SIAM REVIEW
Vol. 62, No. 4, pp. 781–816

Stochastic Sensitivity: A Computable Lagrangian Uncertainty Measure for Unsteady Flows*

Sanjeeva Balasuriya†

Abstract. Uncertainties in velocity data are often ignored when computing Lagrangian particle trajectories of fluids. Modeling these as noise in the velocity field leads to a random deviation from each trajectory. This deviation is examined within the context of small (multiplicative) stochasticity applying to a two-dimensional unsteady flow operating over a finite time. These assumptions are motivated precisely by standard availability expectations of realistic velocity data. Explicit expressions for the deviation's expected size and anisotropy are obtained using an Itô calculus approach, thereby characterizing the uncertainty in the Lagrangian trajectory's final location with respect to lengthscale and direction. These provide a practical methodology for ascribing spatially nonuniform uncertainties to predictions of flows, and also provide new tools for extracting fluid regions that remain robust under velocity fluctuations.

Key words. uncertainty quantification, Lagrangian trajectories, Lagrangian coherent structures, subgrid uncertainty, Lagrangian data assimilation

AMS subject classifications. 34F05, 60H10, 76R50, 60H30, 60H40, 76F99

DOI. 10.1137/18M1222922

Contents

*Received by the editors October 25, 2018; accepted for publication (in revised form) December 16, 2019; published electronically November 3, 2020.
https://doi.org/10.1137/18M1222922
Funding: This work was partially supported by the Australian Research Council through grant DP200101764.

†School of Mathematical Sciences, University of Adelaide, Adelaide SA 5005, Australia (sanjeevabalasuriya@yahoo.com, http://maths.adelaide.edu/sanjeeva.balasuriya/).

1. Introduction. Global ocean/atmospheric models rely on velocity data obtained on a low resolution grid; this data is *Eulerian* in the nomenclature of fluid mechanics because it is given in terms of spatial location and time. The impact of these features on *Lagrangian* ("following-the-flow") trajectories is of particular interest, as it defines where assorted quantities of interest (heat, pollutants, plankton, spores, ozone, energy, etc.) go. "Coherent" regions of these quantities, and their movement, profoundly impact the environment and climate and are therefore a significant area of study. Computation of these "Lagrangian coherent structures" using deterministic methods for given unsteady Eulerian velocity data—ignoring uncertainties in the data—are well established [47, 62, 68, 5, 64, 12, 45, 71]. However, an issue that has come to the fore recently [19, 41] is the fact that *subgrid* effects (since data is only available on a spatial grid) must impact any predictions made. Incorporating these uncertainties into Lagrangian conclusions is an aspect of the so-called stochastic parametrization problem, which attempts to parametrize the uncertainties in a probabilistic way and feed this information into the grid-scale deterministic model. Within existing coherent structure work, however, none explicitly characterizes the effect of velocity uncertainties on trajectories. The main goal of this article is to address this deficiency by proposing a method for quantifying the impact of uncertainty in Eulerian velocity data on Lagrangian trajectories.

Inspired by uncertainty in velocity data, this article provides a set of rigorous computational tools for assessing the uncertainty of Lagrangian trajectories under stochastic variation in the Eulerian velocity field. The principal measure provided—an uncertainty field on the set of initial conditions—can be viewed as an uncertainty lengthscale prediction of eventual Lagrangian locations, which is also useful in ascribing uncertainties to any conclusions reached (e.g., as a confidence weighting if the predicted trajectories are used in a Lagrangian data assimilation algorithm [66, 55, 70, 59, 3]). Alternatively, this field quantifies the stability of trajectories to random ongoing perturbations, enabling the separation of the flow domain according to a measure associated with such robustness. The anisotropy (directionality in the uncertainty) is also captured.

A natural way to model the velocity uncertainty and consequent stochastic Lagrangian trajectories is to use an Itô stochastic ordinary differential equation whose solutions x correspond to particle trajectories. While such equations have been studied extensively within the context of financial mathematics, and in theoretical analyses of noise-induced bifurcation and related phenomena [16, 15, 21, 23, 17], escape/passage times from sets [33, 60, 32], and the control of such systems [61, 1, 2], they have been seldom used in the Lagrangian coherent structure community. However, there is emerging interest in this modeling approach [26, 6, 9, 50, 20]. The drift term would be the deterministic velocity, and the form of the additional diffusion term could be chosen based on any information available in the application of interest and would relate to the uncertainty of the Eulerian velocity. For example, it may be known that the velocity data in certain areas or over certain times has larger uncertainties than other data (e.g., cloud cover impacts on satellite observations of oceanic data). In the absence of any insight into the choice of a particular model, the generic choice could be canonical Brownian motion (see, e.g., [24]). Having made such a choice, the deviation of the stochastic trajectories from the deterministic one, scaled in the limit of small noise, is clearly the random variable of interest. Unlike in the classical Freidlin–Wentzell large deviation approaches [33, 22, 25], the aim here is not to theoretically analyze the probabilities of large deviations, but rather to assign to each

initial condition x *computable* measures of sensitivity toward stochasticity which make straightforward intuitive sense. This circumvents brute-force Monte Carlo numerics on stochastic systems, which are well known to be computationally expensive [31]. There are closely related analyses to this paper [6, 14, 9] that address uncertainties in advected curves and stable/unstable manifolds using formal calculations. In this article, however, an uncertainty measure across the global flow domain, as well as a description of its anisotropy, are developed and rigorously justified.

Section 2 develops the theory; in this first analysis using these ideas, the development is restricted to two-dimensional flows where data is confined to a finite time interval and the stochasticity is assumed to be small. These assumptions are motivated by the facts that most available velocity data is two-dimensional and is available only over a finite time, and, that if one is to have any faith in the data, the uncertainties *must* be considered to be small. However, generality is allowed within this framework the velocity field does not need to be area-preserving nor to possess any particular form of time-dependence, and the diffusion matrix of the perturbing stochasticity can depend on the location and time, while being anisotropic. Thus, the noise is permitted to be multiplicative. Again, these conditions are necessary in our application area; velocity data often has deviations from area-preservation, has fluctuations in time, and has uncertainties which may depend on position and time (e.g., because of cloud cover, nonuniformity in measurement error across a camera's field of vision, or certainty at gridpoints but uncertainty elsewhere). Under generic smoothness assumptions on the velocity and diffusion matrices, a *computable* expression for the scaled variance, projected in a general direction, is obtained. By maximizing this *anisotropic uncertainty* across all directions, an explicit formula is obtained for the *stochastic sensitivity*, in the sense of quantifying the sensitivity of the deterministic system toward stochasticity. This is a scalar field in relation to all initial conditions x, thereby providing an analytical estimate for the spatial-dependence of the stochastic impact. In this way, information on the solutions corresponding to Dirac-delta initial conditions of the Fokker–Planck evolution equation emerge from this analysis. Moreover, if an estimate for the diffusive/noise parameter ε is available, the stochastic sensitivity field allows for defining *robust sets* at time zero: regions of initial conditions such that the expected uncertainty of Lagrangian trajectories with respect to the flow over the given time $[0, T]$ is less than a specified threshold lengthscale L. These sets identify regions in which certainty of Lagrangian conclusions can be ascribed in relation to specified tolerance levels of the diffusion ε, lengthscale L, and time-of-flow T.

Section 3 demonstrates the usage of the theoretical results in a model that is often used as a testbed for Lagrangian coherent structure analysis: the double-gyre [69]. All theoretical expressions—the stochastic sensitivity, its anisotropy, its relationship to a probability density function obtained by evolving the relevant Fokker–Planck equation, and robust sets—are computed and validated in this section. Monte Carlo simulations are performed which additionally verify the theory, and other investigations relating these results to known properties of the double-gyre (e.g., finite-time Lyapunov exponent calculations) are also presented. Section 4 concludes with some comments on theoretical extensions, as well as the immense potential for applying the theory to the problem of identifying Lagrangian coherence subject to inevitable uncertainties in velocity data sets. Moreover, since the stochastic sensitivity is an uncertainty field over the set of initial conditions, it is a new tool which can assign uncertainty levels to each deterministic Lagrangian trajectory calculation, offering a confidence weighting of trajectories which can be useful in Lagrangian data assimila-

tion [66, 55, 70, 59, 3]. In summary, a fundamental new theoretical—but computable for realistic data—set of tools which addresses the pressing need to quantify uncertainty in Lagrangian trajectories is developed in this article.

2. Stochastic Sensitivity Measures. Suppose velocity data u is available in a two-dimensional spatial domain over a finite time, chosen here to be $[0, T]$. This is *Eulerian* data, in the sense that u is available as a spatio-temporal function. Typically, data will be both spatially and temporally discrete, and it will have uncertainties. The goal is to examine *Lagrangian* trajectories generated by u from time 0 to T, which is governed by

$$\frac{dx}{dt} = u(x,t)\,, \tag{2.1}$$

with initial conditions chosen in the spatial domain Ω_0 (an open connected subset of $\mathbb{R}^2$). An equivalent integral representation of (2.1) is in terms of the flow map $F_{t_1}^{t_2}$, which takes an initial condition at time t_1 to its final location at time t_2. An initial point $x \in \Omega_0$ is mapped to its location at a general time $t \in [0, T]$ by

$$F_0^t(x) = x + \int_0^t u\left(F_0^\tau(x), \tau\right)\,\mathrm{d}\tau\,, \tag{2.2}$$

where $F_0^t(x) \in \Omega_t := F_0^t(\Omega_0)$. In writing (2.2), there is an implicit understanding that u is Lipschitz continuous and thus trajectories are well-defined. In practical applications—because data is available on a spatial grid—there is usually implicit smoothing/interpolation of u to all relevant x, thereby hiding any such issue. Using trajectories of (2.1), or equivalently taken from the flow map $F_0^T(\boldsymbol{.})$, there is a multitude of methods for detecting spatial sets which are important with respect to the finite-time flow. These include finite-time Lyapunov exponents [69], sets which are associated with extremal attraction or repulsion [47], coherent sets defined in terms of the transfer operator [40, 35], stable/unstable manifolds obtained by appropriate time extensions [5], and assorted diagnostics derived from trajectories of (2.1). Extensive reviews of these methods are available [5, 47, 64, 12].

Here, the intent is to specifically examine the impact of uncertainty in u, which exists because of observational errors as well as interpolation errors to subgrid levels. Consider instead the stochastic differential equation

$$\mathrm{d}y_t = u(y_t, t)\mathrm{d}t + \varepsilon\sigma(y_t, t)\mathrm{d}W_t\,, \tag{2.3}$$

where y_t evolves in Ω_t for $t \in [0, T]$. The nondimensional noise parameter ε satisfies $0 < \varepsilon \ll 1$, and the noise is permitted to be *multiplicative* in that the 2×2 diffusion matrix σ is permitted to depend on space and time. This may be *specified* based on any additional information that one has (e.g., in oceanographic flows, dispersion depends on the bottom topography as well as the presence of nearby land boundaries [52]). If there is no a priori insight into the nature of the diffusion, a default choice could be to take $\sigma = \mathrm{Id}$ (but retaining dimensions) to address a fairly generic situation [24]. The Wiener process $\mathrm{d}W_t$ is the canonical two-dimensional process, composed of independent one-dimensional Brownian motions in the two coordinate directions. The stochastic differential equation (2.3) should be thought of in an Itô sense, and it is noted that (2.3) could equivalently be formulated in terms of the corresponding Fokker–Planck equation [56, 67]

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho u) = \frac{\varepsilon^2}{2}\nabla\cdot\nabla\cdot\left(\rho\,\sigma\sigma^\top\right)\,, \tag{2.4}$$

which describes how a density field $\rho(y,t)$ evolves. A direct connection between (2.3) and (2.4) is that if a fixed initial condition x is assigned to (2.3), this corresponds to a Dirac-delta distribution centered at this location as an initial condition for (2.4) [32]. The stochastic spread of y_T therefore provides information on the deterministic density function $\rho(y,T)$.

Broadly speaking, the goal is to quantify the uncertainty of Lagrangian trajectories of (2.1) if stochasticity in the form of (2.3) is accounted for. As a field on the set of initial conditions, is it possible to find an uncertainty measure for the location of each Lagrangian trajectory at the final time T? While this is clearly of importance in the study of Lagrangian coherent structures, it can also enable the assignment of a *weighting* to each initial condition based on how confident one is of its eventual location, and hence it can be a useful new tool for Lagrangian data assimilation [66, 55, 70, 59, 3]. Additionally, is it possible to quantify the *anisotropic* spread that is to be anticipated? Can one identify sets ("robust sets") in which the uncertainty lengthscale is less than a stipulated threshold value? Precise ways of answering these questions will be sought in what follows.

The following convention will be used in expressing smoothness/boundedness conditions. If the norm symbol $\|.\|$ is used without additional qualification, it will mean that: (i) for a vector, it is the standard Euclidean norm ("vector norm"); (ii) for a matrix, it is the spectral norm induced by the vector norm ("matrix norm"); and (iii) for a rank-3 tensor, it is the spectral norm induced by the matrix norm ("tensor norm"). Boundedness assumptions will be related to the spatio-temporal set

$$\tilde{\Omega} := \bigcup_{t\in[0,T]} (\Omega_t, t) \ ,$$

and the gradient symbol, when used, refers only to spatial derivatives on the sets Ω_t.

Hypothesis 2.1 (smoothness/boundedness).

(a) The velocity u is globally Lipschitz in $\tilde{\Omega}$, i.e., there exists a constant $\eta > 0$ such that for any any $t \in [0,T]$ and x_1, $x_2 \in \Omega_t$,

$$|u(x_1,t) - u(x_2,t)| < \eta\,|x_1 - x_2| \ . \tag{2.5}$$

Additionally, there exists a constant K_u such that for all $(x,t) \in \tilde{\Omega}$,

$$\max\left\{\|u(x,t)\| \ , \|\nabla u(x,t)\| \ , \|\nabla\nabla u(x,t)\|\right\} \leq K_u \, . \tag{2.6}$$

(b) The flow map $F_{t_1}^{t_2} : \Omega_{t_1} \to \Omega_{t_2}$ is well-defined and invertible for any t_1 and t_2 in $[0,T]$ and, moreover, there exists a constant K_F such that for any $t \in [0,T]$ and any $w \in \Omega_T$,

$$\left\|\nabla F_T^t(w)\right\| \leq K_F \, . \tag{2.7}$$

(c) The diffusion matrix σ and its spatial derivative are uniformly bounded in the sense that there exists K_σ such that for any $(x,t) \in \tilde{\Omega}$,

$$\max\left\{\max_{i,j}|\sigma_{ij}(x,t)| \ , \|\sigma(x,t)\| \ , \|\nabla\sigma(x,t)\|\right\} \leq K_\sigma \, . \tag{2.8}$$

Let $x \in \Omega_0$ be an initial condition chosen for (2.1), which then evolves deterministically according to (2.2). In contrast, let y_t be a solution to (2.3) which satisfies the

identical initial condition $y_0 = x$ at time 0. For any $\varepsilon > 0$, this enables the definition of the random variable $z_\varepsilon(x,t)$ on $\Omega_0 \times [0,T]$ as

$$z_\varepsilon(x,t) := \frac{y_t - F_0^t(x)}{\varepsilon}, \quad \text{subject to} \quad y_0 = x\,. \tag{2.9}$$

This represents the *deviation* of the stochastic trajectory from the deterministic trajectory, at any time $t \in [0,T]$, scaled by the noise parameter. Now,

$$y_t = x + \int_0^t u\left(y_\tau, \tau\right)\,\mathrm{d}\tau + \varepsilon \int_0^t \sigma\left(y_\tau, \tau\right)\,\mathrm{d}W_\tau\,,$$

which is the weak formulation of the stochastic differential equation (2.3), subject to the initial condition $y_0 = x$. Subtracting (2.2) from this gives

$$y_t - F_0^t(x) = \int_0^t \left[u(y_\tau,\tau) - u\left(F_0^\tau(x),\tau\right)\right]\,\mathrm{d}\tau + \varepsilon \int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\,, \tag{2.10}$$

and consequently

$$z_\varepsilon(x,t) = \int_0^t \frac{u(y_\tau,\tau) - u\left(F_0^\tau(x),\tau\right)}{\varepsilon}\,\mathrm{d}\tau + \int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\,. \tag{2.11}$$

This is exact, but the determination of statistics of z_ε is intractable in this form. A *formal* approach could be to expand the y_τ-dependent functions in Taylor series around $F_0^\tau(x)$ and discard higher-order terms in ε, which would result in

$$z_\varepsilon(x,t) = \int_0^t \nabla u\left(F_0^\tau(x),\tau\right) z_\varepsilon\left(x,\tau\right)\,\mathrm{d}\tau + \int_0^t \sigma\left(F_0^\tau(x),\tau\right)\,\mathrm{d}W_\tau\,,$$

or the equivalent differential formulation

$$\frac{\partial}{\partial t} z_\varepsilon(x,t) = \nabla u\left(F_0^t(x),t\right) z_\varepsilon\left(x,t\right) + \frac{\partial}{\partial t}\int_0^t \sigma\left(F_0^\tau(x),\tau\right)\,\mathrm{d}W_\tau\,. \tag{2.12}$$

The *linearity* of this equation in $z_\varepsilon(x,t)$ might promise the ability to explicitly solve for $z_\varepsilon(x,t)$, but there are several impediments: (i) the coefficient matrix is *nonautonomous* and hence the homogeneous solution cannot be written explicitly, (ii) the stochastic nature of the inhomogeneity needs to be accounted for, and (iii) the legitimacy of discarding higher-order terms in ε—i.e., the effective understanding that $z_\varepsilon = \mathcal{O}(1)$ with respect to ε—has not been justified. The first of these impediments is a major hindrance to seeking analytical expressions for $z_\varepsilon(x,t)$, and further complications also exist because u has explicit time-dependence and is not necessarily divergence-free. Relaxing (some of) these conditions, for example, if (2.12) were a Langevin equation [67, 56, 28, 37], would enable an analytical solution. The second impediment points to the need to use genuinely stochastic methods, rather than formal variation-of-parameters manipulations. Finally, the third issue of discarding "lower-order" terms in ε is clearly unjustifiable, because Brownian motion can possess unbounded variation even over a finite time. Therefore, a more careful development is essential, and the formal expression (2.12) will be rejected. Instead, using rigorous methods, it is possible to prove the preliminary fact (useful for later results) that while z_ε is not necessarily $\mathcal{O}(1)$, the *expectation* of z_ε is, in a strong sense.

