Filtering a statistically exactly solvable test model for turbulent tracers from partial observations

B. Gershgorin*, A.J. Majda*

Department of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY, 10012

Abstract

A statistically exactly solvable model for passive tracers is introduced as a test model for the authors’ Nonlinear Extended Kalman Filter (NEKF) as well as other filtering algorithms. The model involves a Gaussian velocity field and a passive tracer governed by the advection-diffusion equation with an imposed mean gradient. The model has direct relevance to engineering problems such as the spread of pollutants in the air or contaminants in the water as well as climate change problems concerning the transport of greenhouse gases such as carbon dioxide with strongly intermittent probability distributions consistent with the actual observations of the atmosphere. One of the attractive properties of the model is the existence of the exact statistical solution. In particular, this unique feature of the model provides an opportunity to design and test fast and efficient algorithms for real-time data assimilation based on rigorous mathematical theory for a turbulence model problem with many active spatiotemporal scales. Here, we extensively study the performance of the NEKF which uses the exact first and second order nonlinear statistics without any approximations due to linearization. The role of partial and sparse observations, the frequency of observations and the observation noise strength in recovering the true signal, its spectrum, and fat tail probability distribution are the central issues discussed here. The results of our study provide useful guidelines for filtering realistic turbulent systems with passive tracers through partial observations.

Keywords: turbulent advection-diffusion equation, exactly solvable model, nonlinear Kalman filter, data assimilation

1. Introduction

Turbulent diffusion is a physical process that describes the transport of a tracer in a turbulent velocity field. Very often the tracer itself has very little or no influence on the...
background flow, in which case it is referred to as a passive tracer. Practically important examples include engineering problems such as the spread of hazardous plumes or pollutants in the atmosphere and contaminants in the ocean. Another class of problems that involve turbulent diffusion are climate science problems concerning the transport of greenhouse gases such as carbon dioxide and others. One of the characteristics of these systems is their complex multiscale structure in both time and space. For example, the spatial scales of atmospheric flows span from planetary scale Rossby waves to local weather patterns with the size of kilometers. Similarly, temporal scales involve both slow dynamics on the scales of decades as well as the fast dynamics on the scales of hours. Another remarkable property of many tracers in the atmosphere is their highly intermittent probability distributions with long exponential tails [1]. Many contemporary applications in science and engineering involve real time filtering of such turbulent non-Gaussian signals from nature with multiple scales in time and space. Filtering combines partially observed features of the true signal together with a dynamic model to obtain a statistical estimate for the state of the system. Such filtering algorithms are based on the classical Kalman filter [2–13]. A contemporary view on filtering complex turbulent signals from nature is given in an overview paper [14]. Important examples involve the real time filtering in weather and climate where data assimilation can play an important role in recalibrating the current climate models and improving them “on the fly” through empirical information theory [15, 16]. Real time tracking of a chemical plume released into the atmosphere or a contaminant injected into the ocean is another extremely important and practical example where real time data assimilation plays a crucial role. Major difficulties in accurate filtering and prediction of noisy turbulent spatially extended signals are typically caused by imperfect partial and sparse observations and model error due to inadequate numerical resolution or incomplete physical understanding. One way to evaluate the consequences of such complications and to build accurate and efficient filtering algorithms is to design a mathematically stringent test model that has the features of the realistic physical system but on the other hand allows for exact statistical solution. Such model not only would help us to understand the underlying physics but also could be used for studying the role of observations and model error such as the use of eddy diffusivity models in data assimilation.

In this paper, we introduce a test model for filtering turbulent tracers which has direct relevance to actual physical systems with the additional attractive feature of exactly solvable statistics for the mean and covariance [17–19] despite the inherent statistical nonlinearity. The model for the velocity field is incompressible two-dimensional with a time dependent cross-sweep, $U(t)$, in one direction and time and spatially dependent shear flow, $v(x, t)$, in the transverse direction. In the atmosphere it would correspond to the zonal jet (cross-sweep in the east-west direction) and the transverse Rossby waves (shear flow in the north-south direction), however different interpretations would arise in other physical contexts. We model the velocity field by a combination of the prescribed deterministic and Gaussian stochastic components. Such simplified representation of the velocity field exhibits some empirical features of turbulent flows which are given by the Navier-Stokes equation, however, they are more manageable mathematically [20]. Transport of a passive tracer, $T$, by the fluid flow with velocity $\vec{v} = (U, v)^T$, can be
modeled by the advection-diffusion equation [20, 21].

\[ \frac{\partial T}{\partial t} + \vec{v} \cdot \vec{\nabla} T = \kappa \Delta_{\vec{x}} T - \gamma T + S(\vec{x}, t), \]  

(1)

where \( \kappa \) is molecular or eddy diffusivity, \( \gamma \) is a uniform damping, and \( S \) is a source for the tracer. Typically the statistical solution for (1) cannot be obtained in a closed form because of the nonlinear advection term \( \vec{v} \cdot \vec{\nabla} T \). This is a manifestation of the “turbulence moment closure problem”. However, in certain simplified situations the exact analytical treatment is possible. For example, in this paper we consider a one dimensional version of the system when the tracer has a prescribed mean gradient in one direction and the one dimensional fluctuations of the tracer around this mean gradient. The atmospheric analogy would include the north-south mean gradient of a tracer gas and zonal fluctuations restricted to a narrow band in the midlatitudes. Despite all the simplifications, this test model carries such important characteristics of a realistic physical system as the turbulent energy cascade of the tracer and strongly intermittent probability distributions for the tracer with long exponential tails that are observed for tracers in the real atmosphere [1, 21].

Following [17, 18, 22–24], a Nonlinear Extended Kalman Filter (NEKF) is designed for this model utilizing exact statistics. We test the performance of the NEKF on system (1) in various dynamical regimes. In particular, we are interested in the case of nondispersive waves, which are often relevant in engineering applications, and dispersive waves such as Rossby waves that mimic the turbulent field in the atmosphere. We study how the type and sparseness of observations impact the filtering skill and how one can improve the performance of the NEKF by adding just one observation on the spatially extended grid. We find that even in very tough regimes when only few observations are available, the NEKF has excellent skill in recovering the spatial structure of the original truth signal as well as its spectral properties and the fat-tailed tracer probability distribution. We also advocate here that the model in (1) has a lot of potential in studying the subtle issues of model error due to eddy diffusivity parameterization and other approximations.

The paper is organized as follows. In Section 2, we provide a detailed description of the test model with a general Gaussian velocity field and a tracer. In Section 3, we show how to compute exactly the first and second order nonlinear statistics of the tracer for a general prescribed Gaussian velocity field. Then, we propose a particular example of a velocity field that is modeled using an Ornstein-Uhlenbeck process and discuss the choice of parameters. For this particular model, we study the properties of the system, such as the correlation time of different modes, the energy spectrum, and pdfs with long exponential tails. Moreover, we obtain a statistically exact equation for eddy diffusivity which makes the model very attractive for studying model error. In Section 4, we give a brief description of the classical Kalman filter as well as the procedure of constructing the NEKF [17, 18] based on the exact statistics obtained in Section 2. We describe how sparse and partial observations can be incorporated in the NEKF. In Section 5, we present the numerical results for filtering, including examples of filtering particular trajectories, recovery of the turbulent spectra and the pdf of the tracer. Furthermore, we study how the filter performance is improved by adding one observation, and discuss the dependence of the filter skill on the observation noise variance, observation time, sparsity of observations and the types of observations used in the NEKF. We end the paper with concluding discussion in Section 6.
2. Model Description

We consider a model for a passive tracer \( T(\vec{x}, t) \) which is advected by a velocity field \( \vec{v}(\vec{x}, t) \). In general, the dynamics of such tracer can be described by the following advection-diffusion equation [20]

\[
\frac{\partial T}{\partial t} + \vec{v} \cdot \nabla_x T = \kappa \Delta_x T - \gamma T + S(\vec{x}, t),
\]  

(2)

where \( \kappa \) is molecular diffusivity, \( \gamma \) represents the uniform damping, \( S(\vec{x}, t) \) is an external source for a tracer, and the velocity field is considered to be incompressible, \( \nabla_x \vec{v} = 0 \). We assume that the model for the velocity field is two dimensional and periodic in space

\[
\vec{v}(\vec{x}, t) = (U(t), v(x, t))^T,
\]  

(3)

which automatically satisfies the incompressibility condition. Moreover, each component, the cross-sweep, \( U(t) \), and the shear flow, \( v(x, t) \), are Gaussian fields with prescribed dynamics. Suppose, that a tracer has a known mean gradient

\[
T(\vec{x}, t) = \alpha_1 x + \alpha_2 y + T'(x, y, t),
\]  

(4)

where \( T'(x, y, t) \) denotes fluctuations around the mean gradient. Then, from (2), we find an equation for \( T'(x, y, t) \)

\[
\frac{\partial T'(x, y, t)}{\partial t} + U(t) \frac{\partial T'(x, y, t)}{\partial x} + v(x, t) \frac{\partial T'(x, y, t)}{\partial y} = \kappa \left( \frac{\partial^2 T'(x, y, t)}{\partial x^2} + \frac{\partial^2 T'(x, y, t)}{\partial y^2} \right) - \gamma(t) T'(x, y, t) - \alpha_1 U(t) - \alpha_2 v(x, t) + S(x, y, t).
\]  

(5)

It often appears that the fluctuations \( T' \) of the tracer only depend on the \( x \) spatial variable. In this situation (5) simplifies to a one-dimensional ODE

\[
\frac{\partial T'(x, t)}{\partial t} + U(t) \frac{\partial T'(x, t)}{\partial x} = -\alpha v(x, t) + \kappa \frac{\partial^2 T'(x, t)}{\partial x^2} - \gamma T'(x, t),
\]  

(6)

where for simplicity we also assumed that \( \alpha_1 \equiv 0 \), \( S(x, y, t) \equiv 0 \) and \( \alpha \equiv \alpha_2 \). In the remainder of the paper, we drop the prime superscript in denoting (6). For a given velocity field, (6) is solved using Fourier transformation. Moreover, as we also show below the statistics of the system \( (U(t), v(x, t), T(x, t))^T \) can be computed exactly for a Gaussian velocity field despite the nonlinearity from the advection term, \( U(t) \frac{\partial T'(x, t)}{\partial x} \). This remarkable feature makes the model extremely attractive for filtering via NÉKF, which uses exact statistics to predict the next state of the system. Next, we show how to compute the first and second order statistics of this one-dimensional model for a general linear Gaussian velocity field, followed by a specific example that has direct relevance to the atmospheric flows. Furthermore, in Section 3.4, we introduce a two-dimensional version of (5) with a deterministic cross-sweep, \( U(t) \), which is also exactly statistically solvable.
3. System statistics

3.1. Tracer statistics for a general linear Gaussian velocity field

Here we show how to find the statistics of the tracer for a general Gaussian velocity field (3). It is convenient to rewrite (6) in Fourier space

\[
\frac{\partial T_k(t)}{\partial t} + ikU(t)T_k(t) = -\kappa k^2 T_k(t) - \gamma T_k(t) - \alpha v_k(t), \quad \text{for } k \in [0, K]
\]

(7)

where the subscript \(k\) denotes the \(k\)th Fourier mode of a corresponding variable. Moreover, unless stated otherwise, we will refer to the fluctuations of the tracer around the mean gradient as the “tracer”. From (7), we can find that the tracer \(T_k(t)\) is nonlinearly dependent on the Gaussian cross-sweep, \(U(t)\). Therefore, the statistics of the tracer are non-Gaussian. Nevertheless, the special structure of the governing equation in (7) allows for exact analytical formulas for the first and second (and in principle, any order) moments. Each Fourier mode \(T_k(t)\) acts as a nonlinear forced-damped oscillator with a random frequency [17, 18, 22, 23, 25]. The path-wise solution of (7) becomes

\[
T_k(t) = D_k(t_0, t)T_k(t_0)e^{-ikJ_U(t_0, t)} - \alpha \int_{t_0}^{t} D_k(s, t)v_k(s)e^{-ikJ_U(s, t)} ds,
\]

(8)

where

\[
D_k(s, t) = e^{-(\gamma + \kappa k^2)(t-s)},
\]

(9)

\[
J_U(s, t) = \int_{s}^{t} U(s')ds'.
\]

(10)

Note that \(D_k(s, t)\) is deterministic while \(J_U(s, t)\) is Gaussian. We also assume a Gaussian initial condition \(T_k(t_0)\) for the tracer. Then, we compute the mean \(\langle T_k(t) \rangle\)

\[
\langle T_k(t) \rangle = D_k(t_0, t)\langle T_k(t_0) \rangle e^{-ikJ_U(t_0, t)} - \alpha \int_{t_0}^{t} D_k(s, t)\langle v_k(s) \rangle e^{-ikJ_U(s, t)} ds.
\]

(11)

