of so-called 'fluctuation theorems' in non-equilibrium statistical mechanics, which imply exponential asymmetry in the probability of entropy production with positive and negative signs. See Gawędzki (2013) and Schuster *et al.* (2013) for recent reviews. These results are deeply related to traditional fluctuation–dissipation theorems in statistical physics, but we have been unable to discover any connection with our Lagrangian FDR. More recently, a time asymmetry has been established in the very-short-time dispersion of nearby Lagrangian trajectories by Falkovich & Frishman (2013) and Jucha *et al.* (2014). However, these results hold only for times of order $\sim (r_0^2/\varepsilon)^{1/3}$ and therefore cannot explain the long-time Richardson behaviour or the observed time asymmetry therein.

The most important implication of the present work is the additional support provided to the concept of Lagrangian spontaneous stochasticity. Exploiting our Lagrangian FDR, we have shown that any empirical evidence for anomalous scalar dissipation, either for passive or for active scalars, and away from walls, must be taken as evidence also for spontaneous stochasticity. There are profound implications of this phenomenon for many Lagrangian aspects of turbulent flows. For example, Constantin & Iyer (2008) have shown that the classical Kelvin–Helmholtz theorems for vorticity dynamics in smooth solutions of the incompressible Euler equations

generalize within their stochastic framework to solutions of the incompressible Navier–Stokes equation with a positive viscosity. In fact, similarly to the case of the advected scalars discussed in the present work, Constantin & Iyer (2008) proved that circulations around stochastically advected loops are martingales backward in time for the Navier–Stokes solution and also proved that this property completely characterizes these solutions. This 'stochastic Kelvin theorem' demonstrates again that the stochastic Lagrangian approach is the natural generalization to non-ideal fluids of the Lagrangian methods for ideal fluids. Furthermore, if there is spontaneous stochasticity, then vortex motion must remain stochastic for arbitrarily high Reynolds numbers. Contrary to the traditional arguments of Taylor & Green (1937), vortex lines in the ideal limit will not be 'frozen into' the turbulent fluid flow in the usual sense. Similar results holds also for magnetic-field-line motion in resistive magnetohydrodynamics (Eyink 2009), and spontaneous stochasticity then implies the possibility of fast magnetic reconnection in astrophysical plasmas for arbitrarily small electrical conductivity (Eyink *et al.* 2013). In Parts II and III, we extend the derivation of our Lagrangian FDR to wall-bounded flows, and derive similar relations between anomalous scalar dissipation and spontaneous stochasticity, as well as new Lagrangian relations for Nusselt–Rayleigh scaling in turbulent convection.

Acknowledgements

We would like to thank K. Sreenivasan for providing us with several important references, as well as S. Punshon-Smith and C. C. Lalescu for useful discussions. We would also like to thank the Institute for Pure and Applied Mathematics (IPAM) at UCLA, where this paper was partially completed during the fall 2014 long program on 'Mathematics of Turbulence'. We also acknowledge the Johns Hopkins Turbulence Database for the numerical turbulence data employed in this work. G.E. is partially supported by a grant from NSF CBET-1507469 and T.D. was partially supported by the Duncan Fund and a Fink Award from the Department of Applied Mathematics and Statistics at the Johns Hopkins University.

Appendix A. Mathematical proofs

A.1. *Existence of limiting transition probabilities*

To make the arguments in § 4 rigorous, we note that the transition probabilities $p^{\nu,\kappa}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ discussed there are well defined for any sequence of continuous (or even just bounded) velocity fields $\boldsymbol{u}^\nu$. However, we shall generally assume that these fields are even smooth for $\nu > 0$ and their energies are bounded uniformly in ν. Because of the latter assumption, we can always extract a subsequence $\nu_j \to 0$ such that $\boldsymbol{u}^{\nu_j} \to \boldsymbol{u}$, with $\boldsymbol{u}$ a finite energy or $L^2(\Omega \times [0, T])$ velocity field, where convergence is in the weak sense,

$$\lim_{j\to\infty} \int_\Omega \mathrm{d}^d x \int_0^T \mathrm{d}t\, \boldsymbol{u}^{\nu_j}(\boldsymbol{x}, t) \cdot \boldsymbol{w}(\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d x \int_0^T \mathrm{d}t\, \boldsymbol{u}(\boldsymbol{x}, t) \cdot \boldsymbol{w}(\boldsymbol{x}, t), \tag{A 1}$$

for all $\boldsymbol{w} \in L^2(\Omega \times [0, T])$. This is a consequence of the Banach–Alaoglu theorem (Rudin 2006). Thus, we consider limits in which there is a definite fixed fluid velocity $\boldsymbol{u}$. If the $\boldsymbol{u}^\nu$ are solutions of the incompressible Navier–Stokes equation indexed by viscosity ν, then we can furthermore select the subsequence $\nu_k \to 0$ so that the limiting velocity $\boldsymbol{u}$ is a 'dissipative Euler solution' in the sense of Lions (1996, § 4.4).

We must now show that a further subsequence $\nu_k = \nu_{j_k}$ can be selected together with a corresponding subsequence $\kappa_k \to 0$, so that the transition probabilities $p^{\nu_k,\kappa_k}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ satisfy the following conditions.

(i) There is a transition density $p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ which is measurable in $\boldsymbol{x}$ so that

$$\lim_{k\to\infty} \int_\Omega \mathrm{d}^d x_0 \int_\Omega \mathrm{d}^d x f(\boldsymbol{x}_0, \boldsymbol{x}) p^{\nu_k,\kappa_k}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d x_0 \int_\Omega \mathrm{d}^d x\ f(\boldsymbol{x}_0, \boldsymbol{x})\ p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) \tag{A 2}$$

for all continuous functions $f \in C(\Omega \times \Omega)$.

(ii) (Normalization) $\int_\Omega \mathrm{d}^d x_0\, p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) = 1$ for a.e. $\boldsymbol{x} \in \Omega$.

(iii) (Volume conservation) $\int_\Omega \mathrm{d}^d x_0 \int_\Omega \mathrm{d}^d x\, g(x_0) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d x_0\, g(\boldsymbol{x}_0)$ for all continuous $g \in C(\Omega)$.

To prove the properties (i)–(iii), the key fact we shall use is that the transition probability densities for $\nu, \kappa > 0$ can be regarded as Young measures,

$$\mu_{\boldsymbol{x}}^{\nu,\kappa,t}(\mathrm{d}\boldsymbol{x}_0) = \mathrm{d}^d x_0\, p^{\nu,\kappa}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t); \tag{A 3}$$

that is, as probability measures $\mu_{\boldsymbol{x}}^{\nu,\kappa,t}$ on Ω which are measurably parameterized by $\boldsymbol{x} \in \Omega$. Fluid dynamicists will be familiar with Young measures from theories of long-time statistical equilibria for two-dimensional fluids (Robert 1991; Sommeria *et al.* 1991). A good introduction is provided by the lectures of Valadier (1994), and a comprehensive treatment can be found in the monograph of Florescu & Godet-Thobie (2012).

Here, we briefly review the necessary theory. In the context of our problem, Young measures may be defined as families of probability measures μ_x, defined on a compact set $Y \subseteq \mathbb{R}^m$, measurably parameterized by $x \in X \subset \mathbb{R}^n$, with X also compact. This uniquely defines a positive Radon measure μ over $X \times Y$ given on product sets by

$$\mu(A \times B) = \int_A \mu_x(B)\, \mathrm{d}x. \tag{A 4}$$

By construction, μ satisfies the following identity:

$$\langle \mu, f \rangle \equiv \int_{X\times Y} f(x, y) \mu(\mathrm{d}x, \mathrm{d}y) = \int_X \left(\int_Y f(x, y) \mu_x(\mathrm{d}y) \right) \mathrm{d}x, \tag{A 5}$$

for any continuous function $f \in C(X \times Y)$. Moreover, for $f \in C(X)$, one has

$$\langle \mu, f \rangle = \int_X f(x)\, \mathrm{d}x; \tag{A 6}$$

that is to say, the projection of μ on X is $\mathrm{d}x$, the Lebesgue measure. One may alternatively take these last two properties as the definition of a Young measure. That is, for any positive Radon measure μ on $X \times Y$ whose projection on X is $\mathrm{d}x$, there is a mapping $x \mapsto \mu_x$ satisfying (A 5). This is the content of the so-called disintegration theorems (Jiřina 1959; Valadier 1973). The mapping $x \mapsto \mu_x$ is unique Lebesgue almost everywhere.

