Stochastic modeling and numerical simulation of ocean dynamics

Long Li

DOCTORAT BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Mathématiques et leurs interactions*

Par
Long LI

Stochastic modeling and numerical simulation of ocean dynamics

Thèse présentée et soutenue à Rennes, le 23 mars 2021
Unité de recherche : INRIA / IRMAR

Rapporteurs avant soutenance :

Dan CRISAN	Professor, Imperial College London
Christian FRANZKE	Associate Professor, Pusan National University

Composition du Jury :

Président :	Arnaud DEBUSSCHE	Professeur, ENS Rennes
Examinateurs :	Dan CRISAN	Professor, Imperial College London
	Christian FRANZKE	Associate Professor, Pusan National University
	Guillaume LAPEYRE	Directeur de recherche, CNRS
	Anne Marie TREGUIER	Directrice de recherche, CNRS
Dir. de thèse :	Etienne MÉMIN	Directeur de recherche, INRIA
Co-dir. de thèse :	Werner BAUER	Research associate, Imperial College London

To my parents and my family.

Acknowledgement

First of all, I would like to express my gratitude to Pr. Dan Crisan and Pr. Christian Franzke for their attentive review of this manuscript. I also wish to show my appreciation to Pr. Arnaud Debussche, Pr. Guillaume Lapeyre and Pr. Anne Marie Treguier for their insightful comments and suggestions of this work.

Foremostly, I would like to give deepest gratitude to my primary supervisor, Pr. Etienne Mémin, who offered me this interesting research topic and guided me throughout this project. Without his invaluable advice, continuous support, and constant patience, this thesis could not have reached its present form. It was a great pleasure to work with him during those years. I would like to extend my sincere thanks to my co-supervisor, Dr. Werner Bauer, who made me understand how rigorous the scientific research should be. I learned especially from him how to write properly conference abstracts and academic articles. In addition, I would like to especially thank Pr. Clémentine Prieur for bringing me into the "stochastic" field in my Master's research projects.

I would like to express my warm thanks to the following collaborators for the current and previous joint works: Bruno Deremble, Noé Lahaye, Camilla Fiorini, Pr. Bertrand Chapron, Valentin Resseguier, Pranav Chandramouli, Pierre Dérian and Rüdiger Brecht. I benefited many knowledge of different research areas from them. I would also like to thank my officemates through the years, Gilles Tissot, Carlo Cintolesi and Mohamed Yacine Ben Ali, for the amicable exchanges and interesting discussions. Besides, I was very grateful to the STUOD team, and in particular to Pr. Darryl Holm and Dr. Wei Pan for their helpful discussions and suggestions.

I also wish to acknowledge the technical support provided by Charles Deltel, and the administrative assistance provided by Huguette Bechu and Caroline Tanguy.

Finally, I would like to thank my parents and my family for their infinite support and encouragement during the entire period of my study in France. Particularly, I am deeply grateful to Lisa for always accompanying me, from Saint Etienne to Grenoble and then to Rennes. I really appreciate the wonderful time that we have together.

Table of Contents

List of Figures

List of Tables

Résumé en français

Cette thèse explore une représentation stochastique des effets des petits tourbillons sur la circulation océanique à grande échelle. Ce cadre stochastique, appelé modélisation sous incertitude de position (en anglais: location uncertainty – LU), est dérivé des lois de conservation physique. Ce modèle repose principalement sur les deux motivations suivantes: fournir des prévisions d'ensemble plus fiable et plus efficace pour un système d'assimilation de données et améliorer les variabilités océaniques pour des simulations à l'échelle climatique.

Cette thèse est principalement construite autours de trois articles publiés (Bauer et al., 2020a,b; Resseguier et al., 2020a), d'un manuscrit soumis (Brecht et al., 2021) et d'un manuscrit en préparation (Li et al., 2021). Elle est organisée en huit chapitres avec une introduction générale et une conclusion. Nous résumons brièvement ci-dessous le contenu de ces chapitres.

Résumé du chapitre 1 – Préliminaires: calcul stochastique

Dans le premier chapitre, nous présentons brièvement quelques définitions et propriétés fondamentales du calcul stochastique. Ces formules mathématiques sont largement utilisées dans cette étude afin de dériver les modèles physiques décrits dans les chapitres suivants. Des notions générales sur les processus stochastiques et les intégrales stochastiques en dimensions finies sont d'abord présentées. Ensuite, nous décrivons rapidement l'extension de ces notions en dimensions infinies.

Résumé du chapitre 2 – Transport sous incertitude de position

Afin de tenir compte des effets des petites échelles sur l'évolution de l'écoulement à grande échelle, une représentation stochastique basée sur un principe de transport physique a d'abord été introduite par Mémin (2014). Ce principe provient d'une décomposition de la vitesse lagrangienne en une composante lisse (en temps) et un terme très oscillant, nommé comme "bruit" de façon générale dans ce travail.

Ce chapitre explique en détail le modèle aléatoire avec ses propriétés principales en

utilisant les outils de calcul stochastique rappelés dans le premier chapitre. Nous montrons tout d'abord que le transport d'un traceur aléatoire par l'écoulement stochastique consiste en un forçage aléatoire multiplicatif correspondant à l'advection du traceur par le bruit (appelé également bruit de transport), une diffusion hétérogène liée à la variance du bruit et une advection effective associée à l'inhomogénéité du bruit.

Une caractéristique importante de ce modèle aléatoire est qu'il conserve l'énergie du traceur pour chaque réalization. Dans ce processus, l'énergie apportée par le bruit est exactement contrebalancée par la perte d'énergie associée à la diffusion. Le terme d'advection effective rend compte de l'action de l'inhomogénéité spatiale de la composante de bruit sur le courant à grande échelle en redistribuant localement l'énergie à grande échelle.

Ensuite, un théorème de transport de Reynolds stochastique est également dérivé afin de décrire le transport eulérien d'un traceur aléatoire dans un volume de contrôle transporté arbitrairement par l'écoulement.

Le travail présenté dans ce chapitre a été publié dans l'article Bauer et al. (2020a).

Résumé du chapitre 3 – Modèles stochastiques tridimensionnels

Dans ce chapitre, nous decrivons certains modèles d'écoulements tridimensionnels dans le cadre LU. La dérivation de ces modèles s'appuie essentiellement sur l'expression du théorème de transport de Reynolds présentée dans le chapitre précédent. Nous présentons d'abord la dérivation des équations stochastiques pour les écoulements à partir des conservations stochastiques de la masse et de la quantité de mouvement. Afin de faire quelques comparaisons avec le cadre classique de la simulation aux grandes échelles de la turbulence, nous présentons un modèle pseudo-stochastique reposant sur une hypothèse forte de séparation d'échelle. Ensuite, nous dérivons une représentation stochastique de l'évolution de la vorticité en utilisant le calcul vectoriel classique.

Pour les circulations atmosphériques et océaniques, deux systèmes stochastiques sont présentés en utilisant l'équilibre hydrostatique classique et les approximations de Boussinesq. Ces modèles seront utilisés afin de dériver plusieurs équations stochastiques quasi-bidimensionnelles dans le chapitre suivant.

De plus, nous établissons un lien entre le modèle LU et le système Craik-Leibovich. En particulier, nous montrons que le modèle LU introduit une structuration de l'écoulement à grande échelle liée à l'action de la composante de petite échelle. Cet effet est généré par la vitesse statistique (inclus dans le terme d'advection effective) induite par les petits tourbillons, qui peut être interprétée comme une généralisation de la dérive de Stokes,

et est donc appelée dérive de Itô-Stokes dans ce travail. En fait, de façon similaire à la dérive de Stokes, cette dérive de Itô-Stokes peut être interprétée comme la différence entre la moyenne d'ensemble de la vitesse lagrangienne et la moyenne d'ensemble de la vitesse effective eulérienne.

Enfin, nous effectuons également une connexion avec la paramétrisation Gent-McWilliams (GM) qui est largement adoptée dans les modèles de circulation générale de l'océan. En projetant le bruit le long des surfaces isopycnales, la dérive de Itô-Stokes dans le modèle LU joue un rôle similaire à la vitesse "bolus" introduite dans le schéma GM et la diffusion du traceur dans le modèle LU est liée à la diffusion iso-neutre introduite dans le schéma de Redi.

Le travail présenté dans ce chapitre a en parti été publié dans l'article Bauer et al. (2020a).

Résumé du chapitre 4 – Approximations géostrophiques stochastiques

Dans ce chapitre, nous présentons quelques systèmes stochastiques quasi-bidimensionnelles décrivant l'évolution des courants mésoéchelle. Outre l'équilibre hydrostatique sur la verticale qui est présenté dans le chapitre précédent, ce chapitre se concentre sur une autre approximation importante entre le gradient de pression et les forces de Coriolis. Cet approximation appélé l'équilibre géostrophique, est principalement lié à l'effet dominant de la rotation terrestre dans les circulations atmosphériques et océaniques à grande échelle.

Nous dérivons d'abord le système stochastique en rotation des écoulements en eaux peu profonde (en anglais: rotating shallow water – RSW), qui est une bonne approximation des équations primitives. Ensuite, nous démontrons la conservation d'énergie pour un système RSW simplifié. Puis, nous spécifions quelques nombres sans dimension pour adimensionner le système RSW et le système Boussinesq présenté dans le chapitre précédent.

Par la suite, nous adoptons des approches asymptotiques pour dériver les équations géostrophiques planétaires (PG) barotropes et quasi-geostrophiques (QG) barotropes en utilisant le système RSW sans dimension, ainsi que les équations PG et QG primitives continues sur verticale en utilisant le système Boussinesq sans dimension. Sous le régime QG, certaines sources et puits de vorticité proviennent de l'interaction des déformations entre le bruit à petite échelle et le courant à grande échelle. Nous montrons que ces termes sont importantes pour préserver l'énergie totale de l'écoulement.

Les modèles quasi-bidimensionnels stochastiques ont été testés avec succès dans plusieurs simulations numériques et certains résultats intéressants sont montrés dans les chapitres suivants.

Les travaux présentés dans ce chapitre reposent sur l'article publié Bauer et al. (2020a) et sur le manuscrit Brecht et al. (2021) qui a été soumis.

Résumé du chapitre 5 – Modélisation numérique de l'incertitude

Afin de réaliser des simulations numériques des modèles LU, la structure spatiale (possiblement dynamique en temps) du bruit doit être modélisée *à priori*. En particulier, nous explorons la méthode de décomposition spectrale à partir de la base des fonctions propres de l'opérateur covariance spatiale. Ce chapitre présente quelques méthodes numériques pour estimer ces fonctions propres empiriques. Nous décrivons d'abord les approches reposant sur les données, puis détaillons certaines méthodes de paramétrisations auto-adaptée aux différentes échelles. Le bruit qui en résulte peut être homogène ou hétérogène en espace et stationnaire ou non-stationnaire en temps.

Nous utilisons en premier lieu l'analyse en composantes principales pour apprendre efficacement la base des fonctions propres stationnaires à partir des données d'une simulation à haute résolution. En particulier, une dérive de correction sous-maille a été identifiée en raison du biais provenu de la procédure de sous-échantillonnage des données. Ensuite, pour relâcher l'hypothèse de stationnarité, nous proposons de mettre à jour en temps les valeurs propres associées aux fonctions propres contraignant les modes principaux temporels du bruit avec la dynamique grande échelle.

Nous avons par ailleurs proposé d'estimer les fonctions propres dépendant du temps à partir des fluctuations de vitesse locales de la simulation courante, avec une renormalization de l'amplitude du bruit s'appuyant sur une hypothèse de similitude.

Une autre méthode de paramétrisation vise à adapter à l'équation de transport stochastique à un opérateur de dissipation sous-maille donné tel que l'hypervisocité. L'idée principale repose sur l'hypothèse que la diffusion dans les modèles LU peut être identifiée avec une dissipation sous-maille spécifiée à un facteur près. Ainsi, pour une base orthonormée fixée comme les ondelettes, les valeurs propres associées sont mises à jour selon le bilan de dissipation d'énergie. Par conséquent, le forçage aléatoire multiplicatif dans le transport contrebalance la dissipation numérique au facteur près.

Les travaux présentés dans ce chapitre ont été publiés dans les articles Bauer et al. (2020a,b); Resseguier et al. (2020a).

Résumé du chapitre 6 – Vérification des prévisions d'ensemble pour la dynamique QG surfacique

Dans ce chapitre, nous détaillons d'abord quelques métriques importantes pour quantifier les compétences de prédiction des prévisions d'ensemble, comme le diagramme de Talagrand, le "continuous ranked proper score" et le score d'énergy. Nous considérons ici le modèle QG surfacique comme cas de test. En particulier, nous utilisons les paramétrisations stochastiques présentées dans le chapitre précédent afin de comparer leur performances en terme de prévisions d'ensemble à temps court.

Nous montrons que le modèle aléatoire proposé, que ce soit avec incertitude homogène ou hétérogène, fournit des ensembles avec une dispersion plus efficace et réaliste qu'un modèle classique avec une perturbation aléatoire de la condition initiale. Cette capacité est surtout importante pour les applications d'assimilation des données. En termes de fiabilité d'ensemble, nous montrons en sus que les modèles avec des bruits hétérogène sont meilleurs que ceux avec des bruits homogènes et que les modèles associés à des bruits non-stationnaires sont meilleurs que ceux avec des bruits stationnaires.

Le travail présenté dans ce chapitre a été publié dans l'article Resseguier et al. (2020a). Cependant, ce chapitre apporte des élements plus précis quant à l'efficacité de différents modèles de bruit.

Résumé du chapitre 7 – Etudes numériques du modèle QG barotrope stochastique

Ce chapitre fournit quelques résultats numériques du système QG barotrope stochastique dérivé dans le chapitre 4. En utilisant un modèle barotrope périodique non visqueux, nous montrons d'abord que le modèle LU avec des bruits à la fois homogènes et hétérogènes préserve l'amplitude et la vitesse de propagation de l'onde de Rossby. De plus, nous illustrons numériquement que l'introduction du bruit inhomogène induit une structuration de l'écoulement à grande échelle avec des tourbillons secondaires.

Ensuite, la performance de ce modèle barotrope stochastique est évaluée au moyen d'une simulation numérique dans le cadre d'une configuration idéalisée – double gyres entraînés par un vent stationnaire dans un bassin fermé peu profond à moyenne latitude. Nous nous intéressons ce contexte à la capacité des modèles stochastiques proposés à représenter correctement les quatre premiers moments statistiques (moyenne, variance, asymétrie et kurtosis) des écoulements à une résolution grossière. La comparaison de

distribution statistique par ses quatres moments avec celle prédite par les données de haute résolution, nous permet de qualifier et de quantifier la prise en compte, par notre représentation stochastique, des effets des tourbillons de mésoéchelle sur la circulation grande échelle.

Les travaux présentés dans ce chapitre ont été publiés dans les articles Bauer et al. (2020a,b).

Résumé du chapitre 8 – Etudes numériques du modèle QG multicouche stochastique

Un modèle QG multicouche ne nécessite que quelques couches pour générer des instabilités baroclines. Cependant, les effets des tourbillons à mésoéchelle peuvent encore être absents pour des configurations grossièrement résolues dans le plan horizontal. Dans ce chapitre, nous continuons nos études sur une circulation double-gyre idéalisée avec un modèle QG multicouche. En particulier, nous nous concentrons sur la reproduction de l'écoulement de jet ainsi que sur la prédiction de la variabilité basse fréquence pour le modèle aléatoire défini sur une maille grossière.

La corrélation spatiale empirique du bruit a d'abord été estimée à partir de données de simulation de haute résolution. Nous montrons que dans ce contexte la dérive de correction sous-maille est très importante pour reproduire sur une maille grossière le jet zonale de la circulation doubles gyres. Une nouvelle méthode de projection a par ailleurs été proposée pour contraindre le bruit à vivre le long des iso-surfaces de la stratification verticale, et ainsi améliorer le transfert de l'énergie potentielle vers l'énergie cinétique. Ce bruit non-stationnaire nous permet en particulier d'améliorer la variabilité intrinsèque du courant à grande échelle. Cette amélioration est révélée par le biais d'une analyse des transferts d'énergie et au moyen de critères statistiques.

Le travail présenté dans ce chapitre provient du manuscrit Li et al. (2021) en préparation de soumission.

Enfin, nous terminons ce mémoire par une conclusion générale en rappelant les résultats principaux de cette thèse. Quelques perspectives sur les travaux de recherche futurs sont également données.

INTRODUCTION

Context

Mesoscale eddies contain a significant proportion of ocean energy and have an important impact on large-scale circulations. They are found everywhere in the ocean, and are particularly intensive in the western boundary currents like the Gulf Stream and the Antarctic Circumpolar Current. Unfortunately, to fully resolve these eddies in numerical simulations, a horizontal resolution of $\sim$ 10km is required, which is far too expensive for a large ensemble of realizations or simulations over a long time duration. Neglecting mesoscale eddy effects may lead to strong errors in the evolution of the large-scale dynamics. Therefore, they need to be properly modeled or parametrized.

A classical parametrization approach is to introduce eddy viscosity in coarse models to mimic the action of the computationally unresolved scales while simultaneously ensuring numerical stability by avoiding pile up of energy at the cutoff scale. The explicit dissipation mechanism is often represented either by a harmonic or biharmonic friction term with uniform coefficient, or through functional operators (Smagorinsky, 1963; Leith, 1971; Griffies and Hallberg, 2000) that depend on the resolved flow. However, encoding only large-scale dissipation in coarse models often leads to an excessive decrease of the resolved kinetic energy (Arbic et al., 2013; Kjellsson and Zanna, 2017). A more widely adopted approach in global ocean models is the Gent-McWilliams parametrization (Gent and McWilliams, 1990; Gent et al., 1995; Treguier et al., 1997; Griffies, 1998), in which an eddy-induced velocity is introduced to flatten isopycnal surfaces and to produce efficient transfers of mean available potential energy to eddy kinetic energy.

An alternative approach is based on stochastic parametrization (Berloff, 2005; Grooms and Majda, 2014; Jansen and Held, 2014; Porta Mana and Zanna, 2014; Cooper and Zanna, 2015; Grooms et al., 2015; Zanna et al., 2017), which aims to introduce energy backscattering across scales. These models provide a marked benefit in improving the internal ocean variability, which can be paramount in ensemble forecasting and data assimilation. As a matter of fact, it is well known that models with poor variability usually lead to very low spread of the ensemble (Karspeck et al., 2013; Franzke et al.,

2015). Hence, assimilation systems tend to be over-confident in the model as compared to the observations (Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013). On the other hand, to overcome numerical instability brought by introducing random forcing, specific tuning parameters are often included in these parameterized models. The success of such tuning methods often do not extend into new flow regimes.

Stochastic parameterization techniques have been proposed for reduced order climate models based on rigorous homogenization techniques (Franzke and Majda, 2006; Franzke et al., 2015; Gottwald et al., 2017). These models rely on a scale-separation principle and introduce a linear stochastic *Ansatz* model with damping terms for the nonlinear small-scale evolution equation. The resulting homogenized dynamics are cubic with correlated additive and multiplicative (CAM) noises. In the absence of scale-separation, the system usually becomes non-Markovian and incorporates memory terms, as shown in the Mori-Zwanzig equation (Gottwald et al., 2017).

Alternatively, Mémin (2014) proposed a consistent stochastic framework defined from physical conservation laws. This derivation keeps the full nonlinearity of the system yet relies on a strong temporal scale-separation assumption. Within this framework, the Lagrangian velocity is decomposed into a smooth component and a highly oscillating random field. A stochastic transport principle is subsequently derived using stochastic calculus. Notably, the resulting evolution of a random tracer includes a multiplicative random forcing, a heterogeneous diffusion and an advection correction due to inhomogeneity of the random flow component. With these additional terms, a remarkable energy conservation property along time for any realization of the advected tracer still holds (Resseguier et al., 2017a). This stochastic transport principle has been used as a fundamental tool to derive stochastic representations of large-scale geophysical dynamics (Resseguier et al., 2017a,b,c; Chapron et al., 2018) in which the missing contributions of unresolved processes are explicitly taken into account. Similar approaches based on the same decomposition have also been recently proposed by Holm (2015); Cotter et al. (2019a,b); Crisan et al. (2019); Gugole and Franzke (2019); Holm (2019); Cotter et al. (2020); Gugole and Franzke (2020).

The performance of such a random model has been evaluated and analyzed in terms of uncertainty quantification and ensemble forecasting (Resseguier et al., 2020a) for a surface quasi-geostrophic (SQG) flow. A more efficient spread is produced by the proposed model compared to a deterministic model with perturbed initial condition. As discussed above, this ability is essential for data assimilation applications. Recently, a stochastic

barotropic quasi-geostrophic (QG) model within this setting has been proposed (Bauer et al., 2020a) to study the structuration effect of the random field on the large-scale flow. Numerical results illustrate that, the introduction of an inhomogeneous random component into a propagating monochromatic Rossby wave induces the formation of extra large vortices. Furthermore, in Bauer et al. (2020b), this stochastic representation is tested for the coarse simulation of a barotropic circulation in a shallow ocean basin, driven by a symmetric double-gyre wind forcing. After reaching a turbulent equilibrium state, a statistical analysis of tracers shows that the proposed random model enables us to reproduce accurately, on a coarse mesh, the local structures of the first four statistical moments (mean, variance, skewness and kurtosis) of the high-resolution eddy-resolved data. Such a numerical evaluation has been extended recently to a wind driven layered QG system in a closed basin. For different noise models very good results have been obtained at coarse climatic scale.

Preview of chapters

This thesis is mainly based on three published papers (Bauer et al., 2020a,b; Resseguier et al., 2020a), one submitted manuscript (Brecht et al., 2021) and one working manuscript (Li et al., 2021). This thesis is organized in eight chapters with a general introduction and conclusion. We briefly resume below the content of those chapters.

Summary of chapter 1 – Preliminaries: Stochastic calculus

This chapter reviews some fundamental definitions and properties of stochastic calculus. These mathematical formulas are widely adopted in this study to derive the physical models described in the following chapters. Since all these results are well-known, very few proofs are presented in this chapter. We present general notions on stochastic processes and integrals in finite dimensions and then describe the extension of this calculus to infinite dimensions.

Summary of chapter 2 – Transport under location uncertainty

In order to account for the effects of the small scales on the evolution of the large-scale flow, a stochastic representation based on a physical transport principle has been first proposed by Mémin (2014) and then applied by Resseguier et al. (2017a) to different

geophysical fluid dynamics. This framework arises from a decomposition of the Lagrangian velocity into a smooth in time component and a highly oscillating term. One important characteristic of this random model is that it conserves the energy of any transported tracer. This chapter fully explains the proposed random model with its main properties using the stochastic calculus tools reviewed in Chapter 1. The work presented in this chapter has been published in Bauer et al. (2020a).

Summary of chapter 3 – Three-dimensional stochastic models

This chapter reviews some important three-dimensional equations under the location uncertainty framework. The core of these models is based on the stochastic Reynolds transport theorem, presented in the previous chapter. We first describe briefly the derivation of the stochastic governing equations of fluid motion, including the stochastic conservation of mass and of linear momentum. In order to do some comparisons with the classical large-eddy simulation framework, we present a pseudo-stochastic model based on a scale-separation assumption. Then a stochastic representation of the vorticity evolution is derived using classical vector calculus. Later, for large-scale atmospheric and oceanic circulations, two stochastic primitive systems are presented under the classical hydrostatic balance and the Boussinesq approximations. These equations will be used to derive several two-dimensional stochastic equations in the subsequent chapter. In addition, we build a connection between the proposed random model and the Craik-Leibovich system. In particular, we show that the effective advection due to the effect of statistical inhomogeneity of the small-scale flow on the large-scale current, can be considered as a generalization of the Stokes drift. Finally, a connection with the Gent-McWilliams parametrization is performed. The work presented in this chapter has been partly published in Bauer et al. (2020a).

Summary of chapter 4 – Stochastic geostrophic approximations

This chapter reviews some important quasi-two-dimensional stochastic equations. Apart from the hydrostatic balance in vertical direction presented in the previous chapter, this chapter focuses on another important approximation in the horizontal direction between the pressure gradient and the Coriolis forces, the so-called geostrophic balance, which is mainly due to the dominant rotating effect of the large-scale atmospheric and oceanic circulations. We first derive the stochastic rotating shallow water (RSW) system, which is

a good approximation of the primitive equations. Then, the energy conservation of a simplified stochastic RSW system is demonstrated. Latter, we specify some scaling numbers to adimensionalize the stochastic RSW system and the stochastic simple Boussinesq system presented in the previous chapter. Subsequently, asymptotic approaches are adopted to develop the barotropic planetary geostrophic (PG) and quasigeostrophic (QG) equations using the non-dimensional RSW system, as well as the stochastic primitive PG and vertically continuous version of the QG equations using the non-dimensional Boussinesq system. These resulting two-dimensional models have been successfully tested in several numerical simulations and some interesting results are outlined in chapters 6, 7 and 8. The work presented in this chapter are based on the published paper Bauer et al. (2020a) and on the manuscript Brecht et al. (2021) that has been submitted.

Summary of chapter 5 – Numerical modeling of uncertainty

In order to perform numerical simulations of the proposed random models, the uncertainty field has to be *a priori* modeled. We explore in particular methods based on the spectral decomposition defined from the eigenfunction basis of the spatial covariance. This chapter presents some numerical methods to estimate the empirical orthogonal functions (EOF). The work presented in this chapter have been published in Bauer et al. (2020a,b); Resseguier et al. (2020a).

Summary of chapter 6 – Ensemble forecasts verification for SQG dynamics

In this chapter, we first detail some important metrics to quantify ensemble forecasts prediction skills, such as the Talagrand diagram, the continuous ranked proper score and the energy score. As a test case, a simple geophysical fluid dynamics model – the stochastic surface quasigeostrophic (SQG) model – is considered. Several stochastic parameterizations presented in the previous chapter are also compared for short-terms ensemble forecasts. We show that the proposed random model, under both homogeneous and heterogeneous uncertainty, provides more efficient ensemble spread than a deterministic model with a perturbation of the initial condition. This ability is in particular essential for data assimilation applications. The work presented in this chapter has been published in Resseguier et al. (2020a).

Summary of chapter 7 – Numerical studies of stochastic barotropic QG model

This chapter provides some numerical results of the stochastic barotropic QG system derived in Chapter 4. Using a simple inviscid barotropic model, we first show that the introduction of inhomogeneous noise induces a structuration of the large-scale flow with strong secondary vortices. Later, the performance of this stochastic barotropic model is assessed for the numerical simulation of an idealized wind-driven double-gyre configuration within an enclosed shallow basin at mid-latitude. We focus then on the ability of the proposed stochastic models to accurately represent at a coarse resolution the four first statistical moments (mean, variance, skewness and kurtosis) of the flow. Comparing this statistical distribution through its four moments to that predicted by the eddy-resolving data enables us to qualify and quantify the accuracy of our stochastic representation of mesoscale eddy effects on large–scale circulation. The work presented in this chapter have been published in Bauer et al. (2020a,b).

Summary of chapter 8 – Numerical studies of stochastic multi-layer QG model

An approximative QG model requires only few layers to capture the baroclinic instabilities. However, the effects of mesoscale eddies can still be missing for coarse configurations in the horizontal direction. In this chapter, we continue our studies on the idealized double-gyre circulation yet now with a multi-layer QG model. In particular, we focus on the reproduction of the meandering jet as well as the prediction of low-frequency variability for the proposed random model on a coarse-grid. The work presented in this chapter is based on the manuscript Li et al. (2021) in preparation for submission.

A general conclusion ends this document in recalling the principal results of this thesis. Some perspectives on future research works are provided.

Part I

Stochastic models

Chapter 1

Preliminaries: Stochastic calculus

Abstract

This chapter reviews some fundamental definitions and properties of stochastic calculus. These mathematical formulas are widely adopted to derive the physical models described the following chapters. Since all these results are well-known, very few proofs are presented in this work. We will first present the stochastic processes and integrals in finite dimensions and then describe the extension of this calculus to infinite dimensions based on linear analysis.

1.1 Finite-dimensional case

This section describes the basics of stochastic calculus in finite dimensions. We first briefly describe the stochastic processes, paying particular attentions to the standard Brownian motions. Then, we review a general class of random processes – the real-valued continuous semimartingales. Later, we adopt such processes to define the Itô integrals. In particular, we point out some important properties of Itô calculus such as the chain rule. Finally, we briefly describe the Stratonovich integrals with its related formulas. These results mainly come from Le Gall (2016); Kunita (1997).

1.1.1 Stochastic process

Let us first recall the general definition of a stochastic process. In the following, let $(\Xi, \mathbb{P})$ be a probability space with Ξ the sample space and $\mathbb{P}$ the probability measure. In addition, let T be an index set and S be a measurable state space.

A *stochastic process* X is defined as a collection of random variables (Kunita, 1997) such that

$$\begin{aligned} X\colon (T \times \Xi) &\to S \\ (t,\ \xi) &\mapsto X_t(\xi), \end{aligned} \tag{1.1}$$

in which $\xi \mapsto X_t(\xi)$ is a random variable for each $t \in T$ given, and $t \mapsto X_t(\xi)$ is called a realization (or trajectory, path, sample, particle) for each $\xi \in \Xi$ given. For simplicity of notation, the randomness symbol ξ is often dropped in literature; this will be the case in this work.

A stochastic process can be continuous ($T \subseteq \mathbb{R}^+$) or discrete ($T \subseteq \mathbb{N}^+$) in time. A simple class of stochastic process is defined in real space ($S \subseteq \mathbb{R}$). In particular, let us focus on the continuous real-valued Gaussian process. To this end, we first recall quickly the real Gaussian random variable. A real random variable X follows the *Gaussian* (or *normal*) distribution $\mathcal{N}(\mu, \sigma^2)$ with $\mu \in \mathbb{R}, \sigma > 0$, if its probability density function (pdf) is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}. \tag{1.2a}$$

In fact, the *expectation* and the *variance* of such Gaussian variable X are respectively given by

$$\mathbb{E}[X] \triangleq \int_\Xi X(\xi)\,\mathrm{d}\mathbb{P}(\xi) = \int_{\mathbb{R}} xp(x)\,\mathrm{d}x = \mu, \tag{1.2b}$$

$$\mathrm{Var}(X) \triangleq \mathbb{E}\Big[(X - \mathbb{E}[X])^2\Big] = \sigma^2, \tag{1.2c}$$

and the quantity $\sigma = \sqrt{\mathrm{Var}(X)}$ is known as the *standard derivation*. In particular, if $\mu = 0$, then X is said to be centered. Further, if $\sigma = 1$, then X is termed as a standard Gaussian variable. Later, a $\mathbb{R}^d$-valued random variable $\boldsymbol{X} = (X_1, \ldots, X_d)^T$ is Gaussian if and only if the marginal $\boldsymbol{\alpha}^T\boldsymbol{X} = \sum_{i=1}^d \alpha_i X_i$ is a real-valued Gaussian variable for all $\boldsymbol{\alpha} \in \mathbb{R}^d$. Hereafter, a continuous real-valued stochastic process $\{X_t\}_{t\in\mathbb{R}^+}$ is Gaussian if $\boldsymbol{X} = (X_{t_1}, \ldots, X_{t_n})^T$ is a $\mathbb{R}^n$-valued Gaussian variable for any $t_1, \ldots, t_n \in \mathbb{R}^+$ and any $n \in \mathbb{N}$. As a result, all the marginals of a Gaussian process are Gaussian, and any linear combination of marginals of a Gaussian process is still Gaussian.

A typical example of the continuous real-valued Gaussian processes is the so-called standard Brownian motion, which is particularly important in the study of stochastic

differential equations (SDEs). A *standard Brownian motion* B satisfies the following characteristics:

(i) B starts from zero almost surely (*a.s.*), *i.e.* $\mathbb{P}(B_0 = 0) = 1$;

(ii) Each path $t \mapsto B_t$ is continuous;

(iii) Each time-increment of B is normally distributed, *i.e.* $B_t - B_s \sim \mathcal{N}(0, t-s)$ for $0 \leq s \leq t$;

(iv) The time-increment of B over disjoint intervals are independent, *i.e.* $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are mutually independent for $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$.

Note that two Gaussian variables X and Y are *independent* if and only if

$$\mathrm{Cov}(X, Y) \triangleq \mathbb{E}\Big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\Big] = 0. \tag{1.3}$$

Later, a finite dimensional ($S \subseteq \mathbb{R}^d$) Brownian motion, $\boldsymbol{B} = (B^1, \ldots, B^d)^\intercal$, is defined by supposing that each component B^i is a standard Brownian motion and that both components are independent.

1.1.2 Continuous semimartingale

In the following, let us focus on a general class of the real-valued stochastic processes – the continuous semimartingale. First of all, as functions of time, some stochastic processes can be of *bounded variation* (BV). Let $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ be a partition of the interval $[0, t]$ for $t \in \mathbb{R}^+$. A random process X is of BV if

$$\forall\, \xi \in \Xi, \qquad \sup_{\{p_n\}} \sum_{i=1}^{p_n} \left| X_{t_i^n}(\xi) - X_{t_{i-1}^n}(\xi) \right| < +\infty. \tag{1.4}$$

In general, most random processes are quite irregular, hence are not of BV. Nevertheless, they could be of bounded quadratic variation. We recall that a stochastic process X is of bounded *quadratic variation* if

$$\langle X, X \rangle_t \overset{\mathbb{P}}{=} \lim_{n \to +\infty} \sum_{i=1}^{p_n} \left(X_{t_i}^n - X_{t_{i-1}}^n \right)^2, \tag{1.5}$$

where the limit, if it exists, is defined in the sense of convergence in probability (Le Gall, 2016). A famous class of such random processes is the continuous martingale. A stochastic

process X is said to be a continuous *martingale* with respect to (*w.r.t.*) the process Y if

$$\mathbb{E}\Big[|X_t|\Big] < +\infty, \quad \forall\, t \in \mathbb{R}^+ \tag{1.6a}$$

$$\mathbb{E}\Big[X_t \,|\, \{Y_r\}_{r\leq s}\Big] = X_s, \quad \forall\, s \leq t, \tag{1.6b}$$

where the last equation means that the expected value of the state at time t, under a given set of past states up to time s, is equal to the state at time s. In particular, a random process can be a continuous martingale *w.r.t.* itself when $Y = X$. A typical example of such continuous martingale is the Brownian motion B with its well-known quadratic variation (Le Gall, 2016):

$$\langle B, B\rangle_t = t. \tag{1.7}$$

Now, we can introduce the continuous *semimartingale* X, which is defined as a sum of the process A of BV and the continuous martingale M:

$$X_t = X_0 + A_t + M_t. \tag{1.8}$$

A remarkable property of such processes is that they are of bounded quadratic covariation (or cross variance), which equals to that of their martingale components. Let $X_t = X_0 + A_t + M_t$ and $Y_t = Y_0 + A'_t + M'_t$ be two continuous semimartingales with two continuous martingales M, M' and two processes A, A' of BV, then their *quadratic covariation* is given by

$$\begin{aligned}\langle X, Y\rangle_t &\overset{\mathbb{P}}{=} \lim_{n\to+\infty} \sum_{i=1}^{p_n} \Big(X^n_{t_i} - X^n_{t_{i-1}}\Big)\Big(Y^n_{t_i} - Y^n_{t_{i-1}}\Big) \\ &= \langle M, M'\rangle_t. \end{aligned} \tag{1.9}$$

To better understand such quadratic-covariation-rule, we show briefly that if any of the two processes X or Y is of BV, then their quadratic covariation reduces to zero. For instance, let us assume that $X = A$ is of BV, then

$$\begin{aligned}\Big|\sum_{i=1}^{p_n} \Big(A^n_{t_i} - A^n_{t_{i-1}}\Big)\Big(Y^n_{t_i} - Y^n_{t_{i-1}}\Big)\Big| &\leq \sup_{i\leq p_n} \Big|Y^n_{t_i} - Y^n_{t_{i-1}}\Big| \sum_{i=1}^{p_n} \Big|A^n_{t_i} - A^n_{t_{i-1}}\Big| \\ &\leq C(t) \sup_{i\leq p_n} \Big|Y^n_{t_i} - Y^n_{t_{i-1}}\Big| \xrightarrow[n\to+\infty]{\mathbb{P}} 0, \end{aligned} \tag{1.10}$$

where the second inequality results from (1.4) and the last convergence (in probability) is based on the continuity of Y.

Let us highlight that if the decomposition (1.8) exists, then it is unique up to indistinguishability. More precisely, supposing that there exists a second decomposition of X with $X_t = X_0 + A'_t + M'_t$, then $\mathbb{P}(A_t = A'_t,\ \forall\, t \in \mathbb{R}^+) = 1$ and $\mathbb{P}(M_t = M'_t,\ \forall\, t \in \mathbb{R}^+) = 1$. This is the so-called *canonical decomposition* (Le Gall, 2016) of the semimartingale X. Indeed, this results from the following theorem:

Theorem 1.1.1 (Canonical decomposition)**.** *Let M be a continuous martingale. If M is also a process of BV with $M_0 = 0$, then $M_t = 0,\ \forall\, t \in \mathbb{R}^+$* a.s.

Hereafter, we can construct the stochastic integrals using the continuous semimartingales. In the following, we drop the word "continuous" for simplicity.

1.1.3 Itô integrals and formulas

This section describes the Itô representation of stochastic integrals. To this end, we first define it *w.r.t.* the martingales and then extended to that *w.r.t.* the semimartingales.

Let Θ be a (locally) bounded process (Kunita, 1997) and M be a martingale, then the following integral is well defined and it is also a martingale:

$$\int_0^t \Theta_s \,\mathrm{d}M_s \overset{\mathbb{P}}{=} \lim_{n\to+\infty} \sum_{i=1}^{p_n} \Theta^n_{t_{i-1}} \big(M^n_{t_i} - M^n_{t_{i-1}}\big). \tag{1.11}$$

Under such definition, Itô integrals satisfy the following properties:

(*Associativity*)

$$\Big\langle \int_0^{\cdot} \Theta_s \,\mathrm{d}M_s,\ \int_0^{\cdot} \Theta'_s \,\mathrm{d}M'_s \Big\rangle_t = \int_0^t \Theta_s \Theta'_s \,\mathrm{d}\langle M, M'\rangle_s, \tag{1.12a}$$

(*Centrality*)

$$\mathbb{E}\Big[\int_0^t \Theta_s \,\mathrm{d}M_s\Big] = 0, \tag{1.12b}$$

(*Isometry*)

$$\mathbb{E}\Big[\Big(\int_0^t \Theta_s \,\mathrm{d}M_s\Big)\Big(\int_0^t \Theta'_s \,\mathrm{d}M'_s\Big)\Big] = \mathbb{E}\Big[\int_0^t \Theta_s \Theta'_s \,\mathrm{d}\langle M, M'\rangle_s\Big], \tag{1.12c}$$

where Θ' is another bounded process and M' is another martingale.

In particular, the quadratic variation of the Itô integrals driven by Brownian motions reduces to

$$\left\langle \int_0^{\cdot} \Theta_s \, \mathrm{d}B_s, \int_0^{\cdot} \Theta_s \, \mathrm{d}B'_s \right\rangle_t = \int_0^t \Theta_s^2 \, \mathrm{d}s. \tag{1.13a}$$

Besides, Equations (1.12b) and (1.12c) give immediately the variance of such process, namely

$$\mathrm{Var}\left[\int_0^t \Theta_s \, \mathrm{d}B_s \right] = \mathbb{E}\left[\int_0^t \Theta_s^2 \, \mathrm{d}s \right]. \tag{1.13b}$$

More generally, the following Itô integral defined *w.r.t.* a semimartingale $X_t = X_0 + A_t + M_t$ (with A of BV and M a martingale) is also a semimartigale:

$$\int_0^t \Theta_s \, \mathrm{d}X_s = \int_0^t \Theta_s \, \mathrm{d}A_s + \int_0^t \Theta_s \, \mathrm{d}M_s, \tag{1.14}$$

where the first component on the right-hand-side (RHS) is a Stieltjes integral (Kunita, 1997) of BV, whereas the latter component is an Itô integral as defined in (1.11). Furthermore, the quadratic-covariation-rule (1.9) provides the associativity of the previous Itô integral, that is

$$\left\langle \int_0^{\cdot} \Theta_s \, \mathrm{d}X_s, \int_0^{\cdot} \Theta'_s \, \mathrm{d}Y_s \right\rangle_t = \int_0^t \Theta_s \Theta'_s \, \mathrm{d}\langle M, M' \rangle_s, \tag{1.15}$$

where $Y_t = Y_0 + A'_t + M'_t$ is another semimartingale.

Afterwards, we give two major results of Itô calculus, namely the integration-by-part formula and the chain rule. Compared to the classical rules for deterministic functions, the *Itô's integration-by-part formula* of two semimartingales includes an additional quadratic covariation term, it reads:

Theorem 1.1.2 (Itô's integration-by-part formula)**.** *Let X and Y be two semimartingales, then*

$$X_t \, Y_t = X_0 \, Y_0 + \int_0^t X_s \, \mathrm{d}Y_s + \int_0^t Y_s \, \mathrm{d}X_s + \langle X, Y \rangle_t. \tag{1.16}$$

Moreover, the chain rule of a function f composed by a semimartingale is provided by the famous *Itô's formula*. In the following, we give these formulas for multi-dimensional semimartingales . In a simple case when f is smooth and deterministic, this reads:

Theorem 1.1.3 (Itô's formula I)**.** *Let $\boldsymbol{X}$ be a $\mathbb{R}^d$-valued semimartingale. Let $f \in C^1([0,T])$*

for any $T \in \mathbb{R}^+$ *and* $f \in C^2(\mathbb{R}^d)$*, then*

$$f(\boldsymbol{X}_t, t) = f(\boldsymbol{X}_0, 0) + \int_0^t \frac{\partial f}{\partial s}(\boldsymbol{X}_s, s)\,\mathrm{d}s + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(\boldsymbol{X}_s, s)\,\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{X}_s, s)\,\mathrm{d}\langle X^i, X^j\rangle_s. \tag{1.17}$$

Taking its differential yields a SDE. Let us outline that the previous formula only applies to deterministic functions. Now, if the function f itself is random, a *generalized Itô's formula* (or *Itô-Wentzell formula*) is required instead. This formula from Kunita (1997) states that

Theorem 1.1.4 (Generalized Itô's formula I)**.** *Let* $f(\boldsymbol{x}, t)$ *be* C^2*-process over* $\boldsymbol{x} \in \mathbb{R}^d$ *and* C^1*-semimartingale over* $t \in \mathbb{R}^+$*, and let* $\boldsymbol{X}$ *be a semimartingale taking values in* $\Omega \subset \mathbb{R}^d$*. Then,* $f(\boldsymbol{X}_t, t)$ *is a semimartingale satisfying*

$$f(\boldsymbol{X}_t, t) = f(\boldsymbol{X}_0, 0) + \int_0^t \mathrm{d}_s f(\boldsymbol{X}_s, s) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(\boldsymbol{X}_s, s)\,\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{X}_t, t)\,\mathrm{d}\langle X^i, X^j\rangle_s + \sum_{i=1}^d \mathrm{d}\Big\langle \int_0^{\cdot} \frac{\partial f}{\partial x_i}(\boldsymbol{X}_s, \mathrm{d}s), X^i \Big\rangle_t. \tag{1.18}$$

where the second term on the RHS, $\mathrm{d}_s f(\boldsymbol{x}, s) \triangleq f(\boldsymbol{x}, s + \mathrm{d}s) - f(\boldsymbol{x}, s)$*, stands for the forward time-increment of the random function* f *at a fixed point* $\boldsymbol{x} \in \Omega$*.*

We remark that the last quadratic covariation constitutes an additional term compared to the classical Itô's formula (1.1.3). This term describes how the random function's gradient covaries with the composed process in time.

To summarize this section, the quadratic (co-)variations play a very important roles in Itô calculus. The generalized Itô's formula must be adopted as the chain rule when the composition function itself is random.

1.1.4 Stratonovich integrals and formulas

For comparison reason, we provide here an alternative representation of the stochastic integrals, that is the Stratonovich integral. Unlike the Itô integral (1.11), a *Stratonovich*

integral defined *w.r.t.* a martingale is no more a martingale but a semimartingale:

$$\int_0^t \Theta_s \circ \mathrm{d}M_s \overset{\mathbb{P}}{=} \lim_{n\to+\infty} \sum_{i=1}^{p_n} \frac{\Theta^n_{t_{i-1}} + \Theta^n_{t_i}}{2} \Big(M^n_{t_i} - M^n_{t_{i-1}} \Big), \tag{1.19}$$

where the symbol $\circ$ is often used to denote Stratonovich integrals and highlights explicitly the difference from Itô integrals. We remark that in most cases, it is possible to convert Stratonovich integral into Itô integral (and vice versa) when one of those integrals is given, using the formula (Kunita, 1997):

Theorem 1.1.5 (Stratonovich-Itô-integral-conversion rule)**.** *Let Y be a semimartingale, if in addition, X is also a semimartingale, then the following Stratonovich integral is well defined:*

$$\int_0^t X_s \circ \mathrm{d}Y_s = \int_0^t X_s \, \mathrm{d}Y_s + \frac{1}{2}\langle X, Y \rangle_t. \tag{1.20}$$

Unlike Itô integrals, the Stratonovich integral is defined such that the integration-by-part formula as well as the chain rule of the classical calculus hold (Kunita, 1997):

Theorem 1.1.6 (Integration-by-part formula II)**.** *Let X and Y be two semimartingales, then*

$$X_t \, Y_t = X_0 \, Y_0 + \int_0^t X_s \circ \mathrm{d}Y_s + \int_0^t Y_s \circ \mathrm{d}X_s. \tag{1.21}$$

Theorem 1.1.7 (Itô's formula II)**.** *Let $\boldsymbol{X}$ be a $\mathbb{R}^d$-valued semimartingale. Let $f \in C^2([0,T])$ for any $T \in \mathbb{R}^+$ and $f \in C^3(\mathbb{R}^d)$, then*

$$f(\boldsymbol{X}_t, t) = f(\boldsymbol{X}_0, 0) + \int_0^t \frac{\partial f}{\partial s}(\boldsymbol{X}_s, s) \, \mathrm{d}s + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(\boldsymbol{X}_s, s) \circ \mathrm{d}X^i_s. \tag{1.22}$$

Theorem 1.1.8 (Generalized Itô's formula II)**.** *Let $f(\boldsymbol{x}, t)$ be C^3-process over $\boldsymbol{x} \in \mathbb{R}^d$ and C^2-semimartingale over $t \in \mathbb{R}^+$, and let $\boldsymbol{X}$ be a semimartingale taking values in $\Omega \subset \mathbb{R}^d$. Then, $f(\boldsymbol{X}_t, t)$ is a semimartingale satisfying*

$$f(\boldsymbol{X}_t, t) = f(\boldsymbol{X}_0, 0) + \int_0^t \mathrm{d}_s \circ f(\boldsymbol{X}_s, s) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(\boldsymbol{X}_s, s) \circ \mathrm{d}X^i_s. \tag{1.23}$$

where the second term on the RHS, $\mathrm{d}_s \circ f(\boldsymbol{x}, s) \triangleq f(\boldsymbol{x}, s + \mathrm{d}s/2) - f(\boldsymbol{x}, \mathrm{d}s/2)$, stands for the central time-increment of the random function f at a fixed point $\boldsymbol{x} \in \Omega$.

Due to these properties the Stratonovich integral is often easier to manipulate in formal calculation. However, it requires more regular assumptions in both time and space for the

random function. In this work, from time to time, it will be interesting to use Stratonovich notation. For instance, we will adopt it to show the energy conservation for the derived stochastic models in following chapters. Nevertheless, this work focus mainly on the Itô calculus setting taking the advantage of martingale property and of explicit expansion when deriving new stochastic models. As shown in the subsequent section, each term in a stochastic transport equation can be given a physical interpretation thanks to the Itô representation.

1.2 Infinite-dimensional case

The previous sections allowed us to define the real-valued stochastic integrals and formulas in finite dimensions. However, a random process may possibly take values in a functional space of infinite dimensions. In particular, this section focuses on two typical Hilbert space-valued processes, the Q-Wiener process and the cylindrical Wiener process, which play an important role in studying the stochastic partial differential equations (SPDEs). Before introducing them, let us first review some useful results of spectral theory of linear operators. These results mainly come from Da Prato and Zabczyk (2014); Lord et al. (2014).

1.2.1 Spectral theory of linear operators

We recall that a *Hilbert space* $\mathbb{H}$ is a vector space with inner product $[\![\cdot,\cdot]\!]_{\mathbb{H}}$ and is complete *w.r.t.* the induced norm $\|u\|_{\mathbb{H}} \triangleq [\![u,u]\!]_{\mathbb{H}}^{1/2}$. In particular, any Hilbert space is a *Banach space* (complete normed vector space). An interesting example of Hilbert space is $L^2(\Omega)$ or $(L^2(\Omega))^d$. For simplicity, we focus on $L^2(\Omega)$. Let $\Omega \subset \mathbb{R}^d$ be a domain, then $L^2(\Omega)$ is a real Hilbert space with the inner product $[\![u,v]\!]_{\mathbb{H}} \triangleq \int_\Omega u(\boldsymbol{x})v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

An important property of Hilbert spaces is the orthogonal projection. Let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal set in $\mathbb{H}$, *i.e.* $[\![\phi_n,\phi_m]\!]_{\mathbb{H}} = \delta_{nm}$ [1]. In addition, if the linear subspace spanned by $\{\phi_n\}_{n\in\mathbb{N}}$ is dense in $\mathbb{H}$, then the set $\{\phi_n\}_{n\in\mathbb{N}}$ is a *complete orthonormal basis* for $\mathbb{H}$ such that

$$u = \sum_{n=1}^{\infty} [\![u,\phi_n]\!]_{\mathbb{H}}\,\phi_n, \qquad \|u\|_{\mathbb{H}}^2 = \sum_{n=1}^{\infty} [\![u,\phi_n]\!]_{\mathbb{H}}^2, \qquad \forall\, u \in \mathbb{H}. \tag{1.24}$$

1. δ denotes the Kronecker symbol with $\delta_{nm} = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$

Further, a Hilbert space is *separable* if it contains a countable dense subset and every separable Hilbert space has an orthonormal basis.

We now briefly review the bounded linear operators. Let X and Y be two vector spaces. A function $\mathrm{L} : X \to Y$ is said to be a *linear operator* if, $\forall\, u, v \in X,\ \forall\, \alpha \in \mathbb{R}$, $\mathrm{L}(\alpha u+v) = \alpha \mathrm{L}u+\mathrm{L}v$. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces. A linear operator $\mathrm{L} : X \to Y$ is *bounded* if for some $C > 0$, $\|\mathrm{L}u\|_Y \leq C\|u\|_X,\ \forall\, u \in X$. Later, $\mathcal{L}(X, Y)$ denotes the set of bounded linear operators. In the case $X = Y$, we write simply $\mathcal{L}(X)$. In fact, the set $\mathcal{L}(X, Y)$ is a Banach space with norm

$$\|\mathrm{L}\|_{\mathcal{L}(X,Y)} \triangleq \sup_{u\neq 0} \frac{\|\mathrm{L}u\|_Y}{\|u\|_X}. \tag{1.25}$$

Therefore, bounded linear operators are continuous: $\|\mathrm{L}u\|_Y \leq \|\mathrm{L}\|_{\mathcal{L}(X,Y)}\|u\|_X$.

An important class of bounded linear operators are the Hilbert-Schmidt operators. Let $\mathbb{G}$ and $\mathbb{H}$ be two separable Hilbert spaces with norms $\|\cdot\|_{\mathbb{G}}$ and $\|\cdot\|_{\mathbb{H}}$ respectively. Let us then define the *Hilbert-Schmidt norm* for an orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$ of $\mathbb{G}$ by

$$\|\mathrm{L}\|_{\mathrm{HS}(\mathbb{G},\mathbb{H})} \triangleq \Big(\sum_{n=1}^{\infty} \|\mathrm{L}\phi_n\|_{\mathbb{H}}^2\Big)^{1/2}. \tag{1.26}$$

Subsequently, the set $\mathrm{HS}(\mathbb{G}, \mathbb{H}) \triangleq \{\mathrm{L} \in \mathcal{L}(\mathbb{G}, \mathbb{H}) : \|\mathrm{L}\|_{\mathrm{HS}(\mathbb{G},\mathbb{H})} < \infty\}$ is a Banach space with the Hilbert-Schmidt norm. As such, an $\mathrm{L} \in \mathrm{HS}(\mathbb{G}, \mathbb{H})$ is known as a *Hilbert-Schmidt operator*. In the case $\mathbb{G} = \mathbb{H}$, we denote its norm by $\|\mathrm{L}\|_{\mathrm{HS}}$ for simplicity. In particular, Hilbert-Schmidt operators are bounded with $\|\mathrm{L}\|_{\mathcal{L}(\mathbb{H})} \leq \|\mathrm{L}\|_{\mathrm{HS}}$.

We now take a look to the Hilbert-Schmidt operators on $L^2(\Omega)$. Let us first recall that an *integral operator* K on $L^2(\Omega)$ with a bounded *kernel* $k \in L^2(\Omega \times \Omega)$ is defined by

$$(\mathrm{K}u)(\boldsymbol{x}) \triangleq \int_\Omega k(\boldsymbol{x}, \boldsymbol{y})u(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \qquad \boldsymbol{x} \in \Omega,\ u \in L^2(\Omega). \tag{1.27}$$

In fact, the relationship between the integral operator and the Hilbert-Schmidt operator on $L^2(\Omega)$ is provided by the following theorem:

Theorem 1.2.1 (Hilbert-Schmidt integral operator)**.** *Any integral operator with kernel $k \in L^2(\Omega \times \Omega)$ is a Hilbert-Schmidt operator on $L^2(\Omega)$. On the other hand, any Hilbert-Schmidt operator* K *on $L^2(\Omega)$ can be written as* (1.27) *with $\|\mathrm{K}\|_{HS} = \|k\|_{L^2(\Omega\times\Omega)}$.*

Let us then present the spectral theory of compact self-adjoint operators. Let X and Y be two Banach spaces. A linear operator $L : X \to Y$ is *compact* if the closure of

$\mathrm{L}(D) = \{\mathrm{L}u : u \in D\}$ is compact in Y for any bounded set D in X. In particular, the integral operator K on $L^2(\Omega)$ with kernel $k \in L^2(\Omega \times \Omega)$ defined by (1.27) is compact. Let $\mathbb{H}$ be a Hilbert space, an $\mathrm{L} \in \mathcal{L}(\mathbb{H})$ is *self-adjoint* on $\mathbb{H}$ if $[\![\mathrm{L}u, v]\!]_{\mathbb{H}} = [\![u, \mathrm{L}v]\!]_{\mathbb{H}}$ for $u, v \in \mathbb{H}$. Back to $L^2(\Omega)$, if the kernel $k \in L^2(\Omega \times \Omega)$ is symmetric, *i.e.* $k(\boldsymbol{x}, \boldsymbol{y}) = k(\boldsymbol{y}, \boldsymbol{x})$ for $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, then the integral operator K in (1.27) is self-adjoint.

Let $\mathrm{L} \in \mathcal{L}(\mathbb{H})$ and if there exists a non-zero $\phi \in \mathbb{H}$ such that $\mathrm{L}\phi = \lambda\phi$, we call λ an *eigenvalue* and ϕ an *eigenfunction* of L. The major result concerning eigenvalues of compact self-adjoint operators, such as integral operators on $L^2(\Omega)$, is the following Hilbert-Schmidt spectral theorem:

Theorem 1.2.2 (Hilbert-Schmidt spectral)**.** *Let* $\mathrm{L} \in \mathcal{L}(H)$ *be self-adjoint and compact. Let us consider that* L *admits a set of eigenfunctions* $\{\phi_n\}_{n\in\mathbb{N}}$ *on* $\mathbb{H}$*, associated with a set of decaying eigenvalues,* i.e. $\forall\, n\ |\lambda_n| \geq |\lambda_{n+1}|$*. Then, all eigenvalues are real with* $\lambda_n \xrightarrow[n\to+\infty]{} 0$*, the eigenfunctions* $\{\phi_n\}_{n\in\mathbb{N}}$ *form an orthonormal basis for* L *and*

$$\mathrm{L}u = \sum_{n=1}^{\infty} \lambda_n \, [\![u, \phi_n]\!]_{\mathbb{H}} \, \phi_n, \qquad \forall\, u \in \mathbb{H}. \tag{1.28}$$

Indeed, this infinite-dimensional result can be considered as a generalization of the spectral decomposition for finite-dimensional matrices.

Finally, we describe the operators of trace class on a Hilbert space $\mathbb{H}$ together with the Mercer's theorem on $L^2(\Omega)$. Recall that a linear operator $\mathrm{L} \in \mathcal{L}(\mathbb{H})$ is called *non-negative definite* if $[\![u, \mathrm{L}u]\!]_{\mathbb{H}} \geq 0$ for any $u \in \mathbb{H}$. In particular, if $k \in L^2(\Omega \times \Omega)$ is a non-negative definite function, *i.e.* $\sum_{i,j=1}^{N} \alpha_i \alpha_j k(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq 0$, $\forall\, \boldsymbol{x}_i \in \Omega$, $\forall\, \alpha_i \in \mathbb{R}$, then the integral operator K on $L^2(\Omega)$ with kernel k is non-negative definite. Let $\mathbb{H}$ be separable and let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal basis. A non-negative definite operator $\mathrm{L} \in \mathbb{H}$ is of *trace class* if

$$\operatorname{Tr} \mathrm{L} \triangleq \sum_{n=1}^{\infty} [\![\mathrm{L}\phi_n, \phi_n]\!]_{\mathbb{H}} < \infty, \tag{1.29}$$

where this sum is independent of the choice of the orthonormal basis. In particular, an operator $\mathrm{L} \in \mathcal{L}(\mathbb{H})$ is of trace class if L is equal to the composition of two Hilbert-Schmidt operators.

Theorem 1.2.3 (Mercer)**.** *Let* Ω *be a bounded domain, let* $k \in C(\bar{\Omega} \times \bar{\Omega})$ *be a symmetric and non-negative definite function and let* K *be the corresponding integral operator* (1.27)*.*

Then, K *admits eigenfunctions* ϕ_n *with eigenvalues* $\lambda_n > 0$ *such that* $\phi_n \in C(\bar{\Omega})$ *and*

$$k(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=1}^{\infty} \lambda_n \phi_n(\boldsymbol{x}) \phi_n(\boldsymbol{y}), \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \bar{\Omega}, \tag{1.30a}$$

where the sum converges in $C(\bar{\Omega} \times \bar{\Omega})$ *and* $\bar{\Omega}$ *is the closure of* Ω*. Furthermore, the integral operator* K *is of trace class such that*

$$\operatorname{Tr} \mathrm{K} = \int_{\Omega} k(\boldsymbol{x}, \boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \sum_{n=1}^{\infty} \lambda_n. \tag{1.30b}$$

1.2.2 Hilbert space-valued process

This section describes the Gaussian processes taking values in a Hilbert space $\mathbb{H}$ using the spectral theory presented in the previous section. In particular, we focus on the Q-Wiener process and the cylindrical Wiener process.

In a general way, an $\mathbb{H}$-valued stochastic process $\{X_t\}_{t \in \mathbb{R}^+}$ is *Gaussian* if for any $u \in \mathbb{H}$, the $\{[\![X_t, u]\!]_{\mathbb{H}}\}_{t \in \mathbb{R}^+}$ is a real-valued Gaussian process. In particular, Hilbert space-valued Gaussian processes can be defined by Hilbert space-valued Gaussian variables. An $\mathbb{H}$-valued random variable X is Gaussian if $[\![X, u]\!]_{\mathbb{H}}$ is a real-valued Gaussian random variable for all $u \in \mathbb{H}$. To introduce the probability distribution of such random variables, let us define a Banach space $L^2(\Xi, \mathbb{H})$ and the covariance operator. $L^2(\Xi, \mathbb{H})$ is the set of $\mathbb{H}$-valued random variables X satisfying

$$\|X\|_{L^2(\Xi,\mathbb{H})} \triangleq \Big(\int_{\Xi} \|X(\xi)\|_{\mathbb{H}}^2 \, \mathrm{d}\mathbb{P}(\xi) \Big)^{1/2} = \mathbb{E}\Big[\|X\|_{\mathbb{H}}^2 \Big]^{1/2} < \infty. \tag{1.31}$$

Further, $L^2(\Xi, \mathbb{H})$ is a Hilbert space with the inner product

$$[\![X, Y]\!]_{L^2(\Xi,\mathbb{H})} \triangleq \int_{\Xi} [\![X(\xi), Y(\xi)]\!]_{\mathbb{H}} \, \mathrm{d}\mathbb{P}(\xi) = \mathbb{E}\Big[[\![X, Y]\!]_{\mathbb{H}} \Big]. \tag{1.32}$$

A linear operator $Q : \mathbb{H} \to \mathbb{H}$ is the *covariance* of the $\mathbb{H}$-valued random variables X and Y if

$$[\![Qu, v]\!]_{\mathbb{H}} = \operatorname{Cov}\Big([\![X, u]\!]_{\mathbb{H}}, [\![Y, v]\!]_{\mathbb{H}} \Big), \qquad \forall\, u, v \in \mathbb{H}. \tag{1.33}$$

Now, let $\mathbb{H}$ be separable and let X be an $\mathbb{H}$-valued Gaussian variable with $\mu = \mathbb{E}[X]$. Then $X \in L^2(\Xi, \mathbb{H})$ and the covariance operator Q of X is symmetric non-negative definite and of trace class with $\operatorname{Tr} Q = \mathbb{E}\Big[\|X - \mu\|_{\mathbb{H}}^2 \Big]$. As such, we denote $X \sim \mathcal{N}(\mu, Q)$. Hereafter, an $\mathbb{H}$-valued process $\{X_t\}_{t \in \mathbb{R}^+}$ is Gaussian if each $[\![X_t, u]\!]_{\mathbb{H}}$ is an $\mathbb{H}$-valued Gaussian variable

for all $u \in \mathbb{H}$ and $t \in \mathbb{R}^+$.

We now focus on a particular example of Hilbert space-valued Gaussian processes, that is the Q-Wiener process. To this end, we assume in the following that $Q \in \mathcal{L}(\mathbb{H})$ is a symmetric non-negative definite operator of trace class in $\mathbb{H}$. From Section 1.2.1, the linear operator Q admits an orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$ of eigenfunctions with corresponding eigenvalues $\lambda_n \in \mathbb{R}^+$ such that $\sum_{n\in\mathbb{N}} \lambda_n < \infty$. As such, an $\mathbb{H}$-valued stochastic process B is referred to as a *Q-Wiener process* if

(i) $B_0 = 0$ *a.s.*;

(ii) Each path $t \mapsto B_t$ is continuous;

(iii) The time increments of B over disjoint intervals are independent;

(iv) $B_t - B_s \sim \mathcal{N}\big(0, (t-s)Q\big)$ for all $0 \le s \le t$.

A remarkable property of the Q-Wiener process is that it can be represented as a linear combination of the eigenfunctions of Q and standard Brownian motions. This is provided by the Karhunen-Loève expansion as follows:

Theorem 1.2.4 (Karhunen-Loève expansion)**.** *If $Q \in \mathcal{L}(\mathbb{H})$ is a self-adjoint non-negative definite operator of trace class in $\mathbb{H}$, then B is a Q-Wiener process if and only if*

$$B_t = \sum_{n=1} \sqrt{\lambda_n}\phi_n\beta_t^n \text{ a.s.}, \tag{1.34}$$

where $\beta_t^n = \frac{1}{\sqrt{\lambda_n}}[\![B_t, \phi_n]\!]_{\mathbb{H}}$ are the independent and identically distributed (i.i.d.) standard Brownian motions, the series on the RHS converges in $L^2(\Xi, \mathbb{H})$ and the equality converges in $L^2(\Xi, C([0,T]), \mathbb{H})$ for any $T > 0$.

Therefore, one can easily show that $B \sim \mathcal{N}(0, tQ)$. For instance, we have

$$\mathrm{Cov}\Big([\![B_t, \phi_n]\!]_{\mathbb{H}}, [\![B_t, \phi_m]\!]_{\mathbb{H}}\Big) = \sqrt{\lambda_n}\underbrace{\mathbb{E}[\beta_t^n\beta_t^m]}_{\delta_{nm}\,t}\sqrt{\lambda_m} = t\lambda_n. \tag{1.35}$$

Besides, the quadratic variation of such process can be specified by

$$\langle B, B\rangle_t = \sum_{n,m=1}^{\infty} \sqrt{\lambda_n\lambda_m}\underbrace{\langle\beta^n, \beta^m\rangle_t}_{\delta_{nm}\,t} = t\ \mathrm{Tr}\,Q. \tag{1.36}$$

In particular, for $\mathbb{H} = L^2(\Omega)$ with a bounded domain Ω, thanks to Theorem 1.2.2, Q can

be written as an integral operator with a kernel $q \in L^2(\Omega \times \Omega)$, namely

$$(Qu)(\boldsymbol{x}) = \int_\Omega q(\boldsymbol{x}, \boldsymbol{y}) u(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}, \qquad u \in L^2(\Omega). \tag{1.37}$$

Therefore, the two-points covariance of the Q-Wiener process reads

$$\mathrm{Cov}\Big(B_t(\boldsymbol{x}), B_t(\boldsymbol{y})\Big) = t \, q(\boldsymbol{x}, \boldsymbol{y}). \tag{1.38}$$

In the case that $Q = I$ (*i.e.* $\lambda_n = 1$, $\forall\, n \in \mathbb{N}$), Q is no more of trace class on the infinite-dimensional space $\mathbb{H}$. Thus, the previous series (1.34) does not converge in $L^2(\Xi, \mathbb{H})$. Nevertheless, we can extend such Q-Wiener process to the cylindrical Wiener process by introducing a "larger" Hilbert space $\mathbb{I}$ such that $\mathbb{H} \subset \mathbb{I}$. As such, by a specific inclusion $\iota : \mathbb{H} \to \mathbb{I}$, the covariance operator $Q = I$ could be of trace class in $\mathbb{I}$ (Da Prato and Zabczyk, 2014). Formally, we define the *cylindrical Wiener process* B as a $\mathbb{H}$-valued process satisfying

$$B_t = \sum_{n=1}^{\infty} \phi_n \beta_t^n, \tag{1.39}$$

where $\{\phi_n\}_{n \in \mathbb{N}}$ is any orthonormal basis of $\mathbb{H}$ and the series on RHS converges in $L^2(\Xi, \mathbb{I})$ if the inclusion $\iota : \mathbb{H} \to \mathbb{I}$ is a Hilbert-Schmidt operator.

Afterwards, one can define the Itô integrals *w.r.t.* the Hilbert space-valued Wiener process, starting from elementary processes and ending up with most complicated processes. This procedure is fully developed by Da Prato and Zabczyk (2014). Finally, most of the properties and formulas in the finite-dimensional case, as presented in Section 1.1.3, hold in infinite dimensions, yet under modified assumptions.

Chapter 2

Transport under location uncertainty

Abstract

In order to account for the effects of the small-scales on the evolution of the large-scale flow, a stochastic representation based on a physical transport principle, has been first proposed by Mémin (2014) and then applied by Resseguier et al. (2017a) to different geophysical fluid dynamics. This framework arises from a decomposition of the Lagrangian velocity into a smooth in time component and a highly oscillating term. One important characteristic of this random model is that it conserves the energy of any transported tracer. This chapter fully explains the proposed random model with its main properties using the stochastic calculus reviewed in Chapter 1. The work presented in this chapter has been published in Bauer et al. (2020a).

2.1 Stochastic flow

This section provides a stochastic representation of the flow dynamics termed modeling under location uncertainty (LU). Such random model is based on a temporal-scale-separation assumption of the following stochastic flow:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{v}(\boldsymbol{X}_t, t)\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\,\mathrm{d}\boldsymbol{B}_t, \tag{2.1}$$

where $\boldsymbol{X}$ is the Lagrangian particle trajectory defined within the bounded domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3), $\boldsymbol{v}$ is the large-scale velocity that is both spatially and temporally cor-

related, and $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is the small-scale uncertainty (also called noise) term that is only correlated in space. This term explicitly describes the flow location uncertainty. The randomness of such noise is driven by the cylindrical $\mathbf{I}_d$-Wiener process $\boldsymbol{B}_t$, meaning that each component B_t^i, $(i = 1, \dots, d)$ is a cylindrical Wiener process (see Definition 1.39) and the components are independent from each other. At each instant t, the spatial structure of such noise is specified through a deterministic integral operator $\boldsymbol{\sigma}(\cdot, t) : (L^2(\Omega))^d \to (L^2(\Omega))^d$ with a bounded [1] matrix kernel $\breve{\boldsymbol{\sigma}} = (\breve{\sigma}_{ij})_{i,j=1,\dots,d}$ such that

$$\boldsymbol{\sigma}(\boldsymbol{x}, t)\boldsymbol{f} \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\boldsymbol{f}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \qquad \forall \boldsymbol{f} \in (L^2(\Omega))^d. \tag{2.2}$$

Besides, let us introduce another integral operator $\boldsymbol{Q} \triangleq \boldsymbol{\sigma}\boldsymbol{\sigma}^T$ with a matrix kernel $\breve{\boldsymbol{q}}$ in the sense that

$$\breve{\boldsymbol{q}}(\boldsymbol{x}, \boldsymbol{y}, t) \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{x}', t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y}, \boldsymbol{x}', t)\,\mathrm{d}\boldsymbol{x}', \tag{2.3a}$$

$$\boldsymbol{Q}(\boldsymbol{x}, t)\boldsymbol{f} = \int_\Omega \breve{\boldsymbol{q}}(\boldsymbol{x}, \boldsymbol{y}, t)\boldsymbol{f}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \qquad \forall \boldsymbol{f} \in (L^2(\Omega))^d. \tag{2.3b}$$

Thanks to Theorem (1.2.1), the fact that the kernel $\breve{\boldsymbol{\sigma}}(\cdot, \cdot, t)$ is bounded at time t, implies that the integral operator $\boldsymbol{\sigma}(\cdot, t)$ is Hilbert-Schmidt on $(L^2(\Omega))^d$. Thus, the integral operator $\boldsymbol{Q}(\cdot, t)$ is of trace class on $(L^2(\Omega))^d$ with $\mathrm{Tr}\,\boldsymbol{Q}(\cdot, t) = \int_\Omega \breve{\boldsymbol{q}}(\boldsymbol{x}, \boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x} < +\infty$. From Definition (2.3a), the integral kernel $\breve{\boldsymbol{q}}$ is symmetric and is a non-negative definite function, hence the integral operator $\boldsymbol{Q}$ is self-adjoint, non-negative definite and compact within the bounded domain Ω.

