Model Error in Filtering Random Compressible Flows Utilizing Noisy Lagrangian Tracers

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

ANDREW J. MAJDA
Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, New York, and Center for Prototype Climate Modeling, NYU Abu Dhabi, Saadiyat Island, Abu Dhabi, United Arab Emirates

(Manuscript received 17 December 2015, in final form 21 April 2016)

ABSTRACT

Lagrangian tracers are drifters and floaters that collect real-time information of fluid flows. This paper studies the model error in filtering multiscale random rotating compressible flow fields utilizing noisy Lagrangian tracers. The random flow fields are defined through random amplitudes of Fourier eigenmodes of the rotating shallow-water equations that contain both incompressible geostrophically balanced (GB) flows and rotating compressible gravity waves, where filtering the slow-varying GB flows is of primary concern. Despite the inherent nonlinearity in the observations with mixed GB and gravity modes, there are closed analytical formulas for filtering the underlying flows. Besides the full optimal filter, two practical imperfect filters are proposed. An information-theoretic framework is developed for assessing the model error in the imperfect filters, which can apply to a single realization of the observations. All the filters are comparably skillful in a fast rotation regime ($\varepsilon < 0.5$). In a moderate rotation regime ($\varepsilon < 1$), significant model errors are found in the reduced filter containing only GB forecast model, while the computationally efficient 3D-Var filter with a diagonal covariance matrix remains skillful. First linear then nonlinear coupling of GB and gravity modes is introduced in the random Fourier amplitudes, while linear forecast models are retained to ensure the filter estimates have closed analytical expressions. All the filters remain skillful in the $\varepsilon < 0.5$ regime. In the $\varepsilon = 1$ regime, the full filter with a linear forecast model has an acceptable filtering skill, while large model errors are shown in the other two imperfect filters.

1. Introduction

Lagrangian tracers are drifters and floaters that collect real-time information of fluid flows, especially at the center of oceans where Eulerian measurements are inaccessible (Griffa et al. 2007; Gould et al. 2004). An important application of Lagrangian data is to recover the current underlying velocity field. To this end, many approximate filters have been developed for assimilation of Lagrangian data (Molcard et al. 2003; Kuznetsov et al. 2003; Apte et al. 2008) and the properties of these filters are studied through numerical experiments (Salman et al. 2006, 2008; Slivinski et al. 2015).

However, because of the complexity and highly nonlinear nature of Lagrangian data assimilation, there was little systematic analysis of the approximate filters based on rigorous theory. Recently, an analytically tractable nonlinear filtering framework for Lagrangian data assimilation was developed (Chen et al. 2014b, 2015), which allows for the study of random incompressible/compressible flow field with full mathematical rigor. In this framework, the turbulent flow field is defined through a finite number of random Fourier modes that are coupled through the tracer observations in a highly nonlinear way.

Denotes Open Access content.

Corresponding author address: Nan Chen, Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012.
E-mail: chennan@cims.nyu.edu

DOI: 10.1175/MWR-D-15-0438.1
Despite the inherent nonlinearity in measuring noisy Lagrangian tracers, the signal–observation process forms a conditional Gaussian system, which allows exact closed analytical formulas for the optimal filter in filtering the velocity field. In Chen et al. (2014b), this Lagrangian data assimilation framework was applied to random incompressible flows, where a practical information barrier in increasing the number of tracers was revealed. In Chen et al. (2015), the filtering framework was applied to a multiscale random compressible flow field that is a linear combination of random incompressible geostrophically balanced (GB) flows and random rotating compressible gravity waves. Rigorous theorems through suitable stochastic fast-wave-averaging techniques and explicit formulas demonstrated that a practical suboptimal filter with mode reduction is as skillful as the full filter in recovering the slow GB flows in the limit of small Rossby number $\varepsilon \rightarrow 0$ for any bounded time interval (Chen et al. 2015).

Since simplifications and approximations are ubiquitous in designing filters, a central practical issue is to understand the model error by utilizing imperfect filters for assimilation of Lagrangian data (Majda 2012; Majda and Harlim 2012). This requires assessing the lack of information in the filter estimate utilizing imperfect filters related to that utilizing perfect one. Yet, despite the application of recursive Bayesian estimation in Lagrangian data assimilation, the filtering skill in the previous works (Salman et al. 2006, 2008; Slivinski et al. 2015; Chen et al. 2014b, 2015) was evaluated mostly based on the pathwise RMS error in the posterior mean estimation, where the uncertainty represented by the posterior covariance was completely ignored. Clearly, a moderate error in the posterior mean estimation utilizing imperfect filters with a tiny posterior covariance is of particular danger since it implies that the biased estimation is falsely trusted with high certainty. Likewise, a strongly overestimated posterior covariance utilizing imperfect filters provides little information even if the posterior mean is quite close to that utilizing the perfect one. Therefore, it is important to develop a systematic framework for assessing the model error in imperfect filters based on the lack of information in the full posterior distribution. For clarification, the terminology “model error” below always refers to the error in the posterior distribution.

Below, an information-theoretic framework (Branicki et al. 2013; Majda and Wang 2006; Majda and Branicki 2012; Branicki and Majda 2014) is developed to assess the model error in imperfect filters for filtering the multiscale random rotating compressible flows, which can apply to a single realization of the observations. The lack of information in the posterior distribution utilizing imperfect filters related to that utilizing perfect filter is measured through an information metric called the relative entropy (Majda and Wang 2006; Majda et al. 2002), which takes into account not only the error in the mean state estimation, but the uncertainty in the filter estimates as well.

Following the general nonlinear filtering framework (Chen et al. 2014b, 2015), the idealized flow fields of the multiscale random rotating compressible flows studied here are defined through random amplitudes of Fourier eigenmodes of the rotating shallow-water equations, which involve both the incompressible GB flows and the rotating compressible gravity waves. To ensure the filter estimates of the perfect full filter having closed analytic expressions that facilitates the study of the model error in the information-theoretic framework, linear and independent stochastic dynamics are adopted for the random amplitudes of different modes. These assumptions are often utilized in tests for Lagrangian data assimilation (Kuznetsov et al. 2003; Apte et al. 2008; Slivinski et al. 2015). Despite such decoupling in the true underlying flow fields and thus in the perfect forecast model, the GB and gravity modes are nevertheless coupled in a highly nonlinear way through the tracer observations. Note that many geophysical scenarios involve fast rotating flows, where the Rossby number $\varepsilon \ll 1$ (Vallis 2006). Thus, the random rotating shallow-water equations become a slow–fast system and the primary objective in practice is to recover the GB component that dominates the slow-varying geophysical flows (Rossby 1937; Gill 1982; Majda 2003; Cushman-Roisin and Beckers 2011) from the noisy Lagrangian tracer observations.

In addition to the full optimal filter, an idealized GB filter involving only the GB dynamics in the forecast model and artificial noisy observations associated with the GB flow is developed, serving as a reference for filtering the slow-varying GB flows (Chen et al. 2015). Two practical imperfect filters are proposed. First, formally applying the mode reduction (Majda et al. 2003, 1999) to the gravity waves results in a suboptimal filter that contains only the GB dynamics in the forecast model, while the noisy observations nevertheless include the coupled GB and gravity modes as in the perfect full filter. This dimension-reduction strategy in the forecast model simplifies the filter structure and saves the computational cost. Another practical reduced filter includes the full GB and gravity dynamics in the forecast model, but the posterior covariance is assumed to be diagonal. As will be shown, the special structure of such a reduced filter leads to a constant diagonal covariance matrix when the system arrives at its statistical steady state after a short time and, therefore, it becomes a 3D-Var type of filter (Navon...
Since this diagonal-reduced 3D-Var filter only requires the update of the posterior mean, it is computationally efficient. Below, the comparison of the filtering skill and the information model error by utilizing these two reduced imperfect filters will be extensively studied in different dynamical regimes.

Another central issue in this paper involves studying a more complicated and realistic flow field. Recall that the random Fourier amplitudes associated with the GB and gravity modes as discussed above are assumed to evolve independently with each other. Yet, both the mathematical theory of the slow–fast geophysical flows (Embid and Majda 1998; Majda 2003; Gershgorin and Majda 2008) and high resolution of turbulent simulations in slow–fast geophysical regimes (Smith 2001; Smith and Waleffe 2002; Waite and Bartello 2004) indicate the interactive effect between the GB and gravity modes. Therefore, following the theory in Embid and Majda (1998), Majda (2003) and Gershgorin and Majda (2008), a quadratic nonlinear interaction between the GB mode and the two gravity modes with the same Fourier wavenumber is incorporated into the underlying dynamics of the random amplitudes associated with the gravity modes, while the GB flow remains evolving independently.

However, the perfect filter including the nonlinear forecast model for the random Fourier amplitudes breaks the conditional Gaussian-filtering framework in Chen et al. (2014b, 2015). Thus, the same linear stochastic forecast models where different modes evolve independently as described above are utilized for filtering the nonlinearly coupled flow field, which ensure that the filter estimates have closed analytical expressions. Despite this intrinsic model error, such simplification is a common strategy for filtering large-dimensional turbulent systems in many more realistic issues, such as utilizing the extended Kalman filter (Haykin 2004) or adopting the mean stochastic forecast model in filtering (Majda and Harlim 2012; Harlim and Majda 2013). Note that the observational process here remains highly nonlinear and thus the coupled signal–observation system still forms a nonlinear filter. It is of practical importance to understand the effect of model error by dropping the nonlinear coupling between different modes in the forecast models for filtering the random rotating compressible flows with nonlinearly coupled GB and gravity modes in different dynamical regimes.