Let us denote by $\mathcal{Y}$ the set of Young measures μ on the product set $X \times Y$. This set has the important property that it is a closed subset of the space $M(X \times Y)$ of Radon measures on $X \times Y$ in the topology of narrow convergence. The narrow topology is the coarsest topology on $M(X \times Y)$ for which the maps $\mu \mapsto \langle \mu, f \rangle$ are continuous for all $f \in C_b(X \times Y)$, the space of bounded continuous functions. Since $X \times Y$ is compact, this topology coincides with the so-called vague topology which is the coarsest for which the maps $\mu \mapsto \langle \mu, f \rangle$ are continuous for all $f \in C_c(X \times Y)$, the

space of compactly supported continuous functions. Furthermore, it coincides with the topology defined by the maps $\mu \mapsto \langle \mu, f\rangle$ for all $f \in C(X \times Y)$. For a detailed discussion of these different topologies, see Florescu & Godet-Thobie (2012). Here, we note only that these make $M(X \times Y)$ into a compact metrizable topological space for compact X, Y. That $\mathcal{Y}$ is a closed subspace of $M(X \times Y)$ may then easily seen by noting that for any sequence $\mu^n \in \mathcal{Y}$ with $\mu^n \to \mu$ narrowly,

$$\int_X f(x)\,\mathrm{d}x = \langle \mu^n, f\rangle \to \langle \mu, f\rangle, \quad \text{for all } f \in C(X), \tag{A 7}$$

so that the projection of μ onto X is $\mathrm{d}x$ and $\mu \in \mathcal{Y}$. A further closed subset $\mathcal{Y}_m \subset \mathcal{Y}$ is the set of measure-preserving Young measures, which satisfy the additional condition that

$$\langle \mu, g\rangle = \int_X \left(\int_Y g(y)\mu_x(\mathrm{d}y) \right) \mathrm{d}x = \int_Y g(y)\,\mathrm{d}y, \quad \text{for all } g \in C(Y), \tag{A 8}$$

which may be stated formally as $\int_X \mathrm{d}x\mu_x(\mathrm{d}y) = \mathrm{d}y$. That $\mathcal{Y}_m$ is closed in the narrow topology is shown by an argument exactly like that for $\mathcal{Y}$ above.

From these basic results, we can easily derive the consequences (i)–(iii), taking $X = Y = \Omega$, where Ω is the closure of a bounded open set with a smooth boundary. Then, with the definition (A 3), one has $\mu^{\nu,\kappa,t} \in \mathcal{Y}_m$ for fixed t and all $\nu, \kappa > 0$. Since $\mathcal{Y}_m$ is a closed subset of the compact metrizable space $M(X \times Y)$, it is itself (sequentially) compact. Hence, given the subsequence ν_j, there is a further subsequence $\nu_k = \nu_{j_k}$ and a corresponding sequence κ_k such that $\mu^{\nu_k\kappa_k,t} \to \mu^{*t} \in \mathcal{Y}_m$ in the narrow topology. It should be noted that the limit μ^{*t} need not be unique and may depend upon the selected subsequence. The narrow convergence $\mu^{\nu_k\kappa_k,t} \to \mu^{*t}$ is equivalent to (i), with the definition

$$\mathrm{d}^d x_0 p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) = \mu_{\boldsymbol{x}}^{*,t}(\mathrm{d}\boldsymbol{x}_0), \tag{A 9}$$

where in general $p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ is a distribution in the variable $\boldsymbol{x}_0$, not an ordinary function. Then, (ii) is a restatement that $\mu^{*t} \in \mathcal{Y}$ and (iii) is a restatement that $\mu^{*t} \in \mathcal{Y}_m$. These observations complete the proof of properties (i)–(iii) above.

With these results in hand, we now rigorously prove the equivalence of spontaneous stochasticity and anomalous dissipation. We do this in two ways: first, by exploiting our general FDR and, second, by the original argument of Bernard *et al.* (1998) for the case of scalars without sources.

A.2. *Proofs using the FDR*

As in the main text, we first consider the case without a scalar source ($S = 0$). Our starting point is the FDR (2.12), with formula (4.1) for the variance $\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}^{\nu_k,\kappa_k}(\boldsymbol{x}))]$. It follows from (i)–(ii) of appendix A.1 that a subsequence $\nu_k = \nu_{j_k}$ can be selected together with a corresponding subsequence $\kappa_k \to 0$, so that the space-averaged variance will satisfy

$$\begin{aligned}\lim_{k\to\infty} \langle \mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}^{\nu_k,\kappa_k})]\rangle_\Omega = \int \mathrm{d}^d x \int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0'\, \theta_0(\boldsymbol{x}_0)\theta_0(\boldsymbol{x}_0') \\ \times [p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)],\end{aligned} \tag{A 10}$$

for all $\theta_0 \in C(\Omega)$, where

$$p_2^*(\boldsymbol{x}_0, 0, \boldsymbol{x}_0', 0|\boldsymbol{x}, t) \equiv \delta^d(\boldsymbol{x}_0 - \boldsymbol{x}_0')p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t). \tag{A 11}$$

It should be noted that p_2^* is a Young measure on $Y = \Omega \times \Omega$ measurably indexed by elements $\boldsymbol{x}$ of $X = \Omega$, since it is a narrow limit of the Young measures $p_2^{\nu_k,\kappa_k}$. We shall not use the property (iii) from appendix A.1 in our argument, although volume conservation was, of course, used in the derivation of the FDR (2.12). Since that FDR holds for all $\nu, \kappa > 0$, it follows that the limit of the cumulative global scalar dissipation exists and must coincide with the limiting variance,

$$\lim_{k\to\infty} \kappa_k \int_0^t \mathrm{d}s \langle |\nabla \theta^{\nu_k,\kappa_k}(s)|^2 \rangle_\Omega = \int \mathrm{d}^d x \int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' \, \theta_0(\boldsymbol{x}_0)\theta_0(\boldsymbol{x}_0') \times [p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)], \tag{A 12}$$

for all $\theta_0 \in C(\Omega)$. It follows immediately that anomalous scalar dissipation requires spontaneous stochasticity since, by the exact formula (A 12), a non-vanishing cumulative dissipation necessitates non-factorization on a finite-measure set of $\boldsymbol{x}$.

The argument that spontaneous stochasticity implies anomalous dissipation is a little more involved. We need to show that if non-factorization holds on a finite-measure set of $\boldsymbol{x}$, then there exists a smooth choice of θ_0 such that both sides of (A 12) are positive. Thus, we assume the opposite, that both sides vanish for all smooth θ_0. The right-hand side then also vanishes for all continuous θ_0, since $C^\infty(\Omega)$ is dense in $C(\Omega)$ in the uniform norm. For example, this density follows by the Stone–Weierstrass theorem (Rudin 2006), since $C^\infty(\Omega)$ is a subalgebra of $C(\Omega)$ containing the constant 1, closed under complex conjugation, and separating points of Ω. Since the integrand with respect to $\boldsymbol{x}$ is a variance, it is non-negative, so that the vanishing of the integral over $\boldsymbol{x}$ implies that there is a subset $\Omega_0 \subset \Omega$ of full measure, such that

$$\int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' \, \theta_0(\boldsymbol{x}_0)\theta_0(\boldsymbol{x}_0')[p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)] = 0, \tag{A 13}$$

for all $\boldsymbol{x} \in \Omega_0$ and $\theta_0 \in C(\Omega)$. It should be noted furthermore that the quantity in the square brackets '[·]' in the equation above is symmetric in $\boldsymbol{x}_0$, $\boldsymbol{x}_0'$. Thus, for any pair of functions g, h, one can take $\theta_0 = g + h$ to infer that

$$\int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' \, g(\boldsymbol{x}_0)h(\boldsymbol{x}_0')[p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)] = 0 \tag{A 14}$$

for all $\boldsymbol{x} \in \Omega_0$ and $g, h \in C(\Omega)$. Since the product functions $(g \otimes h)(\boldsymbol{x}, \boldsymbol{x}_0') = g(\boldsymbol{x}_0)h(\boldsymbol{x}_0')$ form a subalgebra of $C(\Omega^2)$ that satisfies all of the conditions of the Stone–Weierstrass theorem, we can use this theorem again to extend the equality to

$$\int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' f(\boldsymbol{x}_0, \boldsymbol{x}_0')[p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)] = 0 \tag{A 15}$$

for all $\boldsymbol{x} \in \Omega_0$ and $f \in C(\Omega^2)$. The parameterized measure $\nu_{\boldsymbol{x}}$ defined by

$$\nu_{\boldsymbol{x}}(\mathrm{d}\boldsymbol{x}_0, d\boldsymbol{x}_0') = \mathrm{d}^d x_0 \, \mathrm{d}^d x_0' [p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) - p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t)] \tag{A 16}$$

is a difference of two Young measures, and, thus, there is a continuous linear functional on $C(\Omega^2)$ for all $\boldsymbol{x} \in \Omega_0$, also denoted $\nu_{\boldsymbol{x}}$, defined by $\langle \nu_{\boldsymbol{x}}, f \rangle = \int_{\Omega^2} f \, \mathrm{d}\nu_{\boldsymbol{x}}$. Since

$$\langle \nu_{\boldsymbol{x}}, f \rangle = 0, \quad \text{for all } f \in C(\Omega^2) \quad \text{and} \quad \boldsymbol{x} \in \Omega_0, \tag{A 17}$$

it follows for all $\boldsymbol{x} \in \Omega_0$ that $\nu_{\boldsymbol{x}} \equiv 0$, as an element of the dual Banach space $C(\Omega^2)^*$. A direct consequence is that

$$p_2^*(\boldsymbol{x}_0, 0; \boldsymbol{x}_0', 0|\boldsymbol{x}, t) = p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) p^*(\boldsymbol{x}_0', 0|\boldsymbol{x}, t) \tag{A 18}$$

as distributions in $\boldsymbol{x}_0$, $\boldsymbol{x}_0'$, for all $\boldsymbol{x} \in \Omega_0$. However, this contradicts our starting assumption that factorization fails on a set of full measure. Hence, there must be a smooth choice of θ_0 that makes the right-hand side of (A 12) positive, and thus also the left-hand side.