Using properties of the characteristic function of a Gaussian random variable, we obtain the following equality for any complex Gaussian \(z\) and any real Gaussian \(x\) [17, 18, 26, 27]

\[
\langle ze^{ix} \rangle = \left( \langle z \rangle + iCov(z, x) \right)e^{i\langle x \rangle + \frac{1}{2}Var(x)}.
\]

(12)

Applying (12) to (11), we find

\[
\langle T_k(t) \rangle = D_k(t_0, t)\left(\langle T_k(t_0) \rangle - ikCov(T_k(t_0), J_U(t_0, t)) \right)e^{-ik\langle J_U(t_0, t) \rangle - \frac{k^2}{2}Var(J_U(t_0, t))}
\]

\[
-\alpha \int_{t_0}^{t} D_k(s, t)\left(\langle v_k(s) \rangle - ikCov(v_k(s), J_U(s, t)) \right)e^{-ik\langle J_U(s, t) \rangle - \frac{k^2}{2}Var(J_U(s, t))} ds.
\]

(13)

Note that the right-hand side of (13) only depends on the statistics of the prescribed Gaussian velocity field (3) and initial value of the tracer at time \(t_0\). Moreover, as we will demonstrate on a particular example below, the statistics of the velocity field at any time are also given by their initial values. Therefore, (13) gives an exact analytical formula for the mean \(\langle T_k(t) \rangle\) as a functional of the Gaussian initial statistics of the velocity field and the tracer \((U(t_0), v_k(t_0), T_k(t_0))^T\). In Appendix A, we derive similar formulas for the second order statistics of the tracer. Next, we consider a particular choice of the Gaussian velocity field (3) and obtain specific formulas for its mean and covariance.
3.2. A particular choice of the Gaussian velocity field and its statistics

Suppose that the cross-sweep, \( U(t) \), is given by the real part of a complex Ornstein-Uhlenbeck (OU) process

\[
U(t) = \text{Re}[V(t)],
\]

\[
\frac{dV(t)}{dt} = \lambda_U V(t) + f_U(t) + \sigma_U \dot{W}_U(t),
\]

where \( \lambda_U = -\gamma_U + i\omega_U \), \( f_U(t) \) is deterministic forcing which for simplicity is considered to be periodic

\[
f_U(t) = A_f e^{i(\eta t + \phi_f)} + B_f,
\]

and \( \dot{W}_U \) is a complex white noise process [5, 22, 23]. Note that by choosing the cross-sweep to be a real part of the complex OU process as opposed to the real OU process, we allow it to have highly oscillatory correlation functions, which provides a very interesting test case with negative effective diffusivity [20]. Next, the dynamical equation for the velocity component \( v(x, t) \) is

\[
\frac{dv(x, t)}{dt} = P\left( \frac{\partial}{\partial x} \right) v(x, t) + f_v(x, t) + \sigma_v(x) \dot{W}_v.
\]

Here again the function \( f_v(x, t) \) is a known deterministic functions, while \( \dot{W}_v \) represents random white noise fluctuations in forcing arising from hidden nonlinear interactions and other processes [11, 14]. In Fourier space, operator \( P\left( \frac{\partial}{\partial x} \right) \) is given by the symbol \( \hat{P}_k \) that has the form

\[
\hat{P}_k = -\gamma_k + i\omega_k = \lambda_k.
\]

We are looking for the solution of our model that consists of a finite number of Fourier modes \( k \in [0, 2K] \) which is equivalent to the discretization of the \( 2\pi \)-periodic spatial domain into \( 2K + 1 \) points, \( x_j = jh \), with \( h = 2\pi/(2K + 1) \), \( j \in [1, 2K + 1] \). Then, each Fourier component \( v_k(t) \) is governed by the equation

\[
\frac{dv_k(t)}{dt} = \lambda_k v_k(t) + f_k(t) + \sigma_k \dot{W}_k(t) \quad \text{for } k \in [0, K],
\]

where \( f_k(t) \) is Fourier decomposition of \( f_v(x, t) \), the white noise of each Fourier mode \( v_k(t) \) is independent complex white noise [5, 22, 23]. The reality condition is expressed by

\[
v_{-k} = v_k^*,
\]

where the asterisk "*" denotes complex conjugate. In this paper, we discuss and compare two prototype models for the waves \( v_k(t) \):

- nondispersive waves with

\[
\omega_k = -ck, \quad \gamma_k = d_v + \mu k^2,
\]

where \( d_v \) represents the uniform part of the dissipation and \( \mu \) is the strength of the selective part of the dissipation.
• dispersive Rossby waves with

\[ \omega_k = \frac{\beta k}{k^2 + F_s}, \quad \gamma_k = \nu (k^2 + F_s), \]  

(20)

where \( \beta \) represents rotation and \( F_s \) represents stratification.

Note that the case of nondispersive waves may be more interesting for the engineering community while the case of Rossby waves has more implications for atmosphere-ocean science modeling. The strength of the white noise forcing for each mode \( v_k \) is given by \( \sigma_k \) which can be found if the energy spectrum \( E_k \) is prescribed [3, 11, 14]. The energy of a wave \( v_k(t) \) is given by its equilibrium variance

\[ E_k \equiv \text{Var}_{eq}(v_k) = \frac{\sigma_k^2}{2\gamma_k}, \]  

(21)

Therefore, we find

\[ \sigma_k = \sqrt{2\gamma_k E_k}. \]  

(22)

Here, we consider the energy spectrum for \( v_k \) with the following form

\[ E_k = \begin{cases} E, & |k| \leq k_0, \\ E \left( \frac{k}{k_0} \right)^{-5/3}, & |k| > k_0, \end{cases} \]  

(23)

where we chose the equipartition of energy spectrum to mimic energetic large scale waves and a Kolmogorov spectrum for smaller scales. In Fig. 1(a), we illustrate such an energy spectrum for \( v_k \). Note that we can as well choose any other form of a spectrum for the waves \( v_k(t) \) since it is an input parameter in our model and this flexibility is useful for other applications. For the external forcing \( f_k(t) \), we choose an oscillating function

\[ f_k(t) = A_k e^{i(\xi_k t + \phi_k)}. \]  

(24)

Now, we demonstrate the statistical solution for the specific Gaussian velocity field in (14) and (16). The velocity field is given by Gaussian random processes (14) which is uniquely defined by its mean and covariance. Note that these statistics together with the first and second order statistics of the tracer given by (13) and the formulas in the Appendix will be used below to build the NEKF for this system. Therefore, we will assume that all three components of the system \((U(t), v_k(t), T_k(t))^T\) are correlated at some initial time \( t_0 \) which always happens after a data assimilation time step. We start with the cross-sweep \( U(t) \). From (14), we find

\[ V(t) = e^{\lambda_U(t-t_0)}V(t_0) + \int_{t_0}^t e^{\lambda_U(t-s)} f_U(s) ds + \sigma_U \int_{t_0}^t e^{\lambda_U(t-s)} dW_U(s), \]  

\[ U(t) = \text{Re}[V(t)]. \]  

(25)

It is a simple exercise to show that the path-wise solution in (25) with the special form of forcing in (15) is equivalent to

\[ U(t) = \bar{U}(t) + \text{Re}[V'(t)], \]  

\[ \bar{U}(t) = U_0 + A_U \sin(\eta t), \]  

(26)
where \( U_0 \) and \( A_U \) are some constants and \( V'(t) \) solves the unforced Langevin equation
\[
\frac{dV'(t)}{dt} = \lambda_U V'(t) + \sigma_U W_U(t),
\]
with the same parameters as in (14). The form of the cross-sweep, \( U(t) \), that is given in (26) is much simpler to use in our further calculations, therefore, we will assume that the constants \( U_0, A_U, \) and \( \eta \) are given parameters of our model and (26) will be used as the definition of the cross-sweep although a more physically relevant but equivalent definition is in (14). The first and second order statistics of \( V'(t) \) become
\[
\langle V'(t) \rangle = e^{\lambda_U (t-t_0)} \langle V'(t_0) \rangle, \tag{28}
\]
\[
\text{Var}(V'(t)) = e^{-2\gamma_U (t-t_0)} \text{Var}(V'(t_0)) + \frac{\sigma_U^2}{2\gamma_U} (1 - e^{-2\gamma_U (t-t_0)}), \tag{29}
\]
\[
\text{Cov}(V'(t), V'(t)^*) = e^{2\lambda_U (t-t_0)} \text{Cov}(V'(t_0), V'(t_0)^*), \tag{30}
\]
where for the cross-covariance, we used the definition \( \text{Cov}(a, b) = \langle ab^* \rangle - \langle a \rangle \langle b \rangle^* \).

Next, for Gaussian waves \( v_k(t) \), we find the path-wise solution
\[
v_k(t) = e^{\lambda_k (t-t_0)} v_k(t_0) + F_k(t_0, t) + \sigma_k \int_{t_0}^{t} e^{\lambda_k (t-s)} dW_k(s), \tag{31}
\]
where
\[
F_k(t_0, t) = \int_{t_0}^{t} e^{\lambda_k (t-s)} f_k(s) ds. \tag{32}
\]
The first and second order of \( v_k(t) \) statistics become
\[
\langle v_k(t) \rangle = e^{\lambda_k (t-t_0)} \langle v_k(t_0) \rangle + F_k(t_0, t), \tag{33}
\]
\[
\text{Cov}(v_k(t), v_l(t)) = e^{(\lambda_k + \lambda_l) (t-t_0)} \left( \text{Cov}(v_k(t_0), v_l(t_0)) + \frac{\sigma_k^2}{\gamma_k} \delta_{kl} (e^{2\gamma_k (t-t_0)} - 1) \right). \tag{34}
\]

We note that the components of the velocity field \( U(t) \) and \( v_k(t) \) are governed by independent equations, however, they can be correlated provided that there is correlation between them at the initial moment \( t_0 \)
\[
\text{Cov}(V'(t), v_k(t)^*) = e^{(\lambda_U + \lambda_k) (t-t_0)} \text{Cov}(V'(t_0), v_k(t_0)^*). \tag{35}
\]

Now, with a specific form of the Gaussian velocity field, we can specify the mean and variance of \( J_U \) from (10) as well as other statistics from (13) for the mean of the tracer, \( \langle T_k(t) \rangle \). Using (26), we decompose \( J_U \) into deterministic and stochastic parts
\[
J_U(s, t) = \bar{J}(s, t) + \tilde{J}(s, t), \tag{36}
\]
\[
\bar{J}(s, t) = \int_{s}^{t} \tilde{U}(s') ds' = (t-s) U_0 - \frac{A_U}{\eta} (\cos(\eta t) - \cos(\eta s)), \tag{37}
\]
\[
J(s, t) = \int_{s}^{t} \text{Re}[V'(s')] ds'. \tag{38}
\]
Then the mean of the tracer (13) becomes

\[
(T_k(t)) = D_k(t_0, t) e^{-ikJ(t_0, t)} \left( (T_k(t_0)) - ik \text{Cov}(T_k(t_0), J(t_0, t)) \right)
\]

\[
\times e^{-ik(J(t_0, t)) - \frac{4}{\gamma} \text{Var}(J(t_0, t))}
\]

\[
- \alpha \int_0^t D_k(s, t) e^{-ikJ(s, t)} e^{\lambda s(s-t_0)} \left( (v_k(t_0)) - ik \text{Cov}(v_k(t_0), J(s, t)) \right)
\]

\[
\times e^{-ik(J(s, t)) - \frac{4}{\gamma} \text{Var}(J(s, t))} ds
\]

\[
- \alpha \int_0^t D_k(s, t) e^{-ikJ(s, t)} F_k(t_0, s) e^{-ik(J(s, t)) - \frac{4}{\gamma} \text{Var}(J(s, t))} ds,
\]  

(39)

where

\[
(J(s, t)) = \text{Re} \left[ (V'(t_0)) \frac{1}{\lambda U} \left( e^{\lambda U (t-t_0)} - e^{\lambda U (s-t_0)} \right) \right]
\]

\[
\text{Var}(J(s, t)) = \text{Var}(V'(t_0)) \frac{\gamma}{\gamma_U + \omega_U^2} e^{-\gamma_U (s+t-2t_0)} \left( \cosh(\gamma_U (s-t)) - \cos(\omega_U (s-t)) \right)
\]

\[
+ \frac{1}{2} \text{Re} \left[ \text{Cov}(V'(t_0)) \left( \frac{e^{\lambda U (s-t_0)} - e^{\lambda U (t-t_0)}}{\lambda_U} \right)^2 \right]
\]

\[
+ \frac{\sigma^2_U}{4} \left[ 2 \text{Re} \left( \frac{e^{\lambda U (t-s)} - 1}{\lambda_U} \right) \right]
\]

\[
+ e^{-2\gamma_U (t-t_0)} \left( e^{\lambda U (s-t)} - \gamma_U (s-t) \right) + e^{-2\gamma_U (s-t_0)} \left( e^{\lambda U (t-s)} - \gamma_U (t-s) \right) + 2\gamma_U (t-s)
\]

\[
, \gamma_U + \omega_U^2
\]

\[
\text{Cov}(T_k(t_0), J(s, t)) = \frac{1}{2} \left( \text{Cov}(V'(t_0), T_k(t_0)) e^{\lambda U (t-t_0)} - e^{\lambda U (s-t_0)} \right)
\]

\[
+ \text{Cov}(V'(t_0), T_k(t_0)) e^{\lambda U (s-t_0)} \left( \frac{e^{\lambda U (t-t_0)} - e^{\lambda U (s-t_0)}}{\lambda_U} \right),
\]

\[
\text{Cov}(v_k(t_0), J(s, t)) = \frac{1}{2} \left( \text{Cov}(V'(t_0), v_k(t_0)) e^{\lambda U (t-t_0)} - e^{\lambda U (s-t_0)} \right)
\]

\[
+ \text{Cov}(V'(t_0), v_k(t_0)) e^{\lambda U (s-t_0)} \left( \frac{e^{\lambda U (t-t_0)} - e^{\lambda U (s-t_0)}}{\lambda_U} \right).
\]  

(40)

3.3. Closed equation for the eddy diffusivity

Parameterization by eddy diffusivity is often used in practice, in engineering and atmosphere-ocean science to account for unresolved scales [20]. In this situation, one parameterizes the collective effect of the nonlinear interaction of the turbulent velocity field on smaller scales on the tracer as effective or eddy diffusivity. Here, we unambiguously demonstrate the accuracy of such an eddy diffusivity approximation in the simplified model.