On the other hand, since $\boldsymbol{B}_t$ is a cylindrical $\mathbf{I}_d$-Wiener process and $\boldsymbol{\sigma}$ is deterministic, the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is a centered (*i.e.* $\mathbb{E}[(\sigma \mathrm{d}B_t)^i] = 0$, $\forall\, i = 1, \dots, d$) Gaussian process. Furtherore, the two-points (for $\boldsymbol{x}, \boldsymbol{y} \in \Omega$) two-components (for $i, j = 1, \dots, d$) covariance of the noise can be deduced from the Itô isometry (1.12c), it reads:

$$\begin{aligned}
&\mathrm{Cov}\Big((\sigma(\boldsymbol{x}, t)\,\mathrm{d}B_t)^i, (\sigma(\boldsymbol{y}, t)\,\mathrm{d}B_t)^j\Big) \\
&= \mathbb{E}\Big[\Big(\int_\Omega \sum_{k=1}^d \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{x}', t)\,\mathrm{d}B_t^k(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'\Big)\Big(\int_\Omega \sum_{l=1}^d \breve{\sigma}_{jl}(\boldsymbol{y}, \boldsymbol{y}', t)\,\mathrm{d}B_t^l(\boldsymbol{y}')\,\mathrm{d}\boldsymbol{y}'\Big)\Big] \\
&= \mathbb{E}\Big[\int_\Omega \int_\Omega \sum_{k,l=1}^d \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{x}', t)\breve{\sigma}_{jl}(\boldsymbol{y}, \boldsymbol{y}', t)\,\mathrm{d}\underbrace{\langle B^k(\boldsymbol{x}'), B^l(\boldsymbol{y}')\rangle_t}_{\delta_{kl}\,\delta(\boldsymbol{x}'-\boldsymbol{y}')\,t}\,\mathrm{d}\boldsymbol{x}'\mathrm{d}\boldsymbol{y}'\Big],
\end{aligned} \tag{2.4}$$

1. In practice, one may consider that $\forall\, i, j = 1, \dots, d$, $\sup_{(\boldsymbol{x},\boldsymbol{y})\in\Omega} |\breve{\sigma}_{i,j}(\boldsymbol{x}, \boldsymbol{y}, t)| < +\infty$, which implies that $\breve{\sigma}_{ij}(\cdot, \cdot, t) \in L^2(\Omega \times \Omega)$ within the bounded domain Ω.

and using the definition (2.3a) of the kernel $\breve{\boldsymbol{q}}$, the last equality reduces to

$$\operatorname{Cov}\Big(\big(\sigma(\boldsymbol{x},t)\,\mathrm{d}B_t\big)^i,\big(\sigma(\boldsymbol{y},t)\,\mathrm{d}B_t\big)^j\Big)=\breve{q}_{ij}(\boldsymbol{x},\boldsymbol{y},t)\,\mathrm{d}t. \tag{2.5}$$

Therefore, as discussed above, the integral operator $\boldsymbol{Q}$ defined in (2.3b) is a well-defined covariance operator of the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$, namely

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{Q}\,\mathrm{d}t). \tag{2.6}$$

We remark that the above differential forms are informal notations of the Itô integrals. Formally, the stochastic Lagrangian trajectory $\boldsymbol{X}$ is a continuous semimartingale such that

$$\boldsymbol{X}_t=\boldsymbol{X}_0+\underbrace{\int_0^t\boldsymbol{v}(\boldsymbol{X}_s,s)\,\mathrm{d}s}_{\boldsymbol{A}_t}+\underbrace{\int_0^t\boldsymbol{\sigma}(\boldsymbol{X}_s,s)\,\mathrm{d}\boldsymbol{B}_s}_{\boldsymbol{M}_t}, \tag{2.7}$$

where the first integral $\boldsymbol{A}_t$ is a random process of bounded variation (see Definition 1.4) and the latter, $\boldsymbol{M}_t\sim\mathcal{N}\big(\boldsymbol{0},\int_0^t\boldsymbol{Q}(\cdot,s)\,\mathrm{d}s\big)$, is a continuous martingale (see Definition 1.6). In particular, if the integral operator $\boldsymbol{\sigma}$ (as well as its kernel $\breve{\boldsymbol{\sigma}}$) is stationary (*i.e.* independent of time), then $\boldsymbol{M}_t\sim\mathcal{N}(\boldsymbol{0},t\boldsymbol{Q})$ reduces to the $\boldsymbol{Q}$-Wiener process as presented in Section 1.2.2. In addition, we assume that the large-scale velocity $\boldsymbol{v}$ itself is a continuous semimartingale.

Under the LU framework, the noise strength is measured by its variance (or auto-covariance), denoted as $\boldsymbol{a}$, which is defined by

$$\boldsymbol{a}(\boldsymbol{x},t)\triangleq\breve{\boldsymbol{q}}(\boldsymbol{x},\boldsymbol{x},t). \tag{2.8}$$

As a result, the global variance over the domain is bounded, $\int_\Omega\boldsymbol{a}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}=\operatorname{Tr}\boldsymbol{Q}(\cdot,t)<+\infty$. We remark from Equation (2.5) that the variance tensor $\boldsymbol{a}$ has the same unit as a diffusion tensor ($\mathrm{m}^2\cdot\mathrm{s}^{-1}$). In addition, the density of the turbulent kinetic energy (that has a unit of $\mathrm{m}^2\cdot\mathrm{s}^{-2}$) can be specified by [2]$\mathrm{tr}(\boldsymbol{a})/(2\mathrm{d}t)$.

Note that the previous representation (2.2) is a general way to define the noise in LU models. In particular, the Mercer's theorem (1.2.3) ensures that the covariance operator $\boldsymbol{Q}(\cdot,t)$ at one instant t, admits an orthonormal eigenfunction basis $\{\boldsymbol{\phi}_n(\cdot,t)\}_{n\in\mathbb{N}}$ with the corresponding eigenvalues $\lambda_n(t)\geq 0$ such that $\operatorname{Tr}\boldsymbol{Q}(\cdot,t)=\sum_{n\in\mathbb{N}}\lambda_n(t)<\infty$. Therefore, one may equivalently define the noise and its variance based on the following spectral

2. Here, $\mathrm{tr}(\boldsymbol{a})=\sum_{i=1}^d a_{ii}$ denotes the trace of the matrix $\boldsymbol{a}=(a_{ij})_{i,j=1,\dots,d}$.

decomposition:

$$\boldsymbol{\sigma}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{B}_t = \sum_{n\in\mathbb{N}} \sqrt{\lambda_n}(t)\boldsymbol{\phi}_n(\boldsymbol{x},t)\,\mathrm{d}\beta_t^n, \tag{2.9a}$$

$$\boldsymbol{a}(\boldsymbol{x},t) = \sum_{n\in\mathbb{N}} \lambda_n(t)\boldsymbol{\phi}_n(\boldsymbol{x},t)\boldsymbol{\phi}_n^T(\boldsymbol{x},t), \tag{2.9b}$$

where β^n denotes the *i.i.d.* standard Brownian motions (see Section 1.1.1). In fact, this spectral decomposition can be considered as an extension of the Karhunen-Loève expansion (1.34) for a Q-Wiener process.

After introducing mathematically the LU model, let us now outline some remarks from a physical point of view. Contrary to traditional large-eddy simulation (LES) settings, the decomposition (2.1) corresponds to a temporal decomposition, but not to a spatial decomposition formulated through spatial filters and/or decimation operators. In (2.1), the large-scale resolved velocity component $\boldsymbol{v}$ corresponds to a smooth Lagrangian quantity. The time derivative of the noise term can be informally denoted by $\boldsymbol{\sigma}\dot{\boldsymbol{B}}_t/\mathrm{d}t$ in a distribution sense. It represents the fast and highly oscillating unresolved velocity component. In turbulent flows, time and spatial scales are related. In the inertial range, for three-dimensional turbulent flows, the turnover time ratio for two different scales, $\tau_L/\tau_l \propto (L/l)^{2/3}$, exhibits a direct relation between a change of time scale and a change of spatial resolution. A coarsening in time yields thus a space dilation. Efficient LES schemes based on Lagrangian averaging (Meneveau and Katz, 2000) or more specifically on temporal decomposition (2.1) have been assessed on several prototypical flows (Bauer et al., 2020a,b; Chandramouli et al., 2018; Resseguier et al., 2017b; Yang and Mémin, 2019). Therefore, for ease of understanding, we will adhere to the vocabulary of LES in the following chapters and refer to the first term as the large-scale/resolved velocity component of the fluid motion while the second term will be designated as the small-scale/unresolved velocity component.

Although the unresolved random component is simply a Gaussian field (2.6) by construction, it leads to a multiplicative non-Gaussian noise when incorporated in the transport equations. This will be shown in the subsequent section.

2.2 Stochastic transport operator

This section describes the evolution law of a conserved tracer with extensive property (*e.g.* temperature, salinity, buoyancy) transported by the stochastic flow (2.1). In the

following, we first interpret the result in terms of transport operator and then give its formal derivation.

Under the LU framework, a random tracer Θ transported along the stochastic trajectory $\boldsymbol{X}$ means that

$$\forall\, t \in \mathbb{R}^+, \qquad \Theta(\boldsymbol{X}_{t+\delta t}, t+\delta t) = \Theta(\boldsymbol{X}_t, t), \tag{2.10}$$

with δt an infinitely small time variation.

The evolution law of such random tracer can be described by the following stochastic partial differential equation (SPDE), namely

$$\mathbb{D}_t\Theta \triangleq \mathrm{d}_t\Theta + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\Theta - \frac{1}{2}\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}t = 0, \tag{2.11}$$

where $\mathrm{d}_t\Theta(\boldsymbol{x}) \triangleq \Theta(\boldsymbol{x}, t+\delta t) - \Theta(\boldsymbol{x}, t)$ stands for the (forward) local increment in time of the tracer Θ at a fixed point $\boldsymbol{x} \in \Omega$. Note that this differs from the usual notation "$\mathrm{d}\Theta_t$" used in stochastic (ordinary) differential equations varying only through a single variable (usually time).

The SPDE (2.11) encompasses physically meaningful terms. For instance, the third RHS term is a random forcing related to the tracer's advection by the unresolved small-scale flow. This term continuously backscatters random energy to the system (through its quadratic variation). Let us outline that this resulting advection noise, $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\Theta$, is multiplicative (as it depends on the tracer Θ) and non-Gaussian in general.

The last term in (2.11) will be very useful in studying turbulence at large scales. As defined in (2.8), $\boldsymbol{a}$ is the variance of the noise, hence it is a symmetric non-negative definite tensor. Therefore, under suitable boundary conditions, the following term describes the tracer's dissipation:

$$\int_\Omega \Theta\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}\boldsymbol{x} = -\int_\Omega (\boldsymbol{\nabla}\Theta)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\,\mathrm{d}\boldsymbol{x} \leq 0. \tag{2.12}$$

This depicts the mixing mechanism due to the action of the unresolved scales. As shown in Mémin (2014), under a spatially heterogeneous and temporally non-stationary random field in general, the last term in (2.11) plays a role similar to the functional eddy viscosity as introduced in many large-scale circulation models (Smagorinsky, 1963; Redi, 1982). In particular, for a homogeneous, isotropic and stationary random field, in which the variance tensor $\boldsymbol{a}$ becomes $a_0\mathbf{I}_d$, the diffusive term boils down immediately to a harmonic

friction term, $\frac{1}{2}a_0\nabla^2\Theta$, with a uniform coefficient a_0 to be specified.

As an additional feature of interest, there exists an "effective" advection velocity $\boldsymbol{v}^\star$ in (2.11) which is defined as

$$\boldsymbol{v}^\star \triangleq \boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,\boldsymbol{a} + \boldsymbol{\sigma}^T(\boldsymbol{\nabla\cdot}\boldsymbol{\sigma}). \tag{2.13}$$

This effective drift captures the action of inhomogeneity of the random field on the transported tracer and the possible small-scale velocity divergence. It is a *statistical eddy-induced velocity* of crucial importance, as shown in Chapter 3. Such a correction on the advection corresponds to the so-called *turbophoresis* phenomenon associated with small-scale inhomogeneity. This phenomenon drives inertial particles toward regions of lower turbulent diffusivity (Reeks, 1983). In our work, it is characterized by the turbophoresis term $\frac{1}{2}\,\boldsymbol{\nabla\cdot}\,\boldsymbol{a}$. As shown in Section 3.4, this term can be interpreted as a generalization of the Stokes drift, which occurs, for example, in the Langmuir circulation (Craik and Leibovich, 1976; Leibovich, 1980). It is also akin to the velocity correction introduced for tracer mean transport in oceanic or atmospheric circulation models (Andrews and McIntyre, 1978). Recently, this correction was observed to play a crucial role in the transition from the viscous layer regime to the logarithmic layer regime in wall bounded turbulent flows (Pinier et al., 2019).

To sum up, a new stochastic transport operator $\mathbb{D}_t$ involving all the above terms has been introduced in Mémin (2014). Some useful properties of such operator (such as the product rule) have been well described in Resseguier et al. (2017a). In particular, under incompressible noise, *i.e.* $\boldsymbol{\nabla\cdot}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0$, the stochastic transport operator per unit of time, $\mathbb{D}_t/\mathrm{d}t$ (defined in a distribution sense), is shown to be coincident with the stochastic material derivative.

In the following, we derive formally this stochastic transport operator. In order to develop the total variation of the tracer Θ composed with the stochastic trajectory $\boldsymbol{X}$, an adequate chain rule needs to be specified. As discussed in Section 1.1.3, if Θ is a smooth deterministic function, its total differentiation is driven by the Itô formula (1.1.3). However, in our case the random scalar Θ itself is a (continuous) semimartingale. This requires, therefore, to compute the differentiation of the composition of two stochastic processes, hence the generalized Itô formula (1.1.4) is adopted. Applying the differential

form of Equation (1.18) for the random tracer Θ, we have

$$\begin{aligned}\mathrm{d}\,\Theta(\boldsymbol{X}_t,t) = \mathrm{d}_t\Theta(\boldsymbol{X}_t,t) + \sum_{i=1}^{d}\frac{\partial\Theta}{\partial x_i}(\boldsymbol{X}_t,t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2\Theta}{\partial x_i\partial x_j}(\boldsymbol{X}_t,t)\,\mathrm{d}\Big\langle X^i,X^j\Big\rangle_t \\ + \sum_{i=1}^{d}\mathrm{d}\Big\langle\frac{\partial\Theta}{\partial x_i}(\boldsymbol{X},\cdot),X^i\Big\rangle_t. \qquad (2.14)\end{aligned}$$

We remark that $\mathrm{d}\,\Theta(\boldsymbol{X}_t,t) \triangleq \Theta(\boldsymbol{X}_{t+\delta t},t+\delta t) - \Theta(\boldsymbol{X}_t,t)$ denotes the total differentiation of Θ, whereas $\mathrm{d}_t\Theta(\boldsymbol{X}_t,t) \triangleq \Theta(\boldsymbol{X}_t,t+\delta t) - \Theta(\boldsymbol{X}_t,t)$ stands only for its time-increment as explained before.

The first quadratic variation on the RHS of Equation (2.14) can be immediately determined from decomposition (2.1), as follows

$$\begin{aligned}\Big\langle X^i,X^j\Big\rangle_t &= \Big\langle\int_0^{\cdot}\Big(\sigma(\boldsymbol{X}_s,s)\,\mathrm{d}B_s\Big)^i,\int_0^{\cdot}\Big(\sigma(\boldsymbol{X}_s,s)\,\mathrm{d}B_s\Big)^j\Big\rangle_t \\ &= \Big\langle\int_0^{\cdot}\int_\Omega\sum_{k=1}^{d}\breve{\sigma}_{ik}(\boldsymbol{X}_s,\boldsymbol{y},s)\,\mathrm{d}B_s^k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},\int_0^{\cdot}\int_\Omega\sum_{l=1}^{d}\breve{\sigma}_{jl}(\boldsymbol{X}_s,\boldsymbol{z},s)\,\mathrm{d}B_s^l(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}\Big\rangle_t \\ &= \int_0^t\int_\Omega\int_\Omega\sum_{k,l=1}^{d}\breve{\sigma}_{ik}(\boldsymbol{X}_s,\boldsymbol{y},s)\breve{\sigma}_{jl}(\boldsymbol{X}_s,\boldsymbol{z},s)\,\mathrm{d}\underbrace{\Big\langle B^k(\boldsymbol{y}),B^l(\boldsymbol{z})\Big\rangle_s}_{\delta_{kl}\,\delta(\boldsymbol{y}-\boldsymbol{z})\,s}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\ &= \int_0^t\int_\Omega\sum_{k=1}^{d}\breve{\sigma}_{ik}(\boldsymbol{X}_s,\boldsymbol{y},s)\breve{\sigma}_{jk}(\boldsymbol{X}_s,\boldsymbol{y},s)\,\mathrm{d}\boldsymbol{y}\mathrm{d}s \\ &= \int_0^t a_{ij}(\boldsymbol{X}_s,s)\,\mathrm{d}s, \qquad \forall\, i,j=1,\dots,d. \qquad (2.15)\end{aligned}$$

More precisely, the first equality results from the quadratic-covariation-rule (1.9); the second one comes from the definition of the noise (2.2); the third one is based on the associativity of Itô integrals (1.12a), and the last one derives directly from the definition of the variance (2.8).

The last quadratic covariation on the RHS of Equation (2.14) is an additional term compared to the classical Itô formula (1.1.3). Here, it describes the interaction between the stochastic flow and the tracer's gradient. To evaluate this term, let us consider that the semimartingale Θ itself can be decomposed into (written in differential form):

$$\mathrm{d}_t\Theta(\boldsymbol{X}_t,t) = f(\boldsymbol{X}_t,t)\,\mathrm{d}t + \int_\Omega\sum_{k=1}^{d}g_k(\boldsymbol{X}_t,\boldsymbol{y},t)\,\mathrm{d}B_t^k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \qquad (2.16)$$

where f and $g_k (k = 1, \dots, d)$ are assumed to be (locally) bounded. As a result, the quadratic covariation between the stochastic flow and the tracer's gradient can be specified in a similar way as in (2.15), namely

$$\Big\langle \frac{\partial \Theta}{\partial x_i}, X^i \Big\rangle_t = \int_0^t \int_\Omega \sum_{j=1}^d \breve{\sigma}_{ij}(\boldsymbol{X}_t, \boldsymbol{y}, s) \frac{\partial g_j}{\partial x_i}(\boldsymbol{X}_t, \boldsymbol{y}, s)\, \mathrm{d}\boldsymbol{y}\mathrm{d}s, \quad \forall\, i = 1, \dots, d. \tag{2.17}$$

Substituting the expressions (2.15) and (2.17) in Equation (2.14), we obtain the total variation of the tracer Θ, that is

$$\begin{aligned} \mathrm{d}\,\Theta = \mathrm{d}_t\Theta + \Big(\sum_{i=1}^d v_i \frac{\partial \Theta}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^d a_{ij} \frac{\partial^2 \Theta}{\partial x_i \partial x_j} + \int_\Omega \sum_{i,j=1}^d \breve{\sigma}_{ij} \frac{\partial g_j}{\partial x_i}\, \mathrm{d}\boldsymbol{y} \Big)\, \mathrm{d}t \\ + \int_\Omega \sum_{i,j=1}^d \breve{\sigma}_{ij} \mathrm{d}B_t^j \frac{\partial \Theta}{\partial x_i}\, \mathrm{d}\boldsymbol{y}, \end{aligned} \tag{2.18}$$

in which the space-time variables $(\boldsymbol{X}_t, t)$ are dropped for the sake of simplicity.

In particular, as interpreted by Equation (2.10), the fact that the random tracer Θ is conserved along the stochastic trajectory leads to $\mathrm{d}\,\Theta(\boldsymbol{X}_t, t) = 0,\ \forall\, t \in \mathbb{R}^+$. Thus, one may specify explicitly the expressions of f and g_j by identifying Equations (2.18) and (2.16), namely

$$f = -\sum_{i=1}^d v_i \frac{\partial \Theta}{\partial x_i} - \frac{1}{2} \sum_{i,j=1}^d a_{ij} \frac{\partial^2 \Theta}{\partial x_i \partial x_j} - \int_\Omega \sum_{i,j=1}^d \breve{\sigma}_{ij} \frac{\partial g_j}{\partial x_i}\, \mathrm{d}\boldsymbol{y} \tag{2.19a}$$

$$g_j = -\sum_{k=1}^d \breve{\sigma}_{jk} \frac{\partial \Theta}{\partial x_k}. \tag{2.19b}$$

Indeed, these results are provided by the canonical decomposition (see Theorem 1.1.1) of the semimartingale Θ. Substituting the expression of g_j in Equation (2.18) for the last term, we have

$$\begin{aligned} \int_\Omega \breve{\sigma}_{ij} \frac{\partial g_j}{\partial x_i}\, \mathrm{d}\boldsymbol{y} &= -\Big(\int_\Omega \breve{\sigma}_{ij} \breve{\sigma}_{jk}\, \mathrm{d}\boldsymbol{y} \Big) \frac{\partial^2 \Theta}{\partial x_i \partial x_k} - \Big(\int_\Omega \breve{\sigma}_{ij} \frac{\partial \breve{\sigma}_{jk}}{\partial x_i}\, \mathrm{d}\boldsymbol{y} \Big) \frac{\partial \Theta}{\partial x_k} \\ &= -a_{ik} \frac{\partial^2 \Theta}{\partial x_i \partial x_k} - \Big(\underbrace{\int_\Omega \frac{\partial}{\partial x_i} (\breve{\sigma}_{ij} \breve{\sigma}_{jk})\, \mathrm{d}\boldsymbol{y}}_{\partial a_{ik}/\partial x_i} \Big) \frac{\partial \Theta}{\partial x_k} + \Big(\underbrace{\int_\Omega \breve{\sigma}_{jk} \frac{\partial \breve{\sigma}_{ij}}{\partial x_i}\, \mathrm{d}\boldsymbol{y}}_{\triangleq \sigma_{jk}(\partial \sigma_{ij}/\partial x_i)} \Big) \frac{\partial \Theta}{\partial x_k} \\ &= -\frac{1}{2} a_{ik} \frac{\partial^2 \Theta}{\partial x_i \partial x_k} - \frac{1}{2} \frac{\partial}{\partial x_i} \Big(a_{ik} \frac{\partial \Theta}{\partial x_k} \Big) - \frac{1}{2} \frac{\partial a_{ik}}{\partial x_i} \frac{\partial \Theta}{\partial x_k} + \sigma_{jk} \frac{\sigma_{ij}}{\partial x_i} \frac{\partial \Theta}{\partial x_k}, \end{aligned} \tag{2.20}$$

in which the summation are again dropped for the sake of simplicity. Sequentially substituting (2.20) in (2.18), the transport equation of Θ reduces to

$$0 = \mathrm{d}_t\Theta + \sum_{k=1}^{d}\Big(v_k - \frac{1}{2}\sum_{i=1}^{d}\frac{\partial a_{ik}}{\partial x_i} + \sum_{i,j=1}^{d}\sigma_{jk}\frac{\sigma_{ij}}{\partial x_i}\Big)\frac{\partial\Theta}{\partial x_k}\,\mathrm{d}t + \sum_{k=1}^{d}(\sigma\mathrm{d}B_t)^k\frac{\partial\Theta}{\partial x_k} - \frac{1}{2}\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\Big(\sum_{k=1}^{d}a_{ik}\frac{\partial\Theta}{\partial x_k}\Big)\,\mathrm{d}t \tag{2.21}$$

or, equivalently, in a vector form to

$$\mathrm{d}_t\Theta + \Big(\Big(\boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,\boldsymbol{a} + \boldsymbol{\sigma}^T(\boldsymbol{\nabla\cdot}\boldsymbol{\sigma})\Big)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\Big)\,\boldsymbol{\cdot\nabla}\,\Theta - \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,(\boldsymbol{a}\boldsymbol{\nabla}\Theta) = 0. \tag{2.22}$$

2.3 Energy conservation

This section reviews one remarkable property of the previously derived stochastic transport operator. That is, under an isochoric stochastic flow, it preserves along time the global energy of the random tracer Θ (for any realizations):

$$\mathrm{d}_t\int_\Omega\frac{1}{2}\Theta^2\,\mathrm{d}\boldsymbol{x} = 0. \tag{2.23}$$

In order to compute the evolution of the tracer's energy in the stochastic framework, the Itô's integration-by-part formula (1.1.2) is used:

$$\mathrm{d}_t\int_\Omega\frac{1}{2}\Theta^2\,\mathrm{d}\boldsymbol{x} = \int_\Omega\Big(\Theta\mathrm{d}_t\Theta + \frac{1}{2}\mathrm{d}\langle\Theta,\Theta\rangle_t\Big)\,\mathrm{d}\boldsymbol{x}, \tag{2.24}$$

where the tracer's quadratic variation can be easily evaluated from (2.22), in the same way as in (2.15), and reads

$$\begin{aligned}\langle\Theta,\Theta\rangle_t &= \Big\langle\int_0^{\cdot}(\boldsymbol{\nabla}\Theta)^T\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s, \int_0^{\cdot}(\boldsymbol{\nabla}\Theta)^T\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s\Big\rangle_t \\ &= \int_0^t\sum_{i,j=1}^{d}\frac{\partial\Theta}{\partial x_i}\Big(\underbrace{\int_\Omega\sum_{i,j=1}^{d}\breve{\sigma}_{ik}\breve{\sigma}_{jk}\,\mathrm{d}\boldsymbol{y}}_{a_{ij}}\Big)\frac{\partial\Theta}{\partial x_j}\,\mathrm{d}s = \int_0^t(\boldsymbol{\nabla}\Theta)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\,\mathrm{d}s.\end{aligned} \tag{2.25}$$

In fact, this quadratic variation can be interpreted as the increase of tracer energy due to the advection by the noise.

At the same time, to ensure an isochoric stochastic flow, incompressibility constraints

on the effective drift $\boldsymbol{\nabla}\cdot\boldsymbol{v}^\star = 0$ and on the small-scale velocity $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0$ are required. These two constraints are detailed in Section 3.1.1. Here, they help us to establish the strong energy conservation property, that is

$$\mathrm{d}_t \int_\Omega \frac{1}{2}\Theta^2 \,\mathrm{d}\boldsymbol{x} = -\int_\Omega \boldsymbol{\nabla}\cdot\Big(\frac{1}{2}\Theta^2(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\Big)\,\mathrm{d}\boldsymbol{x} + \Big(\underbrace{\int_\Omega \frac{1}{2}\Theta\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}\boldsymbol{x}}_{\text{Energy loss by diffusion}} + \underbrace{\int_\Omega \frac{1}{2}(\boldsymbol{\nabla}\Theta)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\,\mathrm{d}\boldsymbol{x}}_{\text{Energy intake by the noise}}\Big)\,\mathrm{d}t = 0, \tag{2.26}$$

in which the divergence theorem and integration-by-parts (over $\boldsymbol{x} \in \Omega$) have been used under ideal boundary conditions (periodic, null normal velocity, *etc.*). Indeed, Equation (2.26) can be interpreted as a process where the energy brought by the noise is exactly counter-balanced by that dissipated by the diffusion term. This result conserves the main specificity of a transport equation. It explains why we refer to $\mathbb{D}_t$ as a stochastic transport operator.

Moreover, the last equation (2.26) shows energy conservation for each realization of the random tracer. In particular, the ensemble mean of tracer's random energy is also conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Big[\int_\Omega \frac{1}{2}\Theta^2\,\mathrm{d}\boldsymbol{x}\Big] = 0. \tag{2.27}$$

Note that the derivative $\mathrm{d}/\mathrm{d}t$ make senses here, since it is applied to the expectation of a random process which is deterministic. From the variance's definition, Equation (2.27) implies that the loss of energy of the ensemble mean always balances the ensemble variance:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \frac{1}{2}\mathrm{Var}(\Theta)\,\mathrm{d}\boldsymbol{x} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \frac{1}{2}\mathbb{E}[\Theta]^2\,\mathrm{d}\boldsymbol{x}. \tag{2.28}$$

As illustrated in Chapter 6, this process is quite useful for uncertainty quantification (UQ). Since the tracer is continuously randomized while the tracer interacts with the unresolved scales.

2.4 Stratonovich representation

In this section, we give an equivalent expression of the stochastic flow (2.1) and of the stochastic transport operator (2.22) in Stratonovich form. To this end, the general Stratonovich-Itô-integral conversion rule (1.20) is first adopted. Applying this formula for

each component $(i = 1, \dots, d)$ of the noise yields

$$(\sigma(\boldsymbol{X}_t, t) \circ \mathrm{d}B_t)^i = (\sigma(\boldsymbol{X}_t, t)\, \mathrm{d}B_t)^i + \underbrace{\frac{1}{2} \int_\Omega \sum_{i=1}^{d} \mathrm{d}\langle \breve{\sigma}_{ij}(\boldsymbol{X}, \boldsymbol{y}, \cdot), B^j(\boldsymbol{y}) \rangle_t \, \mathrm{d}\boldsymbol{y}}_{\mathrm{I}}. \tag{2.29}$$

Similarly to the previous calculation, to evaluate the above quadratic covariation, I, only the martingale component M_t of the semimartingale $\breve{\sigma}_{ij}(\boldsymbol{X}, \boldsymbol{y}, t)$ is required. This can be determined by the Itô formula (1.1.3), namely

$$M_t = \int_0^t \sum_{k=1}^{d} \frac{\partial \breve{\sigma}_{ij}}{\partial x_k}(\boldsymbol{X}_s, \boldsymbol{y}, s) \Big(\underbrace{\int_\Omega \sum_{l=1}^{d} \breve{\sigma}_{kl}(\boldsymbol{X}_s, \boldsymbol{z}, s)\, \mathrm{d}B_s^l(\boldsymbol{z})\, \mathrm{d}\boldsymbol{z}}_{(\sigma(\boldsymbol{X}_s, s)\, \mathrm{d}B_s)^k} \Big). \tag{2.30}$$

Substituting (2.30) for the quadratic covariation I in (2.29), we obtain

$$\begin{aligned} \mathrm{I} &= \frac{1}{2} \int_\Omega \int_\Omega \sum_{j,k,l=1}^{d} \frac{\partial \breve{\sigma}_{ij}}{\partial x_k} \breve{\sigma}_{kl}\, \mathrm{d} \underbrace{\big\langle B^l(\boldsymbol{z}), B^j(\boldsymbol{y}) \big\rangle_t}_{\delta_{jl}\, \delta(\boldsymbol{y}-\boldsymbol{z})\, t} \mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\ &= \frac{1}{2} \Big(\sum_{k=1}^{d} \frac{\partial}{\partial x_k} \Big(\underbrace{\int_\Omega \sum_{j=1}^{d} \breve{\sigma}_{ij} \breve{\sigma}_{kj}\, \mathrm{d}\boldsymbol{y}}_{a_{ik}} \Big) - \sum_{j=1}^{d} \int_\Omega \breve{\sigma}_{ij} \Big(\sum_{k=1}^{d} \frac{\partial \breve{\sigma}_{kj}}{\partial x_k} \Big) \mathrm{d}\boldsymbol{y} \Big) \mathrm{d}t. \end{aligned} \tag{2.31}$$

Subsequently, substituting (2.31) in (2.29), then in (2.1), we deduce an equivalent Stratonovich representation of the stochastic flow that reads in vector form as

$$\mathrm{d}\boldsymbol{X}_t = \Big(\underbrace{\boldsymbol{v} - \frac{1}{2}\, \boldsymbol{\nabla} \boldsymbol{\cdot}\, \boldsymbol{a} + \frac{1}{2} \boldsymbol{\sigma}^T (\boldsymbol{\nabla} \boldsymbol{\cdot} \boldsymbol{\sigma})}_{\boldsymbol{v}^\star - \frac{1}{2} \boldsymbol{\sigma}^T (\boldsymbol{\nabla} \cdot \boldsymbol{\sigma})} \Big) \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t. \tag{2.32}$$

Afterwards, for a random tracer Θ transported by a Stratonovich flow, one can apply the generalized Itô's formula (1.1.8). Thus, substituting (2.32) in (1.23), we deduce a Stratonovich representation of the stochastic transport operator, that is

$$\mathbb{D}_t \circ \Theta \triangleq \mathrm{d}_t \circ \Theta + \Big(\Big(\boldsymbol{v}^\star - \frac{1}{2} \boldsymbol{\sigma}^T (\boldsymbol{\nabla} \boldsymbol{\cdot} \boldsymbol{\sigma}) \Big)\, \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t \Big) \boldsymbol{\cdot} \boldsymbol{\nabla}\, \Theta = 0, \tag{2.33}$$

where $\mathrm{d}_t \circ \Theta(\boldsymbol{X}_t, t) \triangleq \Theta(\boldsymbol{X}_t, t + \delta t/2) - \Theta(\boldsymbol{X}_t, t - \delta t/2)$ stands now for the central time-increment. In particular, for incompressible small-scale flows, the Stratonovich transport

operator reads

$$\mathbb{D}_t{\circ}\Theta \triangleq \mathrm{d}_t{\circ}\Theta + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}{\circ}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\Theta = 0. \tag{2.34}$$

Let us emphasize again that Equation (2.34) is equivalent to Equation (2.22) yet only under a different representation of the stochastic integrals, in which the Itô advection noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\Theta$ is a martingale of null ensemble mean, whereas the Stratonovich advection noise $\boldsymbol{\sigma}{\circ}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\Theta$ is not.

2.5 Conservation of tracer's moments

This section shows that if a random tracer Θ is transported by the stochastic flow, as given in (2.11), then the local moments Θ^p (for $p\in\{1,2\}$ or $p\in\mathbb{R}^+$, $p\geq 3$) of the tracer are equally transported:

$$\mathbb{D}_t\Big(\frac{1}{p}\Theta^p\Big) = 0. \tag{2.35}$$

In addition, one can consider that, under some incompressible constraints and ideal boundary conditions for both large and small scales, the global moments of tracer are also conserved:

$$\mathrm{d}_t\int_\Omega \frac{1}{p}\Theta^p\,\mathrm{d}\boldsymbol{x} = 0. \tag{2.36}$$

For the sake of simplicity, we first adopt the Stratonovich notation. Since the function $x\mapsto x^p$ is itself deterministic and of class C^3 for $p\in\{1,2\}$ or $p\in\mathbb{R}^+$, $p\geq 3$, the classical Itô's formula (1.1.7) can be applied:

$$\mathrm{d}_t{\circ}\Big(\frac{1}{p}\Theta^p\Big) = \Theta^{p-1}\,\mathrm{d}_t{\circ}\Theta. \tag{2.37}$$

From the Stratonovich transport equation (2.33) of the random tracer Θ, we deduce

$$\begin{aligned}\mathrm{d}_t{\circ}\Big(\frac{1}{p}\Theta^p\Big) &= -\Theta^{p-1}\Big((\boldsymbol{v}^\star - \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}{\cdot}\boldsymbol{\sigma}))\,\mathrm{d}t + \boldsymbol{\sigma}{\circ}\mathrm{d}\boldsymbol{B}_t\Big)\cdot\boldsymbol{\nabla}\,\Theta\\ &= -\Big((\boldsymbol{v}^\star - \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}{\cdot}\boldsymbol{\sigma}))\,\mathrm{d}t + \boldsymbol{\sigma}{\circ}\mathrm{d}\boldsymbol{B}_t\Big)\cdot\boldsymbol{\nabla}\,\Big(\frac{1}{p}\Theta^p\Big),\end{aligned} \tag{2.38}$$

which leads to

$$\mathbb{D}_t{\circ}\Big(\frac{1}{p}\Theta^p\Big) = 0. \tag{2.39}$$

Therefore, the equivalent Itô transport (2.35) of the local moments can be recovered in the same way as shown in Section 2.4.

To demonstrate the conservation of the global moments (2.36), let us denote that $\widetilde{\Theta} \triangleq \frac{1}{p}\Theta^p$ and assume that $\boldsymbol{\nabla}\cdot\boldsymbol{v}^\star = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0$ everywhere in the domain Ω, together with $\boldsymbol{v}^\star\cdot\boldsymbol{n} = \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{n} = 0$ on the boundary $\partial\Omega$, where $\boldsymbol{n}$ stands for the outward pointing unit normal. Hereafter, the evolution of the global moments is given by

$$\begin{aligned}\mathrm{d}_t\int_\Omega \widetilde{\Theta}\,\mathrm{d}\boldsymbol{x} &= \int_\Omega \Big(\frac{1}{2}\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\widetilde{\Theta})\,\mathrm{d}t - (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\widetilde{\Theta}\Big)\,\mathrm{d}\boldsymbol{x}\\ &= \int_\Omega \boldsymbol{\nabla}\cdot\Big(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\widetilde{\Theta}\,\mathrm{d}t - (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\widetilde{\Theta}\Big)\,\mathrm{d}\boldsymbol{x}\\ &= \oint_{\partial\Omega}\boldsymbol{n}\cdot\Big(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\widetilde{\Theta}\,\mathrm{d}t - (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\widetilde{\Theta}\Big)\,\mathrm{d}\boldsymbol{x}',\end{aligned} \tag{2.40}$$

where the last equality comes from the divergence theorem and the latter terms are null according to the ideal boundary conditions. Moreover, the following argument provides us the conservation (2.36) of the global moments:

$$\begin{aligned}\boldsymbol{n}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\widetilde{\Theta})\,\mathrm{d}t &= \sum_{i,j=1}^{d} n_i a_{i,j}\frac{\partial\widetilde{\Theta}}{\partial x_j}\,\mathrm{d}t\\ &= \sum_{j=1}^{d}\mathbb{E}\Big[\underbrace{\sum_{i=1}^{d} n_i(\sigma\mathrm{d}B_t)^i}_{=\,0}(\sigma\mathrm{d}B_t)^j\Big]\frac{\partial\widetilde{\Theta}}{\partial x_j},\end{aligned} \tag{2.41}$$

where the last equality emerges from the definition (2.8) of the noise variance and the noise boundary condition.

We remark that the conservation of tracer's energy, as shown in Section 2.3, can be then considered as a particular case of (2.36) for $p = 2$. For future works, it would be interesting to verify if this conservation laws holds for any real number $p \in (0, 3)$, in which the classical Itô's formula could not be directly applied to have (2.37).

2.6 Stochastic Reynolds transport theorem

This section reviews the stochastic version of the Reynolds transport theorem (SRTT), introduced by Mémin (2014), to express time differentiation of integrals over arbitrarily moving and deforming volumes. In fact, this can be considered as an extension of the stochastic transport equation derived in Section 2.2.

Let us first remark that the stochastic transport equation (2.11) is derived for Lagrangian coordinates, *i.e.* following a fluid particle trajectory. It can be translated in Eulerian coordinates noticing that they are valid for any $\boldsymbol{X}_t \in \Omega$; then, fixing a point in

the fluid domain such that $\boldsymbol{x} = \boldsymbol{X}$, it holds:

$$\mathbb{D}_t\Theta(\boldsymbol{x},t) \triangleq \mathrm{d}_t\Theta + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\Theta - \frac{1}{2}\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}t = 0. \tag{2.42}$$

Hereafter, we adopt such Eulerian equation to derive the SRTT in weak formulation. Let $\mathcal{V}(t)$ be a control-volume at any given time $t \in \mathbb{R}^+$, and $\varphi(\cdot,t) \in C_c^\infty(\Omega)$ be a test function with compact support on $\mathcal{V}(t)$ *i.e.* φ vanishes outside (including the boundaries) of $\mathcal{V}(t)$. As such, the time-differentiation of the inner product between Θ and φ within the control-volume can be written as

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \Theta\varphi\,\mathrm{d}\boldsymbol{x} = \int_\Omega \Big(\varphi\,\mathrm{d}_t\Theta + \Theta\,\mathrm{d}_t\varphi + \mathrm{d}\langle\Theta,\varphi\rangle_t\Big)\,\mathrm{d}\boldsymbol{x}, \tag{2.43}$$

This equality is deduced from the Itô's integration-by-part formula (1.1.2). Note that both quantities Θ and φ can be assumed to be conserved, *i.e.* Equation (2.42) holds also for φ. This allows us to specify, in the same way as in (2.25), the last quadratic covariation term in (2.43), namely

$$\langle\Theta,\varphi\rangle_t = \int_0^t \sum_{k,l=1}^{d} \frac{\partial\Theta}{\partial x_k} a_{kl} \frac{\partial\varphi}{\partial x_l}\,\mathrm{d}s = \int_0^t (\boldsymbol{\nabla}\Theta)^T \boldsymbol{a}\boldsymbol{\nabla}\Theta\,\mathrm{d}s. \tag{2.44}$$

Subsequently, Equation (2.43) reduces to

$$\begin{aligned}
&\int_\Omega \left(\varphi\,\mathrm{d}_t\Theta - \Theta(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\varphi\right)\mathrm{d}\boldsymbol{x} + \mathrm{d}t \underbrace{\int_\Omega \left(\frac{1}{2}\Theta\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}\boldsymbol{\nabla}\varphi) + (\boldsymbol{\nabla}\Theta)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)\mathrm{d}\boldsymbol{x}}_{\mathrm{I}} \\
&= \int_\Omega \varphi\Big(\mathrm{d}_t\Theta + \boldsymbol{\nabla}\cdot\big(\Theta(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\big)\Big)\,\mathrm{d}\boldsymbol{x} + \mathrm{I}\,\mathrm{d}t,
\end{aligned} \tag{2.45}$$

where the last equality comes from the integration-by-part formula (over space) and the null boundary condition of the test function. Similarly, the second term I can be expanded as

$$\begin{aligned}
\mathrm{I} &= \int_\Omega \sum_{i,j=1}^{d} \left(\frac{1}{2}\Theta\frac{\partial a_{ij}}{\partial x_i}\frac{\partial\varphi}{\partial x_j} + \frac{1}{2}\Theta a_{ij}\frac{\partial^2\varphi}{\partial x_i\partial x_j} + \frac{\partial\Theta}{\partial x_i}a_{ij}\frac{\partial\varphi}{\partial x_j}\right)\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega \sum_{i,j=1}^{d} \varphi\Big(-\frac{1}{2}\frac{\partial}{\partial x_j}\Big(\Theta\frac{\partial a_{ij}}{\partial x_i}\Big) + \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_j}(\Theta a_{ij}) - \frac{\partial}{\partial x_j}\Big(a_{ij}\frac{\partial\Theta}{\partial x_i}\Big)\Big)\,\mathrm{d}\boldsymbol{x} \\
&= -\int_\Omega \frac{\varphi}{2} \underbrace{\sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\Big(a_{ij}\frac{\partial\Theta}{\partial x_i}\Big)}_{\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)}\,\mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{2.46}$$

Substituting (2.46) in (2.45) then in (2.43), and combing with (2.42), we obtain

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \Theta\varphi \,\mathrm{d}\boldsymbol{x} = \int_{\Omega} \varphi\Big(\mathbb{D}_t\Theta + \Theta\, \boldsymbol{\nabla}\cdot\, (\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\Big)\, \mathrm{d}\boldsymbol{x}. \tag{2.47}$$

As this relation holds for any test function φ with compact support on $\mathcal{V}(t)$, the integral over Ω on the RHS reduces to an integral over $\mathcal{V}(t)$. Therefore, the rate of change of a random scalar quantity Θ transported by the random flow (2.1) within a control-volume can be written in Eulerian coordinates as follows:

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \Theta \,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{V}(t)} \left(\mathbb{D}_t\Theta + \Theta\, \boldsymbol{\nabla}\cdot\, (\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\right) \mathrm{d}\boldsymbol{x}. \tag{2.48}$$

Let us highlight that such SRTT is a fundamental tool to develop the conservation equations for fluid motions within the LU framework. This is shown later in Section 3.1.

As a supplement, an equivalent Stratonovich form of the SRTT (2.48) can be derived using the Stratonovich transport operator (2.33) and repeating the above procedure. The final result reads

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \Theta \,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{V}(t)} \Big(\mathbb{D}_t\circ\Theta + \Theta\, \boldsymbol{\nabla}\cdot\, ((\boldsymbol{v}^\star - \frac{1}{2}\boldsymbol{\sigma}^T\, \boldsymbol{\nabla}\cdot\, \boldsymbol{\sigma})\,\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\Big)\, \mathrm{d}\boldsymbol{x}. \tag{2.49}$$

2.7 General remarks

At the end of this chapter, we address several points, which in our opinion characterizes the LU framework:

(i) The transport equations derived under the LU framework involves an effective advection $\boldsymbol{v}^\star$ independent of the stochastic integral applied, *i.e.* in both Stratonovich and Itô setting (see 2.33 and 2.13 respectively).

(ii) The LU framework based on the decomposition (2.1) includes a centered noise by the definition of Itô integral. Therefore, the mean of the Lagrangian large-scale velocity component $\mathbb{E}[\boldsymbol{v}]$ corresponds to the mean of the Lagrangian fluid velocity $\mathbb{E}[\mathrm{d}\boldsymbol{X}_t/\mathrm{d}t]$ (defined in distribution sense).

(iii) The LU framework conserves the energy for every realization. As shown later for QG models, this allows us to build dynamics and draw realizations that have exactly the same energy conservation properties as their deterministic counterpart. This strong asset of the LU framework enables us to propose efficient schemes for numerical simulation, analysis and data assimilation of 3D turbulent flows (Chandramouli

et al., 2020; Cintolesi and Mémin, 2020; Kadri Harouna and Mémin, 2017; Resseguier et al., 2017d; Tissot et al., 2020).

(iv) The derivation of the stochastic models under the LU framework follows exactly the same path as the deterministic derivation. As shown in subsequent chapters, only the noise and its amplitude needs to be properly scaled. To oceanographers, this provides a very interesting and pratical tool for investigating the implication of the small-scales.

Finally, we summarize the main points of this chapter in the following shaded box.

Summary

- Noise

$$\begin{aligned}\boldsymbol{\sigma}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{B}_t &= \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{y},t)\,\mathrm{d}\boldsymbol{B}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\\ &= \sum_{n\in\mathbb{N}} \sqrt{\lambda_n}(t)\boldsymbol{\phi}_n(\boldsymbol{x},t)\,\mathrm{d}\beta_t^n\end{aligned}$$

- Variance

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{x},t) &= \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{y},t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{x},\boldsymbol{y},t)\,\mathrm{d}\boldsymbol{y}\\ &= \sum_{n\in\mathbb{N}} \lambda_n(t)\Big(\boldsymbol{\phi}_n\boldsymbol{\phi}_n^T\Big)(\boldsymbol{x},t)\end{aligned}$$

- Stochastic Reynolds transport theorem

$$\mathrm{d}_t\int_{\mathcal{V}(t)}\Theta\,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{V}(t)}\Big(\mathbb{D}_t\Theta + \Theta\,\boldsymbol{\nabla}\boldsymbol{\cdot}\,(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\Big)\,\mathrm{d}\boldsymbol{x}$$

- Stochastic transport operator

$$\mathbb{D}_t\Theta \triangleq \mathrm{d}_t\Theta + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\boldsymbol{\cdot}\boldsymbol{\nabla}\,\Theta - \frac{1}{2}\,\boldsymbol{\nabla}\boldsymbol{\cdot}\,(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}t$$

- Effective advection drift

$$\boldsymbol{v}^\star \triangleq \boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla}\boldsymbol{\cdot}\,\boldsymbol{a} + \boldsymbol{\sigma}^T(\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{\sigma})$$

Chapter 3

Three-dimensional stochastic models

Abstract

This chapters reviews some important three-dimensional equations under the location uncertainty framework. The core of these models is based on the stochastic Reynolds transport theorem, presented in the previous chapter. We first describe briefly the derivation of the stochastic governing equations of fluid motion, including the stochastic conservation of mass and of linear momentum. In order to do some comparisons with the classical large-eddy simulation framework, we present a pseudo-stochastic model based on a scale-separation assumption. Then a stochastic representation of the vorticity's evolution is derived using classical vector calculus. Latter, for large-scale atmospheric and oceanic circulations, two stochastic primitive systems are presented under the classical hydrostatic balance and the Boussinesq approximations. These equations will be used to derive several two-dimensional stochastic equations in the subsequent chapter. In addition, we build a connection between the proposed random model and the Craik-Leibovich system. In particular, we show that the effective advection due to the effect of statistical inhomogeneity of the small-scale flow on the large-scale current, can be considered as a generalization of the Stokes drift. Finally, a connection with the Gent-McWilliams parametrization is performed. The work presented in this chapter has been partly published up in Bauer et al. (2020a).

3.1 Stochastic Navier-Stokes equations

The derivation of governing equations for stochastic flows following a similar strategy as in the classical framework. In particular, the physical arguments used in the classical derivation are also adopted here, while the stochastic Reynolds transport theorem (SRTT) derived in Section 2.6 is applied for the mathematical description of the conservative laws.

3.1.1 Conservation of mass

To derive the stochastic mass equation, we apply the SRTT (2.48) to the random density ρ, under the conservation constraint $\mathrm{d}_t \int_{\mathcal{V}(t)} \rho \, \mathrm{d}\boldsymbol{x} = 0$. For an arbitrary control-volume $\mathcal{V}(t)$, it reads:

$$0 = \int_{\mathcal{V}(t)} \Big(\mathbb{D}_t \rho + \rho \, \boldsymbol{\nabla} \cdot \left(\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t \right) \Big) \, \mathrm{d}\boldsymbol{x}, \tag{3.1}$$

where $\boldsymbol{v}^\star = \boldsymbol{v} - \frac{1}{2} \boldsymbol{\nabla} \cdot \boldsymbol{a} + \boldsymbol{\sigma}^T (\boldsymbol{\nabla} \cdot \boldsymbol{\sigma})$. As the previous equation holds for any control volume $\mathcal{V}(t)$, one can infer that

$$\mathbb{D}_t \rho + \rho \, \boldsymbol{\nabla} \cdot \left(\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t \right) = 0. \tag{3.2}$$

In particular, for an incompressible (isochoric) flow, the density variation is quite small in both time and space, hence the influence of the stochastic transport $\mathbb{D}_t \rho$ on the mass balance (3.2) can be ignored. Thus, the stochastic continuity equation can be approximately written as

$$\boldsymbol{\nabla} \cdot \left(\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t \right) = 0. \tag{3.3}$$

Seperating subsequently the process of bounded variation and the martingale based on the canonical decomposition (1.1.1), the following continuity equation is recovered:

$$\boldsymbol{\nabla} \cdot \Big(\boldsymbol{v} - \frac{1}{2} \boldsymbol{\nabla} \cdot \boldsymbol{a} \Big) = 0, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t = 0. \tag{3.4}$$

The second condition is intuitive and enforces a divergence free random component, whereas the first constraint imposes a divergence-free condition on the effective advection. This latter constraint provides a relation between the smooth resolved velocity component and the divergence of the variance tensor. For homogeneous random fields (such as an isotropic turbulence) this equation boils down to a classical divergence-free condition on

the resolved velocity component (as the variance tensor is constant in that case). In some particular cases, the first condition can be substituted by the sufficient stronger condition $\boldsymbol{\nabla}\cdot\boldsymbol{v} = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a} = 0$.

We remark that Equation (3.4) ignores the role of the term $\mathbb{D}_t\rho$ in the mass balance (3.2). For the sake of simplicity, we assume that $\mathbb{D}_t\rho = 0$ in this work. Nevertheless, in the general case, this term itself can be restricted by thermodynamic laws. Furthermore, only an incompressible flow is adopted in the following, hence the effective advection drift reduces to

$$\boldsymbol{v}^\star = \boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla}\cdot\boldsymbol{a}. \tag{3.5}$$

3.1.2 Conservation of momentum

This section provides an informal derivation of the stochastic momentum equation for an isochoric flow. Extended derivations can be found in Mémin (2014); Mikulevicius and Rozovskii (2004); Resseguier (2017).

According to Newton's second law, the rate of change of the linear momentum of a fluid particle equals to the force acting on it. On the other hand, this work only focuses on the effects of the small scales on the evolution of the large-scale flow, whereas the transport of the small scales by the large-scales as well as by themself remains unknown. As such, the change of the large-scale momentum $\rho\boldsymbol{v}$ within an arbitrary control-volume $\mathcal{V}(t)$ can be written as

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{V}(t)} (\boldsymbol{F}\,\mathrm{d}t + \boldsymbol{G}\,\mathrm{d}\boldsymbol{B}_t)\,\mathrm{d}\boldsymbol{x}, \tag{3.6}$$

where $\boldsymbol{F}$ and $\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t/\mathrm{d}t$ (in a distribution sense) are respectively the time-smooth components and the highly irregular terms of the applied force. From the SRTT (2.48) and the incompressible constraints (3.4), the LHS of Equation (3.6) simplifies to

$$\mathrm{d}_t \int_{\mathcal{V}(t)} \rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{V}(t)} \mathbb{D}_t(\rho\boldsymbol{v})\,\mathrm{d}\boldsymbol{x}. \tag{3.7}$$

Subsequently, dropping the arbitrary control-volume, we obtain

$$\mathbb{D}_t(\rho\boldsymbol{v}) = \boldsymbol{F}\,\mathrm{d}t + \boldsymbol{G}\,\mathrm{d}\boldsymbol{B}_t. \tag{3.8}$$

In particular, if ρ is constant, the LHS reduces to $\rho\mathbb{D}_t\boldsymbol{v}$. In addition, the forces on the

RHS can be developed following the same physical argument as in the classical derivation. In a rotating frame, it reads :

$$\boldsymbol{F} = -\rho(\boldsymbol{f} \times \boldsymbol{v}) - \rho\boldsymbol{g} - \boldsymbol{\nabla} p + \mu\nabla^2\boldsymbol{v}, \tag{3.9a}$$

$$\boldsymbol{G}\,\mathrm{d}\boldsymbol{B}_t = -\rho(\boldsymbol{f} \times \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) - \boldsymbol{\nabla}\mathrm{d}p_t^\sigma + \mu\nabla^2\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t. \tag{3.9b}$$

These forces are due to the gravitation potential, pressure and molecular friction forces (with the dynamic viscosity μ). The pressure term is split into a continuous pressure p and a time-uncorrelated random part $\dot{p}_t^\sigma = \mathrm{d}p_t^\sigma/\mathrm{d}t$. This latter term defined in a distribution sense describes the pressure fluctuations due to the random velocity component. Note that the gravity force is continuous in time, whereas the friction force applies both on the smooth and random velocity components (with $\nabla^2 = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}$ the Laplacian operator). For a fixed observer in a rotating frame, the rate of change of the fluid velocity incorporates the centripetal acceleration and the Coriolis acceleration as additional terms. The centrifugal force is included within an effective gravity $\boldsymbol{g}$. The Coriolis term with parameter $\boldsymbol{f}$ applies both to the large-scale component of the velocity and to the random small-scale field.

As remarked above, only the large scale momentum $\rho\boldsymbol{v}$ is applied in Newton's law in (3.6). Once the small-scales $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\mathrm{d}t$ are included in addition, a weak distributional form must be considered. A more general case can be found in Mémin (2014). Let us then summarized in the following the stochastic Navier-Stokes equations for isochoric flow, involving conservation of linear momentum, conservation of mass and the continuity equation.

(Momentum equation)

$$\mathbb{D}_t(\rho\boldsymbol{v}) + \rho\boldsymbol{f} \times (\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \rho\boldsymbol{g}\,\mathrm{d}t - \boldsymbol{\nabla}(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma) + \mu\nabla^2(\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{3.10a}$$

(Mass conservation)

$$\mathbb{D}_t\rho = 0, \tag{3.10b}$$

(Continuity)

$$\boldsymbol{\nabla}\cdot\boldsymbol{v}^\star = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0. \tag{3.10c}$$

This system corresponds to a large-scale description of the flow in which the effect of the random component is explicitly taken into account. Contrary to traditional Reynolds decomposition techniques, this method does not rely on the time differentiability assumption of the velocity fluctuations. The use of stochastic calculus to characterize the random component introduces naturally additional terms in the momentum equation. As discussed in

Section 2.2, these terms inherently account for several interesting phenomena associated with fluid flows such as backscattering via the multiplicative noise, large-scale dissipation through the variance tensor, and turbophoresis effect with the effective advection. We remark that System (3.10) is incomplete to describe the ocean dynamics. In addition, equations of state such as potential temperature and salinity are required. A stochastic version of such equations are derived by Resseguier et al. (2017a), using the SRTT and the product rule of the stochastic transport operator. For future works, it would be interesting to derive the stochastic governing equations for more complex flows such as a compressible fluid supporting fast acoustic modes (Vallis, 2017). These models can be derived in the same way using the general version of the SRTT associated to compressible random component as presented in Chapter 2.

3.2 Derived stochastic models

In this section, we present two systems of equations derived from the stochastic Navier-Stokes system (3.10). First, we describe a hybrid model where the stochastic contribution on the governing equations is modeled by a partial differential equation (PDE) with random forcing. Then, the three-dimensional stochastic vorticity equation is developed from the stochastic momentum equation (3.10a) using classical vector calculus.

3.2.1 Pseudo-stochastic equations

This section reviews a LES-like model presented in Bauer et al. (2020a); Cintolesi and Mémin (2020); Resseguier et al. (2020a). In the stochastic momentum equation (3.10a), the multiplicative random term, responsible for energy backscattering, ensues from a scale separation principle between the random fluctuation and the large-scale component. In fact, if the large-scale component is assumed to be regular enough, formally a process of bounded variation, Equation (3.10a) can be safely split in terms of a bounded variation process (see Definition 1.4) and in terms of a martingale (see Definition 1.6). Then, the previous system simplifies to the following stochastic LES system with a random forcing:

(Momentum equations)

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}^\star \cdot \boldsymbol{\nabla}\, \boldsymbol{v} - \frac{1}{2}\, \boldsymbol{\nabla} \cdot \boldsymbol{\tau} + \boldsymbol{f} \times \boldsymbol{v} = \boldsymbol{g} - \frac{1}{\rho} \boldsymbol{\nabla} p + \nu \nabla^2 \boldsymbol{v}, \tag{3.11a}$$

(Effective advection drift)

$$\boldsymbol{v}^\star = \boldsymbol{v} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}, \tag{3.11b}$$

(Subgrid diffusion tensor)

$$\boldsymbol{\tau} = \boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{v}, \tag{3.11c}$$

(Random contributions)

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{v} + \boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}p_t^\sigma + \nu\nabla^2\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \tag{3.11d}$$

(Mass equation)

$$\mathbb{D}_t\rho = 0, \tag{3.11e}$$

where $\nu = \mu/\rho$ is the kinematic viscosity. Under such framework, the momentum equation (3.11a) is a classical PDE with a random forcing through the density, whereas the mass conservation (3.11e) remains a SPDE. For the particular case of constant density, together with the incompressibility constraints (3.10c), the momentum equation is purely deterministic without any random forcing. Nevertheless, the system incorporates a stochastic balance equation on the pressure contribution $\mathrm{d}p_t^\sigma$ associated to the unresolved random terms. Also, this system includes the effective advection and the large-scale diffusion brought by the random component, but loses in the momentum equation the multiplicative random term responsible for energy backscattering. The shape of the dissipation operator is a priori known. Thus, there is no need to invoke the Boussinesq assumption to model the Reynolds stress tensor. Note that the variance tensor (and the noise, if required) can be constructed by drawing inspiration from known LES sub-grid scale models such as the Smagorinsky model (Chandramouli et al., 2018; Cintolesi and Mémin, 2020; Kadri Harouna and Mémin, 2017). Such pseudo-stochastic model has been also successively used to define efficient reduced order models (Resseguier et al., 2017d), in which the subgrid dissipation is directly defined from the neglected modes. This representation has the advantage to provide also new diagnosis enabling to quantify local energy dissipation as well as the effect of the turbulence inhomogeneity on the large-scale flow (Pinier et al., 2019). This ability is a strong asset of the LU formalism for reduced order modeling.

3.2.2 Stochastic vorticity equation

In this section, we derive the evolution of the three-dimensional vorticity $\omega = \boldsymbol{\nabla}\times\boldsymbol{v}$ for the large-scale drift $\boldsymbol{v}$ driven by the stochastic momentum equation (3.10a). For the

sake of simplicity, we drop here the Coriolis terms. In addition, we assume that $\rho = 1$ and $\boldsymbol{g} = g\boldsymbol{k}$ with g the gravity acceleration and $\boldsymbol{k} = (0, 0, 1)^T$ the vertical unit vector.

Let us first denote $\mathrm{d}\boldsymbol{V}_t \triangleq \boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and $\mathrm{d}P_t \triangleq p\,\mathrm{d}t + \mathrm{d}p_t^\sigma$. Thus, Equation (3.10a) can be re-written as

$$\mathrm{d}_t\boldsymbol{v} + \mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{v})\,\mathrm{d}t = g\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}\mathrm{d}P_t + \nu\nabla^2(\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t). \tag{3.12}$$

Taking the curl ($\boldsymbol{\nabla}\times$) of this equation, we obtain

$$\mathrm{d}_t\boldsymbol{\omega} + \underbrace{\boldsymbol{\nabla}\times(\mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{v})}_{\mathrm{I}} - \frac{1}{2}\underbrace{\boldsymbol{\nabla}\times\Big(\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{v})\Big)}_{\mathrm{II}}\,\mathrm{d}t = \nu\nabla^2(\boldsymbol{\omega}\,\mathrm{d}t + \boldsymbol{\omega}^\sigma\mathrm{d}\boldsymbol{B}_t), \tag{3.13}$$

where $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\mathrm{d}P_t = 0$ and $\boldsymbol{\omega}^\sigma \triangleq \boldsymbol{\nabla}\times\boldsymbol{\sigma}$. In order to expand the term I, let us recall the following vector calculus identity (for two vectors $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^3$):

$$\boldsymbol{\nabla}(\boldsymbol{f}\cdot\boldsymbol{g}) = (\boldsymbol{f}\cdot\boldsymbol{\nabla})\boldsymbol{g} + (\boldsymbol{g}\cdot\boldsymbol{\nabla})\boldsymbol{f} + \boldsymbol{f}\times(\boldsymbol{\nabla}\times\boldsymbol{g}) + \boldsymbol{g}\times(\boldsymbol{\nabla}\times\boldsymbol{f}). \tag{3.14}$$

Thus, we have

$$\mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{v} = \boldsymbol{\nabla}(\mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{v}) - \boldsymbol{v}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t - \mathrm{d}\boldsymbol{V}_t\times\boldsymbol{\omega} - \boldsymbol{v}\times\mathrm{d}\boldsymbol{\Omega}_t, \tag{3.15}$$

where $\mathrm{d}\boldsymbol{\Omega}_t \triangleq \boldsymbol{\nabla}\times\mathrm{d}\boldsymbol{V}_t$. then applying another vector calculus identity

$$\boldsymbol{\nabla}\times(\boldsymbol{f}\times\boldsymbol{g}) = \boldsymbol{f}(\boldsymbol{\nabla}\cdot\boldsymbol{g}) - \boldsymbol{g}(\boldsymbol{\nabla}\cdot\boldsymbol{f}) + (\boldsymbol{g}\cdot\boldsymbol{\nabla})\boldsymbol{f} + (\boldsymbol{f}\cdot\boldsymbol{\nabla})\boldsymbol{g}, \tag{3.16}$$

the term I in (3.13) reduces to

$$\mathrm{I} = \mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{\omega} - \boldsymbol{\omega}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t - \boldsymbol{\nabla}\times(\boldsymbol{v}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t) - \boldsymbol{\nabla}\times(\boldsymbol{v}\times\mathrm{d}\boldsymbol{\Omega}_t), \tag{3.17}$$

in which $\boldsymbol{\nabla}\cdot\boldsymbol{\omega} = 0$, $\boldsymbol{\nabla}\cdot\mathrm{d}\boldsymbol{V}_t = 0$ have been used. Let us now expand the last term using the Levi-Civita symbol, it reads:

$$\begin{aligned}\boldsymbol{\nabla}\times(\boldsymbol{v}\times\mathrm{d}\boldsymbol{\Omega}_t) &= \sum_{j,k=1}^{3}\varepsilon_{\cdot jk}\frac{\partial}{\partial x_j}\Big(\sum_{l,m=1}^{3}\varepsilon_{klm}v^l\sum_{n,p=1}^{3}\varepsilon_{mnp}\frac{\partial\mathrm{d}\boldsymbol{V}_t^p}{\partial x_n}\Big)\\ &= \sum_{j,k=1}^{3}\varepsilon_{\cdot jk}\sum_{l,n,p=1}^{3}\Big(\underbrace{\sum_{m=1}^{3}\varepsilon_{mkl}\varepsilon_{mnp}}_{\delta_{kn}\,\delta_{lp}\,-\,\delta_{kp}\,\delta_{ln}}\Big)\Big(\frac{\partial v^l}{\partial x_j}\frac{\partial\mathrm{d}\boldsymbol{V}_t^p}{\partial x_n} + v^l\frac{\partial^2\mathrm{d}\boldsymbol{V}_t^p}{\partial x_j x_n}\Big),\end{aligned} \tag{3.18}$$

or,

$$\begin{aligned}\boldsymbol{\nabla}\times(\boldsymbol{v}\times\mathrm{d}\boldsymbol{\Omega}_t) &= \sum_{j,k=1}^{3}\varepsilon_{\cdot jk}\sum_{l=1}^{3}\Big(\frac{\partial v^l}{\partial x_j}\frac{\partial \mathrm{d}\boldsymbol{V}_t^l}{\partial x_k}+v^l\frac{\partial^2\mathrm{d}\boldsymbol{V}_t^l}{\partial x_j\partial x_k}-\frac{\partial}{\partial x_j}(v^l\frac{\partial\mathrm{d}\boldsymbol{V}_t^k}{\partial x_l})\Big)\\ &= \sum_{l=1}^{3}(\boldsymbol{\nabla}v^l\times\boldsymbol{\nabla}\mathrm{d}\boldsymbol{V}_t^l+v^l\underbrace{\boldsymbol{\nabla}\times\boldsymbol{\nabla}\mathrm{d}\boldsymbol{V}_t^l}_{=0})-\boldsymbol{\nabla}\times(\boldsymbol{v}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t). \qquad (3.19)\end{aligned}$$

Substituting Equation (3.19) in Equation (3.17), we have

$$\mathrm{I}=\mathrm{d}\boldsymbol{V}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{\omega}-\boldsymbol{\omega}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t-\sum_{l=1}^{3}\boldsymbol{\nabla}v^l\times\boldsymbol{\nabla}\mathrm{d}\boldsymbol{V}_t^l. \qquad (3.20)$$

Let us now expand the term II in (3.13) by

$$\begin{aligned}\mathrm{II} &= \sum_{j,k=1}^{3}\varepsilon_{\cdot jk}\frac{\partial}{\partial x_j}\Big(\sum_{m=1}^{3}\frac{\partial}{\partial x_m}(\sum_{n=1}^{3}a_{mn}\frac{\partial v^k}{\partial x_n})\Big)\\ &= \sum_{m=1}^{3}\Big(\frac{\partial}{\partial x_m}\sum_{n=1}^{3}a_{mn}\frac{\partial}{\partial x_n}(\underbrace{\sum_{j,k=1}^{3}\varepsilon_{\cdot jk}\frac{\partial v^k}{\partial x_j}}_{\boldsymbol{\nabla}\times\boldsymbol{v}=\boldsymbol{\omega}})\Big)+\sum_{j,k,m=1}^{3}\frac{\partial}{\partial x_m}\Big(\varepsilon_{\cdot jk}\sum_{n=1}^{3}\frac{\partial a_{mn}}{\partial x_j}\frac{\partial v^k}{\partial x_n}\Big)\\ &= \boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{\omega})+\sum_{j,k=1}^{3}\boldsymbol{\nabla}\cdot(\varepsilon_{\cdot jk}\frac{\partial\boldsymbol{a}}{\partial x_j}\boldsymbol{\nabla}v^k). \qquad (3.21)\end{aligned}$$

Substituting subsequently the expressions of I (3.20) and II (3.21) in Equation (3.13), the stochastic vorticity equation can be written as

$$\begin{aligned}\mathbb{D}_t\boldsymbol{\omega} = \boldsymbol{\omega}\cdot\boldsymbol{\nabla}\,\mathrm{d}\boldsymbol{V}_t-\sum_{k=1}^{3}\boldsymbol{\nabla}\mathrm{d}\boldsymbol{V}_t^k\times\boldsymbol{\nabla}v^k+\frac{1}{2}\sum_{j,k=1}^{3}\boldsymbol{\nabla}\cdot\Big(\varepsilon_{\cdot jk}\frac{\partial\boldsymbol{a}}{\partial x_j}\boldsymbol{\nabla}v^k\Big)\\ +\nu\nabla^2(\boldsymbol{\omega}\,\mathrm{d}t+\boldsymbol{\omega}^\sigma\mathrm{d}\boldsymbol{B}_t). \qquad (3.22)\end{aligned}$$

The first source term on the RHS represents vortex stretching due to both the effective advection drift $\boldsymbol{v}^\star$ and the small-scale random flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$. The second source term describes rotating interactions between the strain rate tensors of the small-scale flow and the large-scale component. This term has the same form as the baroclinic term in compressible flows. The third term comes mainly from the inhomogeneity of the subgrid scale mixing. In particular, this term vanishes in an isotropic turbulent model. The last term is the enstrophy (defined as the energy of the large-scale vorticity) dissipation for both large and small scales. Note that under such sources and sinks, the enstrophy integrated over

the domain is not preserved in time a priori. For future work, it is interesting to verify if the enstrophy driven by (3.22) remains bounded over a long time interval.

3.3 Hydrostatic and Boussinesq approximations

In general, the oceanic and atmospheric flows are stratified in such a way that the horizontal density variations are quite small compared to its vertical mean component. In this case the stochastic Navier-Stokes system (3.10) can be simplified with some approximations. Analogously to the classical framework (Vallis, 2017), the hydrostatic balance and the Boussinesq approximation will be adopted in this section.

3.3.1 Stochastic hydrostatic primitive equations

This section presents a stochastic version of the hydrostatic primitive equations (Vallis, 2017). First, in large-scale atmospheric and oceanic circulations, the horizontal scale of the motion is usually much larger than the vertical one. As a result, the horizontal and vertical momentums are often treated separately. Under LU framework, this reads:

(Horizontal momentum equation)

$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times (\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t) = -\frac{1}{\rho}\boldsymbol{\nabla}_{\mathcal{H}}(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}) + \nu\nabla^2\big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t\big), \tag{3.23a}$$

(Vertical momentum equation)

$$\mathbb{D}_t w = -\frac{1}{\rho}\frac{\partial}{\partial z}(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}) - g\,\mathrm{d}t + \nu\nabla^2(w\,\mathrm{d}t + \sigma_z \mathrm{d}B_t), \tag{3.23b}$$

where $\boldsymbol{u}$ (*resp.* $\boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t$) and w (*resp.* $\sigma_z \mathrm{d}B_t$) are the horizontal and vertical components of the three-dimensional large-scale flow $\boldsymbol{v}$ (*resp.* the small-scale flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$); f is the Coriolis parameter which is defined either by $f = 2\tilde{\Omega}\sin\theta$ in spherical coordinates (with the Earth's angular rotation rate $\tilde{\Omega}$ and latitude variable θ), or by $f = f_0 + \beta y$ using the beta-plane approximation (with constant f_0 and β). The horizontal gradient is denoted by $\boldsymbol{\nabla}_{\mathcal{H}} = [\partial_x, \partial_y]^T$.

The hydrostatic balance is a first coarse approximation of the large-scale circulation such that the acceleration term $\mathbb{D}_t w$ on the LHS of Equation (3.23b) has a lower order of magnitude than the terms on the RHS, hence the vertical momentum equation reduces

to

$$\frac{\partial}{\partial z}(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma) = -g\,\mathrm{d}t. \tag{3.24}$$

According to the canonical decomposition of a semimartingale, the previous equation is equivalent to

$$\frac{\partial p}{\partial z} = g, \qquad \frac{\partial \mathrm{d}p_t^\sigma}{\partial z} = 0. \tag{3.25}$$

Together with the mass continuity equations, we summarize here the stochastic hydrostatic primitive equations:

(Horizontal momentum)

$$\mathbb{D}_t\boldsymbol{u} + \boldsymbol{f}\times(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t) = -\frac{1}{\rho}\boldsymbol{\nabla}_{\mathcal{H}}(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma) + \nu\nabla^2(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t), \tag{3.26a}$$

(Hydrostatsy)

$$\frac{\partial p}{\partial z} = g, \qquad \frac{\partial \mathrm{d}p_t^\sigma}{\partial z} = 0, \tag{3.26b}$$

(Mass)

$$\mathbb{D}_t\rho = 0, \tag{3.26c}$$

(Continuity)

$$\boldsymbol{\nabla}\cdot\boldsymbol{v}^\star = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0. \tag{3.26d}$$

The fact that the martingale pressure component $\mathrm{d}p_t^\sigma$ is vertically independent is due to a strong hydrostatic approximation associated to an hypothesized weak vertical (random) acceleration $\mathbb{D}_t w$. For future works, it would be worth modifying this hydrostatic balance by scaling the vertical momentum equation (3.23b) properly based on different levels of noise. For instance, this procedure is performed for the geostrophic approximation in Chapter 4. In this work, we keep the simplified primitive system (3.26) to derive the stochastic shallow water equations. This is detailed in Section 4.1.