Let us next consider the case with $\theta_0 \equiv 0$, but with the source S non-vanishing. In this circumstance, the FDR (2.12) becomes

$$\kappa \int_0^t \mathrm{d}s \langle |\nabla \theta(s)|^2 \rangle_\Omega = \frac{1}{2} \left\langle \mathrm{Var} \left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}^{\nu,\kappa}(s)) \, \mathrm{d}s \right] \right\rangle_\Omega , \tag{A 19}$$

with expression (4.6) for the variance. We show first that there is a suitable subsequence $\nu_k = \nu_{j_k} \to 0$ and $\kappa_k \to 0$ such that

$$\begin{aligned} &\lim_{k\to\infty} \int_\Omega \mathrm{d}^d x \, \mathrm{Var} \left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}^{\nu_k,\kappa_k}(\boldsymbol{x}), s) \, \mathrm{d}s \right] \\ &\quad = \int_\Omega \mathrm{d}^d x \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d y' S(\boldsymbol{y}, s) S(\boldsymbol{y}', s') \\ &\qquad \times [p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) - p^*(\boldsymbol{y}, s|\boldsymbol{x}, t) p^*(\boldsymbol{y}', s'|\boldsymbol{x}, t)] \end{aligned} \tag{A 20}$$

for any $S \in C(\Omega \times [0, t])$ and for suitable limiting transition probabilities p_2^* and p^*. To show this, we note that

$$\mu_{s,s',\boldsymbol{x}}^{\nu,\kappa}(\mathrm{d}\boldsymbol{y}, d\boldsymbol{y}') = \mathrm{d}^d y \, \mathrm{d}^d y' p_2^{\nu,\kappa}(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \tag{A 21}$$

defines a set of Young measures on $Y = \Omega \times \Omega$ measurably indexed by elements $(s, s', \boldsymbol{x})$ of $X = [0, t] \times [0, t] \times \Omega$. Since these spaces X and Y are both compact, we can appeal to the general results on Young measures discussed in appendix A.1 to infer that a subsequence ν_k, κ_k exists so that, for all $f \in C(X \times Y)$,

$$\begin{aligned} &\lim_{k\to\infty} \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d y' \int_\Omega \mathrm{d}^d x f(\boldsymbol{y}, s; \boldsymbol{y}', s'; \boldsymbol{x}) p_2^{\nu_k,\kappa_k}(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \\ &\quad = \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d y' \int_\Omega \mathrm{d}^d x f(\boldsymbol{y}, s; \boldsymbol{y}', s'; \boldsymbol{x}) p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \end{aligned} \tag{A 22}$$

for some limit Young measure with distributional density p_2^*, which it is easy to show inherits the symmetry of $p_2^{\nu_k,\kappa_k}$ in $(\boldsymbol{y}, s)$ and $(\boldsymbol{y}', s')$. Choosing the function f to be of the form $f(\boldsymbol{y}, s; \boldsymbol{y}', s'; \boldsymbol{x}) = h(s') g(\boldsymbol{y}, s; \boldsymbol{x})$ gives also

$$\begin{aligned} &\lim_{k\to\infty} \int_0^t \mathrm{d}s \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d x \, g(\boldsymbol{y}, s; \boldsymbol{x}) p^{\nu_k,\kappa_k}(\boldsymbol{y}, s|\boldsymbol{x}, t) \\ &\quad = \int_0^t \mathrm{d}s \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d x \, g(\boldsymbol{y}, s; \boldsymbol{x}) p^*(\boldsymbol{y}, s|\boldsymbol{x}, t) \end{aligned} \tag{A 23}$$

for all continuous g, with

$$p^*(\boldsymbol{y}, s|\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d y' \; p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \tag{A 24}$$

constant in s' for almost every $s, \boldsymbol{x}$ and defining a consistent one-time Young measure. We can also establish volume-preserving properties of these limiting Young measures, although that will not be necessary to our argument. From these results, (A 20) follows by taking the limit along the subsequence ν_k, κ_k of the formula (4.6) for the variance.

The proof that spontaneous stochasticity is both necessary and sufficient for anomalous scalar dissipation now follows by arguments almost identical to the situation with $\theta_0 \neq 0$, $S \equiv 0$ that was first considered in this section. Necessity is immediate from (A 19), (A 20). The proof of sufficiency is very similar to that given before, by showing that vanishing of the space-integrated variance (A 20) for all smooth source fields S implies the factorization

$$p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) = p^*(\boldsymbol{y}, s|\boldsymbol{x}, t)p^*(\boldsymbol{y}', s'|\boldsymbol{x}, t) \tag{A 25}$$

for almost every $\boldsymbol{x} \in \Omega$. The non-negativity of the $\boldsymbol{x}$-integrand requires some argument, because it is no longer obviously a variance. However, it is the limit of a variance in the sense that

$$\begin{aligned}\lim_{k\to\infty} \int_\Omega \mathrm{d}^d x u(\boldsymbol{x}) \operatorname{Var}\left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}^{\nu_k,\kappa_k}(\boldsymbol{x}), s)\,\mathrm{d}s\right] \\ = \int_\Omega \mathrm{d}^d x u(\boldsymbol{x}) \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d y' S(\boldsymbol{y}, s)S(\boldsymbol{y}', s') \\ \times [p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) - p^*(\boldsymbol{y}, s|\boldsymbol{x}, t)p^*(\boldsymbol{y}', s'|\boldsymbol{x}, t)]\end{aligned} \tag{A 26}$$

for all $u \in C(\Omega)$ and $S \in C(\Omega \times [0, t])$. If also $u \geqslant 0$, then the left-hand Feynman–Kac side is non-negative, and thus so is the right-hand side. This is enough to infer that

$$\begin{aligned}\int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \int_\Omega \mathrm{d}^d y \int_\Omega \mathrm{d}^d y' S(\boldsymbol{y}, s)S(\boldsymbol{y}', s') \\ \times [p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) - p^*(\boldsymbol{y}, s|\boldsymbol{x}, t)p^*(\boldsymbol{y}', s'|\boldsymbol{x}, t)] \geqslant 0\end{aligned} \tag{A 27}$$

for all $\boldsymbol{x} \in \Omega_0$, a set of full measure in Ω. The remainder of the argument uses the same strategy as before, with $\theta_0 \to S$ and the Banach space $C(\Omega^2) \to C((\Omega \times [0, t])^2)$.

The argument when both $\theta_0 \neq 0$ and $S \neq 0$ has already been given in the main text. We only add here the technical detail that a single subsequence may be selected so that one has narrow convergence both of the two-time Young measure

$$\mu_{s,s',\boldsymbol{x}}^{\nu_k,\kappa_k}(\mathrm{d}\boldsymbol{y}, \mathrm{d}\boldsymbol{y}') = \mathrm{d}^d y\, \mathrm{d}^d y' p_2^{\nu_k,\kappa_k}(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \to \mathrm{d}^d y\, \mathrm{d}^d y' p_2^*(\boldsymbol{y}, s; \boldsymbol{y}', s'|\boldsymbol{x}, t) \tag{A 28}$$

and also of the one-time Young measure at time $t_0 = 0$

$$\mu_{\boldsymbol{x}}^{\nu_k,\kappa_k}(\mathrm{d}\boldsymbol{x}_0) = \mathrm{d}^d x_0\, p^{\nu_k,\kappa_k}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) \to \mathrm{d}^d x_0\, p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t). \tag{A 29}$$

The second statement does not follow from the narrow convergence

$$\mu_{s,\boldsymbol{x}}^{\nu_k,\kappa_k}(\mathrm{d}\boldsymbol{y}) = \mathrm{d}^d y\, p^{\nu_k,\kappa_k}(\boldsymbol{y}, s|\boldsymbol{x}, t) \to \mathrm{d}^d y\, p^*(\boldsymbol{y}, s|\boldsymbol{x}, t) \tag{A 30}$$

because $\{0\}$ is a subset of $[0, t]$ with zero Lebesgue measure. However, after extracting a subsequence for which the two-time Young measure converges, one can extract a further subsequence so that the one-time Young measure at time $t_0 = 0$ also converges.