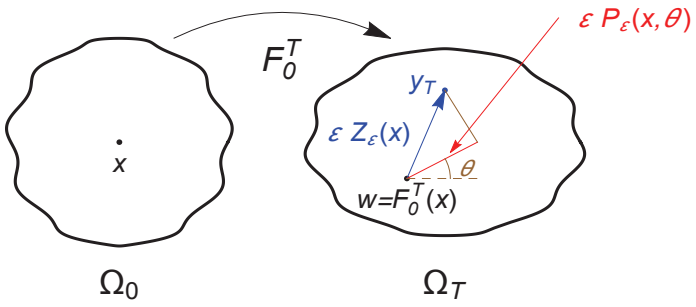


Fig. 1 *Map from times 0 to T associated with the deterministic flow (2.1) (black), and relevant entities resulting from a stochastic simulation of (2.3) (blue).*

LEMMA 2.2 (bounds on z_ε). *Let $\mathbb{E}$ be the expectation with respect to different realizations of (2.3). For z_ε as defined in (2.9) and any $q \geq 1$, there exists a constant K_z^q independent of $x \in \Omega_0$ and ε such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |z_\varepsilon(x,t)|^q\right] \leq K_z^q\,. \tag{2.13}$$

Proof. See Appendix A; this is a relatively straightforward usage of the Burkhölder–Davis–Gundy and Gronwall inequalities. ∎

The main desire is to understand the statistics of the size and orientation of $z_\varepsilon(x,t)$ *at the final time* T. To this end, define

$$Z_\varepsilon(x) := z_\varepsilon(x,T)\,. \tag{2.14}$$

Let w be the image of x under the flow of (2.1) at time T, that is,

$$w := F_0^T(x), \quad \text{and so} \quad x = F_T^0(w)\,. \tag{2.15}$$

Next, consider the signed projections of $Z_\varepsilon(x)$ onto a ray emanating from w in a direction associated with an angle $\theta \in [-\pi/2, \pi/2)$, which are defined by

$$P_\varepsilon(x,\theta) := \hat{n}(\theta)^\top Z_\varepsilon(x) \quad \text{where} \quad \hat{n}(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}; \tag{2.16}$$

see Figure 1. The angle θ is akin to that in polar coordinates centered at w, but is restricted to $[-\pi/2, \pi/2)$ because the projection is permitted to take on a sign: positive if $Z_\varepsilon(x)$'s orthogonal projection falls onto the ray associated with the angle $\theta \in [-\pi/2, \pi/2)$, or negative if it falls onto the extended ray emanating in the opposite direction. Now, the first—perhaps unsurprising—result on the statistics of $Z_\varepsilon(x)$ is as follows.

THEOREM 2.3 (expected location is deterministic). *For all $x \in \Omega_0$ and any angle $\theta \in [-\pi/2, \pi/2)$, $\lim_{\varepsilon\downarrow 0} \mathbb{E}\,[P_\varepsilon(x,\theta)] = 0$, and thus*

$$\lim_{\varepsilon\downarrow 0} \mathbb{E}\,[Z_\varepsilon(x)] = 0\,. \tag{2.17}$$

Proof. See Appendix B. ∎

Theorem 2.3 asserts that, to leading order, the mean of the random quantities y_T lies exactly at the point $F_0^T(x)$ corresponding to deterministic advection. The potential anisotropy of σ suggests that the result may not be true for nonvanishing ε (for an analogous result in the context of applying volume-preserving deterministic dynamics followed by diffusion with zero mean and covariance equal to the identity, see [35]).

Of particular interest is the behavior of the statistics of $P_\varepsilon(x,\theta)$ as $\varepsilon \downarrow 0$. Specifically, uncertainty in the eventual Lagrangian locations is related to the *variance* of $P_\varepsilon(x,\theta)$ (for which the notation $\mathbb{V}$ will be used) over many realizations of (2.3). Explicit analytical expressions for the following important quantities will therefore be sought.

Definition 2.4 (uncertainty measures).

(a) *The* anisotropic uncertainty *is a scalar field defined on* $\Omega_0 \times [-\pi/2, \pi/2)$ *given by*

$$A(x,\theta) := \sqrt{\lim_{\varepsilon \downarrow 0} \mathbb{V}\left[P_\varepsilon(x,\theta)\right]}\,. \tag{2.18}$$

(b) *The* stochastic sensitivity *is a scalar field defined on* Ω_0 *given by*

$$S^2(x) := \lim_{\varepsilon \downarrow 0} \sup_{\theta} \mathbb{V}\left[P_\varepsilon(x,\theta)\right]\,. \tag{2.19}$$

Clearly, the quantity $\varepsilon A(x,\theta)$ is a leading-order estimate for the distance uncertainty of y_T along the ray of angle θ; by computing this for all θ, the one-standard-deviation level of uncertainty in all possible directions can be determined. The anisotropy (i.e., dependence of the uncertainty on θ) is a consequence of *both* the deterministic flow u and the stochastic model σ. Even if $\sigma = \mathrm{Id}$, there will generically be an anisotropy (as will be demonstrated in section 3).

The stochastic sensitivity $S^2(x)$ uses a variance measurement for the leading-order spread, optimized across all directions θ (much like the optimization over all directions used in computing, for example, finite-time Lyapunov exponents). Since S^2 provides a measure of the eventual Lagrangian position's uncertainty, it quantifies the *sensitivity toward stochasticity* occurring due to the accumulated impact of the uncertainty in the Eulerian velocity field. Operationally, $\varepsilon\sqrt{S^2(x)}$ gives the leading-order estimate for the uncertainty of the deviation of y_T from the location $F_0^T(x)$, at each initial location x. The notation S^2 is used to capture both the alliteration in the terminology (Stochastic Sensitivity) and the fact that it is meant to estimate a standard deviation squared.

In establishing computable formulae for both A and S^2, first define

$$J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{2.20}$$

which operates on vectors in $\mathbb{R}^2$ as a $+\pi/2$-rotation. Furthermore, for $w = F_0^T(x) \in \Omega_T$ and $t \in [0,T]$, define the matrix

$$\Lambda(w,t) := e^{\int_t^T [\nabla \cdot u]\left(F_T^\xi(w),\xi\right) \mathrm{d}\xi}\, \sigma\left(F_T^t(w),t\right)^\top J\, \nabla F_T^t(w)\,. \tag{2.21}$$

Note that the gradient operator in (2.21) is with respect to positions w on the spatial domain at the *final* time T. If $\sigma = \mathrm{Id}$ (a potentially default choice) and if u were

divergence-free (as would occur in an incompressible flow), Λ takes on a particularly simple form:

$$\Lambda(w,t) = J\,\nabla F_T^t(w)\,. \tag{2.22}$$

However, generality will be maintained in what follows. Given an Eulerian velocity field u (either explicitly or as data on a grid), all components of Λ can be numerically computed by (if necessary) using interpolation when computing the evolving flow map F_T^t of the deterministic system (2.1). Note that the calculation will be done in *backwards* time from the location $w \in \Omega_T$.

THEOREM 2.5 (anisotropic uncertainty). *For $w \in \Omega_T$ and $\theta \in [-\pi/2, \pi/2)$, define*

$$\tilde{A}(w,\theta) := \left(\int_0^T \left| \Lambda(w,t)\, J\, \hat{n}(\theta) \right|^2 \mathrm{d}t \right)^{1/2}. \tag{2.23}$$

The field $A(\centerdot,\theta)$ on $\Omega_0 \times [-\pi/2, \pi/2)$ can then be obtained by ascribing the value $\tilde{A}(w,\theta)$ to the point $x = F_T^0(w)$, i.e., $A(x,\theta) = \tilde{A}\left(F_0^T(x),\theta\right)$.

Proof. See Appendix C. ∎

For a chosen point $x \in \Omega_0$, computing $A(x,\theta)$ for each θ allows the determination of a leading-order (in ε) one-standard-deviation curve around the point $w = F_0^T(x)$, which elucidates the anisotropy of the spread around w. The stochastic sensitivity (2.19), on the other hand, is the maximum variance across all directions. Let $\Lambda_{ij}(w,t)$ be the components of Λ in (2.21), and further define

$$\begin{aligned}
L(w) &:= \frac{1}{2}\int_0^T \left[\sum_{i=1}^{2} \Lambda_{i2}^2(w,t) - \sum_{i=1}^{2} \Lambda_{i1}^2(w,t) \right] \mathrm{d}t\,, \\
M(w) &:= \int_0^T \sum_{i=1}^{2} \left[\Lambda_{i1}(w,t)\, \Lambda_{i2}(w,t) \right] \mathrm{d}t\,, \\
N(w) &:= \sqrt{L^2(w) + M^2(w)}, \quad \text{and} \\
P(w) &:= \left| \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \int_0^T \Lambda_{ij}^2(w,t)\, \mathrm{d}t \right|.
\end{aligned}$$

THEOREM 2.6 (maximal spreading direction). *Given any $x \in \Omega_0$, the maximal spreading direction from $w = F_0^T(x)$ in Ω_T is given by*

$$\theta_{\max} := \operatorname*{argmax}_{\theta} A\,(x,\theta) = -\frac{\alpha}{2}\,, \tag{2.24}$$

where $\alpha \in [-\pi, \pi)$ is obtained by solving the pair of equations

$$\cos\alpha = \frac{L(w)}{N(w)} \quad \textit{and} \quad \sin\alpha = \frac{M(w)}{N(w)}\,.$$

Proof. See Appendix D; this theorem is proved in conjunction with the next. ∎

In numerically determining the angle $\theta_{\max}$ using (2.24), the four-quadrant inverse tangent (which is built into most computational systems) can be employed to find

$\alpha \in [-\pi, \pi)$ using the expressions for $\cos\alpha$ and $\sin\alpha$. The subsequent value for $\theta_{\max}$ in (2.24) yields a value in $[-\pi/2, \pi/2)$ as desired. Next, noting from Definition 2.4 that

$$S^2(x) = A^2\left(x, \theta_{\max}\right) \tag{2.25}$$

(Lemma 2.2 along with estimates given in Appendix D provide the conditions for using the dominated convergence theorem to move the $\varepsilon \downarrow 0$ limit with impunity), the stochastic sensitivity is easily computed.

THEOREM 2.7 (stochastic sensitivity). *Define*

$$\tilde{S}^2(w) := P(w) + N(w) \tag{2.26}$$

for $w \in \Omega_T$. Then $S^2(\cdot)$ on Ω_0 can be obtained by ascribing the value $\tilde{S}^2(w)$ to each point $x = F_T^0(w)$, i.e., $S^2(x) = \tilde{S}^2\left(F_0^T(x)\right)$.

Proof. See Appendix D. ∎

In the spirit of Lagrangian coherent structure analysis [47, 62, 68, 5, 64, 12], the importance of Theorem 2.7 is that it provides a *field* on the initial space Ω_0, which will help demarcate distinct regions based on the value of the field. Given an unsteady Eulerian velocity field u (taken from experimental, observational, or computational fluid dynamics data or analytically) and any given any model σ for the stochastic perturbation, (2.26) is readily computable. It should be noted that an additional generality in the development is that there is no restriction on the velocity field being compressible; $\nabla \cdot u$ can be computed numerically from the velocity data, and can be directly utilized in the S^2-formula.

The S^2 field on Ω_0 contains intrinsic information on the potential for uncertainty in Lagrangian trajectories, and is a *theoretical* field; that is, it is legitimate at the stochastic differential equations level—assuming a smooth u defined entirely on $\tilde{\Omega}$—without reference to any particular parameter values on spatial resolution or diffusivity. Next, two nonlinear scalings of the field which can interrogate these two physical considerations—and thereby be useful in providing *quantifications* in numerical computations based on experimental/observational data—are now suggested. Both scalings—and their advantages in the data-driven context—will be examined in greater depth in section 3. The first scaling compares the uncertainty lengthscale $\sqrt{S^2}$ with the spatial resolution lengthscale of the available data and then applies a logarithmic scaling to ensure that variations at lower values of S^2 are not brushed over by enormous values of S^2 elsewhere.

DEFINITION 2.8 (resolution-scaled stochastic sensitivity). *Given the spatial resolution lengthscale L_r of the Eulerian velocity field, the* resolution-scaled stochastic sensitivity *is defined on Ω_0 by*

$$S_r(x) := \ln \frac{\sqrt{S^2(x)}}{L_r}\,. \tag{2.27}$$

Given the value of L_r, the nondimensional field S_r—just like S^2—can be computed using the available velocity data and nothing else. Either field can be used in a data-driven situation, bearing in mind that they are simply nonlinearly scaled versions of one another. Using one rather than the other can help in highlighting the structure of the field near certain regions. In particular, uncertainty lengthscales below the resolution lengthscale—which are meaningless in experimental/observational

situations because resolving them is futile—will be exemplified as negative values of S_r.

The second scaling that is suggested is based on additional knowledge: an estimate for the noise level in the data. Small values of S^2 imply that trajectories beginning at corresponding locations have high predictability. In this sense, they are coherent. Conversely, large stochastic sensitivity is associated with *incoherence*. One way of extracting *robust sets* in this sense from S^2 is the following. Given any flow feature in Ω_0, there is a diffusive timescale (associated with the diffusive flow) and an advective timescale (associated with the deterministic flow u) which capture the timescales over which diffusion and advection impact the feature. The Péclet number Pe is the ratio of the diffusive to the advective timescales. From (2.4), it is clear that $\mathrm{Pe} = 2/\varepsilon^2$. Now, since S^2 captures the variance of $\left[y_T - F_0^T(x)\right]/\varepsilon$, a measure of how far y_T is from $F_0^T(x)$ is quantified by the *noise-scaled stochastic sensitivity*

$$S_n(x) := \varepsilon\sqrt{S^2(x)} = \sqrt{2S^2(x)/\mathrm{Pe}}\,. \tag{2.28}$$

The quantity S_n is a physical (dimensional) lengthscale. If this lengthscale is less than an acceptable uncertainty lengthscale of predictability L, the corresponding initial condition $x \in \Omega_0$ can be considered *robust*.

Definition 2.9 (robust sets). *Given a Péclet number* $\mathrm{Pe} = 2/\varepsilon^2$ *and a lengthscale* L*, then the set*

$$R\,(L, \mathrm{Pe}) := \left\{ x \in \Omega_0 \;:\; S^2(x) < \frac{L^2\,\mathrm{Pe}}{2} \right\} \tag{2.29}$$

shall be called robust at the lengthscale L and Péclet number Pe *with respect to a noise model (i.e.,* σ*) operating over the time* $[0, T]$*. Maximal connected subsets* $R_i(L, \mathrm{Pe})$ *of* $R(L, \mathrm{Pe})$ *will be called* robust connected sets *(at this lengthscale and level of diffusion).*

Thus, the stochastic sensitivity field provides a way of extracting sets from Ω_0 (the time-0 spatial locations, i.e., the allowable initial conditions) which are robust as defined in the following sense: the expected uncertainty in Lagrangian predictions by the time T will be less than L. Pertinent to applications is the choice of the scales L (relevant to identifying features one cares about, or the resolution of the data) and Pe (associated with anticipated uncertainty, e.g., estimates of oceanic eddy diffusivity [52, 18, 72, 76, 24] or measurement error). Assigning robustness *with respect to* the uncertainty lengthscale L and diffusion scale Pe is a significant new idea, taking into account precisely the factors which influence the *level* of robustness one seeks. Extracting connected robust sets using Definition 2.9—being mindful of the precise interaction model chosen to relate L to the available spatial resolution and Pe to available diffusive estimates—might be considered one way of quantifying *coherence* of Lagrangian predictions.