3.3.2 Stochastic simple Boussinesq equations

This section reviews a stochastic representation of the simple Boussinesq system (Vallis, 2017), which has a bigger validity range than the previous system (3.26). To this end, let us first decompose the time-smooth pressure component and the density into

$$p(\boldsymbol{x},t) = \overline{p}(z) + p'(\boldsymbol{x},t), \tag{3.27a}$$

$$\rho(\boldsymbol{x},t) = \overline{\rho}(z) + \rho'(\boldsymbol{x},t), \tag{3.27b}$$

in which the scalar fields $\overline{\rho}(z) \triangleq \rho_b + \rho_0(z)$ and $\overline{p}(z)$ are stationary components in equilibrium that depend only on height. The fluctuating components are random functions, *i.e.* they depend on the random velocity component. From the random momentum equation (3.10a), it can be readily inferred that the stationary components are related by the hydrostatic balance:

$$\frac{\partial \overline{p}}{\partial z} = -g\Big(\rho_b + \rho_0(z)\Big). \tag{3.28}$$

From the Boussinesq approximation, we assume that the density fluctuation is much smaller than the stationary component, *i.e.* $|\rho'| \ll \rho_b$. As such, the random density anomalies are transported by the stochastic flow such that

$$\mathbb{D}_t(\rho - \rho_b) = 0. \tag{3.29}$$

Furthermore, the dynamics of density fluctuations can be expressed through the buoyancy variable $b = -g(\rho/\rho_b)$, namely

$$\mathbb{D}_t b = 0. \tag{3.30}$$

Substituting expression (3.27b) for b and introducing the Brunt-Väisälä stratification frequency $N^2(z) = -g(\partial\rho_0/\partial z)/\rho_b$, the previous equation can written as

$$\mathbb{D}_t b + N^2(w^\star \,\mathrm{d}t + \sigma_z \mathrm{d}B_t) = \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,(\boldsymbol{a}_{\cdot z} N^2)\,\mathrm{d}t, \tag{3.31}$$

where $w^\star$ is the vertical component of the effective drift $\boldsymbol{v}^\star$ and $\boldsymbol{a}_{\cdot z}$ stands for the z column vector of $\boldsymbol{a}$.

Let us then summarize the so-called stochastic simple Boussinesq equations:

(Momentum equations)

$$\mathbb{D}_t\boldsymbol{v} + f\boldsymbol{k}\times(\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = b\boldsymbol{k}\,\mathrm{d}t - \frac{1}{\rho_b}\boldsymbol{\nabla}(p'\,\mathrm{d}t + \mathrm{d}p_t^\sigma) + \nu\nabla^2(\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{3.32a}$$

Thermodynamic equation

$$\mathbb{D}_t b + N^2(w^\star\,\mathrm{d}t + \sigma_z \mathrm{d}B_t) = \frac{1}{2}\,\boldsymbol{\nabla\cdot}\,(\boldsymbol{a}_{\cdot z} N^2)\,\mathrm{d}t, \tag{3.32b}$$

(Continuity equation)

$$\boldsymbol{\nabla\cdot}\,\boldsymbol{v}^\star = 0, \qquad \boldsymbol{\nabla\cdot}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0. \tag{3.32c}$$

Neglecting the friction force terms in the momentum equation yields a stochastic Euler-Boussinesq model. In comparison to the stochastic Navier-Stokes system (3.10), the buoy-

ancy term in the momentum equation (3.32a) constitutes an additional random forcing of the vertical large-scale velocity component. In a similar way as shown in Section 3.2.1, this system can also be turned into an LES system with a scale-separation assumption (with the large-scale component of bounded variation). However, in this case the momentum equation retains its random nature due to the buoyancy forcing. From this system of equations a diverse set of approximated models can be obtained through nondimensionalization and power series expansions in terms of small Rossby number (ratio of the inertial magnitude to Coriolis magnitude) with proper scalings. These developments follow the same path as in the deterministic setting. However, the noise introduces an additional degree of freedom that must be appropriately accounted for (Bauer et al., 2020b; Resseguier et al., 2017b,c). Instances of these approximated models include the planetary geostrophic (PG) model, the quasi-geostrophic (QG) model and the surface quasi-geostrophic (SQG) model.

3.4 Connection with Craik-Leibovich equations

In ocean dynamics, the wave-current interactions are described by the Craik-Leibovich system, in which the action of a wave induced velocity, the Stokes drift, produces a so-called "vortex force" that causes streaking in the flow. This section shows that these results can be generalized as coming from the statistical inhomogeneity effect of the small-scale flow component. To this end, we first review briefly the Craik-Leibovich theory (Craik and Leibovich, 1976; Leibovich, 1980; Holm, 1996) delineating the role played by the Stokes drift. We then draw a parallel framework between the Craik-Leibovich system and our stochastic framework leading to a general stochastic Craik-Leibovich system.

3.4.1 Craik-Leibovich equations and Stokes drift

The Craik-Leibovich (CL) equations (see (Craik and Leibovich, 1976; Holm, 1996; Leibovich, 1980; McWilliams et al., 1997) for different elegant derivations) parametrize the effect of surface gravity waves on the evolution of the mean current $\overline{\boldsymbol{v}}$. It reads:

$$\frac{\partial \overline{\boldsymbol{v}}}{\partial t} + \overline{\boldsymbol{v}} \cdot \boldsymbol{\nabla}\, \overline{\boldsymbol{v}} = -\boldsymbol{\nabla} \widetilde{p} + \boldsymbol{v}_S \times \boldsymbol{\omega}, \qquad \boldsymbol{\nabla} \cdot \overline{\boldsymbol{v}} = 0, \tag{3.33a}$$

$$\widetilde{p} = p + \frac{1}{2}\|\overline{\boldsymbol{v}} + \boldsymbol{v}_S\|^2 - \frac{1}{2}\|\overline{\boldsymbol{v}}\|^2, \tag{3.33b}$$

by recognizing the existence of a vortex force $\boldsymbol{v}_S \times \boldsymbol{\omega}$, where $\boldsymbol{\omega} = \boldsymbol{\nabla} \times \overline{\boldsymbol{v}}$ denotes the curl of the mean current (i.e. the mean flow vorticity) while $\widetilde{p}$ is a modified pressure that includes the pressure p as well as a correction term due to Stokes drift (velocity) $\boldsymbol{v}_S$. For a divergence-free large-scale velocity $\overline{\boldsymbol{v}}$, the pressure p is the solution of the following Poisson equation:

$$-\nabla^2\Big(p + \frac{1}{2}\|\overline{\boldsymbol{v}} + \boldsymbol{v}_S\|^2 - \frac{1}{2}\|\overline{\boldsymbol{v}}\|^2\Big) = \boldsymbol{\nabla}\cdot\Big((\overline{\boldsymbol{v}}\cdot\boldsymbol{\nabla})\overline{\boldsymbol{v}} - \boldsymbol{v}_S \times \boldsymbol{\omega}\Big). \tag{3.34}$$

The Stokes drift is related to the velocity of the surface waves $\boldsymbol{v}'$. In CL theory, it is given by

$$\boldsymbol{v}_S = \overline{\left(\int_{t_0}^{t} \boldsymbol{v}' \mathrm{d}s\right) \cdot \boldsymbol{\nabla}\, \boldsymbol{v}'}, \tag{3.35}$$

where the overbar represents a time average over fast variation at fixed Eulerian position. Leibovich demonstrated that for the specific case of nearly rapid irrotational oscillations and surface waves with small slope, the Eulerian mean velocity $\overline{\boldsymbol{v}}$ is related to the Lagrangian mean $\boldsymbol{v}_L$ through the Stokes drift velocity $\boldsymbol{v}_S$ (Andrews and McIntyre, 1978; Leibovich, 1980):

$$\overline{\boldsymbol{v}} = \boldsymbol{v}_L - \boldsymbol{v}_S + O(\epsilon^4). \tag{3.36}$$

The vortex force $\boldsymbol{v}_S \times \boldsymbol{\omega}$ acts as a Lamb vector between the mean flow vorticity and the Stokes drift. The Stokes drift and the associated vortex force term in the momentum equation have been shown to be the main players in the interaction between the gravity waves and the mean current. Note that in its various derivations, the Craik-Leibovich system requires to assume irrotational waves with a weak slope and a divergence free Stokes drift (Holm, 1996; Leibovich, 1980; McWilliams et al., 2004). Formally, it does not strictly apply to general interactions between small-scale velocity fluctuations and the resolved flow. In the following, we show how a stochastic representation enables us to model such interactions by expressing them as an effect of the inhomogeneity of velocity fluctuations leading to the emergence of an associated vortex force.

3.4.2 Itô-Stokes drift

In this section we focus on the effective advection drift (3.5) involved in the LU model. This term unveils the contribution of inhomogeneity at the small-scales as a driver to trigger large-scale structures in the flow.

In order to interpret the action of the turbophoresis term in (3.5) on the large-scale

flow, we rewrite the stochastic momentum equation (3.32a) in an equivalent form outlining the contribution of the effective advection. For reasons that will be made clear in the following, the turbophoresis term

$$\boldsymbol{v}_s \triangleq \frac{1}{2}\,\boldsymbol{\nabla}\cdot\boldsymbol{a} \tag{3.37}$$

is, henceforth, referred to as the "Itô-Stokes" drift. In its expanded form, the stochastic momentum equation (3.32a) reads:

$$\begin{aligned}&\mathrm{d}_t\boldsymbol{v} + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\boldsymbol{v} - \frac{1}{2}\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{v})\,\mathrm{d}t + \boldsymbol{f}\times(\boldsymbol{v}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\\ &\quad = b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}(p'\,\mathrm{d}t + \mathrm{d}p_t^\sigma).\end{aligned} \tag{3.38}$$

By applying a change of variable in (3.38) from $\boldsymbol{v}$ to $\boldsymbol{v}^\star$, assuming a stationary Itô-Stokes drift, we get,

$$\begin{aligned}&\mathrm{d}_t\boldsymbol{v}^\star + \underbrace{\big((\boldsymbol{v}^\star - \boldsymbol{v}_s)\,\mathrm{d}t}_{\text{turbophoresis}} + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)\cdot\boldsymbol{\nabla}\,\boldsymbol{v}^\star - \underbrace{\frac{1}{2}\,\boldsymbol{\nabla}\cdot\big(\boldsymbol{a}\boldsymbol{\nabla}(\boldsymbol{v}^\star + \boldsymbol{v}_s)\big)\,\mathrm{d}t}_{\text{subgrid scales (SGS)}} - \underbrace{\boldsymbol{v}_s\cdot\boldsymbol{\nabla}\,\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}_{\text{Itô-Stokes advection}}\\ &\quad = b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}\mathrm{d}\widetilde{p}_t^\sigma - \underbrace{\boldsymbol{\omega}_s\times(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{Itô-Stokes force}} - \boldsymbol{f}\times(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\\ &\quad - \underbrace{\boldsymbol{f}\times\boldsymbol{v}_s\,\mathrm{d}t}_{\text{Coriolis Stokes}} + \underbrace{\boldsymbol{v}_s\times(\boldsymbol{\omega}^\star\,\mathrm{d}t + \boldsymbol{\nabla}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{vortex force}},\end{aligned} \tag{3.39}$$

where the curl of effective advection and Itô-Stokes drift are denoted as $\boldsymbol{\nabla}\times\boldsymbol{v}^\star = \boldsymbol{\omega}^\star$ and $\boldsymbol{\nabla}\times\boldsymbol{v}_S = \boldsymbol{\omega}_s$, respectively, and a modified pressure including a Bernoulli head term and a noise term has been introduced:

$$\mathrm{d}\widetilde{p}_t^\sigma = \frac{1}{\rho_b}(p'\,\mathrm{d}t + \mathrm{d}p_t^\sigma) + \frac{1}{2}\Big((\|\boldsymbol{v}^\star\|^2 + \|\boldsymbol{v}_s\|^2) - \|\boldsymbol{v}^\star - \boldsymbol{v}_s\|^2\Big)\,\mathrm{d}t + \boldsymbol{v}_s\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \tag{3.40}$$

by using the vector identity (3.14). On the RHS of equation (3.39), a Craik-Leibovich vortex force appears. This force depicts the statistical contribution of the inhomogeneity carried by the variance of the random field on the large-scale current. Hence, this momentum equation may be seen as a general stochastic expression of the Craik-Leibovich system in which the turbophoresis term replaces the Stokes drift associated to wave motion. This is the reason why we designate this term with the more general descriptive name of Itô-Stokes drift. This equation still includes a turbophoresis term and a large-scale dissipation. The Coriolis force now includes a correction term that depends on the Itô-Stokes drift. An advection of the small-scale component by the Itô-Stokes drift also emerges. In

addition to the vortex force, this stochastic formulation includes another force, referred to here as the *Itô-Stokes force*, related to the interaction of the flow with the vorticity of the Itô-Stokes drift. This force could also be gathered with the Coriolis term to form a Coriolis force corrected by the Itô-Stokes vorticity. This allows us to notice an interesting particular case. Given a downward curl of the Itô-Stokes drift and a negligible Coriolis force in front of the Itô-Stokes force, a correction term appears acting opposite to the Coriolis force. Note that the Itô-Stokes drift depends on the variance tensor $\boldsymbol{a}$, which is defined as the variance of the random displacement ($\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$) divided by a decorrelation time τ (i.e. $a \propto L^2/\tau$). Thus, the divergence and curl of the Itô-Stokes drift scales as the inverse of this decorrelation time (i.e. $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot a \propto 1/\tau$). They are small for sufficiently large decorrelation times such as infra gravity or long waves. For gravity waves, the divergence and curl of the Itô-Stokes drift are negligible only for very regular small-scale inhomogeneity such as waves with small slopes. For isotropic random fluctuations, as in isotropic turbulence, the variance tensor is constant in space and the Itô-Stokes drift cancels. In that case, the effective advection and the large-scale velocity are identical and there is no vortex force nor Itô-Stokes force to structure the large-scale flow components. In such a scenario, the only interaction between the large-scale velocity component and the fluctuations is the large-scale diffusion term and the random advection term. Both of them are in equilibrium to ensure energy conservation.

3.4.3 Stochastic Craik-Leibovich equations

In the following we show how Equation (3.39) can be connected with the Craik-Leibovich equation. First, Equation (3.39) can be simplified with additional assumptions on the inhomogeneity of the small-scale component. Given the incompressibility condition on the effective advection, $\boldsymbol{\nabla}\cdot\boldsymbol{v}^\star = \boldsymbol{\nabla}\cdot(\boldsymbol{v} - \boldsymbol{v}_S) = 0$, and an incompressible large-scale flow, we infer an incompressible Itô-Stokes drift $\boldsymbol{\nabla}\cdot\boldsymbol{v}_S = 0$. In this case, the summation of the SGS contribution on $\boldsymbol{v}^\star$ and the turbophoresis term gives:

$$-\frac{1}{2}\,\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{v}^\star\right) - (\boldsymbol{v}_S\cdot\boldsymbol{\nabla})\boldsymbol{v}^\star = -\frac{1}{2}\sum_{i,j=1}^{d}\partial_{x_i}\partial_{x_j}(a_{ij}\boldsymbol{v}^\star), \tag{3.41}$$

and the Itô-Stokes diffusion term reads:

$$-\frac{1}{2}\,\boldsymbol{\nabla}\cdot\left((\boldsymbol{a}\boldsymbol{\nabla})\boldsymbol{v}_s\right) = -\frac{1}{2}\sum_{i,j=1}^{d}\partial_{x_i}\partial_{x_j}(a_{ij}\boldsymbol{v}_s) + (\boldsymbol{v}_s\cdot\boldsymbol{\nabla})\boldsymbol{v}_s.\boldsymbol{\omega}_s \tag{3.42}$$

The last term of this equation can be written:

$$\Big(\boldsymbol{v}_s\cdot\boldsymbol{\nabla}\Big)\boldsymbol{v}_s = \frac{1}{2}\boldsymbol{\nabla}(\|\boldsymbol{v}_s\|^2) - \boldsymbol{v}_s\times\boldsymbol{\omega}_s. \tag{3.43}$$

Equation (3.39) can thus be written as

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{v}^\star + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\boldsymbol{v}^\star - \underbrace{\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}\Big(a_{ij}(\boldsymbol{v}^\star+\boldsymbol{v}_s)\Big)\,\mathrm{d}t}_{\text{SGS}} - \underbrace{\boldsymbol{v}_s\cdot\boldsymbol{\nabla}\,\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}_{\text{noise advection}}\\
&= b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}\Big(\mathrm{d}\widetilde{p}_t^{\sigma} + \frac{1}{2}\|\boldsymbol{v}_s\|^2\,\mathrm{d}t\Big) - \underbrace{\boldsymbol{\omega}_s\times\Big((\boldsymbol{v}^\star+\boldsymbol{v}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\Big)}_{\text{Itô-Stokes force}}\\
&- \boldsymbol{f}\times(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) - \underbrace{\boldsymbol{f}\times\boldsymbol{v}_s\,\mathrm{d}t}_{\text{Coriolis Stokes}} + \underbrace{\boldsymbol{v}_s\times(\boldsymbol{\omega}^\star\,\mathrm{d}t + \boldsymbol{\nabla}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{vortex force}}.
\end{aligned} \tag{3.44}$$

Coriolis and Itô-Stokes forces can be gathered to get a Coriolis effect modified by the Itô-Stokes drift yielding,

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{v}^\star + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\boldsymbol{v}^\star - \underbrace{\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}\Big(a_{ij}(\boldsymbol{v}^\star+\boldsymbol{v}_s)\Big)\,\mathrm{d}t}_{\text{SGS}} - \underbrace{v_s\cdot\boldsymbol{\nabla}\,\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}_{\text{noise advection}}\\
&= b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}\Big(\mathrm{d}\widetilde{p}_t^{\sigma} + \frac{1}{2}\|\boldsymbol{v}_s\|^2\,\mathrm{d}t\Big) - \underbrace{(\boldsymbol{\omega}_s+\boldsymbol{f})\times\Big((\boldsymbol{v}^\star+\boldsymbol{v}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\Big)}_{\text{Coriolis and Itô-Stokes force}}\\
&+ \underbrace{\boldsymbol{v}_s\times(\boldsymbol{\omega}^\star\,\mathrm{d}t + \boldsymbol{\nabla}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{vortex force}}.
\end{aligned} \tag{3.45}$$

When the Itô-Stokes drift term is sufficiently smooth (with a negligible curl) – which together with the null divergence, implies that $\boldsymbol{v}_s$ is quasi-harmonic, i.e. $\nabla^2(\boldsymbol{v}_s)\approx 0$, we obtain:

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{v}^\star + (\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,\boldsymbol{v}^\star - \underbrace{\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}\Big(a_{ij}(\boldsymbol{v}^\star+\boldsymbol{v}_s)\Big)\,\mathrm{d}t}_{\text{SGS}} - \underbrace{\boldsymbol{v}_s\cdot\boldsymbol{\nabla}\,\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}_{\text{noise advection}}\\
&= b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}\Big(\mathrm{d}\widetilde{p}_t^{\sigma} + \frac{1}{2}\|\boldsymbol{v}_s\|^2\,\mathrm{d}t\Big) - \underbrace{\boldsymbol{f}\times\Big((\boldsymbol{v}^\star+\boldsymbol{v}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\Big)\,\mathrm{d}t}_{\text{Coriolis and Stokes}}\\
&+ \underbrace{\boldsymbol{v}_s\times(\boldsymbol{\omega}^\star\,\mathrm{d}t + \boldsymbol{\nabla}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{vortex force}}.
\end{aligned} \tag{3.46}$$

As previously described in Section 3.2.1, these equations can be written in an LES deterministic form (assuming $\boldsymbol{v}^\star$ is a finite variation process) through a decomposition in terms of martingale and finite variation terms. For the large-scale velocity component this yields (from (3.45)) a momentum equation of the form,

$$\begin{aligned}\frac{\partial \boldsymbol{v}^\star}{\partial t} + \boldsymbol{v}^\star \cdot \boldsymbol{\nabla}\, \boldsymbol{v}^\star - \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\Big(\boldsymbol{a}(\boldsymbol{v}^\star + \boldsymbol{v}_s)\Big) \\ = b\,\boldsymbol{k} - \boldsymbol{\nabla}\Big(\widetilde{p} + \frac{1}{2}\|\boldsymbol{v}_s\|^2\Big) - \underbrace{\boldsymbol{f} \times (\boldsymbol{v}^\star + \boldsymbol{v}_s)}_{\text{Coriolis and Stokes}} + \underbrace{\boldsymbol{v}_s \times \boldsymbol{\omega}^\star}_{\text{vortex force}},\end{aligned} \tag{3.47}$$

which recovers the CL form (3.33) with a pressure term

$$\widetilde{p} = p' + \frac{1}{2}\|\boldsymbol{v}^\star + \boldsymbol{v}_s\|^2 - \frac{1}{2}\|\boldsymbol{v}^\star\|^2 = p' + \frac{1}{2}\|\boldsymbol{v}\|^2 - \frac{1}{2}\|\boldsymbol{v}^\star\|^2, \tag{3.48}$$

accounting for the kinetic energy of the Stokes drift. This pressure term corresponds to the pressure correction terms found in (Harcourt, 2015; Holm, 1996; McWilliams et al., 1997).

In the original CL theory, the Stokes drift is identified with the residual velocity of the fast orbital motion of the waves. It arises from a linear perturbation theory and can be defined (in the assumption of small surface wave slope and nearly irrotational wave motion) as the difference between the Lagrangian flow velocity of a fluid element and the averaged Eulerian flow velocity at a fixed point. In our stochastic framework, it corresponds, instead, to a statistical correction resulting from inhomogeneity of the small-scale. The quasi-harmonic assumption used to get the original CL system does not truly correspond to the usual small slope wave assumption supporting the CL derivation (Leibovich, 1980). However, the small slopes assumption can be shown to yield negligible Stokes drift curl for idealized monochromatic linear waves – for long period waves the Itô-Stokes drift curl is negligible and for high frequency the waves amplitude must be very small.

As discussed previously, this quasi-harmonic assumption considers either a sufficiently large decorrelation time for the variance of the velocity fluctuations or velocity fields with sufficiently smooth inhomogeneity. Within the hypotheses pertaining to the derivation of the Craik-Leibovich equations from the LU framework (i.e deterministic $\boldsymbol{v}^\star$ and stationary quasi-harmonic Itô-Stokes drift), the effective advection $\boldsymbol{v}^\star$ can be associated to the mean Eulerian component whereas $\boldsymbol{v}$ is, by definition, a smooth Lagrangian velocity component.

Without these assumptions, we get the more generalized system (3.39) in which the small-scale component interacts with itself through a Lamb vector involving the curl of the Itô-Stokes drift. It is important to stress that this general stochastic CL equation is equivalent to the Euler momentum equations (3.38). This link between the stochastic Euler equations and the Craik-Leibovich equations clearly shows the potential impact of inhomogeneity of the small-scale random field in shaping large-scale structures such as Langmuir cells. Let us remark that for the buoyancy equation (3.32b) or the transport of any scalar no particular changes needs to be done in our setting. These transport equations involve the effective advection $\boldsymbol{v}^\star$ which includes the Itô-Stokes drift correction.

The principal consequence of the connection between the Craik-Leibovich system and the stochastic system (3.32a) is that in the latter, the explicit inclusion of the vortex force is not necessary to trigger secondary circulations. In the stochastic setting, such circulations require only a small-scale velocity component with an appropriately defined Itô-Stokes drift. This constitutes a simpler procedure than considering an LES representation with an explicit vortex force and the accompanying associated modifications (Coriolis modification; fluctuation-fluctuation interaction; modified pressure, etc.).

Furthermore, for accurate noise models (learned, for instance, from high resolution data), one might expect not only to reproduce complex interactions between the mean current and the surface waves but also interactions coming from sufficiently persistent small-scale inhomogeneity (for example, these small-scales could be triggered by the internal waves that arise due to interaction of the bathymetry with tidal waves). In that respect, in our setting, the traditional Stokes drift of the CL system can be interpreted as a particular instance of small-scale inhomogeneity arising from gravity surface waves and wind forcing.

3.5 Connection with Gent-McWilliams parametrization

This section shows briefly how the LU model can be connected with the Gent-McWilliams (GM) parametrization (Gent and McWilliams, 1990; Gent et al., 1995) which is widely adopted in global ocean models. To this end, let us first explain the principles of such scheme. As we known, the deformation radius in the ocean is at most of the order of 100 km. This makes big challenges for global ocean models in resolving the baroclinic instabilities. In particular, the effects of the mesoscale eddies must be modeled or parametrized

in coarse-resolution ocean models. However, many SGS processes such as most of the diffusion-like parameterizations are not successful in this case. In fact, baroclinic instability produces the major transfers of available potential energy (APE) to eddy kinetic energy (EKE) by flattening density surface. This is quite different from pure mixing procedure. Instead of a diffusion-type parametrization, the so-called Gent-McWilliams parametrization is more recommended in ocean modeling. This scheme proceeds by including a so-called bolus velocity $\boldsymbol{v}_{\rm GM} \triangleq (\boldsymbol{u}_{\rm GM}, w_{\rm GM})^T$ to the large-scale currents. This bolus velocity is basically defined from the isoneutral slope vector $\boldsymbol{s} \triangleq -(\kappa_{\rm GM}\boldsymbol{\nabla}_{\mathcal{H}}\rho)/(\partial\rho/\partial z)$ with a diffusion coefficient $\kappa_{\rm GM}$ (of unit $\mathrm{m}^2\cdot\mathrm{s}^{-1}$) to be tuned (could be space-time depended). It reads

$$\boldsymbol{u}_{\rm GM} = -\frac{\partial(\kappa_{\rm GM}\boldsymbol{s})}{\partial z}, \qquad w_{\rm GM} = \boldsymbol{\nabla}_{\mathcal{H}}\cdot(\kappa_{\rm GM}\boldsymbol{s}). \tag{3.49}$$

As such, the bolus velocity effectively advects the density field and with the chosen signs, leads to a reduction of the frontal slope, and substitutes the effect of baroclinic instability. The strength of the effect is controlled by the parameter $\kappa_{\rm GM}$. Furthermore, one can verify that the bolus velocity is divergence-free, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{v}_{\rm GM} = 0$. Indeed, this adiabatic property should be preserved by numerical discretization. Under the GM scheme, the governing equation of a transported tracer Θ (could be potential temperature, salinity or passive tracers) is given by

$$\frac{\partial\Theta}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{v}\Theta) = -\boldsymbol{\nabla}\cdot(\boldsymbol{v}_{\rm GM}\Theta), \qquad \boldsymbol{\nabla}\cdot\boldsymbol{v} = 0. \tag{3.50}$$

Furthermore, as shown in Griffies (1998), the choice (3.49) corresponds to a specific antisymmetric tensor $\boldsymbol{a}_{\rm GM}$ satisfying $\boldsymbol{v}_{\rm GM} = -\boldsymbol{\nabla}\cdot\boldsymbol{a}_{\rm GM}$ such that

$$\boldsymbol{a}_{\rm GM} = \kappa_{\rm GM}\begin{pmatrix} 0 & 0 & -s_x \\ 0 & 0 & -s_y \\ s_x & s_y & 0 \end{pmatrix}, \tag{3.51}$$

where $(s_x, s_y)^T = \boldsymbol{s}$. Subsequently, Equation (3.50) can be re-written as

$$\frac{\partial\Theta}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{v}\Theta) = \boldsymbol{\nabla}\cdot(\boldsymbol{a}_{\rm GM}\boldsymbol{\nabla}\Theta), \tag{3.52}$$

in which the antisymmetric properties $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a}_{\rm GM} = 0$ and $\mathrm{tr}(\boldsymbol{a}_{\rm GM}\boldsymbol{H}_\Theta) = 0$ (with $\boldsymbol{H}$ the Hessian matrix) have been used. In addition, both divergence-free and antisymmetry

ensure that all tracer moments are conserved.

Let us now go back to our stochastic framework. We first define an isopycnal projector as follows:

$$\mathbf{P}(\boldsymbol{x},t) \triangleq \Big(\mathbf{I}_d - \frac{\boldsymbol{\nabla}\rho(\boldsymbol{\nabla}\rho)^T}{|\boldsymbol{\nabla}\rho|^2}\Big)(\boldsymbol{x},t). \tag{3.53}$$

Such projector satisfies immediately

$$\mathbf{P}\boldsymbol{\nabla}\rho = 0, \qquad \mathbf{P}^T = \mathbf{P}, \qquad \mathbf{P}^2 = \mathbf{P}. \tag{3.54}$$

Let us then consider an initial noise $\boldsymbol{\sigma}_0 \mathrm{d}\boldsymbol{B}_t$ with $\boldsymbol{\sigma}_0$ *a priori* modeled or parametrized. Applying the projector $\mathbf{P}$ for such noise $\boldsymbol{\sigma}_0 \mathrm{d}\boldsymbol{B}_t$ leads to a new noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ aiming at flattening density surfaces, that is

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \mathbf{P}(\boldsymbol{\sigma}_0\mathrm{d}\boldsymbol{B}_t), \qquad \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot\boldsymbol{\nabla}\,\rho = 0. \tag{3.55}$$

From the variance's definition (2.8) and the projection property (3.54), one can show that the diffusive flux of density cancels, *i.e.* $\boldsymbol{a}\boldsymbol{\nabla}\rho = 0$. Thus, the evolution of density reduces to

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\Big(\rho(\boldsymbol{v}-\boldsymbol{v}_s)\Big) = 0, \qquad \boldsymbol{\nabla}\cdot(\boldsymbol{v}-\boldsymbol{v}_s) = 0. \tag{3.56}$$

Unlike the GM scheme, the Itô-Stokes drift in our settings is not divergence-free after projection, yet the stochastic continuity equation ensures the adiabatic background. On the other hand, the transport of tracer Θ remains a SPDE, namely

$$\mathrm{d}_t\Theta + \Big((\boldsymbol{v}-\boldsymbol{v}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\Big)\cdot\boldsymbol{\nabla}\,\Theta - \frac{1}{2}\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}\boldsymbol{\nabla}\Theta) = 0. \tag{3.57}$$

As shown in Section (2.2), the last term diffuses the tracer, since the variance $\boldsymbol{a}$ is purely symmetric and positive semidefinite. Nevertheless, energy lost by the diffusion balances with the energy brought by the third random forcing term (see Section 2.3). The statistical-induced velocity $\boldsymbol{v}_s$ is more general than the bolus velocity, in the sense that it is not necessary divergence-free, yet the effective advection drift $\boldsymbol{v}-\boldsymbol{v}_s$ is. Note that the above procedure holds for the buoyancy variable:

$$\frac{\partial b}{\partial t} + \boldsymbol{\nabla}\cdot(b(\boldsymbol{v}-\boldsymbol{v}_s)) + N^2(w-w_s) = 0. \tag{3.58}$$

In the classical case, Treguier et al. (1997) suggested that the GM parameterization at the oceanic boundaries should include along-boundary eddy-induced velocity and residual flux

of buoyancy. Quasigeostrophic theory is used to satisfy these constraints in their works.

When there is more than one active/passive tracer, GM scheme is often combined with Redi isoneutral diffusion (Redi, 1982; Griffies et al., 1998). Under a small angle approximation, $|\boldsymbol{s}| \ll 1$, the Redi diffusion tensor $\boldsymbol{a}_{\mathrm{R}}$ takes the form

$$\boldsymbol{a}_{\mathrm{R}} = \kappa_{\mathrm{R}} \begin{pmatrix} 1 & 0 & s_x \\ 0 & 1 & s_y \\ s_x & s_y & |\boldsymbol{s}|^2 \end{pmatrix}. \tag{3.59}$$

One can immediately see that $\boldsymbol{a}_{\mathrm{R}}$ is symmetric positive semidefinite with positive coefficient κ_{R}. Combining such isoneutral diffusion with the GM scheme gives the following transport equation for tracer Θ:

$$\frac{\partial \Theta}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{v}\Theta) = \boldsymbol{\nabla}\cdot\Big((\boldsymbol{a}_{\mathrm{GM}} + \boldsymbol{a}_{\mathrm{R}})\boldsymbol{\nabla}\Theta\Big), \tag{3.60}$$

in which the term with the antisymmetric $\boldsymbol{a}_{\mathrm{GM}}$ flattens the isopycnals and release the APE while conserving all tracer moment, whereas the term with the symmetric $\boldsymbol{a}_{\mathrm{R}}$ only diffuses along the isopycnals but does not enforce them to flatten out, this term does not change the APE and dissipates all tracer moment except the mean.

On the other hand, the link between the LU model and the isoneutral diffusion have been described by Mémin (2014). For instance, one can consider a simple case in which the initial noise $\boldsymbol{\sigma}_0 \mathrm{d}\boldsymbol{B}_t$ is isotropic and divergence-free, hence the corresponding variance tensor reduces to a constant diagonal matrix $\boldsymbol{a}_0 = \alpha \mathbf{I}_d$ with $\alpha > 0$. As such, after applying the projector $\mathbf{P}$ for $\boldsymbol{\sigma}_0 \mathrm{d}\boldsymbol{B}_t$, the variance becomes

$$\boldsymbol{a} = \mathbf{P}\boldsymbol{a}_0\mathbf{P}^T = \alpha\mathbf{P}. \tag{3.61}$$

Subsequently making the small angle approximation $|\boldsymbol{s}| \ll 1$, the Redi diffusion tensor (3.59) is recovered, *i.e.* $\boldsymbol{a} = \boldsymbol{a}_{\mathrm{R}}$ with $\kappa_{\mathrm{R}} = \alpha$ related to the strength of the initial noise. Let us outline that, unlike the GM-Redi scheme (3.60) (which is based on a decomposition of a mixing tensor into antisymmetric and symmetric parts), by constraining an isopycnal noise (3.55), the LU scheme (3.57) includes separately a Redi-like diffusion through the noise variance and a GM-like advection through the Itô-Stokes drift. For future works, it would be important to clarify that such parametrization leads to an efficient conversion of APE to EKE in a slimily way as in Gent and McWilliams (1990).

Chapter 4

Stochastic geostrophic approximations

Abstract

This chapter reviews some important two-dimensional stochastic equations. Apart from the hydrostatic balances in vertical direction presented in the previous chapter, this chapter focus on another important approximation in the horizontal direction between the pressure gradient and the Coriolis forces, this is the so-called geostrophic balance, which is mainly due to to the dominant rotating effect of the large-scale atmospheric and oceanic circulations. In the following, we first derive the stochastic rotating shallow water (RSW) system, which is a good approximation of the primitive equations. Then, the energy conservation of a simplified stochastic RSW system is demonstrated. Latter, we specify some scaling numbers to adimensionalize the stochastic RSW system and the stochastic simple Boussinesq system presented in the previous chapter. Subsequently, asymptotic approaches are adopted to develop the barotropic planetary geostrophic (PG) and quasigeostrophic (QG) equations using the non-dimensional RSW system, as well as the stochastic primitive PG and vertically continuous version of the QG equations using the non-dimensional Boussinesq system. These resulting two-dimensional models have been successfully tested in several numerical simulations and some interesting results are outlined in Chapters 6, 7 and 8. The work presented in this chapter are based on the published paper Bauer et al. (2020a) and on the manuscript Brecht et al. (2021) that has been submitted.

4.1 Stochastic rotating shallow water equations

In this section, we derive the rotating shallow water equations under LU, denoted as RSW-LU, following the classical strategy presented in Vallis (2017). In particular, we demonstrate one important characteristic of the RSW-LU, namely it preserves the total energy of the large-scale flow.

Let us note that a pseudo-stochastic formulation of the shallow water equations under LU in a non-rotating frame was first proposed by Mémin (2014). Here we provide a full stochastic model.

4.1.1 Derivation of single layer model

The RSW-LU is derived mainly from the stochastic hydrostatic primitive equations (3.26) dropping the viscous terms. First, integrating vertically the hydrostatic balance (3.26b) from 0 to z (see Figure 4.1) under a constant density ρ_0, we have

$$p(x,y,z,t) = p_0(x,y,t) - \rho_0 g z, \tag{4.1}$$

$$\mathrm{d}p_t^\sigma(x,y,z,t) = \mathrm{d}p_t^\sigma(x,y,0,t), \tag{4.2}$$

where p_0 denotes the pressure at the bottom of the basin ($z=0$). Following Vallis (2017), we assume that the weight of the overlying fluid is negligible, *i.e.* $p(x,y,\eta,t) \approx 0$ with η the height of the free surface, leading to $p_0 = \rho_0 g \eta$. This allows us to rewrite Equation (4.1) such that for any $z \in [0,\eta]$ we have

$$p(x,y,z,t) = \rho_0 g\Big(\eta(x,y,t) - z\Big). \tag{4.3a}$$

Subsequently, the pressure gradient forces in the horizontal momentum equation (3.26a) reduce to

$$-\frac{1}{\rho_0}\boldsymbol{\nabla}_{\!\mathcal{H}}\Big(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\Big) = -g\boldsymbol{\nabla}_{\!\mathcal{H}}\eta - \frac{1}{\rho_0}\boldsymbol{\nabla}_{\!\mathcal{H}}\mathrm{d}p_t^\sigma, \tag{4.3b}$$

which do not depend on z according to Equations (4.3a) and (4.2). Therefore, the acceleration terms on the LHS of Equation (3.23a) must not depend on z, hence the shallow water momentum equation can be written as

$$\mathbb{D}_t^{\mathcal{H}}\boldsymbol{u} + \boldsymbol{f}\times\Big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_{\!\mathcal{H}}\mathrm{d}\boldsymbol{B}_t\Big) = -g\boldsymbol{\nabla}_{\!\mathcal{H}}\eta\,\mathrm{d}t - \frac{1}{\rho_0}\boldsymbol{\nabla}_{\!\mathcal{H}}\mathrm{d}p_t^\sigma, \tag{4.4}$$

where $\mathbb{D}_t^{\mathcal{H}}$ is the horizontal stochastic transport operator defined as

$$\mathbb{D}_t^{\mathcal{H}} u \triangleq \mathrm{d}_t u + \Big((\boldsymbol{u} - \boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t\Big)\cdot\boldsymbol{\nabla}_{\mathcal{H}} u - \frac{1}{2}\boldsymbol{\nabla}_{\mathcal{H}}\cdot\Big(\boldsymbol{a}_{\mathcal{H}}\boldsymbol{\nabla}_{\mathcal{H}} u\Big)\,\mathrm{d}t, \tag{4.5}$$

with $\boldsymbol{u}_s \triangleq \frac{1}{2}\boldsymbol{\nabla}_{\mathcal{H}}\cdot\boldsymbol{a}_{\mathcal{H}}$ the two-dimensional Itô-Stokes drift. Note that the three-dimensional variance tensor $\boldsymbol{a}$ consists of a horizontal component $\boldsymbol{a}_{\mathcal{H}}$, a vertical component a_{zz} and a cross component $\boldsymbol{a}_{\mathcal{H}z}$ such that $\boldsymbol{a} = \begin{bmatrix} \boldsymbol{a}_{\mathcal{H}} & \boldsymbol{a}_{\mathcal{H}z} \\ \boldsymbol{a}_{\mathcal{H}z}^T & a_{zz} \end{bmatrix}$. Let us outline that Equation (4.4) is valid only when the cross component $\boldsymbol{a}_{\mathcal{H}z}$ is vertically independent, *i.e.* $\partial\boldsymbol{a}_{\mathcal{H}z}/\partial z = 0$. This assumption is satisfied for noise that do not depend on the vertical coordinate. One may also consider that the horizontal small-scale flow $\boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t$ is spatially uncorrelated with the vertical small-scale flow $\sigma_z \mathrm{d}B_t$, *i.e.* $\boldsymbol{a}_{\mathcal{H}z} = 0$.

In order to derive the shallow water mass equation, let us first integrate vertically the continuity equation (3.26d) from the bottom topography η_b to the free surface η (see Figure 4.1):

$$(w - w_s)|_{z=\eta} - (w - w_s)|_{z=\eta_b} = -h\boldsymbol{\nabla}_{\mathcal{H}}\cdot(\boldsymbol{u} - \boldsymbol{u}_s), \tag{4.6a}$$

$$\sigma\mathrm{d}B_t|_{z=\eta} - \sigma\mathrm{d}B_t|_{z=\eta_b} = -h\boldsymbol{\nabla}_{\mathcal{H}}\cdot\boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t, \tag{4.6b}$$

where $h = \eta - \eta_b$ denotes the thickness of the water column. On the other hand, a small vertical (Eulerian) displacement at the top and the bottom of the fluid leads to a variation of the position of a particular fluid element (Vallis, 2017):

$$\big((w - w_s)\,\mathrm{d}t + \sigma\mathrm{d}B_t\big)\big|_{z=\eta} = \mathbb{D}_t^{\mathcal{H}}\eta, \tag{4.6c}$$

$$\big((w - w_s)\,\mathrm{d}t + \sigma\mathrm{d}B_t\big)\big|_{z=\eta_b} = \mathbb{D}_t^{\mathcal{H}}\eta_b. \tag{4.6d}$$

Combining Equations (4.6), we deduce a stochastic mass equation, then together with the stochastic momentum equation (4.4), we fully obtain a general formulation of the RSW-LU system, namely

(Momentum equation)

$$\mathbb{D}_t\boldsymbol{u} + \boldsymbol{f}\times(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = -g\boldsymbol{\nabla}\eta\,\mathrm{d}t - \frac{1}{\rho_0}\boldsymbol{\nabla}\mathrm{d}p_t^{\sigma}, \tag{4.7a}$$

(Mass equation)

$$\mathbb{D}_t h + h\,\boldsymbol{\nabla}\cdot\big((\boldsymbol{u} - \boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big) = 0. \tag{4.7b}$$

Note that the horizontal symbol $\mathcal{H}$ is dropped in (4.7), as well as in the following (except

for three-dimensional equations), for the sake of simplicity.

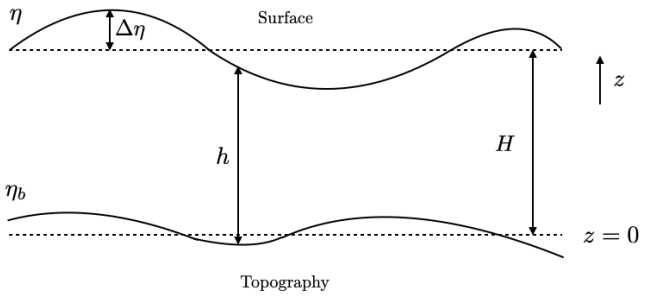

Figure 4.1 – Illustration of a single-layered shallow water system (inspired by Vallis (2017)), h is the thickness of a water column, η is the height of the free surface, η_b is the height of the bottom topography, H is the background thickness and $\Delta\eta$ is the fluctuation of free surface height *w.r.t.* the background thickness. As a result, we have $h = \eta - \eta_b$ and $\eta = H + \Delta\eta$.

4.1.2 Extension to multi-layer model

Following Vallis (2017), the previous single layered RSW-LU system can be extended to a multi-layered model. A layered model is an idealized model for the stratified ocean dynamic, in which the fluid flow is represented as a finite number of moving layers, stacked one upon another and each having a uniform density. It is also assumed that the pressures and velocities are effectively layer-averaged quantities. Suppose now we have N moving layers, the dynamic pressure in the n-th layer can be determined by vertical integrations of hydrostatic balances, namely

$$p_n = -\sum_{i=n}^{N} \rho_0 g_i' \eta_i, \qquad \eta_n = -\sum_{i=1}^{n} h_i, \qquad g_n' = g\frac{\rho_{n+1} - \rho_n}{\rho_0}, \tag{4.8}$$

where g_n' denotes the reduced gravity across the interface between the n-th layer and $(n+1)$-th layer; η_n stands for the interface displacements and h_n is the layer thickness. Let us consider that in the interior of each layer, the ocean dynamic is modeled by the RSW-LU (4.7), then by integration over each layer, the following multi-layer stochastic shallow water system is recovered:

(Momentum equation)

$$\mathbb{D}_t^n \boldsymbol{u}_n + f\boldsymbol{k} \times (\boldsymbol{u}_n \,\mathrm{d}t + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_n) = -\frac{1}{\rho_n}\boldsymbol{\nabla}(p_n \,\mathrm{d}t + (\mathrm{d}p_t^{\sigma})_n), \tag{4.9a}$$

(Mass equation)

$$\mathbb{D}_t^n h_n + h_n \,\boldsymbol{\nabla}\boldsymbol{\cdot}\, ((\boldsymbol{u}_n - (\boldsymbol{u}_s)_n)\,\mathrm{d}t + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_n) = 0. \tag{4.9b}$$

In particular, this stacked model can be considered as a vertical discretization of the three-dimensional stochastic Boussinesq equation (3.32) with the buoyancy variable at each interface defined by

$$b_n = \frac{p_{n+1} - p_n}{\rho_0 (H_n + H_{n+1})/2}, \qquad \forall\, n = 1, \dots, N-1. \tag{4.10}$$

4.1.3 Energy conservation

The single layer system (4.7) is a general formulation to describe the shallow water fluid motion transported by the stochastic flow. However, it is quite difficult to handle the energy and higher order moments of such system without any information of the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and of the random pressure component $\mathrm{d}p_t^\sigma$. In the following, let us consider a simplified RSW-LU system in which the noise is in geostrophic balance with the random pressure term. This type of constraint can be rigorously derived considering a strong hydrostatic balance for the martingale pressure term with $\partial \mathrm{d}p_t^\sigma / \partial z = 0$. The considered system reads:

(Momentum equation)

$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u}\, \mathrm{d}t = -g \boldsymbol{\nabla} \eta\, \mathrm{d}t, \tag{4.11a}$$

(Mass equation)

$$\mathbb{D}_t h + h\, \boldsymbol{\nabla}\cdot \boldsymbol{u}\, \mathrm{d}t = 0, \tag{4.11b}$$

(Continuity equations)

$$\boldsymbol{\nabla}\cdot \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \qquad \boldsymbol{\nabla}\cdot \boldsymbol{u}_s = 0, \tag{4.11c}$$

(Geostrophic noise)

$$\boldsymbol{f} \times \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \approx -\frac{1}{\rho_0} \boldsymbol{\nabla} \mathrm{d}p_t^\sigma. \tag{4.11d}$$

Note that for a sufficiently small noise ($\boldsymbol{\sigma} \approx 0$), this system reduces to the classical RSW system, in which the stochastic transport operator weighted by the unit of time, $\mathbb{D}_t/\mathrm{d}t$, reduces to the material derivative. More importantly, this simplified RSW-LU system preserves the global energy of the large-scale flow (for any realizations):

$$\mathrm{d}_t \int_{\mathcal{A}} \mathrm{E}(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} = 0, \qquad \mathrm{E} \triangleq \underbrace{\frac{\rho_0}{2} h |\boldsymbol{u}|^2}_{\mathrm{KE}} + \underbrace{\frac{\rho_0}{2} g h^2}_{\mathrm{PE}}, \tag{4.12}$$

where $|\boldsymbol{u}|^2 = \boldsymbol{u}\cdot\boldsymbol{u}$ and $\boldsymbol{x} = (x, y) \in \mathcal{A}$ with $\mathcal{A} \subset \mathbb{R}^2$ a bounded horizontal area. We recall that for the shallow water system (Vallis, 2017), the kinetic energy density (KE) and the

potential energy density (PE) are derived from the vertical integrations over the whole water column, *i.e.* $\mathrm{KE} = \int_0^h (\rho_0/2)|\boldsymbol{u}|^2 \,\mathrm{d}z$ and $\mathrm{PE} = \int_0^h (\rho_0/2) g z \,\mathrm{d}z$.

In order to demonstrate the energy conservation more concisely, we adopt an equivalent Stratonovich representation of the simplified RSW-LU system (4.11) using the conversion presented in Section 2.4. It reads:

$$\mathbb{D}_t \circ \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} \,\mathrm{d}t = -g \boldsymbol{\nabla} h \,\mathrm{d}t, \tag{4.13a}$$

$$\mathbb{D}_t \circ h + h \,\boldsymbol{\nabla} \cdot \boldsymbol{u} \,\mathrm{d}t = 0, \tag{4.13b}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = 0, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{u}_s = 0, \tag{4.13c}$$

$$\boldsymbol{f} \times \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t \approx -\boldsymbol{\nabla} \mathrm{d}p_t^\sigma, \tag{4.13d}$$

where the Stratonovich horizontal transport operator is defined by

$$\mathbb{D}_t \circ u \triangleq \mathrm{d}_t \circ u + \Big((\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t \Big) \cdot \boldsymbol{\nabla}\, u. \tag{4.13e}$$

For algebraic simplicity, the constant density ρ_0 is assumed to be 1, as well as the bottom is assumed to be flat, *i.e.* $h = \eta$.

As shown in Section 1.1.4, Stratonovich integrals are defined such that the integration-by-part formula (1.1.6) of ordinary calculus holds. In particular, for two random tracers f and g, we have

$$\mathrm{d}_t \circ (fg) = f \,\mathrm{d}_t \circ g + g \,\mathrm{d}_t \circ f. \tag{4.14}$$

Therefore, the antisymmetric Stratonovich transport operator (4.13e) satisfies immediately

$$\mathbb{D}_t \circ (fg) = g \,\mathbb{D}_t \circ f + f \,\mathbb{D}_t \circ g. \tag{4.15}$$

Applying this rule for PE (4.12) together with the mass equation (4.13b), we obtain

$$\mathbb{D}_t \circ \mathrm{PE} = gh \,\mathbb{D}_t \circ h = -2\,\mathrm{PE}\ \boldsymbol{\nabla} \cdot \boldsymbol{u} \,\mathrm{d}t. \tag{4.16}$$

Similarly, from both mass equation and momentum equation in (4.13), we derive the evolution of KE (4.12):

$$\begin{aligned} \mathbb{D}_t \circ \mathrm{KE} &= h\boldsymbol{u} \cdot \mathbb{D}_t \boldsymbol{u} + \frac{1}{2}|\boldsymbol{u}|^2 \,\mathbb{D}_t \circ h \\ &= -\frac{1}{2}\boldsymbol{u} \cdot \boldsymbol{\nabla}\, (gh^2) \,\mathrm{d}t - \frac{1}{2} h |\boldsymbol{u}|^2 \ \boldsymbol{\nabla} \cdot \boldsymbol{u} \,\mathrm{d}t, \end{aligned} \tag{4.17}$$

noting that $\boldsymbol{u}\cdot(\boldsymbol{f}\times\boldsymbol{u}\,\mathrm{d}t)=0$ yields

$$\mathbb{D}_t\circ\mathrm{KE}+(\boldsymbol{u}\cdot\boldsymbol{\nabla}\,\mathrm{PE}+\mathrm{KE}\;\boldsymbol{\nabla}\cdot\boldsymbol{u})\,\mathrm{d}t=0. \tag{4.18}$$

Subsequently, we deduce the evolution of the density of total energy:

$$\mathbb{D}_t\circ\mathrm{E}+\Big(\mathrm{E}\,\boldsymbol{\nabla}\cdot\boldsymbol{u}+\boldsymbol{\nabla}\cdot(\boldsymbol{u}\,\mathrm{PE})\Big)\,\mathrm{d}t=0. \tag{4.19}$$

Expanding the Stratonovich transport operator (4.13e) together with the continuity equations (4.13c), the previous equation can be re-written as

$$\mathrm{d}_t\circ\mathrm{E}+\boldsymbol{\nabla}\cdot\Big(((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t+\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\,\mathrm{E}+\boldsymbol{u}\,\mathrm{PE}\Big)=0. \tag{4.20}$$

If the fluid domain has zero boundary conditions (e.g. the normal velocities vanishes on each wall or there are no boundaries at all as on the sphere), then one can show that the total energy is invariant in time:

$$\mathrm{d}_t\circ\int_{\mathcal{A}}\mathrm{E}\,\mathrm{d}\boldsymbol{x}=-\oint_{\partial\mathcal{A}}\Big(\mathrm{E}\,((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t+\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{n}+\mathrm{PE}\,\boldsymbol{u}\cdot\boldsymbol{n}\,\mathrm{d}t\Big)\,\mathrm{d}l=0, \tag{4.21}$$

where $\partial\mathcal{A}$ and $\boldsymbol{n}$ denote the area boundaries and the unit normal vector, respectively.

4.2 Geostrophic approximations

In order to derive properly the stochastic geostrophic equations, some scaling numbers are required. This section details the typical scales of variables for both the stochastic shallow water equations (4.7) and the stochastic Boussinesq equations (3.32). In particular, a scaling number for the random noise and its variance is necessary in order to quantify the additional degree of freedom brought by the noise and its variance. The resulting non-dimensional RSW-LU system is useful to develop the barotropic PG and QG equations, whereas the non-dimensional primitive system allows us to derive the three-dimensional PG and QG equations in the subsequent sections.

4.2.1 Non-dimensional stochastic shallow water equations

Let us now scale properly the geostrophic balance for the stochastic rotating shallow water equations (4.7). To this end, we first adimensionalize the basic variables as

$$\boldsymbol{x}=L\,\hat{\boldsymbol{x}},\qquad \boldsymbol{u}=U\,\hat{\boldsymbol{u}},\qquad t=T\,\hat{t},\qquad f=f_0\,\hat{f}, \tag{4.22a}$$

where the capital letters are used for the scales of variables and $\hat{\bullet}$ stand for non-dimensional variables. In the following, only the advection time scale $T = L/U$ is adopted. In addition to these classical scaling numbers, the horizontal variance tensor $\boldsymbol{a}$, which characterizes the strength of uncertainty, is also required and scaled as

$$\boldsymbol{a} = A\,\hat{\boldsymbol{a}}. \tag{4.22b}$$

As mentioned in Section 2.1, since $\boldsymbol{a}$ has the unit of a diffusion tensor ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$), one may consider that A is proportional to UL up to a factor ϵ, *i.e.* $A = \epsilon\, UL$. This factor ϵ, first introduced by Resseguier et al. (2017b), is defined as

$$\epsilon = \frac{T^\sigma}{T}\frac{\mathrm{TKE}}{\mathrm{MKE}}, \tag{4.22c}$$

where T^σ is the correlation time scale of the small-scale component. The mean kinetic energy scale (MKE) is given by U^2 and the turbulent kinetic energy scale (TKE) is defined by A/T^σ. As such, the horizontal noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ can be scaled as

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon} L\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}. \tag{4.22d}$$

In fact, the greater this scaling number ϵ, the larger the variance tensor, hence the stronger the uncertainty. As shown in subsequent sections, using different levels of noise in the stochastic system allows us to model different physical regimes of the large-scale flow.

Let us now review two important non-dimensional numbers – the Rossby number ($\mathrm{R_O}$) and the Burger number ($\mathrm{B_u}$). These numbers are particularly useful in geostrophic approximations. The Rossby number is introduced to measure the ratio between the magnitude of the inertial acceleration ($\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}$) and the magnitude of the Coriolis acceleration ($f\boldsymbol{k}\times\boldsymbol{u}$):

$$\mathrm{R_O} = \frac{U}{f_0 L}. \tag{4.22e}$$

Therefore, a small Rossby number characterizes the dominant effect of earth rotation on the fluid. In particular, it is the case for most of large-scale oceanic circulations with the typical value $\mathrm{R_O} \approx 10^{-2}$. The Burger number measures the squared ratio between the scale of deformation radius L_d and the scale of motion:

$$\mathrm{B_u} = \frac{L_d^2}{L^2}, \tag{4.22f}$$

where the deformation radius in shallow water fluids is given by $L_d = \sqrt{gH}/f_0$, and measures the ratio between buoyancy and rotation effects. The deformation radius gives the typical size of oceanic eddies.

In order to ensure the geostrophic approximation in the stochastic momentum equation (4.7a), $f\boldsymbol{k} \times (\boldsymbol{u}\ \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \approx -\boldsymbol{\nabla}_{\mathcal{H}}(g\eta + \mathrm{d}p_t^\sigma)$, the free surface height $\eta = H + \Delta\eta$ (see Figure 4.1) and the random pressure $\mathrm{d}p_t^\sigma$ necessarily scale as

$$\eta = H + \frac{f_0 U L}{g}\hat{\eta} = H\Big(1 + \frac{\mathrm{R_O}}{\mathrm{B_u}}\hat{\eta}\Big), \tag{4.22g}$$

$$\mathrm{d}p_t^\sigma = \sqrt{\epsilon} f_0 L^2\, \mathrm{d}\hat{p}_{\hat{t}}^\sigma. \tag{4.22h}$$

Substituting these expressions (4.22) in (4.7), we obtain the non-dimensional RSW-LU system, that is

(Momentum equation)

$$\mathrm{R_O}\,\hat{\mathbb{D}}_{\hat{t}}^\epsilon \hat{\boldsymbol{u}} + \hat{f}\boldsymbol{k} \times (\hat{\boldsymbol{u}}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) = -g\hat{\boldsymbol{\nabla}}\hat{\eta}\,\mathrm{d}\hat{t} - \sqrt{\epsilon}\,\hat{\boldsymbol{\nabla}}\mathrm{d}\hat{p}_{\hat{t}}^\sigma, \tag{4.23a}$$

(Mass equation)

$$\mathrm{B_u^{-1}R_O}\,\hat{\mathbb{D}}_{\hat{t}}^\epsilon \hat{\eta} + (1 + \mathrm{B_u^{-1}R_O}\,\hat{\eta})\hat{\boldsymbol{\nabla}} \cdot ((\hat{\boldsymbol{u}} - \epsilon\,\hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) = 0, \tag{4.23b}$$

where $\hat{\mathbb{D}}_{\hat{t}}^\epsilon u \triangleq \mathrm{d}_{\hat{t}} u + \Big((\hat{\boldsymbol{u}} - \epsilon\,\hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\Big) \cdot \hat{\boldsymbol{\nabla}} u - (\epsilon/2)\hat{\boldsymbol{\nabla}} \cdot \Big(\hat{\boldsymbol{a}}\hat{\boldsymbol{\nabla}} u\Big)\,\mathrm{d}\hat{t}$. These non-dimensional equations will be latter used to derive both the stochastic barotropic PG and QG equations.

4.2.2 Non-dimensional stochastic primitive Boussinesq equations

We now describe the non-dimensional representation of the stochastic Boussinesq equations (3.32) without viscous terms. Apart from the scaling numbers presented in (4.22), let us introduce some additional scales for the vertical variables and the cross tensor components in the three-dimensional equations. The scaling number of the large-scale vertical velocity can be simply deduced from the continuity equation (3.32c), namely

$$w = \delta\, U\, \hat{w}, \qquad \delta \triangleq \frac{H}{L} \ll 1. \tag{4.24a}$$

To ensure the geostrophic approximation, the time-smooth pressure component needs to be scaled as

$$p = f_0 U L\, \hat{p}. \tag{4.24b}$$

The classical hydrostatic relation gives us a specific buoyancy scale, that is

$$b = \frac{f_0 U L}{H}\,\hat{b}, \qquad \frac{\partial \hat{b}/\partial \hat{z}}{N^2} = \mathcal{O}(\mathrm{B_u^{-1}R_O}). \tag{4.24c}$$

In particular, to keep the assumption of flat isopycnal (small variation of stratifications) in the QG theory, Resseguier et al. (2017b) proposed to scale the ratio between the vertical scale and the horizontal scale of the noise as follows:

$$\frac{\hat{\sigma}_z \mathrm{d}B_{\hat{t}}}{\hat{\boldsymbol{\sigma}}_{\mathcal{H}} \mathrm{d}\boldsymbol{B}_{\hat{t}}} = \mathcal{O}(\delta\,\mathrm{B_u^{-1}R_O}), \tag{4.24d}$$

which implies

$$\sigma_z \mathrm{d}B_t = (\sqrt{\epsilon}\,\mathrm{B_u^{-1}R_O} H)\,\hat{\sigma}_z \mathrm{d}B_{\hat{t}}. \tag{4.24e}$$

Subsequently, one can deduce the following scales for the cross components and vertical part of the variance tensor. They read, respectively

$$\boldsymbol{a}_{\mathcal{H}z} = (\epsilon\,\mathrm{B_u^{-1}R_O} U H)\,\hat{\boldsymbol{a}}_{\mathcal{H}z}, \qquad a_{zz} = \left(\epsilon\,\delta\,(\mathrm{B_u^{-1}R_O})^2 U H\right)\hat{a}_{zz}. \tag{4.24f}$$

Substituting these expressions (4.24) together with (4.22) in (3.32), we obtain the non-dimensional stochastic Boussinesq system, which reads:

(Momentum equations)

$$\mathrm{R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon}\hat{\boldsymbol{u}} + \hat{f}\boldsymbol{k}\times\left(\hat{\boldsymbol{u}}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\right) = -\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\left(\hat{p}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{d}\hat{p}_{\hat{t}}^{\sigma}\right), \tag{4.25a}$$

(Hydrostatic equation)

$$\delta^2\,\mathrm{R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon}\hat{w} = \hat{b}\,\mathrm{d}\hat{t} - \frac{\partial}{\partial\hat{z}}\left(\hat{p}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{d}\hat{p}_{\hat{t}}^{\sigma}\right), \tag{4.25b}$$

(Continuity equations)

$$\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\cdot(\hat{\boldsymbol{u}} - \epsilon\,\hat{\boldsymbol{u}}_s) + \frac{\partial}{\partial\hat{z}}\left(\hat{w} - \epsilon\,\mathrm{B_u^{-1}R_O}\,\hat{w}_s\right) = 0, \tag{4.25c}$$

$$\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\cdot\hat{\boldsymbol{\sigma}}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_{\hat{t}} + \mathrm{B_u^{-1}R_O}\,\frac{\partial}{\partial\hat{z}}\hat{\sigma}_z\mathrm{d}B_{\hat{t}} = 0, \tag{4.25d}$$

(Thermodynamic equation)

$$\mathrm{B_u^{-1}R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon}\hat{b} + \left(\left(\hat{w} - 2\epsilon\,\mathrm{B_u^{-1}R_O}\,\hat{w}_s\right)\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{B_u^{-1}R_O}\,\hat{\sigma}_z\mathrm{d}B_{\hat{t}}\right) = 0, \tag{4.25e}$$

where the non-dimensional stochastic transport operator $\hat{\mathbb{D}}_{\hat{t}}^{\epsilon}$ in three dimensions is defined as

$$\begin{aligned}\hat{\mathbb{D}}_{\hat{t}}^{\epsilon}u \triangleq \mathrm{d}_{\hat{t}}u &+ \Big((\hat{\boldsymbol{u}} - \frac{\epsilon}{2}(\boldsymbol{\nabla}_{\mathcal{H}}\cdot\hat{\boldsymbol{a}}_{\mathcal{H}} + \mathrm{B_u^{-1}R_O}\,\frac{\partial\hat{\boldsymbol{a}}_{\mathcal{H}z}}{\partial\hat{z}}))\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\Big)\cdot\hat{\boldsymbol{\nabla}}_{\mathcal{H}}u \\ &+ \Big((\hat{w} - \frac{\epsilon}{2}\mathrm{B_u^{-1}R_O}\,(\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\cdot\hat{\boldsymbol{a}}_{\mathcal{H}z} + \mathrm{B_u^{-1}R_O}\,\frac{\partial\hat{a}_{zz}}{\partial\hat{z}}))\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{B_u^{-1}R_O}\,\hat{\sigma}_z\mathrm{d}B_{\hat{t}}\Big)\frac{\partial u}{\partial\hat{z}} \\ &- \frac{\epsilon}{2}\Big(\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\cdot(\hat{\boldsymbol{a}}_{\mathcal{H}}\hat{\boldsymbol{\nabla}}_{\mathcal{H}}u) + \mathrm{B_u^{-1}R_O}\,(\hat{\boldsymbol{\nabla}}_{\mathcal{H}}\cdot(\hat{\boldsymbol{a}}_{\mathcal{H}z}\frac{\partial u}{\partial\hat{z}}) + \frac{\partial}{\partial\hat{z}}(\hat{\boldsymbol{a}}_{\mathcal{H}z}\hat{\boldsymbol{\nabla}}_{\mathcal{H}}u) + \mathrm{B_u^{-1}R_O}\,\frac{\partial}{\partial\hat{z}}(\hat{a}_{zz}\frac{\partial u}{\partial\hat{z}}))\Big)\,\mathrm{d}\hat{t}.\end{aligned}$$

As such, one can use the non-dimensional system (4.25) to derive the stochastic, vertically continuous, three-dimensional PG, QG and SQG equations under specific assumptions.

4.3 Stochastic planetary geostrophic equations

This section describes the formal derivation of the stochastic PG equations following the same strategy as in Vallis (2017). To this end, we make the following assumptions:

(i) The Rossby number is small, *i.e.* $\mathrm{R_O} \ll 1$;

(ii) The scale of motion is significantly larger than the deformation scale, *i.e.* $\mathrm{B_u} \ll 1$. In particular, we consider $\mathrm{B_u} = \mathcal{O}(\mathrm{R_O})$;

(iii) The location uncertainty is moderate, *i.e.* $\epsilon = \mathcal{O}(1)$ or $\epsilon\,\mathrm{R_O} = \mathcal{O}(\mathrm{R_O})$.

Let us show the derivation in the barotropic case using the non-dimensional RSW-LU (4.23). We first expand the prognostic variables $\boldsymbol{u}$ and η in power series of a small Rossby number:

$$\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}^{(0)} + \mathrm{R_O}\,\hat{\boldsymbol{u}}^{(1)} + \mathrm{R_O^2}\,\hat{\boldsymbol{u}}^{(2)} + \cdots, \qquad \hat{\eta} = \hat{\eta}^{(0)} + \mathrm{R_O}\,\hat{\eta}^{(1)} + \mathrm{R_O^2}\,\hat{\eta}^{(2)} + \cdots, \tag{4.26}$$

where $\bullet^{(i)}$ denotes the i-th order quantity in the Rossby number. Substituting subsequently these expressions in (4.23) leads to

$$\begin{aligned}\mathrm{R_O}\,(\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\boldsymbol{u}}^{(0)} + \hat{f}\boldsymbol{k}\times\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}t) + \hat{f}\boldsymbol{k}\times(\hat{\boldsymbol{u}}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) \\ = -\hat{\boldsymbol{\nabla}}(\hat{\eta}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{d}\hat{p}_{\hat{t}}^{\sigma}) - \mathrm{R_O}\,\hat{\boldsymbol{\nabla}}\hat{\eta}^{(1)}\,\mathrm{d}\hat{t} + \mathcal{O}(\mathrm{R_O^2}),\end{aligned} \tag{4.27a}$$

$$\begin{aligned}\mathrm{B_u^{-1}R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\eta}^{(0)} + (1 + \mathrm{B_u^{-1}R_O}\,\hat{\eta}^{(0)})\,\hat{\boldsymbol{\nabla}}\cdot((\hat{\boldsymbol{u}}^{(0)} - \epsilon\,\hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) \\ + \mathrm{R_O}\,\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t} = \mathrm{B_u^{-1}}\mathcal{O}(\mathrm{R_O^2}),\end{aligned} \tag{4.27b}$$

where $\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)} \triangleq \mathrm{d}_{\hat{t}} u + \big((\hat{\boldsymbol{u}}^{(0)} - \epsilon\, \hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\big)\cdot\hat{\boldsymbol{\nabla}} u\,\mathrm{d}\hat{t} - (\epsilon/2)\hat{\boldsymbol{\nabla}}\cdot(\hat{\boldsymbol{a}}\hat{\boldsymbol{\nabla}} u)\,\mathrm{d}\hat{t}$. Using the PG assumption (i) and (iii) together with the canonical decomposition (1.1.1), the lowest order system of equations (4.27) reduce to

$$\hat{f}\boldsymbol{k}\times\hat{\boldsymbol{u}}^{(0)} = -\hat{\boldsymbol{\nabla}}\hat{\eta}^{(0)}, \qquad \hat{f}\boldsymbol{k}\times\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}} = -\hat{\boldsymbol{\nabla}}\mathrm{d}\hat{p}^{\sigma}_{\hat{t}}, \tag{4.28a}$$

$$\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\eta}^{(0)} + (1+\hat{\eta}^{(0)})\,\hat{\boldsymbol{\nabla}}\cdot\big((\hat{\boldsymbol{u}}^{(0)} - \hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\big) = 0. \tag{4.28b}$$

Note that with a given random pressure $\mathrm{d}p^{\sigma}_t$, the previous equations constitute a closed system. Supposing now that the Rossby number goes to zero ($\mathrm{R_O}\to 0$) and using the scalings (4.22), we recover the following (dimensional) stochastic PG equations:

(Mass equation)

$$\mathbb{D}_t h + h\,\boldsymbol{\nabla}\cdot\big((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big) = 0, \tag{4.29a}$$

(Geostrophic balances)

$$f\boldsymbol{k}\times\boldsymbol{u} = -\boldsymbol{\nabla}\eta, \qquad f\boldsymbol{k}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\boldsymbol{\nabla}\mathrm{d}p^{\sigma}_t, \tag{4.29b}$$

where $\eta = h+\eta_b$ (see Figure 4.1). In particular, if $f = f_0$ is constant, then the geostrophic balances (4.29b) provides divergence-free flow components, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{u} = \boldsymbol{\nabla}\cdot\boldsymbol{u}_s = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0$. Thus, the mass equation (4.29a) simplifies to a pure (stochastic) transport equation, $\mathbb{D}_t h = 0$.

Let us outline that the previous system (4.29) is only valid under moderate uncertainty. When the uncertainty becomes strong, both large-scale geostrophic balance and mass equation must be modified. In particular, let us consider that $\epsilon = \mathcal{O}(\mathrm{R_O})$. From the non-dimensional equations (4.27), we obtain a modified geostrophic balance including a diffusion term and the Itô-Stokes drift:

$$f\boldsymbol{k}\times\boldsymbol{u} - \frac{1}{2}\sum_{i,j=1,2}\frac{\partial^2(a_{ij}u)}{\partial x_i\partial x_j} = -g\boldsymbol{\nabla}\eta, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}h) = 0. \tag{4.30}$$

These equations do no longer evolve in time due to the strong uncertainty and are thus of limited interest. However, they resemble to the equations of the Ekman layer and might find some applications in such a context, with a different interpretation of such stationary model. On the other hand, if the uncertainty is sufficiently small, in particular for $\epsilon = \mathcal{O}(\mathrm{R_O^2})$, then the classical PG equations will be recovered, namely

$$f\boldsymbol{k}\times\boldsymbol{u} = -g\boldsymbol{\nabla}\eta, \qquad \frac{\partial h}{\partial t} + \boldsymbol{\nabla}\cdot(h\boldsymbol{u}) = 0. \tag{4.31}$$

This explains why we only adopt the moderate uncertainty assumption (iii) in this work. We remark that following a similar strategy as above, a three-dimensional PG model can be derived from the non-dimensional stochastic Boussinesq equations (4.25). The final governing equations under moderate uncertainty read:

(Thermodynamic equation)

$$\mathbb{D}_t b + N^2\Big((w - w_s)\,\mathrm{d}t + \sigma_z \mathrm{d}B_t\Big) = \frac{1}{2}\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{a}_{\cdot z} N^2)\,\mathrm{d}t, \tag{4.32a}$$

(Geostrophic balances)

$$f\boldsymbol{k}\times\boldsymbol{u} = -\boldsymbol{\nabla}_{\mathcal{H}}\eta, \qquad f\boldsymbol{k}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\boldsymbol{\nabla}_{\mathcal{H}}\mathrm{d}p_t^{\sigma}, \tag{4.32b}$$

(Hydrostatic balances)

$$\frac{\partial p}{\partial z} = b, \qquad \frac{\partial \mathrm{d}p_t^{\sigma}}{\partial z} = 0, \tag{4.32c}$$

(Continuity equations)

$$\frac{\partial}{\partial z}(w - w_s) = \boldsymbol{\nabla}_{\mathcal{H}}\cdot(\boldsymbol{u} - \boldsymbol{u}_s), \qquad \frac{\partial}{\partial z}\sigma_z \mathrm{d}B_t = -\boldsymbol{\nabla}_{\mathcal{H}}\cdot\boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t, \tag{4.32d}$$

where the first buoyancy equation is driven by a three-dimensional stochastic transport operator $\mathbb{D}_t$.