Let us obtain a differential equation for the mean \((T_k(t))\) by differentiating (13).

\[
\frac{\partial (T_k(t))}{\partial t} + ik \langle U(t) \rangle \langle T_k(t) \rangle = -\kappa k^2 \langle T_k(t) \rangle - \gamma \langle T_k(t) \rangle - \alpha \langle v_k(t) \rangle + G_k(t_0, t),
\]  

(41)
where

$$G_k(t_0, t) = -\frac{k^2}{2} D_k(t_0, t) \left( T_k(t_0) \right) \frac{\partial}{\partial t} \left( \text{Var}(J_U(t_0, t)) \right) e^{-ik(J_U(t_0, t))} - \frac{k^2}{2} \text{Var}(J_U(t_0, t))$$

and vanishing initial correlations between the tracer and the velocity field were assumed at the initial time \( t_0 \) for simplicity. Note that (41) is an exact ODE for the mean tracer for the \( k \)th Fourier mode. The last term in (41), \( G_k(t_0, t) \), is actually an enhancement to the diffusion term with a diffusion coefficient given by a kernel

$$\kappa(s, t) = \frac{1}{2} \frac{\partial}{\partial t} \left( \text{Var}(J_U(s, t)) \right).$$

Unlike standard diffusion, this kernel has memory in time. An extremely interesting question is how well can this non-local in time diffusion be approximated by a standard local in time diffusivity? To answer this question, we approximate the first term in the integrand by some constant

$$\kappa_{eddy}(t) \approx \frac{1}{2} \frac{\partial}{\partial t} \left( \text{Var}(J_U(s, t)) \right),$$

then (41) becomes

$$\frac{\partial (T_k(t))}{\partial t} + ik \langle U(t) \rangle T_k(t) \approx (\kappa + \kappa_{eddy}) k^2 (T_k(t)) - \gamma \langle T_k(t) \rangle - \alpha \langle v_k(t) \rangle.$$ (45)

Here, this constant, \( \kappa_{eddy} \), exactly represents the eddy diffusivity that enhances the diffusivity of the system due to smaller scale nonlinear interactions. The model error due to eddy diffusivity approximation can be quantified by comparing the exact mean given by (39) and its approximation given by the solution of a linear ODE (45). The important practical issue of model error due to eddy diffusivity approximation in the filtering context can be addressed unambiguously using this test case. This is an active research area which the authors intend to address in future work.

### 3.4. Two-dimensional problem with deterministic cross-sweep

Now, we consider the more general case of the two-dimensional tracer. We return to the general equation for the tracer in the mean gradient given by (5). Suppose, the cross-sweep is given by a deterministic function \( \bar{U}(t) \), e.g. as in (26). We also assume that the molecular diffusion only exists in the \( y \)-direction. Then, the equation for the tracer, \( T'(x, y, t) \), becomes

$$\frac{\partial T'(x, y, t)}{\partial t} + \bar{U}(t) \frac{\partial T'(x, y, t)}{\partial x} + v(x, t) \frac{\partial T'(x, y, t)}{\partial y} = \kappa \frac{\partial^2 T'(x, y, t)}{\partial y^2} - \gamma(t) T'(x, y, t) - \alpha_2 v(x, t) - \alpha_1 \bar{U}(t) + S(x, y, t).$$ (46)

Next, for simplicity, we assume that \( \alpha_1 = 0 \) and \( S(x, y, t) \equiv 0 \) and perform a Fourier transform in the \( y \)-direction

$$\frac{\partial \hat{T}(x, t)}{\partial t} + \bar{U}(t) \frac{\partial \hat{T}(x, t)}{\partial x} = a_1(x, t) \hat{T}(x, t) + b_1(x, t).$$ (47)
where \( a_l(x,t) = -ilv(x,t) - \kappa l^2 - \gamma(t) \) and \( b_l(x,t) = -\alpha_2 v(x,t) \). We solve (47) by the method of characteristics. We define the characteristics as the solutions of the ODE

\[
\frac{dx(t)}{dt} = \bar{U}(t),
\]

with an initial condition \( x(t_0) = x_0 \). Then, the characteristics have the form

\[
x(t) = x_0 + \int_{t_0}^{t} \bar{U}(s)ds.
\]

Along these characteristics, (47) becomes an ODE

\[
\frac{dT_l(x(t),t)}{dt} = a_l(x(t),t)T_l(x(t),t) + b_l(x(t),t)
\]

with the solution

\[
T_l(x(t),t) = \exp\left(\int_{t_0}^{t} a_l(x(s),s)ds\right)T_l(x_0,t_0)
\]

\[
+ \int_{t_0}^{t} \exp\left(\int_{s}^{t} a_l(x(s'),s')ds'\right)b_l(x(s),s)ds.
\]

Next, we utilize the equation for the characteristics in order to find the solution at any given point, \( x \), and time, \( t \). We find

\[
x_0 = x - \int_{t_0}^{t} \bar{U}(s)ds,
\]

\[
x(s) = x - \int_{s}^{t} \bar{U}(s')ds'.
\]

Next, we substitute these expressions back to (51) to obtain the solution for each vertical mode of the scalar.

We note that this two-dimensional case of the turbulent tracers is also statistically exactly solvable using similar techniques that were used in Section 3 for the one-dimensional case. We present this two-dimensional case here for the purpose of demonstration that the test models for filtering turbulent tracers do not have to be only one dimensional. In fact, it is even possible to generalize the current version of the two dimensional case to have random component in the cross-sweep \( U(t) \) using the method of random characteristics. However, for the rest of the paper, we will use the one-dimensional system from Section 3 as a test model for filtering.

3.5. Properties of the model

In this Section, we describe the general statistical properties of the model (14), (16), (6), such as the energy spectrum, correlation time in equilibrium, and probability distributions. We focus on two different examples of the waves \( v_k(t) \), i.e., non-dispersive waves given by (19) and dispersive Rossby waves given by (20). We use the following physical intuition to choose the parameters for the model
the zonal jet $U(t)$ has positive mean and small enough standard deviation that ensures an eastward direction of the jet.

- the correlation time of the jet $U(t)$ is of the same order as the correlation time of the large scale waves $v_k(t)$ and larger than the correlation time of the tracer $T_k(t)$.

- the dissipation parameters of both the waves and the tracer are chosen relatively small to mimic a realistic fully turbulent system.

- the mean gradient is large enough to make the energy spectrum of the tracer be of the same order as the energy spectrum of the waves, i.e., the tracer is strongly forced by the waves through the mean gradient; moreover, in the strong mean gradient regime, the tracer becomes strongly non-Gaussian with long exponential tails in the pdfs which very well mimics the observed statistics of various tracers in the atmosphere [1].

In Table 1, we give most of the parameters for the model (14), (16), (6). Note that for the forcing $f_k(t)$ in (24), we used the following parameters: $A_k = 0.5 + 0.2\text{Uniform}(-0.5, 0.5)$, $\xi_k = 1.5 + 0.5\text{Uniform}(-0.5, 0.5)$, $\phi_k = \text{Uniform}(0, 2\pi)$, where Uniform$(a, b)$ is the uniform distribution in the interval $(a, b)$. It is important to emphasize that although these parameters are chosen randomly from these distributions, we fix them for each test trajectory and consider constants. Moreover, for in the numerical simulations the number of different Fourier modes for the waves, $v_k(t)$, and the tracer, $T_k(t)$, was either $K = 52$ which corresponds to $2K + 1 = 105$ spatial locations or $K = 10$ which corresponds to $2K + 1 = 21$ spatial locations. In the former case we define the energy spectrum of $v_k(t)$ in (23) with $k_0 = 10$ and in the latter case, we use $k_0 = 5$.

In Fig. 1, we show the spectra of the waves, $v_k$, and the spectra of the tracer for non-dispersive and Rossby waves. The spectrum is given by the equilibrium variance of the corresponding variable. The spectrum of the waves is prescribed by (23) and is independent of time. On the other hand, the spectrum of the tracer is set by the parameters of the model and the spectrum of the waves. Here, the spectrum of the tracer is time dependent because of the nonlinear dependence of the tracer on the time dependent cross-sweep, $U(t)$. We compute a long time average of the spectrum of the tracer, $T_k(t)$, in equilibrium and in the Appendix, we provide analytical formulas for $\text{Var}(T_k(t))$ in equilibrium. The spectrum of the tracer for the case of non-dispersive waves (Fig. 1(b)) turns out to have three distinct parts where it has approximately power law behavior with different exponents. The first two parts correspond to the first 10 modes, where the velocity field has constant spectrum (Fig. 1(a)). The third part has a steep slope in the log-log scale with the exponent approximately $-9/2$. The spectrum of the tracer for the case of Rossby waves (Fig. 1(c)) appears to have two distinct parts. The first part only spans over the 10 Fourier modes and has the exponent of approximately $-1$ which is the Batchelor spectrum for a tracer slowly forced at large scales [20]. The second part of the spectrum of the tracer is much steeper with the same exponent $-4.5$. Although analytical formulas for the spectra of the tracer are available (in the Appendix), the particular form and properties of the spectra utilizing asymptotics of these formulas will be discussed elsewhere by the authors. One of the important tests and practical uses for a filter for turbulent systems is to recover the spectrum of the original truth signal. Below we will use this test to study how well the filter performs in various situations.
In Fig. 2, we demonstrate the equilibrium autocorrelation functions of the cross-sweep, $U(t)$, and the first six Fourier modes of the waves, $v_k(t)$, and the tracer, $T_k(t)$ for the case of nondispersive waves given by (19). We used the following definition of the equilibrium autocorrelation function of a variable $u$

$$C_u(\tau) \equiv \frac{1}{\text{Var}(u(t))} \langle \hat{u}(t + \tau)\hat{u}^*(t) \rangle,$$

(54)

where the averaging is taken over an ensemble of trajectories in equilibrium, and the tilde denotes the variable with its mean removed. The autocorrelation functions of the Gaussian variables $U(t)$ and $v_k(t)$ were computed analytically from (14) and (17)

$$C_U(\tau) = \cos(\omega_U \tau)e^{-\gamma_U \tau},$$

$$C_{v_k} = e^{(-\gamma_k+i\omega_k)\tau},$$

(55)
(56)

For the tracer, $T_k$, we used an ensemble of $L = 400$ trajectories generated via (8) to compute the correlation function. We note that for $k = 3$ and higher the corresponding Fourier mode $T_k(t)$ decorrelates very fast while for the first three modes, the correlation time is significantly longer. It is also interesting to note that there are noticeable peaks in the correlation functions of the tracer for the first few modes at the times that are multiples of $2\pi$ time units. We attribute these peaks to the strongly non-Gaussian statistics of the tracer due to nonlinear advection of the tracer by the cross-sweep, $U(t)$, and strong forcing by the waves through the mean gradient. In Fig. 3, we demonstrate the correlation time of both the velocity field and the tracer for the case of nondispersive waves. We compute the correlation time of a variable $u(t)$ as an integral of the absolute value of the correlation function $C_u(\tau)$

$$\tau_{\text{corr},u} = \int_0^\infty |C_u(\tau)|d\tau.$$ 

(57)

For the Gaussian velocity field, we have

$$\tau_{\text{corr},U} = \frac{1}{\gamma_U},$$

$$\tau_{\text{corr},v_k} = \frac{1}{\gamma_k},$$

(58)
(59)

On the other hand, for the tracer, we use the correlation function from Fig. 2 to find the correlation time. For our choice of parameters from Table 1, the correlation time of the tracer is significantly shorter than the correlation time of the velocity field for each Fourier mode for the case of nondispersive waves. Similarly, in Figs. 4 and 5, we show the correlation functions and correlation times of the velocity field and the tracer for the case of Rossby waves given in (20). We note that a different dispersion relation, $\omega_k$, leads to different oscillations of $C_{v_k}$ for nondispersive and Rossby waves. Moreover, the correlation functions of the tracer have similar pattern but different detailed structure especially for the first few Fourier modes. Again, we note the presence of the peaks at the times that are multiples of $2\pi$, which are even more pronounced for this case of Rossby waves. On the other hand, the correlation times for both cases of the linear dispersion $\omega_k$ are almost the same for the tracer and exactly the same for the velocity field since
the linear dispersion does not affect the correlation time of a linear wave because of the absolute value in definition (57).