A.3. *Rigorous Bernard–Gawędzki–Kupiainen argument*

We first demonstrate that spontaneous stochasticity implies anomalous dissipation of passive scalars, using the same ideas as for the Kraichnan model. We give an indirect proof, supposing that there is no anomalous dissipation and showing that there can then be no spontaneous stochasticity, in contradiction to the starting assumption. Thus, we assume for some strictly convex function h that

$$\int h(\theta^*(\boldsymbol{x}, t))\,\mathrm{d}^d x = \int h(\theta_0(\boldsymbol{x}_0))\,\mathrm{d}^d x_0, \tag{A 31}$$

for all smooth initial data θ_0. Using the volume-conserving property (4.10), we can write this equality as

$$\int \mathrm{d}^d x \left[\int \mathrm{d}^d x_0\, h(\theta_0(\boldsymbol{x}_0)) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) - h(\theta^*(\boldsymbol{x}, t)) \right] = 0. \tag{A 32}$$

Because of Jensen's inequality (4.9), the integrand in the square bracket is non-negative, and thus

$$h(\theta^*(\boldsymbol{x}, t)) = \int \mathrm{d}^d x_0\, h(\theta_0(\boldsymbol{x}_0)) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) \tag{A 33}$$

holds pointwise in space for Lebesgue almost every $\boldsymbol{x}$. We can rewrite this equality in terms of the PDF of the random variable $\tilde{\theta}(\boldsymbol{x}, t) = \theta_0(\tilde{\boldsymbol{\xi}}^*_{t,0}(\boldsymbol{x}))$ to assume the value ψ or

$$p^*_\theta(\psi|\boldsymbol{x}, t) = \int \mathrm{d}^d x_0 \delta(\psi - \theta_0(\boldsymbol{x}_0)) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{A 34}$$

so that (A 33) becomes

$$h(\theta^*(\boldsymbol{x}, t)) = \int \mathrm{d}\psi h(\psi) p^*_\theta(\psi|\boldsymbol{x}, t), \tag{A 35}$$

with $\theta^*(\boldsymbol{x}, t) = \int \mathrm{d}\psi\, \psi p^*_\theta(\psi|\boldsymbol{x}, t)$. Because of strict convexity of h, Jensen's inequality together with (A 35) immediately implies that $p^*_\theta(\psi|\boldsymbol{x}, t)$ is a delta distribution, or

$$p^*_\theta(\psi|\boldsymbol{x}, t) = \delta(\psi - \theta^*(\boldsymbol{x}, t)), \tag{A 36}$$

and thus $\tilde{\theta}(\boldsymbol{x}, t)$ is deterministic. It should be noticed that this conclusion holds for active as well as passive scalars. However, (A 36) by itself does not necessarily contradict spontaneous stochasticity, because particle positions $\tilde{\boldsymbol{\xi}}^*_{t,0}(\boldsymbol{x})$ may remain random but sample only one isosurface of θ_0 for a fixed value θ^*! This latter possibility seems very unlikely to be true, even for an active scalar, for every choice of θ_0. However, we cannot presently rule out a possible 'conspiracy' for an active scalar in which changing θ_0 would always alter the velocity field $\boldsymbol{u}$ so that the limiting particle positions $\tilde{\boldsymbol{\xi}}^*_{t,0}(\boldsymbol{x})$ would remain in an isosurface of θ_0. (Although our proof does not work for active scalars, we conjecture that if there is some θ_0 for which spontaneous stochasticity occurs for an active scalar, then there shall be anomalous scalar dissipation for 'generic' perturbations of θ_0. More formally, in a neighbourhood of θ_0, there shall be a dense G_δ set of scalar initial data which produce anomalous dissipation. It should be noted that for passive scalars also we

expect anomalous dissipation for generic θ_0, although our proofs only guarantee the existence of one initial datum leading to anomalous dissipation.) For a passive scalar, fortunately, we are free to choose $\theta_0(\boldsymbol{x})$ in a completely arbitrary manner without altering the velocity field $\boldsymbol{u}$, and we can then conclude that the particle positions themselves must be deterministic. For example, one argument is to take $\theta_0(\boldsymbol{x}) = x_i$, the ith coordinate of $\boldsymbol{x}$ in the periodic domain, which implies that each ith coordinate of $\tilde{\boldsymbol{\xi}}^*_{t,0}(\boldsymbol{x})$ must be deterministic. (If the periodic domain has diameter L_i in the ith direction, we can take $-L_i/2 \leqslant x_i < L_i/2$. Then, $\theta_0(\boldsymbol{x}) = x_i$ is clearly not continuous at $x_i = \pm L_i/2$! This technical difficulty can be overcome, if $p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ is a measurable function of $\boldsymbol{x}_0$, by choosing sequences of continuous functions $\theta_0(\boldsymbol{x})$ that converge to x_i pointwise. A completely different and fully general approach is to integrate (A 36) over ψ^2, or, equivalently, to take $h(\theta) = \theta^2$, which gives $\int \mathrm{d}^d x_0 \theta_0^2(\boldsymbol{x}_0) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) = \int \mathrm{d}^d x_0 \int \mathrm{d}^d x_0' \theta(\boldsymbol{x}_0)\theta(\boldsymbol{x}_0') p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$, and to use the argument of the preceding appendix A.2.) Since this clearly contradicts the assumed spontaneous stochasticity of the limiting particle positions, we conclude that for each strictly convex function h, there indeed must be anomalous scalar dissipation for some initial data θ_0.

Now, we assume instead that there is anomalous scalar dissipation, which means that the 'deficit' in the ideally conserved integral

$$\Delta^{\nu,\kappa}(t) \equiv \int h(\theta_0(\boldsymbol{x}_0))\, \mathrm{d}^d x_0 - \int h(\theta^{\nu,\kappa}(\boldsymbol{x}, t))\, \mathrm{d}^d x, \quad t > 0 \tag{A 37}$$

converges to some limiting value $\Delta(t) > 0$ as $\nu, \kappa \to 0$. Here, we have explicitly indicated the dependence of the solution $\theta^{\nu,\kappa}(\boldsymbol{x}, t)$ of the scalar advection–diffusion equation upon ν, κ. More precisely, let $\theta_0(\boldsymbol{x}_0)$ be continuous on Ω and consider

$$\theta^{\nu_k,\kappa_k}(\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d x_0\, \theta_0(\boldsymbol{x}_0) p^{\nu_k,\kappa_k}(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{A 38}$$

which is measurable in $\boldsymbol{x}$ and with scalar energies uniformly bounded as

$$\int_\Omega \mathrm{d}^d x |\theta^{\nu_k,\kappa_k}(\boldsymbol{x}, t)|^2 \leqslant \int_\Omega \mathrm{d}^d x_0\ |\theta_0(\boldsymbol{x}_0)|^2 \tag{A 39}$$

by Jensen's inequality and volume conservation. From (i) of appendix A.1, we see that

$$\lim_{k\to\infty} \int_\Omega \mathrm{d}^d x f(\boldsymbol{x}) \theta^{\nu_k,\kappa_k}(\boldsymbol{x}, t) = \int_\Omega \mathrm{d}^d x f(\boldsymbol{x}) \theta^*(\boldsymbol{x}, t) \tag{A 40}$$

for all $f \in C(\Omega)$, where we have defined

$$\theta^*(\boldsymbol{x}, t) \equiv \int_\Omega \mathrm{d}^d x_0\ \theta_0(\boldsymbol{x}_0) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{A 41}$$

which satisfies

$$\int_\Omega \mathrm{d}^d x |\theta^*(\boldsymbol{x}, t)|^2 \leqslant \int_\Omega \mathrm{d}^d x_0 |\theta_0(\boldsymbol{x}_0)|^2 \tag{A 42}$$

again by Jensen's inequality using (ii) and the volume-conservation property (iii) of appendix A.1. Because $C(\Omega)$ is dense in $L^2(\Omega)$, the uniform L^2-bounds (A 39), (A 42) and the convergence (A 40) imply that $\theta^{\nu_k,\kappa_k} \to \theta^*$ weak in L^2. Integral functionals

$$H[\theta] = \int_\Omega \mathrm{d}^d x h(\theta(\boldsymbol{x})) \tag{A 43}$$

for convex functions h and finite-measure sets Ω are weakly lower-semicontinuous on $L^2(\Omega)$ (e.g. Berkovitz (1974) or Braides (2002, § 2.2)), so that one has

$$\int_\Omega \mathrm{d}^d x h(\theta^*(\boldsymbol{x}, t)) \leqslant \liminf_{k\to\infty} \int_\Omega \mathrm{d}^d x h(\theta^{\nu_k,\kappa_k}(\boldsymbol{x}, t)). \tag{A 44}$$

Now, we note that if $\Delta^{\nu,\kappa}(t) \to \Delta(t)$ as $\nu, \kappa \to 0$, then for all sufficiently small ν, κ, one has $\Delta^{\nu,\kappa}(t) > \Delta(t)/2$, say, or, in other words,

$$\int_\Omega h(\theta^{\nu,\kappa}(\boldsymbol{x}, t))\, \mathrm{d}^d x < \int_\Omega h(\theta_0(\boldsymbol{x}_0))\, \mathrm{d}^d x_0 - \frac{1}{2}\Delta(t). \tag{A 45}$$

Combining this with the above results, we thus obtain

$$\int_\Omega h(\theta^*(\boldsymbol{x}, t))\, \mathrm{d}^d x \leqslant \int_\Omega h(\theta_0(\boldsymbol{x}_0))\, \mathrm{d}^d x_0 - \frac{1}{2}\Delta(t) < \int_\Omega h(\theta_0(\boldsymbol{x}_0))\, \mathrm{d}^d x_0. \tag{A 46}$$

There must therefore be at least a positive-measure set (non-zero volume) of points $\boldsymbol{x}$ for which

$$h(\theta^*(\boldsymbol{x}, t)) < \int h(\theta_0(\boldsymbol{x}_0)) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)\, \mathrm{d}^d x_0, \tag{A 47}$$

since otherwise the inequality (A 46) would be violated. Because the analogue of (4.8) holds for the limiting transition probability, i.e.

$$\theta^*(\boldsymbol{x}, t) = \int \mathrm{d}^d x_0\, \theta_0(\boldsymbol{x}_0) p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t), \tag{A 48}$$

and when h is strictly convex, one can conclude that the limiting transition probabilities $p^*(\boldsymbol{x}_0, 0|\boldsymbol{x}, t)$ obtained along the particular subsequence $\nu_n, \kappa_n \to 0$ are not delta distributions of type (3.4). Thus, spontaneous stochasticity must hold for at least this positive-measure set of space points $\boldsymbol{x}$. It should be noted that this direction of the proof did not assume a passive scalar.