The remainder of this paper is organized as follows. In section 2, the multiscale random rotating compressible shallow-water flows are described and the analytically tractable nonlinear Lagrangian data assimilation framework is introduced. The description of the four filters is also included in the same section. In section 3, a general information-theoretic framework for assessing the model error in imperfect filters is developed. Section 4 starts with describing a simple setup of the GB flow field with diverse flow structures varying in time, which is followed by the filtering skill and information model error in filtering the multiscale random rotating compressible flows. Specifically, section 4c deals with the situation where the GB and gravity modes evolve independently, while section 4d handles the flow field where the underlying dynamics of the random Fourier coefficients contains the nonlinear interaction between GB and gravity modes. Section 5 contains a summary and the conclusions.

2. Basic setup

a. Random rotating compressible shallow-water flows

Motivated from the linearized shallow-water equations in the nondimensional form [section 4.4 of Majda (2003)], the two-dimensional (2D) random rotating compressible shallow-water flows are described in the following way:

\[
\begin{align*}
\bar{v}(x,t) &= \sum_{k \in K, \alpha \in \{0, \pm 1\}} \tilde{v}_{k,\alpha}(t) \exp(ik \cdot x) r_{k,\alpha}, \\
\bar{h}(x,t) &= \sum_{k \in K, \alpha \in \{0, \pm 1\}} \tilde{h}_{k,\alpha}(t) \exp(ik \cdot x) r_{k,\alpha},
\end{align*}
\]

where \(\bar{v}\) is the 2D velocity field and \(\bar{h}\) is the height function. In (1), \(K\) is some finite symmetric subset of \(\mathbb{Z}^2\) (i.e., the 2D integer lattice), while modes with \(\alpha = 0\) represent the GB part and modes with \(\alpha = \pm 1\) represent the two gravity waves. The vectors \(r_{k,\alpha}\) are the eigenvectors associated with different modes, where the projection of \(r_{k,0}\) on the velocity components is perpendicular to \(k\) due to the incompressibility of the GB part (Majda 2003; Embid and Majda 1998; Majda and Embid 1998) and \(r_{k,\pm}\) indicate the direction of the compressible gravity waves. The turbulent nature of the underlying flow field is reflected in the wave amplitudes \(\tilde{v}_{k,\alpha}(t)\) with stochastic forcing and damping terms as adopted in many applications (Majda and Harlim 2012; Chen et al. 2015):

\[
\begin{align*}
\dot{\tilde{v}}_{k,0}(t) &= \left[ -d_B \tilde{v}_{k,0}(t) + f_{k,0}(t) \right] dt + \sigma_{k,0} dW_{k,0}(t), \\
\dot{\tilde{v}}_{k,\pm}(t) &= \left[ \left( -d_s + i\alpha_k \pm iy_{k,0} \right) \tilde{v}_{k,\pm}(t) + f_{k,\pm}(t) \right] dt \\
&\quad + \sigma_{k,\pm} dW_{k,\pm}(t),
\end{align*}
\]

where the GB modes \(\tilde{v}_{k,0}\) are assumed to be real and the gravity modes \(\tilde{v}_{k,\pm}\) are complex. In (2), \(\alpha_{k,\pm}\) are the oscillation frequencies of the gravity modes, the
details of which will be given in (3); \(b_r, d_r > 0\) are damping coefficients; \(\sigma_{k,0}, \sigma_{k,1} > 0\) are stochastic forcing amplitudes and \(f_{k,0}, f_{k,1}\) are deterministic forcing; and \(W_{k,0}(t)\) and \(W_{k,1}(t)\) are independent Gaussian white noise with zero mean and variance \(t\).

To guarantee the full flow fields in (1) to be real valued, we require that \(\mathbf{r}_{k,0} = -\mathbf{r}_{-k,-0}\) and \(\mathbf{v}_{k,0} = -\mathbf{v}_{-k,-0}\). For a detailed description of this requirement, we refer to appendix A.1 of Chen et al. (2014b). Note that such a modeling strategy for random turbulence has been widely applied in many other situations (Majda and Harlim 2012). The effect of the slow GB mode on the fast gravity modes is reflected on the nonlinear coupling term with coefficient \(\gamma\) in (2b) (the term that is underlined), which is motivated directly from the mathematical theory of the slow–fast geophysical flows (Embid and Majda 1998; Majda 2003; Gershgorin and Majda 2008) and high resolution of turbulent simulations in slow–fast geophysical regimes (Smith and Waleffe 2002; Waite and Bartello 2004). The situation with \(\gamma = 0\) in (2b) implies utilizing the linear stochastic model to describe the random Fourier coefficients, where the GB and gravity modes are independent from each other. Despite the linear dynamics associated with each Fourier mode, the stochastic forcing and damping terms compensate the nonlinearity in nature in the sense that the small scales represented by the stochasticity interact with the large-scale modes and the full velocity field remains highly turbulent. Some pathwise behaviors of the situation with \(\gamma = 0\) were discussed in Chen et al. (2015). In this paper, the model error in the information-theoretic framework and the pathwise filtering skill of both linearly independent \((\gamma = 0)\) and nonlinearly coupled \((\gamma \neq 0)\) GB and gravity modes will be studied.

The connection between (1) and (2) and the linearized shallow-water equations and the choices of the eigenvectors \(\mathbf{r}_{k,\alpha}\) are described in detail in section 2 of Chen et al. (2015). Note that the modes with \(\alpha = 0\) represent the GB modes \(\omega_{k,0} = 0\), also known as the vortical waves, where the geostrophically balance relation holds. The modes with \(\alpha = \pm\) represent the gravity modes also known as the Poincaré waves (Majda 2003) with a nonzero phase speed of order \(\epsilon^{-1}\):

\[
\omega_{k,\alpha} = \epsilon^{-1} \sqrt{\delta |k|^2 + 1},
\]

where the nondimensional parameters \(\epsilon = Ro\) and \(\delta = Ro^2 Fr^{-2}\). Here \(Ro\) is the Rossby number representing the ratio between the Coriolis term and the advection term and \(Fr\) is the Froude number.

b. Observation process from noisy Lagrangian tracers

The observations are from the trajectories of \(L\) Lagrangian tracers, \(\mathbf{X}_i\), transported by the underlying velocity field with additional noise. The observation process is given by

\[
d\mathbf{X}_i(s) = \mathbf{v}[\mathbf{X}_i(s), s] ds + \sigma_s d\mathbf{W}^s_i(s)
= \sum_{k,\alpha \in \{0, \pm\}} \mathbf{v}_k \cdot \langle \mathbf{X}_i(s) \rangle_{\alpha} \mathbf{r}_{k,\alpha} ds
+ \sigma_s d\mathbf{W}^s_i(s), \quad l = 1, \ldots, L,
\]

where Newton’s law is applied in the first row of (4) and the second row is due to (1). The operator \(\mathcal{R}_s\) is the projection of a 3D vector on its first two dimension entries [i.e., the directions of the 2D velocity field as seen in (1)]. The noise amplitude \(\sigma_s\) in different tracers is assumed to be the same but the Gaussian white noise \(\mathbf{W}^s_i(s)\) (with zero mean and variance \(s\)) itself is independent for different \(l\), where the superscript \(s\) stands for the noise in the observed tracer trajectories.

c. Filters for noisy Lagrangian tracers

Given the observations from the noisy Lagrangian tracers in (4), the goal is to filter the underlying flow field \(\mathbf{v}(\mathbf{x}, t)\) in (1), or equivalently the Fourier coefficients \(\mathbf{v}_{k,\alpha}(t)\) for all \(k\) and \(\alpha\). For simplicity of notations, we define \(\mathbf{k} = \{k, \alpha\} \in \mathbb{K}\) such that the Fourier coefficient and the eigenvector in (1) can be written as \(\mathbf{v}_{k}(\mathbf{x}, t)\) and \(\mathbf{r}_k\), respectively. Furthermore, we define \(\mathbf{k}_0 = \{k, 0\} \in \mathbb{K}_0\) and \(\mathbf{k}_s = \{k, \pm\} \in \mathbb{K}_s\) representing the GB and gravity modes, respectively.