In Figs. 6 and 7 we demonstrate the probability density functions (pdfs) of the first four Fourier modes of the tracer, $T_k$, and the pdfs of the tracer in the physical space, $T(x)$, along with the corresponding Gaussian fits to these pdfs with the same mean and variance. In order to obtain these pdfs, we used a trajectory of length $t = 30000$ time units in the equilibrium to construct the histograms of the corresponding Fourier modes and normalized them so that they integrate to unity. We note that both non-dispersive and Rossby waves lead to strong intermittency that is expressed in long exponential tails. However, in the case of Rossby waves (Fig. 7) not only the tails of the pdfs are non-Gaussian, but also the peaks of the pdfs are narrow. These non-Gaussian strongly intermittent pdfs are very similar to the practically observed pdfs for tracers in the atmosphere [1] which makes this test model extremely interesting and physically relevant. A similar phenomenon was observed and explained in [21] for the case of purely deterministic cross-sweep $U(t)$ through a different intermittency mechanism involving zeros of $U(t)$ as a source of intermittency.

4. Nonlinear Extended Kalman Filter

4.1. Classical Kalman Filter

The linear Kalman filter [2, 4] is used for linear systems when the best possible approximation of the true signal is needed. The algorithm for finding this optimal approximation incorporates the assumed linear dynamics with the imperfect observations of the true signal and uses least squares minimization of the error. Both dynamics and observations are characterized by deterministic and stochastic components. For the Gaussian statistics, the evolution of the ensemble can be fully described by the propagation of the mean and covariance.

Let us discuss the classical Kalman Filter algorithm. Suppose, the evolution model for a vector $\vec{u}_m \equiv \vec{u}(t_m) \in \mathbb{C}^N$ as a function of discrete time $t_m = m\Delta t$ is given by

$$\vec{u}_{m+1} = \vec{F}_{m+1}(\vec{u}_m) + \vec{f}_{m+1} + \vec{\sigma}_{m+1}, \quad (60)$$

where $\vec{F} \in \mathbb{C}^N$ is a linear function, $\vec{f} \in \mathbb{C}^N$ is a deterministic forcing, and $\vec{\sigma} \in \mathbb{C}^N$ is a complex Gaussian noise. Suppose also, that at each time $t_m$ we obtain an observation, $\vec{z}_m \in \mathbb{C}^M$, of the true value of the signal, $\vec{u}_m$

$$\vec{z}_{m+1} = G\vec{u}_{m+1} + \vec{\sigma}^o_{m+1}, \quad (61)$$

where $G \in \mathbb{C}^{M \times N}$ is a linear operator (a matrix) and $\vec{\sigma}^o \in \mathbb{C}^M$ is a complex Gaussian noise with mean zero and covariance $\sigma^o \in \mathbb{C}^{M \times M}$. The classical Kalman filter for (60) and (61) produces the mean and covariance of $\vec{u}_{m+1}$ prior and posterior to knowing observations $\vec{z}_{m+1}$. The prior mean and covariance are denoted as $(\vec{u})_{m+1|m}$ and $r_{m+1|m}$ while the posterior mean and covariance are $(\vec{u})_{m+1|m+1}$ and $r_{m+1|m+1}$. The **prior update** of the mean and covariance is done using (60)

$$\langle \vec{u} \rangle_{m+1|m} \rightarrow \langle \vec{u} \rangle_{m+1|m+1},$$

$$r_{m+1|m} \rightarrow r_{m+1|m+1}, \quad (62)$$
where the model error can also be introduced through the incorrect specification of the model in \( (60) \).

The posterior update is performed using the Kalman formulas

\[
\begin{align*}
\langle \bar{u} \rangle_{m+1|m+1} &= \langle \bar{u} \rangle_{m+1|m} + K_{m+1}(\bar{z}_{m+1} - G(\bar{u})_{m+1|m}), \\
r_{m+1|m+1} &= (\mathcal{I} - K_{m+1}G)r_{m+1|m}, \\
K_{m+1} &= r_{m+1|m}G^r(Gr_{m+1|m}G^r + r^s)^{-1},
\end{align*}
\]

(63)

where \( K_m \in \mathbb{C}^{N \times M} \) is the Kalman gain matrix and \( \mathcal{I} \) is the identity matrix of the size \( N \times N \). The two-step procedure outlined in (62) and (63) provides the classical Kalman filter algorithm that produces the best possible approximation in the least squares sense to the true signal when the model for the signal is linear Gaussian and the observations are also linear with Gaussian noise.

4.2. Nonlinear Extended Kalman Filter

Suppose that the governing model for true dynamics of the system \( (60) \) is nonlinear and non-Gaussian. Then, the Nonlinear Extended Kalman Filter \([14, 17, 18]\) provides a way to extend a classical Kalman filter algorithm to nonlinear models. If the first and second order statistics of the nonlinear model \( (60) \) can be computed exactly, then they are used in prior update step of the NEKF as in \( (62) \). The NEKF was first introduced in \([17, 18]\) and then successfully used in \([22–24]\) to filter both a prototype and more realistic atmosphere and ocean models with a large number of degrees of freedom. The advantage of having exactly solvable statistics avoids linearizations as in the standard Extended Kalman filter which is important in strongly non-Gaussian regimes.

The system \( (14), (16), (6) \) has exactly solvable first and second order statistics as was shown in Section 2. Therefore, this system provides an extremely interesting and physically relevant test case for the NEKF. In this situation, the state vector \( \bar{u} \) has \( 4K + 4 \) components

\[
\bar{u}(t) = \left( \text{Re}[V'(t)], \text{Im}[V'(t)], v(x_1, t), \ldots, v(x_{2K+1}, t), T(x_1, t), \ldots, T(x_{2K+1}, t) \right)^T.
\]

(64)

Note that we need to keep the imaginary part of \( V' \), \( \text{Im}[V'(t)] \), to update the statistics of the cross-sweep, even though the tracer only depends on the real part of \( V' \). We use \( (28), (33) \), and \( (13) \) to update \( \langle u \rangle \) from time \( t_0 = t_m \) to time \( t = t_{m+1} \). Similarly, we use \( (29), (30), (34) \) and equation from the Appendix to update the covariance matrix \( r \in \mathbb{R}^{4K+4 \times 4K+4} \) of \( \bar{u} \). This procedure becomes a prior update step of the NEKF for our model. The posterior update is given by the same Eq. \( (63) \) as in the classical KF.

4.3. Observations

The spatially extended structure of the model \( (14), (16), (6) \) provides an opportunity to consider virtually any configuration of observations ranging from plentiful observations of each variable at every grid point to very sparse observations at only one grid point. In particular, we will focus our attention on the following observations

- plentiful observations at every grid point
- partial observations, when only one or two variables out of the triplet \( V', v, T \) are observed,
• sparse observations, when only every \( p \)th point of the total \( 2K + 1 \) points are observed,

• extremely sparse observations, with only 1 grid point where either \( v \) or \( T \) or both are observed.

Next, we demonstrate the typical forms of the observation matrix \( G \) for various types of observations. Note that for the cross-sweep \( U \), we only observe the real part of \( V' \). For the case of plentiful observations we have \( G \in \mathbb{R}^{4K+3 \times 4K+4} \)

\[
G = \begin{pmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\tag{65}
\]

Next, for the case of observation of the cross-sweep and sparse observation for both \( v \) and \( T \) with \( p = 3 \), the observation matrix becomes \( G \in \mathbb{R}^{2Q+1 \times 4K+4} \) for \( 2K + 1 = pQ \)

\[
G = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\tag{66}
\]

Similarly, one can write the observation matrix for any other configuration of observations. The observation noise covariance matrix \( r_o \) is diagonal with the entries on the diagonal equal to \( r_U^o \) for the observations of the cross-sweep, \( r_v^o \) for the observations of the waves, and \( r_T^o \) for the observations of the tracer. The scalars \( r_U^o \), \( r_v^o \), and \( r_T^o \) are chosen with standard deviation based on the attractor size of the corresponding variable, e.g., 10% of the attractor size. These numbers will be specified below in the next Section.

We note that in the current setup, the observations could be taken on an irregular grid as well, which can be practically relevant for the situations when there are more measuring devices in some areas of the domain of interest than in others. However, in this paper we only consider uniform observational grids or the radical extreme limit of observations at a single point.

5. Filter performance

In this Section, we discuss the performance of the NEKF filter described in Section 4.2 on the model (14), (16), (6). We first show the behavior of the filter on a typical trajectory of this system in physical space. Then, we study how well the filter performs in Fourier space by comparing the spectra of a truth signal with the spectra of the corresponding filtered signal. We consider a suite of various observational schemes as discussed in Section 4.3. Then, we study the performance of the NEKF as a function of observation noise strength and observation time. Finally, we consider a situation of extremely sparse observations and discuss how much of the truth signal can be recovered.
if the observations are only taken at one spatial location. The results of our study are compared for the cases of dispersive Rossby waves and nondispersive waves.

We describe the performance of the filter by the filter skill, which measures the proximity of the filtered signal to the truth signal. We use the root mean square error (RMSE) to measure the filter skill

$$RMSE(\vec{z} - \vec{w}) = \sqrt{\frac{1}{N} \sum_{j=1}^{N} |z_j - w_j|^2},$$

(67)

where $\vec{z}$ and $\vec{w}$ are the complex vectors to be compared and $N$ is the length of each vector. The ratio of the RMSE and the typical magnitude of the signal gives the normalized percentage error. Moreover, to quantify how well the filtered signal recovers the pattern of the truth signal, we use the normalized cross-correlation function given by

$$XC(\vec{z}, \vec{w}) = \frac{1}{2} \left( XC(Re[\vec{z}], Re[\vec{w}]) + XC(Im[\vec{z}], Im[\vec{w}]) \right),$$

(68)

where the cross-correlation between real vectors $\vec{x}$ and $\vec{y}$ is given by

$$XC(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\sqrt{||\vec{x}||^2 ||\vec{y}||^2}}.$$

(69)

5.1. Filtering individual trajectories in physical space

Figure 8 demonstrates a segment of a typical trajectory as a function of time. The first panel in Fig. 8 shows the dynamics of the cross-sweep, the second and third panels show the dynamics of the waves and the tracer at a fixed spatial location, $x_{ob}$, where observations are available. On the other hand, the last two panels show the dynamics of the waves and the tracer at a different spatial location, $x_{no\,ob}$, where observations are not available. To generate this trajectory, we used the parameters from Table 1 for the Rossby waves case with $K = 52$ Fourier modes and $2K + 1 = 105$ spatial locations. For filtering, we use the observations of the cross-sweep (the real part of the fluctuations of the cross-sweep, $V'(t)$, around its mean state, $\bar{U}(t)$, as in (26) were actually observed and filtered), and sparse observations of the waves and the tracer with every third spatial location observed, i.e., with $p = 3$ and the observation matrix $G$ given by (66). We use the observation time $\Delta t = 0.125$, which is shorter than the correlation times of all modes of the system that are shown in Fig. 5. The trajectory is $L = 2000$ assimilation cycles long. We set the standard deviation of the observation noise to be equal 10\% of the typical (attractor) size of the corresponding variable. In Fig. 8, the observations are shown wherever they are available. Moreover, we show both the prior and the posterior filtered signals. We note the exceptionally good skill of the NEKF: the filtered dashed line is almost always on top of the truth solid line. Note that in the first three panel of Fig. 8, where observations are available, the filtered signal tends to be between the prior forecast and the observations, which is in accord with the general theory of Kalman filters [2, 4, 5, 14]. By comparing the second and the fourth panels of Fig. 8 for the waves, we note that having observations at a particular location, $x_{ob}$, improves the filter skill, however, the general shape of the truth signal is recovered even at the location, $x_{no\,ob}$, where observations are not available. Similar conclusions can be drawn from comparing
the third and the fifth panels of Fig. 8 for the tracer. However, here the skill of the filter is excellent even at the non-observable location, \( x_{\text{no\,ob}} \). This is explained below by the fact that observation of the cross-sweep improve the filter skill for the tracer.