Appendix B. Averages of the FDR over random sources and initial data

We derive here the specific consequences of our FDR mentioned in § 2.

B.1. *Steady-state relations (2.14) and (2.15)*

We begin with our general steady-state formula (2.15). We consider compact space domains Ω without boundary, although it is worth observing that identical results hold for wall-bounded domains with no scalar flux through the wall (see Part II). We first note that the contributions from the initial data θ_0 all vanish in the limit $t \to \infty$, so that

$$\langle \kappa |\nabla\theta|^2 \rangle_{\Omega,\infty} = \lim_{t\to\infty} \frac{1}{2t} \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \,\mathrm{Cov}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s), S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')). \tag{B 1}$$

To see this, we observe that the finite-time FDR (2.12) can be expressed as

$$\frac{1}{t}\int_0^t \mathrm{d}s\langle\kappa|\nabla\theta(s)|^2\rangle_\Omega = \frac{1}{2t}\langle\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))]\rangle_\Omega + \frac{1}{t}\left\langle \mathrm{Cov}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})), \int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right]\right\rangle_\Omega + \frac{1}{2t}\left\langle \mathrm{Var}\left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right]\right\rangle_\Omega. \tag{B 2}$$

For bounded initial data, one has $(1/2t)\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))] \leqslant (\max|\theta_0|)^2/t \overset{t\to\infty}{\longrightarrow} 0$. The contribution of the covariance between the forcing and the initial term likewise gives a vanishing contribution, since by the Cauchy–Schwartz inequality,

$$\left|\frac{1}{t}\mathrm{Cov}\left[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x})), \int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right]\right| \leqslant \sqrt{\frac{1}{t}\mathrm{Var}[\theta_0(\tilde{\boldsymbol{\xi}}_{t,0}(\boldsymbol{x}))]\cdot\frac{1}{t}\mathrm{Var}\left[\int_0^t S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)\,\mathrm{d}s\right]}. \tag{B 3}$$

Only the final variance term in (B 2) survives in the limit $t\to\infty$. Rewriting this using the bilinearity of the covariance function gives (B 1).

Next, using the symmetry in s, s' of the integrand, we can restrict the integration range in (B 1) to $s' < s$,

$$\langle\kappa|\nabla\theta|^2\rangle_{\Omega,\infty} = \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}s\int_0^s \mathrm{d}s'\,\mathrm{Cov}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s), S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')). \tag{B 4}$$

We then divide the triangular region $R = \{(s, s') : 0 < s' < s < t\}$ into three subregions,

$$\left.\begin{array}{c} R_I = \{(s, s') : 0 < s' < s < t - n\tau\}, \\ R_{II} = \{(s, s') : 0 < s' < t - 2n\tau,\ \ t - n\tau < s < t\}, \\ R_{III} = R\backslash(R_I \cup R_{II}), \end{array}\right\} \tag{B 5}$$

where τ is the scalar mixing time and n is a positive integer. Region R_{III} gives a contribution that is $O(n^2\tau^2/t)$ and can be neglected in the limit $t\to\infty$. In region R_{II}, we can write

$$\mathrm{Cov}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s), S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')) = \mathbb{E}[(\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')|\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})) - \mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')))S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}), s)], \tag{B 6}$$

where $\mathbb{E}(\cdot|\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}))$ is the conditional average over the Brownian motion given the value of $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})$. In region R_{II}, both $t - s' > 2n\tau$ and $s - s' > n\tau$, so that for $n \gg 1$ one may use the ergodicity of the Lagrangian flow in physical space to obtain

$$\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')|\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})) \simeq \langle S(s')\rangle_\Omega, \quad \mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}), s')) \simeq \langle S(s')\rangle_\Omega, \tag{B 7}$$

which give nearly cancelling contributions in (B 6). Thus, this region makes an arbitrarily small contribution for sufficiently large n. Finally, in region R_I, we instead write

$$\begin{aligned}&\mathrm{Cov}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s),S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}),s'))\\&\quad=\mathbb{E}[S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s)S(\tilde{\boldsymbol{\xi}}_{s,s'}(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})),s')]-\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s))\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}),s'))\end{aligned}\tag{B 8}$$

using $\tilde{\boldsymbol{\xi}}_{t,s'}=\tilde{\boldsymbol{\xi}}_{s,s'}\circ\tilde{\boldsymbol{\xi}}_{t,s}$. Since $t-s'>t-s>n\tau$ in region R_I, the random variables $\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x})$, $\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x})$ are nearly uniformly distributed over the domain Ω for $n\gg 1$, by the ergodicity of the stochastic Lagrangian flow. By the strong Markov property, one thus obtains in region R_I that

$$\mathrm{Cov}(S(\tilde{\boldsymbol{\xi}}_{t,s}(\boldsymbol{x}),s),S(\tilde{\boldsymbol{\xi}}_{t,s'}(\boldsymbol{x}),s'))\simeq\langle S(s)\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{s,s'},s'))\rangle_\Omega-\langle S(s)\rangle_\Omega\langle S(s')\rangle_\Omega.\tag{B 9}$$

The right-hand side can rewritten more compactly as a 'truncated correlation function'

$$\begin{aligned}&\langle S(s)\mathbb{E}(S(\tilde{\boldsymbol{\xi}}_{s,s'},s'))\rangle_\Omega-\langle S(s)\rangle_\Omega\langle S(s')\rangle_\Omega\\&\quad=\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s')\rangle_{\mathbb{E},\Omega}-\langle\tilde{S}_L(s,s)\rangle_{\mathbb{E},\Omega}\langle\tilde{S}_L(s,s')\rangle_{\mathbb{E},\Omega}\\&\quad:=\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s')\rangle^{\mathrm{T}}_{\mathbb{E},\Omega},\end{aligned}\tag{B 10}$$

where we have defined the Lagrangian source field $\tilde{S}_L(\boldsymbol{x},s,s')=S(\tilde{\boldsymbol{\xi}}_{s,s'}(\boldsymbol{x}),s')$ as sampled along stochastic trajectories for $s'<s$ and we have also introduced the notation $\langle\cdot\rangle_{\mathbb{E},\Omega}$ for the joint average over Brownian motion and the space domain. When the difference of the two sides of (B 9) is a function integrable over infinite ranges and vanishing as $n\to\infty$, dominated convergence gives

$$\langle\kappa|\nabla\theta|^2\rangle_{\Omega,\infty}=\lim_{n\to\infty}\lim_{t\to\infty}\frac{1}{t}\int_0^{t-n\tau}\mathrm{d}s\int_0^s\mathrm{d}s'\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s')\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}.\tag{B 11}$$

This non-vanishing contribution from region R_I can be combined with vanishing contributions from regions R_{II}, R_{III} by the reverse of the preceding argument, to give

$$\langle\kappa|\nabla\theta|^2\rangle_{\Omega,\infty}=\lim_{t\to\infty}\frac{1}{t}\int_0^t\mathrm{d}s\int_0^s\mathrm{d}s'\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s')\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}.\tag{B 12}$$

To obtain the final result, we make the change of variables $s'\to\sigma=s'-s$, giving

$$\begin{aligned}\langle\kappa|\nabla\theta|^2\rangle_{\Omega,\infty}&=\lim_{t\to\infty}\frac{1}{t}\int_0^t\mathrm{d}s\int_{-s}^0\mathrm{d}\sigma\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s+\sigma)\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}\\&=\lim_{t\to\infty}\frac{1}{t}\int_{-t}^0\mathrm{d}\sigma\int_{-\sigma}^t\mathrm{d}s\langle\tilde{S}_L(s,s)\tilde{S}_L(s,s+\sigma)\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}\end{aligned}\tag{B 13}$$

after switching the order of integration. This can be rewritten as

$$\langle\kappa|\nabla\theta|^2\rangle_{\Omega,\infty}=\lim_{t\to\infty}\int_{-t}^0\mathrm{d}\sigma\left(1+\frac{\sigma}{t}\right)\langle\langle\tilde{S}_L(0,0)\tilde{S}_L(0,\sigma)\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}\rangle_{[-\sigma,t]},\tag{B 14}$$

where we have introduced the time average over the interval $[-\sigma,t]$, $\sigma<0$,

$$\langle f(0)\rangle_{[-\sigma,t]}:=\frac{1}{t+\sigma}\int_{-\sigma}^t\mathrm{d}s f(s).\tag{B 15}$$

Assuming that the integrand in (B 14) is absolutely integrable uniformly in t, then the dominated convergence theorem applies and we obtain

$$\langle \kappa |\nabla\theta|^2\rangle_{\Omega,\infty} = \int_{-\infty}^{0} \mathrm{d}\sigma\, \langle \langle \tilde{S}_L(0,0)\tilde{S}_L(0,\sigma)\rangle^{\mathrm{T}}_{\mathbb{E},\Omega}\rangle_\infty. \tag{B 16}$$

Expression (B 16) is equivalent to formula (2.15) in the main text.