4.4 Stochastic quasi-geostrophic equations

We describes in this section the formal derivation of the stochastic QG equations. To this end, we make the following assumptions:

(i) The Rossby number is small;

(ii) The scale of motion is similar to that of the deformation scale. In particular, we assume $\mathrm{B_u} = \mathcal{O}(1)$ or $\mathrm{B_u^{-1}R_o} = \mathcal{O}(\mathrm{R_o})$;

(iii) The location uncertainty is moderate;

(iv) The variations of Coriolis parameter are small, namely $|\beta L| < |f_0|$.

Note that the second assumption (ii) differs from that in the PG case. In fact, from Equation (4.22g), it means that the variations of stratification under the QG approximation are small.

4.4.1 Derivation of barotropic model

We first derive the barotropic QG equations following a similar strategy as in the PG case 4.3. According to the additional assumption (iv), the non-dimensional Coriolis parameter $\hat{f}$ on beta-plane can be specified by

$$\hat{f} = \hat{f}_0 + \mathrm{R_O}\,\hat{\beta}\hat{y}. \tag{4.33}$$

Expanding this expression in the non-dimensional system (4.23), we have

$$\begin{aligned} \mathrm{R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\boldsymbol{u}}^{(0)} &+ \hat{f}_0\boldsymbol{k}\times(\hat{\boldsymbol{u}}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) \\ &+ \mathrm{R_O}\,\boldsymbol{k}\times\Big(\hat{\beta}\hat{y}(\hat{\boldsymbol{u}}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) + \hat{f}_0\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t}\Big) \\ &= -\hat{\boldsymbol{\nabla}}(\hat{\eta}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\mathrm{d}\hat{p}_{\hat{t}}^{\sigma}) - \mathrm{R_O}\,\hat{\boldsymbol{\nabla}}\hat{\eta}^{(1)}\,\mathrm{d}\hat{t} + \mathcal{O}(\mathrm{R_O^2}), \end{aligned} \tag{4.34a}$$

$$\begin{aligned} \mathrm{B_u^{-1}R_O}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\eta}^{(0)} &+ (1+\mathrm{B_u^{-1}R_O}\,\hat{\eta}^{(0)})\hat{\boldsymbol{\nabla}}\cdot\Big((\hat{\boldsymbol{u}}^{(0)} - \epsilon\,\hat{\boldsymbol{u}}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}\Big) \\ &+ \mathrm{R_O}\,\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t} = \mathrm{B_u^{-1}}\mathcal{O}(\mathrm{R_O^2}), \end{aligned} \tag{4.34b}$$

where $\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}$ is defined as in the PG case. Using the QG assumption (i)–(iii), the zeroth order equations reduce to

$$\hat{f}_0\hat{\boldsymbol{u}}^{(0)} = \hat{\boldsymbol{\nabla}}^{\perp}\hat{\eta}^{(0)}, \qquad \hat{f}_0\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}} = \hat{\boldsymbol{\nabla}}^{\perp}\mathrm{d}\hat{p}_{\hat{t}}^{\sigma} \tag{4.35a}$$

$$\hat{\boldsymbol{\nabla}}\cdot(\hat{\boldsymbol{u}}^{(0)} - \epsilon\,\hat{\boldsymbol{u}}_s) = 0, \qquad \hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}} = 0, \tag{4.35b}$$

where $\boldsymbol{\nabla}^{\perp} \triangleq (-\partial/\partial y,\ \partial/\partial x)^T$. Combining these two results, we obtain

$$\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}^{(0)} = \hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}_s = \hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}} = 0. \tag{4.35c}$$

Subsequently, by differencing (4.34) and (4.35), we deduce the first order equations, namely

$$\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\boldsymbol{u}}^{(0)} + \boldsymbol{k}\times\Big(\hat{\beta}\hat{y}(\hat{\boldsymbol{u}}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\boldsymbol{\sigma}}\mathrm{d}\boldsymbol{B}_{\hat{t}}) + \hat{f}_0\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t}\Big) = -\hat{\boldsymbol{\nabla}}\hat{\eta}^{(1)}\,\mathrm{d}\hat{t}, \tag{4.36a}$$

$$\mathrm{B_u^{-1}}\,\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\eta}^{(0)} + \hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t} = 0. \tag{4.36b}$$

In order to close the system, we need to eliminate the first order quantities $\hat{\boldsymbol{u}}^{(1)}$ and $\hat{\eta}^{(1)}$. Taking first the vertical curl ($\boldsymbol{k}\cdot\boldsymbol{\nabla}\times$) of Equation (4.36a), we deduce an evolution law

for the relative vorticity, defined as $\hat{\omega}^{(0)} \triangleq \partial\hat{v}^{(0)}/\partial\hat{x} - \partial\hat{u}^{(0)}/\partial\hat{y}$. It reads:

$$\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}\hat{\omega}^{(0)} + \hat{\beta}(\hat{v}^{(0)}\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{\sigma}_y \mathrm{d}B_{\hat{t}}) + \hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\hat{t} = \epsilon\,\hat{S}_1\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{S}_2\,\mathrm{d}B_{\hat{t}}, \tag{4.37a}$$

in which some source and sinks of vorticity emerged such that

$$S_1 \triangleq \mathrm{tr}\Big(\boldsymbol{\nabla}^{\perp}\boldsymbol{u}_s^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big) + \frac{1}{2}\,\boldsymbol{\nabla}\cdot\Big(\frac{\partial\boldsymbol{a}}{\partial x}\boldsymbol{\nabla}v - \frac{\partial\boldsymbol{a}}{\partial y}\boldsymbol{\nabla}u\Big), \tag{4.37b}$$

$$S_2\,\mathrm{d}B_t \triangleq -\mathrm{tr}\Big(\boldsymbol{\nabla}^{\perp}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big). \tag{4.37c}$$

Substituting (4.36b) in (4.37a), as well as introducing the Coriolis correction into the transport operator, we obtain a single equation, that is

$$\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}(\hat{\omega}^{(0)} + \hat{\beta}\hat{y} - \hat{f}_0\mathrm{B_u}\,\hat{\eta}^{(0)}) = \epsilon\,(\hat{S}_1 - 2\hat{\beta}\hat{v}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{S}_2\,\mathrm{d}B_{\hat{t}}. \tag{4.38}$$

Moreover, the large-scale geostrophic balance in (4.35a) allows us to define a lowest order stream function $\hat{\psi}^{(0)} \triangleq \hat{\eta}^{(0)}/\hat{f}_0$ such that

$$\hat{\boldsymbol{u}}^{(0)} = \hat{\boldsymbol{\nabla}}^{\perp}\hat{\psi}^{(0)}, \qquad \hat{\omega}^{(0)} = \hat{\nabla}^2\hat{\psi}^{(0)}. \tag{4.39}$$

Therefore, Equation (4.38) can be re-written as

$$\hat{\mathbb{D}}_{\hat{t}}^{\epsilon(0)}(\hat{\nabla}^2\hat{\psi}^{(0)} + \hat{\beta}\hat{y} - \hat{f}_0^2\mathrm{B_u}^{-1}\,\hat{\psi}^{(0)}) = \epsilon\,(\hat{S}_1 - 2\hat{\beta}\hat{v}_s)\,\mathrm{d}\hat{t} + \sqrt{\epsilon}\,\hat{S}_2\,\mathrm{d}B_{\hat{t}}. \tag{4.40}$$

Finally, restoring the dimensions with (4.22), we get the stochastic barotropic QG equations:

(Potential vorticity equations)

$$\mathbb{D}_t q = \sum_{i=1,2}\mathrm{J}\Big((\sigma\mathrm{d}B_t)^i - u_s^i\,\mathrm{d}t, u^i\Big) - \frac{1}{2}\,\boldsymbol{\nabla}\cdot\Big(\frac{\partial\boldsymbol{a}}{\partial x_i^{\perp}}\boldsymbol{\nabla}u^i\Big)\,\mathrm{d}t - \beta\frac{\partial a_{i2}}{\partial x_i}\,\mathrm{d}t, \tag{4.41a}$$

$$q = \nabla^2\psi + \beta y - \frac{\psi}{L_d^2}, \tag{4.41b}$$

(Continuity equations)

$$\boldsymbol{u} = \boldsymbol{\nabla}^{\perp}\psi, \qquad \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^{\perp}\varphi\mathrm{d}B_t, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0, \tag{4.41c}$$

where $\mathrm{J}(f,g)$ denotes the Jacobian determinant of the vector composed of the two functions f and g with $\mathrm{J}(f,g) = (\partial f/\partial x)(\partial g/\partial y) - (\partial g/\partial x)(\partial f/\partial y)$, $\psi = (g/f_0)\eta$, $\varphi\mathrm{d}B_t =$

$\mathrm{d}p_t^\sigma/f_0$ and ψ_s are respectively the large-scale, the small-scale and the Itô-Stokes stream functions. We remark that the stochastic system (4.41) is valid only under moderate uncertainty. Resseguier et al. (2017c) shows that beyond this scaling the geostrophic balance is eventually modified and includes correction terms to isobaric velocities.

In Chapter 7, the stochastic barotropic QG system (4.41) is used to study the structuration effect of the small-scale random field on the large-scale flow. In addition, such random model under additional forcing and damping is also adopted for analyses of long-term statistical predictions.

Unlike the classical QG model, (4.41) involves sources of PV. For instance, the first term on the RHS of (4.41a) has a similar form as the additional term introduced in the barotropic Leray α-model studied in Holm and Nadiga (2003). In fact, it can be interpreted as the rotating interaction between the strain vectors (McWilliams, 1984; Weiss, 1991) of the large-scale flow and the small-scale flow component. This antisymmetric source term cancels when the two strain vectors are collinear. We highlight in the following their contributions to the conservative energy budget.

4.4.2 Energy conservation

To demonstrate the conservation of total energy for (4.41), we first express this system in an equivalent Stratonovich representation in a similar way as presented in Section 2.4. The Stratonovich form of system (4.41) reads:

$$\boldsymbol{u} = \boldsymbol{\nabla}^\perp\psi, \qquad \boldsymbol{u}_s = \boldsymbol{\nabla}^\perp\psi_s, \qquad \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp\varphi\circ\mathrm{d}B_t, \tag{4.42a}$$

$$\mathbb{D}_t\circ(\nabla^2\psi - F\psi) + \beta\frac{\partial}{\partial x}(\psi\,\mathrm{d}t + \varphi\circ\mathrm{d}B_t) = \mathrm{d}_t\circ S, \tag{4.42b}$$

$$\mathrm{d}_t\circ S \triangleq \underbrace{\mathrm{tr}\Big(\boldsymbol{\nabla}^\perp\boldsymbol{u}_s^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big)\,\mathrm{d}t}_{\triangleq \underline{S_1}} \underbrace{-\,\mathrm{tr}\Big(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big)}_{\triangleq S_2\circ\mathrm{d}B_t}, \tag{4.42c}$$

where $F \triangleq 1/L_d^2$. Remember that the total energy of barotropic QG system is defined as

$$E = \int_{\mathcal{A}}\Big(\frac{1}{2}\|\boldsymbol{\nabla}\psi\|^2 + \frac{1}{2}F\psi^2\Big)\,\mathrm{d}\boldsymbol{x}. \tag{4.43}$$

We remark that the terms involving β in (4.42b) make no direct contribution to the energy budget (Vallis, 2017). For the sake of simplicity, we drop them in the following. Let us

now compute the time evolution of the total energy (4.43) for the system (4.42):

$$\begin{aligned}\mathrm{d}_t\circ E &= \int_{\mathcal{A}} \left((\mathrm{d}_t\circ\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}\,\psi + F\psi\mathrm{d}_t\circ\psi\right)\mathrm{d}\boldsymbol{x} = -\int_{\mathcal{A}}\psi\mathrm{d}_t\circ(\nabla^2\psi - F\psi)\,\mathrm{d}\boldsymbol{x}\\ &= \underbrace{\int_{\mathcal{A}}\psi((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\,(\nabla^2\psi - F\psi)\,\mathrm{d}\boldsymbol{x}}_{\mathrm{I}} - \int_{\mathcal{A}}\psi\mathrm{d}_t\circ S\,\mathrm{d}\boldsymbol{x},\end{aligned} \tag{4.44}$$

in which the integration-by-parts formula is used under close impermeable boundary condition (uniform ψ along the boundary) or under periodic boundaries. The first term (I) can be further manipulated using the incompressible constraints (4.41c) such that

$$\begin{aligned}\mathrm{I} &= \int_{\mathcal{A}}\psi\,\boldsymbol{\nabla}\cdot\Big(((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)(\nabla^2\psi - F\psi)\Big)\,\mathrm{d}\boldsymbol{x}\\ &= -\int_{\mathcal{A}}\boldsymbol{\nabla}\psi\cdot((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)(\nabla^2\psi - F\psi)\,\mathrm{d}\boldsymbol{x}\\ &= -\int_{\mathcal{A}}\boldsymbol{\nabla}\psi\cdot(-\boldsymbol{u}_s\,\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)(\nabla^2\psi - F\psi)\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{4.45}$$

Notice that until this step, we show that without any noise, the classical barotropic QG system conserves well the total energy. Let us then expand the contribution of the random source term in (4.44) as follows:

$$\begin{aligned}-\int_{\mathcal{A}}\psi S_2\circ\mathrm{d}B_t\,\mathrm{d}\boldsymbol{x} &= \sum_{i=1,2}\int_{\mathcal{A}}\psi\frac{\partial\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t}{\partial x_i^\perp}\cdot\boldsymbol{\nabla}\,u^i\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{i=1,2}\int_{\mathcal{A}}\boldsymbol{\nabla}\cdot(\psi\frac{\partial\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t}{\partial x_i^\perp}u^i)\,\mathrm{d}\boldsymbol{x}\\ &= -\int_{\mathcal{A}}(\boldsymbol{\nabla}\psi)^T(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp)\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathcal{A}}\boldsymbol{\nabla}^\perp\cdot(\boldsymbol{u}(\boldsymbol{\nabla}\psi)^T)\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathcal{A}}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\nabla^2\psi\,\mathrm{d}\boldsymbol{x} + \underbrace{\int_{\mathcal{A}}(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp\,(\boldsymbol{\nabla}\psi)^T)\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\,\mathrm{d}\boldsymbol{x}}_{\mathrm{II}},\end{aligned} \tag{4.46}$$

where the last term in the previous equation is null through (4.42a). Indeed,

$$\begin{aligned}\mathrm{II} &= \int_{\mathcal{A}}(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp\,(\boldsymbol{\nabla}\psi)^T)\boldsymbol{\nabla}^\perp\varphi\circ\mathrm{d}B_t\,\mathrm{d}\boldsymbol{x}\\ &= -\int_{\mathcal{A}}\boldsymbol{\nabla}^\perp\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp\,(\boldsymbol{\nabla}\psi)^T)\varphi\circ\mathrm{d}B_t\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathcal{A}}\boldsymbol{\nabla}^\perp\cdot(\boldsymbol{u}^\perp\cdot\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\psi)^T)\varphi\circ\mathrm{d}B_t\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathcal{A}}\Big(\mathrm{tr}(\boldsymbol{\nabla}^\perp(\boldsymbol{u}^\perp)^T\boldsymbol{\nabla}(\boldsymbol{u}^\perp)^T) + (\boldsymbol{u}^\perp\cdot\boldsymbol{\nabla})(\boldsymbol{\nabla}^\perp\cdot\boldsymbol{\nabla}\,\psi)\Big)\,\mathrm{d}\boldsymbol{x} = 0.\end{aligned} \tag{4.47}$$

Similarly, one can show that for the other source/sink term $\underline{S_1}$ (of bounded variation) in (4.44), we have

$$-\int_{\mathcal{A}} \psi \underline{S_1}\,\mathrm{d}\boldsymbol{x} = -\int_{\mathcal{A}} \boldsymbol{\nabla}\psi \cdot \boldsymbol{u}_s \nabla^2\psi\,\mathrm{d}\boldsymbol{x}. \tag{4.48}$$

Substituting the expressions (4.48), (4.46) and (4.45) in Equation (4.44), the evolution of the total energy reduces to

$$\begin{aligned}
\mathrm{d}_t{\circ}E &= \int_{\mathcal{A}} F\psi\boldsymbol{\nabla}\psi\cdot(\boldsymbol{\sigma}{\circ}\mathrm{d}\boldsymbol{B}_t - \boldsymbol{u}_s\,\mathrm{d}t)\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathcal{A}} \boldsymbol{\nabla}{\cdot}(F\psi\boldsymbol{\nabla}^\perp\psi)(\varphi{\circ}\mathrm{d}B_t - \psi_s\,\mathrm{d}t)\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathcal{A}} (F\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}^\perp\psi + F\psi\,\boldsymbol{\nabla}{\cdot}\,\boldsymbol{\nabla}^\perp\psi)(\varphi{\circ}\mathrm{d}B_t - \psi_s\,\mathrm{d}t)\,\mathrm{d}\boldsymbol{x} = 0.
\end{aligned} \tag{4.49}$$

Therefore, the stochastic barotropic QG equations (4.42), proposed in this work, conserves along time the total energy for any realizations. As we can see, the energy of the source processes compensates the increase of energy due to the advection of the resolved flow by the unresolved one. Besides, both large-scale and small-scale geostrophic balances and the divergence-free condition of the Itô-Stokes drift play important roles in this energy conservation. This result is consistent to that found for the simplified RSW-LU system (4.11).

4.4.3 Extension to multi-layer and primitive models

In a very similar way as for the multi-layered shallow water system (4.9), there is a straightforward generalization of the stochastic QG equations in terms of a N-layered model. The layered PV are defined as

$$q_1 = \nabla^2\psi_1 + \beta y - \frac{f_0^2}{g_1' H_1}(\psi_1 - \psi_2), \tag{4.50a}$$

$$q_n = \nabla^2\psi_n + \beta y - \frac{f_0^2}{g_n' H_n}(\psi_n - \psi_{n+1}) + \frac{f_0^2}{g_{n-1}' H_n}(\psi_{n-1} - \psi_n),\ \forall\, 1 < n < N, \tag{4.50b}$$

$$q_N = \nabla^2\psi_N + \beta y + \frac{f_0^2}{g_{N-1}' H_N}(\psi_{N-1} - \psi_N). \tag{4.50c}$$

The evolution of such PV and the continuity equations are given by

$$\mathbb{D}_t^n q_n = \sum_{i=1,2} \mathrm{J}((\sigma\mathrm{d}B_t)_n^i - (u_s)_n^i\,\mathrm{d}t, u_n^i) - \frac{1}{2}\,\boldsymbol{\nabla}{\cdot}\,(\frac{\partial \boldsymbol{a}_n}{\partial x_i^\perp}\boldsymbol{\nabla}u_n^i)\,\mathrm{d}t - \beta\frac{\partial (a_{i2})_n}{\partial x_i}\,\mathrm{d}t, \tag{4.50d}$$

$$\boldsymbol{u}_n = \boldsymbol{\nabla}^\perp\psi_n, \qquad (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_n = \boldsymbol{\nabla}^\perp(\varphi\mathrm{d}B_t)_n, \qquad \boldsymbol{\nabla}{\cdot}(\boldsymbol{u}_s)_n = 0. \tag{4.50e}$$

The numerical implementation of a three layer QG model is presented in Chapter 8, to illustrate numerically that the proposed random model (4.50) allows us to better represent high-resolution ocean internal variability on a very coarse mesh. Let us note that the layered PV (4.50) can be considered as a vertical discretization of the (vertically) continuous three-dimensional PV. Following a similar strategy as in the barotropic case, the three-dimensional QG model (Resseguier et al., 2017b) can be derived from the (non-dimensional) stochastic simple Boussinesq equation (4.25). The final governing equations under moderate uncertainty read:

(Potential vorticity equations)

$$\mathbb{D}_t q = \sum_{i=1,2} \mathrm{J}\Big((\sigma \mathrm{d}B_t)^i - u_s^i \,\mathrm{d}t, u^i\Big) - \frac{1}{2}\boldsymbol{\nabla}\cdot\Big(\frac{\partial \boldsymbol{a}}{\partial x_i^\perp}\boldsymbol{\nabla} u^i\Big)\,\mathrm{d}t - \beta\frac{\partial a_{i2}}{\partial x_i}\,\mathrm{d}t, \tag{4.51a}$$

$$q = \nabla^2\psi + \beta y + \frac{\partial}{\partial z}\Big(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\Big), \tag{4.51b}$$

(Continuity equations)

$$\boldsymbol{u} = \boldsymbol{\nabla}^\perp\psi, \qquad \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp\varphi \mathrm{d}B_t, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0. \tag{4.51c}$$

Hereafter, an elliptic PDE can be constituted by assuming that the interior PV in (4.51) is zero and imposing two boundary conditions in a semi-infinite vertical domain:

$$\nabla^2\psi + \frac{\partial}{\partial z}\Big(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\Big) = 0, \tag{4.52a}$$

$$\frac{\partial \psi}{\partial z} = b, \qquad (z = 0), \tag{4.52b}$$

$$\psi = 0, \qquad (z \to -\infty), \tag{4.52c}$$

with a constant stratification N. The above elliptic problem can be solved in horizontal Fourier space, which gives the inversion relationship between the surface buoyancy and the surface stream function:

$$\widehat{\psi} = \frac{f_0}{N}\frac{\widehat{b}}{\|\boldsymbol{k}\|}, \tag{4.53}$$

where $\widehat{\bullet}$ stands for the Fourier transform coefficient and $\|\boldsymbol{k}\| = \sqrt{k_x^2 + k_y^2}$ denotes the modulus of two-dimensional wave vectors. Taking the inverse Fourier transform together with a stochastic transport of buoyancy at the ocean surface as well as the continuity equations, we recover the stochastic surface quasi-geostrophic (SQG) equations studied

by Resseguier et al. (2017b,c, 2020b):

$$\text{(Buoyancy equations)}$$

$$\mathbb{D}_t b = 0, \qquad b = \frac{N}{f_0}(-\nabla^2)^{1/2}\psi, \tag{4.54a}$$

$$\text{(Continuity equations)}$$

$$\boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi, \qquad \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp \varphi \mathrm{d}B_t, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0. \tag{4.54b}$$

Note that this simple model describes surface flow due to a buoyancy transport. One of it's advantages is that it constitutes a two-dimensional model that captures essential properties of three-dimensional solutions on the surface. It was found by Resseguier et al. (2017b) that, compared to classical models, the proposed stochastic SQG model was more accurate in predicting the extreme events and in diagnosing the frontogenesis and filamentogenesis. In Chapter 6, the ensemble forecasting skills of this random model based on different type of noises are compared to that of a classical model randomized through its initial condition. For future work, it would be interesting to derive some new coupled stochastic models based on multiple space-time scales (Pedlosky, 1984; Grooms et al., 2011). For instance, we may expect that a stochastic PG-QG coupled model would enhance the mesoscale eddies representation in the PG model and that a stochastic SQG-QG coupled model would permit to represent some sub-mesoscale eddies in a QG model.

Part II

Numerical applications

Chapter 5

Numerical modeling of uncertainty

Abstract

In order to perform a numerical simulation of the LU models, the uncertainty field has to be *a priori* modeled. We explore in particular methods based on spectral decompositions (2.9) defined from the eigenfunction basis of the spatial covariance. In practice, we work with a finite set of eigenfunctions of the small-scale Eulerian velocity fluctuations rather than with the Lagrangian displacements. This chapter presents some numerical methods to estimate the empirical orthogonal functions (EOF). We first describe the data-driven approaches and then detail some scale-aware parametrization methods. The work presented in this chapter have been partly published in Bauer et al. (2020a,b); Resseguier et al. (2020a).

5.1 Data-driven approaches

Data-driven approaches are presented in this section to estimate empirical basis functions. The first method is based on the so-called proper orthogonal decomposition (POD) method where the covariance is assumed to be quasi-stationary. We propose in Section 5.1.2 a second approach which introduces time-dependent weight coefficients into the spectral decomposition.

5.1.1 Off-line learning of EOF from high-resolution velocity

Let us consider a set of velocity snapshots $\{\boldsymbol{v}_o(\boldsymbol{x}, t_i)\}_{i=1,\dots,N_t}$ that have been *a priori* coarse-grained from high-dimensional data using a low-pass filter, such as the sharp spectral filter or Gaussian filter (Pope, 2000).

We describes briefly the snapshot POD method (Sirovich, 1987). Let $\boldsymbol{v}'_{\mathrm{o}} = \boldsymbol{v}_{\mathrm{o}} - \overline{\boldsymbol{v}}^t_{\mathrm{o}}$ be the fluctuation snapshots with the overbar denoting temporal average. The corresponding temporal covariance tensor is defined as $\boldsymbol{C} = (c_{ij})_{i,j=1,\dots,N_t}$ such that

$$c_{ij} = \frac{1}{N_t}\langle \boldsymbol{v}'_{\mathrm{o}}(\cdot,t_i), \boldsymbol{v}'_{\mathrm{o}}(\cdot,t_j)\rangle_{\Omega} \triangleq \frac{1}{N_t}\int_{\Omega} \boldsymbol{v}'_{\mathrm{o}}(\boldsymbol{x},t_i)\cdot\boldsymbol{v}'_{\mathrm{o}}(\boldsymbol{x},t_j)\mathrm{d}\boldsymbol{x}. \tag{5.1a}$$

The eigenvalues and their associated eigenfunctions can be estimated from the following eigen problem:

$$\boldsymbol{C}\boldsymbol{B} = \boldsymbol{\Lambda}\boldsymbol{B}, \tag{5.1b}$$

where $\boldsymbol{\Lambda} = (\lambda_i)_{i=1,\dots,N_t}$ is the set of decaying eigenvalues, i.e. $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{N_t} \geq 0$, and $\boldsymbol{B} = (b_{ij})_{i,j=1,\dots,N_t}$, $b_{ij} = b_i(t_j)$ is a complete set of orthogonal eigenvectors. The temporal modes $\{b_i\}_{i=1,\dots,N_t}$ are then normalized such that

$$\overline{b_i(t)b_j(t)}^t = \lambda_i\delta_{ij}, \tag{5.1c}$$

While the spatial modes $\{\boldsymbol{\phi}_i\}_{i=1,\dots,N_t}$ given by

$$\boldsymbol{\phi}_i(\boldsymbol{x}) = \overline{b_i(t)\boldsymbol{v}'_{\mathrm{o}}(\boldsymbol{x},t)}^t, \tag{5.1d}$$

are orthonormal:

$$\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j\rangle_{\Omega} = \delta_{ij}. \tag{5.1e}$$

From this spectral decomposition, each snapshot can be reconstructed by

$$\boldsymbol{v}(\boldsymbol{x},t_j) = \overline{\boldsymbol{v}}^t_{\mathrm{o}}(\boldsymbol{x}) + \sum_{i=1}^{N_t} b_i(t_j)\boldsymbol{\phi}_i(\boldsymbol{x}). \tag{5.1f}$$

In addition, we suppose that such a set of empirical eigenfunctions has a complete (or direct) decomposition (Mémin, 2014; Resseguier et al., 2017d) such that the fluctuations $\boldsymbol{v}'$ of the large-scale current lives in a subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=1,\dots,M_0-1}$, and the small-scale random drift $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\Delta t$ with a sufficiently small time step Δt lives in the residual subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=M_0,\dots,M_1}$ with $M_0 < M_1 \leq N_t$ such that

$$\frac{1}{\Delta t}\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t \approx \sum_{m=M_0}^{M_1}\sqrt{\lambda_m}\boldsymbol{\phi}_m(\boldsymbol{x})\xi_m, \tag{5.2a}$$

where ξ_m are i.i.d. standard Gaussian variables. The corresponding variance tensor is then

given by

$$\frac{1}{\Delta t}\boldsymbol{a}(\boldsymbol{x}) \approx \sum_{m=M_0}^{M_1} \lambda_m \boldsymbol{\phi}_m(\boldsymbol{x})\boldsymbol{\phi}_m^T(\boldsymbol{x}). \tag{5.2b}$$

We remark that the resulting eigenfunctions $\boldsymbol{\phi}_m$ in (5.2) are optimal in terms of the kinetic energy of the data (Holmes et al., 1996). Hereafter, such a POD approach depends only on two parameters: M_0 and M_1 pointing to the first and last modes of the noise. The choice of these parameters depends on the energy ratio γ_0, respectively γ_1, with $0 < \gamma_0 < \gamma_1 < 1$, that needs to be captured by the largest, respectively the smallest, spatial scales of the random flow component. More precisely, let us first introduce the so-called relative information content (RIC) of the eigen decomposition:

$$\mathrm{RIC}(m) = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{N_t} \lambda_i}, \qquad m = 1, \ldots, N_t. \tag{5.3a}$$

Suppose that the largest structure of the random flow is required to contain the ratio γ_0 of the total energy of the fluctuations, the first truncated mode is then determined by

$$M_0 = \min\{m \mid \mathrm{RIC}(m) \geq \gamma_0\}, \tag{5.3b}$$

as shown in Figure 5.1. Analogously, the last truncated mode M_1 can be found with a given ratio γ_1 for the smallest structure. In practice, this latter can be defined empirically.

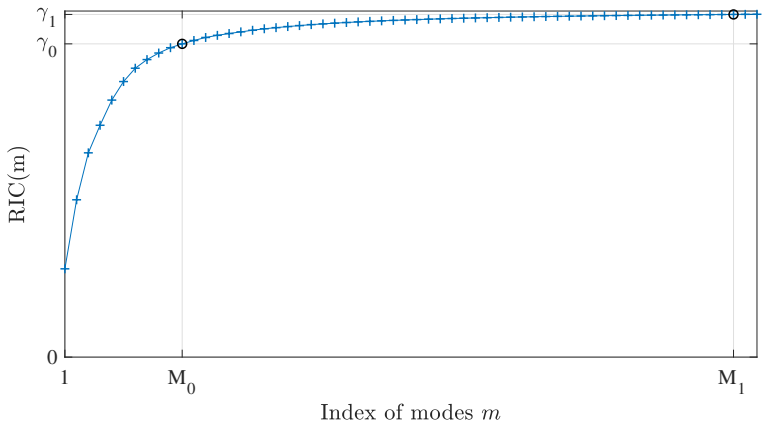

Figure 5.1 – Illustration of the spatial modes truncation for the random velocity, within the spectrum of the corresponding eigenvalues.

Now, the problem boils down to choose adequately the ratio γ_0. One possible solution is to estimate it by comparing the kinetic energy spectrums, between the ensemble of ob-

servation data $\{\boldsymbol{v}_{\mathrm{o}}(\boldsymbol{x}, t_i)\}_{i=1,\dots,N_t}$ and an extra collection of snapshots $\{\boldsymbol{v}_{\mathrm{LR}}(\boldsymbol{x}, t_i)\}_{i=1,\dots,N_t}$, obtained from a simulation of the coarse-resolution benchmark model Δ_{LR}. The parameter γ_0 is approximated by the proportion of the partial energy, accumulated up to the first wavenumber κ_0 for which the two temporally averaged spectrums start to deviate (c.f. Figure 5.2):

$$\gamma_0 \approx \frac{\sum_{\kappa \leq \kappa_0} \overline{\hat{E}_{\mathrm{o}}}^{t}(\kappa)}{\sum_{\kappa \leq \kappa_c} \overline{\hat{E}_{\mathrm{o}}}^{t}(\kappa)}, \tag{5.3c}$$

where $\hat{E}_{\mathrm{o}}$ denotes the instantaneous kinetic energy spectral density of the observations, and $\kappa_c \triangleq \pi/\Delta_{\mathrm{LR}}$ stands for the theoretical effective cutoff.

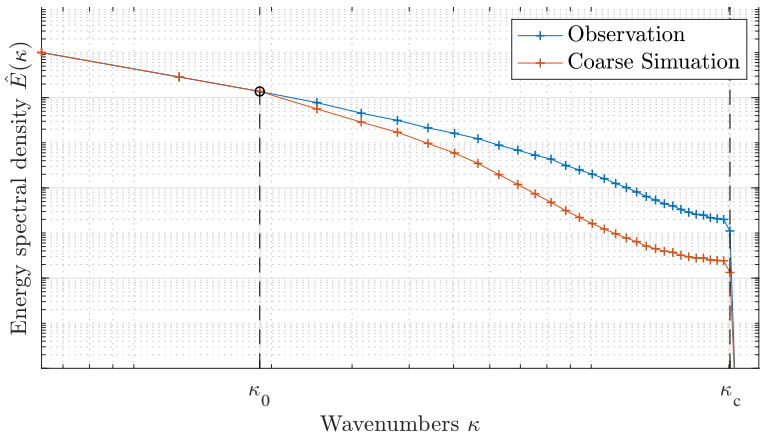

Figure 5.2 – Illustration of the time-averaged kinetic energy spectrums. The wavenumber κ_0 is searched as the first point where the observation and the deterministic coarse-simulation derivate, in order to estimate γ_0 from (5.3c).

Furthermore, Equation (5.1d) reveals that the proposed spatial modes are represented as a linear combination of the instantaneous observed velocity fields. Thus, the resulting small-scale velocity $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\Delta t$ has the same boundary and divergence conditions as the large-scale drift $\boldsymbol{v}$. In the following, we provide an important remark concerning the coarse-graining procedure that need to be considered before performing the previous POD method.

Correction drift derived from coarse-graining procedure

Let us recall from Chapter 2 that the LU model is based on a temporal–scale–separation assumption where the noise term $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ has a much smaller time scale than

the smooth component $\boldsymbol{v}$. Thus, a natural idea would consist in applying the previous POD approach to the temporal fluctuations of the high-resolution snapshots (denoted as $\boldsymbol{v}_l$):

$$\boldsymbol{v}_l' = \boldsymbol{v}_l - \overline{\boldsymbol{v}_l}^t. \tag{5.4}$$

However, this would lead an issue. Considering that $\{\boldsymbol{\phi}_l^i\}_i$ are the resulting EOFs satisfying $\langle \boldsymbol{\phi}_l^i, \boldsymbol{\phi}_l^j \rangle_\Omega = \delta_{ij}$, in order to use these EOFs for the construction of noise on the coarse grid, a coarse-graining procedure is still required, *i.e.* $\boldsymbol{\phi}_L = \mathcal{F}\boldsymbol{\phi}_l$, where $\mathcal{F}$ denotes for a grid-scale filtering operation. As a consequence, we lost the orthogonality of EOFs, *i.e.* $\langle \boldsymbol{\phi}_L^i, \boldsymbol{\phi}_L^j \rangle_\Omega \neq \delta_{ij}$.

An alternative way consists in estimating the basis functions directly on the coarse grid and is described by the following procedure:

$$\begin{aligned}
\boldsymbol{v}_l' &= \boldsymbol{v}_l - \overline{\boldsymbol{v}_l}^t \\
&= \underbrace{\mathcal{F}\boldsymbol{v}_l}_{\boldsymbol{v}_L} + (1-\mathcal{F})\boldsymbol{v}_l - \overline{\mathcal{F}\boldsymbol{v}_l + (1-\mathcal{F})\boldsymbol{v}_l}^t \\
&= \left(\boldsymbol{v}_L - \overline{\boldsymbol{v}_L}^t\right) + (1-\mathcal{F})\boldsymbol{v}_l - \overline{(1-\mathcal{F})\boldsymbol{v}_l}^t,
\end{aligned} \tag{5.5}$$

or,

$$\boldsymbol{v}_L' = \mathcal{F}\boldsymbol{v}_l' = \left(\boldsymbol{v}_L - \overline{\boldsymbol{v}_L}^t\right) + \mathcal{F}\Big((1-\mathcal{F})\boldsymbol{v}_l\Big) - \mathcal{F}\Big(\overline{(1-\mathcal{F})\boldsymbol{v}_l}^t\Big), \tag{5.6}$$

where the first two terms on the RHS depends on space-time, yet the last term depends only on space. Let us highlight that if $\mathcal{F}$ is not a projector (as in the case of Gaussian filtering), *i.e.* $\mathcal{F}^2 \neq \mathcal{F}$, then the last two terms are not null in theory. In practice, one may consider to apply the POD procedure on the subsampled versions of $\boldsymbol{v}_L - \overline{\boldsymbol{v}_L}^t$ to get a set of EOFs defined on the coarse grid, and to keep the subsampled version of $-\mathcal{F}\overline{(1-\mathcal{F})\boldsymbol{v}_l}^t$ as a correction of bias. In Chapter 8, this correction term is included in the numerical simulation of the wind-driven double-gyre circulation. The results show that this term plays a dominant role in reproducing the meandering jet for coarse resolution simulations.

5.1.2 Updated EOF by matching temporal modes

The previous POD procedure is an efficient off-line learning method, yet it relies on a strong stationary assumption, and thus leads to a sequence of random velocity fields with no temporal connection with the resolved dynamics. In the following, we propose

a novel approach that introduces a time-dependent weight coefficient $\alpha_m(t)$ in the POD representation. In this approach, the instantaneous random velocity at each time t is now defined as

$$\frac{1}{\Delta t}\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t \approx \sum_{m=M_0}^{M_1} \sqrt{\lambda_m}\boldsymbol{\phi}_m(\boldsymbol{x})\sqrt{\alpha_m}(t)\xi_m, \tag{5.7a}$$

with the corresponding variance tensor given by

$$\frac{1}{\Delta t}\boldsymbol{a}(\boldsymbol{x},t) \approx \sum_{m=M_0}^{M_1} \lambda_m\boldsymbol{\phi}_m(\boldsymbol{x})\boldsymbol{\phi}_m^T(\boldsymbol{x})\alpha_m(t). \tag{5.7b}$$

Indeed, such a weighting provides an energy re-distribution of the spatial modes at each time step. The weighting principle proposed here consists in selecting from the reference data the set of time instances that match to the large-scale structures of the current simulation. To be more specific, let us consider a current velocity field $\boldsymbol{v}_c(\boldsymbol{x},t)$ at a given time t of the current simulation. The projection coefficient b_1^c of the current fluctuation $\boldsymbol{v}_c'$ on the first spatial mode $\boldsymbol{\phi}_1$ is defined by

$$b_1^c(t) = \langle \boldsymbol{v}_c'(\cdot,t), \boldsymbol{\phi}_1\rangle_\Omega, \tag{5.8a}$$

where the fluctuation $\boldsymbol{v}_c'$ at one position are obtained by subtracting a local average of the current field around that position, and where $\langle \boldsymbol{u},\boldsymbol{v}\rangle_\Omega \triangleq \int_\Omega \boldsymbol{u}\cdot\boldsymbol{v}\mathrm{d}\boldsymbol{x}$ denotes the $L^2(\Omega)$-inner product. As illustrated in Figure 5.3, a collection of matching instants is constructed by identifying the current projection b_1^c to the time series of the first temporal mode $\{b_1(s)\}_{s=1,\dots,N_t}$ subject to a consistent condition of its time increments:

$$S(t) = \Big\{ s \;\Big|\; |b_1(s) - b_1^c(t)| \le \epsilon;\ \Delta_s[b_1]\Delta_t[b_1^c] \ge 0 \Big\}, \tag{5.8b}$$

where ϵ is a sufficiently small threshold and $\Delta_t[b] \triangleq b(t) - b(t-\Delta t)$ stands for the temporal variation of b at time t. This aims at selecting the events corresponding to the same projection coefficient and the same sign of the time increment. The weight coefficient α_m for each mode $m = M_0,\dots,M_1$ is then fixed from the sample variance:

$$\alpha_m(t) = \frac{1}{|S(t)|-1}\sum_{s\in S(t)}(b_m(s)-\mu_m(t))^2,\ \ \mu_m(t) = \frac{1}{|S(t)|}\sum_{s\in S(t)} b_m(s), \tag{5.8c}$$

where $|S|$ stands for the sample size. These time dependent coefficients allows us to slave a set of modes on some dominant modes. Note that we present here only with the first

mode. However, this technique could be extended to a vector of dominant modes in order to select more complex turbulent events. Let us also outline that the boundary and divergence conditions of the random flow (5.7a) remain the same in this weighting method.

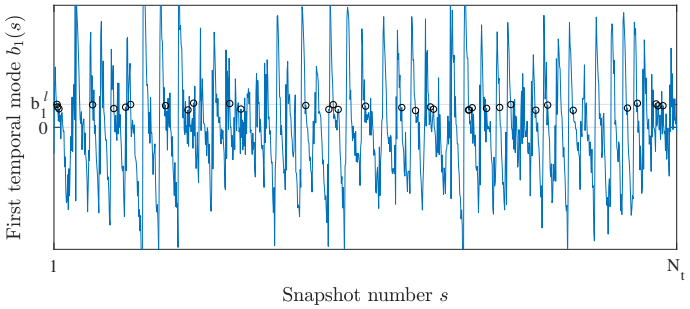

Figure 5.3 – Illustration of mode matching principle: Selection of a sample set of time based on (5.8b) corresponding to potential events matching the large-scale configuration of the current simulation.

5.2 Parameterization methods

We present in this section some noises that are not anymore defined from the data but instead parameterized. To this end, we describe two different kinds of spatial structure for the noise – homogeneous and heterogeneous. The first one is easy-to-implement. It is in particular adapted to turbulent flows with homogeneous small-scale features. On the other hand, the heterogeneous noise has more physical meaning for large-scale atmospheric and oceanic circulations.

5.2.1 The homogeneous stationary model

From the general definitions (2.2) and (2.8), a homogeneous noise means that its correlation operator $\boldsymbol{\sigma}$ is a convolution operator. In addition, to ensure that the homogeneous is incompressible, another convolution by a divergence-free projector, $\mathbf{P} \triangleq \mathbf{I}_d - (\nabla^2)^{-1}\boldsymbol{\nabla}\boldsymbol{\nabla}^T$ ($d = 2$ or 3), is required:

$$\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \Big(\mathbf{P} \star \widetilde{\boldsymbol{\sigma}} \star \mathrm{d}\boldsymbol{B}_t\Big)(\boldsymbol{x}), \tag{5.9}$$

where $\widetilde{\boldsymbol{\sigma}}$ is a convolution kernel of the small-scale flows and the symbol $\star$ denotes the convolution operation. In a simplified case with $d = 2$, Resseguier et al. (2017b) proposed an isotropic model defined through a random stream function:

$$\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}_{\mathcal{H}}^{\perp}\Big(\breve{\varphi} \star \mathrm{d}B_t\Big)(\boldsymbol{x}), \tag{5.10}$$

where $\boldsymbol{\nabla}_{\mathcal{H}}^{\perp} = [-\partial_y, \partial_x]^T$ denotes the perpendicular gradient and $\breve{\varphi} \star \mathrm{d}B_t$ stands for the random stream function with its convolution kernel $\breve{\varphi}$. As shown in Mémin (2014), both isotropy and incompressibility of the noise (5.10) result in a (constant) diagonal variance tensor $a_0\mathbf{I}_2$ with the eddy-viscosity-like coefficient a_0 and the two-dimensional identity matrix $\mathbf{I}_2$. As discussed at the end of Section 4.1.3, for the RSW-LU system (4.7) under geostrophic noise, $\boldsymbol{f} \times \boldsymbol{\sigma}_{\mathcal{H}}\mathrm{d}\boldsymbol{B}_t \approx -\boldsymbol{\nabla}_{\mathcal{H}}\mathrm{d}p_t^{\sigma}$, one can identify, for a constant Coriolis parameter f_0, the random pressure $\mathrm{d}p_t^{\sigma}$ with the proposed random stream function by $\mathrm{d}p_t^{\sigma} = \frac{1}{f_0}\breve{\varphi} \star \mathrm{d}B_t$. In practice, the convolution kernel $\breve{\varphi}$ is specified by three parameters: a fixed omni-directional spectrum slope s, a band-pass filter f_{BP} with support in the range of two wavenumbers κ_m and κ_M, and an eddy-viscosity-like coefficient a_0. In fact, the Fourier transform of the random stream function $\breve{\varphi} \star \mathrm{d}B_t$ can be defined as:

$$\widehat{\breve{\varphi} \star \mathrm{d}B_t}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}} f_{\mathrm{BP}}\left(\|\boldsymbol{k}\|\right) \|\boldsymbol{k}\|^{-\alpha}\, \widehat{\xi}_t(\boldsymbol{k}) \text{ with } \alpha = (3+s)/2, \tag{5.11}$$

where $\widehat{\cdot}$ denotes the Fourier transform coefficient, ξ_t is a space-time white noise, and A is a constant to ensure $\mathbb{E}\left\|\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right\|^2 = 2a_0\Delta t$ with Δt the size of one time stepping. In the simulations, the maximal wavenumber κ_M of the noise can usually be chosen as the effective resolution cutoff, the minimal wavenumber can be set to $\kappa_m = \kappa_M/2$, and the theoretical spectrum slope of two-dimensional (resp. three-dimensional) flow is given by $s = -3$ (resp. $s = -5/3$). The noise strength parameter a_0 is a free parameter that be tuned.

5.2.2 On-line learning of EOF from resolved velocity

The data-driven estimation of EOFs (Section 5.1.1) previously described constitute quite efficient procedures. However, fine-scale observation data, coming either from direct measurements or from high-dimensional simulations, are not always available. Therefore, Bauer et al. (2020a) proposed an alternative approach in which some local fluctuations, called *pseudo-observations* (PSO), are generated directly from a coarse-grid simulation.

Then, the singular value decomposition (SVD) is applied on those PSO to estimate a set of time-varying EOFs; the noise associated with its variance tensor is then built in the same way as in (5.2). Finally, the magnitude of the noise and variance should be scaled down to smaller scales based on a similarity analysis (Kadri Harouna and Mémin, 2017). In the following, we describe in more details both the generation of PSO and the scaling technique.

Generation of pseudo-observations

Without loss of generality, we present here the generation of PSOs in a three-dimensional case. Let us consider a $n_x \times n_y \times n_z$ low-resolution stochastic simulation. The main idea consists in building some PSOs by sliding a local window containing n_w points (with n_w be an odd integer) in each direction (x, y, z) over the spatial grid for each velocity component. More precisely, for each time step t and at every grid point $\boldsymbol{x}_{i,j,k}$, we first list the n_w^3 values of each velocity component contained in the window centered at that point:

$$I(\boldsymbol{x}_{i,j,k}, t) \triangleq \left\{ \boldsymbol{v}(\boldsymbol{x}_{p,q,r}, t) \,\middle|\, |p-i| \leq \frac{n_w-1}{2}, |q-j| \leq \frac{n_w-1}{2}, |r-k| \leq \frac{n_w-1}{2} \right\}. \quad (5.12)$$

Note that appropriate boundary conditions (replication, periodicity, *etc.*) are adopted when looking at a point on the border. Then, for one point $\boldsymbol{x}_{i,j,k}$, we pick randomly (following a discrete uniform law) n_o values (with $n_o \geq n_w^3$) in the set $I(\boldsymbol{x}_{i,j,k}, t)$. These n_o values are the so-called PSOs in the local window centered at the point $\boldsymbol{x}_{i,j,k}$. Repeating subsequently this procedure for all the grid points (*i.e.* $\forall\, i = 1, \ldots, n_x$, $j = 1, \ldots, n_y$, $k = 1, \ldots, n_z$), a global PSO matrix of size $3n_p \times n_o$ (with $n_p \triangleq n_x n_y n_z$ be the total number of grid points) can be construct:

$$\boldsymbol{V}_L \triangleq \left[\begin{array}{ccc} u_1^1 & \cdots & u_1^{n_o} \\ \vdots & \cdots & \vdots \\ u_{n_p}^1 & \cdots & u_{n_p}^{n_o} \\ \hline v_1^1 & \cdots & v_1^{n_o} \\ \vdots & \cdots & \vdots \\ v_{n_p}^1 & \cdots & v_{n_p}^{n_o} \\ \hline w_1^1 & \cdots & w_1^{n_o} \\ \vdots & \cdots & \vdots \\ w_{n_p}^1 & \cdots & w_{n_p}^{n_o} \end{array}\right], \quad (5.13)$$

where u_j^i denotes the i-th observation ($i = 1, \ldots, n_o$) of the u-component at the j-th point ($j = 1, \ldots, n_p$). Thus, each row of $\boldsymbol{V}_L$ contains an ensemble of PSOs at a given point and each column of $\boldsymbol{V}_L$ can be seen as one realization of the vector field. The subscript L denotes the fact that the PSOs are obtained at the window scale L which is coarser than the simulation scale l with $L = n_w\, l$. Further down we will introduce a scaling to "transfer" the observations from L to l.

Spectral representation of observation-scaled fluctuations

Let us now subtract the row-averages of $\boldsymbol{V}_L$ to get the fluctuations, denoted as $\boldsymbol{V}'_L$, within each window:

$$\boldsymbol{V}'_L = \boldsymbol{V}_L - \langle \boldsymbol{V}_L \rangle, \tag{5.14}$$

where $\langle \cdot \rangle$ stands for the ensemble mean over the n_o PSOs. Applying the SVD on these fluctuations $\boldsymbol{V}'_L$, we then obtain a $3n_p \times n_o$ matrix $\boldsymbol{\Phi}$ of the left singular vectors (where the columns of $\boldsymbol{\Phi}$ are mutually orthonormal) and a $n_o \times n_o$ diagonal matrix $\boldsymbol{\Sigma}$ of the singular values such that $\boldsymbol{V}'_L = \boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Psi}^T$ (with $\boldsymbol{\Psi}$ the right singular vectors matrix). As such, the two-points covariances of the fluctuation velocity can be estimated by the following $3n_p \times 3n_p$ matrix:

$$\boldsymbol{C}_L \triangleq \frac{1}{n_o - 1} \boldsymbol{V}'_L (\boldsymbol{V}'_L)^T = \boldsymbol{\Phi}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\Sigma}}^T\boldsymbol{\Phi}^T \tag{5.15}$$

$$\approx \mathbb{E}\Big[\Big(\boldsymbol{v}_L - \mathbb{E}[\boldsymbol{v}_L]\Big)(\boldsymbol{x}, t)\, \Big(\boldsymbol{v}_L - \mathbb{E}[\boldsymbol{v}_L]\Big)^T(\boldsymbol{y}, t)\Big], \tag{5.16}$$

where $\tilde{\boldsymbol{\Sigma}} \triangleq \boldsymbol{\Sigma}/\sqrt{n_o - 1}$ and the matrix $\boldsymbol{C}_L$ is of rank $n_o - 1$ due to the subtraction (5.14). Let us note that (5.14) is based on the assumption that ensemble average can be approximated by local space average. With that assumption, the columns of $\boldsymbol{\Phi}$ are also the eigenvectors of $\boldsymbol{C}_L$ associated to the eigenvalues $\boldsymbol{\Lambda} \triangleq \mathrm{diag}(\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\Sigma}}^T)$. Let $\boldsymbol{\phi}_i$ be the i-th column of $\boldsymbol{\Phi}$ and λ_i be the i-th coefficient of $\boldsymbol{\Lambda}$, for $i = 1, \ldots, n_o - 1$. The random velocity fluctuations and their one-point covariances can be written as

$$\boldsymbol{v}'_L(\boldsymbol{x}) = \sum_{i=1}^{n_o-1} \sqrt{\lambda_i}\boldsymbol{\phi}_i(\boldsymbol{x})\xi_i, \qquad \boldsymbol{c}_L(\boldsymbol{x}) = \sum_{i=1}^{n_o-1} \lambda_i \boldsymbol{\phi}_i(\boldsymbol{x})\boldsymbol{\phi}_i^T(\boldsymbol{x}). \tag{5.17}$$

Let us outline that these fluctuations correspond intrisically to a virtual observation at scale L and must be scaled down to the simulation scale $l = L/n_w$.

Scale-aware modulation of the amplitude

Kadri Harouna and Mémin (2017) proposed a self-adapted scaling method using large-scale information at each time step. This scaling reflects the fact that variances of independent random variables are observed at some large scales L, can be injected at a smaller scale l up to the factor $(l/L)^{2/3}$. In our case, this factor reduces hence to $n^{-2/3}$. We remember that the variance tensor $\boldsymbol{a}$ of the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ corresponds to the variance of fluctuations velocity up to a decorralaton time τ_l, namely

$$\boldsymbol{a} = \tau_l \, n^{-2/3} \boldsymbol{c}_L. \tag{5.18}$$

Now all we need is the small-scale characteristic time τ_l. In order to define this characteristic time, let us first define a (small) velocity scale U_l by

$$U_l \triangleq n^{-1/3} \, \|\mathrm{tr}^{1/2}(\boldsymbol{c}_L)(\boldsymbol{x})\|, \tag{5.19}$$

where the norm could be $\|\cdot\|_2$ or $\|\cdot\|_\infty$. Together with the numerical time-step Δt given by classical Courant-Friedrichs-Lewy (CFL) condition, this gives a numerical small-scale length, $\delta_l = U_l \Delta t$. Applying again the turbulence-power-law (Kadri Harouna and Mémin, 2017), the characteristic time can be estimated by

$$\tau_l = \Big(\frac{l}{\delta_l}\Big)^{2/3} \Delta t = \Big(\frac{l}{U_l}\Big)^{2/3} (\Delta t)^{1/3}. \tag{5.20}$$

Finally, one realization of the noise corresponding to the variance (5.18) is given by

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \sqrt{\tau_l} \, n^{-1/3} \boldsymbol{v}'_L. \tag{5.21}$$

5.2.3 Random forcing derived from subgrid scales dissipations

In this section, we propose a new energy-budget-based stochastic subgrid model. In practice, large-scale fluid dynamics models involve dissipation operators called subgrid scales (SGS) models. They stabilize the numerical simulations and mimic the action of the unresolved small scales by draining the energy at high wavenumbers. Simple SGS models often take the form of classical Laplacian operator or of higher-order hyperviscosity operators (typically, some power of a Laplacian). Let us consider the dynamics of a fluid property q, which is assumed to be transported up to a dissipation operator $\mathcal{L}$ and a

stochastic forcing η:

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \mathcal{L}[q] + \eta, \tag{5.22}$$

where $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{v}\cdot\boldsymbol{\nabla}$ stands for the material derivative operator and the random forcing η is a centered process uncorrelated in time.

Similarly as shown in Section 2.3, through Itô calculus, a conservation of the ensemble mean of energy, $\mathbb{E}[\int_\Omega \frac{1}{2}q^2\,\mathrm{d}\boldsymbol{x}]$, would imply:

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Big[\int_\Omega \frac{1}{2}q^2\,\mathrm{d}\boldsymbol{x}\Big] = \mathbb{E}\int_\Omega \Big(q\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\langle q,q\rangle_t\Big)\,\mathrm{d}\boldsymbol{x} = \mathbb{E}\int_\Omega \Big(q\mathcal{L}[q] + \frac{\mathrm{d}t}{2}(\mathcal{H}[q])^2\Big)\,\mathrm{d}\boldsymbol{x}, \tag{5.23}$$

in which

$$(\mathcal{H}[q])^2(\boldsymbol{x},t) \triangleq \mathbb{E}\Big[\eta^2(\boldsymbol{x},t)\,|\,q(\cdot,t)\Big] \tag{5.24}$$

denotes the variance of random forcing η conditioned on field q at time t and $\langle q,q\rangle_t$ stands for the quadratic variation of q. In order to maintain a desired amount of energy dissipation, we introduce a scaling factor $\zeta > 0$ in the previous balance, namely

$$\int_\Omega q\Big(-\frac{2\zeta}{\mathrm{d}t}\mathcal{L}[q]\Big)\,\mathrm{d}\boldsymbol{x} = \int_\Omega (\mathcal{H}[q])^2\,\mathrm{d}\boldsymbol{x}. \tag{5.25}$$

Now, for a given dissipation operator $\mathcal{L}$, we aim at building a noise η which satisfies the desired balance (5.25). If $\mathcal{L}$ is a negative auto-adjoint operator (which is generally the case), we can define an operator H such that

$$-\frac{2\zeta}{\mathrm{d}t}\mathcal{L} = \mathrm{H}\mathrm{H}^\star, \tag{5.26}$$

where $\mathrm{H}^\star$ is the adjoint of operator H. More precisely, for the Laplacian, biharmonic or higher-order hyperviscosity operators, this operator H is defined respectively as

$$\mathrm{H}[q] \triangleq \begin{cases} \tilde{\zeta}\boldsymbol{\alpha}\boldsymbol{\nabla}q \text{ if } \mathcal{L}[q] = \boldsymbol{\nabla}\cdot(\boldsymbol{\alpha}\boldsymbol{\alpha}^T\boldsymbol{\nabla}q), \\ \tilde{\zeta}\,\boldsymbol{\nabla}\cdot(\boldsymbol{\alpha}\boldsymbol{\nabla}q) \text{ if } \mathcal{L}[q] = -\,\boldsymbol{\nabla}\cdot\Big(\boldsymbol{\alpha}\boldsymbol{\nabla}\Big(\boldsymbol{\nabla}\cdot(\boldsymbol{\alpha}\boldsymbol{\nabla}q)\Big)\Big), \\ \tilde{\zeta}\alpha\boldsymbol{\nabla}(\nabla^2)^p q \text{ if } \mathcal{L}[q] = \alpha^2(\nabla^2)^{2p+1}q, \\ \tilde{\zeta}\alpha(\nabla^2)^p q \text{ if } \mathcal{L}[q] = -\alpha^2(\nabla^2)^{2p}q, \end{cases} \tag{5.27}$$

where $\tilde{\zeta} = \sqrt{2\zeta/\mathrm{d}t}$ and p is a positive integer, α is a constant and $\boldsymbol{\alpha}$ stands for a matrix which depends possibly on space.

In particular, for a specific deterministic subgrid tensor $\mathcal{L}$, if the conditional variance

integral of random forcing, $\int_\Omega (\mathcal{H}[q])^2 \,\mathrm{d}\boldsymbol{x}$, can be set to the integral $\int_\Omega \|\mathrm{H}[q]\|^2 \,\mathrm{d}\boldsymbol{x}$, it is possible then to control the energy dissipation satisfying the desired balance (5.25). Indeed, in that case, we have:

$$\int_\Omega (\mathcal{H}[q])^2 \,\mathrm{d}\boldsymbol{x} = \int_\Omega \|\mathrm{H}[q]\|^2 \,\mathrm{d}\boldsymbol{x} = \int_\Omega q(\mathrm{H}^\star\mathrm{H})[q] \,\mathrm{d}\boldsymbol{x} = \int_\Omega q\big(-\frac{2\zeta}{\mathrm{d}t}\mathcal{L}[q]\big) \,\mathrm{d}\boldsymbol{x}. \tag{5.28}$$

Nevertheless, to simulate the random forcing, knowing its global variance is not enough. We also need to model its spatial structure (*e.g.* the local variation of the variance, its spatial correlations). In this purpose, we express the noise on a convenient orthonormal basis of $L^2(\Omega)$ denoted $\{e_k\}_{k\geq 0}$:

$$\eta(\boldsymbol{x},t) = \sum_{k\geq 0} \lambda_k(t) e_k(\boldsymbol{x}) \xi_k, \qquad \xi_k \sim \mathcal{N}(0,1), \tag{5.29}$$

where the eigenvalues are updated in time by

$$\lambda_k(t) = \int_\Omega \|\mathrm{H}[q]\|(\boldsymbol{x},t) e_k(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \tag{5.30}$$

Thus, by definition of $\mathcal{H}$ and by the Parseval theorem, the conditional variance integral of random forcing is given by

$$\int_\Omega (\mathcal{H}[q])^2 \,\mathrm{d}\boldsymbol{x} = \int_\Omega \mathbb{E}\big[\eta^2(\boldsymbol{x},t) \,|\, q(\cdot,t)\big] \,\mathrm{d}\boldsymbol{x} = \sum_{k\geq 0} |\lambda_k|^2 = \int_\Omega \|\mathrm{H}[q]\|^2 \,\mathrm{d}\boldsymbol{x}. \tag{5.31}$$

Therefore, Equation (5.28) is valid and the random forcing η defined by (5.29) and (5.30) satisfies the objective energy balance (5.25). We remark that like the advection noise $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot \boldsymbol{\nabla} q)$ in the LU transport equation (2.11), the random forcing η defined by (5.29) is also multiplicative and non-Gaussian, as their eigenvalues in (5.30) depends on each realizations of q.

As a consequence, given a numerical dissipation $\mathcal{L}$ together with a chosen orthonormal basis $\{e_k\}_{k\geq 0}$, it is always possible to define a noise (5.29) that respects the assumed energy balance (5.25). However, Equation (5.25) constitutes only a global balance. Indeed, locally in space, the variance of random forcing is

$$(\mathcal{H}[q])^2 = \sum_{k\geq 0} |\lambda_k|^2 |e_k|^2 \neq \Big(\sum_{k\geq 0} \lambda_k e_k\Big)^2 = \|\mathrm{H}[q]\|^2. \tag{5.32}$$

Moreover, we have

$$\|\mathrm{H}[q]\|^2 \neq q\mathrm{H}^\star\mathrm{H}[q] = q\Big(-\frac{2\zeta}{\mathrm{d}t}\mathcal{L}[q]\Big). \tag{5.33}$$

The choice of the fixed basis $\{e_k\}_{k\geq 0}$ in this model is of crucial importance, as it influences in particular the variance heterogeneity and the correlation lengths of the random forcing. In order to be close to a local energy balance, basis functions with small supports are more adpated. This enables us to restrain the inequality (5.32) and makes it closer to an equality. For instance, Fourier modes have a large support and would lead to a homogeneous random forcing with a constant variance. In this case, the inequality (5.32) would be $(\mathcal{H}[q])^2 = \frac{1}{|\Omega|}\int_\Omega \|\mathrm{H}[q]\|^2 \,\mathrm{d}\boldsymbol{x} \neq \|\mathrm{H}[q]\|^2(\boldsymbol{x},t)$. In contrast a basis of regularized Dirac functions $e_k(\boldsymbol{x}) = \delta(\boldsymbol{x}-\boldsymbol{x}_k)$, defined on the grid points $\boldsymbol{x}_k$, have near zero-measure supports. The inequality (5.32) would hence become an equality. Nevertheless, such an infinitesimally small basis function support would induce an almost zero correlation length of random forcing, which makes no sense physically. According to such analysis, wavelet basis seems a promising trade-off and will be used in Chapter 6.

Let us summarize here the above method from the LU viewpoint. The goal is to define the multiplicative advection noise $\eta = -\frac{1}{\Delta t}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla} q$ in such way that the corresponding LU diffusion term $\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}q)$ corresponds (up to a scaling factor ζ) to the numerical dissipation $\mathcal{L}[q]$. In this purpose, the noise η is expressed by (5.29) with a given orthonormal basis $\{e_k\}_{k\geq 0}$. The eigenvalues λ_k in (5.29) are adapted along time according to (5.30) and (5.27). The stochastic transport of q is finally described by Equation (5.22). As such, the energy brought by the noise amounts to ζ times the one losts by the dissipation $\mathcal{L}[q]$. From the construction of the multiplicative noise, it is difficult to recover explicitly the additive velocity noise $\frac{1}{\Delta t}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and the variance tensor $\boldsymbol{a}$, hence it is hard to specify the Itô-Stokes drift $\frac{1}{2}\,\boldsymbol{\nabla}\cdot\,\boldsymbol{a}$ (included in $\boldsymbol{v}^\star$) in general. However, this term vanishes in some particular cases. For instance, this is true if the coefficient $\boldsymbol{\alpha}$ of the Laplacian/biharmonic operator in (5.27) is constant in space. We remark that, so far, we used this energy-budget-based method for a transported scalar. For future work, this method could be further explored in more general situations such as the momentum equation in the Boussinesq system (3.11a) or for the potential vorticity in the QG equation (4.51a) including sources and sinks terms.

Chapter 6

Ensemble forecasts verification for SQG dynamics

Abstract

In this chapter, we first review some important metrics to quantify ensemble forecasts skills, such as the Talagrand diagram, the continuous ranked proper score and the energy score. As a test case, a simple geophysical fluid dynamics – the stochastic surface quasi-geostrophic (SQG) model – will be considered. After presenting it, several stochastic parameterizations described in the previous chapter will be compared for short-term ensemble forecasts. We will show that the proposed random model, under both homogeneous and heterogeneous uncertainty, is more efficient than a deterministic model with a perturbation of the initial condition. This ability is in particular essential for data assimilation applications. The work presented in this chapter has been partly published in Resseguier et al. (2020a). This chapter brings more precise views on the forecast efficiency of different noise models.

6.1 Metrics for ensemble forecasts

This section is a short summary about existing verification ensemble tools. More detailed review can be found in (Hamill and Colucci, 1997; Weigel, 2012). In the following, let $\{q_1^o, \dots, q_N^o\}$ be a set of N-observations and $\{(q_1^1, \dots, q_1^M),\ \dots,\ (q_N^1, \dots, q_N^M)\}$ be an ensemble of M-members with q_n^m being the m-th ensemble member of the n-th forecast. Note that N can represent either the spatial points or times steps of the observations.

6.1.1 Talagrand diagram

A Talagrand diagram (TD, also called rank histogram) is a technique used to check the reliability of an ensemble forecast or a set of quantile. The idea consists in checking the availability of the N observations in a set of rank r. The sets of rank are built from rank statistics which is the sorted ensemble $\{(q_1^{(1)}, \dots, q_1^{(M)}), \dots, (q_N^{(1)}, \dots, q_N^{(M)})\}$ (Anderson, 1996; Hamill and Colucci, 1997):

$$\forall\,(m,n) \in \{1,\dots,M\} \times \{1,\dots,N\}, \qquad r_m = \frac{1}{N}\sum_{n=1}^{N} \widehat{\mathbb{P}}\Big(q_n^{(m-1)} \leq q_n^o < q_n^{(m)}\Big), \tag{6.1}$$

where $\widehat{\mathbb{P}}$ is the estimated probability of ranking an observation between two sorted ensemble members.

A calibrated ensemble should result in a flat histogram. However, a flat histogram does not guarantee a calibrated ensemble (Hamill, 2001). A flat histogram mostly indicates that the ensemble and observation are sampled from a common distribution. A U-shaped rank histogram corresponds to an underdispersion or to a conditional bias (Hamill, 2001) of the ensemble members. A dome-shaped histogram shows, on the contrary, an overdispersion of the ensemble. A non-symmetric histogram is the footprint of bias.

6.1.2 Mean squared error and mean ensemble variance

A necessary condition for ensemble reliability is that the mean squared error (MSE) of the ensemble mean forecast is close to the mean ensemble variance (MEV, also called mean squared dispersion), up to an ensemble size-dependent scaling factor (Stephenson and Dolas-Reyes, 2000; Weigel, 2012):

$$\underbrace{\frac{1}{N}\sum_{n=1}^{N}\big(\widehat{\mathbb{E}}[q_n] - q_n^o\big)^2}_{\text{MSE}} \approx \frac{M+1}{M}\underbrace{\Big(\frac{1}{N}\sum_{n=1}^{N}\widehat{\mathrm{Var}}[q_n]\Big)}_{\text{MEV}} \tag{6.2}$$

where $\widehat{\mathbb{E}}$ and $\widehat{\mathrm{Var}}$ are the empirical estimators of the ensemble mean and variance:

$$\widehat{\mathbb{E}}[q] \triangleq \frac{1}{M}\sum_{m=1}^{M} q^m, \qquad \widehat{\mathrm{Var}}(q) \triangleq \frac{1}{M-1}\sum_{m=1}^{M}\Big(q^m - \widehat{\mathbb{E}}[q]\Big)^2. \tag{6.3}$$

In fact, Equation (6.2) is used to diagnose if the ensemble mean error matches the ensemble spread dispersion. However, it is a necessary but not sufficient condition of ensemble reliability in the sense that any forecast system not satisfying (6.2) is unreliable, yet a

forecast system satisfying (6.2) is not necessarily reliable. Further metrics have to be used to determine the ensemble reliability.

6.1.3 Continuous ranked proper score

Let us now review some forecast proper scores. A proper score is negatively oriented, such that a lower score indicated a better ensemble forecast. A proper score converging to zero means that the observation cannot be distinguished from the ensemble members.

A typical proper score used for ensemble model evaluation is the continuous ranked proper score (CRPS). The CRPS is defined as the integrated squared difference between the cumulative forecast and observation distribution (Hersbach, 2000):

$$\mathrm{CRPS}(F, q^o) = \int_{\mathbb{R}} \Big(F(\mathcal{B}) - H(\mathcal{B} - q^o)\Big)^2 \mathrm{d}\mathcal{B}, \tag{6.4}$$

where F is the cumulative distribution function (CDF) of the point-wise random process q (represented by the ensemble) and H is the Heaviside function (equals to one if $q^o \leq \mathcal{B}$ and zero otherwise). Thus, the CRPS represents the distance between the Heaviside function and the distribution of the random variable q.

This Heaviside function represents the inequality between an ensemble member and the observation. In particular, Hersbach (2000) proposed an easy-to-implement estimator of the CRPS, namely

$$\widehat{\mathrm{CRPS}}(\hat{F}, q^o) = \sum_{m=1}^{M} \alpha^{(m)} \Big(\frac{m}{M}\Big)^2 + \sum_{m=0}^{M-1} \beta^{(m)} \Big(1 - \frac{m}{M}\Big)^2 \tag{6.5a}$$

with

$$\alpha^{(m)} = \begin{cases} 0 & \text{if } q^o \leq q^{(m)} \\ q^o - q^{(m)} & \text{if } q^{(m)} < q^o \leq q^{(m+1)} \\ q^{(m+1)} - q^{(m)} & \text{if } q^{(m+1)} < q^o \end{cases} \tag{6.5b}$$

and

$$\beta^{(m)} = \begin{cases} q^{(m+1)} - q^{(m)} & \text{if } q^o \leq q^{(m)} \\ q^{(m+1)} - q^o & \text{if } q^{(m)} < q^o \leq q^{(m+1)} \\ 0 & \text{if } q^{(m+1)} < q^o \end{cases}, \tag{6.5c}$$

where $q^{(0)} = -\infty$, $q^{(M+1)} = \infty$ and $\{q^{(1)}, \ldots, q^{(M)}\}$ is the sorted ensemble forecast.