In Fig. 9 (a,b), we show the snapshot of the same trajectory at the time \( t = 33.375 \) (see Fig. 8 for the temporal evolution). We note that the observations are only given at every third spatial location as discussed above. By comparing the truth and the filtered signals by sight, we note very high skill of the NEKF. Next, in Fig. 9 (c,d), we show the cross correlation given by (68) for both the waves and the tracer as a function of spatial location. The oscillating pattern of the cross-correlation in Fig. 9 (c,d) reflects the presence and the absence of the observations. At the locations where observations are available, the pattern correlation is close to 1, which is a sign of almost perfect recovery of the truth signal by the filtered signal. On the other hand, at the locations where observations are not available, the cross-correlation has the values around 0.85 for the the waves, \( v(x,t) \), which is still significantly high. For the tracer, \( T(x,t) \), the cross-correlation is always very high, of the order 0.98, with slightly higher skill at the locations with observations than at the locations without observations.

### 5.2. Recovery of turbulent spectra

Here, we study how well the NEKF recovers the spectrum of the original truth signal with various observation schemes. We test the NEKF on a sample trajectory that is \( L = 2000 \) assimilation cycles long, with each cycle \( \Delta t = 0.125 \). We use very small observation noise with a standard deviation of approximately 1% of the typical size of the signal. We consider six different observation schemes:

1. observations of \( v \) only
2. observations of \( T \) only
3. observations of both \( U \) and \( v \)
4. observations of both \( U \) and \( T \)
5. observations of both \( v \) and \( T \)
6. observations of \( U, v, \) and \( T \).

Within each of these schemes, we take observations either at each or at every 3d or at every 7th spatial locations to study how well sparse observations are filtered. We note that the NEKF uses the exact nonlinear statistics to build the prior forecast, moreover, we use very small observation time and observation noise. Therefore, most of the errors in the filtered solution can be attributed to the imperfect partial and sparse observations. This set up provides a transparent and unambiguous test case for the role of observations in the NEKF.

In Figs. 10, 11, we compare the truth spectra and the filtered spectra for the case of nondispersive waves for both \( v \) and \( T \). Since the equilibrium statistics are time dependent and we only have one sample trajectory, as opposed to having an ensemble of trajectories, we consider time averaged squared amplitudes, \( \langle |v_k|^2 \rangle \) and \( \langle |T_k|^2 \rangle \), instead of the time averaged variances which require the knowledge of the time dependent values of the means of \( v_k \) and \( T_k \), the information that may not be available. Here, we refer to these time averaged squared amplitudes as the spectrum, although this definition is technically different from the one used in Section 3.5; therefore, the truth spectra in Figs. 10, 11 are a little different from the “ideal” spectra in Fig. 1. In Figs. 12, 13, we compare the
truth spectra and the filtered spectra for the case of Rossby waves for both $v$ and $T$. After studying Figs. 10-13, we draw the following conclusions:

- By decreasing the number of observed points on the grid, we notice a naturally expected decrease in the skill of the filtering. Moreover, this deterioration of the skill appears at the high wave numbers which is also not surprising since the high wave numbers correspond to the small spatial scales in physical space.

- If only the waves, $v$, are observed then the filtered spectrum of waves is very close to the truth spectrum of waves, while the filtered spectrum of the tracer has a large error when compared to the truth spectrum of the tracer. However, the forms of the filtered and the truth spectra of the tracer are similar by sight.

- Observations of both components of the velocity field, i.e., the cross-sweep and the waves, significantly improve the skill in filtering the spectrum of the tracer even though the tracer is not directly observed.

- Observations of the tracer alone provide very good skill in filtering the spectrum of the tracer. Moreover, the filtered spectrum of the waves at small wave numbers is close to the truth spectrum of the waves, i.e., observations of the tracer alone carry information about the large scale spatial structure of the waves. The recovery of the small scale spatial structure of the waves requires additional observations of the waves.

- Additional observations of the cross-sweep when the tracer is observed do not help to improve the recovery of the spectrum of the waves at the large wave numbers which is a consequence of the fact that the cross-sweep and the waves have independent dynamics.

- Plentiful observations of all variables provide the best filtered skill because maximum possible information is utilized in the recovery of the spectra.

- The NEKF performs equally well in recovering the turbulent spectra for both the non-dispersive and dispersive Rossby wave systems for the case of sparse but not extremely sparse observations.

Note that here, we concentrate on the filtering of the waves, $v$, and the tracer, $T$. The role of the observations of the cross-sweep will be discussed next in Section 5.4. The skill of the filter in recovering the cross-sweep will be discussed below in Section 5.5 where we focus on studying the dependence of the filter skill on the observation noise strength and the observation time.

5.3. Recovery of the fat tail tracer probability distributions

Another important statistical characteristics of a stochastic turbulent signal is its equilibrium pdf. As we discussed in the introduction and in Section 3.5, one of the key properties of the model for the tracer with a mean gradient is the strongly intermittent pdfs with long exponential tails as in the tracers observed in the atmosphere [1]. Therefore, it is important that a filtering algorithm recovers the pdf of the original true signal. Here, we study how well the NEKF recovers the truth pdf depending on the
observations. To make the testing fair, we compare the histogram of the filtered signal with the histogram of the sample truth trajectory. Note that due to under-sampling with the shorter time series of the true signal, the histogram of the sample truth trajectory may be different from the ideal pdf of the truth signal. In Fig. 14, we demonstrate how the probability distribution of the tracer is recovered by the filtered signal for the Rossby waves case for different types of sparse observations with $p = 7$. We used the whole trajectory for the tracer in space and time ($2K + 1 = 105$ spatial points $\times$ 2000 assimilation cycles) and 80 bins to construct the histograms in Fig. 14. Note that the normalized histogram of the sample truth trajectory, which is filtered, is different from the ideal pdf of the tracer. However, the sample trajectory still has long exponential tails although with a slope different from the slope of the tails of the ideal pdf. We note that the filtered histogram almost perfectly recovers the histogram of the sample truth trajectory for all types of observations except for the case of observations of the waves only. As we have seen earlier in Section 5.2, the NEKF has also poor skill in recovering the spectrum of the tracer when only the waves are observed, otherwise the spectrum of the tracer was recovered with high skill. Similar results were found for the case of non-dispersive waves.

5.4. Improvement of the filtering skill by adding just one observation

Let us consider a situation when only very sparse observations are available. How much can one improve the filtering skill by adding just one observation at a fixed spatial location? Will an observation of the velocity field or of the tracer provide better filtering skill? To mimic such situations, we carry out the following simulation. Suppose, that in our model only the cross-sweep is observed. Then we fix some spatial location (which can be arbitrary because of the translational symmetry of the system) at which observations of either the waves, $v$, or the tracer, $T$, or both are taken. By comparing the filtering skill in all of these four filtering setups, we find answers to the questions asked in the beginning of this Section.

We start with the case of non-dispersive waves. In Fig. 15, we show the skill of the NEKF in recovering the spatial structure of the waves and of the tracer using both the pattern cross-correlation, $XC$, and the root mean square error, $RMSE$. If only the cross-sweep $U$ is observed, then the NEKF has no skill with the values of $XC \approx 0.1$ and $RMSE \approx 1$. If in addition to the cross-sweep, there is one observation of the waves, $v$, available at each assimilation cycle, we improve the skill of NEKF in filtering waves at the observation point and to the right of the observation point, i.e., in the direction of the zonal cross-sweep. At the observation point, we have almost perfect recovery of the waves with $XC \approx 1$ and $RMSE \approx 0.1$. Then, as we move further away to the right of the observation point, the filtering skill gradually drops to the value of $XC \approx 0.3$ and $RMSE \approx 0.9$. On the other hand, the skill in filtering the tracer improves a lot by adding only one observation of the waves: cross-correlation increases from $XC \approx 0.2$ with observations of the cross-sweep alone to values as high as $XC \approx 0.75$ at each point of the grid, not just around the observation point, and the $RMSE$ decreases from $RMSE \approx 1$ to $RMSE \approx 0.65$ also uniformly around the spatial domain. Next, we study the improvement of the filtering skill by adding one observation of the tracer. In this situation, the skill in filtering waves increases very slightly up to $XC \approx 0.4$ and $RMSE \approx 0.9$ uniformly around the spatial domain. On the other hand, the skill of the NEKF in filtering the tracer increases significantly: at the observation site, $XC \approx 1$
and $RMSE \approx 0.2$, and as we move further away to the right of the observation point, i.e., in the direction of the zonal cross-sweep, the skill decreases to $XC \approx 0.85$ and $RMSE \approx 0.65$, which still indicate significant skill. Finally, if we add observations of both the waves and the tracer at the same spatial location in addition to the observations of the cross-sweep, the filtering skill for both $v$ and $T$ improves just slightly over the situation when only the cross-sweep and the tracer at one spatial location are observed.

In Fig. 16, we compare the truth spectra of both $v$ and $T$ with the filtered spectra when in addition to the observations of the cross-sweep, the observations of either waves, or tracer, or both become available at one spatial location. We note that with the observations of the cross-sweep alone, the filtered spectra have very large errors. However, if the observations of either $v$, or $T$, or both, at one spatial location become available, the NEKF recovers the forms of the spectra of both $v$ and $T$, although the exponents of the turbulent cascades of the filtered signals are different from the corresponding exponents for the truth spectra. We note that for the first part of the spectrum that involve the first 10 modes, the recovery for both waves and tracer is excellent regardless of which variable is observed at the fixed spatial location. However, the high wave number parts of the filtered spectra have steeper decay than in the true spectra. It is natural to expect more skill in the low wave number part of the spectrum with sparse observations than in the high waves number part since sparse observations provide information about the large scale spatial structure of the truth signal. However, it is surprising to find that just one observation can improve the filtered spectrum so much.

We continue with the case of Rossby waves. In Fig. 17, we show the skill of the NEKF in recovering the spatial structure of the waves and of the tracer using both the pattern cross-correlation, $XC$, and the root mean square error, $RMSE$. We note that in this case even with the observations of the cross-sweep only, the filtered signal has some nontrivial skill with the $XC \approx 0.65$ and $RMSE \approx 0.8$ for the waves and $XC \approx 0.55$ and $RMSE \approx 0.8$ for the tracer. Here, similar to the case of non-dispersive waves, the skill in filtering the tracer improves significantly if one observation of the tracer becomes available: $XC$ increases up to 0.9 and $RMSE$ drops to 0.4. On the other hand, unlike in the case of nondispersive waves, here, the skill of filtering the waves does not improve significantly when an observation of either $v$, or $T$, or both, become available at one point, except for the skill at the close proximity to this observation point. Moreover, the observations of the waves do not help to recover the tracer. Similar conclusions can be drawn from Fig. 18, for the spectra of the waves and the tracer.

In Fig. 19, we study the improvement of recovery of the pdf in physical space by the NEKF for the case of nondispersive waves when only one observation of the waves and/or of the tracer become available. We note that with observations of the cross-sweep only, the filtered pdf is much more peaked than the truth pdf. However, when only one observation of either the waves or the tracer or both become available, the filtered pdf almost perfectly recovers the sample truth pdf. For the case of Rossby waves shown in Fig. 20, we note that if only the cross-sweep is observed, the filtered histogram has a higher central peak and much narrower tails that the pdf of the truth trajectory. Adding one observation of the waves does not improve the filtered pdf. However, adding one observation of the tracer makes the filtered pdf very close to the sample truth pdf.

To summarize this Section, we have found that adding one observation of either the waves, or the tracer, or both, can significantly improve the filtering skill in both physical space and Fourier space. In the case of nondispersive waves, the filtered signal is closer
to the truth signal in the direction of the cross-sweep, i.e., the information that the observation carries, propagates in the direction of the cross-sweep. On the other hand, for Rossby waves the filtered signal for the tracer improves if an observation of the tracer becomes available and this improvement is only slightly better in the direction of the cross-sweep.