We now derive as a special case of (B 16) the result (2.14) for a statistical steady state maintained by a random scalar source, delta-correlated in time. In addition to the delta-in-time covariance (2.13), the source is assumed to satisfy the condition that the ensemble average is zero and also that the space integral is zero in every realization,

$$\int_\Omega \mathrm{d}^d x\, \tilde{S}(\boldsymbol{x}, t) = 0 \quad \text{a.s.} \tag{B 17}$$

This latter condition means that there is no net input of scalar by the source. As a consequence, the truncation terms in (B 16) vanish identically. Averaging the formula (B 16) over the random source, the delta covariance (2.13) then implies that

$$\kappa \langle |\nabla\theta|^2\rangle_{\Omega,\infty,S} = \frac{1}{V}\int_\Omega \mathrm{d}^d x\ C_S(\boldsymbol{x},\boldsymbol{x}), \tag{B 18}$$

where integration of the delta function $\delta(\sigma)$ over $\sigma \in [-\infty, 0]$ gives a factor of $1/2$.

B.2. *Free decay relation (2.17)*

Next, we derive from our general FDR the relation (2.17) of Sawford *et al.* (2005) and Buaria *et al.* (2016) for a decaying passive scalar with a random initial linear profile satisfying the statistical isotropy condition (2.16). To obtain this result, we note that for any random variable $\tilde{X}$, $\mathrm{Var}(\tilde{X}) = \mathbb{E}|\tilde{X}^{(1)} - \tilde{X}^{(2)}|^2/2$, where $\tilde{X}^{(1)}$, $\tilde{X}^{(2)}$ are two independent random variables identically distributed as $\tilde{X}$. Therefore, our FDR (2.12) in the case of vanishing scalar sources and random scalar initial values can be re-expressed as

$$\kappa \int_0^t \mathrm{d}s \langle |\nabla\tilde{\theta}(s)|^2\rangle_\Omega = \frac{1}{4}\langle \mathbb{E}|\tilde{\theta}_0(\tilde{\boldsymbol{\xi}}^{(1)}_{t,0}) - \tilde{\theta}_0(\tilde{\boldsymbol{\xi}}^{(2)}_{t,0})|^2\rangle_\Omega. \tag{B 19}$$

Assuming that $\tilde{\theta}_0(\boldsymbol{x}) = \tilde{\boldsymbol{G}}\cdot\boldsymbol{x}$ and averaging over random initial data using (2.16) gives

$$\kappa \int_0^t \mathrm{d}s \langle |\nabla\tilde{\theta}(s)|^2\rangle_{\Omega,\theta_0} = \frac{1}{4}G^2\langle \mathbb{E}^{1,2}|\tilde{\boldsymbol{\xi}}^{(1)}_{t,0} - \tilde{\boldsymbol{\xi}}^{(2)}_{t,0}|^2\rangle_\Omega, \tag{B 20}$$

which is (2.17).

The interest of this relation is that it directly connects the temporal evolution of the mean scalar dissipation to two-particle dispersion of stochastic Lagrangian trajectories. For example, it relates dissipative anomalies of scalar fluctuations and kinetic energy if the Prandtl number is fixed and if the dispersion on the right-hand side of (B 20) exhibits a Richardson scaling $\sim\varepsilon t^3$ with ε independent of ν. This relation can be used to establish equivalence of spontaneous stochasticity and anomalous scalar dissipation for situations that satisfy the specific assumptions under which it is derived. One can obtain such a relation for more general initial data than linear profiles by instead

assuming only that the initial scalar field is smooth and that its second-order structure function satisfies the pair of inequalities for all $\boldsymbol{x}, \boldsymbol{x}' \in \Omega$ that

$$c_{\theta_0}|\boldsymbol{x}-\boldsymbol{x}'|^2 \leqslant \langle|\tilde{\theta}_0(\boldsymbol{x})-\tilde{\theta}_0(\boldsymbol{x}')|^2\rangle_{\theta_0} \leqslant C_{\theta_0}|\boldsymbol{x}-\boldsymbol{x}'|^2, \tag{B 21}$$

for some constants $0 < c_{\theta_0} < C_{\theta_0} < \infty$. Averaging (B 19) over such initial data then yields

$$\frac{1}{4}c_{\theta_0}\langle\mathbb{E}|\tilde{\boldsymbol{\xi}}^{(1)}_{t,0}-\tilde{\boldsymbol{\xi}}^{(2)}_{t,0}|^2\rangle_{\Omega} \leqslant \kappa\int_0^t \mathrm{d}s\langle|\nabla\tilde{\theta}(s)|^2\rangle_{\Omega,\theta_0} \leqslant \frac{1}{4}C_{\theta_0}\langle\mathbb{E}|\tilde{\boldsymbol{\xi}}^{(1)}_{t,0}-\tilde{\boldsymbol{\xi}}^{(2)}_{t,0}|^2\rangle_{\Omega}. \tag{B 22}$$

This gives upper and lower bounds for the cumulative scalar dissipation directly in terms of the two-particle dispersion, which again relate anomalous scalar dissipation to spontaneous stochasticity. If there is a smooth random scalar source whose structure function satisfies bounds similar to (B 21), then one can obtain analogous bounds relating cumulative scalar dissipation to the time-integrated two-particle dispersion.

Appendix C. Numerical methods

C.1. *Methods for § 3*

In order to integrate backward Itô stochastic differential equations (SDEs) (2.3) of the form

$$\hat{\mathrm{d}}\tilde{\boldsymbol{\xi}}(s) = \boldsymbol{u}(\tilde{\boldsymbol{\xi}}(s), s)\,\mathrm{d}s + \sqrt{2\kappa}\,\hat{\mathrm{d}}\widetilde{\boldsymbol{W}}(s), \tag{C 1}$$

we use a reflected time $\hat{s} = t_f - s$ which converts them into forward Itô stochastic differential equations SDEs. The latter are integrated with the standard Euler–Maruyama scheme (Kloeden & Platen 1992). We solved (C 1) with the Euler–Maruyama scheme, which for additive noise is first order in both the weak and strong senses (Kloeden & Platen 1992). The turbulent velocity field in (C 1) was retrieved from the Johns Hopkins Turbulence Database with the `getVelocity` function, which returns velocities at requested points interpolated in space by sixth-order Lagrange polynomials and in time by piecewise-cubic Hermite polynomials. We used a time step of $\Delta s = 6.6 \times 10^{-4}$, or one-third of the time between database frames. We calculated statistics with averages over independent solutions of (C 1) and, to test for weak convergence in the time integration, we doubled Δs, with relative change $<0.1\,\%$.

To estimate particle dispersions and transition probability densities $p_y(y', 0|\boldsymbol{x}, t_f)$, we used N-sample ensembles of stochastic trajectories $\tilde{\boldsymbol{\xi}}_n(s)$, $n = 1, \ldots, N$, solving the above SDE. The particle dispersions were calculated by the unbiased estimators

$$\begin{aligned}\mathbb{E}^{1,2}[|\tilde{\boldsymbol{\xi}}^{(1)}(s)-\tilde{\boldsymbol{\xi}}^{(2)}(s)|^2] &\doteq \frac{2}{N(N-1)}\sum_{n<m}|\tilde{\boldsymbol{\xi}}_n(s)-\tilde{\boldsymbol{\xi}}_m(s)|^2 \\ &= \frac{2}{N-1}\sum_{n=1}^{N}|\tilde{\boldsymbol{\xi}}_n(s)-\overline{\boldsymbol{\xi}}_N(s)|^2,\end{aligned} \tag{C 2}$$

with $\overline{\boldsymbol{\xi}}_N(s) = (1/N)\sum_{n=1}^{N}\tilde{\boldsymbol{\xi}}_n(s)$ as the sample mean. For large N, these are nearly the same as

$$\mathbb{E}^{1,2}[|\tilde{\boldsymbol{\xi}}^{(1)}(s)-\tilde{\boldsymbol{\xi}}^{(2)}(s)|^2] \doteq \frac{2}{N}\sum_{n=1}^{N}|\tilde{\boldsymbol{\xi}}_n(s)-\mathbb{E}[\tilde{\boldsymbol{\xi}}(s)]|^2, \tag{C 3}$$

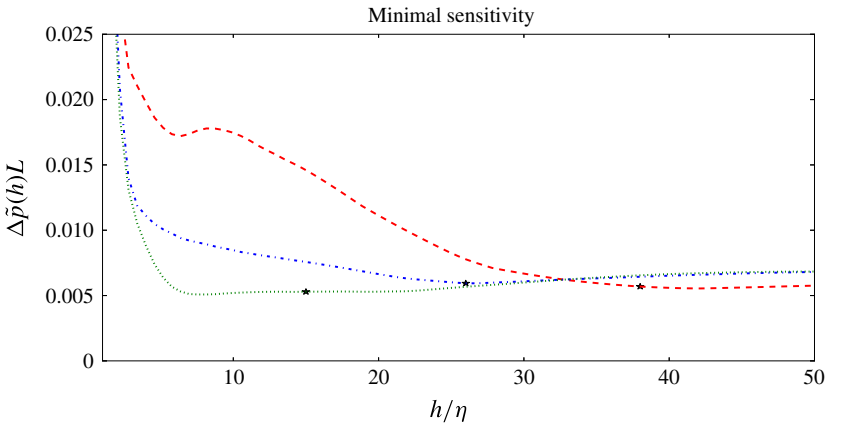

FIGURE 2. (Colour online) The L^1-differences of PDF estimates for successive bandwidths h_j are plotted for $Pr = 0.1$ (green, dot, $\cdots\cdots$), 1.0 (blue, dash–dot, — · —) and 10 (red, dash, – – –). The optimum bandwidths correspond to the local minima at the smallest h_j values, marked with a filled star ($\star$) on the plots.

which is twice the sample average over N independent random variables each with the same distribution as $|\tilde{\boldsymbol{\xi}}(s) - \mathbb{E}[\tilde{\boldsymbol{\xi}}(s)]|^2$. Our error bars for the particle dispersion are thus taken to be twice the s.e.m. for this N-sample average.