Recall that each trajectory of the noisy Lagrangian tracers is given by (4). Let us group all \(\mathbf{X}_i(s)\) into one \(1 \times 2L\) column vector \(\mathbf{X}\). Then the abstract form of the observation process for the \(L\) noisy Lagrangian tracers follows:

\[
d\mathbf{X} = \mathbf{P}_X(\mathbf{X}_i) \mathbf{U}_s ds + \sigma_s d\mathbf{W}_r,
\]

where \(\mathbf{W}_r^i\) is a \(2L \times 2L\) diagonal matrix and \(\mathbf{P}_X(\mathbf{X})\) is given by

\[
\mathbf{P}_X(\mathbf{X}) = \begin{bmatrix}
\mathbf{P}_X(\mathbf{X}_1(s)) & \cdots & \exp(i\mathbf{k} \cdot \mathbf{X}_1(s)) \mathbf{r}_k & \cdots \\
\vdots & \ddots & \vdots & \vdots \\
\mathbf{P}_X(\mathbf{X}_L(s)) & \cdots & \exp(i\mathbf{k} \cdot \mathbf{X}_L(s)) \mathbf{r}_k & \cdots \\
\end{bmatrix} := [\mathbf{P}_X^1(\mathbf{X}), \mathbf{P}_X^L(\mathbf{X})],
\]

according to (4). With a slight abuse of the notation, \(\mathbf{r}_k\) in (6) is the eigenvector that has only the first two entries corresponding to the 2D velocity directions.
On the other hand, by formally applying mode reduction over the gravity waves, it is possible to write down the simplified random flow field that contains only GB part of the flow. The corresponding noisy Lagrangian tracers transported by only the GB flow can be formally constructed as

\[ d\mathbf{X}_k^B = P_k^B(\mathbf{X}_k^B(t)) \mathbf{U}_k^B ds + \sigma_x d\mathbf{W}_k^B. \]  

(7)

Note that (7) is an artificial observation process since it is practically impossible to extract the component that corresponds to the random GB signals from the full noisy observations.

With the observation processes in (5) or (7), what remains is to design the forecast models in filters for the velocity field. Recall the dynamics associated with the true velocity field in (2).

In the situation with uncoupled GB and gravity modes (i.e., \( \gamma = 0 \)), the underlying dynamics of the Fourier coefficients for wavenumber \( k \) associated with the flow field \( \mathbf{v}(x, t) \) in (1) reduces to a linear stochastic system:

\[ \begin{aligned}
\dot{\mathbf{v}}_{k,0}(t) &= [-d_B \mathbf{v}_{k,0} + f_k(t)] dt + \sigma_{k,0} dW_{k,0}(t), \\
\dot{\mathbf{v}}_{k,\pm 1}(t) &= [-(d_g + i\omega_{k,\pm 1}) \mathbf{v}_{k,\pm 1} + f_k(t)] dt + \sigma_{k,\pm 1} dW_{k,\pm 1}(t).
\end{aligned} \]  

(8a, 8b)

As was done for the tracers, the Fourier coefficients for all the GB modes in (8a) and all the gravity modes in (8b) can be grouped into a \( 1 \times |K| \) and a \( 1 \times |2K| \) column vector, respectively. Then, the corresponding dynamics of the GB and gravity modes becomes

\[ \begin{aligned}
\dot{\mathbf{U}}_s^B &= -\Gamma^B \mathbf{U}_s^B ds + \mathbf{F}_s^B ds + \Sigma_s^B d\mathbf{W}_s^B(s), \\
\dot{\mathbf{U}}_s^g &= (\Gamma^g + i\Omega_s) \mathbf{U}_s^g ds + \mathbf{F}_s^g ds + \Sigma_s^g d\mathbf{W}_s^g(s),
\end{aligned} \]  

(9a, 9b)

where the superscript “\( B \)” stands for GB modes and “\( g \)” stands for gravity modes; and the joint form of (9a) and (9b) is given by

\[ \dot{\mathbf{U}}_s = -\Gamma \mathbf{U}_s ds + \mathbf{F}_s ds + \Sigma_u d\mathbf{W}_u(s), \]  

(10)

where \( \Omega_s \) in (9b) is a diagonal matrix and its \( k \)th entry \( (k \in K_s) \) is given by (3), and \( \Gamma \) in (10) involves both the damping \( \Gamma^B \), \( \Gamma^g \) and the oscillation frequency \( d\Omega_s \).

Utilizing the perfect dynamics of the underlying flow field in (10) as the forecast model in the filter, the joint observation–signal system in (5) and (10) becomes a conditional Gaussian system given the observations. For such a system with Gaussian initial conditions, the conditional distribution of the flow field given the observed noisy Lagrangian tracer trajectories, knowing as posterior distribution, is a Gaussian distribution where the evolutions of the conditional mean and conditional covariance have closed analytic formulas (Liptser and Shiryaev 2001). This provides an exact filter for recovering the underlying velocity field.

In the situation where the GB modes affect the gravity modes in a nonlinear way [i.e., \( \gamma \neq 0 \) in (2)], the perfect observation–signal system is no longer a conditional Gaussian system since given the observations the underlying dynamics (2) is a quadratic nonlinear system with non-Gaussian statistics, which breaks the analytically tractable filtering framework. As a result of the high dimensionality of the coupled signal–observation system, it is computationally unaffordable to solve the posterior distribution via direct numerical methods. Thus, in the appearance of the nonlinearly coupled GB and gravity modes (\( \gamma \neq 0 \)) in the true velocity field (2), the linear system in (8) is nevertheless utilized as the forecast model in the designed filters to maintain the analytically solvable feature. Despite this intrinsic model error, such simplification is a common strategy for filtering large-dimensional turbulent systems in many more realistic issues, where the linearized methods such as the extended Kalman filter (Haykin 2004) or the mean stochastic forecast model (Majda and Harlim 2012; Harlim and Majda 2013) are widely adopted. An important practical issue is to understand the effect of model error due to adopting linear stochastic forecast models with independent GB and gravity components to filter the random rotating compressible flows with nonlinearly coupled GB and gravity modes. Note that the tracer trajectories in (5) are transported by the true nonlinear dynamics (2), while the linear stochastic turbulent system in (8) is only utilized as the forecast model in the filters.

In the following, four different filters, which all belong to the conditional Gaussian framework, are designed. Note that approximations and simplifications are made not only in the underlying dynamics of the flow field (10), but in the observation process such as (7) as well. Therefore, in the following, we refer the combination of the underlying dynamics for the flow field and the observation process as the “filter.”

1) FULL FILTER WITH LINEAR FORECAST DYNAMICS

Utilizing the nonlinear observation process (5) and the linear dynamics with independent GB and gravity modes as the forecast model (9), the full filter with linear forecast dynamics is given by

\[ \begin{aligned}
\dot{\mathbf{X}}_s &= P_s(\mathbf{X}_s) \mathbf{U}_s ds + \sigma_x d\mathbf{W}_s, \\
\dot{\mathbf{U}}_s &= -\Gamma \mathbf{U}_s ds + \mathbf{F}_s ds + \Sigma_u d\mathbf{W}_u(s).
\end{aligned} \]  

(11)
The filter (11) is a perfect optimal filter if the underlying flow of the truth (2) is also linear, [i.e., \( \gamma = 0 \) in (2b)]. In such a case, we simply name (11) as the full filter. Otherwise (\( \gamma \neq 0 \)), model error comes from ignoring the nonlinear coupling of GB and gravity modes in (2b). The analytic solution of updating the posterior mean and posterior covariance of \( \mathbf{U} \) given \( \mathbf{X}_{s=t} \) is shown in Liptser and Shiryaev (2001) and is summarized in section 2.2.1 of Chen et al. (2015).

2) IDEALIZED GB FILTER

In many practical issues, the primary practical objective is to recover the GB component that dominates the slow-varying geophysical flows (Rossby 1937; Gill 1982; Majda 2003; Cushman-Roisin and Beckers 2011). To this end, an idealized GB filter is constructed based on the GB forecast model (9a) and the artificial observations from only the GB part of the flow (7):

\[
\begin{align*}
\frac{d \mathbf{X}^B_t}{dt} &= \mathbf{P}_X^B(\mathbf{X}^B_t) \mathbf{U}^B_t ds + \sigma_X^B d\mathbf{W}^B_t, \\
\frac{d \mathbf{U}^B_t}{dt} &= -\Gamma^B \mathbf{U}^B_t ds + \mathbf{F}^B_t ds + \Sigma^B_d d\mathbf{W}^B_t(s).
\end{align*}
\] (12)

This idealized GB filter (12) is a perfect filter regardless of the coupling coefficient \( \gamma \) in (2) as the nonlinearity in the underlying flow appears only in the gravity modes.

Since the underlying GB flow is incompressible, the properties of this idealized GB filter were well studied in Chen et al. (2014b). In addition, without being scrambled by the gravity modes, this perfect GB filter indicates the optimality of filtering the GB flow field. Thus, the results from this idealized GB flow are regarded as the reference for checking the filtering skill utilizing other filters. Note that, despite its optimality, the GB filter is not a practical filter because extracting the observations corresponding only to the random GB part of flow from the full noisy observations is impractical in real applications.

3) REDUCED FILTER WITH ONLY GB FORECAST MODEL THROUGH FULL OBSERVATIONS

Motivated from the idealized GB filter (12), a practical reduced filter for filtering GB part of the flow is formed by adopting only the GB dynamics (9a) as the forecast model, while the coupled GB and gravity observations from noisy Lagrangian tracers are utilized as the input in the observation process. This follows the formal application of the mode reduction strategy (Majda et al. 2003, 1999) to the gravity waves in the forecast model. For consistency, the corresponding dynamics of the observation process contains the modes associated with only the GB flow as well [i.e., replacing \( \mathbf{P}_X \) in (5) by \( \mathbf{P}_X^B \)]. Therefore, such reduced filter reads as

\[
\begin{align*}
\frac{d \mathbf{X}^B}{dt} &= \mathbf{P}_X^B(\mathbf{X}^B) \mathbf{U}^B ds + \sigma_X^B d\mathbf{W}^B, \\
\frac{d \mathbf{U}^B_t}{dt} &= -\Gamma^B \mathbf{U}^B ds + \mathbf{F}^B ds + \Sigma^B_d d\mathbf{W}^B(s).
\end{align*}
\] (13)

Since the gravity parts of the flow are dropped from the forecast model in (13), the dimension of the flow field \( \mathbf{U}^B_t \) in (13) is only 1/3 compared with \( \mathbf{U} \) of the full filter in (11), and in turn the number of the entries in the covariance matrix is only 1/9 of that associated with the perfect filter. Because most of the computational cost lies in the update of the posterior covariance, this reduced filter is more computational efficient than the full filter. Yet, the reduced filter (13) is only a suboptimal filter due to the model error from filtering only GB part of the flow through the full mixed observations.