5.5. Filtering skill as a function of observation time and observation noise strength

In this Section, we study how the NEKF performance depends on the observation noise strength and observation time. We test the statistics of the filter on a long trajectory with \( L = 4000 \) assimilation cycles with observation time \( \Delta t = 0.125 \), i.e., the total length of the test trajectory is \( \tau = 500 \). The test trajectory is generated for a system with \( K = 10 \) Fourier modes or \( 2K + 1 = 21 \) spatial locations and the energy spectrum is given by \( (23) \) with \( k_0 = 5 \). For this fixed test trajectory, we run a series of simulations, when the filtering is performed with various levels of the observation noise strength, various lengths of observation time, six different types of observations as discussed in Section 5.2, and three different values of the sparsity parameter \( p = \{1, 3, 7\} \) which means that every \( p \)th spatial location is observed. For the observation noise level, \( r_0^{\text{ref}} \), we choose the reference value that corresponds to errors of 10% of the attractor size of a corresponding variable. Then we decrease or increase this reference level of the noise by a factor from the set \( a \in \{0.125, 0.25, 0.5, 1, 2, 4, 8\} \) and filter the system with the level of noise given by \( r^o = ar_0^{\text{ref}} \). In this way, we test the performance of the filter for a very broad range of observation noises, \( r^o \), from very small values of the order of 1% of the size of the truth signal to extremely large values that are comparable with the size of the truth signal. Note that \( r^o \) has three components that correspond to \( U, v, \) and \( T \). Similarly, for the observation time, we choose a reference value \( \Delta t_{\text{ref}} = 0.125 \) time units. The series of observation times is produced by multiplying \( \Delta t_{\text{ref}} \) by a factor \( b \in \{1, 2, 4, 8, 16\} \), i.e., \( \Delta t = b\Delta t_{\text{ref}} \). The smallest observation time is shorter than the correlation time of any mode in the system (see Fig. 5 for the first 10 modes). The largest observation time \( \Delta t = 2 \) is between the correlation times of the second and the third Fourier modes of the tracer and is still shorter than the correlation time of all 10 modes of the waves (see Fig. 5 for the first 10 modes).

In Fig. 21, we show the RMSE for the cross-sweep as a function of the noise strength of the observations for the case of Rossby waves. We note that the filter performance deteriorates as the noise strength of the observations increases or if the observation time increases. Moreover, we note that if only the tracer is observed, the cross-sweep can still be filtered very well, while observations of the waves alone provide no information about the cross-sweep and, therefore, the filtering of the cross-sweep from observations of the waves has no skill at all. As the sparsity of the observations increases, the filtering skill decreases. However, even when only 3 out of 21 spatial locations are observed the filtering of the cross-sweep has significant skill for small enough noise level and short enough observation time. In Figs. 22 and 23, we show the RMSE for the waves and the tracer as a function of the noise strength of the observations for the case of Rossby waves. Here again we confirm that with larger observation noise and with larger observation time, the filtering skill deteriorates. It is surprising to find that if the waves are filtered from observations either of the tracer alone, or the tracer and the cross-sweep, then the performance of the filter is more sensitive to the change in observation time for small values of the observation noise than for large values of observation noise. On the other
hand, if the waves are filtered from observations of the waves and possibly other variables, then the filter skill is more sensitive to the changes in observation time for large values of the noise strength. A similar situation is observed for the tracer. The general conclusion here is that if we are interested in filtering a variable from the observations of other variables then it is essential to use the shortest possible observation time. However, if the observations of the variable of interest are available, then increasing the observation time may not hamper the filtering skill especially, if the noise variance of the observations is kept small. Finally, we note that more sparse observations lead to less accurate filtering results as expected. We also confirmed that similar results hold for the test case with nondispersive waves.

6. Conclusions

A statistically exactly solvable test model for filtering turbulent multiscale signals from nature was introduced here. This model is a special case of the advection-diffusion equation with a Gaussian velocity field. Despite the special structure this model has many important features of real physical systems such as turbulent cascade and intermittent probability distribution of the tracer. Moreover, the model is general enough to mimic many engineering as well as climate science applications. The remarkable property of the system is that the exact first and second order statistics can be computed analytically. This enables us to use the model with a Nonlinear Extended Kalman filter (NEKF) which uses the exact first and second order statistics to forecast the next state of the model. We studied the role of sparse and partial observations in the filter performance. Moreover, we studied the skill of the NEKF in recovering the turbulent spectrum of both the velocity field and the tracer and recovering the non-Gaussian pdf of the tracer. We learned that by adding just one observation, we can dramatically improve the performance of the filter which is very interesting from a practical perspective when only few observations are available.

A number of exciting future directions follow from this work since the present model provides an unambiguous test problem for some central contemporary issues in filtering. The existence of analytical formulas for the eddy diffusivity provides an opportunity to address the issues of model error in a transparent fashion for both filtering and climate change science [15]. Another interesting class of problems that can be addressed via information theory is how to optimize targeted observations for such turbulent signals with model error. The mathematical aspects of the model regarding the turbulent tracer spectrum and exponential tail pdfs merit further investigation. The authors are currently pursuing all of these issues and will report on them in the near future.

Acknowledgment

The research of A.J. Majda is partially supported by National Science Foundation grant DMS-0456713, the office of Naval Research grants 25-74200-F6607, and N00014-05-1-0164, and the Defense Advanced Projects Agency grant N0014-07-1-0750. Boris Gershgorin is supported as a postdoctoral fellow through the last two agencies.
A. Second order statistics of the tracer

In Section 2, we showed how to compute first and second order statistics of the Gaussian velocity field \((U(t), v_k(t))\)^T, and the mean of the tracer \(\langle T_k(t) \rangle\). Here, we present the analytical formulas for the second order statistics of the system \((U(t), v_k(t), T_k(t))\)^T given Gaussian initial values \((U(t_0), v_k(t_0), T_k(t_0))\)^T.

\subsection*{A.1 Cov\((U, T)\)}

We start with the covariance of the cross-sweep and the tracer. Note that for simplicity we consider the fluctuations \(V'(t)\) of the function \(V(t)\) around the deterministic mean (see (26) for more details)

\[
\text{Cov}(V'(t), T_k(t)) = \langle V'(t)T_k(t) \rangle - \langle V'(t) \rangle \langle T_k(t) \rangle.
\]

We only need to compute the first correlator and then use (28) and (13) for the second term

\[
\langle V'(t)T_k(t) \rangle = D_k(t_0, t)e^{-ikJ(t_0, t)}\langle V'(t)T_k(t_0) \rangle e^{-ikJ(t_0, t)}
\]

\[
- \alpha \int_{t_0}^{t} D_k(s, t)e^{-ikJ(s, t)}e^{\lambda_s(t-s)} \langle V'(t)v_k(t_0) \rangle e^{-ikJ(s, t)} ds
\]

\[
- \alpha \int_{t_0}^{t} D_k(s, t)e^{-ikJ(s, t)}F_k(t_0, s) \langle V'(t) \rangle e^{-ikJ(s, t)} ds.
\]

The correlator in the first line becomes

\[
\langle V'(t)T_k(t_0) e^{-ikJ(t_0, t)} \rangle = (\text{Cov}(V'(t), T_k^*(t_0)) + \langle V'(t) \rangle \langle T_k(t_0) \rangle)
\]

\[
-ik[\langle V'(t) \rangle \text{Cov}(T_k(t_0), J(t_0, t)) + \langle T_k(t_0) \rangle \text{Cov}(V'(t), J(t_0, t))]
\]

\[
+k^2 \text{Cov}(T_k(t_0), J(t_0, t)) \text{Cov}(V'(t), J(t_0, t))
\]

\[
\times e^{-ikJ(t_0, t)} - \frac{k^2}{2} \text{Var}(J(t_0, t))
\]

\[
e^{\lambda_U(t-t_0)} (\text{Cov}(V'(t_0), T_k^*(t_0)) + \langle V'(t_0) \rangle \langle T_k(t_0) \rangle)
\]

\[
-ik[\langle V'(t_0) \rangle \text{Cov}(T_k(t_0), J(t_0, t)) + \langle T_k(t_0) \rangle \text{Cov}(v_k(t_0), J(t_0, t))]
\]

\[
-k^2 \text{Cov}(v_k(t_0), J(t_0, t)) \text{Cov}(V'(t_0))
\]

\[
e^{-ikJ(t_0, t)} - \frac{k^2}{2} \text{Var}(J(t_0, t))
\]

where

\[
\text{C}_V(s) = e^{-\lambda_U(t-t_0)} \text{Cov}(V'(t), J(s, t))
\]

\[
= \frac{1}{2} \left( \frac{\text{Var}(V'(t_0))}{\lambda_U} e^{\lambda_U(t-t_0)} - e^{\lambda_U(t-t_0)} + \frac{\text{Cov}(V'(t_0), V(t_0))}{\lambda_U} (e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)}) \right)
\]

\[
+ \frac{\sigma^2}{4\gamma_U} \left( -\frac{1}{\lambda_U} (e^{-\lambda_U(t-t_0)} - e^{-\lambda_U(s-t_0)}) - \frac{1}{\lambda_U} (e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)}) \right).
\]

Similarly, we find

\[
\langle V'(t)v_k(t_0) e^{-ikJ(s, t)} \rangle = (\text{Cov}(V'(t_0), v_k^*(t_0)) + \langle V'(t_0) \rangle \langle v_k(t_0) \rangle)
\]

\[
-ik[\langle V'(t_0) \rangle \text{Cov}(v_k(t_0), J(s, t)) + \langle v_k(t_0) \rangle \text{Cov}(V'(s))]
\]

\[
-k^2 \text{Cov}(v_k(t_0), J(s, t)) \text{Cov}(V'(s))
\]

\[
\times e^{-ikJ(s, t)} - \frac{k^2}{2} \text{Var}(J(s, t)).
\]
and

\[ \langle V'(t) e^{-ikJ(s,t)} \rangle = e^{\lambda V (t-t_0)} \left( \langle V'(t_0) \rangle - ikC_V(s) \right) e^{-ik(J(s,t))} - \frac{2}{\gamma} \text{Var}(J(s,t)). \]

A.2 Cov(v, T)

We continue with the covariance of the waves and the tracer.

\[ \text{Cov}(v_k(t), T_l(t)) = \langle v_k(t)T_l^*(t) \rangle - \langle v_k(t) \rangle \langle T_l(t) \rangle^*. \]

We already know how to find \( \langle v_k(t) \rangle \) and \( \langle T_l(t) \rangle \) from (33) and (13), so we just compute the first correlator

\[
\langle v_k(t)T_l^*(t) \rangle = e^{\lambda_k(t-t_0)} D_l(t_0, t) e^{il\tilde{J}(t_0, t)} \langle v_k(t_0)T_l^*(t_0) \rangle e^{il\tilde{J}(t_0, t)}
\]

\[
+ F_k(t_0, t) D_l(t_0, t) e^{il\tilde{J}(t_0, t)} \langle T_l^*(t_0) \rangle e^{il\tilde{J}(t_0, t)}
\]

\[
- \alpha \int_{t_0}^t D_l(s, t) e^{il\tilde{J}(s, t)} \langle v_k(t) v_l^*(s) \rangle e^{il\tilde{J}(s, t)} ds,
\]

where the correlators become

\[
\langle v_k(t_0)T_l^*(t_0) \rangle = \left( \text{Cov}(v_k(t_0), T_l(t_0)) + \langle v_k(t_0) \rangle \langle T_l(t_0) \rangle^* \right)
\]

\[
+ il[\langle v_k(t_0) \rangle \text{Cov}(T_l^*(t_0), J(t_0, t)) + \langle T_l(t_0) \rangle^* \text{Cov}(v_k(t_0), J(t_0, t))]
\]

\[
- l^2 \text{Cov}(v_k(t_0), J(t_0, t)) \text{Cov}(T_l^*(t_0), J(t_0, t)) \right)e^{il\tilde{J}(t_0, t)} - \frac{2}{\gamma} \text{Var}(J(t_0, t)),
\]

and

\[
\langle T_l^*(t_0) e^{il\tilde{J}(t_0, t)} \rangle = \left( \langle T_l(t_0) \rangle^* + il \text{Cov}(T_l^*(t_0), J(t_0, t)) \right) e^{il\tilde{J}(t_0, t)} - \frac{2}{\gamma} \text{Var}(J(t_0, t)),
\]

and

\[
\langle v_k(t) v_l^*(s) e^{il\tilde{J}(s, t)} \rangle = e^{\lambda_k(t-t_0) + \lambda_l^*(s-t_0)} \left( \text{Cov}(v_k(t_0), v_l(t_0)) + \langle v_k(t_0) \rangle \langle v_l(t_0) \rangle^* \right)
\]

\[
+ \frac{\sigma_k^2}{2\gamma} \delta_l \left( e^{2\gamma(s-t_0)} - 1 \right) + il[\langle v_k(t_0) \rangle \text{Cov}(v_l^*(t_0), J(s, t)) + \langle v_l(t_0) \rangle^* \text{Cov}(v_k(t_0), J(s, t))]
\]

\[
- l^2 \text{Cov}(v_k(t_0), J(s, t)) \text{Cov}(v_l^*(t_0), J(t_0, t)) \right) + F_k(t_0, t) e^{\lambda_l^*(s-t_0)} \langle v_l^*(t_0) \rangle + il \text{Cov}(v_l^*(t_0), J(s, t))
\]

\[
+ F_l^*(t_0, s) e^{\lambda_k(t-t_0)} \left( \langle v_k(t_0) \rangle + il \text{Cov}(v_k(t_0), J(s, t)) \right) + F_k(t_0, t) F_l^*(t_0, s) \right) e^{il\tilde{J}(s,t)} - \frac{2}{\gamma} \text{Var}(J(s,t)).
\]

A.3 Cov(T, T)