To estimate the position PDFs, we used kernel density estimator methods (Silverman 1986). For the y-coordinate at time 0 of the particle $\tilde{\boldsymbol{\xi}}(0) = (\tilde{\xi}(0), \tilde{\eta}(0), \tilde{\zeta}(0))$ started at $\boldsymbol{x}$ at time t_f, one has for a grid y_i' of possible y-values

$$\tilde{p}_h(y_i', 0|\boldsymbol{x}, t_f) \doteq \frac{1}{N}\sum_{n=1}^{N} K_h(\tilde{\eta}_n(0) - y_i'), \tag{C 4}$$

where $K_h(y) = (1/h)K(y/h)$ is a filter kernel with bandwidth h. We take K to be a Gaussian with unit variance and we choose the bandwidth h by the 'principle of minimal sensitivity' from renormalization-group theory (Stevenson 1981). The latter procedure is based upon the observation that, when the number N of samples is sufficiently large for the average in (C 4) to be converged to the convolution $(K_h * p)(y_i', 0|\boldsymbol{x}, t_f)$, then the result will be independent of h for any value less than the scale of variation of the limit PDF. Since this is an exact invariance property of the limiting result, the 'principle of minimal sensitivity' selects the optimal bandwidth h_* for finite N so that varying the bandwidth has minimal effect on the PDF estimate. Precisely, one picks h_* by considering a decreasing sequence of candidate values h_j, computing the L_1-difference $\Delta\tilde{p}(h_j) := \|\tilde{p}_{h_j} - \tilde{p}_{h_{j-1}}\|_{L^1}$ for successive bandwidths, and picking h_* where $\Delta\tilde{p}(h_j)$ is most nearly flat. This procedure is illustrated in figure 2 for the particle position PDFs that were presented in figure 1(f), but using $N = 1024$ samples. The L^1-differences are plotted versus h_j in figure 2 for the three choices of Prandtl number. The bandwidths chosen correspond to the local minima for each curve at the smallest h_j value indicated by the star ($\star$) on the graph, i.e. $h_*/\eta \approx 16$ for $Pr = 0.1$, $h_*/\eta \approx 26$ for $Pr = 1$, $h_*/\eta \approx 39$ for $Pr = 10$. In some cases, we did not observe local minima as in figure 2, and in those instances our procedure was to select as 'optimal' bandwidth h_* the smallest h_j in an interval where the plot of $\Delta\tilde{p}(h_j)$ versus h_j had a slope of magnitude less than 0.01.

Finally, after selecting the optimal bandwidth h_*, we obtained the sample-size error for the PDF estimate as the s.e.m. for the sample average (C 4) over N independent random variables identically distributed as $K_{h_*}(\tilde{\eta}(0) - y_i')$, with bandwidth h_* fixed. In addition to this statistical error, an additional source of error arises from the choice of h_*. To assess this, we recalculated the kernel density estimator with a 10 % increase in bandwidth, or $1.1h_*$, and took the absolute difference in the two PDF estimates as a measure of the error associated with small variations in bandwidth. The two types of errors were found to have more or less comparable magnitudes, and the error bars in figure 1(*e*,*f*) represent the total error obtained from their sum.

REFERENCES

ALBEVERIO, S. & BELOPOLSKAYA, YA. 2010 Generalized solutions of the Cauchy problem for the Navier–Stokes system and diffusion processes. *Cubo (Temuco)* **12** (2), 77–96.

BATCHELOR, G. K. 1951 Pressure fluctuations in isotropic turbulence. *Math. Proc. Camb. Phil. Soc.* **47** (2), 359–374.

BERG, J., LÜTHI, B., MANN, J. & OTT, S. 2006 Backwards and forwards relative dispersion in turbulent flow: an experimental investigation. *Phys. Rev.* E **74**, 016304.

BERKOVITZ, L. D. 1974 Lower semicontinuity of integral functionals. *Trans. Am. Math. Soc.* **192**, 51–57.

BERNARD, D., GAWĘDZKI, K. & KUPIAINEN, A. 1998 Slow modes in passive advection. *J. Stat. Phys.* **90**, 519–569.

BIFERALE, L., BOFFETTA, G., CELANI, A., LANOTTE, A. & TOSCHI, F. 2005 Particle trapping in three-dimensional fully developed turbulence. *Phys. Fluids* **17** (2), 021701.

BIFERALE, L., LANOTTE, A. S., SCATAMACCHIA, R. & TOSCHI, F. 2014 Intermittency in the relative separations of tracers and of heavy particles in turbulent flows. *J. Fluid Mech.* **757**, 550–572.

BITANE, R., HOMANN, H. & BEC, J. 2013 Geometry and violent events in turbulent pair dispersion. *J. Turbul.* **14** (2), 23–45.

BOFFETTA, G. & SOKOLOV, I. M. 2002 Relative dispersion in fully developed turbulence: the Richardson's law and intermittency corrections. *Phys. Rev. Lett.* **88**, 094501.

BRAIDES, A. 2002 *Gamma-Convergence for Beginners*. Oxford University Press.

BUARIA, D., YEUNG, P. K. & SAWFORD, B. L. 2016 A Lagrangian study of turbulent mixing: forward and backward dispersion of molecular trajectories in isotropic turbulence. *J. Fluid Mech.* **799**, 352–382.

CHAVES, M., GAWĘDZKI, K., HORVAI, P., KUPIAINEN, A. & VERGASSOLA, M. 2003 Lagrangian dispersion in Gaussian self-similar velocity ensembles. *J. Stat. Phys.* **113** (5–6), 643–692.

CONSTANTIN, P. & IYER, G. 2008 A stochastic Lagrangian representation of the three-dimensional incompressible Navier–Stokes equations. *Commun. Pure Appl. Maths* **61**, 330–345.

CONSTANTIN, P. & IYER, G. 2011 A stochastic-Lagrangian approach to the Navier–Stokes equations in domains with boundary. *Ann. Appl. Probab.* **21**, 1466–1492.

CORRSIN, S. 1951 On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. *J. Appl. Phys.* **22** (4), 469–473.

DAVILA, J. & VASSILICOS, J. C. 2003 Richardson's pair diffusion and the stagnation point structure of turbulence. *Phys. Rev. Lett.* **91** (14), 144501.

DONZIS, D. A., SREENIVASAN, K. R. & YEUNG, P. K. 2005 Scalar dissipation rate and dissipative anomaly in isotropic turbulence. *J. Fluid Mech.* **532**, 199–216.

DRIVAS, T. D. & EYINK, G. L. 2017 A Lagrangian fluctuation–dissipation relation for scalar turbulence. Part II. Wall-bounded flows. *J. Fluid Mech.* (submitted). arXiv:1703.08133.

EYINK, G., VISHNIAC, E., LALESCU, C., ALUIE, H., KANOV, K., BÜRGER, K., BURNS, R., MENEVEAU, C. & SZALAY, A. 2013 Flux-freezing breakdown in high-conductivity magnetohydrodynamic turbulence. *Nature* **497** (7450), 466–469.

EYINK, G. L. 2001 Dissipation in turbulent solutions of 2D Euler equations. *Nonlinearity* **14** (4), 787–802.

EYINK, G. L. 2009 Stochastic line motion and stochastic flux conservation for nonideal hydromagnetic models. *J. Math. Phys.* **50** (8), 083102.

EYINK, G. L. 2010 Stochastic least-action principle for the incompressible Navier–Stokes equation. *Physica* D **239**, 1236–1240.

EYINK, G. L. 2011 Stochastic flux freezing and magnetic dynamo. *Phys. Rev.* E **83** (5), 056405.

EYINK, G. L. & DRIVAS, T. D. 2015*a* Quantum spontaneous stochasticity. arXiv:1509.04941.

EYINK, G. L. & DRIVAS, T. D. 2015*b* Spontaneous stochasticity and anomalous dissipation for Burgers equation. *J. Stat. Phys.* **158** (2), 386–432.

EYINK, G. L. & DRIVAS, T. D. 2017 A Lagrangian fluctuation–dissipation relation for scalar turbulence. Part III. Turbulent Rayleigh–Bénard convection. *J. Fluid Mech.* (submitted). arXiv:1703.09604.

FABER, T. & VASSILICOS, J. C. 2009 Turbulent pair separation due to multiscale stagnation point structure and its time asymmetry in two-dimensional turbulence. *Phys. Fluids* **21** (1), 015106.

FALKOVICH, G. & FRISHMAN, A. 2013 Single flow snapshot reveals the future and the past of pairs of particles in turbulence. *Phys. Rev. Lett.* **110** (21), 214502.

FALKOVICH, G., GAWĘDZKI, K. & VERGASSOLA, M. 2001 Particles and fields in fluid turbulence. *Rev. Mod. Phys.* **73**, 913–975.