The above two scalings will help to highlight different physical considerations, as will be illustrated in the numerical simulations in section 3. Significant applications of these results are anticipated, and are discussed in more detail in section 4.

3. Computability of Uncertainty Measures. This section will illustrate the computability of the theoretical expressions in a situation in which the Eulerian velocity is assumed known in the sense of the data possessing a certain spatial and temporal resolution. While much future work is anticipated, for this article (whose purpose is to establish the theoretical framework for stochastic sensitivity measures),

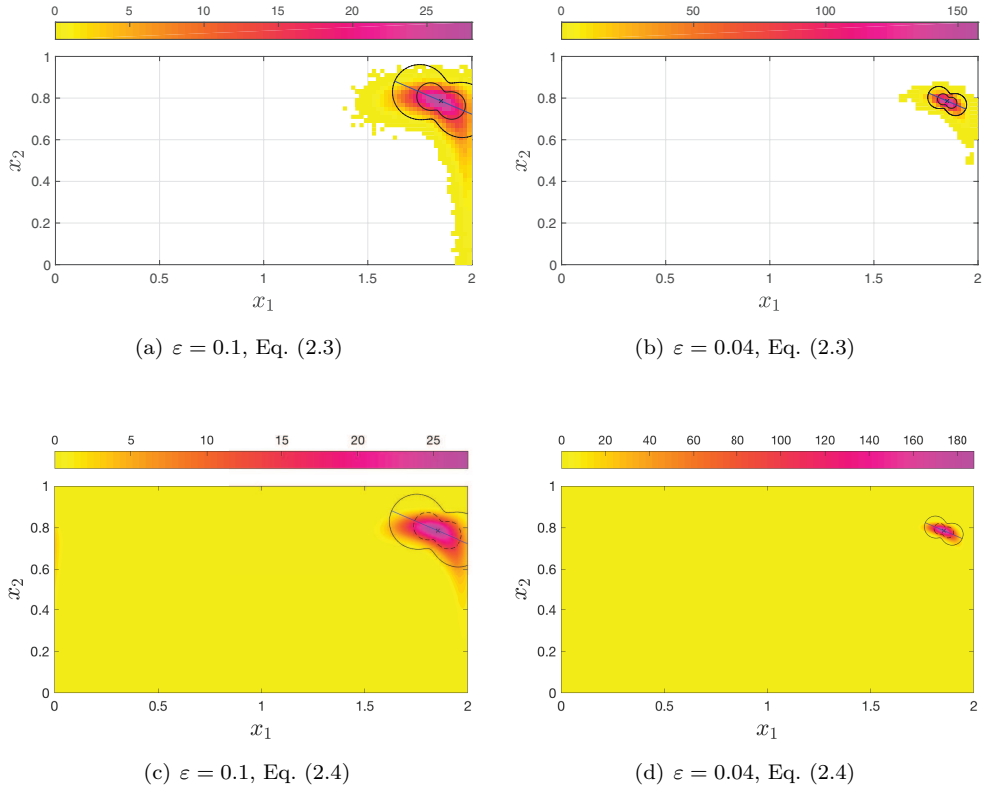
the simplest situations of an incompressible flow and $\sigma = \mathrm{Id}$ will be used. Velocity data will be taken from a model flow which has been extensively used (see the citations in [63, 31]) as a testbed for Lagrangian coherent structure analysis: the double-gyre model introduced by Shadden, Lekien, and Marsden [69]. This choice has the added advantage of having well-understood properties, and can therefore be compared to the theoretical values for A, $\theta_{\max}$, S^2, S_r, and R.

The velocity field u for the double-gyre flow for $x = (x_1, x_2) \in \Omega_0 := [0, 2] \times [0, 1]$ is given by [69]

$$u(x,t) = \begin{pmatrix} -\pi A_{dg} \sin\left[\pi\phi(x_1,t)\right] \cos\left[\pi x_2\right] \\ \pi A_{dg} \cos\left[\pi\phi(x_1,t)\right] \sin\left[\pi x_2\right] \frac{\partial\phi}{\partial x_1}(x_1,t) \end{pmatrix}, \tag{3.1}$$

in which $\phi(x_1, t) := \varepsilon_{dg} \sin(\omega t)\, x_1^2 + (1 - 2\varepsilon_{dg} \sin(\omega t))\, x_1$. Here, $\Omega_t = F_0^t(\Omega_0) = \Omega_0$ for any t. This possesses two counterrotating gyres when $\varepsilon_{dg} = 0$: one in $(0, 1) \times (0, 1)$ and the other in $(1, 2) \times (0, 1)$. When $\varepsilon_{dg} \neq 0$ but is small, chaotic transport occurs between the gyres due to the splitting of the heteroclinic manifold along $x_1 = 1$ into stable and unstable manifolds that intersect infinitely often [63]. The splitting of Ω_0 into separate regions (almost coherent within the gyres, with chaotic mixing in the flange areas related to manifold intersections) has led to the double-gyre being a well-established paradigm in methods for the identification of coherent regions (see [63] for an extensive list of such uses). For the first time, an explicit assessment of the impact *of noise* on the variation from deterministic trajectories can now be performed, without the restriction to any particular definition of "coherence" in the multitude of methods which seek Lagrangian coherent structures. Instead, the intuitively straightforward idea of uncertainty in the eventual Lagrangian location is used. For all calculations, the parameter choices are $A_{dg} = 1$, $\omega = 2\pi$, and—unless specified otherwise—$\varepsilon_{dg} = 0.05$. Velocity data will be taken from the expression in (3.1), but subject to a spatial grid on Ω_0 and a temporal grid on $[0, T]$, to mimic how data would be available in a realistic situation.

First, the anisotropic uncertainty expression (2.23) is investigated in relation to the initial condition $x = (1.4, 0.1)$ subjected to the flow until time $T = 0.4$. Stochastic simulations of (2.3), using the Euler–Maruyama scheme [53] with $\Delta t = 0.0001$, are performed 10,000 times to obtain a cluster of final conditions. Those conditions remaining in Ω_0 are binned into square bins of side 0.02, and the probability density function is computed (simulations exiting Ω_0 are excised when performing the normalization). This is shown for two values of ε in Figure 2(a,b). Superimposed on these Monte Carlo simulation results is a computation of the final deterministic location $w = F_0^T(x)$ (black "x"), the one-standard-deviation region a distance $\varepsilon A(x, \theta)$ computed in each direction θ (dashed black curve), and the two-standard-deviation region $2\varepsilon A(x, \theta)$ (solid black curve). In computing $A(x, \theta)$ using (2.23), it is necessary to deterministically advect—using (2.1) and the standard third-order Runge–Kutta scheme—a "star-grid" of conditions located at w and its "north," "south," "east," and "west" points a distance 0.001 away, in order to numerically compute the flow map gradient as required in (2.23). Moreover, the blue line in Figure 2 is the (theoretical) maximal deviation direction computed using (2.24). The anisotropic uncertainty neighborhoods indicated by the black curves display a "keyhole" structure, elongated in the direction of maximal spreading and pinched inwards in the perpendicular direction; this appears to be generic in situations in which there is large deterministic stretching. The stochastic simulations—whose distribution is shown by the color background—certainly reflect the theoretical predictions on the anisotropic spread of the uncertainty, stretching out in the direction of the (maximal) blue line, with the



(a) $\varepsilon = 0.1$, Eq. (2.3) (b) $\varepsilon = 0.04$, Eq. (2.3)

(c) $\varepsilon = 0.1$, Eq. (2.4) (d) $\varepsilon = 0.04$, Eq. (2.4)

Fig. 2 (a,b) *Probability density functions arising from stochastic simulations of* (2.3) *with initial condition* $x = (1.4, 0.1)$ *to time* $T = 0.4$*, at two different* ε *values.* (c,d) *Density evolution of the Fokker–Planck equation* (2.4) *by time* T*, with initial condition a Dirac mass at* x*.*

one- and two-standard-deviation regions apparently confining the probability density as expected. The agreement is stronger for smaller ε, as expected from the theory.

In Figures 2(c,d), a comparison is made with analogous solutions to the Fokker–Planck equation (2.4). In this case, a Dirac mass initial condition (approximated by a Gaussian with standard deviation 0.001) is positioned at x, and the Fokker–Planck equation is solved numerically to determine the density distribution at time T. Periodic boundary conditions are imposed, thereby allowing for the use of a pseudo-spectral code: fast Fourier transforms in space, and a Crank–Nicolson scheme with $\Delta t = 0.0001$ for time evolution. The spatial discretization was adjusted until a convergent result was obtained. The smaller ε value (0.04) required a high-resolution grid (side length 0.00005) for convergence. The results from the Fokker–Planck simulations (Figures 2(c,d)) are consistent with the corresponding stochastic differential equation statistics (Figures 2(a,b)), in the sense of both scaling of the probability distribution and how it is spread. Moreover, *both* are consistent with the anisotropy indicated theoretically via $A(x, \theta)$ and $\theta_{\max}$.

To examine this more quantitatively, Figure 3(a,b) uses stochastic simulations as in Figures 2(a,b) to compute the projection $P_\varepsilon(x, \theta)$ onto the ray $\theta = \theta_{\max}$ for two values of ε. The probability density distributions of $P_{0.2}$ and $P_{0.04}$ are shown in Figures 3(a,b). While asymmetric for larger ε, symmetry is approached as $\varepsilon \downarrow 0$.

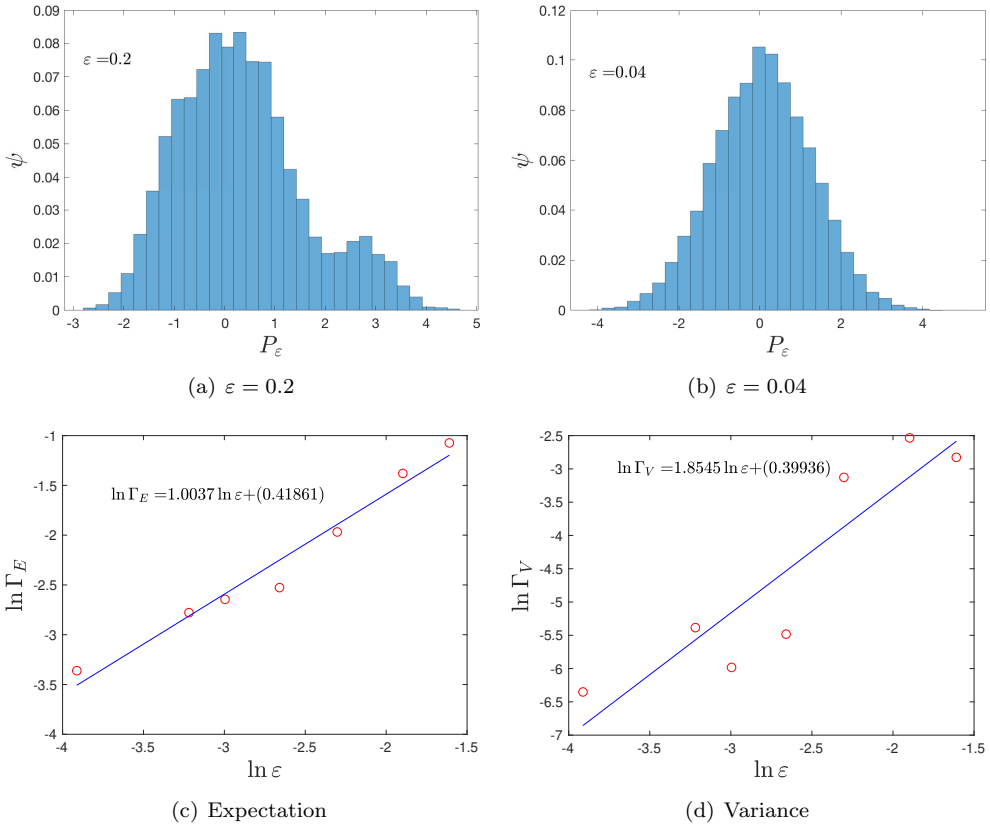


Fig. 3 (a,b) *Probability density* ψ *of* P_ε *in the direction* $\theta_{\max}$ *(as computed from* (2.24)*) from the stochastic simulation data of Figure* 2 *for two values of* ε. (c,d) *Errors* Γ_E *and* Γ_V *between the theoretical and computed standard deviations as* ε *varies.*

However, *normality* is not claimed for the limiting distribution. By Theorems 2.3 and 2.5, the expectation and standard deviation computed in this way must decay to 0 and $A(x, \theta_{\max})$, respectively. Thus, the error measures

$$\Gamma_E(\varepsilon) := \left|\mathbb{E}\left[P_\varepsilon(x, \theta_{\max})\right] - 0\right| \quad \text{and} \quad \Gamma_V(\varepsilon) := \left|\sqrt{\mathbb{V}\left[P_\varepsilon(x, \theta_{\max})\right]} - A(x, \theta_{\max})\right|$$

for $x = (1.4, 0.1)$ can be used. These measures are shown in Figures 3(c,d) by the red circles, based on stochastically simulating (2.3) 10,000 times for different values of ε. The lines of best fit for each log-log plot indicate that $\Gamma_E \sim \varepsilon$ and $\Gamma_V \sim \varepsilon^{1.85}$, validating Theorems 2.3 and 2.5 (which claim that these errors go to zero with ε).

Figure 4 illustrates the same types of plots as in Figures 2(a,b), but with six different initial conditions chosen simultaneously. Three of these remain relatively robust (the scatter is limited) *and* have fairly isotropic neighborhoods. These are initial conditions which are known to be within the regular gyres of the double-gyre. Even when T increases, the scatter of points (from the Monte Carlo simulations) and the anisotropic neighborhoods (from the theory) are fairly well confined. In contrast, the other three initial conditions display a distinct anisotropy, with keyhole neighborhoods emerging by $T = 0.4$. The scatter of the stochastic simulations shows

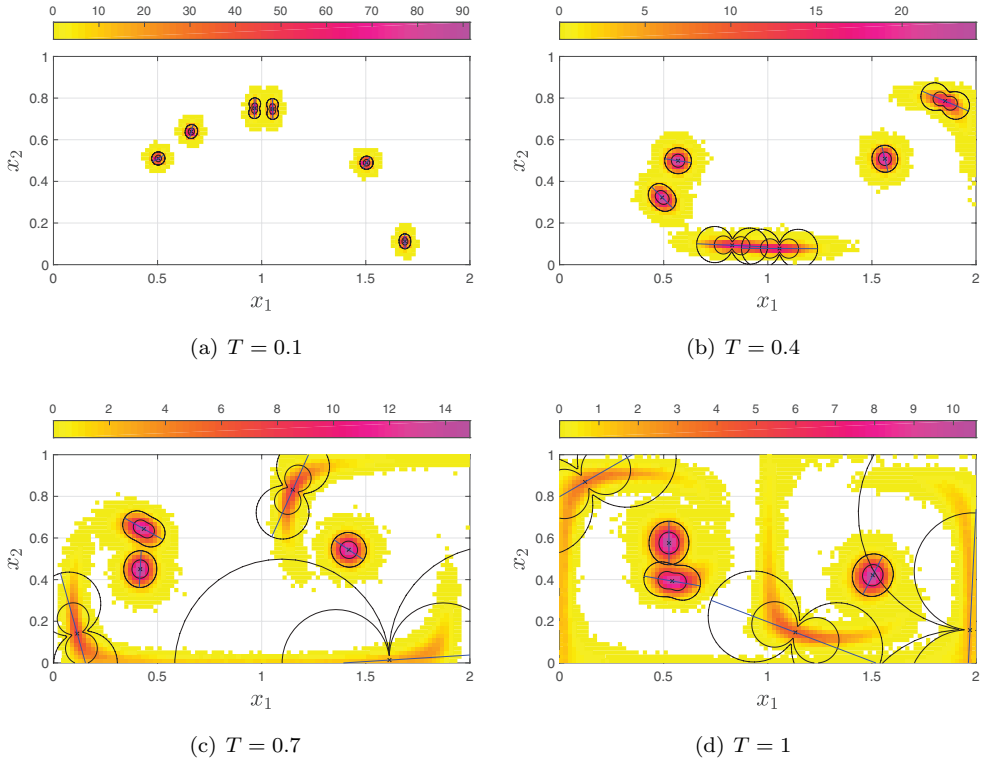
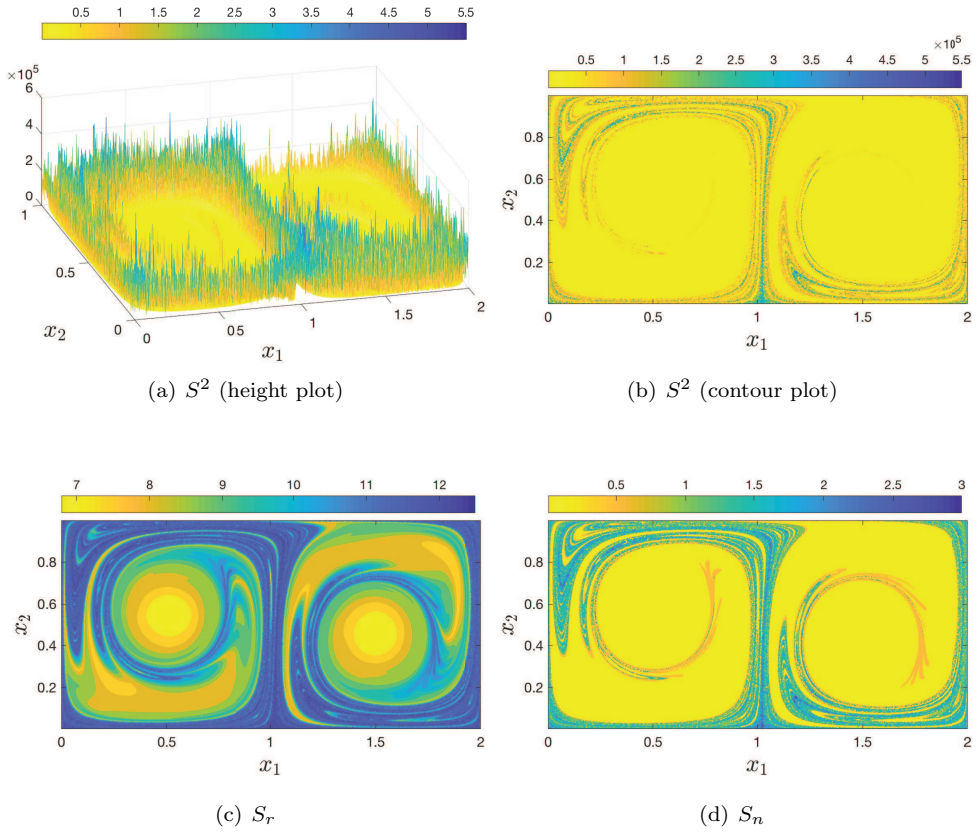