4) DIAGONAL-REDUCED 3D-VAR FILTER

Another practical reduced filter includes both the GB dynamics and the linearized gravity dynamics in the forecast model (9), which are the same as the full filter (11), but the posterior cross covariance is assumed to stay zero and thus it reduces to a diagonal filter. Furthermore, as shown in appendix A, the diagonal entries in the posterior covariance associated with this diagonal-reduced filter converge to constant values after a short relaxation time and, therefore, only the update of the posterior mean is needed afterward. Because of the same behavior as the 3D-Var with a constant background error covariance (Navon 2009), this filter is named as a diagonal-reduced 3D-Var filter:

\[
\begin{align*}
\frac{d \mathbf{X}^B}{dt} &= \mathbf{P}_X^B(\mathbf{X}^B) \mathbf{U}^B ds + \sigma_X^B d\mathbf{W}^B, \\
\frac{d \mathbf{U}^B_t}{dt} &= -\Gamma^B \mathbf{U}^B ds + \mathbf{F}^B ds + \Sigma^B_d d\mathbf{W}^B(s).
\end{align*}
\] (14)

When the true underlying flow field is linear [i.e., \( \gamma = 0 \) in (2)], the only model error in the diagonal reduced 3D-Var filter (14) comes from the ignoring of the off-diagonal entries in the posterior covariance matrix. If the diagonal entries dominate the posterior covariance matrix, then a comparable filtering skill in the diagonal reduced 3D-Var filter (14) is expected as the full filter (11) but (14) is much more efficient. On the other hand, when \( \gamma \neq 0 \) in the true underlying flow fields (2), an extra model error in the diagonal reduced 3D-Var filter (14) comes from utilizing the linear forecast model for the gravity modes, which is the same as in the full filter (11).

3. An information-theoretic framework in assessing model error

As discussed in section 1, due to the inevitable approximations in real-world Lagrangian data assimilation,
it is of practical importance to assess and understand the model error by utilizing imperfect filters with various simplifications. Note that the traditional approach of measuring the filtering skill is based on the pathwise RMS error that takes into account only the pointwise information in the posterior mean, while the information in the posterior covariance that represents the uncertainty in the filter estimate is completely ignored. To assess the lack of information in the posterior distribution of imperfect filters, an information-theoretic framework is developed.

Information theory was widely adopted to measure the lack of information in filtering and prediction utilizing imperfect models (Majda and Gershgorin 2010, 2011b,a; Majda et al. 2002; Kleeman 2002). Recently, a systematic information-theoretic approach was developed in Branicki and Majda (2014) to quantify the statistical accuracy and present the optimality of Kalman filters with model error. Another important application of information theory is for model calibration in predicting the real-time indices of the Madden–Julian oscillation (Chen et al. 2014a; Chen and Majda 2015b,a), which shows the significant skill of capturing the extreme events that cannot be assessed by the pathwise RMS error and pattern correlation.

Here, the information model error is assessed through the relative entropy (Majda and Wang 2006; Majda et al. 2002):

\[
\mathcal{H}(p, q) = \int p \ln \frac{p}{q} \, dq
\]

(15)

which measures the lack of information in the probability distribution function (PDF) associated with the imperfect model \(q\) related to that of the perfect system \(p\). It is nonnegative with \(\mathcal{H} = 0\) only when \(p = q\) and is invariant under nonlinear changes of variables. Note that when both \(p = \mathcal{N}(\mathbf{m}_p, R_p)\) and \(q = \mathcal{N}(\mathbf{m}_q, R_q)\) are Gaussian, the relative entropy has the following closed form:

\[
\mathcal{H}(p, q) = \left[ \frac{1}{2} (\mathbf{m}_p - \mathbf{m}_q)^T R_q^{-1} (\mathbf{m}_p - \mathbf{m}_q) \right]
+ \frac{1}{2} \left[ \text{tr}(R_p R_q^{-1}) - N - \ln \det(R_p R_q^{-1}) \right],
\]

(16)

where \(N\) is the dimension of both the distributions. The first term in brackets in (16) is called the "signal," which measures the lack of information in the mean weighted by model covariance. The second term in brackets is called the "dispersion," which involves only the covariance ratio.

Now we develop an information-theoretic framework to measure the model error in imperfect filters. Consider a coupled system with variables \((\mathbf{u}_f, \mathbf{u}_I)\), where \(\mathbf{u}_f\) stands for observations and \(\mathbf{u}_I\) represents the variables for filtering. Let us denote \(p\) and \(p^M\) as the PDFs of the perfect and the imperfect models, respectively. In a typical scenario, the imperfect model is coarse grained and thus we assume the distribution \(p^M\) is formed only by the conditional moments up to \(L\). Let us further denote \(p_L\) as the PDF that is reconstructed utilizing the \(L\) conditional moments of the perfect model. Then the joint distributions regarding \(\mathbf{u}_f\) and \(\mathbf{u}_I\) can be written as

\[
p(\mathbf{u}_f, \mathbf{u}_I) = p(\mathbf{u}_f | \mathbf{u}_I) \pi(\mathbf{u}_I),
\]

\[
p_L(\mathbf{u}_f, \mathbf{u}_I) = p_L(\mathbf{u}_f | \mathbf{u}_I) \pi(\mathbf{u}_I),
\]

\[
p^M(\mathbf{u}_f, \mathbf{u}_I) = p^M_L(\mathbf{u}_f | \mathbf{u}_I) \pi^M(\mathbf{u}_I).
\]

According to (Branicki et al. 2013), the lack of information in the imperfect model related to the perfect one is given by

\[
\mathcal{H}(p(\mathbf{u}_f, \mathbf{u}_I), p^M(\mathbf{u}_f, \mathbf{u}_I)) = \mathcal{H}(p(\mathbf{u}_f, \mathbf{u}_I), p_L(\mathbf{u}_f, \mathbf{u}_I))
+ \mathcal{H}(p_L(\mathbf{u}_f, \mathbf{u}_I), p^M_L(\mathbf{u}_f | \mathbf{u}_I)),
\]

(17)

where the first term on the right-hand side of (17) is the intrinsic barrier that measures the lack of information in the perfect model due to the coarse-grained effect from the insufficient measurement and the second term is the model error where the imperfect model is compared with the perfect model that possesses the same number of moments. Direct calculation (Branicki et al. 2013) shows that

\[
\text{Model error} = \mathcal{H}(\pi(\mathbf{u}_I), \pi^M(\mathbf{u}_I))
+ \int \pi^M(\mathbf{u}_I) \mathcal{H}(p^M_L(\mathbf{u}_f | \mathbf{u}_I), \pi^M(\mathbf{u}_f | \mathbf{u}_I)) \, d\mathbf{u}_f.
\]

(18)

In filtering the state variables \(\mathbf{u}_I\), we assume the observations in the imperfect model \(\pi^M(\mathbf{u}_I | s \leq t)\) is the same as those in the perfect model \(\pi(\mathbf{u}_I | s \leq t)\). Therefore, the first term in the model error in (18) disappears and \(\pi^M(\mathbf{u}_I)\) in the second term is replaced by \(\pi(\mathbf{u}_I)\), which simplifies the model error in (18) (Branicki et al. 2013):

\[
\text{Model error} = \int \pi(\mathbf{u}_I) \mathcal{H}(p_L(\mathbf{u}_f | \mathbf{u}_I), p^M_L(\mathbf{u}_f | \mathbf{u}_I)) \, d\mathbf{u}_f.
\]

(19)

**Model error for a single realization of the observations**

In filtering the random compressible flow, only one single realization of the observational trajectory associated with each tracer \(\mathbf{u}_i^r(s \leq t), i = 1, \ldots, L\) is given. Thus, we simply need to assess the following model error:
that is, the error in the posterior distribution of the flow field given the observations from noisy Lagrangian tracers.

In section 4c, when the underlying flow field is generated from system (2) with decoupled GB and gravity modes (i.e., \( \gamma = 0 \)), the full filter (11) is a perfect filter. Since we have also assumed the observations in the two reduced filters (13) and (14) are the same as those in the full filter, (20) is applied to compute the model error at each time \( t \), where \( p_L^{\gamma} [u_{t}(t) \mid u_{t}(s)] \) is the posterior distribution of the perfect full filter (11) and \( p_L^{\gamma} [u_{t}(t) \mid u_{t}(s)] \) is that of one of the imperfect filters (13) or (14). Note that all the three filters are conditional Gaussian filters. Thus \( L = 2 \) in (20) and the model error are split into signal and dispersion as described in (16).