Finally, we compute the cross-covariance of the tracer

\[ \text{Cov}(T_k(t), T_l(t)) = \langle T_k(t)T_l^*(t) \rangle - \langle T_k(t) \rangle \langle T_l(t) \rangle^*. \]
We already know how to find \( \langle T_k(t) \rangle \) from (13), so we just compute the first correlator

\[
\langle T_k(t)T^*_k(t) \rangle = D_k(t_0, t)D_l(t_0, t)e^{-i(k-l)\tilde{J}(t_0, t)}\langle T_k(t_0)T^*_l(t_0)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle
\]

\[-\alpha D_k(t_0, t)\int_t^t D_k(s, t)e^{-ik\tilde{J}(s, t)}\langle v_k(s)T^*_k(t_0)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle ds \]

\[-\alpha D_k(t_0, t)e^{-ik\tilde{J}(t_0, t)}\int_t^t D_l(s, t)e^{ik\tilde{J}(s, t)}\langle v_l(s)T^*_k(t_0)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle ds \]

\[+\alpha^2 \int_t^t \int_{t_0}^t D_k(s, t)D_l(r, t)e^{-i(k-l)(\tilde{J}(t_0, t)-\tilde{J}(r, t))}\langle v_k(s)v_l^*(r)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle ds dr.\]

The correlator in the first term is computed by

\[
\langle T_k(t_0)T^*_k(t_0)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle = \left(\text{Cov}(T_k(t_0), T_l(t_0)) + \langle T_k(t_0) \rangle \langle T_l(t_0) \rangle \right) - \frac{k^2}{2}\text{Var}(J(t_0, t)).
\]

The correlator in the second term can be found by

\[
\langle v_k(s)T^*_k(t_0)e^{-i(k-l)\tilde{J}(t_0, t)} \rangle = e^{\lambda_k(s-t_0)}\left(\text{Cov}(v_k(t_0), T_l(t_0)) + \langle v_k(t_0) \rangle \langle T_l(t_0) \rangle \right)
\]

\[+i\langle v_k(t_0) \rangle (-k\text{Cov}(T^*_k(t_0), J(s, t)) + l\text{Cov}(T^*_l(t_0), J(t_0, t)))
\]

\[+i\langle T_l(t_0) \rangle (-k\text{Cov}(v_k(t_0), J(s, t)) + l\text{Cov}(v_k(t_0), J(t_0, t)))
\]

\[+(-k\text{Cov}(T^*_l(t_0), J(s, t)) + l\text{Cov}(T^*_l(t_0), J(t_0, t)))\langle v_k(t_0) \rangle (J(s, t)) + l\text{Cov}(v_k(t_0), J(t_0, t))\)

\[\times e^{-i(k-l)\tilde{J}(t_0, t)) - \frac{k^2}{2}\text{Var}(J(t_0, t)) - \frac{l^2}{2}\text{Var}(J(t_0, t)) + kl\text{Cov}(J(t_0, t), J(t_0, t))}
\]

\[+F_k(t_0, s)\left(\langle T_l(t_0) \rangle + i(-k\text{Cov}(T_l(t_0), J(s, t)) + l\text{Cov}(T_l(t_0), J(t_0, t)))\right)
\]

\[\times e^{-i(k-l)\tilde{J}(t_0, t)) - \frac{k^2}{2}\text{Var}(J(t_0, t)) - \frac{l^2}{2}\text{Var}(J(t_0, t)) + kl\text{Cov}(J(t_0, t), J(t_0, t))}
\]

Using symmetry we only need to find \( \text{Cov}(J(s, t), J(r, t)) \) for \( s > r \)

\[
\text{Cov}(J(s, t), J(r, t)) = \text{Cov}\left(\int_s^t \text{Re}[V'(s')]ds', \int_r^t \text{Re}[V'(r')]dr'\right)
\]

\[= \frac{1}{2} \text{Re} \left[ \int_s^t \int_r^t (\text{Cov}(V'(s'), V'(r')) + \text{Cov}(V'(s'), V'(r')^*)) ds'dr' \right]
\]

\[= \frac{1}{2} \text{Re} \left[ \int_s^t \int_r^s (\cdot) ds'dr' + \int_s^t \int_s^r (\cdot) ds'dr' \right]
\]

\[= \text{Var}(J(s, t)) + \frac{1}{2} \text{Re} \left[ \int_s^t \int_s^r \text{Cov}(V'(s'), V'(r')) + \text{Cov}(V'(s'), V'(r')^*) ds'dr' \right]
\]

(70)
We find
\[\int_s^t \int_r^s \left( \text{Cov}(V'(s'), V'(r')) + \text{Cov}(V'(s'), V'(r')^*) \right) ds' dr' \]
\[= \text{Var}(V'(t_0)) \int_s^t \int_r^s e^{\lambda_U(s'-t_0) + \lambda_U(r'-t_0)} ds' dr' + \frac{\sigma^2}{2\gamma_U} \int_s^t \int_r^s e^{\lambda_U s' + \lambda_U r'} (e^{2\gamma_U r'} - e^{2\gamma_U t_0}) ds' dr' \]
\[+ \text{Cov}(V'(t_0)) \int_s^t \int_r^s e^{\lambda_U (s'+r'-2t_0)} ds' dr', \]
\[
\text{where}\]
\[\int_s^t \int_r^s e^{\lambda_U(s'-t_0)+\lambda_U(r'-t_0)} ds' dr' = \frac{1}{|\lambda_U|^2} \left( e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)} \right) \left( e^{\lambda_U(s-t_0)} - e^{\lambda_U(r-t_0)} \right) \]
\[\int_s^t \int_r^s e^{\lambda_U(s'+r'-2t_0)} ds' dr' = \frac{1}{\lambda_U} \left( e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)} \right) \left( e^{\lambda_U(s-t_0)} - e^{\lambda_U(r-t_0)} \right) \]
\[\int_s^t \int_r^s e^{\lambda_U s' + \lambda_U r'} (e^{2\gamma_U r'} - e^{2\gamma_U t_0}) ds' dr' \]
\[= \frac{1}{\lambda_U} \left( e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)} \right) \left( e^{\lambda_U(s-t_0)} - e^{\lambda_U(r-t_0)} \right) \]
\[= \left( e^{\lambda_U(t-t_0)} - e^{\lambda_U(s-t_0)} \right) \left( -\frac{1}{\lambda_U} \left( e^{-\lambda_U(s-t_0)} - e^{-\lambda_U(r-t_0)} \right) - \frac{1}{|\lambda_U|^2} \left( e^{\lambda_U(s-t_0)} - e^{\lambda_U(r-t_0)} \right) \right) \]

For the last term, we find
\[\langle v_k(s)v_j^*(r)e^{-i(kJ(s,t)-lJ(r,t))} \rangle = \left( \text{Cov}(v_k(s), v_l(r)) + \langle v_k(s) \rangle \langle v_l(r) \rangle \right) \]
\[+ \langle v_k(s) \rangle (-k\text{Cov}(v_j^*(r), J(s,t)) + l\text{Cov}(v_j^*(r), J(r,t))) \]
\[+ \langle v_j^*(r) \rangle (-k\text{Cov}(v_k(s), J(s,t)) + l\text{Cov}(v_k(s), J(r,t))) \]
\[-(-k\text{Cov}(v_k(s), J(s,t)) + l\text{Cov}(v_j^*(r), J(r,t)))(-k\text{Cov}(v_k(s), J(s,t)) + l\text{Cov}(v_k(s), J(r,t))) \]
\[\times \left( e^{-i(kJ(s,t)-lJ(r,t))} - \frac{\sigma^2}{2\gamma_U} \text{Var}(J(s,t)) - i\text{Cov}(J(s,t), J(r,t)) \right) \]
\[+ \left( F_k(t_0, s)e^{\lambda_U (s-t_0)} \langle v_l(t_0) \rangle \right) + F_j^*(t_0, r)e^{\lambda_U (s-t_0)} \langle v_k(t_0) \rangle \]
\[+ iF_j^*(t_0, r)e^{\lambda_U (s-t_0)} (-k\text{Cov}(v_j^*(t_0), J(s,t)) + l\text{Cov}(v_j^*(t_0), J(r,t))) \]
\[+ iF_k(t_0, s)e^{\lambda_U (r-t_0)} (-k\text{Cov}(v_k(t_0), J(s,t)) + l\text{Cov}(v_k(t_0), J(r,t))) \]
\[\times e^{-i(kJ(s,t)-lJ(r,t))} - \frac{\sigma^2}{2\gamma_U} \text{Var}(J(s,t)) - i\text{Cov}(J(s,t), J(r,t)) \right) \]
Here, we used
\[ \text{Cov}(v_k(t), v_l(t)) = e^\lambda_k(s-t_0) + \lambda_l(r-t_0) \text{Cov}(v_k(t_0), v_l(t_0)) + e^\lambda_k(s-t_0) + \lambda_l(r-t_0) J_v, \]
where
\[ J_v = \frac{\sigma^2}{2 \gamma_k} \delta^k_t \left( e^{2 \gamma_k (\min(s,r) - t_0)} - 1 \right). \]

A.4 Invariant statistics

Here, we present the time invariant first and second order statistics of the system by taking the limit \( t_0 \to -\infty \) in the above formulas. For the cross-sweep, we find
\[ \langle V'(t) \rangle = 0, \]
\[ \text{Var}(V'(t)) = \frac{\sigma^2}{2 \gamma U}, \]
\[ \text{Cov}(V'(t), V'(t)^*) = 0. \]

Next, for the waves \( v_k(t) \), we have
\[ \langle v_k(t) \rangle = F_k(t), \]
\[ \text{Cov}(v_k(t), v_l(t)) = \frac{\sigma^2}{2 \gamma_k} \delta^k_t, \]
\[ \text{Cov}(V(t), v_l(t)) = 0, \]
where
\[ F_k(t) = \int_{-\infty}^{t} e^{\lambda_k(t-s)} f_k(s) ds = \frac{A_k}{i \xi - \lambda_k} e^{i(\xi t + \phi_k)}. \]

Next, we find statistics of \( T \). The mean becomes
\[ \langle T_k(t) \rangle = -\alpha \int_{-\infty}^{t} D_k(s, t) e^{-ik J(s,t)} F_k(s) e^{-\frac{i \xi}{2} \text{Var}(J(s,t))} ds, \]
where
\[ \text{Var}(J(s,t)) = \frac{\sigma^2}{2 \gamma U} \left( \text{Re} \left[ \frac{e^{\lambda_U (t-s)} - 1}{\lambda_U^2} \right] + \frac{\gamma_U (t-s)}{\gamma^2_U + \omega_U^2} \right). \]

The cross-covariance of \( T \) and \( V' \) is computed via
\[ \langle V'(t) T_k(t) \rangle = -\alpha \int_{-\infty}^{t} D_k(s, t) e^{-ik J(s,t)} F_k(s) \langle V'(t) \rangle e^{-ik J(s,t)} ds, \]
where
\[ \langle V'(t) \rangle e^{-ik J(s,t)} = i k \frac{\sigma^2}{4 \gamma U \lambda_U} \left( 1 - e^{\lambda_U (t-s)} \right) e^{-\frac{i \xi}{2} \text{Var}(J(s,t))}. \]
The cross-covariance of $T$ and $v$ is computed by

$$
\text{Cov}(v_k(t), T_l(t)) = -\sigma_k^2 2\gamma_k \int_{-\infty}^{t} D_k(s,t) e^{i\lambda_k(t-s)} e^{-\frac{1}{2\gamma_k} \text{Var}(J(s,t))} ds.
$$

Finally, for the covariance of $T$ we find

$$
\text{Cov}(T_k(t), T_l(t)) = \alpha^2 \int_{-\infty}^{t} \int_{-\infty}^{t} D_k(s,t) D_l(r,t) e^{-i(kJ(s,t)-lJ(r,t))} \langle v_k(s)v_l^*(r) e^{-i(kJ(s,t)-lJ(r,t))} \rangle dr ds,
$$

where

$$
\langle v_k(s)v_l^*(r) e^{-i(kJ(s,t)-lJ(r,t))} \rangle = \left( \sigma_k^2 2\gamma_k \right) e^{-\gamma_k(s+r-2 \min(s,r)) + i\omega_k(s-r)} + F_k(s) F_l^*(r)
\times e^{-\frac{1}{2\gamma_k} \text{Var}(J(s,t)) - \frac{1}{2\gamma_k} \text{Var}(J(r,t)) + kl \text{Cov}(J(s,t), J(r,t))},
$$

and for $s > r$

$$
\text{Cov}(J(s,t), J(r,t)) = \text{Var}(J(s,t)) - \frac{\sigma_U^2}{4\gamma_U} \text{Re} \left[ \frac{1}{\lambda_U} (e^{\lambda_U(t-s)} - e^{\lambda_U(t-r)} - 1 + e^{\lambda_U(s-r)}) \right].
$$

References


<table>
<thead>
<tr>
<th>$\gamma_U$</th>
<th>$\omega_U$</th>
<th>$\sigma_U$</th>
<th>$U_0$</th>
<th>$A_U$</th>
<th>$\eta$</th>
<th>$c$</th>
<th>$\beta$</th>
<th>$F_s$</th>
<th>$\nu, \mu$</th>
<th>$d_0$</th>
<th>$\gamma$</th>
<th>$\kappa$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04</td>
<td>1</td>
<td>0.3</td>
<td>1.7</td>
<td>0.5</td>
<td>2</td>
<td>1</td>
<td>8.91</td>
<td>16</td>
<td>0.002</td>
<td>0.03</td>
<td>0.04</td>
<td>0.001</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Parameters for the model (14), (16), (6).
Figure 1: Panel (a): spectrum of the waves, $v_k$, described by (23) with $k_0 = 10$; panel (b): spectrum of the tracer, $T_k$, for the case of nondispersive waves; panel (c): spectrum of the tracer, $T_k$, for the case of Rossby waves.
Figure 2: Real part of the autocorrelation functions of $U(t)$ (dashed line), $v_k(t)$ (thin solid line), and $T_k(t)$ (thick solid line) for the first six Fourier modes for the case of nondispersive waves given by (19)
Figure 3: Correlation time of the velocity field $U(t)$, $v_k(t)$ (circles) and the tracer $T_k(t)$ (crosses) for different Fourier modes for the case of nondispersive waves given by (19). Note that the correlation time for the cross-sweep $U(t)$ is shown at $k = 0$. 