Fig. 4 *Probability density distributions arising from simultaneous stochastic simulations with* $\varepsilon = 0.05$ *for six different initial conditions, along with the theoretical anisotropic uncertainty neighborhoods (see description in Figure* 2*), as the time* T *evolves.*

that they do indeed get stretched out in the direction of $\theta_{\max}$ (blue line). This gives additional credence to the fact that using S^2, which captures exactly the stretching in this direction, is highly effective in quantifying the dominant uncertainty; there is much less uncertainty in other directions because of the "keyhole" structure of the neighborhoods. The observed *curvature* in the dominant spreading—which the current (locally linear) theory is not able to capture—reduces as ε gets smaller. By $T = 1$, three of the anisotropic uncertainty neighborhoods are so large that they have gone well outside Ω_0, and indeed the stochastic simulations appear to have scattered well outside. These are associated with regions which are within the *chaotic region* of the double-gyre, and thus this behavior is to be expected. If the quantity $\varepsilon A(x, \theta)$ is large enough to extend well outside Ω_T, this provides a caution that *deterministic conclusions associated with that initial condition are questionable.* There is currently no such sanity check in standard deterministic Lagrangian coherent structure analysis illustrating, primarily, a "leading-order" usefulness of the current theory.

In contrast with *selected* initial conditions, the stochastic sensitivity S^2 seeks to provide an uncertainty measure *as a field across* Ω_0, thereby helping to identify regions at time 0 which will be associated with robust (small S^2) behavior. To compute this, particles are numerically seeded on $[0, 2] \times [0, 1]$ in a 800×400 uniform grid, corresponding to a resolution of $L_r = 0.0025$ in each direction. Each particle here corresponds to choosing a $w \in \Omega_T$. The grid at time 0 is backward advected using



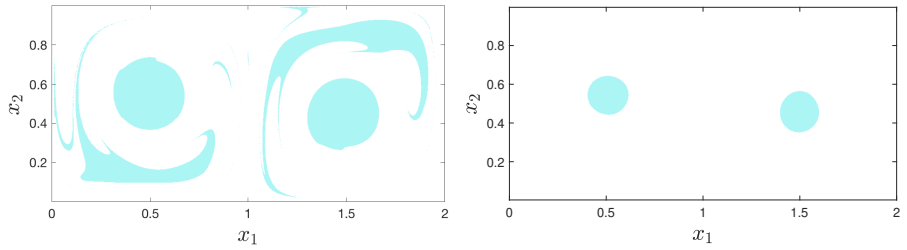
(a) S^2 (height plot) (b) S^2 (contour plot)

(c) S_r (d) S_n

Fig. 5 *The S^2 field with illustrations of its two scalings for the double-gyre flow in the time interval $[0, 5]$: (a) S^2 field as a graph (height plot), (b) S^2 field as a contour plot, (c) resolution-scaled stochastic sensitivity field S_r (2.27) with lengthscale $L_r = 0.0025$, and (d) noise-scaled stochastic sensitivity field S_n with Péclet number* $\mathrm{Pe} = 100{,}000$.

(2.1) and the standard third-order Runge–Kutta scheme until time 0, using a time step $\Delta t = 0.01$. At each general intermediate time $t \in [0, T]$, $\Lambda(w, t)$ in (2.21) is in this case simply $J\nabla F_T^t(w)$ as given in (2.22) because (3.1) is area-preserving and $\sigma = \mathrm{Id}$. This data enables the estimation of $\tilde{S}^2$ on Ω_T using (2.26). Since $S^2(x) = \tilde{S}^2\left(F_T^0(w)\right)$, the computed $\tilde{S}^2$ enables the determination of S^2 on a set of values $x \in \Omega_0$ which are *non-uniform* because they are generated by $x = F_T^0(w)$ rather than being specified. MATLAB's `gridfit` algorithm is then used to extend S^2 uniformly to Ω_0.

Figure 5(a) displays the S^2 field computed in this way for the flow from times 0 to $T = 5$. There are very large values along the peripheries and near $x_1 = 1$, rising sharply from regions of very low values. Figure 5(b) shows the S^2 field as a contour plot, which is a more effective method of illustrating the structures on Ω_0. The observed "wrapping back and forth" of the extremely large value ridges is associated with the well-known behavior of the stable manifold. In contrast, there are large yellow regions in Figures 5(a,b), indicating regions of very low S^2 values. Initial conditions which are chosen in these regions are therefore not highly influenced by velocity uncertainties. These regions are related to the two gyres of the double-gyre. The S^2 field is the *intrinsic* Lagrangian uncertainty field, the computation of which



(a) $R(0.05, 100000)$ (b) $R(0.02, 100000)$

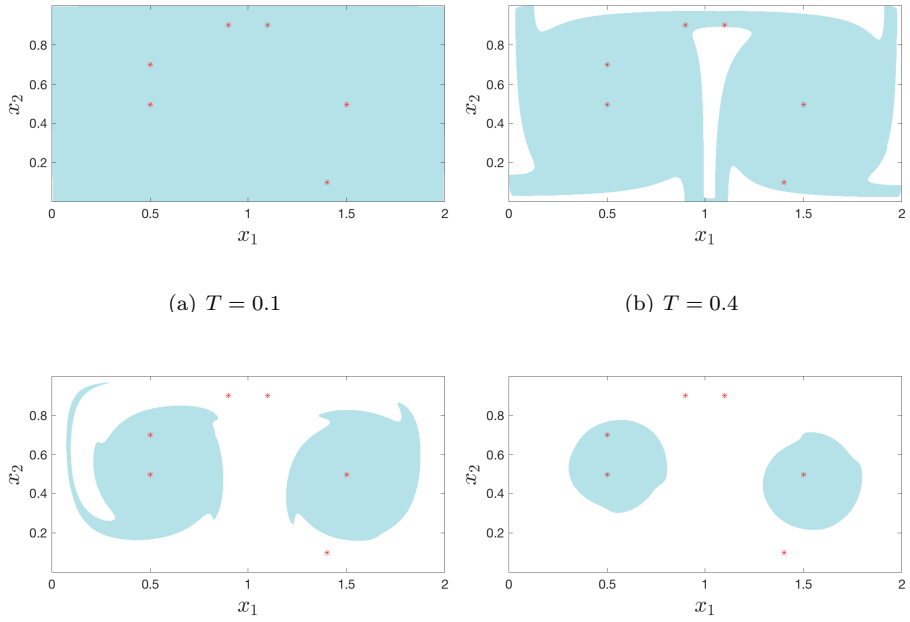
Fig. 6 *Robust sets (shaded) extracted from the stochastic sensitivity field in Figure* 5*, with a noise scale given by the Péclet number* 100,000*:* (a) *robust set with lengthscale* 0.05*, and* (b) *robust set with lengthscale* 0.02*.*

does not require knowledge of the scale ε of the uncertainty in the Eulerian velocity. As such, it is computable for *any* given velocity data u, and is hence suggested as the "go-to" field to compute for any system. It provides a measure of the uncertainty of Lagrangian prediction at each initial condition, *relative* to the rest.

Another relative measure is given by the resolution-scaled stochastic sensitivity field (2.27), in which $\sqrt{S^2}$ is scaled by the resolution lengthscale $L_r = 0.0025$ used in these calculations, and a logarithm is applied. This field, too, can be computed directly from the available velocity data (as long as the spatial resolution is uniform at L_r), without having any additional information. Shown in Figure 5(c), this field elucidates more structure, because moderate values are not smeared out by the presence of extremely large values. Indeed, regions associated with the stable manifold are shown to have an uncertainty level of an order of magnitude higher than the resolution length on a logarithmic scale (blue regions). Moreover, highly robust regions at the center of the gyres are more clearly identified (light yellow blobs). There are no negative values in this instance, because the uncertainties are all well above the resolution lengthscale. Thus, uncertainties are observable (are not at subgrid scales), and need to be taken into account.

The noise-scaled stochastic sensitivity field S_n in (2.28) is an *absolute* uncertainty measure which gives a physical lengthscale of uncertainty. Computing S_n requires an estimate for the size of the diffusion (Eulerian velocity uncertainty) in terms of ε or Pe. With the choice Pe = 100,000, S_n is shown in Figure 5(d). This nonlinear scaling of S^2 offers the physical interpretation that, under this diffusive level, there is an uncertainty lengthscale of over 2 in the chaotic regions. This is larger than the width of Ω_0! Indeed, many stochastic trajectories starting at these high S^2 locations are likely to exit the domain Ω_0 completely, despite the fact that all deterministic trajectories remain within Ω_0. So, if the Péclet number is of this size, computing deterministic trajectories with initial conditions starting in these regions makes no sense whatsoever; inevitable uncertainties mean that the calculated trajectories are spurious. The scaled S^2 fields therefore enable fundamental decisions of this nature.

In Figure 6(a), the robust set $R(0.05, 100000)$ (shaded) is extracted from the S_n field. (These are regions in Figure 5(d) with uncertainty lengthscale $S_n < 0.05$.) In addition to well-defined blobs in the cores of the gyres, there are other regions which are robust at this level. On the other hand, if robustness of lengthscales 0.02 is sought



(a) $T = 0.1$ (b) $T = 0.4$

(c) $T = 0.7$ (d) $T = 1$

Fig. 7 *The robust sets $R(0.1, 800)$ in Ω_0 (shaded) corresponding to flow in $[0, T]$, computed as direct comparison with Figure 4, with the six initial conditions shown by red stars.*

(Figure 6(d)), then this occurs only in the gyre core regions. The robust sets (and their maximal connected subsets) can therefore be identified explicitly in terms of specified physical characteristics (diffusion level and lengthscale) using Definition 2.9, which is easily computed having found the field S^2.

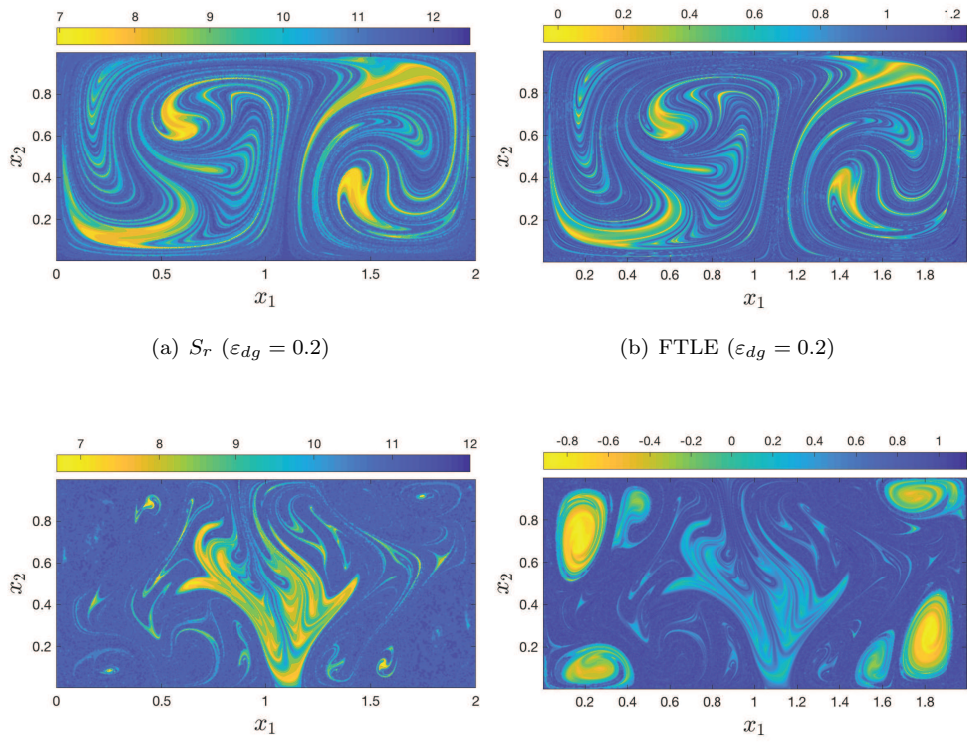
The different behaviors of the six initial conditions of Figure 4 are now examined in terms of robust sets. A lengthscale $L = 0.1$ (representing an estimate threshold size of one-standard-deviation level of the stochastic blobs in Figure 4 before they become too broken apart) is chosen, and here $\varepsilon = 0.05$ (thus, Pe $= 800$). Based on this information S^2 can be computed on each of the four time intervals $[0, T]$ in Figure 4, and the robust set $R(0.1, 800)$ (as in Definition 2.9) at time 0 can be identified for each instance. This is shown by the shaded regions in Figure 7. Note that different sets are identified for the four situations because the time interval over which the stochastic sensitivity is computed is different. As is reasonable, the uncertainty increases if a longer time interval is considered, and as a result the robust sets shrink in size as larger values for the final time T are chosen. All subfigures here display sets *at the initial time* $t = 0$, thereby identifying flow regions at time 0 based on expected uncertainties for Lagrangian trajectories computed over different time durations. The six initial conditions used for generating Figure 4 are indicated by red stars.

Now, Figure 4 showed via stochastic simulations that three of the initial conditions exhibited increasingly spreading (and anisotropic) behavior as T progressed, whereas the other three remained more focused and isotropic. This behavior—obtained by stochastic simulations—can be inferred easily using the location of the initial condi-

tions in relation to the robust sets in Figure 7 (which used exactly the same times of duration associated with Figure 4). Given that all six initial conditions in Figure 4(a) are firmly within the robust set of lengthscale 0.1, the expectation is that the uncertainties of Lagrangian trajectories, if initiated at each such initial condition at time 0 and evolved until time $T = 0.1$, will all be within a size of 0.1. This is confirmed by Figure 4(a). Figure 7(b) shows that if the flow is considered for the time interval $[0, 0.4]$ instead, then three of the initial conditions are well within the robust set, but the other three–while being within the set—are on its outskirts. This would suggest that stochastic simulations over this time duration will exhibit some evidence of spreading beyond a size of 0.1, which is mildly indicated in three of the probability distributions in Figure 4(b). If the flow is considered for the longer time interval $[0, 0.7]$, then Figure 7(c) now has three initial conditions which are *outside* the robust set, indicating that the probability density of stochastic simulations in this case is expected to spread out wider than the specified lengthscale $L = 0.1$, and this is indeed displayed in Figure 4(c). Finally, Figure 7(d) shows that for the flow until time $T = 1$, the same three initial conditions offer a certainty of less than $L = 0.1$, whereas the other three initial conditions give a larger uncertainty—features which are readily seen in Figure 4(d), which clearly shows that three of the densities continue to remain focused, whereas the other three have spread out well beyond the stipulated lengthscale of $L = 0.1$.