On the other hand, in section 4d, when the underlying flow field is generated from the system with nonlinearly coupled GB and gravity modes (i.e., \( \gamma \neq 0 \) in (2)), the full filter with linear dynamics in (11) is no longer a perfect filter. Two alternative approaches are applied to assess the model error in the imperfect filters. In the first method, we assess the model error in the posterior mean estimate of the imperfect filters compared with the true signal. Here, we adopt the general relative entropy equation (15), where \( p \) is the time-averaged PDF of the true signal and \( q \) is the time-averaged PDF associated with the posterior mean estimation from one of the imperfect filters. Although this model error measures the lack of information based only on the posterior mean, it is nevertheless different from the pathwise RMS error. In fact, this information metric takes into account the spread of both the posterior mean time series and the true signal. Therefore, it is able to quantify the skill of the imperfect filters in capturing the extreme events in the true signal, which is not accessible by the pathwise RMS error and pattern correlation (Chen and Majda 2015b). The second approach involves formally applying the posterior distribution of the idealized GB filter (12) to \( p_L \) in (20). Yet, since the observations in the GB filter are different from those in the three imperfect filters, this argument becomes only an approximation in assessing the model error in the filter estimates utilizing the imperfect filters related to that utilizing the perfect one within the information-theoretic framework developed in (20). In appendix B, we compare the information model error by utilizing either the full filter (11) or the GB filter (12) as the reference distribution \( p \) in (20) in the situation with \( \gamma = 0 \) to justify that the approximation error due to adopting GB filter as the reference filter is acceptable in studying the information model error of the imperfect filters in the dynamics regimes of interest.

4. Numerical experiments

a. Simple GB flow with time-varying flow structures

An interesting and realistic GB flow field involves time-varying flow structures. The simplest setup of such GB flow consists of five Fourier wavenumbers, where \( k = (0, 0), (\pm 1, 0), \), and \((0, \pm 1)\). Since the eigenvector corresponding to the GB mode \( k = (0, 0) \) has only nonzero entry in \( h \) direction [see section 2.1 of Chen et al. (2015)], the underlying GB flow is essentially driven by the four modes with \(|k| = 1\), that is,

\[
v = \sum_{|k| = 1} \hat{v}_k(t) \exp(ik \cdot x) \mathcal{P}_v r_k,
\]

where for notation simplicity we have dropped the subscript \(-1\) that distinguishes GB flows from gravity waves. To look at the flow structures of the GB field in (21), we write down the eigenvectors [see section 2.1 of Chen et al. (2015)] projected on the horizontal and vertical velocity directions:

\[
\mathcal{P}_v r_{(1,0)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ i \end{pmatrix}, \quad \mathcal{P}_v r_{(0,1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \end{pmatrix}.
\]

In addition, the four Fourier bases in (21) with \(|k| = 1\) are given by

\[
\begin{align*}
k &= (1, 0): \exp(ix) &= \cos(x) + i \sin(x), \\
& k &= (1, -1): \exp(-ix) &= \cos(x) - i \sin(x), \\
& k &= (0, 1): \exp(iy) &= \cos(y) + i \sin(y), \\
& k &= (0, -1): \exp(-iy) &= \cos(y) - i \sin(y).
\end{align*}
\]

Inserting (22) and (23) into (21), the horizontal and vertical velocities \((v_1, v_2)\) are given by

\[
\sqrt{2} v_1 = \hat{v}_{(1,0)} \times (-i) \times [\cos(y) + i \sin(y)] + \hat{v}_{(-1,0)} \times (i) \times [\cos(y) - i \sin(y)]
= \hat{v}_{(1,0)} [i \cos(y) + \sin(y)] + \hat{v}_{(-1,0)} [i \cos(y) + \sin(y)],
\]

\[
\sqrt{2} v_2 = \hat{v}_{(1,0)} \times (i) \times [\cos(x) + i \sin(x)] + \hat{v}_{(-1,0)} \times (-i) \times [\cos(x) - i \sin(x)]
= \hat{v}_{(1,0)} [i \cos(x) - \sin(x)] + \hat{v}_{(-1,0)} [i \cos(x) - \sin(x)].
\]
Since the Fourier coefficients associated with the GB modes are assumed to be real, we have \( v_{(0,1)} = v_{(0,-1)} \) and \( \tilde{v}_{(1,0)} = \tilde{v}_{(-1,0)} \), which simplify (24),

\[
\begin{align*}
v_1 &= \sqrt{2}\tilde{v}_{(0,1)} \sin(y), \\
v_2 &= -\sqrt{2}\tilde{v}_{(1,0)} \sin(x),
\end{align*}
\]

and the corresponding streamfunction is given by

\[
\psi = -\sqrt{2}\tilde{v}_{(1,0)} \cos(x) + \sqrt{2}\tilde{v}_{(0,1)} \cos(y).
\]

Thus, we only need to look at the amplitude of the two coefficients to determine the structure of the GB flow. With different choices of \( v_{(0,1)} \) and \( \tilde{v}_{(1,0)} \), the streamlines illustrate various profiles that switch between

1) simple shear flow: \( v_{(0,1)} \ll 1, \tilde{v}_{(1,0)} \sim O(1) \) or \( v_{(0,1)} \ll 1, \tilde{v}_{(1,0)} \sim O(1) \),
2) 2D array of swirling eddies: \( v_{(0,1)} \approx \tilde{v}_{(1,0)} \sim O(1) \), and
3) swirling eddies embedded in a shear-flow stream: \( v_{(0,1)} \sim O(1), \tilde{v}_{(1,0)} \sim O(1) \) but \( v_{(0,1)} \approx \tilde{v}_{(1,0)} \).


b. Filter setup

As in section 4a, the underlying flow field contains five Fourier wavenumbers with \(|k| \leq 1\). Thus, the total number of GB and gravity modes is \(|K_6| = 5\) and \(|K_g| = 10\), respectively. In most realistic situations, the number of the observations is typically less than the degree of freedom of the underlying system. Thus, we set the number of the tracers \( L = 5 < 15 = |K| \). The observation noise level is set to be \( \sigma_s = 0.2 \), which is a moderate value, implying that the filters make use of the information in both the forecast models and the observations.

The GB mode at the largest scale \( k = (0, 0) \) is set to be deterministic while the other 4 GB modes with \(|k| = 1\) and all the 10 gravity modes are stochastic. The damping and stochastic forcing coefficients are determined in the situation with uncoupled GB and gravity modes [i.e., \( \gamma = 0 \) in (2)], and the same values are adopted in the coupled case. The energy in each stochastic GB mode is set to be \( E_B = 0.3 \) and that in each gravity mode is \( E_g = 0.1 \). A relatively small damping \( \beta_B = \beta_g = 0.05 \) is utilized for all the stochastic modes, which correspond to a moderately long decorrelation time \( \tau = 20 \) non-dimensional units in the uncoupled flow case. The stochastic forcing in each GB mode is computed by utilizing the formula \( \sigma_k^2 \mu/(2\beta_B) = E_B \) and similar for that in each gravity mode.

In addition to the typical values mentioned above, the filtering skill dependence on different parameters is of particular interest. Below, the filtering dependence on the number of tracers \( L \), the observation noise \( \sigma_s \), and the energy in the gravity modes \( E_g \) will be explored. In each experiment, only one parameter is varied and the others are all set to be their typical values.

The deterministic forcings are chosen in two different ways:

1) Zero deterministic forcing: The flow is purely driven by the stochastic forcing, which makes it possible to study the effect of the random forcing in changing the underlying flow structures.

2) Time-periodic deterministic forcing:

\[
\text{GB modes: } f_{k,0} = a_{k,0} \cos(\phi t) + b_{k,0},
\]

\[
\text{Gravity modes: } f_{k,\pm} = a_{k,\pm} \exp(i\phi t),
\]

where \( a_{k,0} = \sqrt{3}/10 \) and \( b_{k,0} = \sqrt{3}/20 \) for mode \((0, 0)\);
\( a_{k,0} = \sqrt{3}/10 \) and \( b_{k,0} = \sqrt{3}/200 \) for modes \((\pm 1, 0)\);
\( a_{k,0} = -\sqrt{3}/10 \) and \( b_{k,0} = \sqrt{3}/200 \) for modes \((0, \pm 1)\);
and \( a_{k,\pm} = 1/10 \) for all gravity modes. The frequency \( \phi = 20 \). The amplitudes of these large-scale deterministic forcing and stochastic forcing are comparable. This setup implies the flow field has a large-scale background mean and a random part. The flow structure is able to switch between nearly straight streamlines and swirling eddies according to (26). Comparing the two situations helps us understand the effect of the deterministic mean flow on the filtering skill.

For the initialization of the filters, the states of all the stochastic modes are set to be consistent with the value at their statistical equilibrium associated with the forecast models, where the initial uncertainty of the stochastic modes is 0.3 and 0.1 for each GB and gravity mode, respectively.

The tracer trajectories \( X \), utilized in the full filter (11) and the two reduced filters (13) and (14) are identical. On the other hand, the trajectories \( X^B \) in (7) for the GB filter (12) are based only on the GB part of the flow and, therefore, they are different from those in (5). For the sake of comparing the filtering skill, we impose the same observation noise process in (5) and (7) (i.e., \( \mathbf{W}_s = \mathbf{W}_p \)). Furthermore, the initial locations of the tracers utilized in both the full filter and GB filter are the same and are distributed uniformly in the periodic domain \( \mathbb{T}^2 = [-\pi, \pi]^2 \).