FLORESCU, L. C. & GODET-THOBIE, C. 2012 *Young Measures and Compactness in Measure Spaces*. de Gruyter.

FRIEDMAN, A. 2006 *Stochastic Differential Equations and Applications*. Dover.

FRISCH, U. 1995 *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press.

GAWĘDZKI, K. 2008 Soluble models of turbulent transport. In *Non-Equilibrium Statistical Mechanics and Turbulence* (ed. J. Cardy, G. Falkovich & K. Gawędzki), London Mathematical Society Lecture Note Series, vol. 355, pp. 44–107. Cambridge University Press.

GAWĘDZKI, K. 2013 Fluctuation relations in stochastic thermodynamics. arXiv:1308.1518.

GAWĘDZKI, K. & VERGASSOLA, M. 2000 Phase transition in the passive scalar advection. *Physica* D **138**, 63–90.

HARTMAN, P. 2002 *Ordinary Differential Equations*, 2nd edn. SIAM.

IKEDA, N. & WATANABE, S. 1989 *Stochastic Differential Equations and Diffusion Processes*. North-Holland.

ITÔ, K. 1984 *An Introduction to Probability Theory*. Cambridge University Press.

JIŘINA, M. 1959 On regular conditional probabilities. *Czech. Math. J.* **9** (3), 445–451.

JUCHA, J., XU, H., PUMIR, A. & BODENSCHATZ, E. 2014 Time-reversal-symmetry breaking in turbulence. *Phys. Rev. Lett.* **113** (5), 054501.

KANEDA, Y., ISHIHARA, T., YOKOKAWA, M., ITAKURA, K. & UNO, A. 2003 Energy dissipation rate and energy spectrum in high resolution direct numerical simulations of turbulence in a periodic box. *Phys. Fluids* **15**, L21–L24.

KLOEDEN, P. E. & PLATEN, E. 1992 *Numerical Solutions of Stochastic Differential Equations*. Springer.

KOLMOGOROV, A. N. 1941*a* Dissipation of energy in locally isotropic turbulence. *Dokl. Akad. Nauk SSSR* **32** (1), 16–18.

KOLMOGOROV, A. N. 1941*b* Equations of turbulent motion in an incompressible fluid. *Dokl. Akad. Nauk SSSR* **30**, 299–303.

KOLMOGOROV, A. N. 1941*c* The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR* **30** (4), 299–303.

KRAICHNAN, R. H. 1968 Small-scale structure of a scalar field convected by turbulence. *Phys. Fluids* **11**, 945–953.

KUNITA, H. 1997 *Stochastic Flows and Stochastic Differential Equations*, Cambridge Studies in Advanced Mathematics, vol. 24. Cambridge University Press.

KUPIAINEN, A. 2003 Nondeterministic dynamics and turbulent transport. *Ann. Henri Poincaré* **4** (Suppl. 2), S713–S726.

LE JAN, Y. & RAIMOND, O. 2002 Integration of Brownian vector fields. *Ann. Probab.* **30**, 826–873.

LE JAN, Y. & RAIMOND, O. 2004 Flows, coalescence and noise. *Ann. Probab.* **32**, 1247–1315.

LI, Y., PERLMAN, E., WAN, M., YANG, Y., BURNS, R., MENEVEAU, C., CHEN, S., SZALAY, A. & EYINK, G. L. 2008 A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence. *J. Turbul.* **9**, N31.

LIONS, P. L. 1996 *Mathematical Topics in Fluid Mechanics: Volume 1: Incompressible Models*. Oxford University Press.

NOVIKOV, E. A. 1965 Functionals and the random-force method in turbulence theory. *Sov. Phys. JETP* **20** (5), 1290–1294.

OBUKHOV, A. M. 1949 Structure of the temperature field in turbulent flow. *Izv. Akad. Nauk. SSSR, Ser. Geogr. i Geofiz.* **13**, 58–69.

OKSENDAL, B. 2013 *Stochastic Differential Equations: An Introduction with Applications*. Springer Science & Business Media.

ONSAGER, L. 1945 The distribution of energy in turbulence. *Phys. Rev.* **68**, 286.

ONSAGER, L. 1949 *Nuovo Cim. Suppl.* **VI**, 279–287.

PEARSON, B. R., KROGSTAD, P.-A. & VAN DE WATER, W. 2002 Measurements of the turbulent energy dissipation rate. *Phys. Fluids* **14**, 1288–1290.

REZAKHANLOU, F. 2014 Regular flows for diffusions with rough drifts. arXiv:1405.5856.

RICHARDSON, L. F. 1926 Atmospheric diffusion shown on a distance-neighbor graph. *Proc. R. Soc. Lond.* A **110**, 709–737.

ROBERT, R. 1991 A maximum-entropy principle for two-dimensional perfect fluid dynamics. *J. Stat. Phys.* **65** (3–4), 531–553.

ROBERT, R. & SOMMERIA, J. 1991 Statistical equilibrium states for two-dimensional flows. *J. Fluid Mech.* **229**, 291–310.

RUDIN, W. 2006 *Functional Analysis*. McGraw-Hill.

SAFFMAN, P. G. 1960 On the effect of the molecular diffusivity in turbulent diffusion. *J. Fluid Mech.* **8** (02), 273–283.

SALMON, R. 1988 Hamiltonian fluid mechanics. *Annu. Rev. Fluid Mech.* **20** (1), 225–256.

SAWFORD, B. 2001 Turbulent relative dispersion. *Annu. Rev. Fluid Mech.* **33** (1), 289–317.

SAWFORD, B. L., YEUNG, P. K. & BORGAS, M. S. 2005 Comparison of backwards and forwards relative dispersion in turbulence. *Phys. Fluids* **17**, 095109.

SCHUSTER, H. G., KLAGES, R., JUST, W. & JARZYNSKI, C. 2013 *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*. Wiley.

SHLESINGER, M. F., WEST, B. J. & KLAFTER, J. 1987 Lévy dynamics of enhanced diffusion: application to turbulence. *Phys. Rev. Lett.* **58** (11), 1100.

SILVERMAN, B. W. 1986 *Density Estimation for Statistics and Data Analysis*. CRC Press.

SOMMERIA, J., STAQUET, C. & ROBERT, R. 1991 Final equilibrium state of a two-dimensional shear layer. *J. Fluid Mech.* **233**, 661–689.

SREENIVASAN, K. R. 1998 An update on the energy dissipation rate in isotropic turbulence. *Phys. Fluids* **10** (2), 528–529.

SREENIVASAN, K. R. & SCHUMACHER, J. 2010 Lagrangian views on turbulent mixing of passive scalars. *Phil. Trans. R. Soc. Lond.* A **368** (1916), 1561–1577.

STEVENSON, R. M. 1981 Optimized perturbation theory. *Phys. Rev.* D **23** (12), 2916.

TAYLOR, G. I. 1917 Observations and speculations on the nature of turbulent motion. In *Reports and Memoranda of the Advisory Committee for Aeronautics, No. 345* (ed. G. K. Batchelor), Reprinted in The Scientific Papers of Sir Geoffrey Ingram Taylor: Volume 2, Meteorology, Oceanography and Turbulent Flow. Cambridge University Press, 1960.

TAYLOR, G. I. 1922 Diffusion by continuous movements. *Proc. Lond. Math. Soc.* **20** (1), 196–212.

TAYLOR, G. I. & GREEN, A. E. 1937 Mechanism of the production of small eddies from large ones. *Proc. R. Soc. Lond.* A **158** (895), 499–521.

THALABARD, S., KRSTULOVIC, G. & BEC, J. 2014 Turbulent pair dispersion as a continuous-time random walk. *J. Fluid Mech.* **755**, R4.

TRAN, C. V. & DRITSCHEL, D. G. 2006 Vanishing enstrophy dissipation in two-dimensional Navier–Stokes turbulence in the inviscid limit. *J. Fluid Mech.* **559**, 107–116.

TSINOBER, A. 2009 *An Informal Conceptual Introduction to Turbulence*. Springer.

VALADIER, M. 1973 Désintégration d'une mesure sur un produit. *C. R. Acad. Sci.* A **276**, 33–35.

VALADIER, M. 1994 A course on Young measures. *Rend. Istit. Mat. Univ. Trieste, Supplemento* **XXVI**, 349–394.

VANDEN-EIJNDEN, E. W. & VANDEN-EIJNDEN, E. 2000 Generalized flows, intrinsic stochasticity and turbulent transport. *Proc. Natl Acad. Sci. USA* **97**, 8200–8205.

VANDEN-EIJNDEN, E. W. & VANDEN-EIJNDEN, E. 2001 Turbulent Prandtl number effect on passive scalar advection. *Physica* D **152–153**, 636–645.

YAGLOM, A. M. 1949 On the local structure of the temperature field in a turbulent flow. *Dokl. Acad. Sci. USSR* **69** (6), 743.

YEUNG, P. K., DONZIS, D. A. & SREENIVASAN, K. R. 2005 High-Reynolds-number simulation of turbulent mixing. *Phys. Fluids* **17** (8), 081703.

YU, H., KANOV, K., PERLMAN, E., GRAHAM, J., FREDERIX, E., BURNS, R., SZALAY, A., EYINK, G. L. & MENEVEAU, C. 2012 Studying Lagrangian dynamics of turbulence using on-demand fluid particle tracking in a public turbulence database. *J. Turbul.* **12**, N13.