Robust sets offer predictions on the distribution of stochastic Lagrangian trajectories, as evidenced by the comparison between Figures 4 and 7. In this case, since the double-gyre flow is well known, the reasons for this particular behavior can be deduced: the two clearly distinguished robust sets in Figure 7(d) are the "cores of the two gyres" of the double-gyre model, which are well known to move "coherently" according to many different definitions used in analyzing Lagrangian coherent structures [69, 45, 12, 47], whereas the region around $x_1 = 1$ is a chaotic, incoherent region. However, rather than trying to demarcate these characterizations—or define precisely what they mean—using various methods popular in Lagrangian coherent structures, they are identifiable by stipulating *robustness in relation to stochastic perturbations in the Eulerian velocity field* using Definition 2.9. It must be emphasized that this robustness is specified in relation to the time-of-flow $[0, T]$, the diffusive level (encoded in $\mathrm{Pe} = 2/\varepsilon^2$), and the lengthscale L of features to be considered. Standard Lagrangian coherent structure methods *also* often have thresholds or coherence measures in their computations—sometimes explicit (e.g., the minimum threshold if thresholding finite-time Lyapunov exponents to identify chaotic regions [69, 12], the decision on the number of spectral elements to choose and/or the spectral gap [35, 36], or the clustering criterion [39, 65]), but sometimes implicit in the numerical algorithm used (e.g., the distance a trajectory is allowed to venture into a nonallowable set before computations are stopped [30], or the threshold condition for eigenvectors in identifying almost coherent structure when using transfer operators [34]). In this case, the threshold numbers relate explicitly to noise level and uncertainty lengthscale.

Finally, the stochastic sensitivity field will be compared to a well-established coherence measure—the finite-time Lyapunov exponent (FTLE)—computed by forward advection from time 0 to $T = 5$, using the identical grid spacing and time-stepping as used for the stochastic sensitivity calculations. The FTLE specifically measures exponential separation of nearby trajectories following the *deterministic* flow, and it is computed here using standard methods [69, 47]. As for S^2, the FTLE at each point is defined by maximizing a quantity (in this case, exponential separation under the



(a) S_r ($\varepsilon_{dg} = 0.2$) (b) FTLE ($\varepsilon_{dg} = 0.2$)

(c) S_r ($\varepsilon_{dg} = 5$) (d) FTLE ($\varepsilon_{dg} = 5$)

Fig. 8 *The resolution-scaled stochastic sensitivity S_r (left) and forward FTLE field (right) at time* 0 *for the double-gyre flow associated with the flow in the time interval* $[0, 5]$ *with resolution* $L_r = 0.0025$*, at* $\varepsilon_{dg} = 0.2$ *(top) and* 5 *(bottom).*

deterministic flow) over all directions. A difference, though, is that the directions for the FTLE are at the initial time (asking the question "In which direction should initial conditions be perturbed to elicit the greatest eventual stretching?") as opposed to those for S^2 which are at the final time ("In which direction will there be maximal spreading at the final time?"). For Figure 8 the default ε_{dg} value of 0.05 is changed to investigate regimes which are *not* in the typical "separation between gyres and chaotic zones" regime. The S_r field is used (with spatial resolution $L_r = 0.0025$ as in the data) as the version of the scaled S^2 field which most closely resembles the scaling of the FTLE. For $\varepsilon_{dg} = 0.2$, the stochastic sensitivity and the FTLE fields display strong similarities. However, the $\varepsilon_{dg} = 5$ pictures have a difference: the FTLE figure possesses four elliptic regions in the four corners with *negative* stretching (i.e., compression) which are *not* indicated in the S_r figure. The S_r field, in contrast, displays a patchwork of larger uncertainties in the region $x_1 \in [1.6, 2]$, $x_2 < 0.5$, which the FTLE field fails to reveal—indeed, this is an area in which the FTLE is *small*! It is not surprising that the regions most sensitive to stochasticity match up reasonably well with the largest exponential separation regions, since disturbances in an exponentially stretching region will tend to amplify. The fact that there are differences should also not be a surprise: S^2 tracks the impact of any specified stochastic model σ, does so *continuously with time*, and consequently captures the influence of ∇F_0^t

across all $t \in [0, T]$, while in contrast FTLEs only require ∇F_0^T and are confined to deterministic effects. Moreover, in instances in which σ is spatio-temporally dependent, the S^2 field can be significantly different from the FTLE field compared to the $\sigma = \mathrm{Id}$ situation pictured in Figure 8, because the S^2 field incorporates nonuniformity while the FTLE field does not. In general, the S^2 field (not shown, since its scaled version S_r is shown instead) appears to be much sharper than the FTLE field, with anomalously large values being concentrated along curves. Indeed, the ability of the stochastic sensitivity to reveal "ridge-like" objects related to stable/unstable manifolds in a sharper fashion than the FTLE is visible in the dark blue curve emanating upwards from near $(1.05, 0)$ in Figure 8(a), in contrast with its appearing less distinctly in Figure 8(b).

While an explicit expression for the velocity field was available for the computations demonstrated here, it must be emphasized that this procedure has been developed so that it can also be carried out using velocity *data* (obtained from observations, experiments, or computational fluid dynamics simulations), by choosing Δt to be the time spacing of the data, and utilizing the spatial resolution L_r of the data to form the relevant grid. The utility of the expressions for S^2, S_r, S_n, A, $\theta_{\max}$, and R is that they can (and will) be applied in more realistic situations such as those, thereby allowing the quantification of uncertainty in identifying important flow structures.

4. Conclusions and Outlook. Interest in the impact of uncertainties in the Eulerian velocity field has only recently been emerging: finite-time Lyapunov [44, 8] and Lagrangian diagnostics [13] calculations using ensembles of stochastic realizations, fattening of material curves [6], fuzziness imparted on stable/unstable manifolds and consequent mixing [9], surfaces across which diffusive flux is minimal [50], and a numerical method for transfer operator computation which supplants the standard initial- and/or end-time set diffusion by continuous-time diffusion [26]. This article presents a particular framework for quantifying the resulting Lagrangian uncertainties as a physically interpretable *field* which is easily computable using velocity data. The uncertainty in the Eulerian data can be due to unavoidable accuracy issues in the data, measurement error, availability only on a spatial and temporal grid, the presence of turbulence at subgrid levels, and other uncaptured effects. For two-dimensional Eulerian velocity data which is available only for a finite time, the method works independently of whether the data is divergence-free or whether there is complicated time-dependence. To be specific, tools which quantify the uncertainty in the final Lagrangian positions, including the anisotropy and a method for extracting robust sets, have been developed using a stochastic differential equation approach.

One can choose fairly general stochastic models for the Eulerian uncertainties by specifying different spatio-temporally dependent diffusion matrices σ. This is powerful because, for example, diffusion which is nonuniform and anisotropic appears to be present in oceanic flows [18, 72, 76]. Moreover, σ's (x, t)-dependence can be used to model various experimentally or observationally relevant uncertainties: cloud cover over certain regions at certain times when taking satellite measurements; the fact that measurement uncertainties are greater in the periphery of a camera's visual range; having σ identically zero at gridpoints but nonzero away from them to assess resolution error, etc. An alternative view of the role of spatial resolution is gained from investigating the connection between the lengthscale predicted from stochastic sensitivity analysis (i.e., $L_s = \varepsilon\sqrt{S^2(x)}$) and the optimal spatial resolution lengthscale; preliminary results based on turbulence experiments and available oceanographic data suggest that using resolutions finer than L_s is not only unnecessary, but also counter-

productive [29]. Additional investigations on the issue of spatial resolution L_r—also using the resolution-scaled stochastic sensitivity (2.27)—are underway. The ability to deal with two-dimensional non-divergence-free Eulerian velocities is also useful because available data often has these features. For example, oceanic velocity data is mainly deduced from satellite measurements of sea-surface heights, and particle image velocimetry data is easiest to obtain in two dimensions. In such instances, even though the full three-dimensional velocity field is expected to be divergence-free because water is incompressible, the two-dimensional signature is *not*. Thus, the methodology can also be applied to dynamically evolving models in nonfluids contexts in which area-preservation is violated.

These tools are expected to have a significant impact on the research area of "Lagrangian coherent structures," in which there has hitherto been little explicit incorporation of uncertainty. (Numerical diffusivity of course affects any deterministic computations in these approaches, but does so in an uncontrolled fashion.) Many Lagrangian coherent structure methods use *ad hoc* definitions to extract coherent structures. An extensive discussion is provided in [12]; quick examples include ridges of FTLE fields [69], curves/surfaces to which there is extremal attraction [47], stable/unstable manifold curves defined by extensions of time [4], contours which are convex [49]), and clustering methods which attempt to group together sets or particles which have "similar" behavior [65, 39]. Most of these methods are explicitly confined to *deterministic* Lagrangian advection of (2.1). It should be mentioned that entities extracted from each method do not necessarily match those from another method, as is indicated, for example, in comparisons in Figure 1 in [45] and Figure 4 in [12]. Rather than follow the more established (and hence deterministic) Lagrangian coherent structure methods, stochastic sensitivity extracts robust sets explicitly in terms of robustness toward noise, with the essential ability to specify threshold values for the Péclet number and the lengthscale of fluid elements whose robustness is quantified (see Definition 2.9). At different levels of these physical characteristics, different robust sets can be extracted (see Figure 5(c,d)); these identify sets at time 0 associated with the flow from time 0 to time T. There is no claim (or attempt to capture) material advection of such sets, because they are defined precisely in terms of uncertainties in Lagrangian prediction; changing T, for example, will identify a *different* set at time 0, whereas the lengthscale L and Péclet number Pe parametrize the level of robustness in the set. As such, the stochastic sensitivity measures developed here form a natural, physically motivated way of identifying structures which are *robust to velocity perturbations*. Of course, this can be construed as an alternative way of defining "coherence" which is completely different from standard methods in Lagrangian coherent structure analysis.

It is possible that the impact of uncertainties in eventual Lagrangian locations can be incorporated into many standard Lagrangian coherent structure methods, because they are (mostly) assessed based on the deterministic flow map from time 0 to T. Thus, a spatially varying uncertainty can be assigned to this map, and the consequences investigated. A study on the uncertainty inherited by one method—FTLEs—has already been performed [8]. Other studies are anticipated in the future, allowing for uncertainties to be assigned to different types of "coherence" claims.

The models used here—namely, a stochastic differential equation for the velocity field and the Fokker–Planck equation for density evolution—have been used in porous media and contaminant spreading work [27, 28]. In these cases, the first of these models is typically a Langevin equation (with *incompressible* [28] and either *steady* [27]

or *statistically homogeneous* [28] velocity field and *constant isotropic* diffusion [27, 28]), resulting in the second model becoming a classical advection-diffusion equation. In these situations, due to the availability of methods such as Green's functions and manipulations using Fourier transforms [27] or assumed probability models [28], some formal results on the contaminant dispersion and center of mass variance have been obtained. Stochastic fluctuations are a natural way to model the pore-scale variations; however, the setting is somewhat different in these studies because this article relies on available velocity measurements, which is typically difficult in porous flows. More specifically, this work relates to using velocity data to predict spatial structures which are robust under stochasticity. With these observations in mind, relationships to porous media which use stochastic fluctuation models [54, 27, 28] are being pursued.

The approach of this article is strongly connected to ideas in uncertainty quantification. Uncertainties engendered in evolving systems are generically due to two aspects: (i) uncertainty in the initial position in Ω_0, and (ii) uncertainty in the evolving dynamics from times 0 to T. Type (i), in the context of Lagrangian trajectories, possibly has a "simple" solution within the present framework: an uncertainty δx in the location x at time 0 will translate to a leading-order uncertainty $\nabla F_0^T(x)\delta x$ in Ω_T. Indeed, this or derived quantities are what is used in many basic Lagrangian coherent structure works; the idea is to determine what happens when sets are changed slightly at time 0, but the flow map F_0^T is kept fixed. Some typical examples that illustrate how uncertainty type (i) applies to Lagrangian coherent structure methods include the following examples.

- Maximizing $\left|\nabla F_0^T(x)\delta x\right| / |\delta x|$ over all initial directions gives the FTLE field on Ω_0 [69, 47, 10].
- Haller's variational approaches [46, 47, 48] ask questions about varying curves or surfaces in Ω_0, in relation to a deterministic F_0^T.
- Froyland's Perron–Frobenius (transfer) operator approaches [34, 35, 38] sometimes apply "fattening" of sets at time 0 before pushing forward by the deterministic map F_0^T. (However, it must be noted that in *many* implementations of this process [40, 45, 12, 57], such a diffusion/fattening is not explicitly included, and the transfer operator is numerically computed as the push-forward operator on densities in Ω_0 by the *deterministic* map F_0^T. Uncontrolled and implicit *numerical diffusion* is of course present in this process.)

In the first two examples δx is thought of as a deterministic quantity (over which an optimization is performed), and hence is not typically imagined to be "an uncertainty." Nevertheless, the approaches rely on understanding the sensitivity toward changes δx in Ω_0. The third example, when using an *explicit* diffusion, may be viewed as possessing a stochastic δx and so can be considered to fall within the type (i) uncertainty quantification realm. It is highlighted that uncertainties of type (ii)—the impact of *uncertainties in the evolving dynamics* on eventual Lagrangian locations—do not seem to have been explicitly addressed until now (though hints of them appear in some recent diffusive approaches [26, 50, 6, 9]). Thus, a fundamental contribution to uncertainties in Lagrangian trajectories has been made.

The Fokker–Planck equation (2.4) is intimately connected to the idea of uncertainties, because it captures the evolution of a probability density function. Additionally, when the Lagrangian transport of various scalar quantities—as opposed to mere trajectories—is relevant, the governing advection-diffusion equation (see, e.g., [75, 73, 74]) falls within the framework of a Fokker–Planck equation. Therefore, the connection between the stochastic sensitivity measures developed here and the

Fokker–Planck equation (2.4) invites further exploration [7], particularly because simulations (as in Figure 2) confirm this promising connection. An intriguing question is whether both aspects of uncertainty quantification (uncertainty in initial position, and uncertainty in evolving dynamics) can be tackled *together* within the framework of the Fokker–Planck equation. The initial uncertainty can be considered as the smearing out of a Dirac mass initial condition for this equation (consonant with the approaches in [35, 34, 38]), whereas what is done in this article is equivalent to starting from an *exact* Dirac mass initial condition and then quantifying the eventual spread of the evolved density. In this sense, type (i) uncertainties are associated with a choice of initial condition to the Fokker–Planck equation (2.4), while type (ii) are velocity uncertainties which are captured by the diffusion term σ in the Fokker–Planck equation. Investigations of this connection, and possible amalgamation of these two ideas relating to eventual uncertainty, are being pursued in both theoretical and computational senses.

One method of improving uncertainties in predictions is by using data assimilation [66, 55, 70, 59, 3], in which additional existing knowledge is used to decide on the reliability of model predictions and/or to discard questionable information. The stochastic sensitivity field $S^2(x)$ and its scaled versions given here may provide new tools in such data assimilation methods, because $S^2(x)$ can be thought of as assigning a *confidence weight* to each predicted Lagrangian trajectory by identifying with an initial condition $x \in \Omega_0$. Put another way, stochastic sensitivity can quantify the *model error* [3, 58] (in a spatially nonuniform fashion) in using Eulerian velocity measurements to predict Lagrangian particle trajectories. This application is currently being pursued with collaborators with expertise in data assimilation.

Clearly, an important extension of the current theory would be to three-dimensions, which initially seems to provide significant stumbling blocks because the proof presented in Appendix B relies on two-dimensionality. Incorporating aspects of Melnikov-like developments in higher dimensions [42, 11, 78, 77] seems called for; however, the stochastic sensitivity, unlike those developments, cannot have any prescribed geometry and must genuinely provide a *global* scalar field on Ω_0. Extensions to *higher* dimensions have the conceptual benefit of thinking of u in (2.1) not merely as a velocity field, but rather as *any* deterministic model which describes the evolution of multiple species (encoded as components of a state vector x). Then the stochastic sensitivity measures describe the susceptibility of the model's conclusions to uncertainty in the model.

Another theoretical extension would be to pursue higher moments of the random deviation. (There is no reason to expect normality in the statistics of Z_ε, as evidenced by related work on stochastic advection [74, 9] and Figure 3.) This would be an attempt to garner more detail on the statistics of the Lagrangian deviation, and may enable a connection to Freidlin–Wentzell large deviation theory [33, 22, 25].

Finally, the efficacy and accuracy of using the stochastic sensitivity measures—in comparison to other Lagrangian coherent structure methods [12]—to noisy and/or low-resolution velocity data to identify robust structures is currently under investigation in collaboration with experimental colleagues. Using toy models such as the double-gyre is well and good, but realistic flows display many other features (turbulence, energy at different scales, energy cascades, etc.) which cannot possibly be included in any model with *specified* Eulerian velocity. Data obtained from such systems inevitably suffers spatial resolution issues, and any computational work done with this data implicitly or explicitly smoothly interpolates it. In smooth flows, ve-

locities are only correct to leading order in the spatial resolution; in turbulent ones, a smooth velocity field is deduced, ignoring the fact that the Eulerian velocity may not be smooth. Therefore, there are clear model uncertainties whether or not the advection performed is "deterministic." Using stochastic sensitivity measures allows for a practical way to quantify and evaluate the impact of these effects. Specifically, spatial resolution and system diffusion are, respectively, exemplified by the two scalings of S^2 suggested: S_r and S_n. For example, preliminary work indicates that the stochastic sensitivity is able to identify structures even when spurious noise is introduced into an experimental velocity field, significantly more robustly than do FTLEs [29]. Studying the robustness of the stochastic sensitivity measure to such noise, and low-resolution data, is ongoing.