Two dynamical regimes are considered. The first one is a fast rotation regime with small Rossby number \( \varepsilon = 0.1 \), which mimics the motion in the midlatitude atmosphere or ocean (Majda 2003). Another dynamical regime involves moderate rotation with \( \varepsilon = 1 \). Note that the GB flow is kept to be the same in both regimes.
and the only difference lies in the gravity waves according to the rotation frequency and eigenmodes in section 2a and in Chen et al. (2015). The non-dimensional number \(d = 1\) is fixed, which implies that the Rossby and the Froude number are equal with each other. Below, the filtering behavior up to a long time \(t = 200\) is studied.

c. Results for filtering the random flow fields with uncoupled GB and gravity modes

In this subsection, we study the situation where the random GB and gravity modes evolve independently [i.e., \(\gamma = 0\) in (2)]. Thus, the underlying dynamics of the velocity field for Fourier wavenumber \(k\) of nature is given by

\[
d_\gamma \hat{u}_{k,0}(t) = \left[ -R \hat{u}_{k,0} + f_{k,0}(t) \right] dt + \sigma_{k,0} dW_{k,0}(t),
\]

(28a)

\[
d_\gamma \hat{u}_{k,\pm}(t) = \left[ -d_\gamma + i \omega_{k,\pm} \right] \hat{u}_{k,\pm}(t) + f_{k,\pm}(t) \right] dt + \sigma_{k,\pm} dW_{k,\pm}(t).
\]

(28b)

Since the true dynamics (28) and the forecast model (8) in the full filter (11) are the same for all \(k\), the full filter becomes a perfect filter.

First, we look at the tracer behaviors. Figure 1a includes the comparison of the tracer trajectories utilizing the full filter (11) and the GB filter (12) at an initial period from \(t = 0\) to \(t = 10\) in the two dynamics regimes with different \(\varepsilon\), where the large-scale deterministic forcing is set to be zero. For conciseness, only one of the five tracers associated with each filter is shown. The two trajectories starting at the same location almost overlap with each other during this short initial period in \(\varepsilon = 0.1\) regime while the two trajectories diverge quickly in the \(\varepsilon = 1\) regime. Comparing the snapshot of the GB flow (column III) with the full flow (columns I and II) at \(t = 10\), it is clear that the gravity waves have nonnegligible contributions to the total flow at each time instant. Fortunately, because of the fast oscillation nature of the gravity waves in \(\varepsilon = 0.1\) regime, the effect of the gravity waves is averaged out and, therefore, the two trajectories align with each other. Figure 1b is similar to Fig. 1a, but the time-periodic deterministic forcing in the underlying flow in (28) is nonzero as described in (27) in section 4b and the initial period shown is shortened up to \(t = 7\). The same phenomenon is found in Fig. 1b in the two different dynamics regimes, despite the fact that the tracers move faster due to the deterministic background flow velocity. We have also found that the deterministic forcing has no effect on the RMS error and the uncertainty in the filter estimates. In Figs. 1c and 1d, we compare the posterior variance for GB mode (1, 0) as a function of time up to \(t = 25\). The difference by adopting different deterministic forcing is insignificant. Yet, it is obvious that the
relaxation time of the posterior variance toward the statistical equilibrium state is longer in the \( \varepsilon = 1 \) regime. Since the large-scale deterministic forcing only affects the tracer speed while it has little influence on the filtering skill, below we focus on the situation with no large-scale deterministic forcing.

Next, we study the long-term behavior of the tracers’ distribution. In Fig. 2 we show the distributions of the tracers utilizing the full filter (11) and the GB filter (12) in the two dynamical regimes at \( t = 199 \). In addition to showing the distribution with \( L = 5 \) tracers, the results with \( L = 20 \) are included to provide a more clear vision. Since the GB filter deals with only the incompressible GB flow, it has been proved [see Chen et al. (2014b)] that the tracers are uniformly distributed at the statistical equilibrium state. With the interference of the gravity modes, the distribution of tracers at \( t = 199 \) remains nearly uniform in the \( \varepsilon = 0.1 \) regime since the fast oscillation averages out the effect from the random compressible gravity waves. On the other hand, pronounced clustering of tracers is found in the \( \varepsilon = 1 \) regime due to the compressible nature of the underlying flow. In addition, it is clear that with the \( L = 5 \) tracers, the underlying GB flow can be filtered with high accuracy in both dynamical regimes utilizing both the full and the GB filter.

We now focus on the filtering skill utilizing different filters. As stated in (26), the structure of GB flow is controlled by the two Fourier coefficients: \( \hat{v}_{(0,1)} \) and \( \hat{v}_{(1,0)} \). To this end, we show in Fig. 3 the truth and the posterior mean estimates of these two coefficients. In the \( \varepsilon = 0.1 \) regime, the filtered solutions of \( \hat{v}_{(0,1)} \) and \( \hat{v}_{(1,0)} \) utilizing all four filters are quite close to the truth while in the \( \varepsilon = 1 \) regime a significant error with many unexpected oscillations is found (Fig. 3d) in the filter estimate utilizing the reduced filter with only the GB forecast model in (13). To provide an intuitive illustration, the recovered streamlines of the GB flow is demonstrated in Fig. 4 at two time instants, where the true GB streamline at \( t = 142.3 \) is a 2D array of swirling eddies and at \( t = 161.4 \) it becomes a shear-flow stream. Consistent with the filter behaviors as shown in Fig. 3, the filtered streamlines utilizing all the filters are nearly the same as the truth in the \( \varepsilon = 0.1 \) regime. On the other hand, despite the skillful filter estimates utilizing both the GB (12) and the full filter (11), the reduced filter with only the GB forecast model (13) leads to a large disparity in the recovered streamlines in the \( \varepsilon = 1 \) regime, where the true flow at \( t = 142.3 \) is swirling eddies while the recovered field is shear flows (Fig. 4c). Similarly, the weak shear-flow stream at \( t = 161.4 \) becomes strong swirling eddies in the filtered solution (Fig. 4d). In addition, although some inaccuracy is also found in the filter estimate utilizing the diagonal-reduced 3D-Var filter (14), the recovered streamlines are qualitatively similar to the truth.

To understand the dependence of the filters’ behavior on different parameters, we show in Figs. 5–7 the filtering skill as a function of the tracer numbers \( L \), the observation noise \( \sigma_x \), and the energy in the gravity modes \( E_g \), respectively. Both the RMS error in the posterior mean estimate and averaged posterior variance are computed over time interval \( t \in [20, 200] \), where only the statistics of mode \((1, 0)\) is shown for simplicity. The information
model error in filtering the GB flows utilizing the two imperfect reduced filters (13) and (14) compared with the perfect full filter (11) through the relative entropy (20) is computed, where the information model error is split into the signal and dispersion parts utilizing the formula in (16). The model error averaged over time interval $t \in [20, 200]$ in filtering the GB flow field is shown in these figures.

First, we look at the RMS error in filtering the GB mode $(1, 0)$. The RMS error decreases in the filter estimates utilizing the GB filter (12) with the increase of $L$ and the decrease of $\sigma_x$ and $E_g$. In the $\varepsilon = 0.1$ regime, all the filters have almost comparably high filtering skill. In the $\varepsilon = 1$ regime, the filtering skill utilizing both the full filter (11) and the diagonal-reduced 3D-Var filter (14) remains close to that utilizing the idealized GB filter (12). However, the RMS error utilizing the reduced filter with only the GB forecast model (14) is much larger. Note that the RMS error in the reduced filter with only the GB forecast model (13) shoots up with a decrease of

![Random Fourier coefficients in $\varepsilon = 0.1$ regime](image)

![Random Fourier coefficients in $\varepsilon = 1$ regime](image)

**FIG. 3.** Truth and the filtered solutions (posterior mean estimates) of $\hat{u}_{(1,0)}$ and $\hat{v}_{(1,0)}$ of the GB flow in (26). These two Fourier coefficients determine the flow structure. (top) The $\varepsilon = 0.1$ regime. (bottom) The $\varepsilon = 1$ regime.
When $\sigma_x$ is small, which is a different trend compared with the other filters. Clearly, a small $\sigma_x$ means the filter trusts more toward the observations. However, despite observation equation in the filter involving only the GB modes, the actual observations contain both the GB and gravity components. Therefore, large errors come from the interference from the gravity waves in the input of the observations.

Now we focus on the information model error in (20). As shown in Figs. 5c, 6c, and 7c, the model error in the reduced filter with only the GB forecast model (13) is significant in the $\varepsilon = 1$ regime, where the signal part has a dominant portion. In contrast to (13), the lack of information in the diagonal-reduced 3D-Var filter (14) shown in Figs. 5d, 6d, and 7d is much smaller. In addition, the model error in the $\varepsilon = 0.1$ regime utilizing both the imperfect filters is smaller than that in the $\varepsilon = 1$ regime. Note that different trends in large $L$ and small $\sigma_x$ are found in the RMS error and information model error utilizing the diagonal-reduced 3D-Var filter (14). This is because the signal part of the information model error in (16) is proportional to the inverse of the covariance of the imperfect model. With a slowly varying gap in the mean estimates, a smaller covariance implies a more certain estimate of the incorrect state and thus a larger information model error. It is worthwhile pointing out that the information model error has no upper bound and, thus, it is very sensitive when the model covariance becomes extremely small. A bounded measurement for checking the model error in the posterior distribution is the Hellinger distance (Beran 1977; Branicki and Majda 2014), which is, however, not able to be explained as a measure of information gain.