6.1.4 Energy score

The previous CRPS is a point-wise (in space or/and time) score, hence cannot summarize the whole uncertainty of the ensemble in one value in a multivariate case. Therefore, to perform the evaluation of our ensemble in a multivariate case, the CRPS will not be enough. Although the mean CRPS can be used, it would miss some information. Instead, a generalized form of the CRPS is required. Following (Gneiting and Raftery, 2007), a possible generalization of the CRPS is the energy score (ES), which is defined as

$$\mathrm{ES}(F, q^o) = \mathbb{E}_F\|q - q^o\| - \frac{1}{2}\mathbb{E}_F\|q - q'\|, \tag{6.6}$$

where the observation q^o, the random variable q and the CDF F are both multivariate, and $\|q\| = (\sum_{n=1}^{N}(q_n)^2)^{1/2}$ is the Euclidean norm of q. In particular, an estimator of the ES is given by

$$\widehat{\mathrm{ES}}(\hat{F}, q^o) = \frac{1}{M}\sum_{m=1}^{M}\|q^m - q^o\| - \frac{1}{2M^2}\sum_{m=1}^{M}\sum_{m'=1}^{M}\|q^m - q^{m'}\|. \tag{6.7}$$

6.2 Model configurations

High-resolution (HR) deterministic SQG simulations of test flows will provide references to which we will compare random simulations performed at a lower resolution. For this purpose, the high-resolution (512^2) simulation outputs will be projected onto the space of low-resolution (128^2) fields. Each HR buoyancy field will be filtered (through a low-pass spectral filter) and subsampled to a low-resolution. We will refer to these projected fields as "*observations*" and denote them b^o. For all simulations, a standard hyper-viscosity scheme has been introduced (Held et al., 1995):

$$\mathbb{D}_t b = \alpha\nabla^8 b\,\mathrm{d}t, \tag{6.8}$$

with a positive coefficient α proportional to M_x^{-8} where M_x denotes the grid size (*i.e.* 128 or 512). The domain size is a square box $L_x \times L_y = 1000$ km $\times 1000$ km and the boundary conditions are double periodic. We recall from Equation (4.54) that the horizontal velocity $\boldsymbol{u}$ is related to the buoyancy b in Fourier space through the usual SQG relation:

$$\hat{\boldsymbol{u}} = i\boldsymbol{k}^{\perp}\frac{\hat{b}}{N\|\boldsymbol{k}\|}, \tag{6.9}$$

where the constant stratification is set to $N = 3f_0$ with the Coriolis frequency f_0 fixed to be 1.028×10^{-4} s^{-1}, which corresponds to a latitude of 45°.

Several simulations of these models have been performed based on two types of initial condition (IC) and on various parametrizations of uncertainty. A first type of initial buoyancy field, denoted as "*Vortices IC*", is shown in Figure 6.1 (a). As shown in Resseguier et al. (2017b), this field consists of a spatially smooth buoyancy field with two warm elliptical anticyclones (positive buoyancy) and two cold elliptical cyclones (negative buoyancy). Another type of initial buoyancy field, denoted as "*Spectral IC*", is shown in Figure 6.1(b). It is a homogeneous Gaussian random field generated from a spectrum of buoyancy with a prescribed slope equals to $-5/3$. This slope corresponds to the power law of a developed SQG turbulence. In both cases, the amplitude of the buoyancy is set to $B_0 = 10^{-3}$ m s^{-2}.

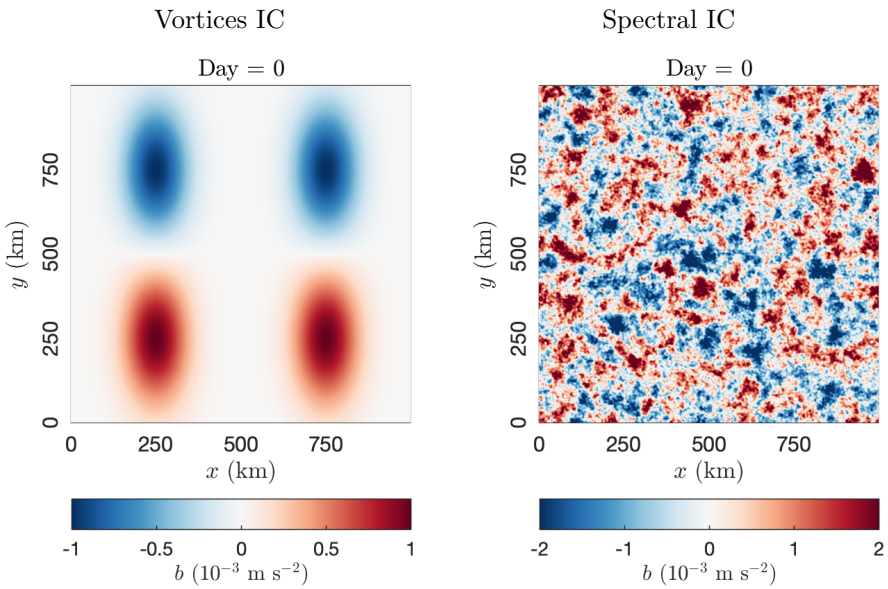

Figure 6.1 – Two initial buoyancy (m s^{-2}) fields. Left: a smooth field with four given vortices (cyclones in blue and anti-cyclones in red); Right: a homogeneous Gaussian field generated from a $-5/3$ spectrum.

As in the study Resseguier et al. (2017b), we focus on the first-month forecast of the smooth initial field (Vortices IC in Figure 6.1). Figure 6.2 shows the reference HR simulation for this first month. After two weeks, filament instabilities (Lapeyre, 2017) create developed turbulence. In this work, we also study the free-decaying SQG turbulence flow initialized by a rough field (Spectral IC in Figure 6.1). The evolution of free-decaying

turbulence can be seen in Figure 6.3. A part of the initial energy is dissipated by the deterministic subgrid tensor. However, a part of this initial energy cascades to the larger scales by creating larger vortices from the merging of small vortices.

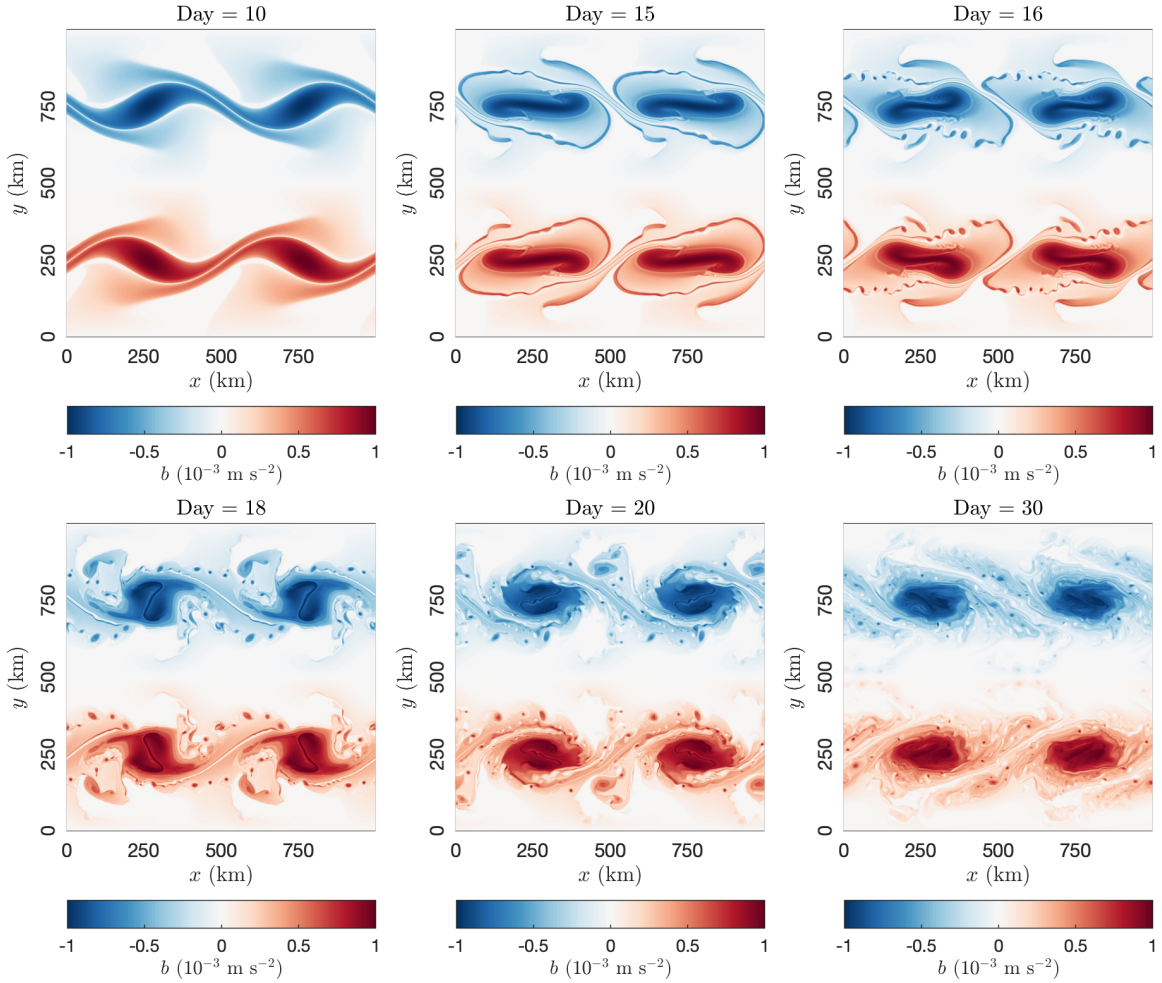

Figure 6.2 – Buoyancy (m s^{-2}) after various days of advection for the usual SQG model at resolution 512^2, based on the Vortices IC.

Several coarse resolution ensembles have been forecasted with both initial buoyancy fields of Figure 6.1. Two ensembles rely on randomly perturbed initial condition (PIC) and four others are driven by different dynamics under LU. Specifically, the small-scale velocity noises in those LU models have been generated using different approaches: a homogeneous isotropic stationary model based on Fourier basis (described in Section 5.2.1), a heterogeneous stationary model based on the POD procedure of the HR velocity (described in Section 5.1.1) and two heterogeneous non-stationary models based on the

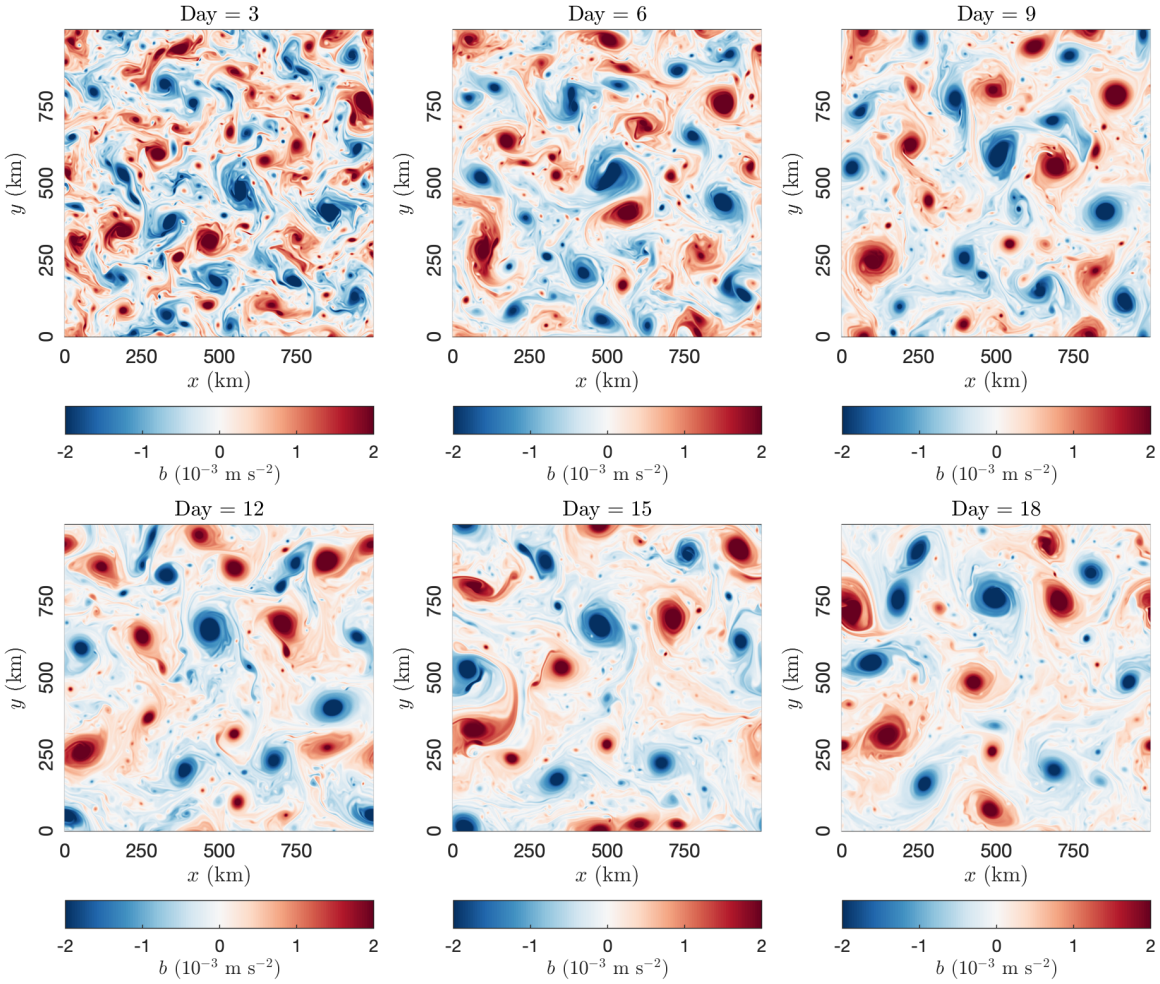

Figure 6.3 – Buoyancy (m s^{-2}) after various days of advection for the usual SQG model at resolution 512^2, based on the Spectral IC.

learning procedure of EOFs from the coarse resolution pseudo-observations (described in Section 5.2.2) and on the subgrid scale dissipation-adapted noise (described in Section 5.2.3). In the following, these four random models are denoted as "*LU-FFT*", "*LU-POD*", "*LU-PSO*" and "*LU-SGS*" respectively. For the LU-FFT model, the small-scale energy is specified by the diffusion coefficient $a_0 = 18$ m^2 s^{-1} (see Equation 5.11). In addition, the spectrum slope s is fixed to be $-5/3$ corresponding to the SQG dynamics. For more, information on the values of these parameters, one can refer to Resseguier et al. (2017b). For the LU-POD model, under Vortices IC (resp. Spectral IC), we used $N_t = 740$ (resp. 500) HR snapshots with uniform time steps of half an hour to train the spatial modes, and we kept only the modes from $M_0 = 5$ (resp. 3) to $M_1 = 105$ (resp. 103) as the EOFs of the noise (see Equation 5.2a). This choice has been operated through the strategy described in Section 5.1.1. We remark that the quasi-stationary assumption of the covariance in the snapshot POD method cannot be fully satisfied under these purely dissipated test cases. This corresponds thus to an inexact stationary assumption on the considered time frame. For the LU-PSO model, we build $n_O = 20$ pseudo-observations by sliding a 3×3 window over each grid point (see Equation 5.13). Besides, a circular boundary condition is adopted when looking at a point on the border. For the LU-SGS model, we fixed the "symmlet" wavelet basis (Daubechies, 1992) with five vanishing moments and updated the eigenvalues in time according to the buoyancy dissipation budget (5.30)–(5.27). In particular, in the Vortices IC case, we have forecasted five ensembles with five different values $(10\%, 20\%, 30\%, 40\%, 50\%)$ of the scaling factor ζ. For instance, Figure 6.4 shows us that for a given resolution, when the scaling factor ζ increases (i.e. when the noise variance counter-balances a larger part of the numerical dissipation), more and more small-scale structures are presented in the physical field. Hereafter, we set $\zeta = 50\%$ in the LU-SGS model when comparing with other random models. For the benchmark ensembles provided by the PIC models, the similar sampling methods FFT and PSO are adopted, yet the obtained small-scale random fields are used only at the initial time.

A first comparaison of the ensemble solutions provided by these random models is given by Figure 6.5 (under Vortices IC), in which we have first computed the buoyancy spectrum of each snapshot and for each ensemble member, then taken the ensemble (over 30 realizations) and the time (over the last 20 days) mean of these spectrums, *i.e.* $\overline{\mathbb{E}\left[|\widehat{b}|^2\right]}^t(\kappa)$. The results show that both LU models produce higher energy of buoyancy than the PIC model at the high wavenumbers, hence better in resolving the small-scale features. In particular, the ensemble mean spectrum of the LU-SGS scheme are much closer to that of

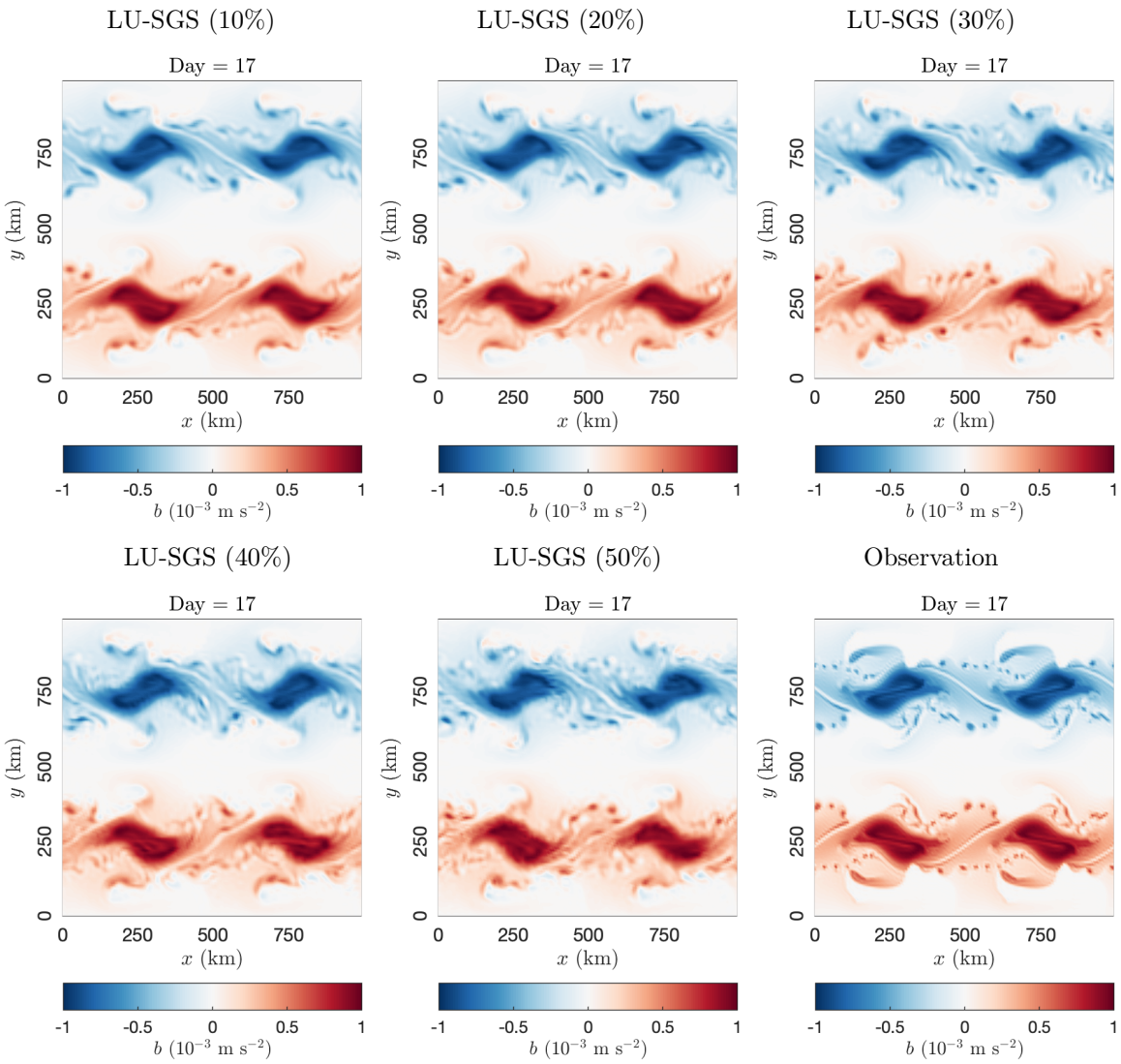

Figure 6.4 – Comparison of one realization of the buoyancy field after 17 days of advection, according to various energy scaling factor in LU-SGS. We remember that the observation is coarse-grained from the HR buoyancy field.

the reference. As expected, when the scaling factor ζ increases, the results become even better.

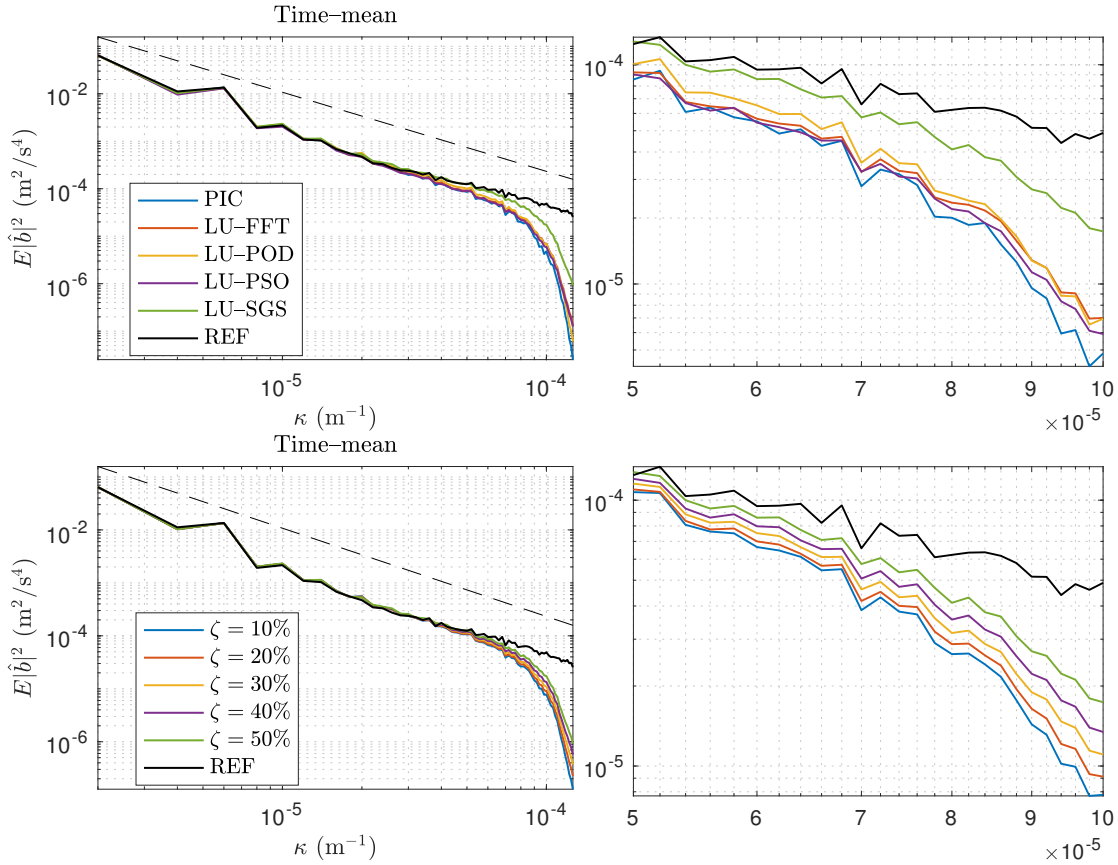

Figure 6.5 – Comparaison of the ensemble (over 30 realizations) and time (over the last 20 days) mean of the buoyancy spectrums for different random models (top) and for the LU-SGS model under various energy factors (bottom); on the right zoom of the left plots for the high wavenumbers. Note that "PIC" (top) stands here for the PIC-FFT model and the dashed lines (left) describe the -5/3 power law.

6.3 Short-term ensemble forecast

Once the ensembles have been produced by the models introduced previously, we can measure the quality of ensemble forecasts by the easy-to-implement criterions presented in Section 6.1. As described by Stephenson and Dolas-Reyes (2000), typically in meteorological applications, the number of ensemble members is of the order of 10–100, whereas the rank of the observation/forecast states is of the order of 10^3–10^7. In our case, we have

$M = 30$ ensemble members for each forecast system. In the following, unless specifically indicated, we consider that each system has $N = 128^2$ forecasts. Indeed, each instantaneous buoyancy field with 128^2 spatial points describes the phase of each system. Thus, $b^m(\boldsymbol{x}_n, t)$ denotes the m-th ensemble member of the n-th forecast at day t.

In Figures 6.6, we compare the criterion (6.2) produced by different random models. For Vortices IC, all models produce low errors during the first 10 days and high errors during Day 17 to Day 24. In this case, LU-SGS and LU-PSO show in particular a better matching between the MSE of mean forecast and the ensemble spread. Furthermore, in the LU-SGS method, when the factor ζ augments, both the spreading and the ensemble errors increase. For Spectral IC, all ensemble models tend to a stationary state much more faster. In this case, LU-SGS and LU-PSO exhibit still better matching between MSE and MEV. In both cases, heterogeneous noise models behave better than the homogeneous ones.

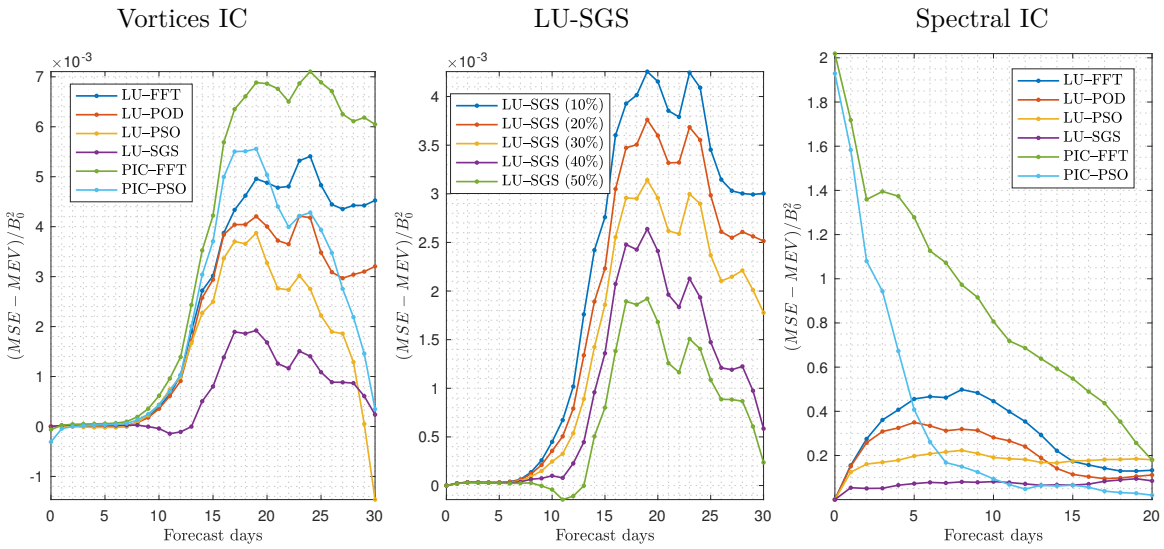

Figure 6.6 – Comparisons of ensemble forecasts for different stochastic models based on the two tested IC. Left: The difference between the MSE of mean forecast and mean ensemble variance; Middle: Using the empirical noise (5.2.3) with wavelet basis and various energy scaling factors under Vortices IC; Right: Results under Spectral IC. Note that all these results are normalized by the amplitude of the buoyancy, *i.e.* $B_0 = 10^{-3}$ m s^{-2}.

Another evaluation of the ensembles dispersion have been carried out through Talagrand diagrams. As shown in Table 6.1, for Vortices IC, the method applying an homogeneous perturbation on the initial condition (PIC-FFT) is extremely underdispersive. LU-FFT, LU-POD and PIC-PSO formed a flatter and higher U-shaped underdispersive

profile. The ensemble dispersion is still underestimated. Nevertheless, LU-PSO estimates show the flatter histograms, and LU-SGS with 50% energy allocated to the numerical dissipation provides a slight overdispersion. From Table 6.2, we observe that for Spectral IC, both LU-PSO and PIC-PSO with only initial perturbations provide almost perfect ensemble spread, but the best one is obtained for LU-SGS. In the same way as for the previous criterion, for this case, the Talagrand diagrams seems to converge more rapidly toward a stationary state.

As explained above, distinct models yield distinct spreading based criterions (6.2) and Talagrand diagrams. Nonetheless, these two criterions are only necessary but not sufficient conditions to assess the ensemble reliability. An objective analysis needs hence to rely on others metrics such as proper scores. To begin, we first evaluate the numerical results with the CRPS (see Section 6.1.3). Applied on ensemble of spatio-temporal fields, maps of CRPS can be represented at each fixed time step as in Figures 6.7 and 6.8 (for the Vortices IC case). As expected we observe on all those maps that the normalized CRPS is relatively high on small-scale structures and low on the center of each vortices. Indeed, turbulent structures are obviously more difficult to reproduce accurately. These four figures suggest that PIC-FFT is the worst model since it has the largest CRPS values. Instead, the LU-PSO and LU-SGS models provide the smallest CRPS; revealing hence that they make fewer local errors at these advection times. Figures 6.9 and 6.10 plot the CRPS maps corresponding to Spectral IC. In this case, the CRPS structures are first small and spread over the spatial domain. Then, these small-scale CRPS structures merge and create larger structures of higher intensity. The merging is due to the inverse energy cascade of SQG turbulence (Lapeyre, 2017) which aggregates together the badly resolved small-scale turbulence structures. The resulting large vortices have chaotic trajectories. The difficulty for the coarse models to track these trajectories yield high CRPS values in the centers of those vortices. Nevertheless, the LU-PSO and LU-SGS models still provide much better results than the other random models.

The previous analysis describes the variability of the CRPS in time and space. However, the global evaluation of a model requires to summarize the CRPS information. Figure 6.11 summarizes the models' performances by representing each CRPS by one box plot. In this way, we can directly check the statistical variability of CRPS values over one CRPS map. For Vortices IC, the model PIC-FFT has the biggest spread of CRPS and it has the highest CRPS median and mean after 30 days of advection. Therefore, Figure 6.11 confirms that the model PIC-FFT has the worst uncertainties quantification (UQ)

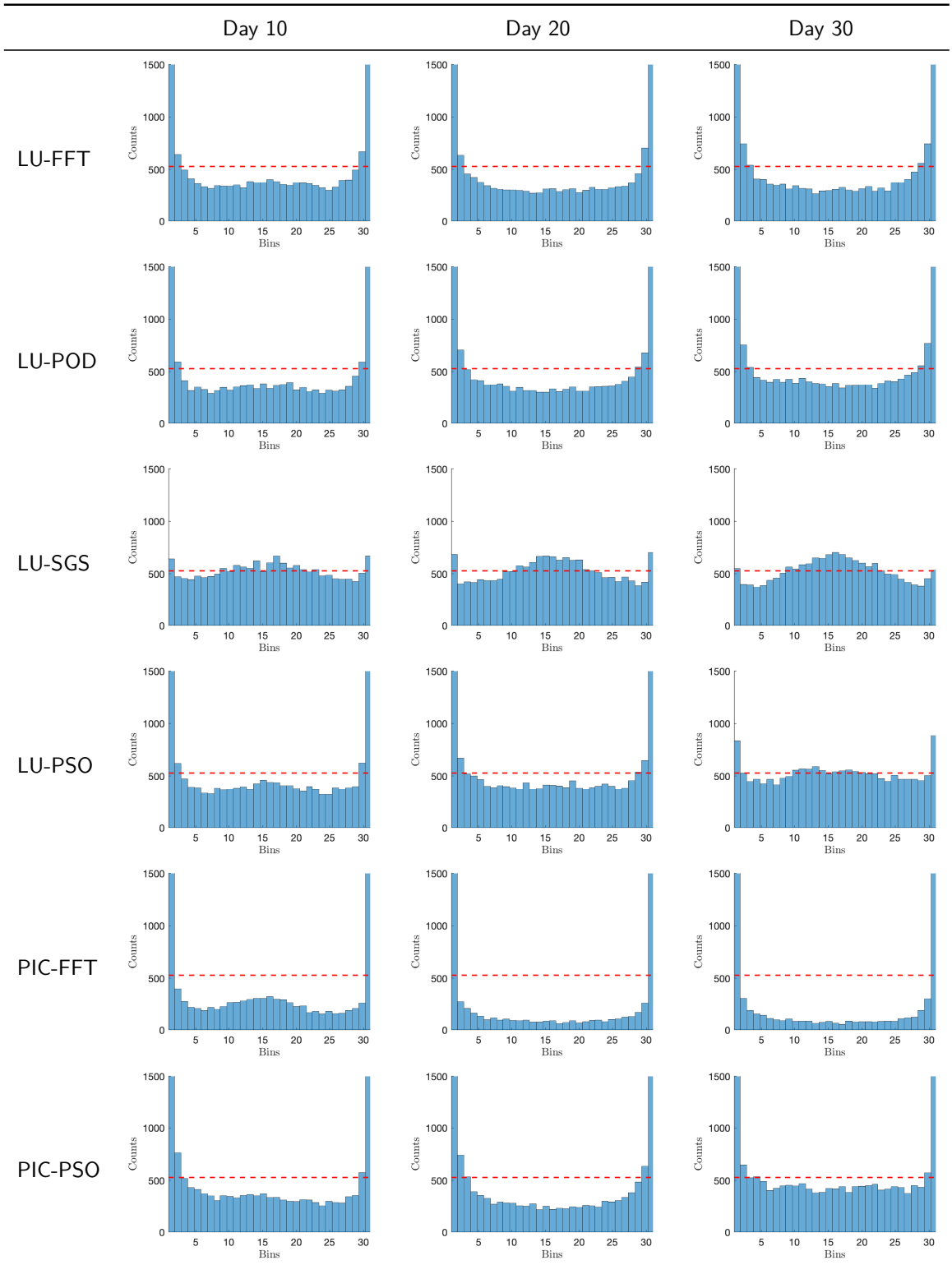

Table 6.1 – Talagrand diagrams of different random models on some forecast days under Vortices IC.

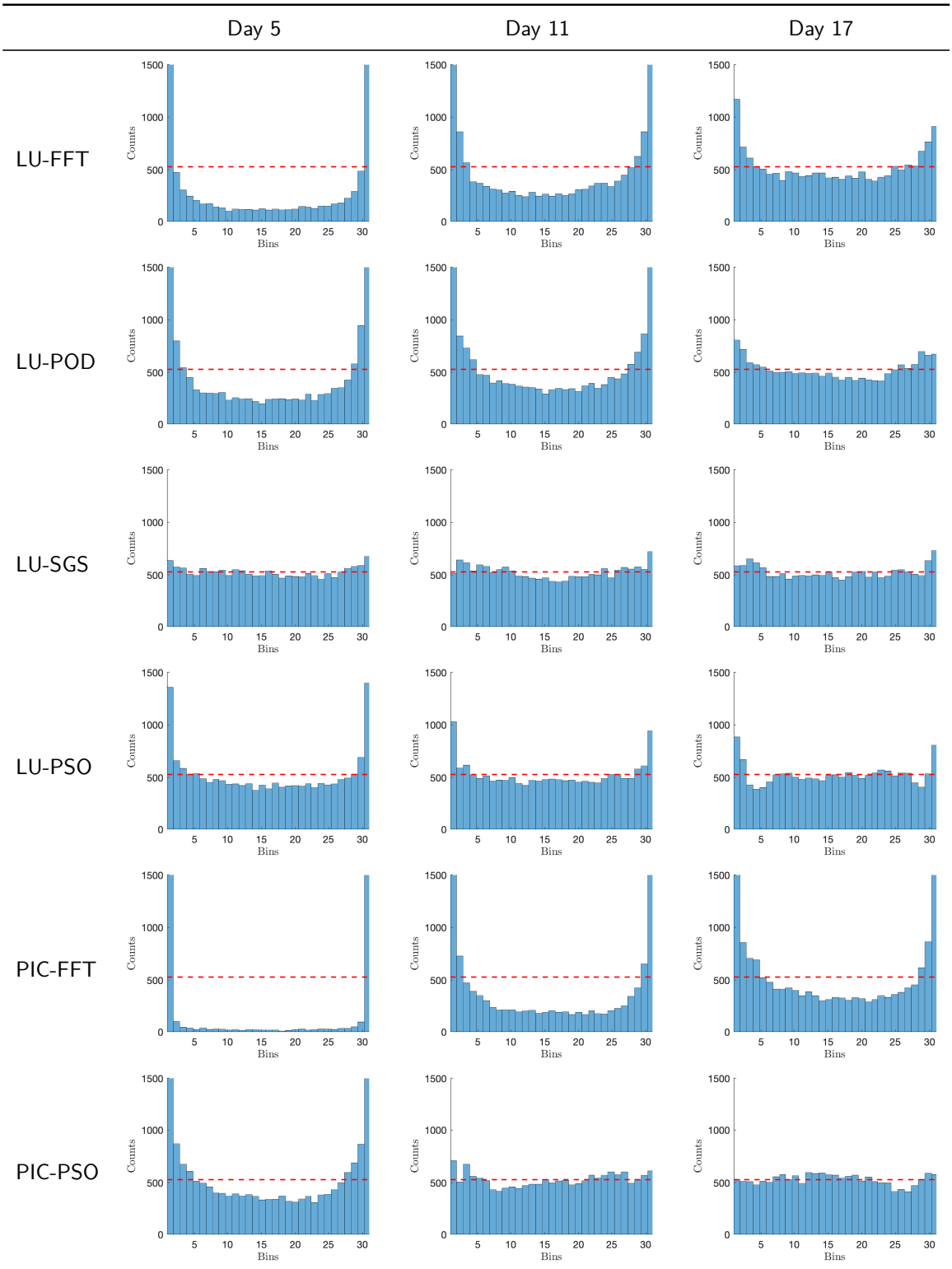

Table 6.2 – Talagrand diagrams of different random models on some forecast days under Spectral IC.

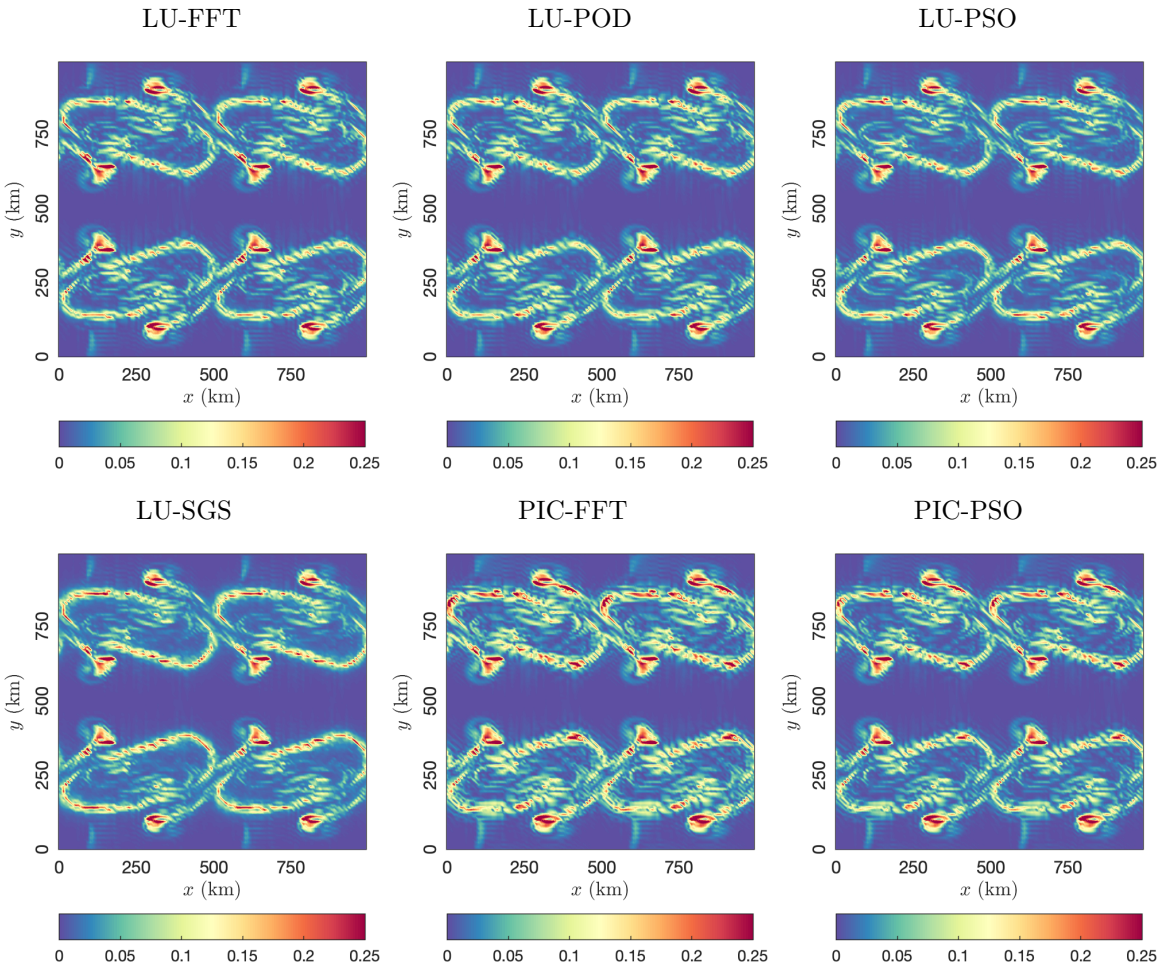

Figure 6.7 – Normalized CRPS of different random models after 15 days of advection under Vortices IC.

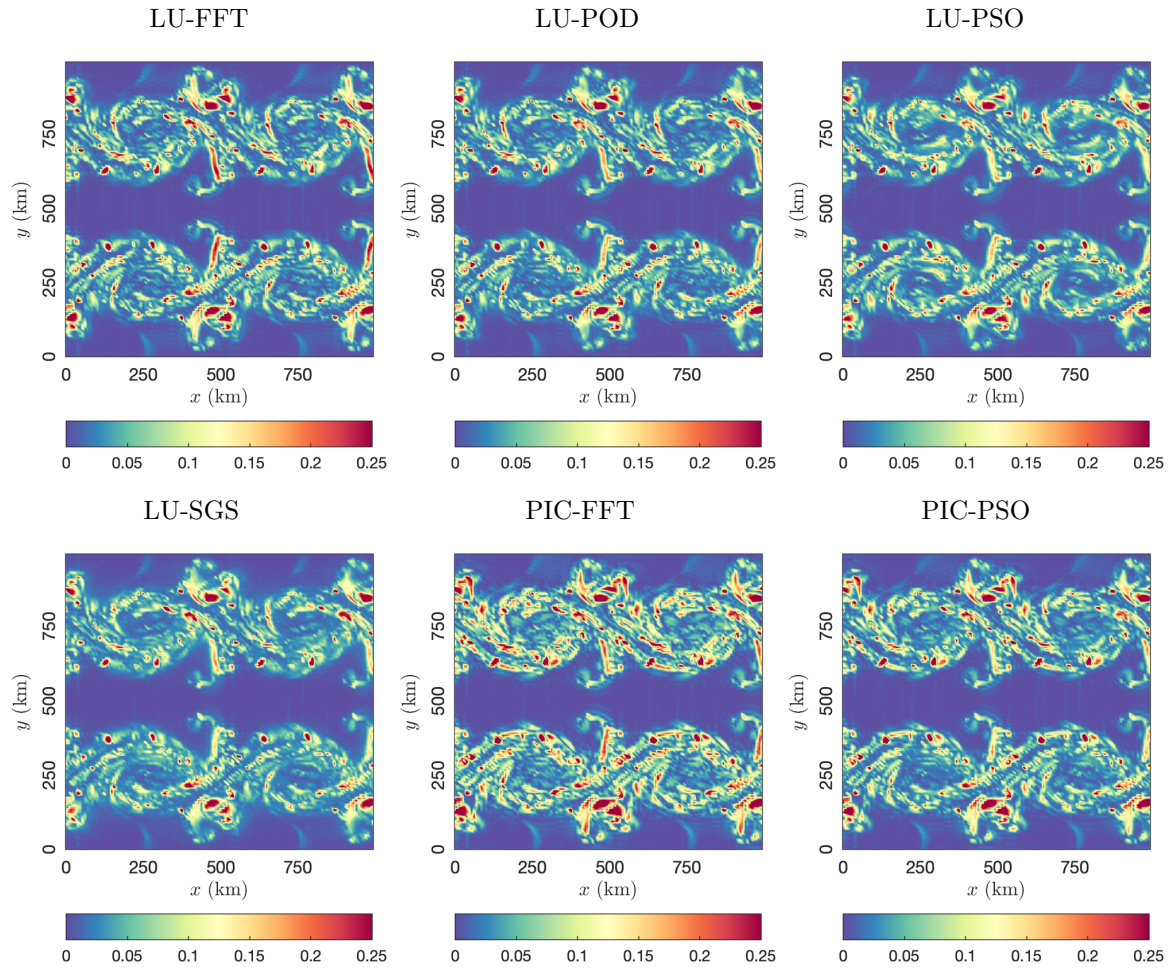

Figure 6.8 – Normalized CRPS of different random models after 20 days of advection under Vortices IC.

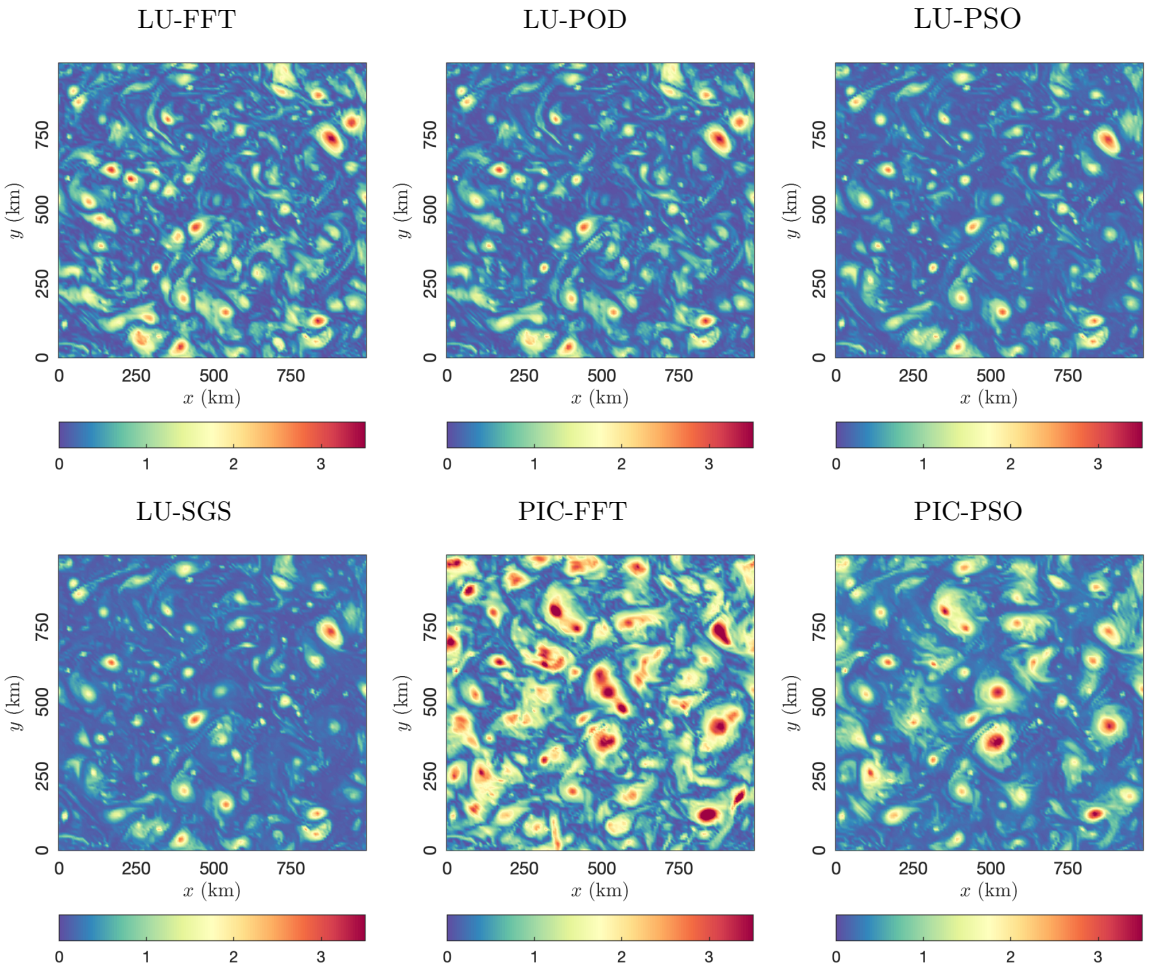

Figure 6.9 – Normalized CRPS of different random models after 5 days of advection under Spectral IC.

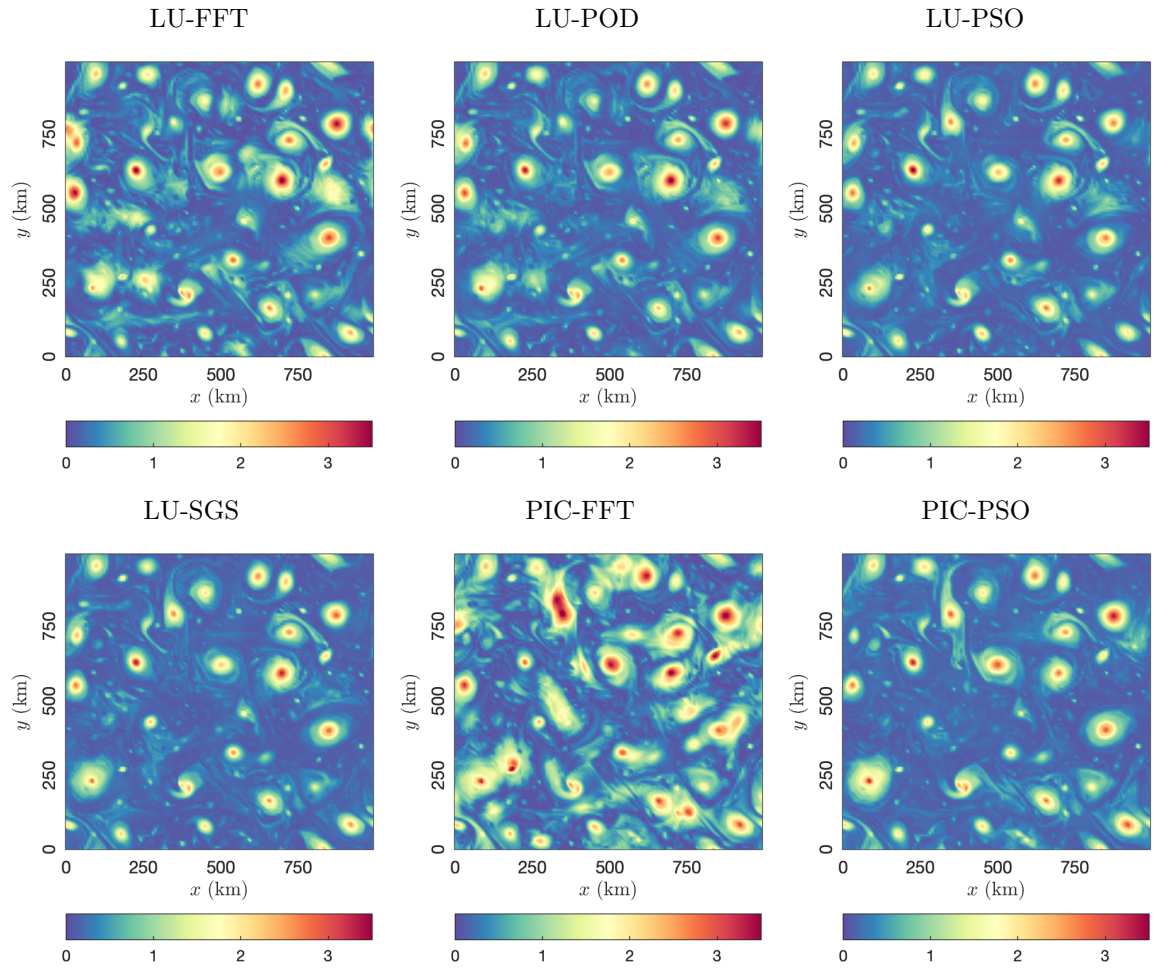

Figure 6.10 – Normalized CRPS of different random models after 10 days of advection under Spectral IC.

skill. After 20 and 30 days of advection, the models LU-FFT and LU-POD obtain the lowest CRPS median, whereas the models LU-PSO and LU-SGS have the lowest CRPS mean. Moreover, LU-SGS provides the lowest spread CRPS. As shown in Figure 6.12, for Spectral IC, PIC-FFT is always the worst with still the highest mean, median and spread at different time steps. After a short advection of 10 days, models LU-PSO and LU-SGS have the lowest mean, median and spread of CRPS. At a later time the other LU techniques show only slightly higher mean, median and spread than LU-SGS.

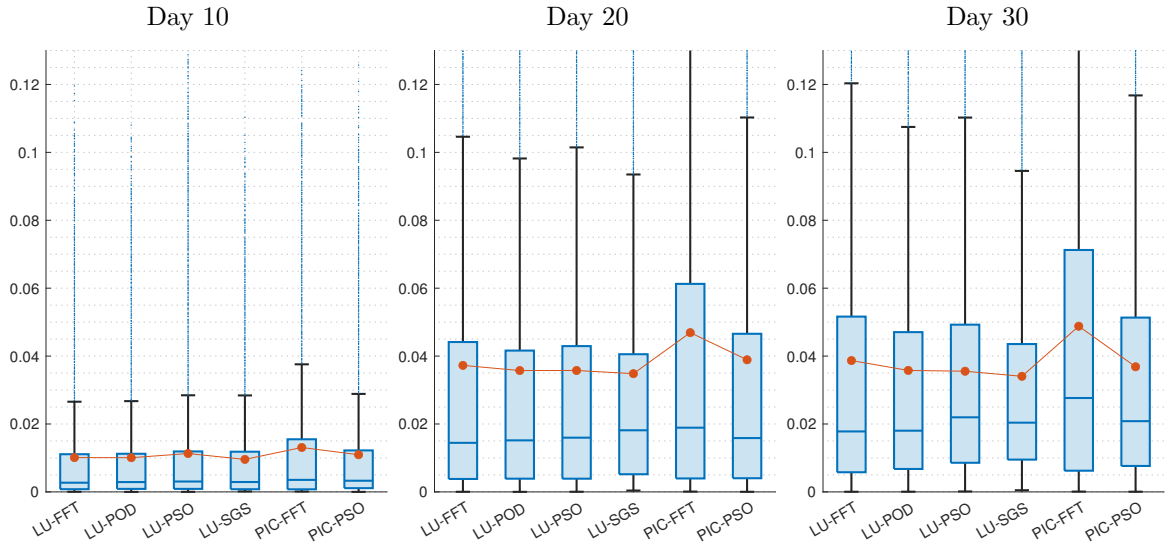

Figure 6.11 – Box plot of the normalized CRPS computed on the ensemble at each spatial points for each models at different time step of advection under Vortices IC. Each box plot displays the following information: the median (blue line inside the box), the lower and upper quartiles (contour blue lines), any outliers (blue points) computed using the interquartile range, and the minimum and maximum values (black lines) that are not outliers. In addition, the mean CRPS (red point) is added for each box.

In order to analyze the multivariate structure of the error between the ensemble and the reference, Figures 6.13 and 6.14 illustrate, respectively, the normalized mean CRPS and the normalized energy scores (see Section 6.1.4) of advection times. There, the analyzed multivariate structure is the spatial structure of the random fields. These two figures show that the model PIC-FFT has the highest mean CRPS and energy score under both initial conditions. Thus, this model provides the worst UQ skill with multivariate spatially ensemble. The model PIC-PSO has lower mean CRPS and energy score after two weeks of advection for Spectral IC and after three weeks for Vortices IC. In fact, for Spectral IC, the large-scale structures has formed after two weeks, hence PIC-PSO makes less error with

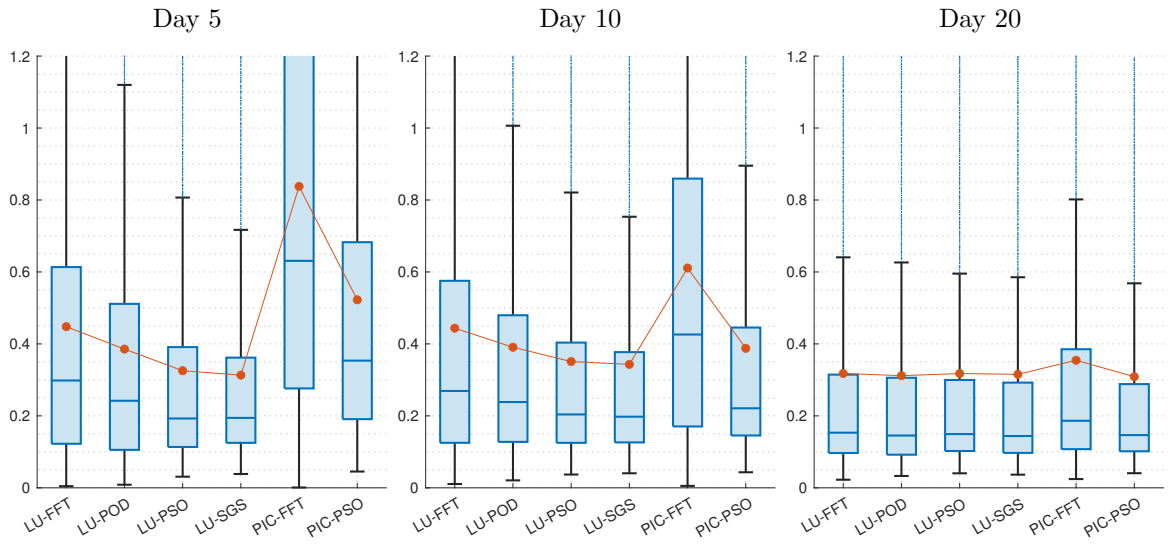

Figure 6.12 – Box plot of the normalized CRPS computed on the ensemble at each spatial points for each models at different time step of advection under Spectral IC.

the reference. Nevertheless, during the first week of advection, PIC-PSO has high CRPS and energy score. Within this period the small-scale structures are present and are not well reproduced by PIC-PSO. Thus, in this case, PIC-PSO is clearly less efficient than the models LU-POD, LU-PSO and LU-SGS. For Vortices IC, the model PIC-PSO has higher CRPS and energy score in the period 16 to 20 days of advection. During this period, some small-scale structures are produced and well resolved by the simulations. After this period, the two proper scores of PIC-PSO decrease probably due to the bifurcation phenomenon described in Resseguier et al. (2020a). Figures 6.13 and 6.14 suggest that the models LU-PSO and LU-SGS provide the weakest error. Indeed, during the major part of the simulation and for both initial conditions, these models show the lowest proper scores. In particular, the two proper scores of the LU-SGS method reveal that it is the most efficient ensemble model according to this metric.

We now consider proper scores where the temporal structures are considered instead of the spatial ones. For Vortices IC, Figure 6.15 reflects that the model PIC-FFT has the highest normalized energy score. On the edges of the centers of the vortices, this model encompasses the worst temporal structures. The models LU-POD, LU-PSO and LU-SGS seem to perform better and the latter is the best one. Moreover, these results hold for Spectral IC as well, which are illustrated by Figure 6.16.

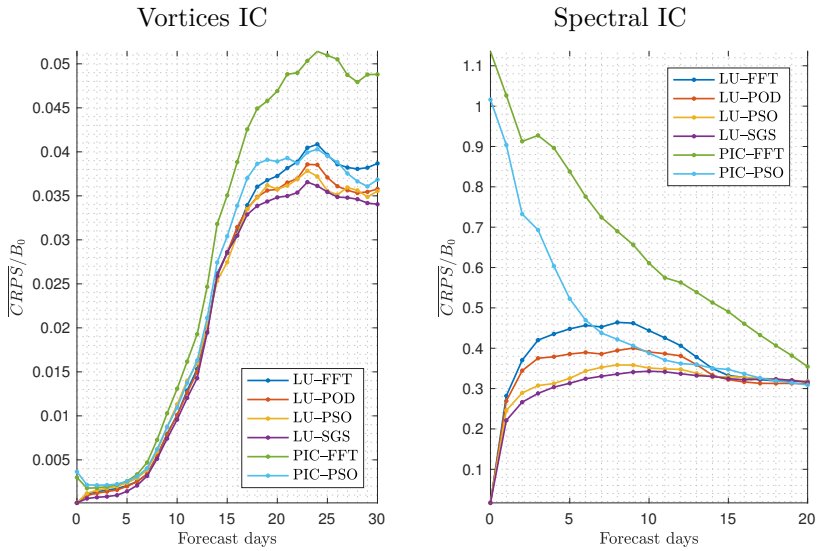

Figure 6.13 – Normalized CRPS of the spatial multivariate ensemble for each models on all time steps under both initial conditions.

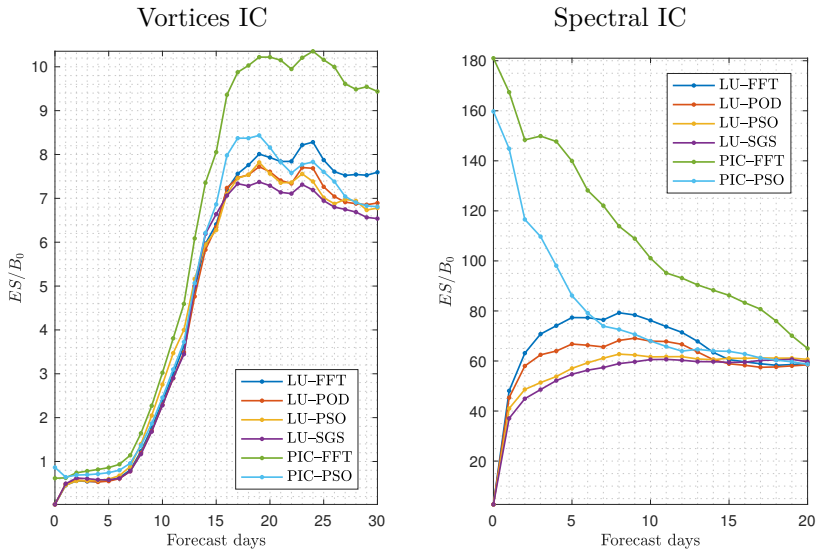

Figure 6.14 – Normalized energy score of the spatial multivariate ensemble for each models on all time steps under both initial conditions.

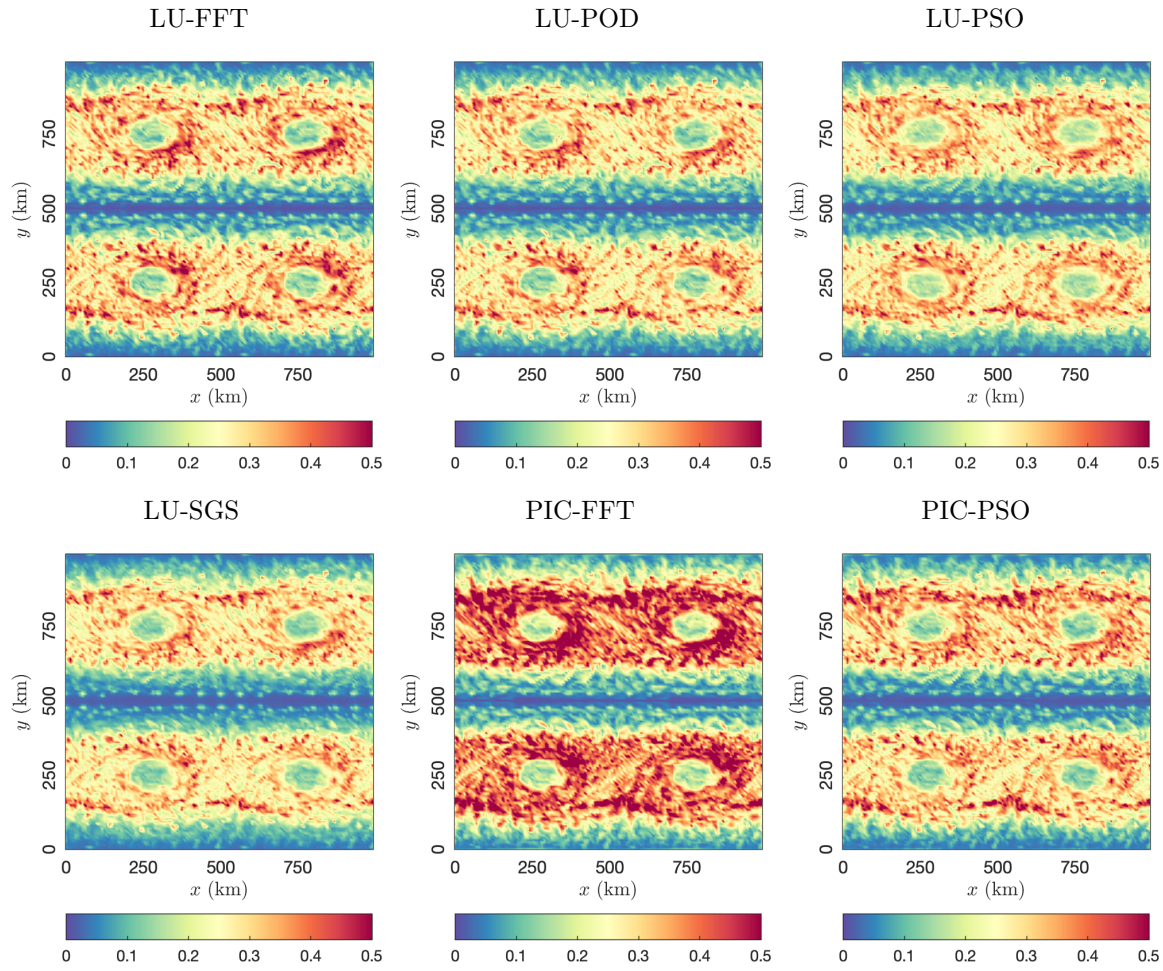

Figure 6.15 – Normalized energy score of the temporal multivariate ensemble for each models at each spatial points under Vortices IC.

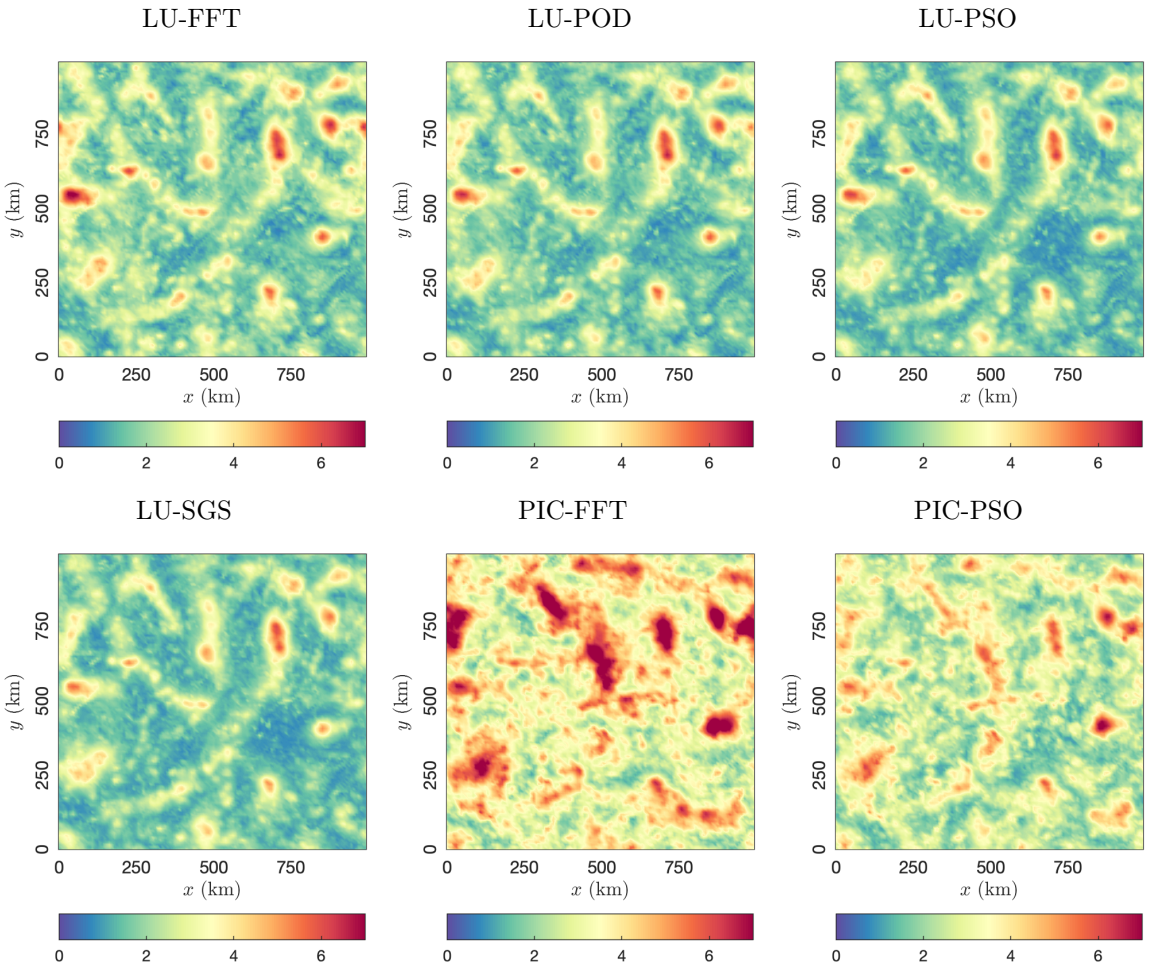

Figure 6.16 – Normalized energy score of the temporal multivariate ensemble for each models at each spatial points under Spectral IC.

6.4 Conclusion

For this numerical study, ensemble forecast skills have been assessed through a set of verification tools. Table 6.3 summarizes validation scores estimated for each UQ model. Talagrand diagram is an evaluation of the calibration behavior of an ensemble, while proper scores focus on dispersion and errors between univariate or multivariate ensembles and references. Each verification score has his own specificity, it is essential to be aware of their properties to select the most suitable measure (see Section 6.1). Also, to avoid miss-election of ensemble methods, the assessment of a number of scoring rules is advised. In Table 6.3, the PIC methods obtained the lowest overall verification scores. In contrast, LU-PSO and LU-SGS models present the best performances. Therefore, we recommend one of this method for UQ tasks, and we strongly advice to avoid relying only on initial conditions randomization.

To summarize this chapter, in terms of UQ skills, the heterogeneous noise models are better than the homogeneous noise models, and the non-stationary noise models are better than the stationary noise models. Nevertheless, the higher UQ skills of these more complicated methods also come with a slightly higher computational cost. The performances of LU-POD should be also relativized as they have been run here on two non-stationary examples. For stationary models at climatic scale such models may perform better. We will see however that the introduction of dynamics-adapted non-stationary features for the noise enables to improve the performances of POD noises.

Model	MSE	TD	CRPS	ES
LU-FFT	5	5	4	4
LU-POD	4	4	3	3
LU-PSO	2	2	2	2
LU-SGS	1	1	1	1
PIC-FFT	6	6	6	6
PIC-PSO	3	3	5	5

Table 6.3 – Rank of model performance by different scores. MSE: Mean squared error; TD: Talagrand diagram; CRPS: Continuous ranked proper score; ES: Energy score. The rank numbers 1 to 6 represent from the best to the worst accuracy.

Chapter 7

Numerical studies of stochastic barotropic QG model

Abstract

This chapter provides some numerical results of the stochastic barotropic quasi-geostrophic (BQG) system derived in Chapter 4. Using a simple inviscid BQG model, we first show that the introduction of inhomogeneous noise induces a structuration of the large-scale flow with strong secondary vortices. This corresponds to a simple numerical illustration of the effects described in Chapter 3. Latter, the performance of this stochastic barotropic model is assessed for the numerical simulation of an idealized wind-driven double-gyre configuration within an enclosed shallow basin at mid-latitude. We focus then on the ability of the proposed stochastic models to accurately represent at a coarse resolution the four first statistical moments (mean, variance, skewness and kurtosis) of the flow. Comparing this statistical distribution through its four moments to that predicted by the eddy-resolving (higher resolution) data enables us to qualify and quantify the accuracy of our stochastic representation of mesoscale eddy effects on large–scale circulation. The work presented in this chapter have been published in Bauer et al. (2020a,b).