In view of these applications, stochastic sensitivity, its anisotropic quantification, and robust sets provide a novel set of tools allowing (for the first time) *uncertainties* to be ascribed to Lagrangian motion in fluid flows. The fact that these uncertainties can be parametrized in terms of both the diffusion parameter and the lengthscale, as well as its nonuniform distribution across initial conditions, will allow for powerful new applications in the areas of Lagrangian coherent structures, Lagrangian data assimilation, and Lagrangian motion in turbulent flows, and, in particular, will provide links to experimental and observational considerations including spatial resolution, diffusion, stochastic parametrization, and nonuniform uncertainties.

Appendix A. Proof of Lemma 2.2.

A preliminary lemma will be necessary.

LEMMA A.1. *If y_τ is a solution of* (2.3) *for $\tau \in [0,T]$, then for any $p > 0$, there exists a constant K_Σ^p such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sigma(y_\tau,\tau)\,\mathrm{d}W_\tau\right|^{2p}\right] \le K_\Sigma^p\,. \tag{A.1}$$

Proof. This proof relies on a special case of the Burkhölder–Davis–Gundy inequality (e.g., Theorem 5.6.3 in [51]) which, given any local martingale N_t, asserts the existence of C_p for any $p > 0$ such that

$$\mathbb{E}\left[\sup_{t\in[0,T]}|N_t|^{2p}\right] \le C_p\,\mathbb{E}\left[N_T^p\right]\,. \tag{A.2}$$

Set

$$N_t := \int_0^t \sigma_\tau^{(1)}\,\mathrm{d}W_\tau\,,$$

in which $\sigma_\tau^{(1)}$ is the first row of the matrix $\sigma_\tau(y_\tau,\tau)$; this is a local martingale because σ is uniformly bounded as given in (2.8). Applying (A.2) with $p = 1$ yields the existence of C_1 such that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sigma_\tau^{(1)}\,\mathrm{d}W_\tau\right|^2\right] \le C_1\,\mathbb{E}\left[\left\{\int_0^T\left|\sigma_\tau^{(1)}\right|^2\,\mathrm{d}\tau\right\}\right] \le 2C_1K_\sigma^2T\,,$$

with the second inequality arising from (2.8). Now let $p > 0$ be arbitrary. An identical

bound can be applied to $\sigma_\tau^{(2)}$, the second row of σ, and so

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^T \sigma(y_\tau,\tau)\,\mathrm{d}W_\tau\right|^{2p}\right] &= \mathbb{E}\left[\sup_{t\in[0,T]}\left\{\left|\int_0^T \sigma_\tau^{(1)}\,\mathrm{d}W_\tau\right|^2+\left|\int_0^T \sigma_\tau^{(2)}\,\mathrm{d}W_\tau\right|^2\right\}^p\right]\\ &\le \mathbb{E}\left[\left\{2C_1K_\sigma^2T+2C_1K_\sigma^2T\right\}^p\right]\\ &= \left(4C_1K_\sigma^2T\right)^p =: K_\Sigma^p\,,\end{aligned}$$

as desired. ∎

Now, let us turn to the proof of Lemma 2.2. In (2.10), the (exact) stochastic differential equation for $y_t - F_0^t(x)$ has already been established. Taking absolute values of both sides,

$$\left|y_t - F_0^t(x)\right| \le \left|\int_0^t u(y_\tau,\tau) - u\left(F_0^\tau(x),\tau\right)\mathrm{d}\tau\right| + \varepsilon\left|\int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\right|. \tag{A.3}$$

Noting that $|a_1+a_2|^q \le 2^{q-1}\left(|a_1|^q+|a_2|^q\right)$ for $q\ge 1$, the qth power of (A.3) yields

$$\begin{aligned}\left|y_t - F_0^t(x)\right|^q &\le 2^{q-1}\left(\left|\int_0^t u(y_\tau,\tau) - u\left(F_0^\tau(x),\tau\right)\mathrm{d}\tau\right|^q + \varepsilon^q\left|\int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\right|^q\right)\\ &\le 2^{q-1}\left(T^{q-1}\int_0^t \left|u(y_\tau,\tau) - u\left(F_0^\tau(x),\tau\right)\right|^q\mathrm{d}\tau + \varepsilon^q\left|\int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\right|^q\right)\\ &\le 2^{q-1}\left(T^{q-1}\eta^q\int_0^t \left|y_\tau - F_0^\tau(x)\right|^q\mathrm{d}\tau + \varepsilon^q\left|\int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\right|^q\right)\end{aligned}$$

by using Hölder's inequality and the Lipschitz condition (2.5) on u. Next, consider applying the supremum over all $t\in[0,T]$ to both sides of the above inequality, followed by taking the expectation. This gives

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_t - F_0^t(x)\right|^q\right] &\le 2^{q-1}T^{q-1}\eta^q\,\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t \left|y_\tau - F_0^\tau(x)\right|^q\mathrm{d}\tau\right]\\ &\qquad + 2^{q-1}\varepsilon^q\,\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sigma\left(y_\tau,\tau\right)\,\mathrm{d}W_\tau\right|^q\right]\\ &\le 2^{q-1}T^{q-1}\eta^q\mathbb{E}\left[\int_0^T \left|y_\tau - F_0^\tau(x)\right|^q\mathrm{d}\tau\right] + 2^{q-1}\varepsilon^q K_\Sigma^{q/2}\\ &\le 2^{q-1}T^{q-1}\eta^q\int_0^T \mathbb{E}\left[\sup_{\xi\in[0,\tau]}\left|y_\xi - F_0^\xi(x)\right|^q\right]\mathrm{d}\tau + 2^{q-1}\varepsilon^q K_\Sigma^{q/2}\end{aligned}$$

by utilizing (A.1). Applying Gronwall's inequality now gives the bound

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_t - F_0^t(x)\right|^q\right] \le 2^{q-1}\varepsilon^q K_\Sigma^{q/2}\exp\left(2^{q-1}T^{q-1}\eta^qT\right),$$

which leads to the desired result

$$\mathbb{E}\left[\sup_{t\in[0,T]}\frac{\left|y_t - F_0^t(x)\right|^q}{\varepsilon^q}\right] \le 2^{q-1}K_\Sigma^{q/2}\exp\left(2^{q-1}T^q\eta^q\right) := K_z^q\,.$$

Appendix B. Proof of Theorem 2.3. The strategy used here builds up expressions to be used toward the main goal (obtaining the *variance* of $P_\varepsilon(x,\theta)$ as per Theorems 2.5 and 2.7), rather than merely establishing Theorem 2.3, which can be obtained by simpler means as well.

While a major impediment in using the (formal) integral equation (2.12) to quantify $Z_\varepsilon(x)$ was the nonautonomous nature of the coefficient matrix, classical (deterministic) Melnikov methods [5, 42, 4, 11, 77] are able to deal with a *specific* projection of $Z_\varepsilon(x)$ associated with stable/unstable manifolds. (For a recent stochastic Melnikov approach—with the goal of establishing intersections between stable and unstable manifolds for a particular realization of noise—see [78].) Here, a Melnikov-like approach is combined with stochastic calculus to obtain the variance of a *general* projection (onto any direction θ) of $Z_\varepsilon(x)$. Since $F_0^t(x)$ gives the flow map of the deterministic flow (2.1),

$$\frac{\partial}{\partial t}F_0^t(x) = u\left(F_0^t(x),t\right)\,. \tag{B.1}$$

Taking the gradient with respect to x yields

$$\frac{\partial}{\partial t}\nabla F_0^t(x) = \nabla u\left(F_0^t(x),t\right)\,\nabla F_0^t(x)\,, \tag{B.2}$$

which states that $\nabla F_0^t(x)$ is a solution to the equation of variations of (2.1). The projection of $Z_\varepsilon(x)$ is sought in the directions defined by $\theta \in [-\pi/2,\pi/2)$. To this end, choose the infinitesimal quantity

$$\delta w = -\delta\, J\,\hat{n}(\theta) = \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix}\,\delta$$

for some small $\delta > 0$. With the observation that the differentials associated with the map $x \to w$ in (2.15) are connected by

$$\delta w = \nabla F_0^T(x)\delta x \quad \text{and hence} \quad J\delta w = J\nabla F_0^T(x)\delta x\,,$$

it is easy to see that the direction θ in which projections of $Z_\varepsilon(x)$ are sought (as given in (2.16)) obeys

$$\hat{n}(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = J\frac{\delta w}{|\delta w|} = \frac{J\nabla F_0^T(x)\delta x}{\left|\nabla F_0^T(x)\delta x\right|}\,.$$

Upon defining the scalar Itô process

$$M_t(y_t) := \left[J\nabla F_0^t(x)\delta x\right]^\top\left[y_t - F_0^t(x)\right] = \varepsilon\left[J\nabla F_0^t(x)\delta x\right]^\top z_\varepsilon(x,t) \tag{B.3}$$

for times $t \in [0,T]$, (2.16) and (B.3) indicate that the required projection is

$$P_\varepsilon(x,\theta) = \frac{M_T(y_T)}{\varepsilon\left|J\nabla F_0^T(x)\delta x\right|}\,. \tag{B.4}$$

Understanding the final outcome of the Itô process M_T is therefore a first step in finding the statistics of the projection $P_\varepsilon(x,\theta)$ (it is clear that $\mathbb{E}\left[P_\varepsilon\right] = \mathcal{O}(1)$ because Lemma 2.2 implies that for fixed δx, $\mathbb{E}\left[M_t(y_t)\right] = \mathcal{O}(\varepsilon)$).

A solvable differential equation for M_t is now sought, following the spirit of the derivation of Melnikov theory in dynamical systems [5, 43], but in the stochastic

context and independent of any invariant manifolds. Itô's lemma [51] applied to (B.3) in conjunction with y_t's evolution equation (2.3) gives that

$$\mathrm{d}M_t = \left\{ \frac{\partial M_t}{\partial t} + (\nabla M_t)^\top u + \frac{1}{2}\,\mathrm{Tr}\,\left[(\varepsilon\sigma)^\top (\nabla\nabla M_t)\,(\varepsilon\sigma) \right] \right\} \mathrm{d}t + (\nabla M_t)^\top (\varepsilon\sigma)\,\mathrm{d}W_t\,,$$

where all the quantities above are evaluated at (y_t, t). The quantity $\nabla\nabla M_t$ is the Hessian with respect to the spatial argument y_t which, given the linearity of M_t with respect to y_t in (B.3), disappears, and so

$$\mathrm{d}M_t = \left\{ \frac{\partial M_t}{\partial t} + (\nabla M_t)^\top u \right\} \mathrm{d}t + (\nabla M_t)^\top (\varepsilon\sigma)\,\mathrm{d}W_t\,. \tag{B.5}$$

Expressions for the various terms in (B.5) will now be built up. First, from M_t as defined in (B.3),

$$\begin{aligned}
\frac{\partial M_t}{\partial t} &= \left[J\nabla F_0^t(x)\delta x \right]^\top \left[-\frac{\partial}{\partial t} F_0^t(x) \right] + \left[J\frac{\partial}{\partial t}\left(\nabla F_0^t(x) \right)\delta x \right]^\top \left[y_t - F_0^t(x) \right] \\
&= \left[J\nabla F_0^t(x)\delta x \right]^\top \left[-u\left(F_0^t(x), t \right) \right] + \varepsilon \left[J\nabla u\left(F_0^t(x), t \right)\, \nabla F_0^t(x)\delta x \right]^\top z_\varepsilon(x,t)\,,
\end{aligned}$$

where (B.1) and (B.2) have been used. The next term in (B.5) can be expressed by first taking the gradient of (B.3) with respect to the spatial variable y_t, such that

$$\begin{aligned}
(\nabla M_t(y_t))^\top u\,(y_t,t) &= \left[J\nabla F_0^t(x)\delta x \right]^\top u\,(y_t,t) \\
&= \left[J\nabla F_0^t(x)\delta x \right]^\top \left[u\left(F_0^t(x), t \right) + \nabla u\left(F_0^t(x), t \right)\left(y_t - F_0^t(x) \right) \right] \\
&\quad + \frac{1}{2}\left[J\nabla F_0^t(x)\delta x \right]^\top \left[\left(y_t - F_0^t(x) \right)^\top \nabla\nabla u\,(\xi_1, t)\left(y_t - F_0^t(x) \right) \right] \\
&= \left[J\nabla F_0^t(x)\delta x \right]^\top \left[u\left(F_0^t(x), t \right) + \varepsilon\nabla u\left(F_0^t(x), t \right) z_\varepsilon(x,t) \right] \\
&\quad + \frac{\varepsilon^2}{2}\left[J\nabla F_0^t(x)\delta x \right]^\top \left[z_\varepsilon(x,t)^\top \nabla\nabla u\,(\xi_1, t)\, z_\varepsilon(x,t) \right],
\end{aligned}$$

where the second equality is obtained by Taylor expanding u around $F_0^t(x)$, and $\nabla\nabla u(\xi_1, t)$ is the second-derivative tensor of u evaluated at some spatial point ξ_1. By a similar argument,

$$\begin{aligned}
(\nabla M_t(y_t))^\top(\varepsilon\sigma(y_t,t)) &= \left[J\nabla F_0^t(x)\delta x \right]^\top \varepsilon \left[\sigma\left(F_0^t(x), t \right) + \nabla\sigma\,(\xi_2, t)\left(y_t - F_0^t(x) \right) \right] \\
&= \varepsilon\left[J\nabla F_0^t(x)\delta x \right]^\top \left[\sigma\left(F_0^t(x), t \right) \right] + \varepsilon^2 \left[J\nabla F_0^t(x)\delta x \right]^\top \nabla\sigma(\xi_2, t) z_\varepsilon(x,t),
\end{aligned}$$

in which ξ_2 is a spatial point. Substituting each of these three expressions into (B.5) gives the expression

$$\begin{aligned}
\mathrm{d}M_t &= \varepsilon \left\{ \left(J\nabla u \nabla F_0^t(x)\delta x \right)^\top z_\varepsilon(x,t) + \left(J\nabla F_0^t(x)\delta x \right)^\top \nabla u\, z_\varepsilon(x,t) \right\} \mathrm{d}t \\
&\quad + \varepsilon \left(J\nabla F_0^t(x)\delta x \right)^\top \sigma\,\mathrm{d}W_t + \frac{\varepsilon^2}{2}\left[J\nabla F_0^t(x)\delta x \right]^\top z_\varepsilon(x,t)^\top \nabla\nabla u\,(\xi_1, t)\, z_\varepsilon(x,t)\mathrm{d}t \\
&\quad + \varepsilon^2 \left[J\nabla F_0^t(x)\delta x \right]^\top \nabla\sigma\,(\xi_2, t)\, z_\varepsilon(x,t)\mathrm{d}W_t\,,
\end{aligned}$$

where, when omitted for u and σ, the arguments are $(F_0^t(x), t)$. For the first line in the final expression above, set $A = \nabla u$, $b = z_\varepsilon(x,t)$, and $c = \nabla F_0^t(x)\delta x$, and invoke the identity

$$(JAc)^\top b + (Jc)^\top Ab = (\mathrm{Tr}\,A)\, b^\top Jc$$

for 2×2 matrices A and 2×1 vectors b and c which is well established in various contexts (see (A5) in [6], (2.39) in [5], (4.5.8) in [43], or 3.1 in [4]), thereby arriving at

$$\begin{aligned}\mathrm{d}M_t = {} & \varepsilon \left[\nabla \cdot u\right] z_\varepsilon(x,t)^\top \left(J\nabla F_0^t(x)\delta x\right) \mathrm{d}t + \varepsilon \left(J\nabla F_0^t(x)\delta x\right)^\top \sigma \, \mathrm{d}W_t \\ & + \varepsilon^2 \left[J\nabla F_0^t(x)\delta x\right]^\top \left(\frac{1}{2} z_\varepsilon(x,t)^\top \nabla\nabla u\left(\xi_1,t\right) z_\varepsilon(x,t)\mathrm{d}t + \nabla\sigma\left(\xi_2,t\right) z_\varepsilon(x,t)\mathrm{d}W_t\right).\end{aligned}$$

However, from (B.3), the term multiplying $\mathrm{Tr}\,(\nabla u) = \nabla \cdot u$ in the first line above is simply $M_t(y_t)$, and so the above reduces to

$$\begin{aligned}\mathrm{d}M_t = {} & \left[\nabla \cdot u\right]\left(F_0^t(x),t\right) M_t \mathrm{d}t + \varepsilon \left[J\nabla F_0^t(x)\delta x\right]^\top \sigma\left(F_0^t(x),t\right) \mathrm{d}W_t \\ & + \varepsilon^2 \left[J\nabla F_0^t(x)\delta x\right]^\top \left(\frac{1}{2} z_\varepsilon(x,t)^\top \nabla\nabla u\left(\xi_1,t\right) z_\varepsilon(x,t)\mathrm{d}t + \nabla\sigma\left(\xi_2,t\right) z_\varepsilon(x,t)\mathrm{d}W_t\right).\end{aligned}$$