34
Figure 4: Real part of the autocorrelation functions of $U(t)$ (dashed line), $v_k(t)$ (thin solid line), and $T_k(t)$ (thick solid line) for the first six Fourier modes for the case of Rossby waves given by (20)
Figure 5: Correlation time of the velocity field $U(t)$, $v_k(t)$ (circles) and the tracer $T_k(t)$ (crosses) for different Fourier modes for the case of Rossby waves given by (20). Note that the correlation time for the cross-sweep $U(t)$ is shown at $k = 0$. 
Figure 6: Four panels on the top demonstrate the pdfs with long exponential tails of the first four Fourier modes of the tracer $T_k$ (solid line) and the Gaussian pdfs with the same mean and variance (dashed line) as the original pdfs. The lower panel shows the pdf of the tracer in the physical space. The case of nondispersive waves given by (19) is shown. Note the logarithmic scale of the y-axis. The pdfs of the real parts only are shown.
Figure 7: Four panels on the top demonstrate the pdfs with long exponential tails of the first four Fourier modes of the tracer $T_k$ (solid line) and the Gaussian pdfs with the same mean and variance (dashed line) as the original pdfs. The lower panel shows the pdf of the tracer in the physical space. The case of Rossby waves given by \((20)\) is shown. Note the logarithmic scale of the y-axis. The pdfs of the real parts only are shown.
Figure 8: A segment of a typical trajectory for the Rossby waves case as a function of time: the truth signal (solid line), the observations (circles), the prior values (pluses), and the posterior values (dashed line). Panel (a) shows the cross-sweep (real part of the fluctuations, $V'(t)$, is shown), panel (b) shows the waves, $v(x_{ob}, t)$, at a fixed spatial location $x_{ob}$ where observations are available, panel (c) is the same as panel (b) but for the tracer $T(x_{ob}, t)$, panel (d) shows the the waves, $v(x_{no\,ob}, t)$, at a fixed spatial location $x_{ob}$ where observations are not available. Panel (e) is the same as panel (d) but for the tracer $T(x_{no\,ob}, t)$. Observations of the cross-sweep and sparse observations of both $v(x, t)$ and $T(x, t)$ with $p = 3$ were used.
Figure 9: Panel (a): a segment of a typical trajectory, $v(x,t_0)$ for the Rossby waves case as a function of space for a fixed time $t_0 = 33.375$: the truth signal (solid line), the observations (circles), the prior values (pluses), and the posterior values (dashed line). Panel (b) is the same as panel (a) but for the tracer, $T(x,t_0)$. Panel (c) shows the pattern correlation as a function of space between the truth signal, $v(x,t)$, and the corresponding posterior values. Panel (d) is the same as panel (c) but for the tracer, $T(x,t_0)$. Observations of the cross-sweep and sparse observations of both $v(x,t)$ and $T(x,t)$ with $p = 3$ were used.
Figure 10: Spectrum recovery via the NEKF for the nondispersive waves case and observations of \( v \) only (first row), \( T \) only (second row), and \( U \) and \( v \) (third row): the truth (thick solid line), the filtered signal with plentiful observations, i.e., \( p = 1 \) (thin solid line), the filtered signal with every 3d point observed (dashed line), and the filtered signal with every 7th point observed (dotted line). Left column shows the spectrum of the waves, \( v_k \), right column shows the spectrum of the tracer, \( T_k \).
Figure 11: Spectrum recovery via the NEKF for the nondispersive waves case and observations of $U$ and $T$ (first row), $v$ and $T$ (second row), and $U$, $v$, and $T$ (third row): the truth (thick solid line), the filtered signal with plentiful observations, i.e., $p = 1$ (thin solid line), the filtered signal with every 3d point observed (dashed line), and the filtered signal with every 7th point observed (dotted line). Left column shows the spectrum of the waves, $v_k$, right column shows the spectrum of the tracer, $T_k$. 
Figure 12: Spectrum recovery via the NEKF for the Rossby waves case and observations of $v$ only (first row), $T$ only (second row), and $U$ and $v$ (third row): the truth (thick solid line), the filtered signal with plentiful observations, i.e., $p = 1$ (thin solid line), the filtered signal with every 3d point observed (dashed line), and the filtered signal with every 7th point observed (dotted line). Left column shows the spectrum of the waves, $v_k$, right column shows the spectrum of the tracer, $T_k$. 
Figure 13: Spectrum recovery via the NEKF for the Rossby waves case and observations of $U$ and $T$ (first row), $v$ and $T$ (second row), and $U$, $v$, and $T$ (third row): the truth (thick solid line), the filtered signal with plentiful observations, i.e., $p = 1$ (thin solid line), the filtered signal with every 3rd point observed (dashed line), and the filtered signal with every 7th point observed (dotted line). Left column shows the spectrum of the waves, $v_k$, right column shows the spectrum of the tracer, $T_k$. 44
Figure 14: Recovery of the pdf of the tracer in physical space via NEKF for the Rossby waves case. Thick solid line shows the ideal pdf of the tracer in physical space as in Fig. 7, dashed line shows the Gaussian with the same mean and variance, thin solid line shows the normalized histogram of the sample trajectory which is filtered, crosses show the normalized histogram of the filtered signal. The observed variables are shown to the right of the corresponding panel, sparse observations with $p = 7$ were used. Note the logarithmic scale of the y-axis.
Figure 15: Improvement of the filter skill in physical space by adding an observation at just one spatial location for the case of nondispersive waves. Top row shows the cross-correlation between the truth and the filtered signals, bottom row shows the percentage RMS error of the filtered signal, left column corresponds to the waves, right column corresponds to the tracer. The solid line corresponds to the skill with the observations of the cross-sweep only, the dashed line corresponds to the skill with observations of the cross-sweep and the waves at one spatial location (denoted by a thick dot), the dotted line corresponds to the skill with observations of the cross-sweep and the tracer at one spatial location, the dash-dotted line corresponds to the skill with observations of the cross-sweep and the waves and the tracer at one spatial location.
Figure 16: Improvement of the filter skill in recovering the turbulent spectrum by adding an observation at just one spatial location for the case of nondispersive waves. Top panel shows the spectrum of the waves, bottom panel shows the spectrum of the tracer. Thick solid line corresponds to the truth spectrum, thin solid line corresponds to the filtered spectrum with observations of the cross-sweep only; dashed line corresponds to the filtered spectrum with observations of the cross-sweep and the waves at one spatial location only; dotted line corresponds to the filtered signal with observations of the cross-sweep and the tracer at one spatial location only; dashed-dotted line corresponds to the observations of the cross-sweep and the waves and the tracer at one spatial location only.
Figure 17: Improvement of the filter skill in physical space by adding an observation at just one spatial location for the case of Rossby waves. Top row shows the cross-correlation between the truth and the filtered signals, bottom row shows the percentage RMS error of the filtered signal, left column corresponds to the waves, right column corresponds to the tracer. The solid line corresponds to the skill with the observations of the cross-sweep only, the dashed line corresponds to the skill with observations of the cross-sweep and the waves at one spatial location (denoted by a thick dot), the dotted line corresponds to the skill with observations of the cross-sweep and the tracer at one spatial location, the dash-dotted line corresponds to the skill with observations of the cross-sweep and the waves and the tracer at one spatial location.
Figure 18: Improvement of the filter skill in recovering the turbulent spectrum by adding an observation at just one spatial location for the case of Rossby waves. Top panel shows the spectrum of the waves, bottom panel shows the spectrum of the tracer. Thick solid line corresponds to the truth spectrum, thin solid line corresponds to the filtered spectrum with observations of the cross-sweep only, dashed line corresponds to the filtered spectrum with observations of the cross-sweep and the waves at one spatial location only, dotted line corresponds to the filtered signal with observations of the cross-sweep and the tracer at one spatial location only, dashed-dotted line corresponds to the observations of the cross-sweep and the waves and the tracer at one spatial location only.
Figure 19: Improving recovery of the pdf of the tracer in physical space via NEKF for the nondispersive waves case by adding an observation at just one spatial location. Thick solid line shows the ideal pdf of the tracer in physical space as in Fig. 7, dashed line shows the Gaussian with the same mean and variance, thin solid line shows the normalized histogram of the sample trajectory which is filtered, crosses show the normalized histogram of the filtered signal. Top left panel corresponds to observations of the cross-sweep only, top right panel corresponds to observations of the cross-sweep and the waves at one spatial location, bottom left panel corresponds to observations of the cross-sweep and the tracer at one spatial location only, bottom right panel corresponds to observations of the cross-sweep and the waves and the tracer at one spatial location only. Note the logarithmic scale of the y-axis.
Figure 20: Improving recovery of the pdf of the tracer in physical space via NEKF for the Rossby waves case by adding an observation at just one spatial location. Thick solid line shows the ideal pdf of the tracer in physical space as in Fig. 7, dashed line shows the Gaussian with the same mean and variance, thin solid line shows the normalized histogram of the sample trajectory which is filtered, crosses show the normalized histogram of the filtered signal. Top left panel corresponds to observations of the cross-sweep only, top right panel corresponds to observations of the cross-sweep and the waves at one spatial location, bottom left panel corresponds to observations of the cross-sweep and the tracer at one spatial location only, bottom right panel corresponds to observations of the cross-sweep and the waves and the tracer at one spatial location only. Note the logarithmic scale of the y-axis.
Figure 21: RMSE of the filtered cross-sweep, $U$, as a function of the observation noise variance, $r_o$, for various values of observation time, $\Delta t$, different observed variables, and different sparsity of observations. The case of Rossby waves with $2K + 1 = 21$ spatial location is shown. The first column corresponds to the observations at each spatial location ($p = 1$), the second column corresponds to the observations at every 3d spatial location ($p = 3$), and the third column corresponds to the observations at every 7th spatial location ($p = 7$). Each row corresponds to a different set of observed variables that are shown at the right of a corresponding column. Different line types correspond to different observation times, see the legend in the first panel for explanation. Note the logarithmic scale of the x-axis and the floating scale of the y-axis.
Figure 22: RMSE of the filtered waves, $v$, as a function of the observation noise variance, $r^2$, for various values of observation time, $\Delta t$, different observed variables, and different sparsity of observations. The case of Rossby waves with $2K + 1 = 21$ spatial location is shown. The first column corresponds to the observations at each spatial location ($p = 1$), the second column corresponds to the observations at every 3rd spatial location ($p = 3$), and the third column corresponds to the observations at every 7th spatial location ($p = 7$). Each row corresponds to a different set of observed variables that are shown at the right of a corresponding column. Different line types correspond to different observation times, see the legend in the first panel of Fig. 21 for explanation. Note the logarithmic scale of the $x$-axis and the floating scale of the $y$-axis.
Figure 23: RMSE of the filtered tracer, $T$, as a function of the observation noise variance, $r_o$, for various values of observation time, $\Delta t$, different observed variables, and different sparsity of observations. The case of Rossby waves with $2K+1 = 21$ spatial location is shown. The first column corresponds to the observations at each spatial location ($p = 1$), the second column corresponds to the observations at every 3rd spatial location ($p = 3$), and the third column corresponds to the observations at every 7th spatial location ($p = 7$). Each row corresponds to a different set of observed variables that are shown at the right of a corresponding column. Different line types correspond to different observation times, see the legend in the first panel of Fig. 21 for explanation. Note the logarithmic scale of the x-axis and the floating scale of the y-axis.