7.1 Deciphering the role of small-scale inhomogeneity for inviscid BQG

In this section we aim at comparing, for a barotropic quasi-geostrophic stochastic model (BQG-LU), the effect of the isotropic homogeneous noise (5.2.1) with a null Itô-Stokes drift and the inhomogeneous noise (5.1.1) built from a scale similarity assumption. Following the conclusions of Chapter 3, for the first noise, no large-scale secondary structuration should be observed as no Itô-Stokes drift is associated whereas in the second case the Itô-Stokes drift is non-zero and should impact the large-scale solution.

The non-dimensionalized simulation of the BQG-LU system (with $\beta = L_d = 1$) for the two types of noise are initialized with the same monochromatic Rossby waves, $\psi(\boldsymbol{x}, t = 0) = 0.1\cos(2x)$ (which is a solution of the deterministic BQG system). The geometry is defined as a uniform 128×128 Cartesian grid within a double periodic domain $[0, 2\pi] \times [0, 2\pi]$. Hence, all prognostic variables, such as the stream function ψ, the relative vorticity $\xi = (\nabla^2 - \mathbf{I}/L_d^2)\psi$, the vector noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and the variance tensor $\boldsymbol{a}$ are double periodic. To discretize spatially the vorticity equation (4.41a), we employ Arakawa's nine-point conservative scheme (Arakawa and Lamb, 1977, 1981), together with second order centered finite differences for the stochastic and diffusion terms. For time-stepping of the large-scale time-correlated terms in (4.41a), we use a strong stability preserving third order Runge Kutta (RK3) scheme (Gottlieb, 2005) with a Courant-Friedrich-Lewy (CFL) number of 1/3. The time-uncorrelated terms in (4.41a) are integrated in the final step of the RK3 scheme using the Euler-Maruyama scheme (Gugole and Franzke, 2019; Pavliotis and Stuart, 2008). To invert the Helmholtz equation associated to the stream-function (4.41b), an efficient Fast Fourier Transform (FFT) solver (Press et al., 2007) is adopted.

The objective here, in addition to the study of the structuration effect by the small-scales, is to assess experimentally the preservation of the large-scale initial Rossby wave for the stochastic system with both noise models. In the inhomogeneous case, at each time we uniformly draw $n_o = 20$ pseudo-observations by sliding a $n_w^2 = 3 \times 3$ window over the grid, with circular boundary conditions. Both noise models are fixed to a similar amplitude. To that end, the homogeneous noise amplitude is fixed with $a_0 \approx 2.85 \times 10^{-5}$, which is determined by the mean amplitude of the inhomogeneous (and time varying) noise. The other parameters are chosen as: $k_M = \pi/\Delta$ with Δ the grid spacing, $k_m = k_M/2$ and $s = -3$. The vorticity evolution for each noise is plotted in Figures 7.1 and 7.2.

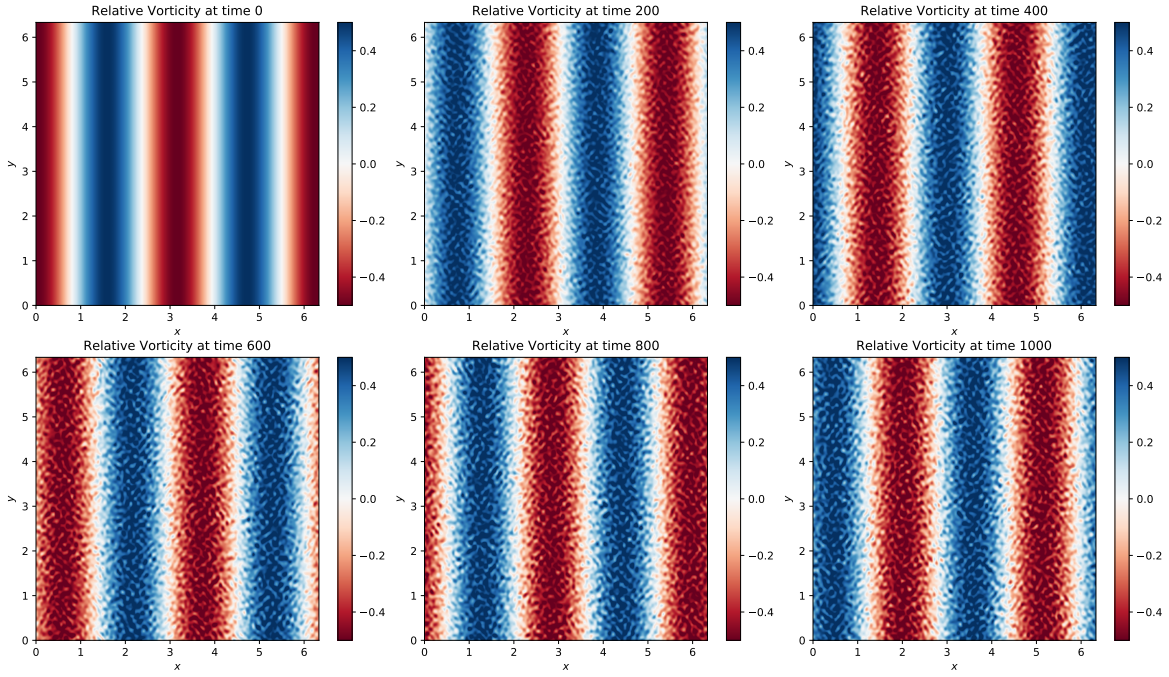

Figure 7.1 – BQG-LU evolution for the homogeneous stationary model over a period of $t = 1000$ adimensional time and a 128×128 spatial grid.

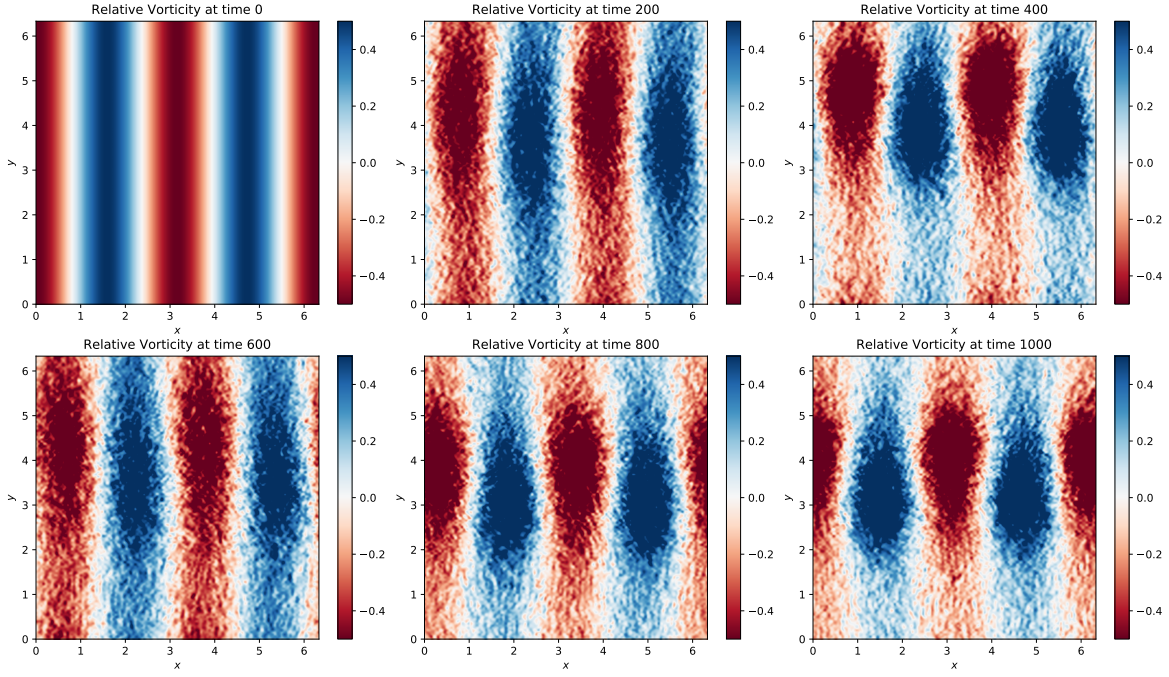

Figure 7.2 – BQG-LU evolution for the heterogeneous non-stationary model over a period of $t = 1000$ adimensional time and a 128×128 spatial grid.

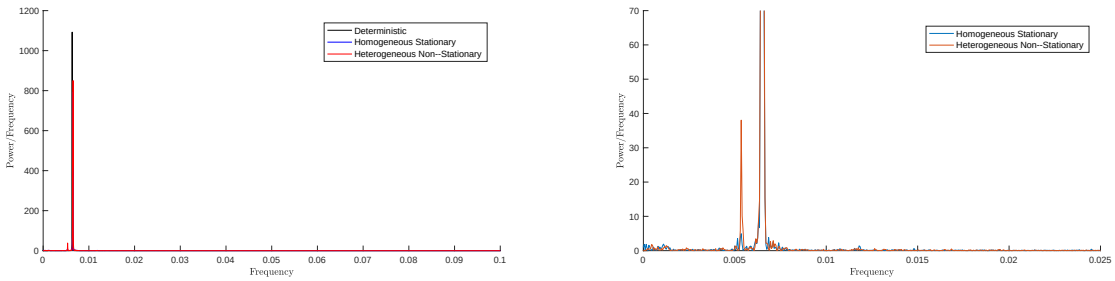

Figure 7.3 – Fourier spectra of the temporal series of a fixed grid point for the deterministic BQG, BQG-LU with homogeneous noise and BQG-LU with inhomogeneous noise; on the right zoom of the left plots.

As can be observed, both noises lead to stable solutions. The initial monochromatic wave can still be seen in both simulations in Figures 7.1 and 7.2. The monochromatic wave can also be seen in the Fourier spectrum of the temporal signal associated to a given point of the grid (see Figure 7.3). The deterministic BQG and the BQG-LU for both noises have a strong common frequency peak corresponding to this monochromatic wave. In the homogeneous case, for which the Itô-Stokes drift is null, we observe a statistically homogeneous solution with no particular structuration of the flow. This can be observed visually in Figure 7.1 or infered from the Fourier spectra in Figure 7.3. We clearly see the superposition of the monochromatic wave with a homogeneous noise pattern active at all scales. In contrast, for the inhomogeneous case, after a spin-up time of approximately $t = 200$, the apparition of large vortices can be observed in Figure 7.2. These patterns correspond to the secondary peak on the energy spectrum at slightly lower frequency than the monochromatic initial wave as observed in the RHS of Figure 7.3. They correspond to slower events. These vortices remain stable for a while then disappear and reappear with a longer time periodicity. This structuration is stable along time and still conserves the large scale Rossby waves. Note that due to energy conservation, a part of the energy of the monochromatic wave is redistributed to the secondary vortex structure (as see in Figure 7.3).

7.2 Stochastic barotropic wind-driven circulations

The wind–driven circulation is a classical simplified problem in oceanography (Vallis, 2017), which produces qualitatively realistic patterns of mesoscale eddies in approximate

geostrophic equilibrium. A particular circulation (Greatbatch and Nadiga, 2000) living in a highly turbulent regime under weak dissipation of potential enstrophy leads to a stationary four–gyre structure in a long–time average sense.

In this work, we use a single-layer QG formulation to study the wind-driven circulation in an oceanic basin following (Vallis, 2017). Under this regime, the dimensional barotropic vorticity equation (BVE) can be written as:

$$\frac{\partial \omega}{\partial t} + \mathrm{J}(\psi, \omega) + \beta \frac{\partial \psi}{\partial x} = F + D, \tag{7.1a}$$

$$\nabla^2 \psi = \omega, \tag{7.1b}$$

where $\omega = \boldsymbol{\nabla} \times \boldsymbol{u} = \partial_x v - \partial_y u$ is the relative (or kinematic) vorticity (henceforth, referred to as vorticity) with $\boldsymbol{k} = [0, 0, 1]^T$. The geostrophic velocity $\boldsymbol{u}$ can be defined by a stream function ψ such that $\boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi = [-\partial_y \psi, \partial_x \psi]^T$. The nonlinear advection is transformed into a Jacobian operator which is defined as $\mathrm{J}(\psi, \omega) = \partial_x \psi \partial_y \omega - \partial_y \psi \partial_x \omega$. The linear term $\beta \partial_x \psi$ describes the advection of β-planetary vorticity. An active tracer in this case is given by the potential vorticity (PV) defined as $q = \omega + \beta y$.

On the right-hand side (RHS) of (7.1a), $F = \boldsymbol{\nabla} \times \boldsymbol{\tau}/(\rho H)$ is a forcing which adds vorticity into the gyres, due to the wind stress $\boldsymbol{\tau}$ over the ocean surface, where ρ and H are respectively (resp.) the basic fluid density and depth of the basin. An idealized double-gyre wind stress (Greatbatch and Nadiga, 2000; San et al., 2011, 2013), defined only in zonal direction, is used in this work within the basin $\Omega = [0, L] \times [-L, L]$, that is

$$\boldsymbol{\tau} = [\tau_0 \cos(\frac{\pi y}{L}), 0]^T, \tag{7.2}$$

where τ_0 is the magnitude of the wind. This form of wind stress (San et al., 2011, 2013) represents the meridional profile of easterly trade winds, mid-latitude westerlies, and polar easterlies from south to north over the ocean basin.

The boundary layer friction D can be interpreted either as a linear drag for the Ekman layer as presented in the Stommel problem (Stommel, 1948), an eddy viscosity term as presented in the Munk problem (Munk, 1950), or a combination of the two (Fox-Kemper, 2005). In this work, we are more interested in the Munk model, by assuming that the ocean has a flat-bottom. The eddy viscosity that we will discuss in the following will be either harmonic $D = \nu_2 \nabla^2 \omega$ or biharmonic $D = -\nu_4 \nabla^4 \omega$, with a uniform coefficient ν_2

(of unit $m^2 s^{-1}$) or ν_4 (of unit $m^4 s^{-1}$).

To simplify the problem, one may scale the equation (7.1a) by comparing each term to the dominant β-effect (Vallis, 2017). The leading order is given by the Sverdrup balance between the rotation and wind forcing, i.e. $\beta \partial_x \psi \approx |F|$, which provides a characteristic size of velocity:

$$V = \frac{\tau_0}{\rho H} \frac{\pi}{\beta L}. \tag{7.3a}$$

This leads to the following scaling of time, vorticity and stream function:

$$t = \frac{L}{V} t', \qquad \omega = \frac{V}{L} \omega', \qquad \psi = VL\psi', \tag{7.3b}$$

where the variables with prime symbol ($'$) are adimensionalized.

The thickness of the Munk boundary layer can be then quantified by the balance between the β-effect and friction (Munk, 1950). For instance, $\beta \partial_x \psi \approx \nu_2 \nabla^2 \omega$ gives us a harmonic-boundary-layer scale, that is

$$\delta_2 = \left(\frac{\nu_2}{\beta} \right)^{1/3}. \tag{7.3c}$$

Similarly, $\beta \partial_x \psi \approx \nu_4 \nabla^4 \omega$ gives us a biharmonic-boundary-layer scale:

$$\delta_4 = \left(\frac{\nu_4}{\beta} \right)^{1/5}. \tag{7.3d}$$

The nonlinear advection term $\mathrm{J}(\psi, \omega)$ is smaller than the linear terms. Nevertheless, the nonlinear effect may still be important in the boundary layer, especially in the western one. To measure its strength, one may define a β-Rossby number (denoted as R_β) as the ratio of the size of the nonlinear term to the β-effect:

$$\mathrm{R}_\beta = \frac{V}{\beta L^2}. \tag{7.3e}$$

Using these scaling numbers (7.3a)–(7.3e) for (7.1a), the dimensional BVE reduces to its adimensional form as:

$$\frac{\partial \omega'}{\partial t'} + \mathrm{J}(\psi', \omega') + \frac{1}{\mathrm{R}_\beta} \frac{\partial \psi'}{\partial x'} = \frac{1}{\mathrm{R}_\beta} \sin(\pi y') + \frac{1}{\mathrm{R}_\beta} D, \tag{7.4}$$

with $D = (\delta_2 / L)^3 \nabla^2 \omega'$ or $D = -(\delta_4 / L)^5 \nabla^4 \omega'$ resulting from (7.3c) or (7.3d), respectively.

The adimensional PV is written as $q' = \mathrm{R}_\beta \omega' + y'$, and the Poisson equation (7.1b) is invariant under this adimensionalization, i.e. $\nabla^2 \psi' = \omega'$. For the sake of readability, in the following we drop the prime for all the adimensional variables.

To close the problem, we need one initial condition – that will be discussed in Section 7.3.1 – and two boundary conditions. The first boundary condition is imposed by the no-normal-flow condition due to the forcing form:

$$\psi|_{\partial\Omega} = 0, \qquad \text{i.e. } u|_{x=0,L} = v|_{y=-L,L} = 0, \tag{7.5a}$$

where $\partial\Omega$ denotes the basin's boundary. The second one depends on the chosen eddy viscosity form. For a harmonic friction, i.e. $D = (\delta_2/L)^3 \nabla^2 \omega$, we impose

$$\omega|_{\partial\Omega} = 0, \tag{7.5b}$$

while for a biharmonic friction, i.e. $D = -(\delta_4/L)^5 \nabla^4 \omega$, we set

$$\omega|_{\partial\Omega} = 0 \quad \text{and} \quad \left.\frac{\partial^2 \omega}{\partial n^2}\right|_{\partial\Omega} = 0, \tag{7.5c}$$

where $\frac{\partial^2}{\partial n^2}$ denotes for the second derivative in normal direction. Note that in both cases, together with the no-normal-flow condition, we get a free-slip condition

$$\left.\frac{\partial^2 \psi}{\partial n^2}\right|_{\partial\Omega} = 0, \qquad \text{i.e. } \left.\frac{\partial v}{\partial x}\right|_{x=0,L} = \left.\frac{\partial u}{\partial y}\right|_{y=-L,L} = 0, \tag{7.5d}$$

with no horizontal shear on each boundary. Finally, we remark that the Munk model (7.4) depends only on two parameters, which are R_β and δ_2/L (resp. δ_4/L).

On the other hand, concerning the stochastic BQG equation (4.41), let us first recall the scaling of variance tensor and small-scale flow (see Section 4.2.1 for details):

$$\boldsymbol{a} = \epsilon V L \boldsymbol{a}', \qquad \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon} L \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}'_t. \tag{7.6}$$

In addition, only moderate uncertainty ($\epsilon \sim 1$) is adopted for the present work. Under such an assumption, the final dimensionless stochastic barotropic vorticity equation (SBVE)

in Stratonovich notation is written as

$$\mathbb{D}_t^\epsilon \circ \omega + \frac{1}{\mathrm{R}_\beta}\frac{\partial}{\partial x}(\psi\,\mathrm{d}t + \sqrt{\epsilon}\,\varphi \circ \mathrm{d}B_t) = \frac{1}{\mathrm{R}_\beta}(F + D)\,\mathrm{d}t + \mathrm{d}_t \circ S^\epsilon, \tag{7.7a}$$

$$\mathbb{D}_t^\epsilon \circ \omega \triangleq \mathrm{d}_t \circ \omega + \mathrm{J}\Big((\psi - \epsilon\,\psi_s)\,\mathrm{d}t + \sqrt{\epsilon}\,\varphi \circ \mathrm{d}B_t, \omega\Big), \tag{7.7b}$$

$$\mathrm{d}_t \circ S^\epsilon = \sum_{i=1,2} \mathrm{J}(\epsilon\,u_s^i\,\mathrm{d}t - \sqrt{\epsilon}\,\sigma \circ \mathrm{d}B_t^i, u^i). \tag{7.7c}$$

To close the problem, we assume that the small-scale component $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and the Itô-Stokes drift $\boldsymbol{u}_s$ have the same boundary conditions as the large-scale current $\boldsymbol{u}$, given in (7.5a) and (7.5d).

7.3 Numerical simulations of SBVE

In the following, we discuss and compare the respective numerical simulations of the BVE (7.4) and the SBVE (7.7). The main motivation here is to numerically assess if the proposed random model reproduces well the long-term statistics of the high resolution (eddy-resolving) simulation.

7.3.1 Model configurations

All the models have been discretized with the same numerical schemes. As detailed in Appendix 8.3, a staggered Arakawa C-grid (Arakawa and Lamb, 1977) has been considered. In that respect, the nonlinear Jacobian terms in the governing equations are discretized using Arakawa's 9-points conservative scheme (Arakawa and Lamb, 1981). To invert the Poisson equation (7.1b) associated to the stream function, an efficient discrete sine transform solver (Press et al., 2007) is adopted. For the time-stepping, the 3rd order Runge-Kutta scheme (Gottlieb, 2005) with a CFL number of 1/3 is considered for BVE. As presented by Cotter et al. (2019a), for the SBVE we used a similar time integration scheme.

In all the configurations we fix the basin length to $L = 1$ and the Rossby number to $\mathrm{R}_\beta = 0.06^2$. For the high resolution eddy-resolving model, a regular mesh with 256×512 cells with uniform grid spacing $\Delta_{\mathrm{HR}} = 0.004$ and a five times wider harmonic boundary layer $\delta_2 = 0.02$ have been used. We consider a quiescent state as the initial condition, that is $\psi(\boldsymbol{x}, t = 0) = 0, \forall \boldsymbol{x} \in \Omega$. For such an initial condition, the dominant Sverdrup balance between the forcing and rotation leads to a symmetric PV field during a short

period. As the nonlinear inertial term becomes more and more important, a symmetry breaking phenomena occurs (at $t \approx 2$), which can be observed from the time series of the global kinetic energy in Figure 7.4. This so-called spin-up period is then followed by a dissipation stage (up to $t \approx 5$) of the very high enstrophy that has been produced during the spin-up. Immediately after, the flow dynamics becomes rapidly turbulent. The three subsequent snapshots in Figure 7.4 illustrate this vigorous eddying nature. At coarse resolution, the subgrid dissipation model is defined through a biharmonic friction term with a grid-dependent uniform coefficient. The estimated values of δ_4 at coarse resolutions 64×128, 32×64 and 16×32 are, respectively, 0.026, 0.040 and 0.049.

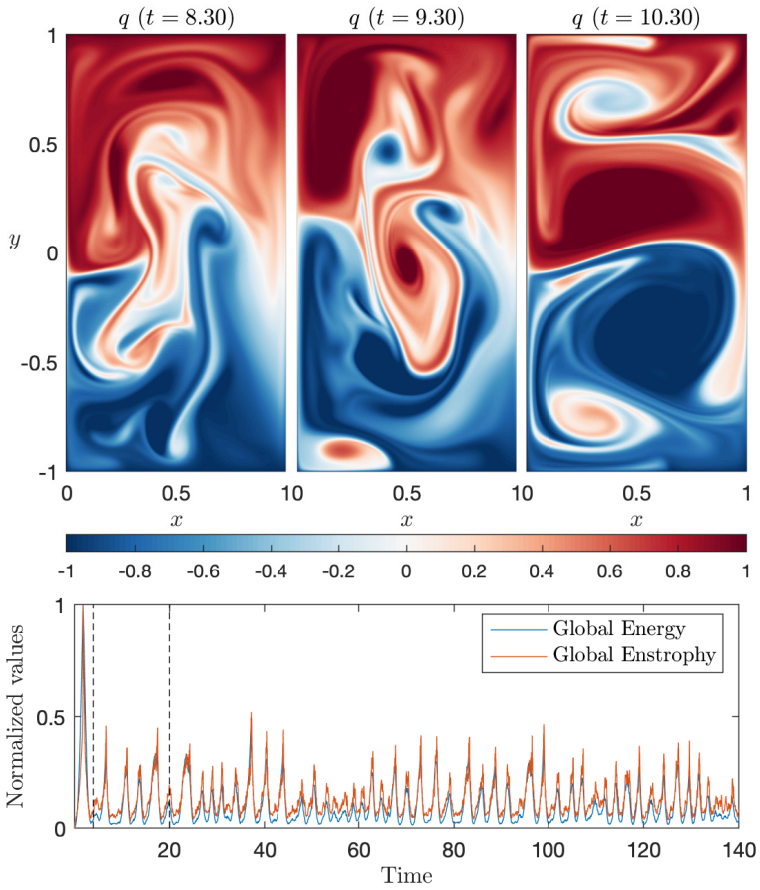

Figure 7.4 – Instantaneous snapshots of PV and time series of the global energy and enstrophy, provided by the eddy-resolving BVE at resolution 256×512. The global energy is defined by $E(t) = \frac{1}{2}\int_\Omega (u^2 + v^2) \mathrm{d}\boldsymbol{x}$ and the global enstrophy is defined by $Z(t) = \frac{1}{2}\int_\Omega \omega^2 \mathrm{d}\boldsymbol{x}$. The plots show their graph normalized by their temporal maxima.

The numerical simulations of the SBVE are performed using both the POD (denoted

as SBVE_{POD}) technique and Mode Matching (denoted as SBVE_{MM}) approach. As shown in Figure 7.5, by introducing randomness into the initial symmetric double-gyre circulation, the symmetry breaking state is reached much earlier for the SBVE simulations, than for the BVE. Hereafter, in order to compare the different models and to reduce the spin-up errors, we use the coarse-grained version of one specific eddy-resolving snapshot (after $t = 5$) as the initial condition for all coarse model runs. In other words, the BVE and the SBVE at each coarse resolution are simulated from the very same initial field, in which the spin-up period is accounted for at the eddy-resolving resolution. An instantaneous illustration of the small-scale random stream function, denoted as $\psi_r \triangleq \frac{1}{\Delta t}\varphi \mathrm{d}B_t$, and the Itô-Stokes stream function ψ_s, is shown in Figure 7.6. It appears that both ψ_r and ψ_s based on MM are stronger and more regular than those based on POD.

7.3.2 Long-term prediction of statistics

Although we are working in a turbulent regime, the statistics of the large-scale tracers ψ and q tend to reach a statistical steady state equilibrium. As shown in Greatbatch and Nadiga (2000), a robust four-gyre structure is characterized in time-averaged circulation, as long as the dissipation is sufficiently weak. Here, a weak dissipation means that the boundary layer size δ_2 or δ_4 has a smaller order than the so-called Rhines scale $\sqrt{\mathrm{R}_\beta}$ (Vallis, 2017). However, this does not indicate that the flow dynamics are under resolved. Note that in under resolved simulations, the contour lines of the averaged tracers would be oscillating. On the other hand, increasing the explicit dissipation up to the order of Rhines scale, would result in a conventional double-gyre.

In this work, apart from the mean structure, we are also interested in the eddy energy distributions and higher order moments of the tracers, such as skewness and kurtosis. These two standard moments of a probability distribution characterize the asymmetry and extreme events, respectively. They are particularly informative when the distribution is non-Gaussian.

More precisely, the first four central moments of ψ are defined by

$$m_1[\psi] = \overline{\psi}^t, \qquad m_k[\psi] = \overline{(\psi - m_1[\psi])^k}^t, \ k = 2, 3, 4, \tag{7.8a}$$

where the superscript $(^k)$ denotes the power, while the subscript $(_k)$ denotes the order of the moment order. Similarly, the central moments of q as function of the prognostic

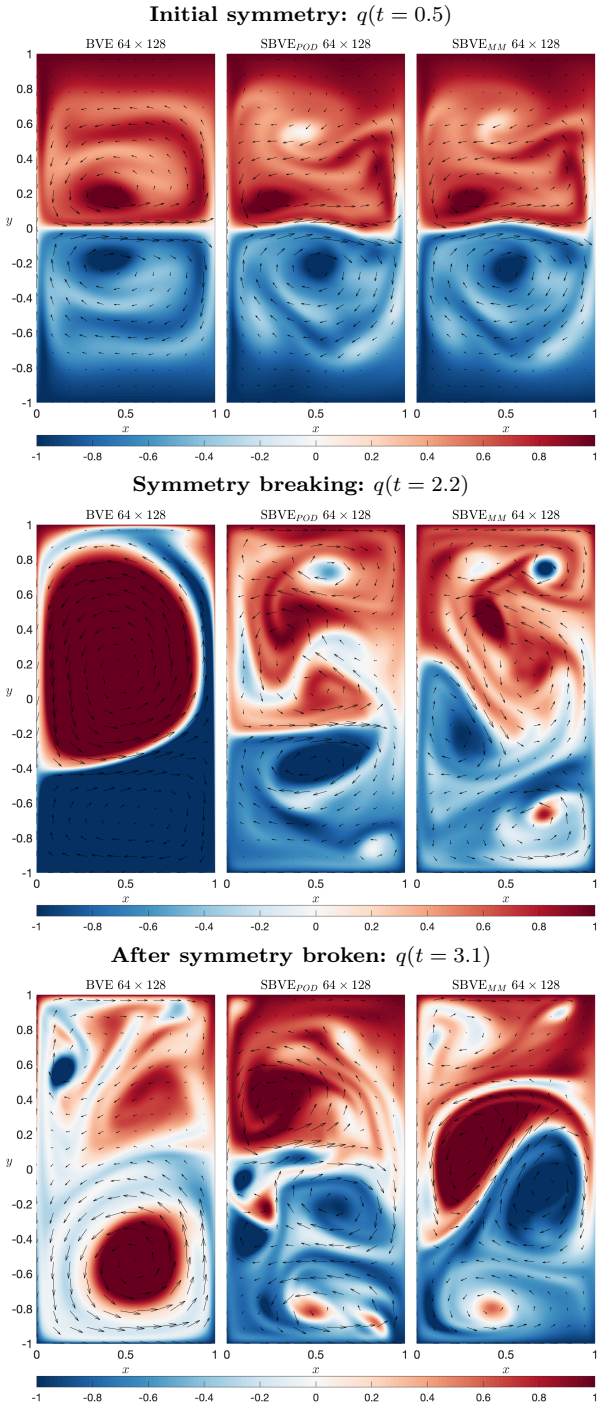

Figure 7.5 – Instantaneous snapshots of PV provided by different models at resolution 64×128. The associated large-scale velocity field is indicated here by the black arrows. Note that these velocity values are located on the PV-grid (see Figure 8.23) through linear interpolations.

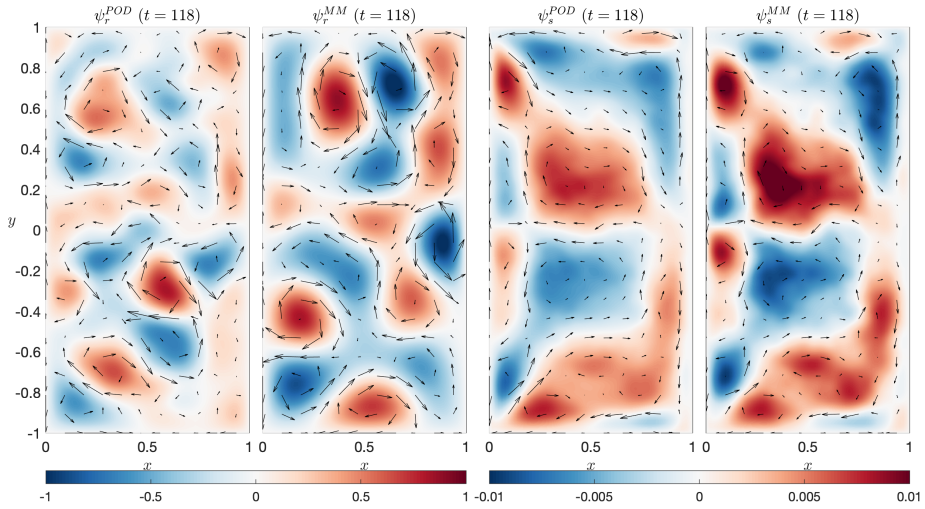

Figure 7.6 – Instantaneous snapshots of the small-scale random stream functions ψ_r^{POD}, ψ_r^{MM} and the Itô-Stokes stream functions ψ_s^{POD}, ψ_s^{MM}, resp. provided by the $\mathrm{SBVE}_{\mathrm{POD}}$ and the $\mathrm{SBVE}_{\mathrm{MM}}$ at resolution 64×128. The associated small-scale random velocity $\frac{1}{\Delta t}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is indicated here by the black arrows. Note that these velocity fields are located on the ψ-grid (see Figure 8.23) through linear interpolations.

variable ω are defined by

$$m_1[q] = \mathrm{R}_\beta m_1[\omega] + y, \qquad m_k[q] = \mathrm{R}_\beta^k m_k[\omega], \;\; k = 2, 3, 4. \tag{7.8b}$$

The skewness s (resp. kurtosis k) of ψ reduces to

$$s[\psi] = \frac{m_3[\psi]}{(m_2[\psi])^{3/2}}, \qquad k[\psi] = \frac{m_4[\psi]}{(m_2[\psi])^2} - 3, \tag{7.8c}$$

where algebraic manipulations ensure that the kurtosis of the Gaussian distribution is zero. The skewness (resp. kurtosis) of q is given by

$$s[q] = s[\omega], \qquad k[q] = k[\omega]. \tag{7.8d}$$

We remark from (7.8) that the skewness and kurtosis of both tracers ψ and q are not defined at boundaries, since the second moments are zero there. In addition, the eddy kinetic energy (EKE) and the eddy potential enstrophy (EPE) are provided through second order moments by:

$$\mathrm{EKE} = \frac{1}{2}(m_2[u] + m_2[v]), \qquad \mathrm{EPE} = \frac{1}{2}m_2[q]. \tag{7.8e}$$

In the following, theses statistics are computed for both BVE and SBVE at resolution 64×128, 32×64 and 16×32. Before discussing the results, the convergence of each statistic at each resolution is quantified. This can be done by progressively increasing the time interval, and computing a global error of the statistics between two adjacent intervals. More precisely, let us consider a point-wise statistic f obtained for a sufficiently long interval $[t_0, t_1]$ (where t_1 depends on the resolution considered) with a uniform partition of increment δt. We propose to measure the convergence by a relative global error $\tilde{\epsilon}$ between the subintervals $[t_0, t]$ and $[t_0, t-\delta t]$:

$$\tilde{\epsilon}(f_t) \triangleq \frac{\|f_t - f_{t-\delta t}\|_2}{\|f_{t_1}\|_2}, \tag{7.9}$$

where $\|\cdot\|_2 = \langle\cdot,\cdot\rangle_\Omega$ stands for the $L^2(\Omega)$-norm, and $f_t(\boldsymbol{x}), \forall \boldsymbol{x} \in \Omega$, denotes the local-in-time point-wise statistics associated to the interval $[t_0, t]$. In practice, we initiate this procedure from a reasonable intermediate instant $t_c \in [t_0, t_1]$, and t_0 is a fixed time after the spin-up (set to $t_0 = 20$ in this work, c.f. Figure 7.4) and the time increment has been fixed to $\delta t = 0.1$. A statistic is considered to be converged, as soon as the time series of relative global errors reaches a stable low error level. As shown in Figure 7.7, we observe that the convergence to an error less than 1% for resolutions $256\times512, 64\times128, 32\times64$, and 16×32 is reached approximatively after the time $140, 250, 350$ and 500, respectively. We note that the coarser the resolution, the longer it takes to get converged statistics. This is even more pronounced for higher moments. This is likely due to higher values of the turbulent viscosity which prevent the flow to visit freely its attractor and enforce it to stay for a much longer time in the attraction basin of the equilibrium points (Chapron et al., 2018). Note also that as observed therein, the convergence time for SBVE is shorter for all resolutions studied here (not shown). Therefore, we choose to use for all simulations the slowest convergence time (i.e. the one computed for BVE).

Hereafter, we focus on the comparisons of the statistics obtained for the different coarse models. To build a reference (REF) for each resolution, we directly subsample the statistics computed on the eddy-resolving data - i.e. we do not smooth them in order not to lower their energy. Figure 7.8 shows that at the coarsest resolution 16×32, the four-gyre structure is captured for both models, yet the two outer gyres predicted by SBVE are more enhanced and closer to the reference, compared to those obtained by BVE. Since the scale parameters are fixed, the major contribution comes from the stochastic representation of the mixing effects incorporated through the eddy-resolving data. A more accurate

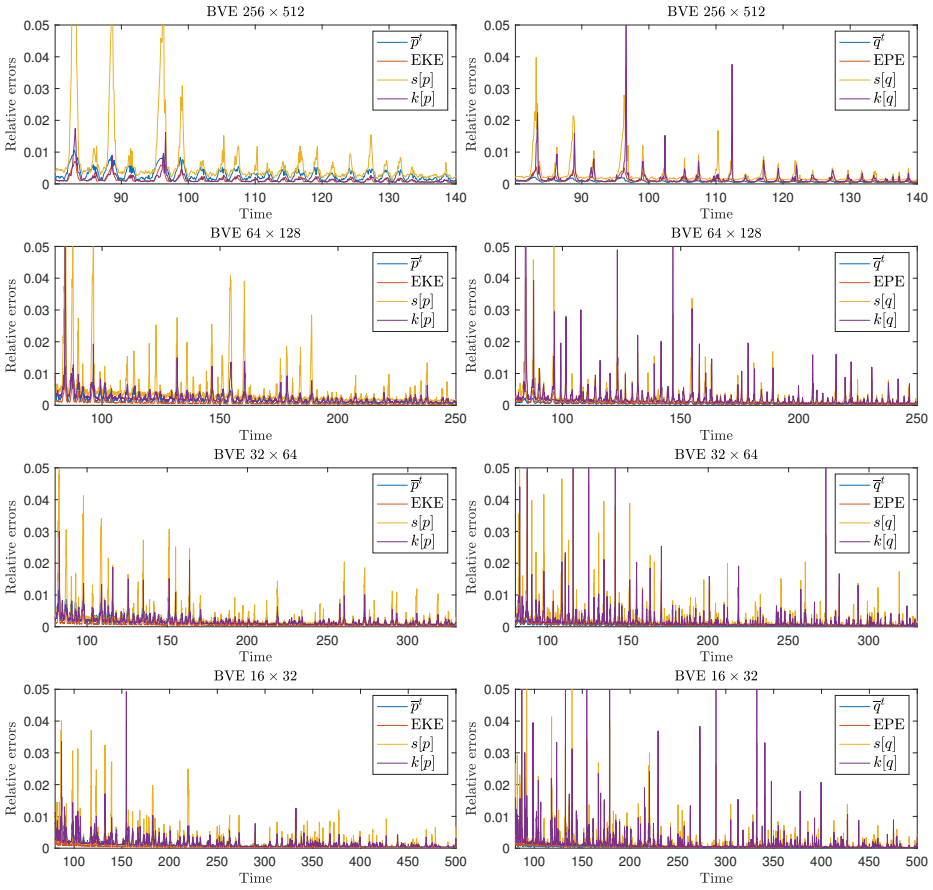

Figure 7.7 – Time series of the relative errors of the statistics by progressively increasing the time interval. In each row, the left plot shows the statistical errors of the stream function (or velocity), and the right one shows that of the PV. In each column, the results correspond, from top to bottom, to resolutions 256×512, 64×128, 32×64, and 16×32. Note that in both cases, the first (adimensioned time) interval on which we compute the statistics is set to be $[20, 80]$; this interval is progressively augmented with a time step of 0.1. The Y-axis values describe the converging percentage of one statistic w.r.t. its global (over the spatial domain) value performed at previous instant.

nonlinearity is produced such that a stronger distortion of the PV field between inner and outer gyres is observed. From Figure 7.9, we observe that compared to BVE, SBVE_{MM} produces higher eddy energy in the front between the outer and inner gyres, and higher eddy enstrophy in the region between the two inner gyres. However, both coarse models do not produce enough energy flux in the western and eastern boundary layers. In particular, the too low tracers' variance in the eastern boundary layers leads to markedly higher skewness and kurtosis than those observed in the reference. Nevertheless, the introduction of randomness enables us to increase the internal variability of the tracers. For instance, as shown in Figures 7.10 and 7.11, the region with extreme values of skewness and kurtosis is significantly reduced for SBVE_{MM} when compared to BVE. As the resolution increases, it can be noticed from Figures 7.12 and 7.13 that the local structures of the PV statistics provided by SBVE_{MM}, qualitatively converges to the reference.

In order to provide a more quantitative comparison, we propose here a global performance index, measured by the root mean squared error (RMSE) with an *a-posteriori* normalization to ensure a similar error level of the different statistics. Given a statistic f with reference f_{REF}, the normalized RMSE is defined as

$$\overline{\text{RMSE}}(f) = \frac{\frac{1}{|\Omega|}\|f - f_{\text{REF}}\|_2}{\max\limits_{\boldsymbol{x}\in\Omega} |f_{\text{REF}}(\boldsymbol{x})|}. \tag{7.10}$$

Table 7.1 compares the results of the different models at the coarsest resolution 16×32. The proposed stochastic model shows a clear improvement of all the statistics w.r.t. the references. This improvement is particularly noticeable for the higher moments. For instance, compared to BVE, SBVE_{MM} has 35.87% and 39.26% less errors in skewness and kurtosis of the stream function (SF), respectively. The mode matching strategy, SBVE_{MM}, performs better than the POD strategy, SBVE_{POD}, for all moments, although the latter already reduces the BVE error of the first and second moments (with an improvement of $9,7\%$ for the SF mean and $12,6\%$ for EKE). Both SBVE_{MM} and SBVE_{POD} reach very similar errors in terms of EKE and EPE (with an improvement above 10% for both quantities) and SBVE_{MM} is more efficient in reducing errors in the third and fourth moments. These results highlight the benefits that are brought by properly incorporating, into large-scale simulations, the effects of the small-scale flow component through its statistical distribution. From Table 7.2 and 7.3 we see that these RMSEs improvements still hold as the resolution is increased. The improvements at resolution 64×128 in terms of EKE and EPE are still noticeable (25%). The third order moment of SF continues to improve (20%)

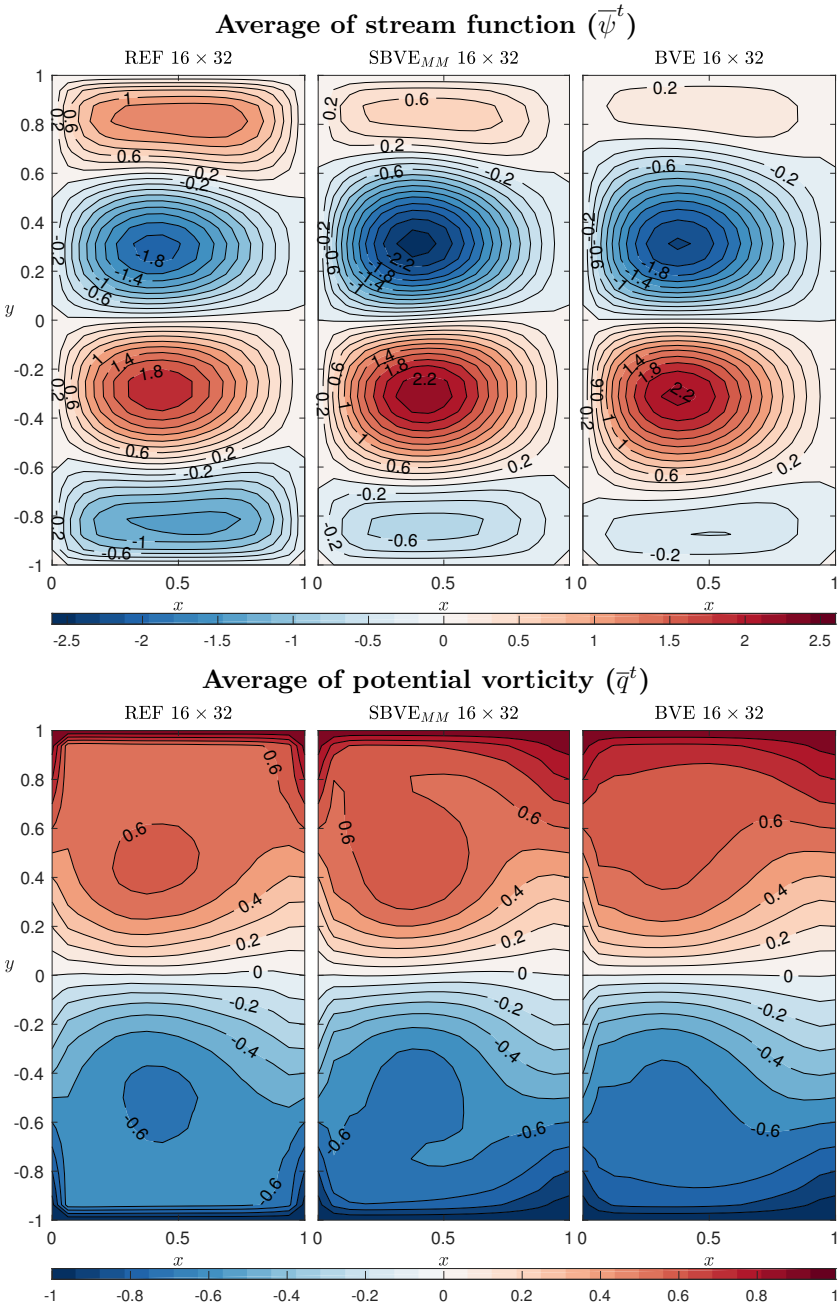

Figure 7.8 – Contour plots of the time-average fields at resolution 16×32. The top three plots depict the SF with contour interval (CI) of 0.2, and the bottom three show PV with CI of 0.1. In each panel, the first one is REF, the second one is SBVE_{MM} and the third one is BVE.

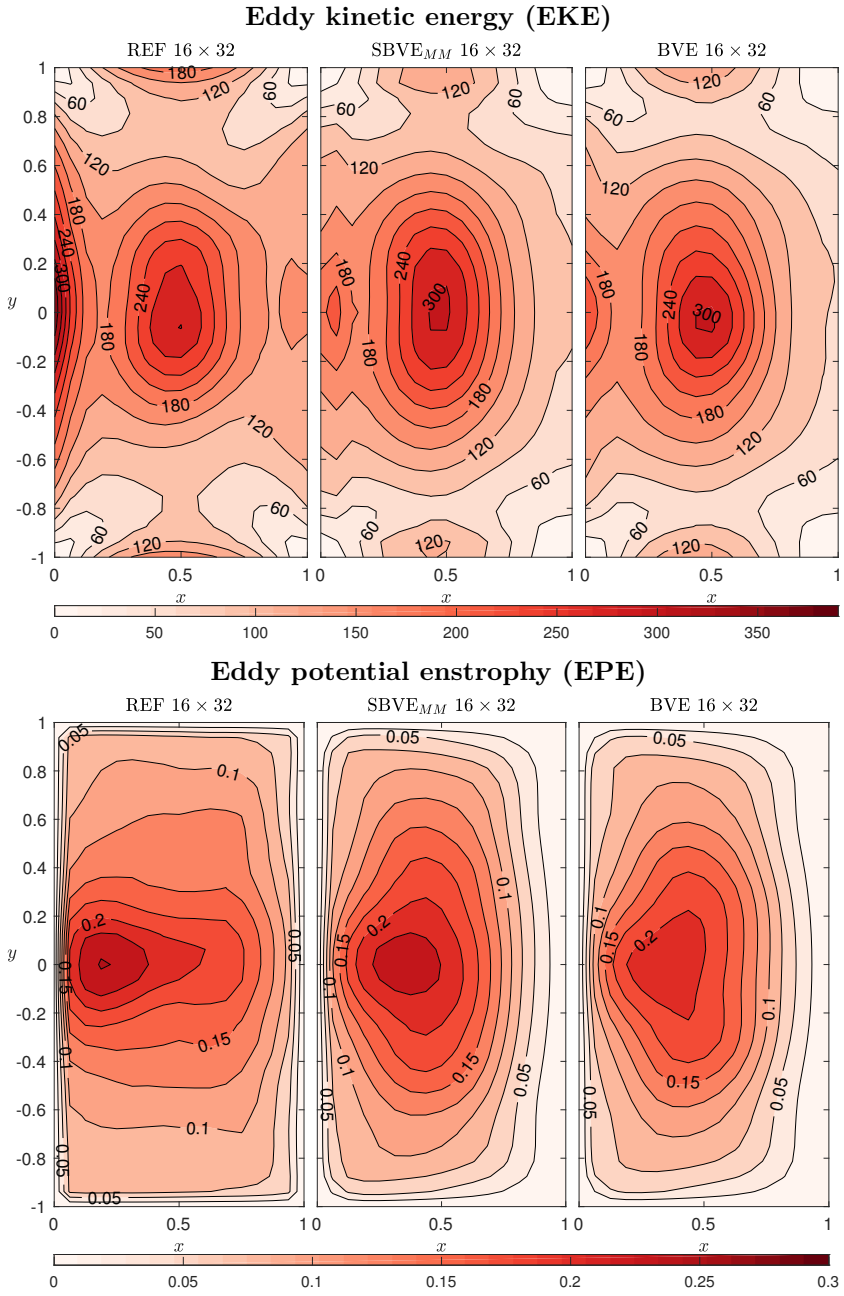

Figure 7.9 – Contour plots of the time-variance fields at resolution 16×32. The top three plots depict EKE with CI of 30, and the bottom three show EPE with CI of 0.025. In each panel, the first one is REF, the second one is SBVE_{MM} and the third one is BVE.

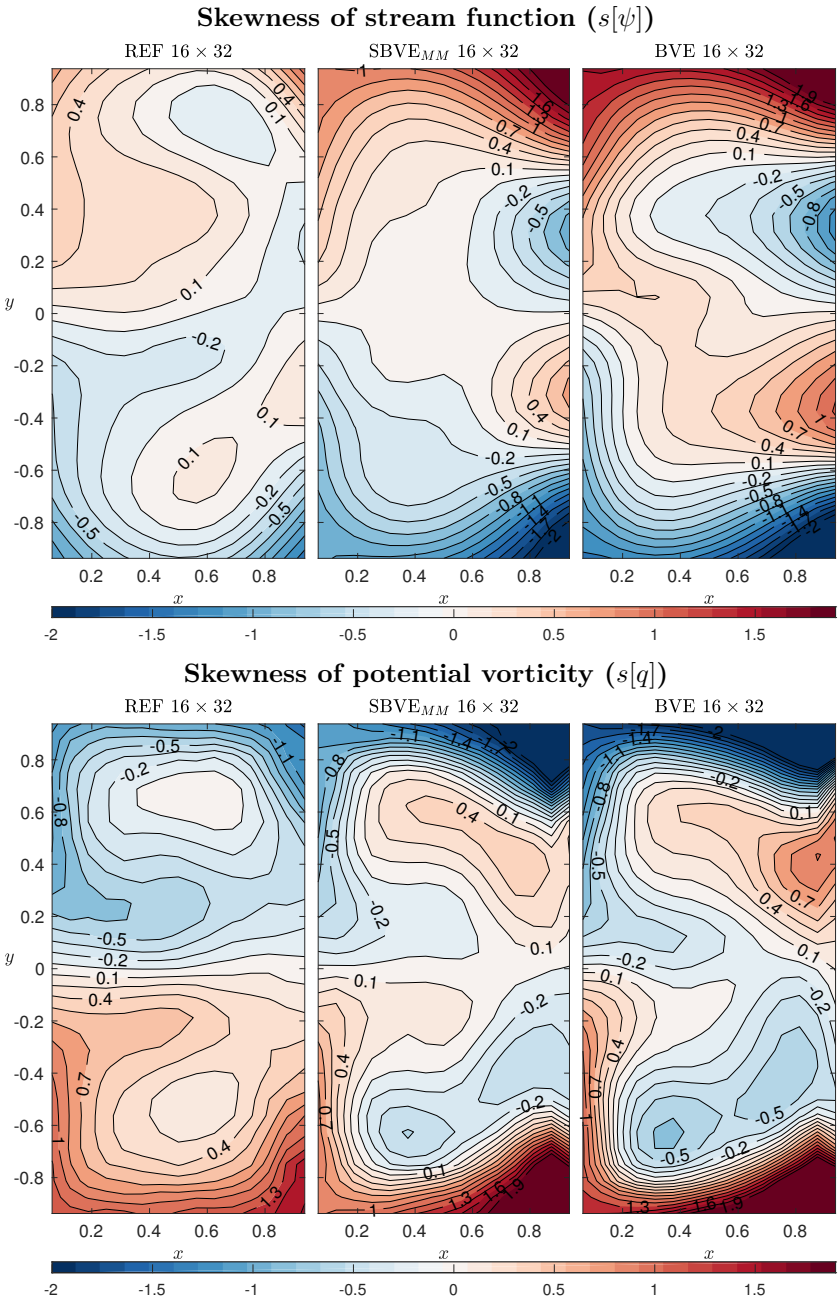

Figure 7.10 – Contour plots of the time-skewness fields at resolution 16 × 32. The top three plots depict third-order SF moment with CI of 0.15, and the bottom three show third-order PV moment with CI of 0.15. In each panel, the first one is REF, the second one is $\mathrm{SBVE}_{\mathrm{MM}}$ and the third one is BVE. The visualized quantity is not defined on the boundary of both fields.

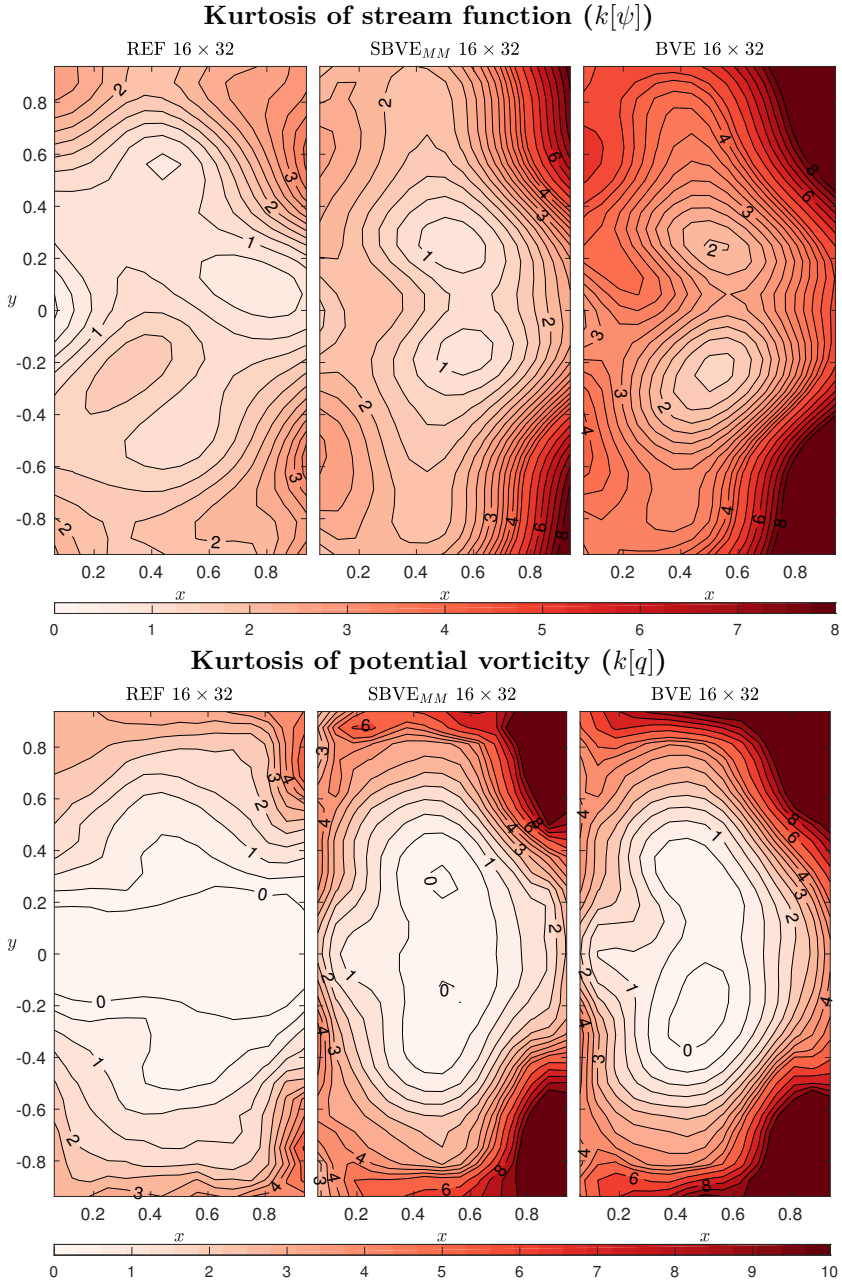

Figure 7.11 – Contour plots of the time-kurtosis fields at resolution 16×32. The top three plots depict fourth-order SF moment with CI of 0.25 within $[0, 4.5]$ and of 0.5 within $[5, 8]$, and the bottom three show fourth-order PV moment with CI of 0.5 within $[0, 4.5]$ and of 1 within $[5, 10]$. In each panel, the first one is REF, the second one is SBVE$_{\text{MM}}$ and the third one is BVE. The visualized quantity is not defined on the boundary of both fields.

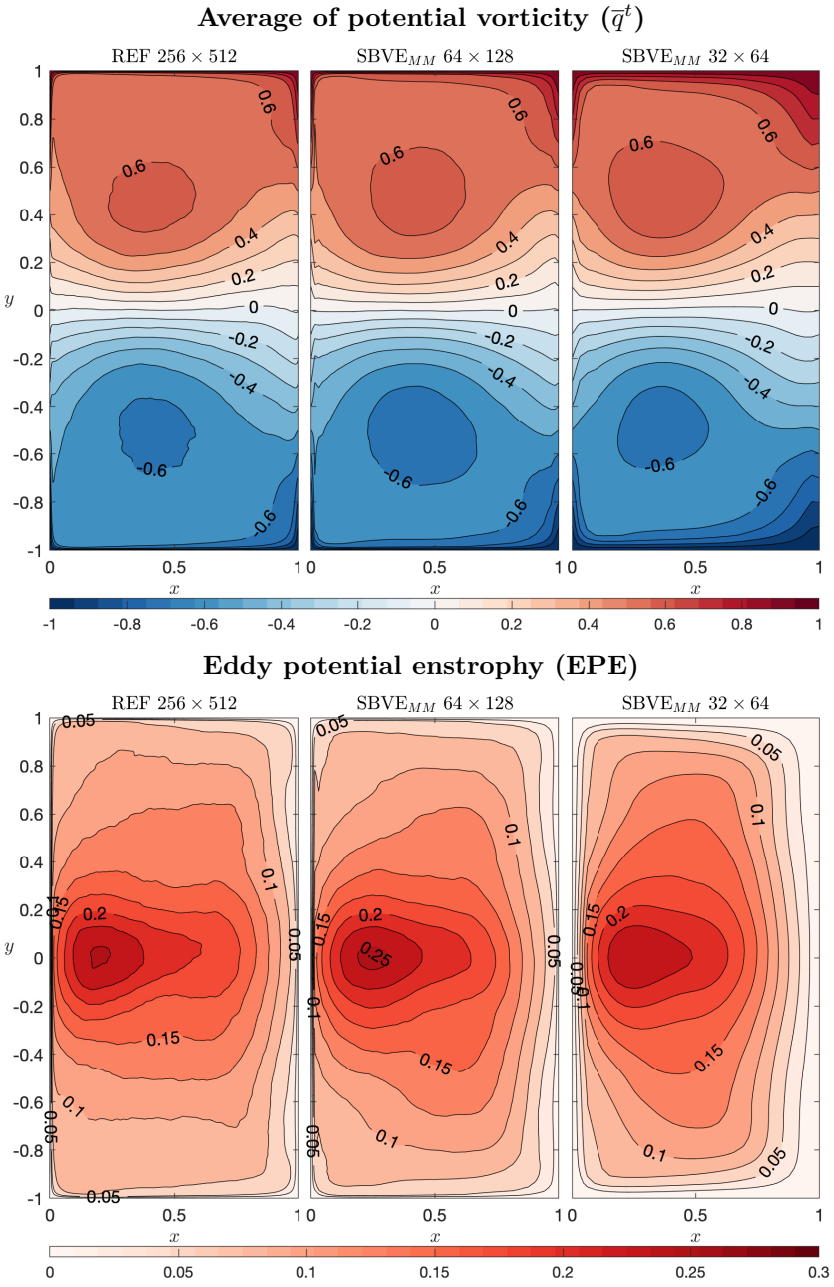

Figure 7.12 – Contour plots showing the qualitative convergence of the statistics for SBVE_{MM}. The top three plots describe the averaged PV with CI of 0.1, and the bottom three show EPE with CI of 0.025. In each panel, the first one stands for BVE 256×512, the second one is SBVE_{MM} 64×128 and the third one is SBVE_{MM} 32×64.

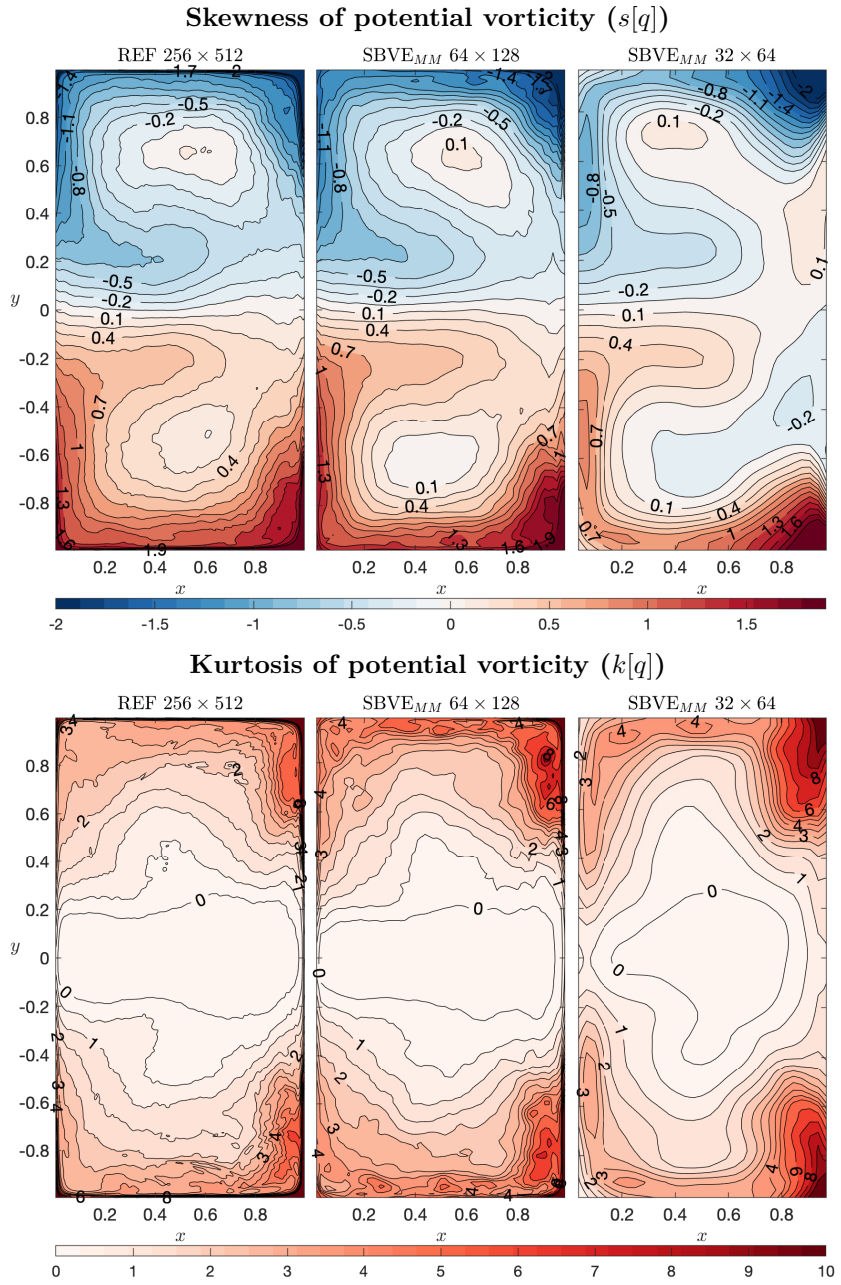

Figure 7.13 – Contour plots showing the qualitative convergence of the statistics for $\mathrm{SBVE}_{\mathrm{MM}}$. The top three plots depict the PV-skewness with CI of 0.15, and the bottom three show the PV-kurtosis with CI of 0.5 within $[0, 4.5]$ and of 1 within $[5, 10]$. In each panel, the first one is BVE 256×512, the second one is $\mathrm{SBVE}_{\mathrm{MM}}$ 64×128 and the third one is $\mathrm{SBVE}_{\mathrm{MM}}$ 32×64. The visualized quantity is not defined on the boundary of both fields.

while for the fourth order moments the improvement is less significant. Both SBVE_{MM} and SBVE_{POD} improve also the first order moments at resolution 32×64 (at almost the same rate as for the coarsest resolution) and 64×128 (with a smaller decreasing of the errors).

$\overline{\text{RMSE}}$ / Model	$\overline{\psi}^t$	$\overline{q}^t$	EKE	EPE	$s[\psi]$	$s[q]$	$k[\psi]$	$k[q]$
BVE	0.245	0.091	0.111	0.148	0.499	0.406	0.782	0.806
SBVE_{POD}	0.221	0.082	**0.097**	**0.132**	0.489	0.390	0.624	0.758
SBVE_{MM}	**0.197**	**0.075**	**0.098**	**0.131**	**0.320**	**0.325**	**0.475**	**0.631**

Table 7.1 – Comparison of the normalized RMSEs between different models at resolution 16×32 with $\text{R}_\beta = 0.06^2$ and $\delta_4 = 0.049$ fixed. The lowest errors are highlighted in bold.

$\overline{\text{RMSE}}$ / Model	$\overline{\psi}^t$	$\overline{q}^t$	EKE	EPE	$s[\psi]$	$s[q]$	$k[\psi]$	$k[q]$
BVE	0.108	0.061	0.073	0.122	0.190	0.166	0.218	0.155
SBVE_{POD}	0.094	**0.056**	0.064	0.116	**0.161**	0.146	**0.182**	0.122
SBVE_{MM}	**0.089**	**0.055**	**0.058**	**0.107**	**0.161**	**0.136**	**0.181**	**0.106**

Table 7.2 – Comparison of the normalized RMSEs between different models at resolution 32×64 with $\text{R}_\beta = 0.06^2$ and $\delta_4 = 0.040$ fixed. The lowest errors are highlighted in bold.

$\overline{\text{RMSE}}$ / Model	$\overline{\psi}^t$	$\overline{q}^t$	EKE	EPE	$s[\psi]$	$s[q]$	$k[\psi]$	$k[q]$
BVE	0.075	0.028	0.036	0.055	0.087	0.039	0.068	0.035
SBVE_{POD}	0.073	**0.024**	0.034	0.047	0.080	0.036	**0.061**	0.031
SBVE_{MM}	**0.069**	**0.023**	**0.027**	**0.041**	**0.068**	**0.034**	**0.061**	**0.029**

Table 7.3 – Comparison of the normalized RMSEs between different models at resolution 64×128 with $\text{R}_\beta = 0.06^2$ and $\delta_4 = 0.026$ fixed. The lowest errors are highlighted in bold.

In addition to the discussions above, it is also important to show if the SBVE on coarse mesh can reproduce the temporal correlation behaviors of the reference (Gugole and Franzke, 2019). To this end, the autocorrelation functions (ACF) for the time series of the global stream function is adopted. More precisely, this ACF is defined as

$$\text{ACF}(\tau) = \frac{\overline{\Big(\Psi(t) - \overline{\Psi}^t\Big)\Big(\Psi(t+\tau) - \overline{\Psi}^t\Big)}^t}{\sigma_\Psi^2}, \tag{7.11}$$

where τ stands for a time-lag, $\Psi(t) = \frac{1}{|\Omega|}\int_\Omega \psi(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x}$ is the global stream function at time t, and σ_Ψ is the (temporal) standard deviation of Ψ. Figure 7.14 shows that compared to the BVE at each coarse resolution, both $\mathrm{SBVE}_{\mathrm{POD}}$ and $\mathrm{SBVE}_{\mathrm{MM}}$ capture better the ACF of the reference. For instance, they have smaller decorrelation time scales compared to the BVE. Besides, the best results are provided by the mode matching method, which is consistent with our previous conclusions.

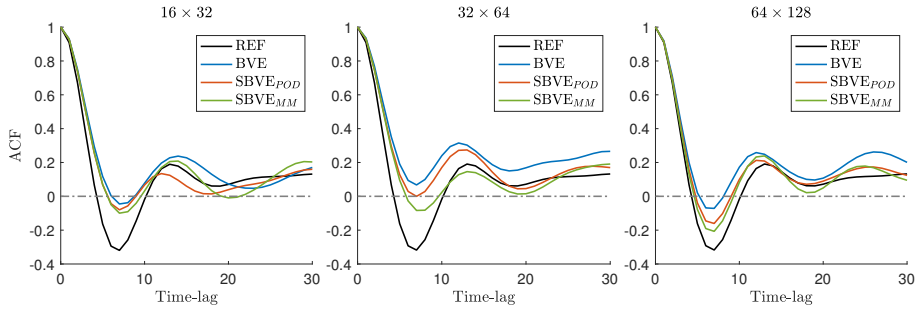

Figure 7.14 – Comparison of the autocorrelation functions (ACF) of the global stream function between different models, at resolution 16×32, 32×64 and 64×128. All the ACFs are calculated from $t = 20$ to $t = 100$.

7.4 Conclusions

The action of the velocity fluctuations towards large-scale flow structuration is first demonstrated on a simple inviscid BQG model. The randomized system still conserves the Rossby wave structure, while introducing secondary vortices. On the contrary, a homogeneous isotropic noise conserves only the primary wave structure. This is a strong indication of the predominant role played by inhomogeneity of the small-scale velocity on shaping coherent large-scale structures in turbulent flows as first wonderfully intuited by Phillips (1977).

In order to quantify the accuracy of the proposed random model, a statistical analysis of the flow tracers has been performed for a single-layered wind-driven QG model. As expected, compared to a classical coarse model, the proposed stochastic model better represents the nonlinearity at the resolved scales while properly dissipating the unresolved scales, leading hence to a balanced correction of excessive dissipation and the continuous increase of internal variability. As a result, it reproduces better on a coarse mesh, the local structures of the distribution of eddy-resolving tracers. Although the idealized barotropic

model used in this work cannot describe quantitatively the real ocean, they do in fact produce qualitatively realistic patterns of large-scale flow in the major basins of the world, as illustrated in Vallis (2017).

Chapter 8

Numerical studies of stochastic multi-layer QG model

Abstract

As we know, unlike the primitive models, an approximative QG model requires only few layers to capture the baroclinic instabilities in practice. However, the effects of mesoscale eddies can still be missing for coarse configurations in the horizontal direction. In this chapter, we continue our studies on the idealized double-gyre circulation yet now with a multi-layer QG model. In particular, we focus on the reproduction of the meandering jet as well as the prediction of low-frequency variability for coarse-grid LU models. The work presented in this chapter is based on the manuscript Li et al. (2021) in preparation for submission.

8.1 Model configurations and simulations

In this study, we apply our stochastic framework to the Quasi-Geostrophic Coupled Model (Q-GCM) described by Hogg et al. (2003), which has been intensively assessed and whose results are well documented in the literature. This Q-GCM differs from many QG models in that it is a coupled model between the ocean and the atmosphere. The model characteristics are illustrated in Figure 8.1. In particular, a mixed layer is included in order to take account the exchanges of heat and momentum between the ocean and the atmosphere. As a result, an explicit diabatic term due to Ekman pumping of sea surface temperature (SST) is involved in the dynamical system.