Finally, to provide a deeper understanding of the two imperfect filters, we include in Figs. 8a–d some time series of the filtered solutions for mode $(1, 0)$. Figures 8a and 8b show the absolute error in the posterior mean estimate of the GB mode $(1, 0)$ utilizing the reduced filter with only the GB forecast model (13). In the $\varepsilon = 0.1$ regime, the error amplitude remains significantly smaller than the true signal. On the other hand, except a small error at the initial period for $t \leq 20$ in the $\varepsilon = 1$ regime, the amplitude of the error is comparable with that of the true signal, which leads to a significant lack of information in the signal part. In Figs. 8c and 8d, the posterior covariance for mode $(1, 0)$ utilizing both the full filter (11) and the diagonal-reduced 3D-Var filter (14) is shown. Clearly, the diagonal components of the covariance matrix of the full filter, that is, both the variance of the GB mode (blue) and that of the gravity mode (black), have much larger amplitudes than the cross covariance between them (magenta). The negligible cross covariance is possibly due to the orthogonality of the eigenvectors associated with the GB and gravity modes. We have also checked the cross covariance between different GB and different gravity modes and they are small as well. These are evident proofs for the skillful behavior of the reduced 3D-Var filter (14). It is also noticeable that the posterior variance of the diagonal-reduced 3D-Var filter (14) becomes a constant after a short initial relaxation time, which is justified in appendix A. Note that the reduced 3D-Var filter (green) results in a smaller variance than the full filter (blue) in the $\varepsilon = 1$ regime, which leads to the increase of the information model error. A natural improvement for the diagonal-reduced 3D-Var filter is to
\[ \varepsilon = 0.1 \]

(a) RMS error in posterior mean for mode (1,0)

(b) Posterior variance for mode (1,0)

(c) Model error in the reduced filter with only GB forecast model

(d) Model error in the diagonal reduced 3D–Var filter

\[ \varepsilon = 1 \]

Fig. 5. (a),(b) The RMS error in the posterior mean estimate and the averaged posterior covariance over time \( t \in [20, 200] \) for the GB mode (1, 0) utilizing different filters. (c),(d) The information model error in filtering the GB flow utilizing the two reduced filters as a function of the number of tracers \( L \). Here the information model error is computed utilizing the formula in (20) and then averaged in time for \( t \in [20, 200] \). In (20), \( p \) is the posterior distribution of the perfect full filter (11) and \( p^M \) is that of the reduced filter with only the GB forecast model (13) [in (c)] and the diagonal-reduced filter (14) [in (d)], respectively.

Inflate its diagonal covariance matrix by a factor \( r \) with \( r \times \mathbf{R}_e \). In Fig. 8e, we show the information model error as a function of the inflation factor \( r \). When \( r = 1.6 \), which is around the ratio of the averaged variance utilizing the full filter over that utilizing the diagonal-reduced filter at the statistical equilibrium, the information model error is reduced by 40%. The lack of information in the dispersion part is nearly zero as expected and that in the signal part is...
also reduced since the signal part is proportional to the inverse of the model covariance.

d. Results for filtering the random flow fields with coupled GB and gravity modes

From now on, we will study the skill of filtering the multiscale random rotating compressible flow in the situation that each GB mode affects the underlying dynamics of the two corresponding gravity modes through quadratic nonlinear interactions, which is motivated directly from mathematical theory of the slow–fast geophysical flows (Embid and Majda 1998; Majda 2003; Gershgorin and Majda 2008) and high resolution of turbulent simulations in slow–fast geophysical regimes (Smith 2001; Smith and Waleffe 2002; Waite and Bartello 2004). Let us recall the governing equations of the underlying flow field for Fourier wavenumber \( k \):

\[ \varepsilon = 0.1 \]

\[ \varepsilon = 1 \]

Fig. 6. As in Fig. 5, but as a function of the observational noise \( \sigma_x \).
\[ \varepsilon = 0.1 \] (a) RMS error in posterior mean for mode (1,0) 
\[ \varepsilon = 1 \] (b) Posterior variance for mode (1,0) 
\[ \text{(c) Model error in the reduced filter with only GB forecast model} \]
\[ \text{(d) Model error in the diagonal reduced 3D-Var filter} \]

Fig. 7. As in Fig. 5, but as a function of the energy \( E_g \) in gravity modes, where \( E_g^0 = 0.1 \) is the standard value in each gravity mode.

\[ d\tilde{v}_{k,0}(t) = \left[ -d_{B,k} \tilde{v}_{k,0} + f_{k,0}(t) \right] dt + \sigma_{k,0} dW_{k,0}(t), \quad (29a) \]

\[ d\tilde{v}_{k,z}(t) = \left[ (-d_{g} + i\omega_{k,z} \pm iy\tilde{v}_{k,0}) \tilde{v}_{k,z}(t) + f_{k,z}(t) \right] dt + \sigma_{k,z} dW_{k,z}(t), \quad (29b) \]

where the coupling coefficient \( y \) is nonzero. On the other hand, such nonlinear coupling between the GB and gravity modes (the underlined term) is dropped in the forecast models of both the full filter (11) and the diagonal-reduced 3D-Var filter (14) and, therefore,
these forecast models each become a linear independent stochastic model (28) as discussed in section 2c. Because of this model error, the full filter is no longer a perfect filter. Note that the observational processes in (11), (13), and (14) remain highly nonlinear with coupled GB and gravity modes.

We first look at the intrinsic change in the coupled flow fields with the coupling effect. In Fig. 9, the sample trajectories and the associated power spectrums of the gravity mode (1, 0) are demonstrated, and those of the GB mode are also shown as comparison. The spectrum of the gravity mode becomes more and more flat with the increase of the coupling coefficient $\gamma$ in both regimes. In the $\epsilon < 0.1$ regime, the spectrums of the GB and gravity modes remain having almost no overlapped band even with $\gamma = 5$, which implies a clear scale separation between them and, therefore, skillful filtering results of the GB flow are expected. On the other hand, the spectrum bands of the GB and gravity modes in the $\epsilon = 1$ regime become completely overlapped with each other for $\gamma > 1$, which indicates that the GB and gravity flows are hard to be distinguished from the mixed observations. Therefore, the filtering skill in the $\epsilon = 1$ regime is expected to deteriorate.

We show in Fig. 10 the filtered GB modes (1, 0) and (0, 1) and the reconstructed streamlines with $\gamma = 2$ in the $\epsilon = 1$ regime. Here the true GB flow is adopted to be the same as that in section 4c and, therefore, the two Fourier modes in Fig. 10 remain the same as those in Fig. 3. The filter estimate of the GB filter (12) has very little change due to the randomness in the observation noise. However, the filter estimates utilizing all three imperfect filters contain evident errors, where the bias utilizing the reduced filter with only the GB forecast model (13) is the most significant. This is reflected in the recovered streamlines at five different time instants. The reduced filter with only the GB forecast model (13) leads to completely wrong flow structures, while the full filter with the linear forecast model (11) at least has some skill at $t = 8.5$ and $t = 105.5$. In addition, the diagonal-reduced 3D-Var filter (14) is skillful for recovering the shear flow at time $t = 8.5$ as well.

In Fig. 11, we show the RMS error in the posterior mean estimation and the averaged posterior variance for mode (1, 0), where the filtering skill in both the GB and one of the gravity modes is included. As motivated from Fig. 9, the nonlinear coupling up to $\gamma = 5$ has little effect on the filtering skill of the GB mode utilizing all the filters in the $\epsilon = 0.1$ regime due to the apparent scale separation. The error in the filtered solution of the gravity mode is also almost unchanged with different $\gamma$, which is possibly due to the fact that its intrinsic oscillation in this fast oscillation regime dominates the stochastic oscillation from the interaction with the GB mode and, therefore, the stochastic oscillation behaves as insignificant random noise. On the other hand, the filtering skill of the GB mode utilizing all the three imperfect filters deteriorates with a gradual increase of $\gamma$ in
the $\varepsilon = 1$ regime. Among the three imperfect filters, the largest RMS error remains in the reduced filter with only the GB forecast model (13). In addition, unlike the uncoupled situation where the full filter and the diagonal-reduced 3D-Var filter always have a comparable filtering skill, with a nonzero $\gamma$ the error utilizing the diagonal-reduced 3D-Var filter (14) becomes more significant than the full filter with the linear forecast model (11). Furthermore, filtering the gravity waves becomes less skillful utilizing both the full filter with the linear forecast model and the diagonal-reduced 3D-Var filter with the increase of $\gamma$ in the $\varepsilon = 1$ regime.