Using the standard integrating factor approach allows this to be rewritten as

$$\begin{aligned}\mathrm{d}\left[e^{-\int_0^t [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} M_t\right] = {} & \varepsilon e^{-\int_0^t [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left(J\nabla F_0^t(x)\delta x\right)^\top \sigma\left(F_0^t(x),t\right) \mathrm{d}W_t \\ & + \frac{\varepsilon^2}{2} e^{-\int_0^t [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left[J\nabla F_0^t(x)\delta x\right]^\top z_\varepsilon(x,t)^\top \nabla\nabla u\left(\xi_1,t\right) z_\varepsilon(x,t)\mathrm{d}t \\ & + \varepsilon^2 e^{-\int_0^t [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left[J\nabla F_0^t(x)\delta x\right]^\top \nabla\sigma\left(\xi_2,t\right) z_\varepsilon(x,t)\mathrm{d}W_t\,.\end{aligned}$$

Next, the above expression is integrated from 0 to T, bearing in mind that $y_0 = x$ and thus $M_0(y_0) = 0$, to yield

(B.6)

$$\begin{aligned}M_T(y_T) = {} & \varepsilon \int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left(J\nabla F_0^t(x)\delta x\right)^\top \sigma\left(F_0^t(x),t\right) \mathrm{d}W_t \\ & + \frac{\varepsilon^2}{2} \int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left[J\nabla F_0^t(x)\delta x\right]^\top z_\varepsilon(x,t)^\top \nabla\nabla u\left(\xi_1,t\right) z_\varepsilon(x,t)\mathrm{d}t \\ & + \varepsilon^2 \int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_0^\xi(x),\xi\right)\mathrm{d}\xi} \left[J\nabla F_0^t(x)\delta x\right]^\top \nabla\sigma\left(\xi_2,t\right) z_\varepsilon(x,t)\,\mathrm{d}W_t\,.\end{aligned}$$

From (B.4), in computing $P_\varepsilon(x,\theta)$, it is necessary to divide the above expression by the quantity $\left|J\nabla F_0^T(x)\delta x\right|$. All three terms above also possess a factor $\left(J\nabla F_0^t(x)\delta x\right)^\top$, and so the generated ratio will be rewritten in a useful way. Since $x = F_T^0(w)$, for any t,

$$F_0^t(x) = F_0^t\left(F_T^0(w)\right) = F_T^t(w)\,,$$

and hence

$$\nabla F_0^t(x)\delta x = \nabla F_T^t(w)\delta w \quad \text{and} \quad J\nabla F_0^t(x)\delta x = J\nabla F_T^t(w)\delta w\,.$$

This means that

$$\frac{\left[J\nabla F_0^t(x)\delta x\right]^\top}{\left|J\nabla F_0^T(x)\delta x\right|} = \frac{\left[J\nabla F_T^t(w)\delta w\right]^\top}{\left|J\nabla F_T^T(w)\delta w\right|} = \left[J\nabla F_T^t(w)\frac{\delta w}{|J\delta w|}\right]^\top = \left[J\nabla F_T^t(w)\begin{pmatrix}\sin\theta\\ -\cos\theta\end{pmatrix}\right]^\top.$$

Defining for $\theta \in [-\pi/2, \pi/2)$ the quantity

$$Q_\theta(w,t) := -J\,\nabla F_T^t(w)\,J\,\hat{n}(\theta) \tag{B.7}$$

allows the rewriting

(B.8)
$$\frac{\left[J\nabla F_0^t(x)\delta x\right]^\top}{\left|J\nabla F_0^T(x)\delta x\right|} = \left[J\nabla F_T^t(w)\begin{pmatrix}\sin\theta\\ -\cos\theta\end{pmatrix}\right]^\top = \left[J\nabla F_T^t(w)\,(-J)\,\hat{n}(\theta)\right]^\top = Q_\theta(w,t)^\top .$$

Substituting (B.6) into the projection definition (B.4) and utilizing (B.8) results in

$$P_\varepsilon(x,\theta) = I_1(\theta) + \varepsilon I_2(\theta) + \varepsilon I_3(\theta)\,, \tag{B.9}$$

in which

$$\begin{aligned} I_1(\theta) &:= \int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi}\, Q_\theta(w,t)^\top \sigma\left(F_T^t(w),t\right)\,\mathrm{d}W_t\,,\\ I_2(\theta) &:= \frac{1}{2}\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi}\, Q_\theta(w,t)^\top z_\varepsilon(x,t)^\top \nabla\nabla u\,(\xi_1,t)\, z_\varepsilon(x,t)\,\mathrm{d}t\,,\\ I_3(\theta) &:= \int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi}\, Q_\theta(w,t)^\top \nabla\sigma\,(\xi_2,t)\, z_\varepsilon(x,t)\,\mathrm{d}W_t\,. \end{aligned} \tag{B.10}$$

In the above, some of the x-dependence has been transformed to w-dependence using $F_0^t(x) = F_T^t(w)$ and $F_0^\xi(x) = F_T^\xi(w)$. Equation (B.9) provides an explicit expression for P_ε in terms of stochastic integrals. Now,

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left[P_\varepsilon(x,\theta)\right] = \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[I_1(\theta)\right] + \lim_{\varepsilon\downarrow 0}\left(\varepsilon\,\mathbb{E}\left[I_2(\theta)\right]\right) + \lim_{\varepsilon\downarrow 0}\left(\varepsilon\,\mathbb{E}\left[I_3(\theta)\right]\right)\,. \tag{B.11}$$

The presence of the prefactor ε in the last two terms in (B.11) indicates that, as long as the expectations are bounded, they can be discarded. For I_2,

$$\begin{aligned} \mathbb{E}\left[|I_2(\theta)|\right] &\le \frac{1}{2}\mathbb{E}\left[\left|\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} Q_\theta(w,t)^\top z_\varepsilon(x,t)^\top \nabla\nabla u\,(\xi_1,t)\, z_\varepsilon(x,t)\,\mathrm{d}t\right|\right]\\ &\le \frac{1}{2}\,\mathbb{E}\left[\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} K_F K_u \left|z_\varepsilon(x,t)\right|^2\,\mathrm{d}t\right]\\ &= \frac{1}{2}\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} K_F K_u \mathbb{E}\left[\left|z_\varepsilon(x,t)\right|^2\right]\,\mathrm{d}t < \infty\,. \end{aligned}$$

In the above, a bound on Q_θ (defined in (B.7)) has been obtained as

$$|Q_\theta(x,t)| = \left|J\nabla F_T^t(w)\, J\hat{n}(\theta)\right| \le \left|\nabla F_T^t(w)\right| \le K_F \tag{B.12}$$

resulting from the boundedness of the flow (2.7). Moreover, the second derivative of u is bounded using the smoothness conditions on u given by (2.6), and the z_ε terms

are bounded using Lemma 2.2 with $q = 2$. For the stochastic integral I_3,

(B.13)

$$\begin{aligned}
\mathbb{E}\left[|I_3(\theta)|\right] &\le \sqrt{\mathbb{E}\left[I_3(\theta)^2\right]} \\
&= \left\{\mathbb{E}\left[\left(\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} Q_\theta(w,t)^\top \nabla\sigma(\xi_2,t) z_\varepsilon(x,t)\mathrm{d}W_t\right)^2\right]\right\}^{1/2} \\
&= \left\{\mathbb{E}\left[\int_0^T e^{2\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} \left|Q_\theta(w,t)^\top \nabla\sigma\left(\xi_2,t\right) z_\varepsilon(x,t)\right|^2 \mathrm{d}t\right]\right\}^{1/2} \\
&\le \left\{\mathbb{E}\left[\int_0^T e^{2\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} K_F^2 K_\sigma^2 \left|z_\varepsilon(x,t)\right|^2 \mathrm{d}t\right]\right\}^{1/2} \\
&= K_F K_\sigma \left\{\int_0^T e^{2\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} \mathbb{E}\left[\left|z_\varepsilon(x,t)\right|^2\right] \mathrm{d}t\right\}^{1/2} < \infty\,.
\end{aligned}$$

In the above, the Itô isometry [51] has been used in the second equality, and the final boundedness result comes from Lemma 2.2 and the fact that the t-integral is bounded over the finite domain $[0,T]$. Thus, the terms I_2 and I_3 in (B.11) do not contribute. Hence,

$$\begin{aligned}
\lim_{\varepsilon\downarrow 0}\mathbb{E}\left[P_\varepsilon(x,\theta)\right] &= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[I_1(\theta)\right] \\
&= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} Q_\theta(w,t)^\top \sigma\left(F_T^t(w),t\right)\,\mathrm{d}W_t\right] \\
&= 0\,,
\end{aligned}$$

because the bounds on Q_θ (B.12) and σ (2.8) ensure that the integrand is square summable over t, and thus the expectation of the stochastic integral is zero [51]. Now, this establishes that, in the limit $\varepsilon \downarrow 0$, the expectation of the projection of $Z_\varepsilon(x)$ in *any* direction is zero. Thus, the expectation of $Z_\varepsilon(x)$ is zero in this limit.

Appendix C. Proof of Theorem 2.5. From (B.9), and using the result of Theorem 2.3,

$$\begin{aligned}
\lim_{\varepsilon\downarrow 0}\mathbb{V}\left[P_\varepsilon(x,\theta)\right] &= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\left|P_\varepsilon(x,\theta)\right|^2\right] - \lim_{\varepsilon\downarrow 0}\left|\mathbb{E}\left[P_\varepsilon(x,\theta)\right]\right|^2 \\
&= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\left|I_1(\theta)+\varepsilon I_2(\theta)+\varepsilon I_3(\theta)\right|^2\right] - 0 \\
&= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\left|I_1(\theta)\right|^2\right] + 2\lim_{\varepsilon\downarrow 0}\varepsilon\mathbb{E}\left[I_1(\theta)(I_2(\theta)+I_3(\theta))\right] \\
&\qquad + \lim_{\varepsilon\downarrow 0}\varepsilon^2\mathbb{E}\left[\left|I_2(\theta)+I_3(\theta)\right|^2\right]\,.
\end{aligned} \tag{C.1}$$

As long as the expectations in the last two terms are bounded, the presence of the prefactors of ε will ensure that those two terms will vanish. Using the Cauchy–Schwarz

inequality,

$$\begin{aligned}\mathbb{E}\Big[|I_2(\theta)|^2\Big] &= \frac{1}{4}\mathbb{E}\left[\left|\int_0^T e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi} Q_\theta(w,t)^\top z_\varepsilon(x,t)^\top \nabla\nabla u(\xi_1,t)\, z_\varepsilon(x,t)\,\mathrm{d}t\right|^2\right] \\ &\le \frac{1}{4}\,\mathbb{E}\left[\left(\int_0^T e^{2\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi}\mathrm{d}t\right)\right. \\ &\qquad \left.\times\left(\int_0^T \left|Q_\theta(w,t)^\top z_\varepsilon(x,t)^\top \nabla\nabla u\,(\xi_1,t)\, z_\varepsilon(x,t)\right|^2\mathrm{d}t\right)\right] \\ &\le \frac{1}{4}\left(\int_0^T e^{2\int_t^T [\nabla\cdot u]\left(F_T^\xi(w),\xi\right)\mathrm{d}\xi}\mathrm{d}t\right)\mathbb{E}\left[\int_0^T K_F^2 K_U^2\,|z_\varepsilon(x,t)|^4\;\mathrm{d}t\right] < \infty\,,\end{aligned}$$

because of Lemma 2.2 with $q = 4$. Since $\mathbb{E}\big[\,|I_3(\theta)|^2\,\big]$'s boundedness has already been established in (B.13),

$$\mathbb{E}\left[|I_2(\theta)+I_3(\theta)|^2\right] \le 2\,\mathbb{E}\left[|I_2(\theta)|^2+|I_3(\theta)|^2\right] < \infty\,.$$

Moreover,

$$\begin{aligned}\mathbb{E}\left[|I_1(\theta)\,(I_2(\theta)+I_3(\theta))|\right] &\le \mathbb{E}\left[|I_1(\theta)I_2(\theta)|\right]+\mathbb{E}\left[|I_1(\theta)I_3(\theta)|\right] \\ &\le \sqrt{\mathbb{E}\left[|I_1(\theta)|^2\right]\mathbb{E}\left[|I_2(\theta)|^2\right]}+\sqrt{\mathbb{E}\left[|I_1(\theta)|^2\right]\mathbb{E}\left[|I_3(\theta)|^2\right]}\,,\end{aligned}$$

which is bounded by the results above. This allows for the discarding of the last two of the three terms in (C.1), and hence

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\mathbb{V}\left[P_\varepsilon(x,\theta)\right] &= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[|I_1(\theta)|^2\right] \\ &= \int_0^T \left|e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w)\right),\xi)\mathrm{d}\xi} Q_\theta(w,t)^\top \sigma\left(F_T^t(w),t\right)\right|^2\,\mathrm{d}t \\ &= \int_0^T \left|e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w)\right),\xi)\mathrm{d}\xi}\,\sigma\left(F_T^t(w),t\right)^\top Q_\theta(w,t)\right|^2\,\mathrm{d}t \\ &= \int_0^T \left|e^{\int_t^T [\nabla\cdot u]\left(F_T^\xi(w)\right),\xi)\mathrm{d}\xi}\,\sigma\left(F_T^t(w),t\right)^\top(-J)\;\nabla F_T^t(w)\,J\,\hat{n}(\theta)\right|^2\,\mathrm{d}t \\ &= \int_0^T |\Lambda(w,t)\,J\,\hat{n}(\theta)|^2\;\mathrm{d}t\,,\end{aligned}$$

where the first step is accomplished by using Itô's isometry on (B.10), the second is obtained by rewriting the magnitude of a row vector as that of the corresponding column vector, the third results by invoking the definition for Q_θ in (B.7), and finally the fourth is obtained by using Λ's definition (2.21). Taking the square root gives $\tilde{A}(w,\theta)$ in (2.23). The fact that this is equal to $A(x,\theta)$ is simply because of the invertible relationship $w = F_0^T(x)$.

Appendix D. Proofs of Theorems 2.6 and 2.7. Theorems 2.6 and 2.7 will be proven together in this appendix. First, the operation $\sup_\theta$ is applied to (C.1). In view of the boundedness of all of the expectations of the quadratic terms I_iI_j as

established in Appendix C, the dominated convergence theorem allows for swapping the $\varepsilon \downarrow 0$ limit and the supremum over θ. Thus,

$$S^2(x) = \lim_{\varepsilon\downarrow 0} \sup_\theta \mathbb{E}\left[|I_1(\theta)|^2\right] = \sup_\theta \lim_{\varepsilon\downarrow 0} \mathbb{E}\left[|I_1(\theta)|^2\right] = \sup_\theta \left[A(x,\theta)\right]^2 ,$$

in which the ε-multiplied terms of (C.1) have been discarded because of the boundedness properties established in Appendix C. Therefore, from (2.23),

$$S^2(x) = \sup_\theta \int_0^T \left|\Lambda(w,t)\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}\right|^2 \mathrm{d}t =: \sup_\theta H(\theta).$$

Expanding in terms of the components of Λ and using standard trigonometric equalities allows $H(\theta)$ to be written as

$$\begin{aligned} H(\theta) &= \frac{1}{2}\int_0^T \sum_{i=1}^2\sum_{j=1}^2 \Lambda_{ij}^2(w,t)\mathrm{d}t \\ &\quad + \frac{\cos 2\theta}{2}\int_0^T\left[\sum_{i=1}^2 \Lambda_{i2}^2(w,t) - \sum_{i=1}^2 \Lambda_{i1}^2(w,t)\right]\mathrm{d}t - \sin 2\theta \int_0^T \sum_{i=1}^2 [\Lambda_{i1}(w,t)\Lambda_{i2}(w,t)]\mathrm{d}t \\ &= \frac{1}{2}\int_0^T \sum_{i=1}^2\sum_{j=1}^2 \Lambda_{ij}^2(w,t)\mathrm{d}t + N(w)\left[\cos 2\theta\cos\alpha - \sin 2\theta\sin\alpha\right] \\ &= \frac{1}{2}\int_0^T \sum_{i=1}^2\sum_{j=1}^2 \Lambda_{ij}^2(w,t)\mathrm{d}t + N(w)\cos\left(2\theta+\alpha\right) , \end{aligned}$$

using the definitions of α and $N(w)$ as given in Theorem 2.6, where $\alpha \in [-\pi,\pi)$. Clearly, $H(\theta)$ achieves its maximum value when $\cos(2\theta+\alpha) = 1$, and it is easily verified that choosing $\theta = -\alpha/2$ achieves this for a value $\theta \in [-\pi/2, \pi/2)$. This completes the proofs of Theorems 2.6 and 2.7.

Acknowledgment. Conversations with Nick Buchdahl and Nick Ouellette are gratefully acknowledged.