This work will only focus on the dynamics of ocean component. Nevertheless, the mixed

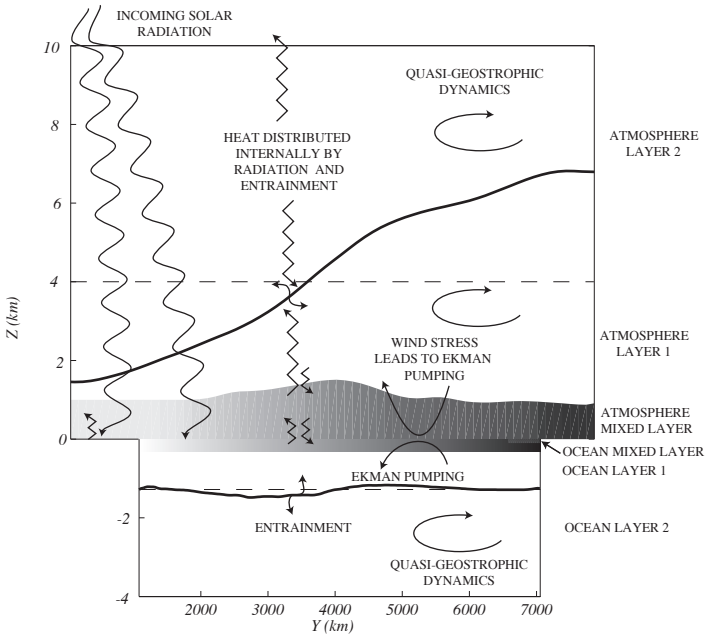

Figure 8.1 – Illustration of the Quasi-Geostrophic Coupled Model (Q-GCM) from Hogg et al. (2003) (see details in `http://www.q-gcm.org`).

layer in the ocean-only mode can still be retained with a simple stationary modeled wind stress. The whole stochastic coupled model will nevertheless be used in future work. Under the LU scheme, the evolution of the mixed layer temperature T_m reads:

$$\mathbb{D}_t T_m = (D_m + F_m)\,\mathrm{d}t, \tag{8.1}$$

where the surface layered noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t,1}$, Itô-Stokes drift $\boldsymbol{u}_{s,1}$ and variance tensor $\boldsymbol{a}_1$, are incorporated in the stochastic transport $\mathbb{D}_t T_m$. The large-scale mixed-layer velocity is given by $\boldsymbol{u}_m = \frac{1}{f_0}\boldsymbol{\nabla}^\perp p_1 + \frac{1}{f_0 H_m}\boldsymbol{\tau}^\perp$. On the RHS of (8.1), the dissipation term is $D_m = K_2\nabla^2 T_m - K_4\nabla^4 T_m$ with K_2 (resp. K_4) the Laplacian (resp. biharmonic) diffusion coefficient. The second term, $F_m = \frac{w_0(T_1+T_m)}{2H_m} - \frac{F_0}{\rho_0 C_p H_m}$, describes the exchange of upper layer heat due to the diabatic effect of Ekman pumping, together with the imposed heat forcing at the surface, where F_0 is the positive upward heat flux and C_p is the specific heat capacity of the ocean.

Once the SST is updated, we then recover a vertical entrainment (or forcing), $w_1 = -\frac{\Delta_m T}{2\Delta_1 T} w_0$, across the mixed layer embedded in the surface layer with the Ekman pumping $w_0 = \frac{1}{f_0}\boldsymbol{\nabla}\times\boldsymbol{\tau}$ with the wind stress $\boldsymbol{\tau}$ *a priori* modeled. Together with a linear drag at the

bottom layer, $w_N = \frac{\delta_{ek}}{2f_0}\nabla^2 p_N$ (with Ekman layer thickness δ_{ek}), the stochastic governing equation of the forced and damped PV of layer $k = 1, \dots, N$ reads

$$\mathbb{D}_t q_k = (D_k + F_k)\,\mathrm{d}t + \mathrm{d}S_{t,k}, \qquad F_k = -\frac{f_0}{H_k}(w_k - w_{k-1}), \tag{8.2}$$

where the source/sink terms are defined in (4.41a) and the dissipation D_k for the vorticity takes the same form as for the SST, *i.e.* $D_k = \frac{A_2}{f_0}\nabla^4 p_k - \frac{A_4}{f_0}\nabla^6 p_k$. The ocean dynamic pressure can be then inverted from the previous updated PV, namely

$$q_k = \frac{1}{f_0}\nabla^2 p_k + \frac{f_0}{H_k}(\eta_k - \eta_{k-1}) + \beta y, \tag{8.3}$$

where η is the perturbed interface height (see Figure 4.1) and the inversion is carried out with a Fast Fourier Transform (FFT) method (see details in Hogg et al. (2003)). In this work, a conservative flux form is adopted for both the evolution of PV (8.2) and of SST (8.1). These are detailed in Appendix 8.3. As for the time integration, we keep the classical Leapfrog scheme for the large-scale variables and then include the additional variations due to the LU flux.

Two types of noise will be studied for this model. The first one is based on the eddy-resolving (5 km) data and built from the off-line learning procedure presented in Section 5.1.1. Let us highlight that the corrected drift derived from the bias ensuing from the coarse-graining process has been included in all the LU coarse models. The following results show the significant improvement in the jet enforcing brought by adding this correction term. Besides, another type of noise defined from a projection is also tested in this work. It is based on similar ideas as those developed for the three-dimensional case (3.54). Under the QG framework, we propose to constraint the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ along the iso-surfaces of the stratification through a layered two-dimensional projector $\mathbf{P}_k$ such that

$$\tilde{\boldsymbol{\sigma}}_k \mathrm{d}\boldsymbol{B}_t = \mathbf{P}_k(\boldsymbol{\sigma}_k \mathrm{d}\boldsymbol{B}_t), \qquad \mathbf{P}_k = \mathbf{I}_2 - \frac{\boldsymbol{\nabla}\Theta_k(\boldsymbol{\nabla}\Theta_k)^T}{\|\boldsymbol{\nabla}\Theta_k\|^2}, \qquad \Theta_k = \frac{f_0}{H_k}(\eta_k - \eta_{k-1}). \tag{8.4}$$

The common parameters for all the simulations are listed in Table 8.1, whereas the parameters corresponding to different configurations are specified separately in 8.2. In Hogg et al. (2004), a set of eddy-resolving simulations of the wind-driven circulation have been presented to diagnose the variability of the model in terms of the parameters. In this work, our aim is to improve the variability of low-order models limited by low Reynolds

numbers associated to coarse grids, particularly in the case that the baroclinic instability could not be resolved. To that end, several coarse-resolution (at least the same order of the deformation radius) models are simulated and compared. Before this, let us first explain the two types of eddy-resolving data on which we rely. As illustrated by Figure 8.2, the dynamics without including the SST evolution is mainly characterized by a long meandering jet in each layers. On the other hand, as shown in Figure 8.3, the inclusion of the additional forcing coming from the SST evolution produces much more variabilities inside the gyres. Furthermore, an opposite PV gradient is observed in the second layer. Both of these two configurations are then interesting for our study of coarse-resolution models. Since we can simply focus on the ability of different coarse models to reproduce the jet with the first configuration and diagnose more sepcifically the variability inside the gyres using the latter test case.

Parameters	Value	Description
$X \times Y$	(3840×4800) km	Domain size
H_k	$(350, 750, 2900)$ m	Mean layer thickness
H_m	100 m	Mixed layer thickness (fixed)
T_k	$(287, 282, 276)$ K	Potential temperature structure
ρ	1000 kg m^3	Density
g'_k	$(0.025, 0.0125)$ m s^{-2}	Reduced gravity
δ_{ek}	2 m	Bottom Ekman layer thickness
τ_0	2×10^{-5} m^2 s^{-2}	Wind stress magantitude
α_{bc}	0.2	Mixed boundary condition coefficient
A_2	0 m^2 s^{-1}	Laplacian viscosity coefficient
K_2	100 m^2 s^{-1}	Temperature diffusion coefficient
f_0	9.375×10^{-5} s^{-1}	Mean Coriolis parameter
β	1.754×10^{-11} (m s)$^{-1}$	Coriolis parameter gradient
L_d	$(39, 22)$ km	Baroclinic Rossby radii

Table 8.1 – Common parameters for all the models.

Resolution Δx (km)	Timestep Δt (s)	Viscosity (m^4 s^{-1})
5	600	2.0×10^9
40	1200	5.0×10^{11}
80	1440	5.0×10^{12}
120	1800	1.0×10^{13}

Table 8.2 – Values of grid varying parameters. Note that the value of K_4 for SST is fixed to the same value as the biharmonic coefficient A_4.

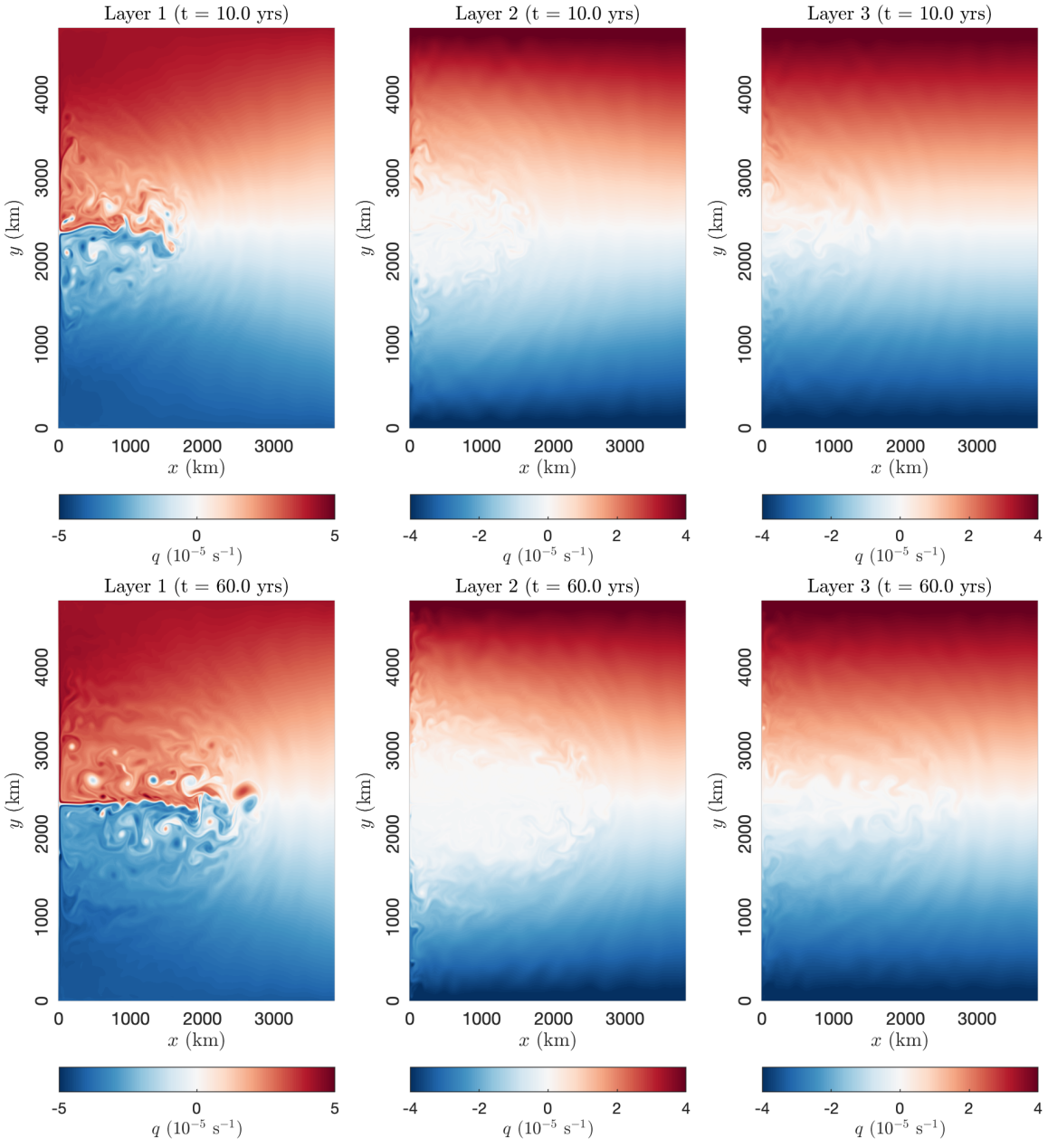

Figure 8.2 – Instantaneous (Year 60) snapshot of PV of the eddy-resolving (5 km) simulation without SST evolution.

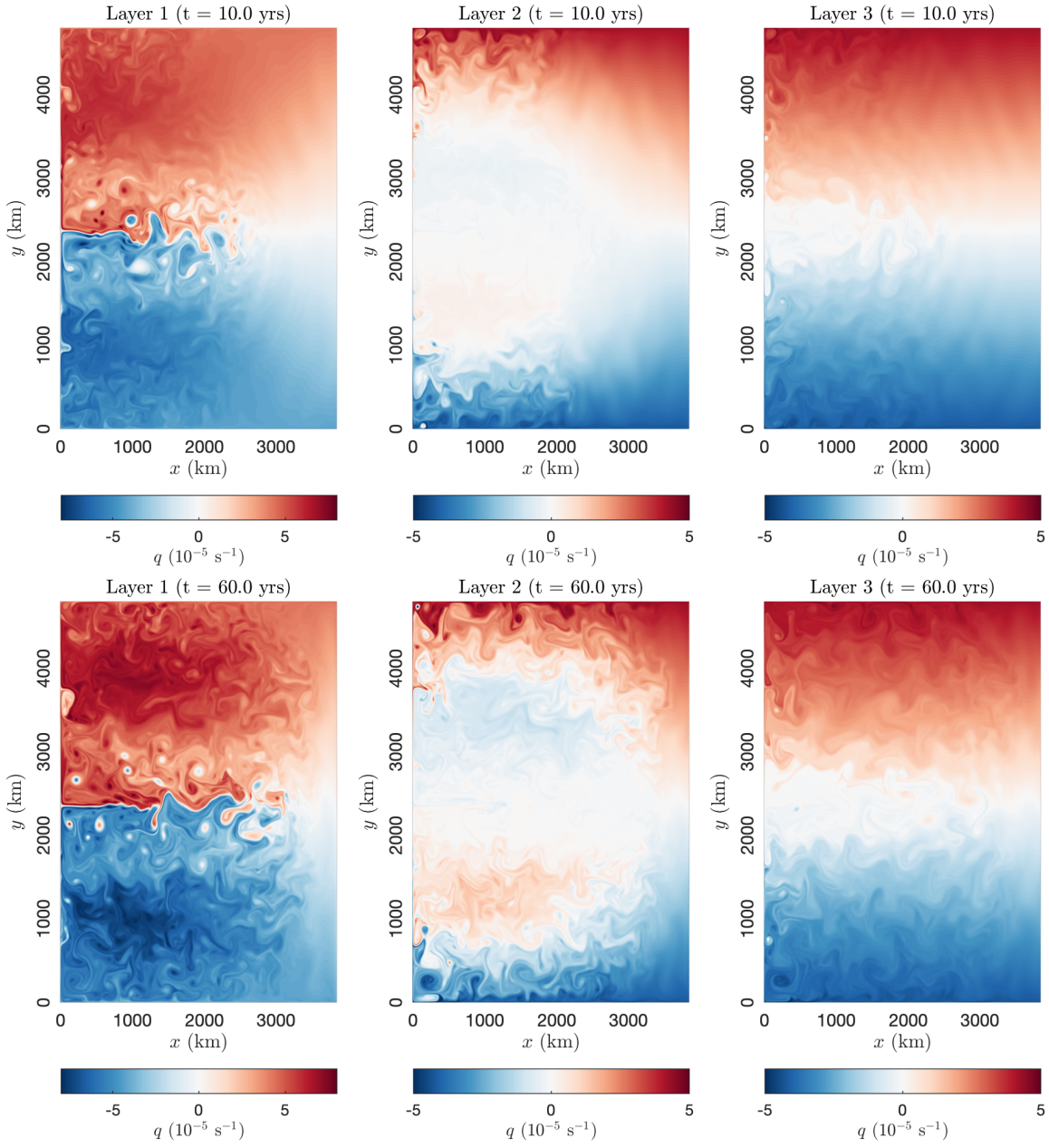

Figure 8.3 – Instantaneous (Year 60) snapshot of PV of the eddy-resolving (5 km) simulation with SST evolution.

For the first configuration, the two components of the POD (velocity) noise in the upper layer are illustrated on the top of Figure 8.4, whereas the projection-based noise is shown on the bottom. Note that in both cases, the correction drift due to the bias associated to the coarse-graining process (see Section 5.1.1) is included. We observe the structure of such noises is mostly localized around the jet area and along the western boundary, which are well informed from the fast fluctuations of the eddy-resolving data. Besides, from the definition of the projection (8.4), the latter noise corresponds to a vertically linear combination of the POD noise, hence it includes combination of barotropic and baroclinic structures. This will be further detailed latter. With such noises, some snapshots of PV provided by the LU coarse simulations (40 km and 80 km) are shown in Figure 8.5. In addition, to enable a simple first comparison, the results of corresponding (*i.e.* with the same parameters) deterministic low-resolution (denoted as LR in the following) simulations are included. As we can see, both LU coarse models enable us to reproduce the zonal jet and the meridional perturbation of the zero contour. The latter result relies majorly on the non-linear eddies effect. On the other hand, the LR simulations produce a symmetric structure *w.r.t.* the basin center. In particular, at the 80 km resolution, we observe only a dominant wave structure stuck to the western boundary layer due to an over-dissipation effect. The deterministic coarse resolution flows are almost stationary. The LU-POD-P model produces more small-scale fluctuations in many places at each coarse resolution. Later, the jet enhancement effect in LU coarse models will be diagnosed and analyzed more precisely.

Similarly for the second configuration, the POD noise with its projected version are build from the data with the SST evolution. As shown in Figure 8.6, these noises are stronger than those in the previous configuration, as the entrainment driven by the evolution of the SST increases the variability of the eddy-resolving data, hence more local fluctuations are transferred to the noise in each layer. Subsequently incorporating these uncertainties in the proposed random dynamical model (8.2) together with the stochastic transport of SST (8.1), some very interesting results are recovered, as demonstrated in Figure 8.7. First, the improvements seen the previous test case still hold here, namely the zonal jet is reproduced and perturbed in both LU coarse models. In addition, the variability of the gyres seems also to be enhanced in this case. The instantaneous structures of the PV field from LU coarse simulations are much closer to the reference eddy-resolving snapshots 8.3, compared to the deterministic LR model.

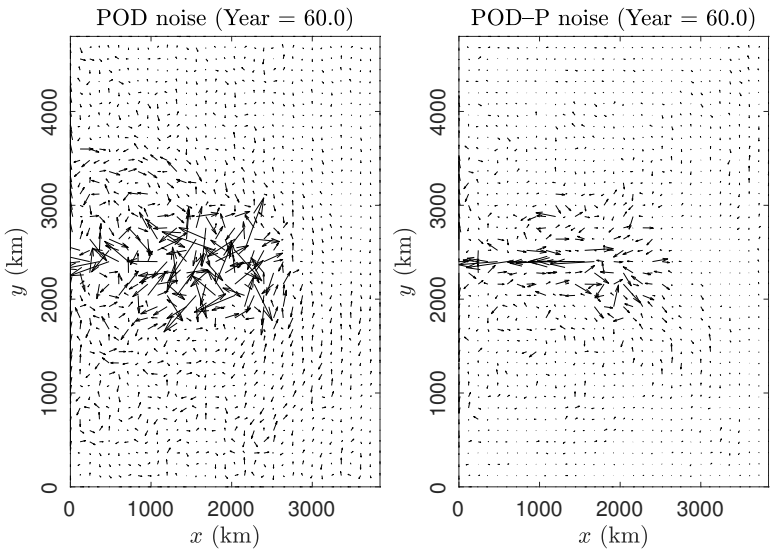

Figure 8.4 – Instantaneous (Year 60) snapshots of upper layer velocity noises (40 km). The LU-POD is learned from eddy-resolving data without SST evolution. The LU-POD-P is build from the projection (8.4) of the LU-POD. The lengths of arrows are adjusted by a factor of 3.

8.2 Diagnostic of low-frequency variability

In this section, we explore in depth the prediction of the low-frequency variability for the LU coarse models. To this end, we perform our studies from two points of view: statistical predictions and energetic analyzes. The first one is inspired from the previous study presented in Section 7.3.2 by qualifying some statistical structures of interest such as the temporal mean, and by quantifying some meaningful statistical measures such as the RMSE of some prognostic variables. The latter one is inspired from the works of Hogg et al. (2004); Hogg and Blundell (2006) with some diagnostics of eddies energy and of energy transfers. The two subsequent sections describe these aspects respectively. For a comparison reason, we keep the classical coarse model (LR) as a benchmark.

8.2.1 Statistical predictions

This section describes the qualitative and the quantitative predictions of the long-term statistics provided by the LU coarse models. Unlike the barotropic test case presented in Section 7.3.2, we only focus on the temporal mean for the multi-layer QG system. However, we will not analyze the results layer by layer. Instead, we prefer to perform

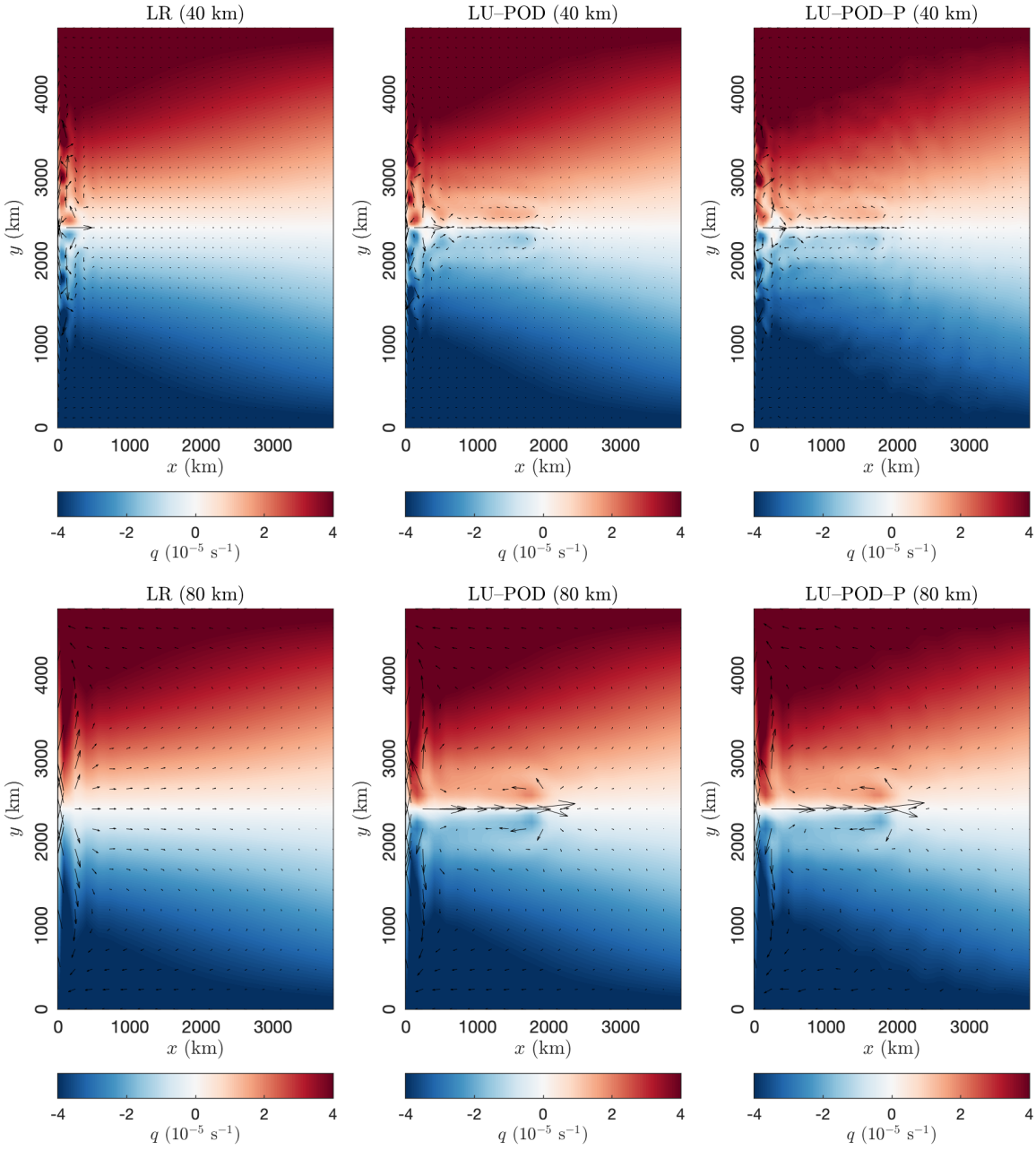

Figure 8.5 – Comparison of instantaneous (Year 60) snapshot of upper-layer PV for coarse-model simulations without SST evolution. The lengths of arrows are adjusted by a factor of 3.

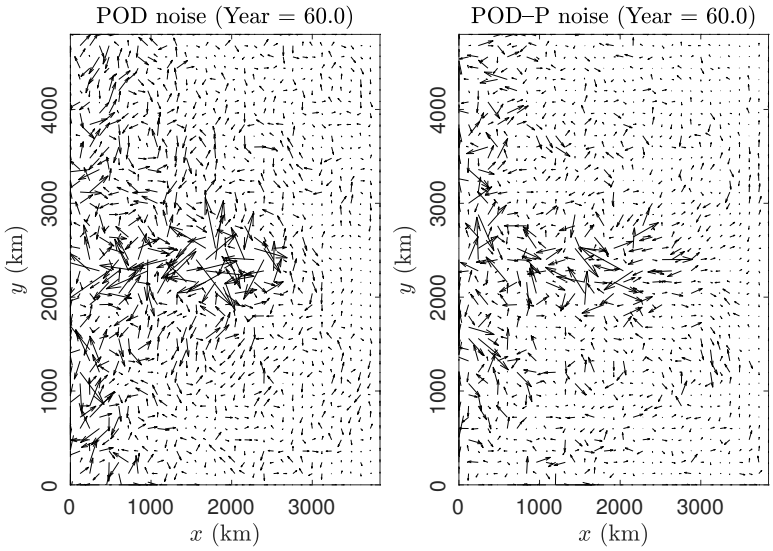

Figure 8.6 – Instantaneous (Year 60) snapshots of upper layer velocity noises (40km). The LU-POD is learned from eddy-resolving data with SST evolution. The lengths of arrows are adjusted by a factor of 3.

a vertical decomposition, since qualifying the barotropic and baroclinic effects are more meaningful in such a system.

As described by Hogg et al. (2003), we can re-formulate the N-layered QG model in vector form with $\boldsymbol{q} = (q_1, \ldots, q_N)^T$ and $\boldsymbol{p} = (p_1, \ldots, p_N)^T$ such that

$$f_0 \boldsymbol{q} = \nabla^2 \boldsymbol{p} + f_0 \beta y - f_0^2 \boldsymbol{A} \boldsymbol{p}, \tag{8.5}$$

where $\boldsymbol{A} \in \mathbb{R}^N \times \mathbb{R}^N$ denotes the vertical stretching tensor defined by

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{H_1 g_1'} & \frac{-1}{H_1 g_1'} & & & \\ \frac{-1}{H_2 g_1'} & \frac{1}{H_2}\left(\frac{1}{g_2'} + \frac{1}{g_1'}\right) & \frac{-1}{H_2 g_2'} & & \\ & \ddots & \ddots & \ddots & \\ & & & & \\ & & & \frac{-1}{H_N g_{N-1}'} & \frac{1}{H_N g_{N-1}'} \end{bmatrix}. \tag{8.6}$$

We next assume that the tridiagonal matrix $\boldsymbol{A}$ admits N right eigenvectors $\boldsymbol{R}_m$, N left

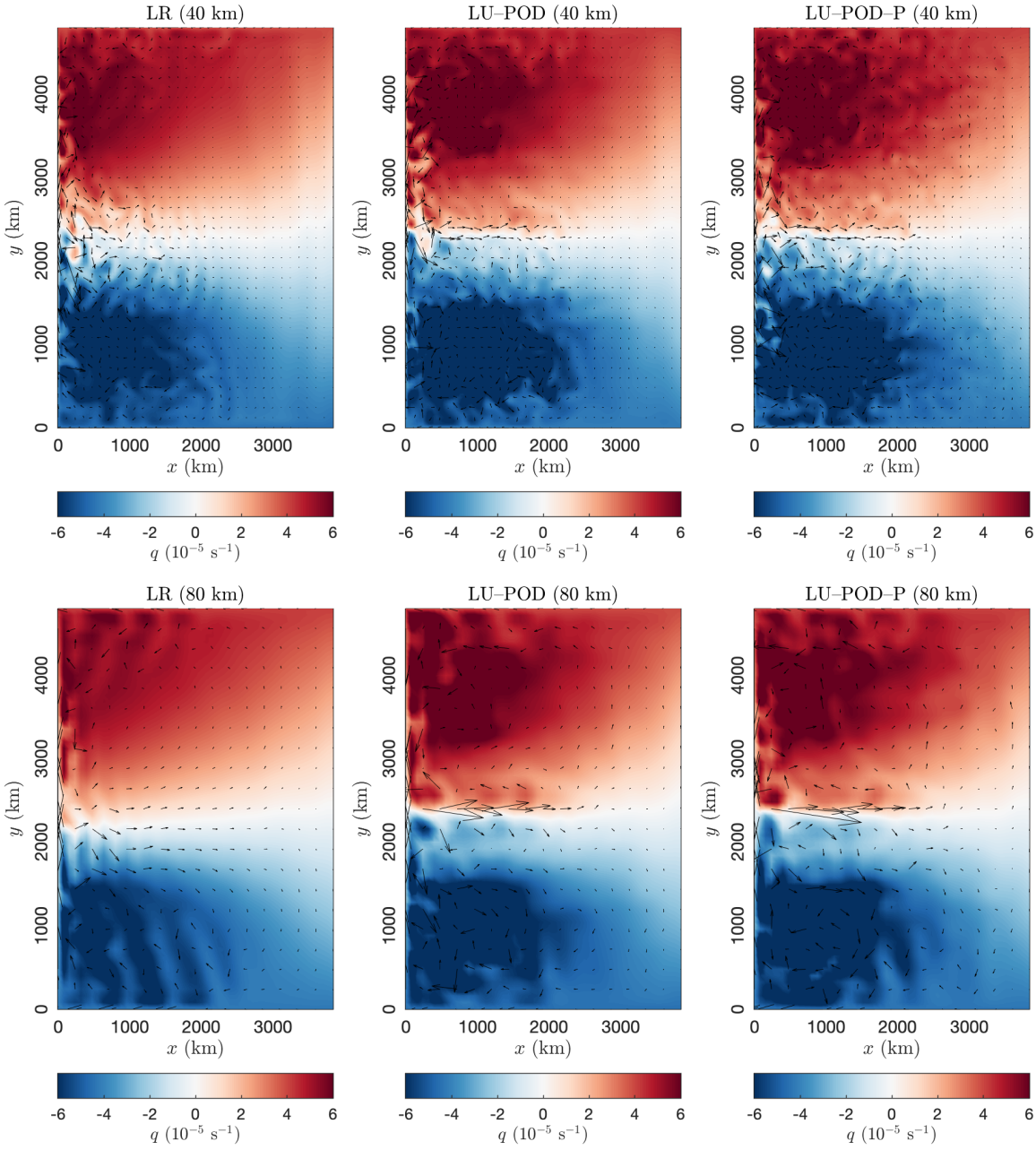

Figure 8.7 – Comparison of instantaneous (Year 60) snapshot of upper-layer PV for coarse-model simulations with SST evolution. The lengths of arrows are adjusted by a factor of 3.

eigenvectors $\boldsymbol{L}_m^T$ and N eigenvalues λ_m $(m = 1, \ldots, N)$ such that

$$\boldsymbol{A}\boldsymbol{R}_m = \lambda_m \boldsymbol{R}_m, \qquad \boldsymbol{L}_m^T \boldsymbol{A} = \lambda_m \boldsymbol{L}_m^T, \tag{8.7}$$

where $\boldsymbol{L}_m$ and $\boldsymbol{R}_m$ are orthogonal, *i.e.* $\boldsymbol{L}_m^T \cdot \boldsymbol{R}_m$ when $m \neq n$. However, $\boldsymbol{L}_m^T \neq \boldsymbol{R}_m$ in general, since $\boldsymbol{A}$ is not symmetric. From the structure of matrix $\boldsymbol{A}$, all the eigenvalues are real and positive, *i.e.* $\lambda_m \geq 0$. Note that the dimensions of λ_m are $1/c_m^2$ with the phase speed c_m and the dimensions of $f_0^2\lambda_m$ are $1/r_m^2$ with the deformation radius r_m. In particular, $\lambda_1 = 0$ corresponds to the barotropic mode and $\lambda_m,\ m > 1$ to the $(m-1)$-th baroclinic mode.

We then expand the layered pressure vector $\boldsymbol{p}$ in terms of the right eigenvectors $\boldsymbol{R}_n$. It reads

$$\boldsymbol{p} = \sum_{n=1}^{N} \tilde{p}_n \boldsymbol{R}_n, \tag{8.8}$$

where $\tilde{p}_n$ denotes the vertical mode coefficients. Multiplying Equation (8.8) by $\boldsymbol{L}_m^T$, we have

$$\boldsymbol{L}_m^T \cdot \boldsymbol{p} = \sum_{n=1}^{N} \tilde{p}_n \boldsymbol{L}_m^T \cdot \boldsymbol{R}_n = \tilde{p}_m (\boldsymbol{L}_m^T \cdot \boldsymbol{R}_m), \tag{8.9}$$

which leads to

$$\tilde{p}_m = \frac{\boldsymbol{L}_m^T}{\boldsymbol{L}_m^T \cdot \boldsymbol{R}_m} \cdot \boldsymbol{p}. \tag{8.10}$$

Let us define the mode-to-layer conversion matrix $\boldsymbol{C}_{m2l}$ and the layer-to-mode conversion matrix $\boldsymbol{C}_{l2m}$ respectively by

$$C_{m2l}(l,m) \triangleq \left[c_{l,m}\right] = R_m(l), \qquad C_{l2m}(m,l) \triangleq \left[\tilde{c}_{m,l}\right] = \frac{L_m^T(l)}{\boldsymbol{L}_m^T \cdot \boldsymbol{R}_m}. \tag{8.11}$$

As a result, one can convert bidirectionally the pressure vector between the layer's and mode's coordinate by

$$\tilde{\boldsymbol{p}} = \boldsymbol{C}_{l2m}\boldsymbol{p}, \qquad \boldsymbol{p} = \boldsymbol{C}_{m2l}\tilde{\boldsymbol{p}}, \tag{8.12}$$

or, equivalently written in component form by

$$\tilde{p}_m = \sum_{l=1}^{N} \tilde{c}_{m,l} p_l, \qquad p_l = \sum_{m=1}^{N} c_{l,m} \tilde{p}_m. \tag{8.13}$$

Therefore, the time-averaged modal pressure $\overline{\tilde{\boldsymbol{p}}}^t$ equals to the projection coefficient of the

time-averaged layered pressure $\overline{\boldsymbol{p}}^t$:

$$\overline{\tilde{p}}_m^t = \sum_{l=1}^{N} \tilde{c}_{m,l} \overline{p}_l^t. \tag{8.14}$$

Hereafter, the time average of the modal pressures for the LU-POD and LU-POD-P models as well as the LR models without the SST evolution are computed from year 60 to 180 using 2920 snapshots. In particular, a directly subsampled temporal mean of the eddy-resolving data is considered as the reference (REF) in the following. From Figure 8.8, we observe from the reference that both barotropic and baroclinic modes characterize the meandering jet. Then, we can clearly see that the LU-POD and LU-POD-P at 40 km enable to reproduce qualitatively the local structures of both vertical modes, predicted by the reference. Conversely, the classical LR model can only capture the stationary solution – the symmetric double-gyre. In fact, the two gyres are in that case completely separated. We remark that these results also hold for resolutions 80 km and 120 km, as demonstrated in Figures 8.9 and 8.10. We believe these results are very encouraging within the perspective of global ocean models run at climatic scales (>100 km). Subsequently, an analogue procedure is performed for all the coarse-resolution models in the configuration including the evolution of SST. For instance, the results of the barotropic mode and of the 1st baroclinic mode (at 40 km, 80 km and 120 km) are provided by Figures 8.11, 8.12, 8.13 and a conclusion consistent with the previous test case is recovered.

Afterwards, we quantify the variability of the coarse models using some metrics proposed by Grooms et al. (2015). In order to evaluate the accuracy of the time-mean $\overline{f}^t$ and the standard derivation σ_f for a variable of interest f, one can first adopt their root-mean-squares (RMS), namely

$$\text{RMS of } \sigma_f = \Big(\frac{1}{|\Omega|}\int_\Omega \sigma_f^2 \,\mathrm{d}\boldsymbol{x}\Big)^{1/2}, \tag{8.15a}$$

where $|\Omega|$ denotes the domain size. Subsequently, with a given reference standard derivation $\sigma_{f_{\text{ref}}}$ (in our case, we considered the one defined from a direct subsampling of the high-resolution standard derivation), one may quantify the root-mean-square errors (RMSE) between σ_f and $\sigma_{f_{\text{ref}}}$, that is

$$\text{RMSE of } \sigma_f = \Big(\frac{1}{|\Omega|}\int_\Omega (\sigma_f - \sigma_{f_{\text{ref}}})^2 \,\mathrm{d}\boldsymbol{x}\Big)^{1/2}. \tag{8.15b}$$

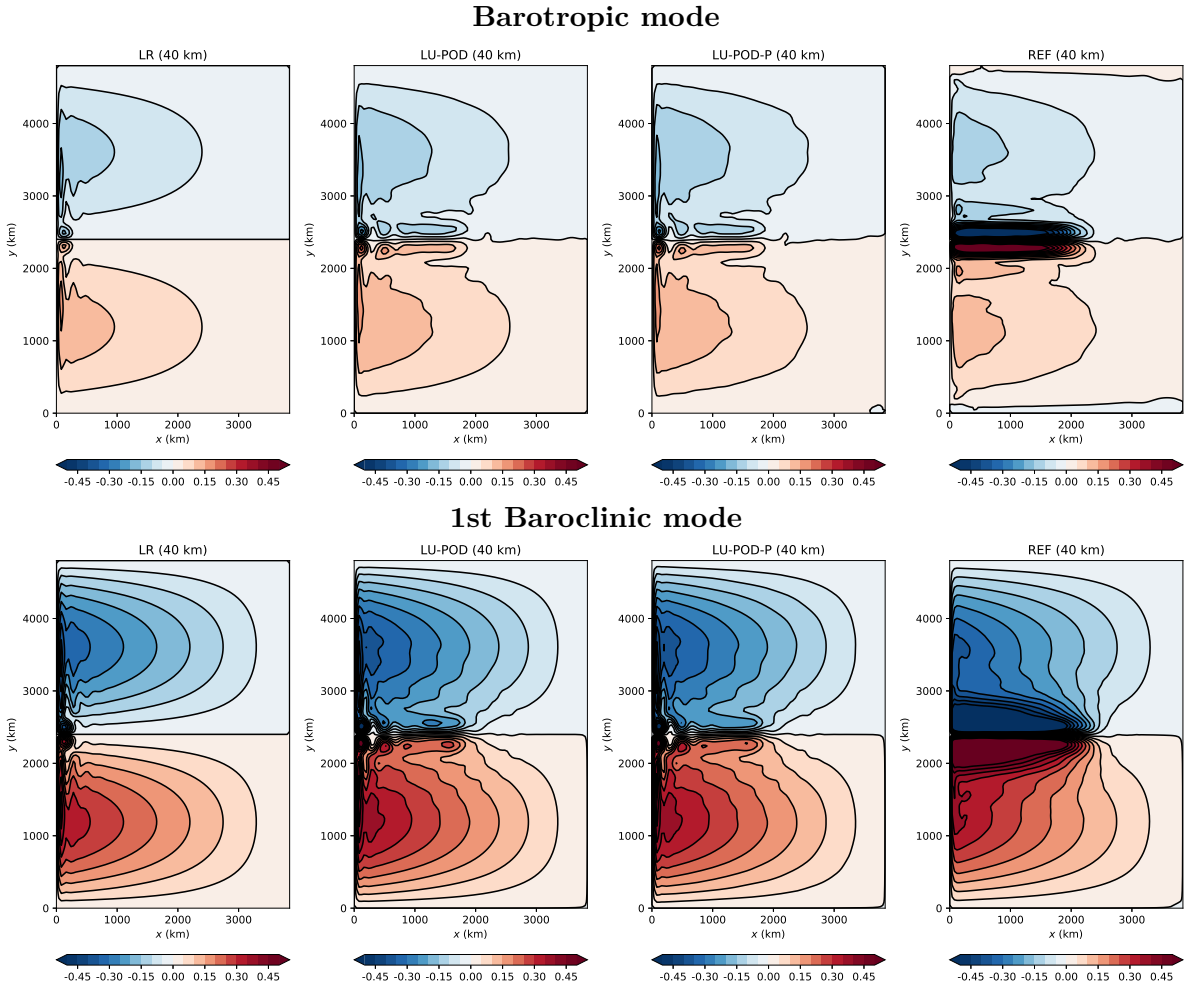

Figure 8.8 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (40 km) simulations without SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.05 (m^2s^{-2}).

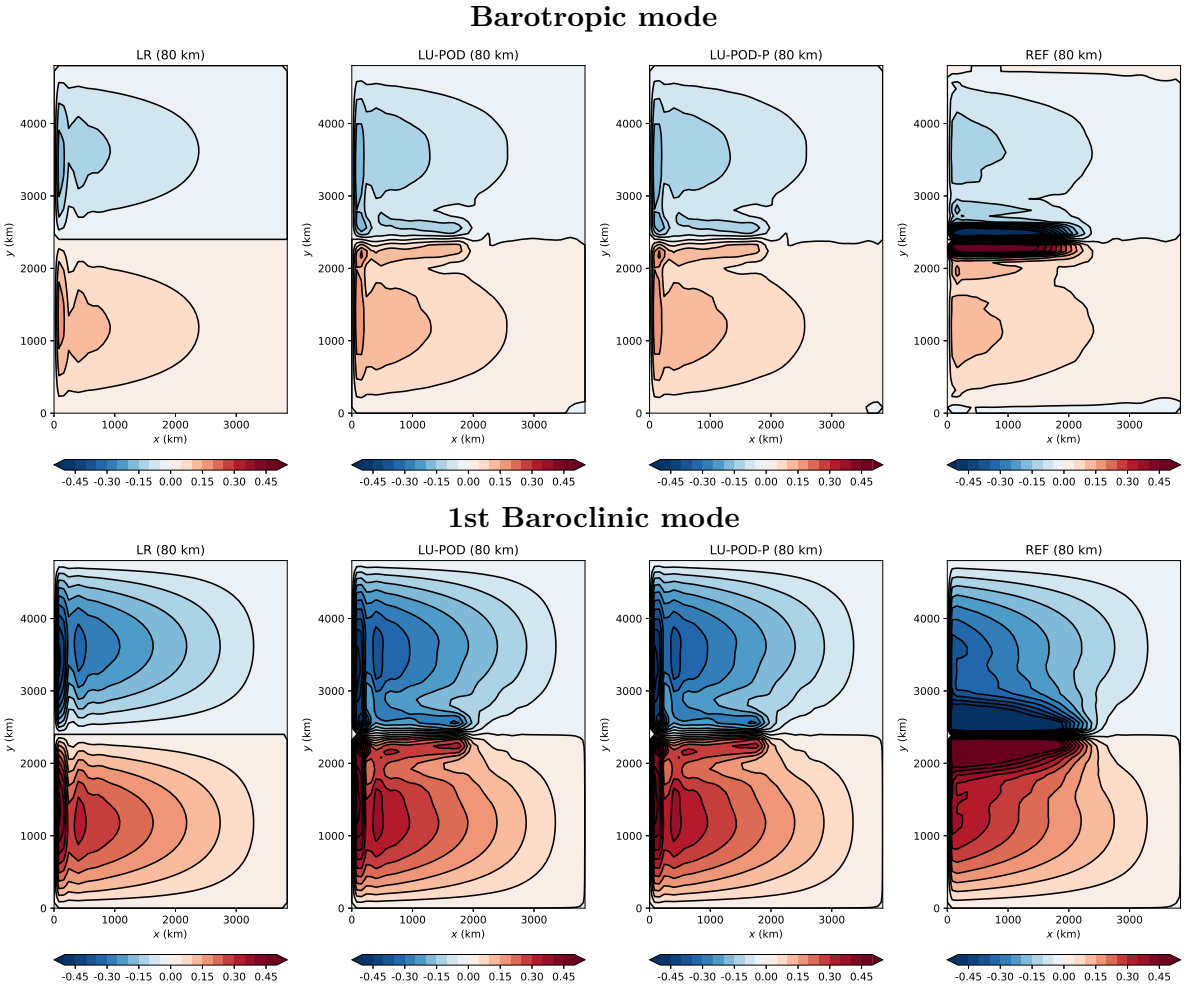

Figure 8.9 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (80 km) simulations without SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.05 (m^2s^{-2}).

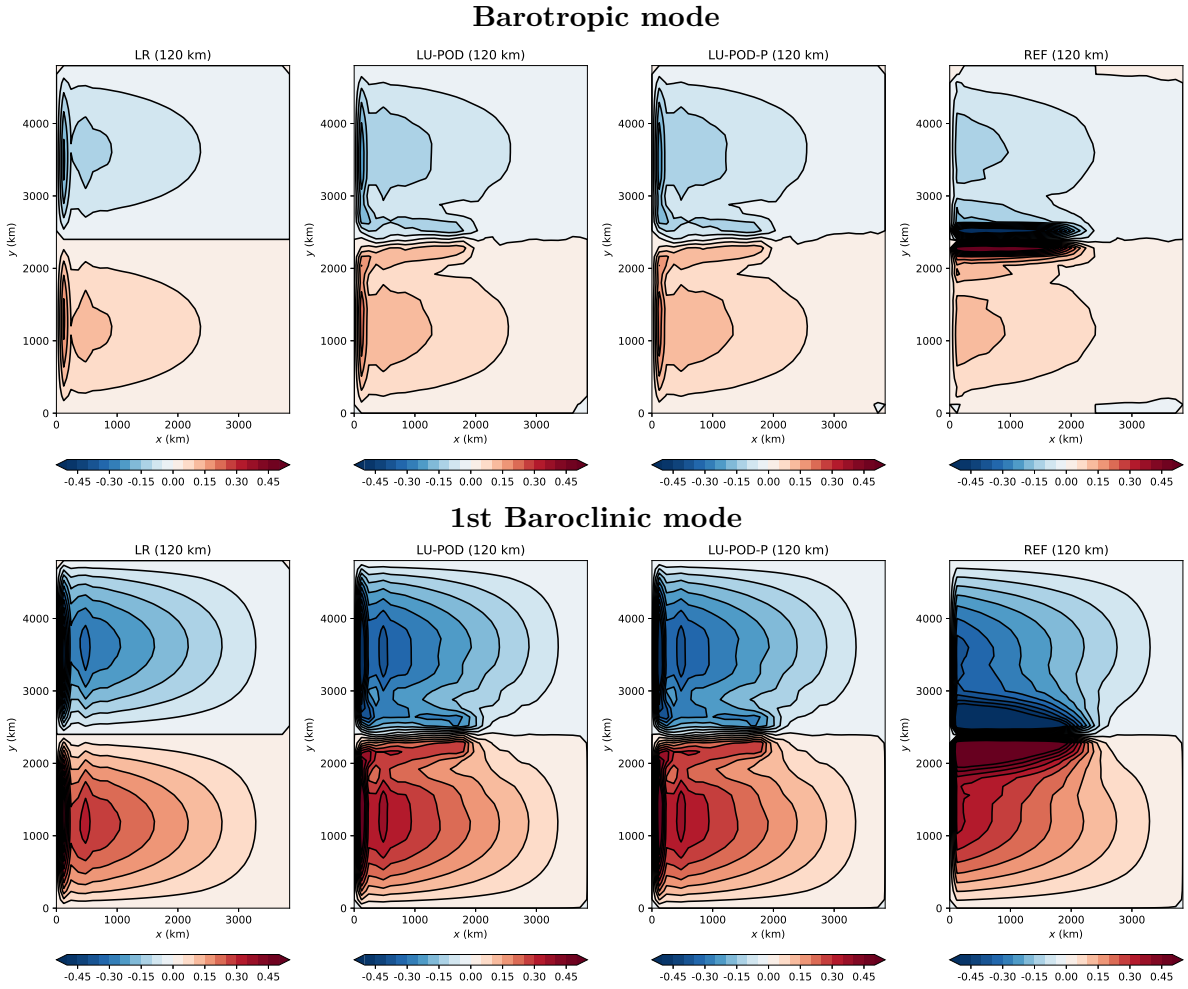

Figure 8.10 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (120 km) simulations without SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.05 (m^2s^{-2}).

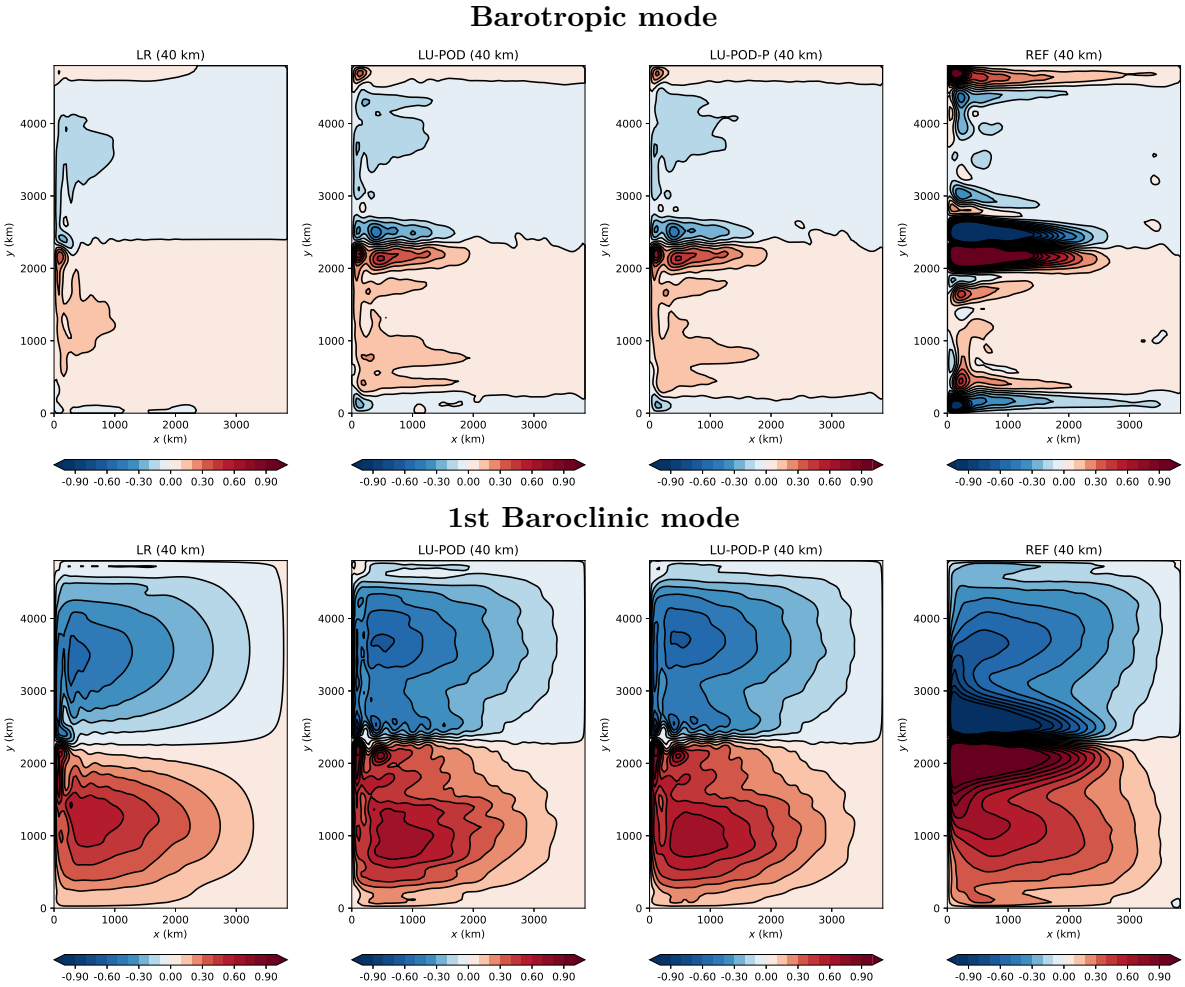

Figure 8.11 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (40 km) simulations with SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.1 (m^2s^{-2}).

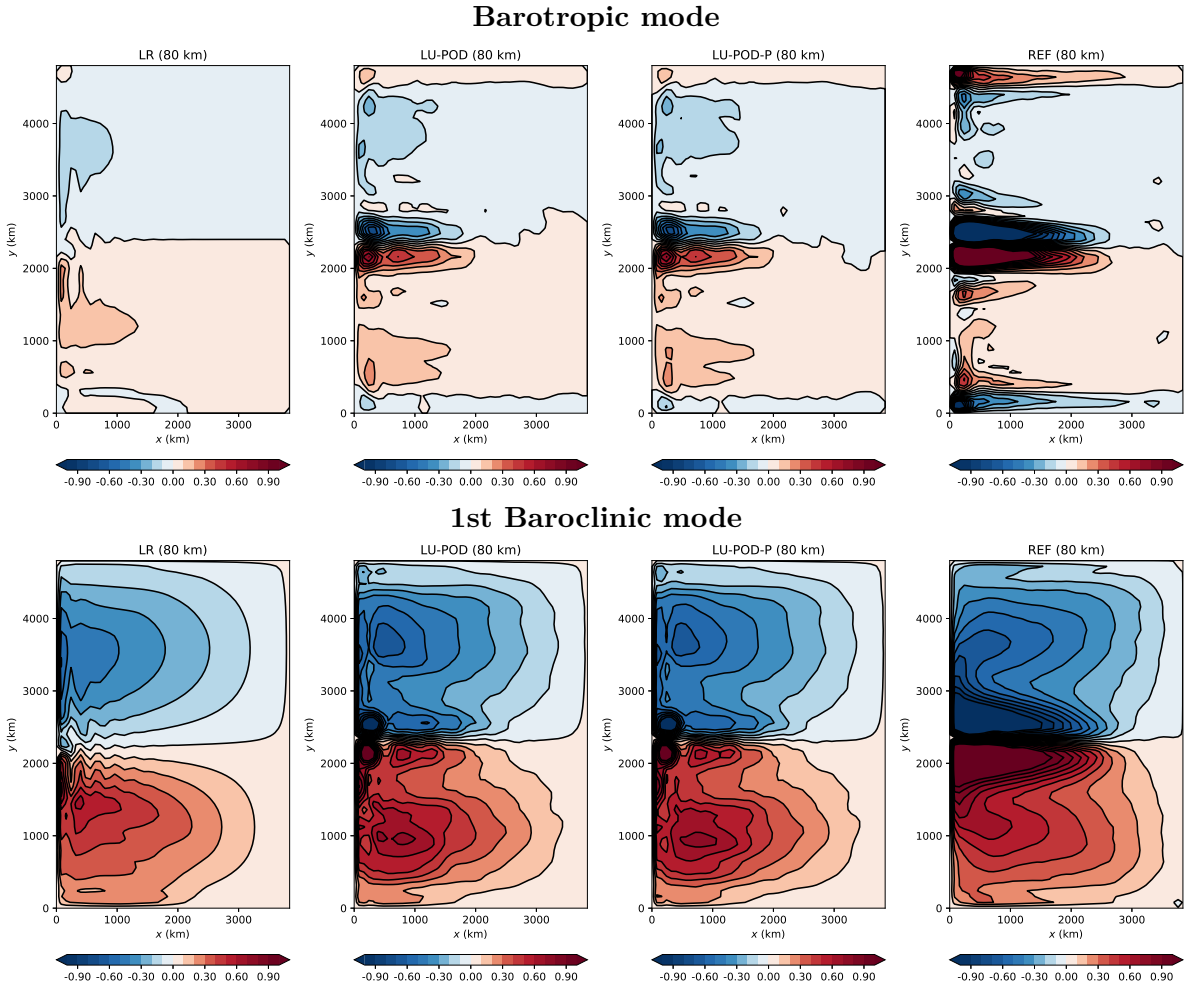

Figure 8.12 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (80 km) simulations with SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.1 (m^2s^{-2}).

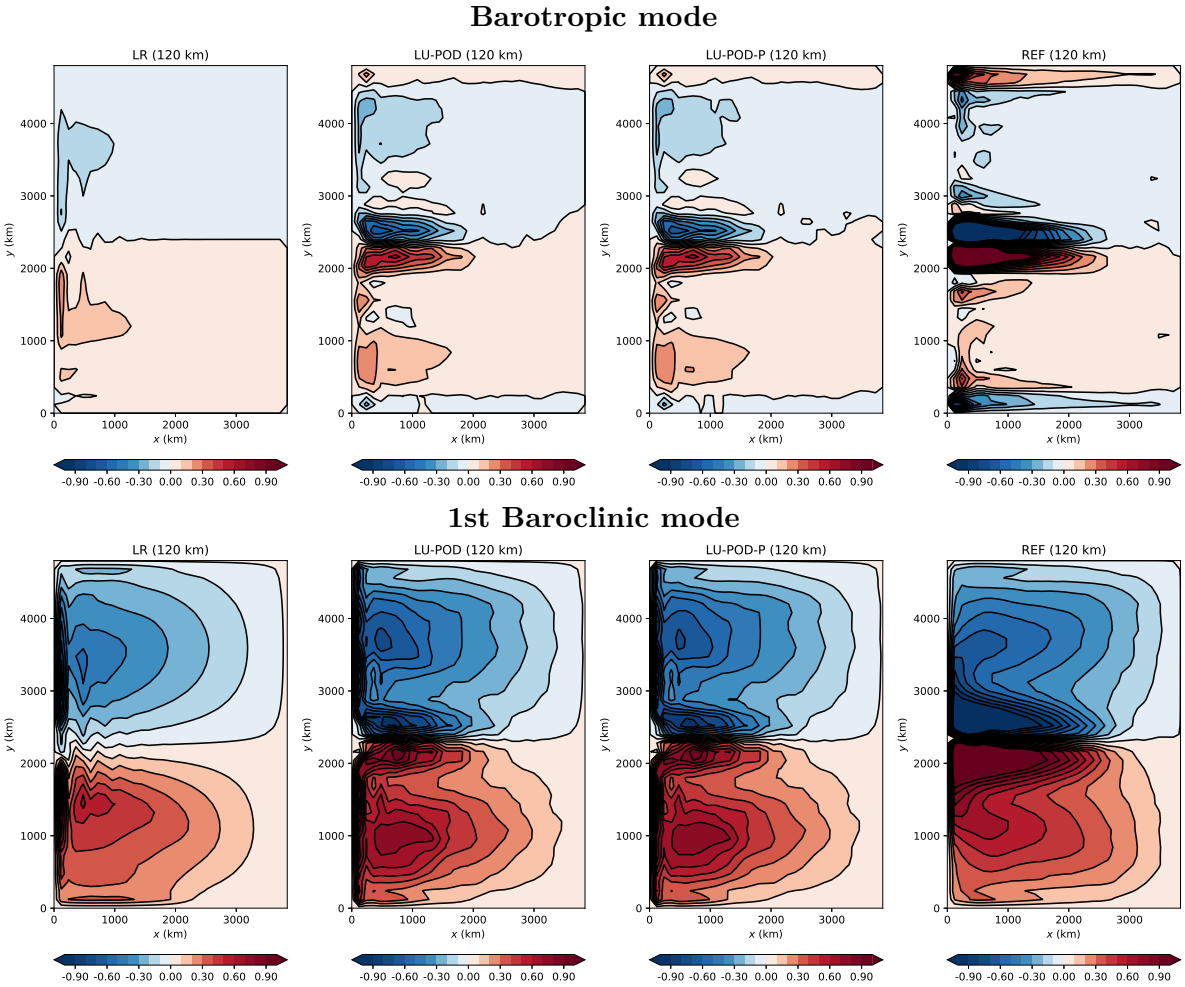

Figure 8.13 – Comparison of mean (from 60 yrs to 180 yrs) contour of vertical pressure modes for coarse-model (120 km) simulations with SST evolution. The matplotlib.colors.SymLogNorm class is used with linthresh fixed to be 0.3 and the contour interval is fixed to be 0.1 (m^2s^{-2}).

Note that such formula holds for the time mean $\overline{f}^t$. Let us then introduce the pattern correlation (PC) and the relative entropy between the coarse models and the reference:

$$\text{PC of } \sigma_f^2 = \frac{\int_\Omega \sigma_f^2 \sigma_{f_{\text{ref}}}^2 \, \mathrm{d}\boldsymbol{x}}{(\int_\Omega \sigma_f^4 \, \mathrm{d}\boldsymbol{x})^{1/2} (\int_\Omega \sigma_{f_{\text{ref}}}^4 \, \mathrm{d}\boldsymbol{x})^{1/2}}, \tag{8.15c}$$

$$\text{Dispersion} = \frac{1}{|\Omega|} \int_\Omega \Big(\frac{\sigma_{f_{\text{ref}}}^2}{\sigma_f^2} - 1 - \log(\frac{\sigma_{f_{\text{ref}}}^2}{\sigma_f^2}) \Big) \, \mathrm{d}\boldsymbol{x}, \tag{8.15d}$$

$$\text{Entropy} = \frac{1}{|\Omega|} \int_\Omega \frac{1}{2} \left(\frac{(\overline{f}^t - \overline{f}^t_{\text{ref}})^2}{\sigma_f^2} + \frac{\sigma_{f_{\text{ref}}}^2}{\sigma_f^2} - 1 - \log\left(\frac{\sigma_{f_{\text{ref}}}^2}{\sigma_f^2} \right) \right) \mathrm{d}\boldsymbol{x}. \tag{8.15e}$$

Basically, from these definitions, we can say that a coarse model with high variability will have high correlation but low entropy for both temporal mean and standard derivation. Conversely, a coarse model with poor variability will lead to a large entropy. In the following, we apply these metrics to the layer pressures and the interface height deviations using the different coarse-resolution simulations. Table 8.3 shows the results for the configuration without the SST evolution. In this case, the LR model has an infinite dispersion and entropy due to an almost zero variability. Compared to the LU-POD, the projected noise enables to increase significantly the internal variability. However, the upper layer correlation remains low. Instead, in the other test case including the SST evolution, the variability of both coarse models are improved. The LU-POD-P always provides the best results. In particular, let us outline that at the coarsest resolution of 120 km, the two LU models provide significant improvements compared to the LR model for all the metrics, as illustrated in Table 8.8 for the pressure fields and in Table 8.9 for the height fields. Similar conclusions can be drawn for the other resolutions. These measure values are then integrated vertically and summarized in Figure 8.14. One can clear see that the improvement of LU models in prediction of variability are resolution-awared.

8.2.2 Energetic diagnosis

This section provides another diagnosis of the low-frequency variability based on some time series analysis of global energy (over space). Note that all the time series in this section are provided by a 5th order Butterworth 2-years low-pass filter.

In the first place, the jet velocity are calculated from the magnitude of the maximum eastward velocity. In the configuration without SST (8.15), both LU-POD and LU-POD-P at resolutions bigger than the deformation radius, produce from the very beginning higher

Layer	Model	RMSE of $\overline{p}^t$	RMSE of σ_p	PC of σ_p^2	Dispersion	Entropy
	LR	0.57	0.34	0.02	∞	∞
1	LU-POD	0.44	0.33	0.04	393	200
	LU-POD-P	0.44	0.31	0.07	37	20
	LR	0.16	0.20	0.09	∞	∞
2	LU-POD	0.14	0.19	0.11	292	147
	LU-POD-P	0.14	0.16	0.30	18	9
	LR	0.05	0.19	0.13	∞	∞
3	LU-POD	0.05	0.18	0.20	293	147
	LU-POD-P	0.05	0.16	0.64	18	9

Table 8.3 – Domain-averaged measures of skill for layer pressures provided by different coarse models ($40\,km$) without SST evolution. The RMS errors are measured in m^2s^{-2}, the remaining columns are dimensionless. The RMS of $\overline{p}^t$ and σ_p for the reference is $(0.96,\ 0.16,\ 0.05)\,m^2s^{-2}$ and $(0.35,\ 0.20,\ 0.20)\,m^2s^{-2}$ respectively. Here, "∞" means that the value is bigger than $\mathcal{O}(10^{16})$.

Layer	Model	RMSE of $\overline{p}^t$	RMSE of σ_p	PC of σ_p^2	Dispersion	Entropy
	LR	1.27	1.20	0.34	179	93
1	LU-POD	0.97	1.18	0.30	71	37
	LU-POD-P	0.97	1.06	0.43	14	8
	LR	0.57	1.01	0.47	394	199
2	LU-POD	0.49	0.97	0.51	76	39
	LU-POD-P	0.48	0.85	0.80	13	7
	LR	0.26	1.01	0.63	639	320
3	LU-POD	0.22	0.96	0.85	96	48
	LU-POD-P	0.22	0.82	0.94	14	7

Table 8.4 – Domain-averaged measures of skill for layer pressures provided by different coarse models ($40\,km$) with SST evolution. The RMS errors are measured in m^2s^{-2}, the remaining columns are dimensionless. The RMS of $\overline{p}^t$ and σ_p for the reference is $(2.32,\ 0.76,\ 0.26)\,m^2s^{-2}$ and $(1.38,\ 1.13,\ 1.07)\,m^2s^{-2}$ respectively.

Interface	Model	RMSE of $\overline{\eta}^t$	RMSE of σ_η	PC of σ_η^2	Dispersion	Entropy
	LR	44	19	0.37	36	275
1	LU-POD	32	19	0.28	33	140
	LU-POD-P	32	18	0.31	14	75
	LR	37	16	0.52	32	169
2	LU-POD	32	16	0.50	27	125
	LU-POD-P	32	15	0.54	12	74

Table 8.5 – Domain-averaged measures of skill for interface height provided by different coarse models ($40\,km$) with SST evolution. The RMS errors are measured in m, the remaining columns are dimensionless. The RMS of $\overline{\eta}^t$ and σ_η for the reference is $(106,\ 54)\,m$ and $(24,\ 23)\,m$ respectively.

Layer	Model	RMSE of $\overline{p}^t$	RMSE of σ_p	PC of σ_p^2	Dispersion	Entropy
	LR	1.32	1.28	0.18	598	312
1	LU-POD	0.84	1.21	0.30	96	50
	LU-POD-P	0.84	1.07	0.48	14	8
	LR	0.60	1.02	0.47	2916	1466
2	LU-POD	0.48	0.98	0.66	95	48
	LU-POD-P	0.47	0.83	0.85	12	6
	LR	0.25	1.04	0.49	9644	4825
3	LU-POD	0.22	0.99	0.91	149	75
	LU-POD-P	0.22	0.82	0.88	15	8

Table 8.6 – Domain-averaged measures of skill for layer pressures provided by different coarse models ($80\,km$) with SST evolution. The RMS errors are measured in m^2s^{-2}, the remaining columns are dimensionless. The RMS of $\overline{p}^t$ and σ_p for the reference is $(2.32,\ 0.76,\ 0.26)\,m^2s^{-2}$ and $(1.38,\ 1.12,\ 1.07)\,m^2s^{-2}$ respectively.

Interface	Model	RMSE of $\overline{\eta}^t$	RMSE of σ_η	PC of σ_η^2	Dispersion	Entropy
	LR	48	21	0.12	140	575
1	LU-POD	30	20	0.16	36	137
	LU-POD-P	30	19	0.18	20	91
	LR	40	19	0.32	103	375
2	LU-POD	32	17	0.47	30	134
	LU-POD-P	32	15	0.49	16	88

Table 8.7 – Domain-averaged measures of skill for interface height provided by different coarse models ($80\,km$) with SST evolution. The RMS errors are measured in m, the remaining columns are dimensionless. The RMS of $\overline{\eta}^t$ and σ_η for the reference is $(106,\ 53)\,m$ and $(24,\ 23)\,m$ respectively.

Layer	Model	RMSE of $\overline{p}^t$	RMSE of σ_p	PC of σ_p^2	Dispersion	Entropy
	LR	1.32	1.31	0.21	1.2×10^4	6240
1	LU-POD	0.84	1.20	0.29	97	50
	LU-POD-P	0.84	1.08	0.43	17	10
	LR	0.59	1.01	0.50	8.7×10^4	4.4×10^4
2	LU-POD	0.47	0.98	0.66	97	49
	LU-POD-P	0.47	0.84	0.84	14	7
	LR	0.25	1.04	0.47	1.9×10^5	9.5×10^4
3	LU-POD	0.22	0.99	0.91	158	79
	LU-POD-P	0.22	0.83	0.88	18	9

Table 8.8 – Domain-averaged measures of skill for layer pressures provided by different coarse models ($120\,km$) with SST evolution. The RMS errors are measured in m^2s^{-2}, the remaining columns are dimensionless. The RMS of $\overline{p}^t$ and σ_p for the reference is $(2.31,\ 0.76,\ 0.26)\,m^2s^{-2}$ and $(1.38,\ 1.12,\ 1.07)\,m^2s^{-2}$ respectively.

Interface	Model	RMSE of $\overline{\eta}^t$	RMSE of σ_η	PC of σ_η^2	Dispersion	Entropy
	LR	48	22	0.10	3376	3426
1	LU-POD	32	21	0.12	45	121
	LU-POD-P	31	20	0.13	28	84
	LR	39	22	0.30	2640	2651
2	LU-POD	33	17	0.42	42	168
	LU-POD-P	32	16	0.44	27	113

Table 8.9 – Domain-averaged measures of skill for interface height provided by different coarse models ($120\,km$) with SST evolution. The RMS errors are measured in m, the remaining columns are dimensionless. The RMS of $\overline{\eta}^t$ and σ_η for the reference is (106, 53) m and (24, 23) m respectively.

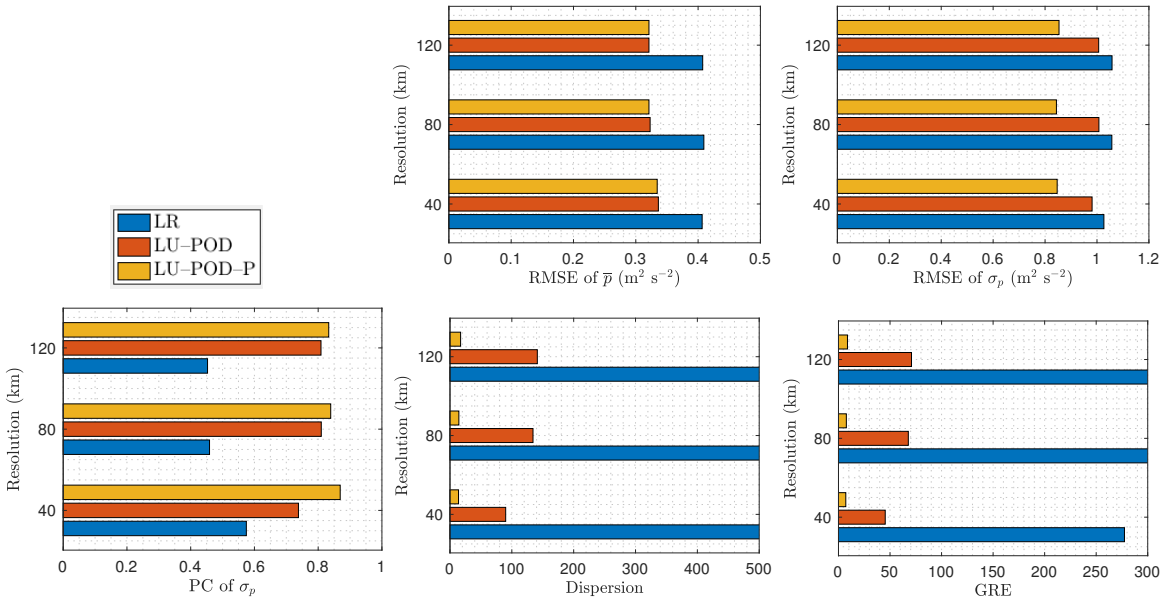

Figure 8.14 – Bar-plots of statistical measures of pressures for different models (with SST) in terms of resolutions.

jet magnitude than the LR. Afterwards, both coarse models converge into a stationary jet state. We remark that the LU-POD and LU-POD-P are often in opposite phase. For the other test case including the SST (8.16), the jet magnitude is increased for all models in comparison to the previous results.