Finally, we study the information model error. Since the full filter with the linear forecast model (11) is no longer a perfect filter, the model error in both (11) and the two reduced filters (13) and (14) are assessed following the discussion at the end of section 3.
Figures 12a and 12b show the model error in the time-averaged PDF of the posterior mean estimation utilizing the imperfect filters related to that of the true signal over time interval $t \in [20, 200]$ for the GB mode $(1, 0)$. In the $\varepsilon = 0.1$ regime, the model error remains small for all the filters. In the $\varepsilon = 1$ regime, the model error of the three imperfect filters becomes large for $\gamma \geq 1$, where the largest model error is found in the diagonal-reduced 3D-Var filter. In Fig. 12c, we compare the time series of the posterior mean estimate and the true signal with $\gamma = 2$ in the $\varepsilon = 1$ regime and the associated PDFs are shown in Fig. 12d. Clearly, the difference in the PDF of the posterior mean estimates compared with the truth, reflecting the lack of information, is obvious utilizing all the three imperfect filters. Particularly, the large information model error in the diagonal-reduced 3D-Var filter (14) is due to the fact that its PDF is more concentrated than that of the truth. This implies that the posterior mean estimation of (14) misses many extreme events, such as those around $t = 140$. Note that with a nonzero coupling coefficient $\gamma$, a nonnegligible cross covariance between the GB and gravity modes appears.
and, therefore, a large model error is expected by dropping the off-diagonal entries in the posterior covariance matrix. It is worthwhile pointing out that the RMS error and the information model error provide different views in assessing the filtering skill in the posterior mean estimation. Despite a smaller RMS error compared with the reduced filter with only the GB forecast model (13), a larger information model error in the diagonal-reduced 3D-Var filter (14) implies the potential danger in utilizing (14) with a moderate or
Figure 13 shows the information model error in the posterior distribution $p_M$ of the GB flow utilizing the three imperfect filters (11), (13), and (14) compared to $p$ utilizing the idealized GB filter (12) averaged over time $t \in [20, 200]$. It is clear that the information model error in all the three imperfect filters remain small in the $\varepsilon = 0.1$ regime while it becomes significant larger in the $\varepsilon = 1$ regime and increases as a function of $\gamma$. Again, the signal part dominates the model error. As expected, the full filter with linear forecast model (11) has the smallest lack of information. Among the two reduced filters, the computational efficient diagonal-reduced 3D-Var filter (14) has a smaller model error than the reduced filter by completely dropping the forecast models associated with the gravity waves in (13). Yet, the lack of information in the diagonal-reduced 3D-Var filter increases much more significantly with $\gamma$ than the full filter with the linear forecast model.

5. Summary and conclusions

In this paper, the filtering skill and the information model error of filtering the multiscale random rotating compressible flows utilizing noisy Lagrangian tracers are extensively studied. The random flow fields are defined through random amplitudes of Fourier eigenmodes of the rotating shallow-water equations, which involve both the random incompressible GB flows and the random rotating compressible gravity waves (section 2a). The GB and gravity modes are coupled in a highly nonlinear way in the tracer observations (section 2b). An information-theoretic framework (section 3) is developed to assess the lack of information and model error in imperfect filters, which applies to a single realization of the observations. Two scenarios of the underlying dynamics of the flow fields are taken into consideration.

First, linear stochastic equations are utilized to model the random amplitudes of Fourier modes, where different modes are assumed to be independent to each other [8]. Despite the highly nonlinear observations, the joint signal–observation process forms a conditional Gaussian system, which allows closed analytical update of the posterior states in the optimal filter. In addition to the full optimal filter, an idealized GB filter (12) is proposed as a reference for filtering the slow-varying GB flow, which is of primary concern in practice, and the two reduced filters (13) and (14) are developed for practical purpose. The true GB flow field is designed based on a simple setup with time-varying flow structures (section 4a). In the dynamical regime with fast rotation (Rossby number $\varepsilon = 0.1$), all four filters have comparably high filtering skill. In a moderate rotation regime ($\varepsilon = 1$), a significant information model error is found utilizing the reduced filter with only the GB forecast model through the full observations in (13). On the other hand, the diagonal-reduced 3D-Var filter (14) is not only computationally efficient but nearly as skillful as the optimal filters in filtering the GB modes as well (section 4c).

In the second part of this paper, a more realistic situation with coupled GB and gravity modes is considered, where each GB mode affects the two gravity...
modes with the same Fourier wavenumber through a quadratic nonlinear interaction in (29) following the mathematical theory of the slow–fast atmosphere flows (Embid and Majda 1998; Majda 2003; Gershgorin and Majda 2008). Since the full filter with the nonlinear forecast model no longer belongs to the conditional Gaussian filtering framework, the linear forecast model is adopted in both the full and the diagonal-reduced 3D-Var filters, which follows the common practical strategy for filtering high-dimensional turbulent systems (Majda and Harlim 2012; Harlim and Majda 2013). Again, all four filters are comparably skillful in the $\varepsilon = 0.1$ regime even in the appearance of a strong nonlinear coupling in the true flow (section 4d). In the $\varepsilon = 1$ regime, the three imperfect filters (i.e., the full filter with linear dynamics and the two reduced filters), lose their filtering skill with the increase of the nonlinear coupling. The filtering skill of the full filter with the linear forecast model remains acceptable. On the other hand, information theory shows that the diagonal-reduced 3D-Var filter fails to recover the extreme events, while the reduced filter with only the GB forecast model suffers from a significant lack of information in the posterior distribution compared to that of the idealized GB filter.

Although the filtering framework proposed here is idealized, it provides some guidelines for filtering the real-world issues. For example, the Rossby number is small for the motion of planetary scale in the midlatitudes. According to the analysis in this work, the effect of gravity waves is averaged out effectively and, therefore, the computationally efficient reduced filters (13) and (14) are good approximations in this regime. On the other hand, linear approximations of highly nonlinear underlying dynamics do not necessarily lead to huge model errors (section 4d). If the scale separation between different modes is evident (top panel of Fig. 9) then the cheap linear approximations such as the various versions of the Kalman filter and 3D-VAR remain skillful.

There are many other sources of the model errors that appear in the real-world applications, which have not been extensively explored in this work. Nevertheless, the effect of some other types of the model error can also be studied in the framework proposed here. For example, a simplified dynamics in the imperfect filter can be utilized to study the model error due to the poor representation of surface layer processes. In addition, adopting an underestimated stochastic forcing coefficient [e.g., $\Sigma_U$ in (11)] mimics the situation in many realistic models for turbulence, which have too much dissipation (Palmer 2001) due to inadequate resolution and deterministic parameterization of unresolved features. Some preliminary results are reported in Chen and Majda (2016) and more extensive studies based on the noisy Lagrangian tracer framework remain as future research work. Note that the multiscale filtering framework proposed here deals
with both balanced and unbalanced flows and is useful for both atmosphere and ocean sciences.

Acknowledgments. This research of A. J. M is partially supported by the Office of Naval Research Grant ONR MURI N00014-12-1-0912. N. C. is supported as a graduate research assistant on this grant. A. J. M also gratefully acknowledges the generous support of the Center for Prototype Climate Modeling of NYU Abu Dhabi Research Institute grant.

APPENDIX A

Filtering Formulas for the Diagonal-Reduced 3D-Var Filter

In the diagonal-reduced 3D-Var filter, the posterior covariance is set to be diagonal. The formulas of updating the posterior mean are the same as that utilizing the full filter. To see the update of the posterior covariance, we denote \( R_{t,i} \) to be the \((i, i)\) th entry of \( R_t \).

Then the update of the posterior covariance \( R_t \) [see section 2.2.1 of Chen et al. (2015)] becomes \(|K|\) independent 1D equations:

\[
\frac{dR_{t,i}}{dt} = \left\{ -\Gamma_t R_{t,i} - R_{t,i} \Gamma_t^* + (\Sigma_u \Sigma_u^*)_{ii} \\
- \sigma^2 \Gamma_t^2 |P_t(X_t)P_t(X_t)|_{ii} \right\} dt,
\]

(A1)

where \((\cdot)_{ii}\) means the \((i, i)\)th entry of the matrix. In each time step, after solving each \( R_{t,i} \), we insert \( R_t \) into the posterior mean update. It is worthwhile noticing that the \((i, i)\)th component of \( P_t(X_t)P_t(X_t) \) is simply \( |\mathbf{r}_i|^2 \), as is seen in (6) due to the fact that \( \exp[-\mathbf{k} \cdot \mathbf{X}_t(s)] \times \exp[\mathbf{k} \cdot \mathbf{X}_t(s)] = 1 \) for \( \mathbf{k} \in \mathbf{K} \). Thus, (A1) is deterministic. In addition, the diagonal entry \( R_{t,i} \) converges to a constant equilibrium value after a short relaxation time.

On the other hand, the \( P_t(X_t)P_t(X_t) \) matrix in the full filter is not a constant matrix because the tracer locations play important roles in the off-diagonal components. Because of the nonlinearity in \( R_t \), the diagonal entries affected by the off-diagonal one also become time dependent.

APPENDIX B

Approximation Error by Utilizing the GB Filter as the Reference in Assessing the Information Model Error

Here, we compare the information model error in (20) in the imperfect filters (13) and (14) in the situation with \( \gamma = 0 \) by choosing different reference perfect filters. The posterior distribution associated with either the full filter (11) or the idealized GB filter (12) is chosen as \( p_M \). The goal is to see the approximation bias in (20) by choosing the GB filter (12) as the reference filter in assessing the information model error. We show the results as a function of the energy in the gravity modes in Fig. B1.
It is clear that the information model error in both the imperfect filters by utilizing the posterior distribution associated with the GB filter (12) as the reference distribution $p$ in (20) is slightly larger than utilizing that associated with the full filter (11) due to the extra lack of information in the observations. Importantly, the qualitative conclusions in assessing the model error in the reduced filters by utilizing different reference distribution $p$ remain the same, which provides the justification of adopting the GB filter as the reference in assessing the model error in the more complicated situation with $\gamma \neq 0$.

REFERENCES