REFERENCES

[1] N. Agram and B. Oksendahl, *Stochastic control of memory mean-field processes*, Appl. Math. Optim., 79 (2019), pp. 181–204. (Cited on p. 782)

[2] D. Andersson and B. Djehiche, *A maximum principle for SDEs of mean-field type*, Appl. Math. Optim., 63 (2011), pp. 341–356. (Cited on p. 782)

[3] A. Apte, M. Hairer, A. Stewart, and J. Voss, *Sampling the posterior: An approach to non-Gaussian data assimilation*, Phys. D, 230 (2007), pp. 50–64. (Cited on pp. 782, 784, 785, 804)

[4] S. Balasuriya, *A tangential displacement theory for locating perturbed saddles and their manifolds*, SIAM J. Appl. Dyn. Syst., 10 (2011), pp. 1100–1126, https://doi.org/10.1137/100814640. (Cited on pp. 802, 807, 809)

[5] S. Balasuriya, *Barriers and Transport in Unsteady Flows: A Melnikov Approach*, Math. Model. Comput. 21, SIAM, Philadelphia, 2016, https://doi.org/10.1137/1.9781611974584. (Cited on pp. 782, 784, 790, 807, 809)

[6] S. Balasuriya, *Stochastic uncertainty of advected curves in finite-time unsteady flows*, Phys. Rev. E, 95 (2017), art. 062201. (Cited on pp. 782, 783, 801, 803, 809)

[7] S. Balasuriya, *Stochastic approaches to Lagrangian coherent structures*, Adv. Stud. Pure. Math., 85 (2020), pp. 95–104. (Cited on p. 804)

[8] S. Balasuriya, *Uncertainty in finite-time Lyapunov exponent calculations*, J. Comput. Dyn., 7 (2020), pp. 313–337. (Cited on pp. 801, 802)
[9] S. Balasuriya and G. Gottwald, *Estimating stable and unstable sets and their role as transport barriers in stochastic flows*, Phys. Rev. E, 98 (2018), art. 013106. (Cited on pp. 782, 783, 801, 803, 804)
[10] S. Balasuriya, R. Kalampattel, and N. Ouellette, *Hyperbolic neighborhoods as organizers of finite-time exponential stretching*, J. Fluid Mech., 807 (2016), pp. 509–545. (Cited on p. 803)
[11] S. Balasuriya, I. Mezić, and C. Jones, *Weak finite-time Melnikov theory and* 3*D viscous perturbations of Euler flows*, Phys. D, 176 (2003), pp. 82–106. (Cited on pp. 804, 807)
[12] S. Balasuriya, N. Ouellette, and I. Rypina, *Generalized Lagrangian coherent structures*, Phys. D, 372 (2018), pp. 31–51. (Cited on pp. 782, 784, 790, 799, 802, 803, 804)
[13] F. Balibrea-Iniesta, C. Lopesino, S. Wiggins, and A. Mancho, *Lagrangian descriptors for stochastic differential equations: A tool for revealing the phase portrait of stochastic dynamical systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 26 (2016), art. 1630036. (Cited on p. 801)
[14] I. Bashkirtseva and L. Ryashko, *Stochastic sensitivity of the closed invariant curves for discrete-time systems*, Phys. A, 410 (2014), pp. 236–243. (Cited on p. 783)
[15] N. Bergland and B. Gentz, *The effect of additive noise on dynamical hysteresis*, Nonlinearity, 15 (2002), pp. 605–632. (Cited on p. 782)
[16] N. Bergland and B. Gentz, *On the noise-induced passage through an unstable periodic orbit* II: *General case*, SIAM J. Math. Anal., 46 (2014), pp. 310–352, https://doi.org/10.1137/120887965. (Cited on p. 782)
[17] N. Bergland, B. Gentz, and C. Kuehn, *Hunting French ducks in a noisy environment*, J. Differential Equations, 252 (2012), pp. 4786–4841. (Cited on p. 782)
[18] P. Berloff and I. Kamenkovich, *On spectral analysis of mesoscale eddies. Part* II: *Nonlinear analysis*, J. Phys. Oceanography, 43 (2013), pp. 2528–2544. (Cited on pp. 791, 801)
[19] J. Berner, U. Achatz, L. Batte et al., *Stochastic parameterization: Toward a new view of weather and climate models*, Bull. Amer. Meteorol. Soc., 98 (2017), pp. 565–587. (Cited on p. 782)
[20] M. Branicki and K. Uda, *Lagrangian Uncertainty Quantification and Information Inequalities for Stochastic Flows*, preprint, https://arxiv.org/abs/1905.08707, 2019. (Cited on p. 782)
[21] M. Callaway, T. Doan, J. Lamb, and M. Rasmussen, *The dichotomy spectrum for random dynamical systems and pitchfork bifurcations with additive noise*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 1548–1574. (Cited on p. 782)
[22] A. Chiarini and M. Fischer, *On large deviations for small noise Itô processes*, Adv. Appl. Prob., 46 (2014), pp. 1126–1147. (Cited on pp. 782, 804)
[23] A. Churubini, J. Lamb, M. Rassmussen, and Y. Sato, *A random dynamical systems perspective on stochastic resonance*, Nonlinearity, 30 (2017), pp. 2835–2853. (Cited on p. 782)
[24] C. Cotter and G. Pavliotis, *Estimating eddy diffusivities from noisy Lagrangian observations*, Commun. Math. Sci., 7 (2009), pp. 805–838. (Cited on pp. 782, 784, 791)
[25] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Springer, New York, 2010. (Cited on pp. 782, 804)
[26] A. Denner, O. Junge, and D. Matthes, *Computing coherent sets using the Fokker–Planck equation*, J. Comput. Dyn., 3 (2016), pp. 163–177. (Cited on pp. 782, 801, 803)
[27] M. Dentz and F. de Barros, *Dispersion variance for transport in heterogeneous porous media*, Water Resources Res., 49 (2013), pp. 3443–3461. (Cited on pp. 802, 803)
[28] M. Dentz and D. Tartakovsky, *Probability density functions for passive scalars dispersed in random velocity fields*, Geophys. Res. Lett., 37 (2010), art. L24406. (Cited on pp. 786, 802, 803)
[29] L. Fang, S. Balasuriya, and N. Ouellette, *Disintangling resolution, precision and inherent stochasticity in nonlinear systems*, Phys. Rev. Res., 2 (2020), art. 023343. (Cited on pp. 802, 805)
[30] M. Farazmand and G. Haller, *Computing Lagrangian coherent structures from their variational theory*, Chaos, 22 (2012), art. 013128. (Cited on p. 799)
[31] F. Feppon and P. F. J. Lermusiaux, *Dynamically orthogonal numerical schemes for efficient stochastic advection and Lagrangian transport*, SIAM Rev., 60 (2018), pp. 595–625, https://doi.org/10.1137/16M1109394. (Cited on pp. 783, 792)
[32] E. Forgoston and R. O. Moore, *A primer on noise-induced transitions in applied dynamical systems*, SIAM Rev., 60 (2018), pp. 969–1009, https://doi.org/10.1137/17M1142028. (Cited on pp. 782, 785)

[33] M. Freidlin and A. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York, 1998. (Cited on pp. 782, 804)

[34] G. Froyland, *An analytical framework for identifying finite-time coherent sets in time-dependent dynamical systems*, Phys. D, 250 (2013), pp. 1–19. (Cited on pp. 799, 803, 804)

[35] G. Froyland, *Dynamic isoperimetry and the geometry of Lagrangian coherent structures*, Nonlinearity, 28 (2015), pp. 3587–3622. (Cited on pp. 784, 788, 799, 803, 804)

[36] G. Froyland and O. Junge, *On fast computation of finite-time coherent sets using radial basis functions*, Chaos, 25 (2015), art. 087409. (Cited on p. 799)

[37] G. Froyland and P. Koltai, *Estimating long-term behavior of periodically driven flows without trajectory integration*, Nonlinearity, 30 (2017), pp. 1948–1986. (Cited on p. 786)

[38] G. Froyland and K. Padberg-Gehle, *Finite-time entropy: A probabilistic approach for measuring nonlinear stretching*, Phys. D, 241 (2012), pp. 1612–1628. (Cited on pp. 803, 804)

[39] G. Froyland and K. Padberg-Gehle, *A rough-and-ready cluster-based approach for extracting finite-time coherent sets from sparse and incomplete trajectory data*, Chaos, 25 (2015), art. 087406. (Cited on pp. 799, 802)

[40] G. Froyland, N. Santitissadeekorn, and A. Monahan, *Transport in time-dependent dynamical systems: Finite-time coherent sets*, Chaos, 20 (2010), art. 043116. (Cited on pp. 784, 803)

[41] I. Grooms and L. Zanna, *A note on "Toward a stochastic parameterization of ocean mesoscale eddies,"* Ocean Model., 113 (2017), pp. 30–33. (Cited on p. 782)

[42] J. Gruendler, *The existence of homoclinic orbits and the method of Melnikov for systems in R^n*, SIAM J. Math. Anal., 16 (1985), pp. 907–931, https://doi.org/10.1137/0516069. (Cited on pp. 804, 807)

[43] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, New York, 1983. (Cited on pp. 807, 809)

[44] H. Guo, W. He, T. Peterka, H.-W. Shen, S. Collis, and J. Helmus, *Finite-time Lyapunov exponents and Lagrangian coherent structures in uncertain unsteady flows*, IEEE Trans. Visual. Comput. Graphics, 22 (2016), pp. 1672–1682. (Cited on p. 801)

[45] A. Hadjighasem, M. Farazmand, D. Blazevski, G. Froyland, and G. Haller, *A critical comparison of Lagrangian methods for coherent structure detection*, Chaos, 27 (2017), art. 053104. (Cited on pp. 782, 799, 802, 803)

[46] G. Haller, *A variational theory of hyperbolic Lagrangian coherent structures*, Phys. D, 240 (2011), pp. 574–598. (Cited on p. 803)

[47] G. Haller, *Lagrangian coherent structures*, Annu. Rev. Fluid Mech., 47 (2015), pp. 137–162. (Cited on pp. 782, 784, 790, 799, 802, 803)

[48] G. Haller and F. Beron-Vera, *Geodesic theory of transport barriers in two-dimensional flows*, Phys. D, 241 (2012), pp. 1680–1702. (Cited on p. 803)

[49] G. Haller, A. Hadjighasem, M. Farazmand, and F. Huhn, *Defining coherent vortices objectively from the vorticity*, J. Fluid Mech., 795 (2016), pp. 136–173. (Cited on p. 802)

[50] G. Haller, D. Karrasch, and F. Kogelbauer, *Material barriers to diffusive and stochastic transport*, Proc. Natl. Acad. Sci. USA, 115/37 (2018), pp. 9074–9079. (Cited on pp. 782, 801, 803)

[51] G. Kallianpur and P. Sundar, *Stochastic Analysis and Diffusion Processes*, Oxford University Press, Oxford, 2014. (Cited on pp. 805, 808, 811)

[52] I. Kamenkovich, I. Rypina, and P. Berloff, *Properties and origins of the anisotropic eddy-induced transport in the north Atlantic*, J. Phys. Oceanography, 45 (2015), pp. 778–791. (Cited on pp. 784, 791)

[53] P. Kloeden, E. Platen, and H. Schurz, *Numerical Solution of SDE through Computer Experiments*, Universitext, Springer-Verlag, Berlin, 1994, https://doi.org/10.1007/978-3-642-57913-4. (Cited on p. 792)

[54] T. Le Borgne, M. Dentz, and J. Carrera, *Lagrangian statistical model for transport in highly heterogeneous velocity fields*, Phys. Rev. Lett., 101 (2008), art. 090601. (Cited on p. 803)

[55] J. Maclean, N. Santitissadeekorn, and C. Jones, *A coherent structure approach for parameter estimation in Lagrangian data assimilation*, Phys. D, 360 (2017), pp. 36–45. (Cited on pp. 782, 784, 785, 804)

[56] B. Matkowsky, *Singular perturbations, stochastic differential equations, and applications*, in Singular Perturbations and Asymptotics, R. Meyer and S. Parter, eds., Academic Press, New York, 1980, pp. 109–147. (Cited on pp. 784, 786)

[57] P. Miron, F. Beron-Vera, M. Olascoaga, J. Sheinbaum, P. Pérez-Brunius, and G. Froyland, *Lagrangian dynamical geography of the Gulf of Mexico*, Sci. Rep., 7 (2017), art. 7021. (Cited on p. 803)

[58] L. Mitchell and A. Carrassi, *Accounting for model error due to unresolved scales within ensemble Kalman filtering*, Q. J. Roy. Meteorol. Soc., 141 (2015), pp. 1417–1428. (Cited on p. 804)

[59] A. Molcard, L. Piterbarg, A. Griffa, T. Ozgokmen, and A. Mariano, *Assimilation of drifter observations for the reconstruction of the Eulerian circulation field*, J. Geophys. Res. Oceans, 108 (2003), art. 3056. (Cited on pp. 782, 784, 785, 804)

[60] S. Monter and Y. Bakhtin, *Normal forms approach to diffusion near hyperbolic equilibria*, Nonlinearity, 24 (2011), pp. 1883–1907. (Cited on p. 782)

[61] Z. Pan and T. Başar, *Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion*, SIAM J. Control Optim., 37 (1999), pp. 957–995, https://doi.org/10.1137/S0363012996307059. (Cited on p. 782)

[62] T. Peacock and J. Dabiri, *Introduction to focus issue: Lagrangian coherent structures*, Chaos, 20 (2010), art. 017501. (Cited on pp. 782, 790)

[63] K. Priyankara, S. Balasuriya, and E. Bollt, *Quantifying the role of folding in nonautonomous flows: The unsteady double-gyre*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 27 (2017), art. 1750156. (Cited on p. 792)

[64] R. Samelson, *Lagrangian motion, coherent structures, and lines of persistent material strain*, Annu. Rev. Mar. Sci., 5 (2013), pp. 137–163. (Cited on pp. 782, 784, 790)

[65] K. Schlueter-Kuck and J. Dabiri, *Coherent structure coloring: Identification of coherent structures from sparse data using graph theory*, J. Fluid Mech., 811 (2017), pp. 468–486. (Cited on pp. 799, 802)

[66] K. Schlueter-Kuck and J. Dabiri, *Model parameter estimation using coherent structure colouring*, J. Fluid Mech., 861 (2018), pp. 886–900. (Cited on pp. 782, 784, 785, 804)

[67] Z. Schuss, *Singular perturbation methods in stochastic differential equations of mathematical physics*, SIAM Rev., 22 (1980), pp. 119–155, https://doi.org/10.1137/1022024. (Cited on pp. 784, 786)

[68] S. Shadden, *Lagrangian coherent structures*, in Transport and Mixing in Laminar Flows: From Microfluidics to Oceanic Currents, R. Grigoriev, ed., Wiley, 2011, pp. 59–89. (Cited on pp. 782, 790)

[69] S. Shadden, F. Lekien, and J. Marsden, *Definitions and properties of Lagrangian coherent structures from Lyapunov exponents in two-dimensional aperiodic flows*, Phys. D, 212 (2005), pp. 271–304. (Cited on pp. 783, 784, 792, 799, 802, 803)

[70] L. Slivinski, E. Spiller, A. Apte, and B. Sandstede, *A hybrid particle-ensemble Kalman filter for Lagrangian data assimilation*, Monthly Weather Rev., 143 (2015), pp. 195–211. (Cited on pp. 782, 784, 785, 804)

[71] M. Speetjens, G. Metcalfe, and M. Rudman, *Lagrangian Transport and Chaotic Advection in Three-Dimensional Laminar Flows*, preprint, https://arxiv.org/abs/1904.07580, 2019. (Cited on p. 782)

[72] K. Stewart, P. Spence, S. Waterman, J. L. Sommer, J.-M. Moines, J. Lilly, and M. England, *Anisotropy of eddy variability in the global ocean*, Ocean Model., 95 (2015), pp. 53–65. (Cited on pp. 791, 801)

[73] W. Tang and C. Luna, *Dependence of advection-diffusion-reaction on flow coherent structures*, Phys. Fluids, 25 (2013), art. 106602. (Cited on p. 803)

[74] W. Tang and P. Walker, *Finite-time statistics of scalar diffusion in Lagrangian coherent structures*, Phys. Rev. E, 86 (2012), art. 045201. (Cited on pp. 803, 804)

[75] J.-L. Thiffeault, *Using multiscale norms to quantify mixing and transport*, Nonlinearity, 25 (2012), pp. R1–R44. (Cited on p. 803)

[76] E. van Sebille, S. Waterman, A. Barthel, R. Punpkin, S. Keating, C. Fogwill, and C. Turney, *Pairwise surface drifter separation in the western Pacific section of the Southern Ocean*, J. Geophys. Res. Oceans, 120 (2015), pp. 6769–6781. (Cited on pp. 791, 801)

[77] M. Wechselberger, *Extending Melnikov theory to invariant manifolds on non-compact domains*, Dyn. Syst., 17 (2002), pp. 215–233. (Cited on pp. 804, 807)

[78] K. Yagasaki, *Melnikov processes and chaos in randomly perturbed dynamical systems*, Nonlinearity, 31 (2018), pp. 3057–3085. (Cited on pp. 804, 807)