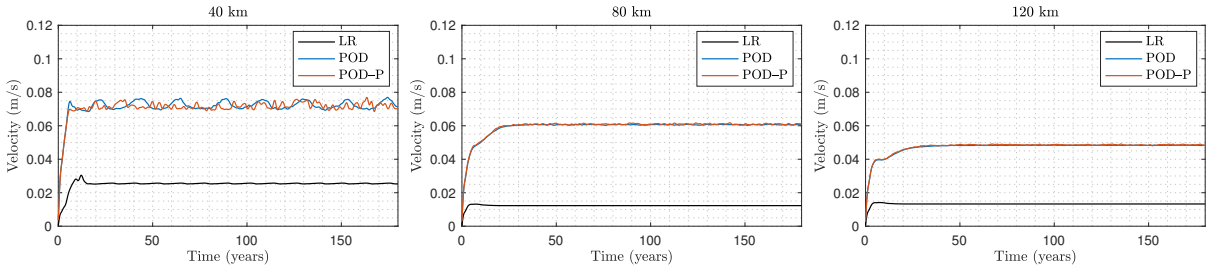

Figure 8.15 – Time series of ocean jet velocity provided by different coarse models (without SST). The typical order of the reference is 0.4.

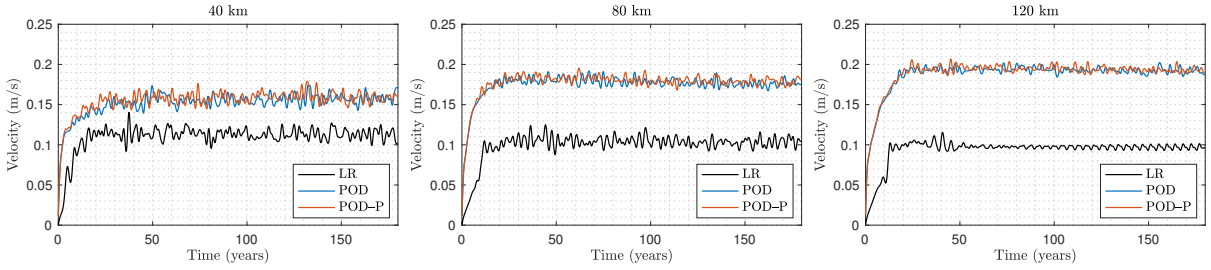

Figure 8.16 – Time series of ocean jet velocity provided by different coarse models (with SST). The typical order of the reference is 0.5.

Further, we use the KE decomposition proposed by Hogg and Blundell (2006), in which the large-scale velocity $\boldsymbol{u}$ is first decomposed into $\overline{\boldsymbol{u}} + \boldsymbol{u}'$ with $\overline{\boldsymbol{u}}$ filtered by the 2-years low-pass filter and $\boldsymbol{u}'$ the residual velocity. Subsequently, the layered kinetic energy is also decomposed into two eddy kinetic energy (EKE) by

$$\mathrm{KE}_k = \underbrace{\frac{\rho H_k}{2|\Omega|}\int_\Omega |\overline{\boldsymbol{u}}|^2 \,\mathrm{d}\boldsymbol{x}}_{\text{Standing EKE}} + \underbrace{\frac{\rho H_k}{2|\Omega|}\int_\Omega |\boldsymbol{u}'|^2 \,\mathrm{d}\boldsymbol{x}}_{\text{Transient EKE}}. \tag{8.16}$$

This procedure is performed in the configuration with the SST evolution. As illustrated on the top of Figure 8.17, when the horizontal resolution is similar to the largest baroclinic deformation radius, the transient EKE are beyond the standing EKE for each model. In

particular, the POD-LU-P behaves the best at such resolutions, in the sense that much more EKE is produced at 40 km compared to the others. As shown on the bottom of this figure, all the systems produce more standing EKE than transient EKE. Nevertheless, the LU-POD-P seems always the best with even higher transient EKE than standing EKE of the LR system at resolution 80 km.

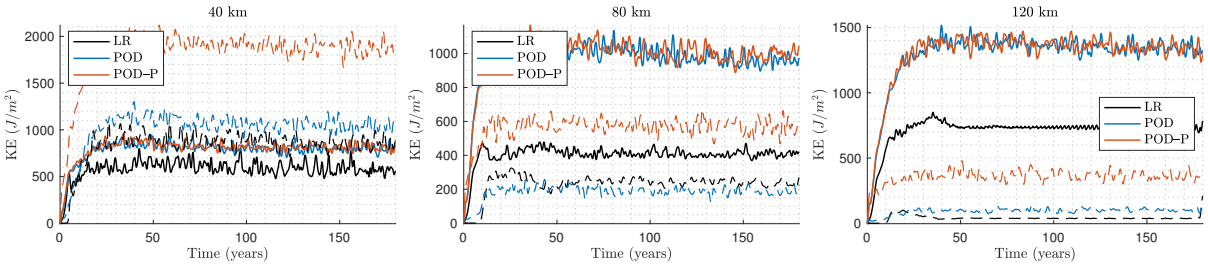

Figure 8.17 – Time series of the standing eddy KE and the transient eddy KE provided by different coarse models (with SST). The standing EKE are described by solid lines whereas the transient EKE are presented by dashed lines.

To understand the source of EKE produced by the LU coarse models (particularly for the LU-POD-P), we further diagnose the transfers of energy between different parts of the system (Hogg and Blundell, 2006):

$$\sum_{k=1}^{N}\frac{\partial \mathrm{KE}_k}{\partial t} = \underbrace{-\sum_{k=1}^{N-1}\frac{\partial \mathrm{PE}_k}{\partial t}}_{\text{PE exchange}} + \underbrace{\rho\int_\Omega u_1\tau^x \mathrm{d}\boldsymbol{x}}_{\text{Wind forcing}} \underbrace{-\rho\sum_{k=1}^{N-1}\int_\Omega g'_k\eta_k w_k\,\mathrm{d}\boldsymbol{x}}_{\text{Buoyancy forcing}}$$
$$\underbrace{-\frac{1}{2}\rho\delta_{ek}f_0\int_\Omega(u_N^2+v_N^2)\mathrm{d}\boldsymbol{x}}_{\text{Bottom drag}} \underbrace{-\sum_{k=1}^{N}A_4\rho H_k\int_\Omega(u_k\nabla^4 u_k + v_k\nabla^4 v_k)\mathrm{d}\boldsymbol{x}}_{\text{Dissipation}} + \cdots, \quad (8.17)$$

where the RHS terms in the first line are the source of KE, whereas the terms in the last line are the sinks. Note that for the configuration without SST, the buoyancy forcing term cancels as $w_1 = 0$. We remark that the previous equation is not complete in describing all the energy transfers for the LU framework. For future work, it would be interesting to estimate numerically these contributions in the energy transfers.

Figure 8.18 shows the results of the previous energy decomposition for both coarse models without the evolution of SST at different resolutions. Compared to the LR model, both LU coarse models provide higher equilibrium between the wind forcing and the dissipation. From Equation (8.17), this mainly results from a higher magnitude of the

dynamic $\boldsymbol{u}$ in each layer. In particular, as demonstrated in Figure 8.19, the LU-POD-P model produces higher energy transfers from PE to KE at each resolution. Otherwise, in the test case with SST forcing, the higher source of KE for LU-POD-P results also from the sum of PE exchange and the buoyancy forcing, as illustrated in Figure 8.20.

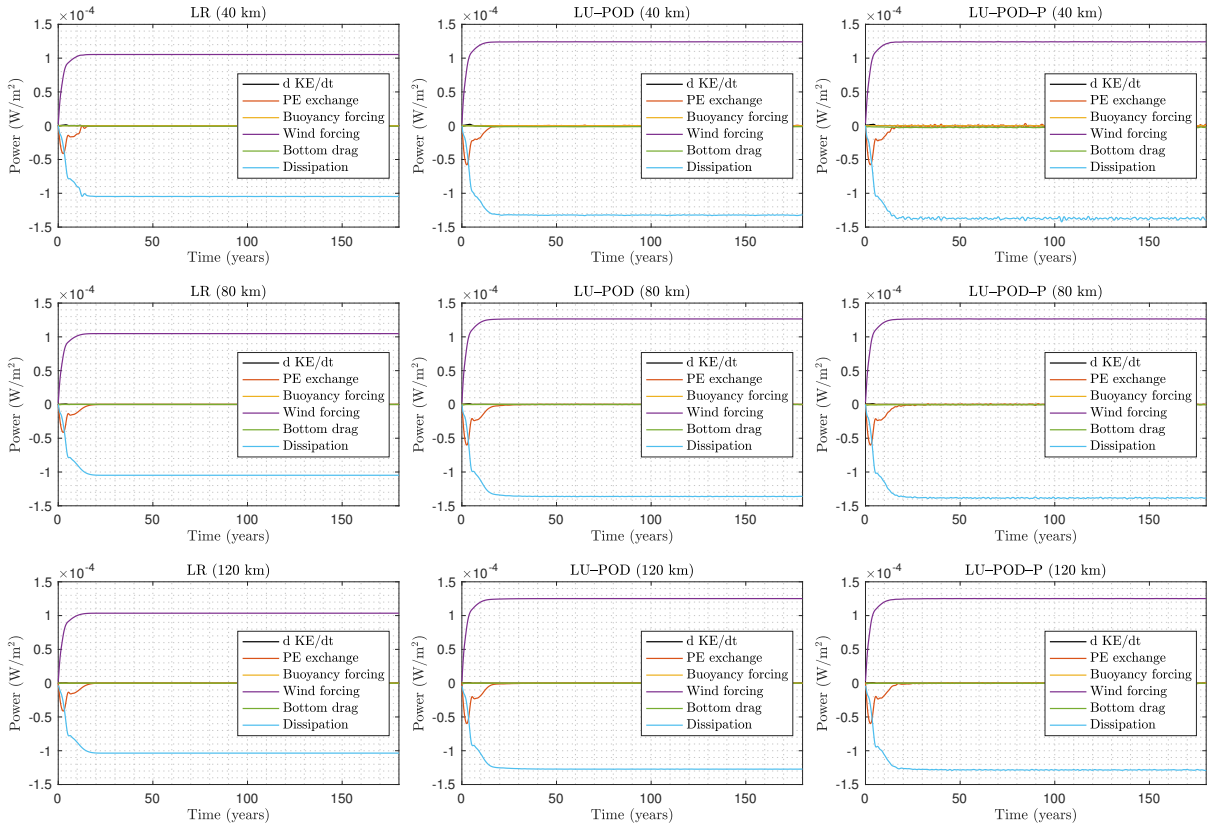

Figure 8.18 – Time series of contributions to the rate of kinetic energy for different coarse models (without SST). From top to bottom: 40 km, 80 km and 120 km; From left to right: LR, LU-POD, LU-POD-P.

Finally, we give an overview of the energy backscattering provided by the LU models. From a basic knowledge of the QG turbulence theory (McWilliams, 2006), like two-dimensional turbulence, there are both inverse-energy and forward-enstrophy inertial-range turbulent cascades under the ideal inviscid framework. However, introducing only a dissipation mechanism like the hyperviscosity while keeping the invariant external forcing like the wind level for coarse models often leads to an excessive decrease of the resolved KE (Arbic et al., 2013; Kjellsson and Zanna, 2017) in numerical simulations. As shown in Figure 8.21, at different resolutions without the SST coupling, both LU models improve the

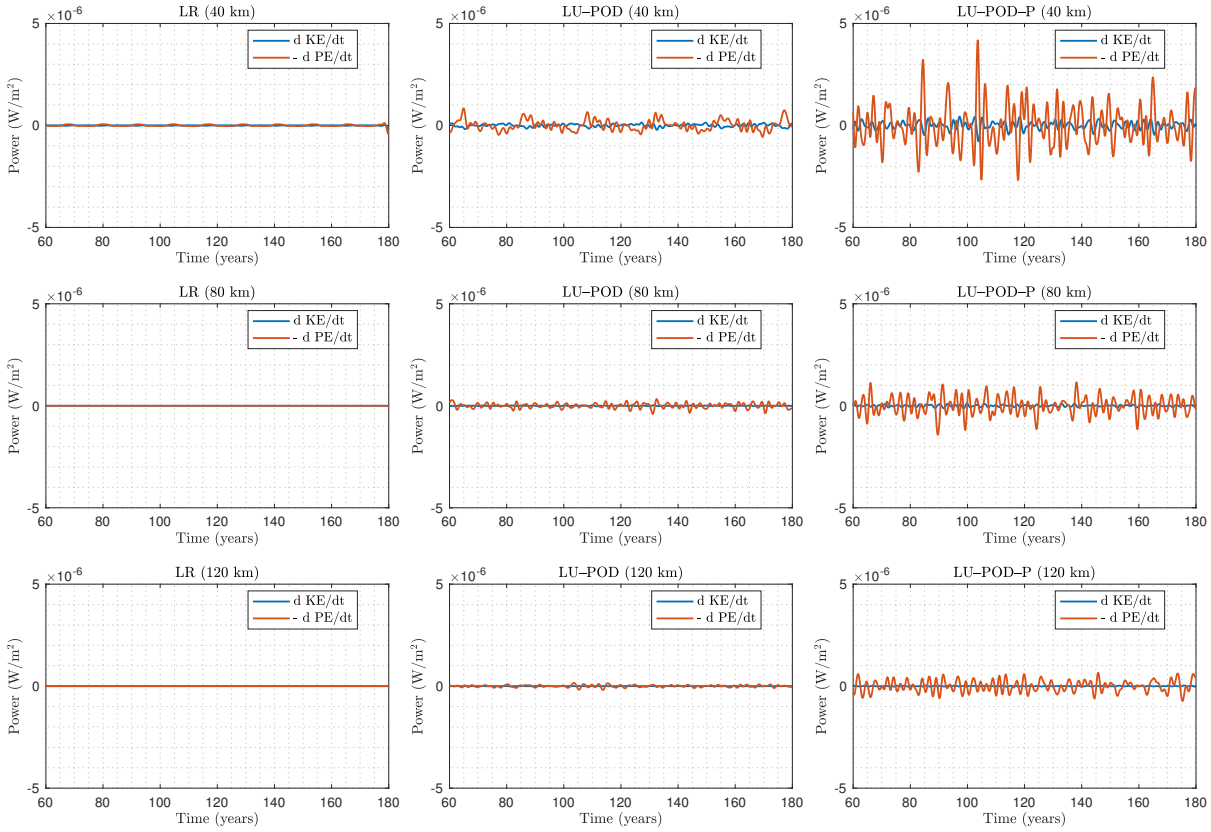

Figure 8.19 – Time series of the rate of change of KE and of the negative rate of change of PE for different coarse models (without SST). From top to bottom: 40 km, 80 km and 120 km; From left to right: LR, LU-POD, LU-POD-P.

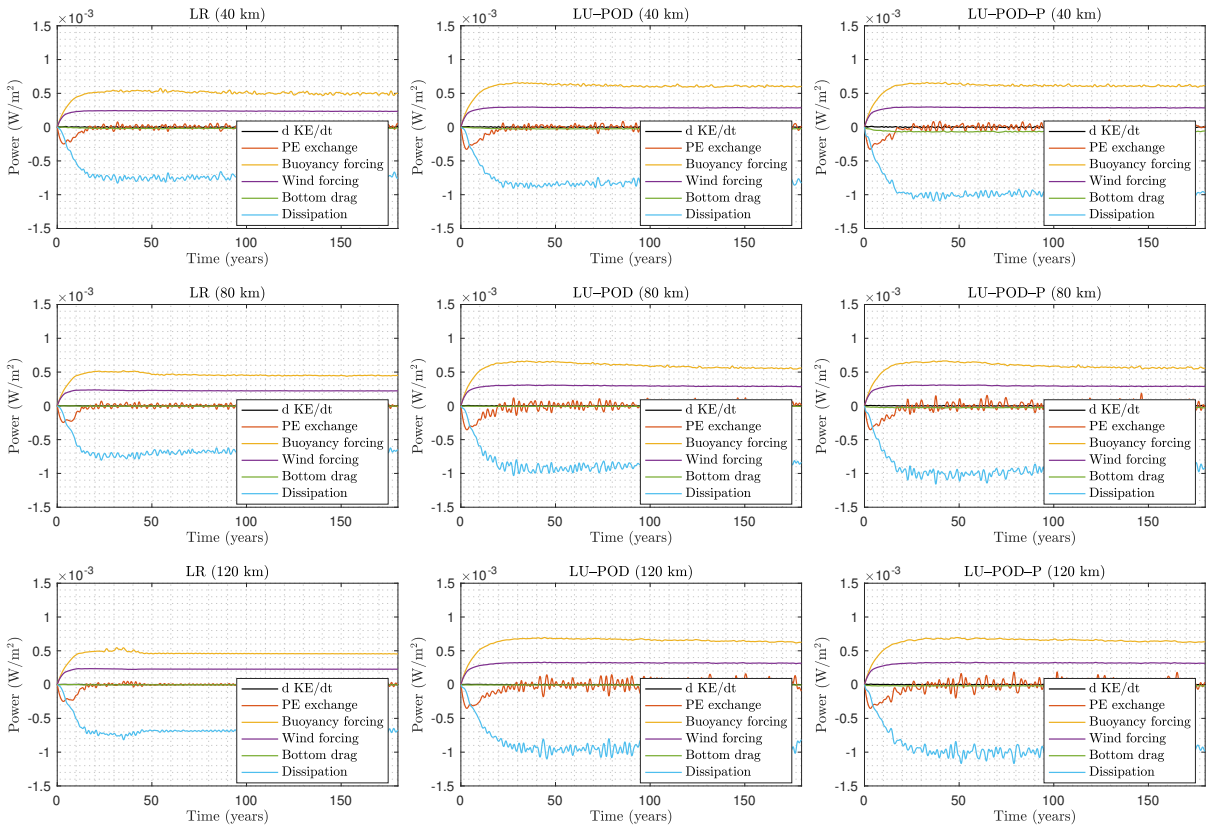

Figure 8.20 – Time series of contributions to the rate of kinetic energy for different coarse models (with SST). From top to bottom: 40 km, 80 km and 120 km; From left to right: LR, LU-POD, LU-POD-P.

time- and layer-averaged spectrums of KE over all wavenumbers compared to those of the LR models. Indeed, the introduction of the non-linear convection by the noise (including the correction drift) reduces the amount of energy dissipated at small scales and allows to backscatter them to the large scales. In particular, the lower the resolution is, the greater amount of large-scale KE is brought by the random models. As illustrated in Figure 8.22, similar results are found for the test case including the SST forcing. In that latter case, the LU-POD-P models, with a noise along the iso-surfaces of vertical stratification, involves an additional SST random forcing that brings an higher KE backscattering than the stationary noises models (LU-POD) at both resolutions. This seems to highlight the importance of the non-stationary characteristic of the noise brought by the iso-stratification projection. Inspired by Scott and Arbic (2007); Zanna et al. (2017), to better interpret the energy transfert across scales, the spectral contributions of each terms in Equation (8.17) can further be explored. This will be performed in a subsequent work.

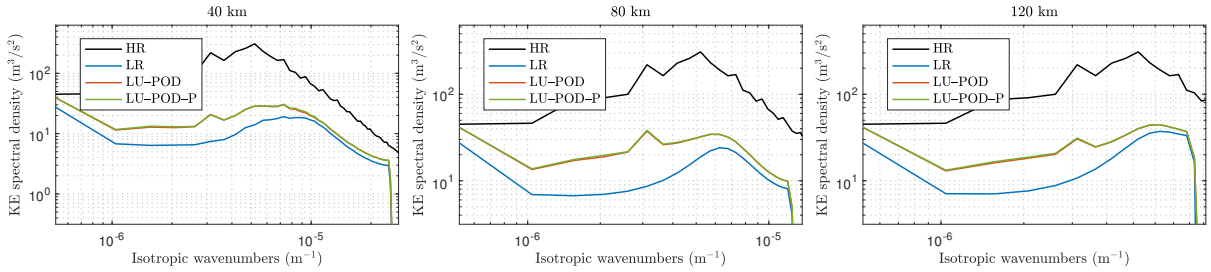

Figure 8.21 – KE spectral density averaged in time (from 60 yrs to 75 yrs using 366 snapshots) and in layers provided by different models (without SST). The HR one is provided by the eddy-resolving (5 km) simulation data. Note that the computational method presented by Durran et al. (2017) is adopted in these two-dimensional spectrums and that the isotropic wavenumbers are normalized by $1/\pi$ while the KE spectral density are normalized by π.

8.3 Conclusion

In this work, the LU model has been successfully implemented in a well established QG dynamical core. We considered here a random flow built from high-resolution data. An additional subgrid correction drift has been introduced in the random flow due to the bias ensuing from the coarse-grained procedure. This non intuitive term seems quite important in the reproduction of the meandering jet within the wind-driven double-gyre

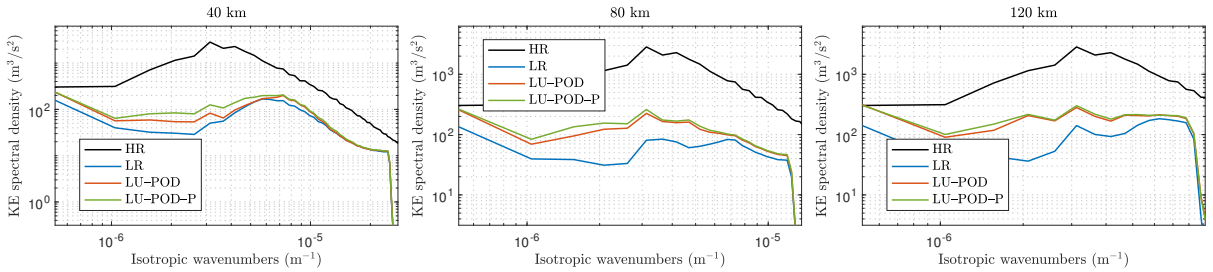

Figure 8.22 – KE spectral density averaged in time (from 60 yrs to 75 yrs) and in layers provided by different models (with SST).

circulation. Further, a projection method has been proposed to constrain the noise to live along the iso-surfaces of the vertical stratification. The resulting noise enables us to improve the intrinsic variability of the large-scale current. From some statistical studies and energy analysis, this improvement is well demonstrated.

Appendix: Spatial discretization of the LU terms

In this appendix we detail the numerical discretization (in space) of the LU additional terms in Equations (8.2) and (8.1). For simplicity, let us denote, $\dot{\boldsymbol{u}} = \frac{1}{\Delta t}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t - \boldsymbol{u}_s$, in the following. The prognostic variables are discretized in the Arakawa C–grid, as illustrated from Fig.A1, where the large–scale geostrophic velocities are discretized by $u_{i,j+1/2} = -\frac{1}{f_0\Delta}(p_{i,j+1} - p_{i,j})$ and $v_{i+1/2,j} = \frac{1}{f_0\Delta}(p_{i+1,j} - p_{i,j})$. The partial free-slip boundary condition of pressure field, $\partial^2_{nn} p = -(\alpha/\Delta)\partial_n p$ (with α positive constant), deduces the ghost southern and northern boundary values of the zonal velocitiy:

$$u^{(0)} = \frac{1-\alpha/2}{1+\alpha/2}u^{(1)}, \qquad u^{(n+1)} = \frac{1-\alpha/2}{1+\alpha/2}u^{(n)}, \tag{8.18}$$

with the same results for the meridional components at western and eastern boundaries. The small-scale velocity, $\dot{\boldsymbol{u}}$, is assumed to have the very same boundaries. As mentioned before, all the LU terms can be re-written in a flux form such that it can be discretized by

$$\boldsymbol{\nabla}\cdot\boldsymbol{F}\,|_{i,j} = \frac{1}{\Delta}\Big(F^x\,|_{i+1/2,j} - F^x\,|_{i-1/2,j} + F^y\,|_{i,j+1/2} - F^y\,|_{i,j-1/2}\Big). \tag{8.19}$$

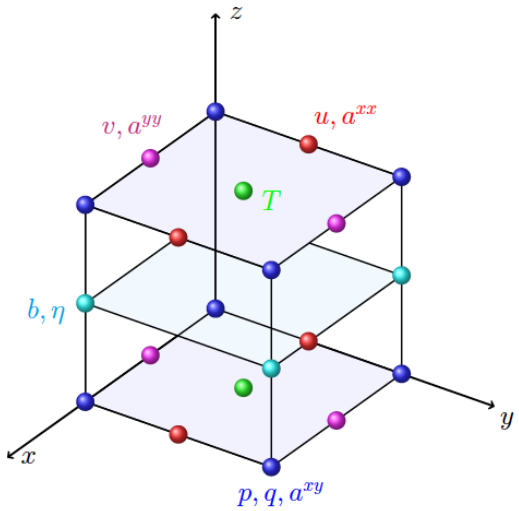

Figure 8.23 – Illustration of the discretized grid. The small-scale zonal (meridional) velocity, $\dot{u}$ ($\dot{v}$) locates at the same position as the large-scale one, u (v).

LU flux in the PV evolution

In the evolution of PV, the LU flux can be decomposed into, $\boldsymbol{F} = \boldsymbol{F}_{\text{ad}} + \boldsymbol{F}_{\text{df}} + \boldsymbol{F}_{\text{so}} + \boldsymbol{F}_{\text{si}}$, where $\boldsymbol{F}_{\text{ad}} = -\dot{\boldsymbol{u}}q$, $\boldsymbol{F}_{\text{df}} = \frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q$, $\boldsymbol{F}_{\text{so}} = -(\boldsymbol{u}\cdot\boldsymbol{\nabla}^{\perp})\dot{\boldsymbol{u}} - \boldsymbol{a}\boldsymbol{\nabla}f$ and $\boldsymbol{F}_{\text{si}} = \frac{1}{2}(\partial_x \boldsymbol{a}\boldsymbol{\nabla}v - \partial_y \boldsymbol{a}\boldsymbol{\nabla}u)$ are, respectively, the flux of advection, diffusion, source and sink. We give the corresponding discretized form of each flux in the following.

Advection flux We propose to use a conservative advection form that corresponds exactly to the 9-points Arakawa Jacobian scheme, i.e. $\boldsymbol{\nabla}\cdot(\boldsymbol{u}q) = \frac{1}{f_0}\mathrm{J}(p,q)$. It reads:

$$F^x_{\text{ad}} = -\left(\frac{2}{3}\overline{\overline{u}^{x}}^{y}\overline{q}^{x} + \frac{1}{3}\overline{\overline{u}^{x}\overline{q}^{d}}^{y}\right), \qquad F^y_{\text{ad}} = -\left(\frac{2}{3}\overline{\overline{v}^{y}}^{x}\overline{q}^{y} + \frac{1}{3}\overline{\overline{v}^{x}\overline{q}^{d}}^{y}\right), \tag{8.20}$$

where the averaging operations are defined by $\overline{q}^{x}_{i+1/2,j} = \frac{1}{2}(q_{i,j} + q_{i+1,j})$, $\overline{q}^{y}_{i,j+1/2} = \frac{1}{2}(q_{i,j} + q_{i,j+1})$ and $\overline{q}^{d}_{i+1/2,j+1/2} = \frac{1}{2}(q_{i,j} + q_{i+1,j+1})$. More precisely, the explicit expressions of our

advection flux are

$$F^x_{\text{ad}}\,|_{i+1/2,j} = -\frac{1}{12}\Big((\dot{u}_{i,j+1/2}+\dot{u}_{i+1,j+1/2}+\dot{u}_{i,j-1/2}+\dot{u}_{i+1,j-1/2})(q_{i,j}+q_{i+1,j}) + \frac{1}{2}(\dot{u}_{i,j+1/2}+\dot{u}_{i+1,j+1/2})(q_{i,j}+q_{i+1,j+1}) + \frac{1}{2}(\dot{u}_{i,j-1/2}+\dot{u}_{i+1,j-1/2})(q_{i,j}+q_{i+1,j-1})\Big), \tag{8.21a}$$

$$F^x_{\text{ad}}\,|_{i-1/2,j} = -\frac{1}{12}\Big((\dot{u}_{i-1,j+1/2}+\dot{u}_{i-1,j-1/2}+\dot{u}_{i,j-1/2}+\dot{u}_{i,j+1/2})(q_{i,j}+q_{i-1,j}) + \frac{1}{2}(\dot{u}_{i-1,j+1/2}+\dot{u}_{i,j+1/2})(q_{i,j}+q_{i-1,j+1}) + \frac{1}{2}(\dot{u}_{i-1,j-1/2}+\dot{u}_{i,j-1/2})(q_{i,j}+q_{i-1,j-1})\Big), \tag{8.21b}$$

$$F^x_{\text{ad}}\,|_{i,j+1/2} = -\frac{1}{12}\Big((\dot{v}_{i-1/2,j+1}+\dot{v}_{i-1/2,j}+\dot{v}_{i+1/2,j}+\dot{v}_{i+1/2,j+1})(q_{i,j}+q_{i,j+1}) + \frac{1}{2}(\dot{v}_{i+1/2,j+1}+\dot{v}_{i+1/2,j})(q_{i,j}+q_{i+1,j+1}) + \frac{1}{2}(\dot{v}_{i-1/2,j+1}+\dot{v}_{i-1/2,j})(q_{i,j}+q_{i-1,j+1})\Big), \tag{8.21c}$$

$$F^x_{\text{ad}}\,|_{i,j-1/2} = -\frac{1}{12}\Big((\dot{v}_{i-1/2,j}+\dot{v}_{i-1/2,j-1}+\dot{v}_{i+1/2,j-1}+\dot{v}_{i+1/2,j})(q_{i,j}+q_{i,j-1}) + \frac{1}{2}(\dot{v}_{i+1/2,j}+\dot{v}_{i+1/2,j-1})(q_{i,j}+q_{i+1,j-1}) + \frac{1}{2}(\dot{v}_{i-1/2,j}+\dot{v}_{i-1/2,j-1})(q_{i,j}+q_{i-1,j-1})\Big). \tag{8.21d}$$

Diffusion flux We discretize our diffusion flux component $F^x_{\text{df}} = \frac{1}{2}\Big(a^{xx}\partial_x q + a^{xy}\partial_y q\Big)$ and $F^y_{\text{df}} = \frac{1}{2}\Big(a^{xy}\partial_x q + a^{yy}\partial_y q\Big)$ by

$$F^x_{\text{df}}\,|_{i+1/2,j} = \frac{1}{8\Delta}\Big((a^{xx}_{i,j+1/2}+a^{xx}_{i,j-1/2}+a^{xx}_{i+1,j+1/2}+a^{xx}_{i+1,j-1/2})(q_{i+1,j}-q_{i,j}) + \frac{1}{2}(a^{xy}_{i,j}+a^{xy}_{i+1,j})(q_{i,j+1}-q_{i,j-1}+q_{i+1,j+1}-q_{i+1,j-1})\Big), \tag{8.22a}$$

$$F^x_{\mathrm{df}}|_{i-1/2,j} = \frac{1}{8\Delta}\Big((a^{xx}_{i-1,j+1/2} + a^{xx}_{i-1,j-1/2} + a^{xx}_{i,j+1/2} + a^{xx}_{i,j-1/2})(q_{i,j} - q_{i-1,j}) + \frac{1}{2}(a^{xy}_{i-1,j} + a^{xy}_{i,j})(q_{i-1,j+1} - q_{i-1,j-1} + q_{i,j+1} - q_{i,j-1})\Big), \tag{8.22b}$$

$$F^y_{\mathrm{df}}|_{i,j+1/2} = \frac{1}{8\Delta}\Big((a^{yy}_{i+1/2,j} + a^{yy}_{i-1/2,j} + a^{yy}_{i+1/2,j+1} + a^{yy}_{i-1/2,j+1})(q_{i,j+1} - q_{i,j}) + \frac{1}{2}(a^{xy}_{i,j} + a^{xy}_{i,j+1})(q_{i+1,j} - q_{i-1,j} + q_{i+1,j+1} - q_{i-1,j+1})\Big), \tag{8.22c}$$

$$F^y_{\mathrm{df}}|_{i,j-1/2} = \frac{1}{8\Delta}\Big((a^{yy}_{i+1/2,j-1} + a^{yy}_{i-1/2,j-1} + a^{yy}_{i+1/2,j} + a^{yy}_{i-1/2,j})(q_{i,j} - q_{i,j-1}) + \frac{1}{2}(a^{xy}_{i,j-1} + a^{xy}_{i,j})(q_{i+1,j-1} - q_{i-1,j-1} + q_{i+1,j} - q_{i-1,j})\Big). \tag{8.22d}$$

We use the same approximation of the advection flux for the sources and the same approximation of the diffusion flux for the sinks.

Discretization of LU flux in the SST evolution

In the evolution of SST, the LU flux is decomposed in two terms, $\boldsymbol{F}_{\mathrm{ad}} = -\dot{\boldsymbol{u}}T$ and $\boldsymbol{F}_{\mathrm{df}} = \frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}T$. In the following, the index (i,j) stands for the T-grid.

Advection flux Only central-winding is adopted for the SST advection:

$$F^x_{\mathrm{ad}}|_{i+1/2,j} = -\frac{1}{2}\dot{u}_{i+1/2,j}(T_{i,j} + T_{i+1,j}), \tag{8.23a}$$

$$F^x_{\mathrm{ad}}|_{i-1/2,j} = -\frac{1}{2}\dot{u}_{i-1/2,j}(T_{i-1,j} + T_{i,j}), \tag{8.23b}$$

$$F^y_{\mathrm{ad}}|_{i,j+1/2} = -\frac{1}{2}\dot{v}_{i,j+1/2}(T_{i,j} + T_{i,j+1}), \tag{8.23c}$$

$$F^y_{\mathrm{ad}}|_{i,j-1/2} = -\frac{1}{2}\dot{v}_{i,j-1/2}(T_{i,j-1} + T_{i,j}). \tag{8.23d}$$

Diffusion flux Only central-interpolation is adopted for the SST diffusion:

$$F^x_{\mathrm{df}}\big|_{i+1/2,j} = \frac{1}{2}\Big(a^{xx}_{i+1/2,j}(T_{i+1,j}-T_{i,j}) + \frac{1}{8}(a^{xy}_{i+1/2,j+1/2}+a^{xy}_{i-1/2,j+1/2}) \\ (T_{i,j+1}-T_{i,j-1}+T_{i+1,j+1}-T_{i+1,j-1})\Big), \tag{8.24a}$$

$$F^x_{\mathrm{df}}\big|_{i-1/2,j} = \frac{1}{2}\Big(a^{xx}_{i-1/2,j}(T_{i,j}-T_{i-1,j}) + \frac{1}{8}(a^{xy}_{i-1/2,j+1/2}+a^{xy}_{i-1/2,j-1/2}) \\ (T_{i-1,j+1}-T_{i-1,j-1}+T_{i,j+1}-T_{i,j-1})\Big), \tag{8.24b}$$

$$F^y_{\mathrm{df}}\big|_{i,j+1/2} = \frac{1}{2}\Big(a^{yy}_{i,j+1/2}(T_{i,j+1}-T_{i,j}) + \frac{1}{8}(a^{xy}_{i-1/2,j+1/2}+a^{xy}_{i+1/2,j+1/2}) \\ (T_{i+1,j+1}-T_{i-1,j+1}+T_{i+1,j}-T_{i-1,j})\Big), \tag{8.24c}$$

$$F^y_{\mathrm{df}}\big|_{i,j-1/2} = \frac{1}{2}\Big(a^{yy}_{i,j-1/2}(T_{i,j}-T_{i,j-1}) + \frac{1}{8}(a^{xy}_{i-1/2,j-1/2}+a^{xy}_{i+1/2,j-1/2}) \\ (T_{i+1,j}-T_{i-1,j}+T_{i+1,j-1}-T_{i-1,j-1})\Big). \tag{8.24d}$$

Conclusion

In this thesis, we have explored a stochastic representation of the small-scale effects on the large-scale oceanic circulation. This stochastic framework, called modeling under location uncertainty (LU), is derived from the classical physical conservation laws and arises from a decomposition of the Lagrangian velocity into a time-smooth component and a highly oscillating noise term. Using some fundamental stochastic calculus, a stochastic partial differential equation (SPDE) is found to describe the evolution law of a random tracer transported along the stochastic Lagrangian trajectory. Compared to the classical transport equation associated to the deterministic material derivative operator, this SPDE involves a new stochastic transport operator that includes three additional terms: A multiplicative random forcing corresponding to the tracer's advection by the noise, which backscatters energy to the system; A tracer's dissipation depicting the mixing mechanism due to the action of the small scales, which plays a role similar to the eddy viscosity introduced in many large-scale circulation models; An effective advection term capturing the small-scale inhomogeneity action over the large-scale current, which redistributes the large-scale energy. One important characteristic of this SPDE is that it preserves in time the global energy of the random tracer for any realizations. In this process, the energy brought by the noise is exactly counter-balanced by the energy loss caused by the diffusion. A stochastic Reynolds transport theorem (SRTT) is derived from this stochastic transport to describe the Eulerian transport of random tracers within arbitrary moving and deforming control-volumes.

The resulting SRTT combined with physical conservation laws allows us to derive the stochastic hydrostatic primitive equations and the stochastic simple Boussinesq equations. Using this latter, we have shown that the proposed LU model introduces large-scale flow structuration caused by the action of the small-scale flow component. This effect is generated by the statistical eddy-induced velocity (appeared in the effective advection term), associated to the noise inhomogeneity, which can be interpreted as a generalization of the Stokes drift and is hence referred to in this work as Itô-Stokes drift. In fact, this Itô-Stokes drift relies on the difference between the ensemble mean of Lagrangian fluid velocity and the mean of the Eulerian effective velocity. With this analogy and proper

assumptions, we have shown that the derived stochastic Boussinesq dynamics encompasses the classical Leibovich system unveiling a vortex force responsible of the emergence of secondary circulations. However, this stochastic system is not restricted to wave current interactions and is capable of accounting for more complex interactions between the large-scale current and the small-scale fluctuations. In addition, we have briefly interpreted the proposed LU model in terms of the Gent-McWilliams (GM) parametrization that is widely used in many global ocean models. By projecting the noise along the isopycnal surfaces (iso-surfaces of density or of buoyancy), the Itô-Stokes drift in LU scheme plays a similar role as the so-called bolus velocity introduced in the GM scheme and the tracer's diffusion through the noise variance can be connected with Redi-like isoneutral diffusion.

In order to simplify the governing equations of the oceanic (and atmospheric) large-scale circulations, some vertical and horizontal approximations are adopted for the derived stochastic primitive and Boussinesq systems. A resulting stochastic shallow-water model under geostrophic noise is demonstrated to conserve the total energy. Through scaling and asymptotic small parameter power expansion different stochastic versions of the planetary geostrophic (PG), the quasigeostrophic (QG) and the surface quasigeostrophic (SQG) models have been explicitly obtained. In these derivations, the noise amplitude needs to be properly scaled. This additional degree of freedom compared to the deterministic case provides an interesting and practical tool for investigating the implication of the small-scale components of the flow. Under the QG regime, additional vorticity sources arise from the interaction of the strains between the small-scale noise and the large-scale current. These terms have shown to be essential for the total energy conservation of the flow. In this thesis, these simplified two-dimensional and/or layered stochastic models have been tested in different configurations through several numerical simulations.

To perform the numerical simulation of the LU models on coarse meshes, we have in particular explored various types of noises – homogeneous/heterogenous and stationary/non-stationary, using different approaches – data-driven/parameterized. In the data-driven models, the proper orthogonal decomposition (POD) procedure has been adopted to learn efficiently the stationary eigenfunction basis of the velocity noise from data of an eddy-resolving simulation. In particular, a subgrid correction drift has been identified due to the bias ensuing from the coarse-graining procedure of the high-resolution data used to perform the POD. To release the strong stationary assumption of this approach, we proposed to update-in-time the noise EOFs by by matching the noise's principal temporal modes with the large-scale dynamics. This novel method consists in selecting from the

reference data the set of time instances that match the large-scale structures of a coarse-resolution simulation. In addition to homogeneous and stationary noise expressed on a set of Fourier modes of high wavenumbers as proposed in Resseguier et al. (2017b), two new parameterizations of heterogeneous non-stationary noises have been used. The first one consists in estimating time-dependent EOFs from local velocity fluctuations generated directly from the coarse on going simulation, together with a rescaling of the noise amplitude based on a similarity assumption. However, this method comes with a slightly higher computational cost, since the singular value decomposition (SVD) procedure has to be performed on the local fluctuations at each time step. The latter method is designed for a stochastic transport equation of passive tracers with a given subgrid scale dissipation operator, such as Laplacian, biharmonic or hyperviscosity operator. The main idea relies on the assumption that the variance diffusion in LU models can be identified with the subgrid dissipation up to a scaling factor. As such, for a given orthonormal basis such as wavelet, the associated eigenvalues are updated according to a desired energy dissipation balance. The resulting transport noise counter-balances then the numerical dissipation removed by the subgrid scales up to a prescribed scaling factor. However, it is difficult in that case to recover explicitly the velocity noise and its variance tensor, only the random advection term is specified.

For numerical validations, we have first used the random SQG model to study short-terms ensemble forecasting skills predicted by different LU noise models, using some classical metrics such as Talagrand diagrams, the continuous ranked proper score and the energy score. As a result, we have shown that the proposed LU models provide higher ensemble spread and more reliable forecast than the classical random model built from perturbation of the initial condition (PIC). This highlights that LU models are able to overcome the underdispersive issue (Mitchell and Gottwald, 2012; Franzke et al., 2015) of the PIC models without tuning any "inflation" parameters classically introduced to increase the ensemble covariance (Anderson and Anderson, 1999). This ability is in particular of crucial importance for data assimilation applications in geophysical sciences such as meteorology or oceanography. Furthermore, in terms of ensemble reliability, we concluded that the LU heterogeneous noise models are better than the homogeneous noise models and the non-stationary noise models are better than the stationary ones.

Numerical simulations of the barotropic QG model under LU has been then assessed. In the test case of a simplified Rossby wave, we have shown that the proposed random model with both homogeneous and heterogeneous noises preserve the magnitude and

the propagation speed of the wave. Besides, we have illustrated numerically that the inhomogeneous noise together with the Itô-Stokes drift induces a structuration of the large-scale flow with strong secondary vortices.

In another viscous double-gyre configuration driven by an ideal steady wind within an enclosed shallow basin, we have performed long-terms statistical analysis for different coarse models. Compared to a deterministic coarse model, the LU model better represents the nonlinearity at the resolved scales while properly dissipating the unresolved scales, hence leading to a balanced correction of excessive dissipation and to an improvement of internal variability. Both stationary and non-stationary noise models enable us to reproduce qualitatively and quantitatively on a coarse mesh the statistical distribution of the large-scale tracers, predicted by the eddy-resolving simulation data. In addition, the non-stationary noise based on the temporal-mode-matching approach provides better performances in most of the statistics.

Finally, the proposed LU model has been implemented in a well established multi-layered QG dynamical core (Hogg et al., 2004). The empirical spatial correlation of the small-scale noise has been first estimated from eddy-resolving simulation data. In particular, the non intuitive subgrid correction drift has been found to be very important in reproducing on a coarse mesh the meandering jet of the double-gyre circulation. In addition, a new projection method has been proposed to constrain the noise living along the iso-surfaces of the vertical stratification, in order to improve the transfer of available potential energy toward the kinetic energy. The resulting non-stationary noise enables us to improve the intrinsic low-frequency variability of the large-scale current. This improvement has been well demonstrated through a transfer analysis and some statistical criterion. We have highlighted that both stationary and non-stationary noise models provide significant improvements at ocean climatic scale (resolution beyond 100 km) in terms of variability metrics, compared to a corresponding deterministic coarse model. At that resolution, the effects of the mesoscale eddies within the ocean basin at the midlatitude (of deformation radius around 30$\sim$40 km) are better represented by the LU models, even though the baroclinic instability can not be fully resolved.

In future work, we believe that this study may have three interesting prospective axes – models, noises and assimilations. First, the encouraging numerical results presented in this thesis encourage us to implement the proposed random model on more complex and realistic flow configurations. A following study for a three-layered atmosphere and three-layered ocean coupled QG model (Hogg et al., 2003) is already in progress. Including

random forcing into the bottom layer of the atmosphere leads to both to an unsteady wind stress and the random diabatic forcing due to Ekman pumping of the sea surface temperature (SST) in both atmospheric and oceanic mixed layers. Thus, it will be interesting to verify the response of the ocean dynamics in terms of variability and energy to these highly variable forcings, and particularly to compare the results to a stationary wind forced ocean dynamics as well as to an unforced dynamics. Instead of the idealized double-gyre circulation, a southern ocean channel configuration (Hogg and Blundell, 2006; Grooms et al., 2015) can be further adopted to investigate the intrinsic variability of the Antarctic Circumpolar Current (ACC). Such configuration could be completed with the topography prescribed by Hogg and Blundell (2006). In that test case, we could compare the statistical prediction skills of the LU coarse models to that of the different GM parameterizations presented in Grooms et al. (2015). The stochastic primitive Boussinesq models, either with the NEMO (Lévy et al., 2010; NEMO team, 2016) or MITgcm (Marshall et al., 1997b,a; Campin et al., 2020) codes could then be tested for global ocean circulations with different noises as presented in this work. However, in that case the random pressure term in the LU model has to be a priori modeled or computed based on proper scalings of hydrostasy. These efforts aim at progressively going toward the study of stochastic IPCC-class ocean models and to confirm whether relevant stochastic flow models may contribute to improve them.

Another possible research axis focuses on the improvement of the noise modeling. New data analysis techniques could be explored in practice to estimate the dynamics of the small-scale noise, first from eddy-resolving simulation data, then from high-resolution satellite data. Techniques relying on the spectral analysis of the Koopman operator (Mezić, 2013; Giannakis, 2019; Gugole and Franzke, 2020) could be contemplated for that purpose. Investigations on the constitution of the random term from high frequency waves generated by surface waves or through internal waves will also be conducted. The performance of the resulting random models will be assessed and analyzed again in terms of forecasting. The objective will consist in evaluating their ability to characterize the long-terms impacts of some small-scale events.

On the other hand, one possible parametrization method based on the classical linear stability analysis (Smith, 2007; Vallis, 2017) of the resolved dynamics could be investigated and tested. The motivation consists always in designing adequate noises that provide efficient conversion from the mean available potential energy (MAPE) to the eddy kinetic energy (EKE) through baroclinic instabilities. For instance, we could first linearize the

QG equation around the local means of velocity and of potential vorticity (PV) gradient. Substituting subsequently a plane-wave solution into that linearized equation and solving the resulting eigenvalue problem, will allow us to obtain a set of normal modes (eigenfunctions) associated with their frequencies (eigenvalues). Then, in order to represent the noise terms, we may select some modes and frequencies of the fastest growing waves according to the rate of conversion from MAPE to EKE. This rate can be calculated from the QG energy budget equation. We expect that such linear stability analysis could provide a good estimate of the characteristic time and locations for the small-scale noise. However, the computational cost associated to such a procedure has to be carefully considered. Nevertheless, this could provide us a very interesting parameterization of the small-scale velocity components. Moreover, we have to keep in mind that data are not always available.

Finally, we have a strong interest on the coupling of random dynamical models with high-resolution satellite image observation through the use of ensemble data assimilation procedures, particularly through the particle filtering (Bain and Crisan, 2008; Cotter et al., 2019b). The data assimilation technique proposed will benefit from the natural multiscale analysis framework associated to data-driven noise models. The objective will be to devise efficient data assimilation procedures in high-dimensional space through the proposed random model. As such, this modeling provides an efficient mean to sample the dynamics along meaningful directions, construct reduced order models but also to smooth the data on a dynamical ground. In the first place, we could adopt the hierarchical particle filter (Wikle and M., 2007; Van Leeuwen, 2009) for the LU coarse simulations using the Q-GCM dynamical cores and the eddy-resolving simulation data. This will allow us to reduces the assimilation complexity and to rely on different algorithms for different parts of hierarchy. Later, for higher-dimensional problem such as the stochastic primitive Boussinesq system, we could consider some Markov chain Monte Carlo (MCMC) methods such as Metropolis-Hastings (Robert et al., 2004; Van Leeuwen et al., 2019).

BIBLIOGRAPHY

Anderson, J., Anderson, S., 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly Weather Review 127, 2741–2758.

Anderson, J.L., 1996. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. Journal of Climate 9, 1518–1530.

Andrews, D., McIntyre, M., 1978. An exact theory of nonlinear waves on a Lagrangian-mean flow. Journal of Fluid Mechanics 89, 609–646.

Arakawa, A., Lamb, V.R., 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. Methods in Computational Physics 17, 173–265.

Arakawa, A., Lamb, V.R., 1981. A potential enstrophy and energy conserving scheme for the shallow water equations. Monthly Weather Review 109, 18–36.

Arbic, B.K., Polzin, K.L., Scott, R.B., Richman, J.G., Shriver, J.F., 2013. On eddy viscosity, energy cascades, and the horizonal resolution of gridded stallite altimeter products. Journal of Physical Oceanography 43, 283–300.

Bain, A., Crisan, D., 2008. Fundamentals of Stochastic Filtering. Stochastic Modelling and Applied Probability, Springer New York.

Bauer, W., Chandramouli, P., Chapron, B., Li, L., Mémin, E., 2020a. Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach. Journal of Physical Oceanography 50, 983–1003.

Bauer, W., Chandramouli, P., Li, L., Mémin, E., 2020b. Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models. Ocean Modelling 151, 101646.

Berloff, P., 2005. Random-forcing model of the mesoscale oceanic eddies. Journal of Fluid Mechanics 529, 71–95.

Brecht, R., Li, L., Bauer, W., Mémin, E., 2021. Rotating shallow water flow under location uncertainty with a structure-preserving discretization. URL: https://hal.inria.fr/hal-03131680. Submitted for Journal of Advances in Modeling Earth Systems.

Campin, J.M., Heimbach, P., Losch, M., Forget, G., Adcroft, A., Menemenlis, D., Hill, G., Jahn, O., Scott, J., Mazloff, M., Fox-Kemper, B., Doddridge, E., Fenty, I., Bates, M., Smith, T., Martin, T., Lauderdale, J., Abernathey, R., Deremble, B., Bourgault, P., 2020. MITgcm/MITgcm: mid 2020 version. National Oceanic and Atmospheric Administration.

Chandramouli, P., Heitz, D., Laizet, S., Mémin, E., 2018. Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty. Computers and Fluids 168, 170–189.

Chandramouli, P., Mémin, E., Heitz, D., 2020. 4D large scale variational data assimilation of a turbulent flow with a dynamics error model. Journal of Computational Physics 412.

Chapron, B., Dérian, P., Mémin, E., Resseguier, V., 2018. Large-scale flows under location uncertainty: a consistent stochastic framework. Quarterly Journal of the Royal Meteorological Society 144, 251–260.

Cintolesi, C., Mémin, E., 2020. Stochastic modelling of turbulent flows for numerical simulations. Fluids 5, 108.

Cooper, F.C., Zanna, L., 2015. Optimisation of an idealised ocean model, stochastic parameterisation of sub-grid eddies. Ocean Modelling 88, 38–53.

Cotter, C., Crisan, D., Holm, D.D., Pan, W., Shevchenko, I., 2019a. Numerically modeling stochastic Lie transport in fluid dynamics. Multiscale Modeling & Simulation 17, 192–232.

Cotter, C., Crisan, D., Holm, D.D., Pan, W., Shevchenko, I., 2019b. A particle filter for stochastic advection by Lie transport (SALT): A case study for the damped and forced incompressible 2D Euler equation. arXiv preprint arXiv:1907.11884 .

Cotter, C., Crisan, D., Holm, D.D., Pan, W., Shevchenko, I., 2020. Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. Foundations of Data Science 2, 173.

Craik, A., Leibovich, S., 1976. A rational model for Langmuir circulations. Journal of Fluid Mechanics 73, 401–426.

Crisan, D., Flandoli, F., Holm, D.D., 2019. Solution properties of a 3D stochastic Euler fluid equation. Journal of Nonlinear Science 29, 813–870.

Da Prato, G., Zabczyk, J., 2014. Stochastic equations in infinite dimensions. Encyclopedia of Mathematics and its Applications. 2 ed., Cambridge University Press.

Daubechies, I., 1992. Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics.

Durran, D., Weyn, J.A., Menchaca, M.Q., 2017. Practical considerations for computing dimensional spectra from gridded data. Monthly Weather Review 145, 3901–3910.

Fox-Kemper, B., 2005. Reevaluating the roles of eddies in multiple barotropic wind-driven gyres. Journal of Physical Oceanography 35, 1263–1278.

Franzke, C.E., Majda, A.J., 2006. Low-order stochastic mode reduction for a prototype atmospheric GCM. Journal of the Atmospheric Sciences 63, 457–479.

Franzke, C.E., O'Kane, T.J., Berner, J., Williams, P.D., Lucarini, V., 2015. Stochastic climate theory and modeling. Wiley Interdisciplinary Reviews: Climate Change 6, 63–78.

Gent, P.R., McWilliams, J.C., 1990. Isopycnal mixing in ocean circulation models. Journal of Physical Oceanography 20, 150–155.

Gent, P.R., Willebrand, J., McDougall, T.J., McWilliams, J.C., 1995. Parameterising eddy-induced tracer transports in ocean circulation models. Journal of Physical Oceanography 25, 463–474.

Giannakis, D., 2019. Data-driven spectral decomposition and forecasting of ergodic dynamical system. Applied and Computational Harmonic Analysis 47, 338–396.

Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102, 359–378.

Gottlieb, S., 2005. On higher order strong stability preserving Runge-Kutta and multi-step time discretizations. Journal of Scientific Computing 25, 105–128.

Gottwald, G., Crommelin, D.T., Franzke, C.E., 2017. Stochastic climate theory, in: Nonlinear and Stochastic Climate Dynamics. Cambridge University Press, pp. 209–240.

Gottwald, G., Harlim, J., 2013. The role of additive and multiplicative noise in filtering complex dynamical systems. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 469, 20130096.

Greatbatch, R.J., Nadiga, B.T., 2000. Four-gyre circulation in a barotropic model with double-gyre wind forcing. Journal of Physical Oceanography 30, 1461–1471.

Griffies, S.M., 1998. The Gent-McWilliams skew flux. Journal of Physical Oceanography 28, 831–841.

Griffies, S.M., Gnanadesikan, A., Pacanowski, R.C., Larichev, V.D., Dukowicz, J.K., Smith, R.D., 1998. Isoneutral diffusion in a z-coordinate ocean model. Journal of Physical Oceanography 28, 805–830.

Griffies, S.M., Hallberg, R.W., 2000. Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. Monthly Weather Review 128, 2935–2946.

Grooms, I., Julien, K., Fox-Kemper, B., 2011. On the interactions between planetary geostrophy and mesoscale eddies. Dynamics of Atmospheres and Oceans 51, 109–136.

Grooms, I., Majda, A.J., 2014. Stochastic superparameterization in quasigeostrophic turbulence. Journal of Computational Physics 271, 78–98.

Grooms, I., Majda, A.J., Smith, K.S., 2015. Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current. Ocean Modelling 85, 1–15.

Gugole, F., Franzke, C.E., 2019. Numerical development and evaluation of an energy conserving conceptual stochastic climate model. Mathematics of Climate and Weather Forecasting 5, 45–64.

Gugole, F., Franzke, C.E., 2020. Spatial covariance modeling for stochastic subgrid-scale parameterizations using dynamic mode decomposition. Journal of Advances in Modeling Earth Systems 12.

Hamill, T.M., 2001. Interpretation of rank histograms for verifying ensemble forecasts. Monthly Weather Review 129, 550–560.

Hamill, T.M., Colucci, S.J., 1997. Verification of Eta-RSM short-range ensemble forecasts. Monthly Weather Review 125, 1312–1327.

Harcourt, R., 2015. An improved second-moment closure model of Langmuir turbulence. Journal of Physical Oceanography 45, 84–103.

Held, I., Pierrehumbert, R., Garner, S., Swanson, K., 1995. Surface quasi-geostrophic dynamics. Journal of Fluid Mechanics 282, 1–20.

Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting 15, 559–570.

Hogg, A.M., Blundell, J.R., 2006. Interdecadal variability of the southern ocean. Journal of Physical Oceanography 36, 1626–1645.

Hogg, A.M., Dewar, W.K., Killworth, P.D., Blundell, J.R., 2003. A quasi-geostrophic coupled model (Q-GCM). Monthly Weather Review 131, 2261–2278.

Hogg, A.M., Killworth, P.D., Blundell, J.R., 2004. Mechanisms of decadal variability of the wind-driven ocean circulation. Journal of Physical Oceanography 35.

Holm, D.D., 1996. The ideal Craik-Leibovich equations. Physica D: Nonlinear Phenomena 98, 415–441.

Holm, D.D., 2015. Variational principles for stochastic fluid dynamics. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 471, 20140963.

Holm, D.D., 2019. Stochastic closures for wave-current interaction dynamics. Journal of Nonlinear Science 29.

Holm, D.D., Nadiga, B.T., 2003. Modeling mesoscale turbulence in the barotropic double-gyre circulation. Journal of Physical Oceanography 33, 2355–2365.

Holmes, P., Lumley, J., Berkooz, G., 1996. Turbulence, coherence structures, dynamical systems and symetry. Cambridge University Press.

Jansen, M.F., Held, I.M., 2014. Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. Ocean Modelling 80, 36–48.

Kadri Harouna, S., Mémin, E., 2017. Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. Computers and Fluids 156, 456–469.

Karspeck, A.R., Yeager, S., Danabasoglu, G., Hoar, T., Collins, N., Raeder, K., Anderson, J., Tribbia, J., 2013. An ensemble ajustment Kalman filter for the CCSM4 ocean component. Journal of Climate 26, 7392–7413.

Kjellsson, J., Zanna, L., 2017. The impact of horizontal resolution on energy transfers in global ocean models. Fluids 2, 45.

Kunita, H., 1997. Stochastic flows and stochastic differential equations. volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press.

Lapeyre, G., 2017. Surface quasi-geostrophy. Fluids 2, 7.

Le Gall, J., 2016. Brownian Motion, Martingales, and Stochastic Calculus. Graduate Texts in Mathematics, Springer International Publishing.

Leibovich, S., 1980. On wave-current interaction theories of Langmuir circulations. Journal of Fluid Mechanics 99, 715–724.

Leith, C.E., 1971. Atmospheric predictability and two-dimensional turbulence. Journal of the Atmospheric Sciences 28, 145–161.

Lévy, M., Klein, P., Tréguier, A.M., Iovino, D., Madec, G., Masson, S., Takahashi, K., 2010. Modifications of gyre circulation by sub-mesoscale physics. Ocean Modelling 34, 1–15.

Li, L., Deremble, B., Lahaye, N., Mémin, E., 2021. Improving internal variability of coarse-resolution ocean models by a stochastic transport. In preparation for Journal of Physical Oceanography .

Lord, G., Powell, C., Shardlow, T., 2014. An Introduction to Computational Stochastic PDEs. Cambridge Texts in Applied Mathematics, Cambridge University Press.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C., 1997a. A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. Journal of Geophysical Research: Oceans 102, 5753–5766.

Marshall, J., Hill, C., Perelman, L., Adcroft, A., 1997b. Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. Journal of Geophysical Research: Oceans 102, 5733–5752.

McWilliams, J.C., 1984. The emergence of isolated coherent vortices in turbulent flow. Journal of Fluid Mechanics 146, 21.

McWilliams, J.C., 2006. Fundamentals of Geophysical Fluid Dynamics. Cambridge University Press.

McWilliams, J.C., Restrepo, J., Lane, E., 2004. An asymptotic theory for the interaction of waves and currents in coastal waters. Journal of Fluid Mechanics 511, 135–178.

McWilliams, J.C., Sullivan, P., Moeng, C.H., 1997. Langmuir turbulence in the ocean. Journal of Fluid Mechanics 334, 1–30.

Mémin, E., 2014. Fluid flow dynamics under location uncertainty. Geophysical & Astrophysical Fluid Dynamics 108, 119–146.

Meneveau, C., Katz, J., 2000. Scale-invariance and turbulence models for large-eddy simulation. Annual Review of Fluid Mechanics 32, 1–32.

Mezić, I., 2013. Analysis of fluid flows via spectral properties of the Koopman operator. Annual Review of Fluid Mechanics 45, 357–378.

Mikulevicius, R., Rozovskii, B., 2004. Stochastic Navier–Stokes equations for turbulent flows. SIAM Journal on Mathematical Analysis 35, 1250–1310.

Mitchell, L., Gottwald, G., 2012. Data assimilation in slow-fast systems using homogenized climate models. Journal of the Atmospheric Sciences 69, 1359–1377.

Munk, W.H., 1950. On the wind-driven ocean circulation. Journal of Meteorology 7, 80–93.

NEMO team, 2016. NEMO ocean engine. Scientific Notes of Climate Modelling Center. Institut Pierre-Simon Laplace.

Pavliotis, G.A., Stuart, A., 2008. Multiscale methods: averaging and homogenization. Springer.

Pedlosky, J., 1984. The equations for geostrophic motion in the ocean. Journal of Physical Oceanography 14, 448–455.

Phillips, O.M., 1977. The Dynamics of the Upper Ocean. Cambridge University Press.

Pinier, B., Mémin, E., Laizet, S., Lewandowski, R., 2019. Stochastic flow approach to model the mean velocity profile of wall-bounded flows. Physical Review E 99, 063101.

Pope, S., 2000. Turbulent flows. Cambridge University Press.

Porta Mana, P., Zanna, L., 2014. Toward a stochastic parameterization of ocean mesoscale eddies. Ocean Modelling 79, 1–20.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical recipes 3rd edition: The art of scientific computing. 3 ed., Cambridge University Press.

Redi, M.H., 1982. Oceanic isopycnal mixing by coordinate rotation. Journal of Physical Oceanography 12, 1154–1158.

Reeks, M., 1983. The transport of discrete particles in inhomogeneous turbulence. Journal of Aerosol Science 14, 729–739.

Resseguier, V., 2017. Mixing and fluid dynamics under location uncertainty. Ph.D. thesis. Université Rennes 1.

Resseguier, V., Li, L., Jouan, G., Derian, P., Mémin, E., Chapron, B., 2020a. New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid computational fluid dynamics. Archives of Computational Methods in Engineering , 1886–1784.

Resseguier, V., Mémin, E., Chapron, B., 2017a. Geophysical flows under location uncertainty, part I: Random transport and general models. Geophysical & Astrophysical Fluid Dynamics 111, 149–176.

Resseguier, V., Mémin, E., Chapron, B., 2017b. Geophysical flows under location uncertainty, part II: Quasi-geostrophic models and efficient ensemble spreading. Geophysical & Astrophysical Fluid Dynamics 111, 177–208.

Resseguier, V., Mémin, E., Chapron, B., 2017c. Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence. Geophysical & Astrophysical Fluid Dynamics 111, 209–227.

Resseguier, V., Mémin, E., Heitz, D., Chapron, B., 2017d. Stochastic modelling and diffusion modes for proper orthogonal decomposition models and small-scale flow analysis. Journal of Fluid Mechanics 826, 888–917.

Resseguier, V., Pan, W., Fox-Kemper, B., 2020b. Data-driven versus self-similar parameterizations for stochastic Lie transport and location uncertainty. Nonlinear Processes in Geophysics 27.

Robert, C.P., , Cassela, G., 2004. Monte Carlo Statistical Methods. Springer Texts in Statistics. 2 ed., Springer-Verlag.

San, O., Staples, A.E., Iliescu, T., 2013. Approximate deconvolution large eddy simulation of a stratified two-layer quasigeostrophic ocean model. Ocean Modelling 63, 1–20.

San, O., Staples, A.E., Wang, Z., Iliescu, T., 2011. Approximate deconvolution large eddy simulation of a barotropic ocean circulation model. Ocean Modelling 40, 120–132.

Scott, R.B., Arbic, B.K., 2007. Spectral energy fluxes in geostrophic turbulence: Implications for ocean energetics. Journal of Physical Oceanography 37, 673–688.

Sirovich, L., 1987. Turbulence and the dynamics of coherent structures, part I: Coherent structures. Quarterly of Applied Mathematics 45, 561–571.

Smagorinsky, J., 1963. General circulation experiments with the primitive equations. Monthly Weather Review 91, 99–164.

Smith, K.S., 2007. The geography of linear baroclinic instability in earth's oceans. Journal of Marine Research 65, 655–683.

Stephenson, D.B., Dolas-Reyes, F.J., 2000. Statistical methods for interpreting Monte Carlo ensemble forecasts. Tellus A: Dynamic Meteorology and Oceanography 52, 300–322.

Stommel, H., 1948. The westward intensification of wind-driven ocean currents. Eos, Transactions, American Geophysical Union 29, 202–206.

Tissot, G., Mémin, E., Cavalieri, A., 2020. Stochastic linear modes in a turbulent channel flow. Journal of Fluid Mechanics .

Treguier, A.M., Held, I.M., Larichev, V.D., 1997. Parameterization of quasigeostrophic eddies in primitive equation ocean models. Journal of Physical Oceanography 27, 567–580.

Vallis, G.K., 2017. Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation. 2 ed., Cambridge University Press.

Van Leeuwen, P.J., 2009. Particle filtering in geophysical systems. Monthly Weather Review 137, 4089–4114.

Van Leeuwen, P.J., Künsch, H.R., Nerger, L., Potthast, R., Reich, S., 2019. Particle filters for high-dimensional geoscience applications: A review. Quarterly Journal of the Royal Meteorological Society 145, 2335–2365.

Weigel, A.P., 2012. Ensemble Forecasts. John Wiley and Sons, Ltd. chapter 8. pp. 141–166.

Weiss, J., 1991. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. Physica D: Nonlinear Phenomena 48, 273–294.

Wikle, C.K., M., B.L., 2007. A Bayesian tutorial for data assimilation. Physica D 230, 1–16.

Yang, Y., Mémin, E., 2019. Estimation of physical parameters under location uncertainty using an ensemble^2-expectation-maximization algorithm. Quarterly Journal Royal Meteorological Society 145, 418–433.

Zanna, L., Porta Mana, P., Anstey, J., David, T., Bolton, T., 2017. Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. Ocean Modelling 111, 66–80.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

Titre : Modélisation stochastique et simulation numérique des dynamiques océaniques

Mot clés : Modélisation stochastique, système dynamique, variabilité océanique, analyse des données, quantification d'incertitude, prévision d'ensemble

Résumé : Cette thèse explore une représentation stochastique de l'effet des petites échelles sur la circulation océanique grande échelle. Ce modèle stochastique, appelé modélisation sous incertitude de position (LU), résulte d'une décomposition de la vitesse en une composante lisse en temps et d'un terme de bruit très oscillant. Trois avantages principaux de ce modèle aléatoire sont soulignés dans cette thèse : il introduit une structuration des écoulements à grande échelle par l'inhomogénéité du bruit ; il fournit un système de prévision d'ensemble fiable et améliore la variabilité interne des modèles océaniques à résolution grossière.
Nous avons d'abord dérivé plusieurs modèles dynamiques aléatoires à partir du calcul stochastique et des lois de conservation physique. Nous avons démontré que certains modèles aléatoires conservent l'énergie. Dans ces modèles, une vitesse statistique induite par les petits tourbillons, et associée à l'inhomogénéité du bruit, est interprétée comme une généralisation de la dérive de Stokes. Nous avons montré que la dynamique stochastique dérivée contient un terme de forçage dit "force vortex" qui conduit à l'émergence de circulations secondaires. Cet effet est illustré numériquement par le cas test simple d'une onde de Rossby barotrope.
Nous avons ensuite proposé différents types de bruits - homogènes / hétérogènes et stationnaires / non stationnaires, en utilisant différentes méthodes - guidées par les données ou paramétrées. Pour les validations numériques nous avons utilisé une dynamique quasi-géostrophique surfacique pour étudier les capacités de prévision d'ensemble à court terme associés à différents modèles de bruit. En utilisant certaines métriques classique de prévision ensembliste, nous avons montré que les modèles LU fournissent une dispersion d'ensemble plus élevée ainsi que des prévisions plus fiables que les modèles aléatoires construits à partir de la perturbation de la condition initiale. De plus, nous avons conclu que les modèles de bruit hétérogènes sont meilleurs que les modèles homogènes et les modèles de bruit non-stationnaires plus performant que les modèles stationnaires.
Dans ce travail, des simulations numériques des circulations double-gyres forcées par le vent sont également évaluées. Dans le cas barotrope, nous avons montré que les modèles LU permettent de reproduire qualitativement et quantitativement sur un maillage grossier la distribution statistique à temps long et à grande échelle des traceurs, prédite par les données de haute résolution. Dans le cas multicouche, nous avons montré que la dérive de correction sous-maille, résultant de la procédure de sous-échantillonnage des données s'avère être très importante pour reproduire le jet zonal sur un maillage grossier. Nous avons mis en évidence que les modèles LU améliorent à l'échelle climatique océanique la variabilité intrinsèque à basse fréquence du courant à grande échelle. Cette amélioration a été démontrée par l'analyse du transfert d'énergie et de certains critères statistiques.

Title: Stochastic modeling and numerical simulation of ocean dynamics

Keywords: Stochastic modeling, dynamical system, ocean variability, data-driven model, uncertainty quantification, ensemble forecasting

Abstract: This thesis explores a stochastic representation of the small-scale effects on the large-scale oceanic circulation. This stochastic model, called modeling under location uncertainty (LU), arises from a velocity decomposition into a time-smooth component and a highly oscillating noise term. Three major benefits of such random model are outlined in this thesis: it introduces large-scale flow structuration by the noise inhomogeneity; it provides a reliable ensemble forecasting system and it improves internal variability of coarse-resolution ocean models.
We first derived several random dynamic models from stochastic calculus and physical conservation laws. Some resulting random models are demonstrated to conserve the energy. In these models, a statistical eddy-induced velocity, associated to the noise inhomogeneity, is interpreted as a generalization of the Stokes drift. We have shown that the derived stochastic dynamics encompasses the vortex force which leads to the emergence of secondary circulations. This effect is numerically illustrated from a simple test case of the barotropic Rossby wave.
We then proposed various types of noises - homogeneous/heterogenous and stationary/non-stationary, using different approaches - data-driven/parameterized. For numerical validations, we used the surface quasi-geostrophic dynamics to study short-terms ensemble forecasting skills predicted by different noises models. Using some classical metrics for ensemble forecasts, we have shown that the LU models provide higher ensemble spread and more reliable forecast than the random models built from perturbation of the initial condition. In addition, we concluded that the heterogeneous noise models are better than the homogeneous ones and the non-stationary noise models are better than the stationary ones.
In this work, numerical simulations of the wind-driven double-gyre circulations are also assessed. In the barotropic case, we have shown that the LU models enable us to reproduce qualitatively and quantitatively on a coarse mesh the long-term statistical distribution of the large-scale tracers, predicted by the eddy-resolving simulation data. In the multi-layered case, we have found that the subgrid correction drift, ensuing from the coarse-graining procedure of the high-resolution data, is very important in reproducing on a coarse mesh the meandering jet. We have highlighted that the LU models improve the intrinsic low-frequency variability of the large-scale current at ocean climatic scale. This improvement has been well demonstrated through the energy transfer analysis and some statistical criterions.